

Sensitivity of measuring the urban form and greenery using street-level imagery: A comparative study of approaches and visual perspectives

Filip Biljecki ^{a,b,*}, Tianhong Zhao ^{a,c}, Xiucheng Liang ^a, Yujun Hou ^a

^a Department of Architecture, National University of Singapore, Singapore

^b Department of Real Estate, National University of Singapore, Singapore

^c School of Architecture and Urban Planning, Shenzhen University, China

ARTICLE INFO

Keywords:

User-generated
Urban sensing
GeoAI
Mapillary
KartaView
Volunteered Geographic Information

ABSTRACT

Street View Imagery (SVI) is crucial in estimating indicators such as Sky View Factor (SVF) and Green View Index (GVI), but (1) approaches and terminology differ across fields such as planning, transportation and climate, potentially causing inconsistencies; (2) it is unknown whether the regularly used panoramic imagery is actually essential for such tasks, or we can use only a portion of the imagery, simplifying the process; and (3) we do not know if non-panoramic (single-frame) photos typical in crowdsourced platforms can serve the same purposes as panoramic ones from services such as Google Street View and Baidu Maps for their limited perspectives. This study is the first to examine comprehensively the built form metrics, the influence of different practices on computing them across multiple fields, and the usability of *normal* photos (from consumer cameras). We overview approaches and run experiments on 70 million images in 5 cities to analyse the impact of a multitude of variants of SVI on characterising the physical environment and mapping street canyons: a few panoramic approaches (e.g. fisheye) and 96 scenarios of perspective imagery with variable directions, fields of view, and aspect ratios mirroring diverse photos from smartphones and dashcams. We demonstrate that (1) disparate panoramic approaches give different but mostly comparable results in computing the same metric (e.g. from R=0.82 for Green View to R=0.98 for Sky View metrics); and (2) often (e.g. when using a front-facing ultrawide camera), single-frame images can derive results comparable to commercial panoramic counterparts. This finding may simplify typical processes of using panoramic data and also unlock the value of billions of crowdsourced images, which are often overlooked, and can benefit scores of locations worldwide not yet covered by commercial services. Further, when aggregated for city-scale analyses, the results correspond closely.

1. Introduction

Largely thanks to panoramas provided by commercial sources such as Google Street View and Baidu Maps, street view imagery (SVI) rose to prominence in urban and environmental studies, introducing efficient, large-scale, and convenient means to sense the built environment and streetscapes (Biljecki and Ito, 2021; Deng et al., 2021; Guan et al., 2022; Seiferling et al., 2017; Qiu et al., 2022; Ki and Lee, 2021; Helbich et al., 2021; Kang et al., 2021b; Liang et al., 2023). Among many applications, it has been monumental for characterising the urban form and measuring the amount of vegetation, serving a variety of disciplines that depend on understanding the physical environment quantitatively.

A good deal of approaches quantifying urban morphology and street greenery were possible before the advent of SVI with other and well established types of data such as vector geospatial datasets and satellite imagery (Miao et al., 2020; Gong et al., 2018; Li et al.,

2018). However, powered by computer vision techniques such as image segmentation and the multiplication of computing resources, SVI was rapidly established as a prominent approach to conduct such studies and it has been complementing, if not overtaking traditionally used data sources, and also introducing new use cases not possible with any other source of data (Gaw et al., 2022). The highly visible arrival of SVI was no accident — the key reasons SVI carved out its own unique role in the urban data pantheon are readily available datasets, global coverage, ease of access, high spatial resolution, scalability, convenience of use, veracity, street-level perspective, and continuous advancements in computer vision (Biljecki and Ito, 2021; Aikoh et al., 2023).

Most studies focus on characterising the omnipresent elements of the urban form — buildings, greenery, and visible sky. Measures that parameterise them, such as the Sky View Factor (SVF), Green View

* Corresponding author at: Department of Architecture, National University of Singapore, Singapore.

E-mail addresses: filip@nus.edu.sg (F. Biljecki), zhaotianhong2016@email.szu.edu.cn (T. Zhao), xiucheng@u.nus.edu (X. Liang), yujun@nus.edu.sg (Y. Hou).

Index (GVI), and Building View Factor (BVF), dominate the multidisciplinary literature (Gong et al., 2018; Lu et al., 2023). The general tenor of the approaches is straightforward — the proportion of an image that is detected to represent a particular urban form element (e.g. greenery, sky) is considered as a proxy of the exposure to it at that particular (ground-level) location. With the large-scale availability of street-level imagery and readily available computing resources, the approach is often scaled using tens of thousands of photos across a city to conduct urban scale studies at a high spatial resolution. These simple but powerful metrics have accelerated a great number of studies to understand their interaction with urban phenomena such as urban heat island effect, comfort, housing prices, mental health, and mobility choices (a comprehensive overview will be given in Section 2).

Despite such studies now being almost routinely conducted around the world, there are two trends and factors that warrant investigation, which are expected to become even more important in the future.

First — the proliferation of such studies exposed diverging approaches, and there is a lack of overview of the different instances used and their comparison. For example, researchers use substantially different perspectives (e.g. a panorama or fisheye image) and various parameters to calculate seemingly the same metrics, and there is a lack of clarity in terminology and measures (e.g. Green View Index and Green View Factor both appear in the literature). While SVI continues to be a staple of numerous studies, researchers tend to follow certain approaches without motivating their choice and discussing their implications, uncertainty, and reliability.

Second — studies predominantly rely on commercial platforms that offer panoramic imagery, but crowdsourced (citizen-led) efforts of collecting street-level images and platforms to store them (e.g. Mapillary, KartaView) have emerged as well, gaining attention for the same research purposes, similarly as other urban data such as climate measurements (Meier et al., 2017). While offering advantages such as flexibility and countering restrictions of commercial services, their notable shortcoming is the field of view and coverage — most volunteers use non-panoramic cameras, e.g. dashcams and smartphones mounted in cars, offering a limited (single-frame) view. It remains unknown and not systematically investigated whether such imagery can rival the well established panoramic imagery that gives a much broader overview of the streetscape. Moreover, this uncertainty is compounded by the large variety of consumer equipment used in practice, e.g. cameras have different fields of view and aspect ratios.

Conflating the two seemingly disconnected aspects, Fig. 1 illustrates the background and motivation for this research. Commercial services offer panoramic imagery (360 degrees horizontally and 180 degrees vertically) enabling a broad perspective and extensive survey of the streetscape and the built environment in the immediate vicinity (Fig. 1A and B). The panorama is often transformed to another view — e.g. a set of perspective images based on a various parameters (Fig. 1C) or fisheye directed to the zenith (Fig. 1D). Images are segmented using computer vision techniques (Fig. 1C, D, and G) to compute urban form indexes such as the Sky View Factor and Green View Index based on the proportion of each class in the image (Fig. 1 bottom right), and sometimes aggregated. These metrics support numerous application domains, but the difference among various approaches to derive them has not been investigated. On the other hand, due to the consumer equipment and heterogeneous approaches used (Fig. 1E), the vast majority of crowdsourced imagery has only a single-frame (limited) and perspective view (Fig. 1F, G) both horizontally and vertically, and in most cases facing forward the direction of the road, offering a limited understanding of the built environment (Fig. 1H) without a panoramic view offered by commercial platforms. It is not known extensively whether perspective (limited, non-panoramic) single images could be sufficient for the same tasks panoramic SVI owes its success in the research community. Finally, not all perspective imagery in crowdsourced services is front-facing and following the direction of driving, but it may include also lateral perspectives (Fig. 1I), and many

others we investigate in this paper, of which usability and purpose remain unexplored.

In this paper, for the first time, we build a structured overview of approaches and a mosaic of indicators found in research. We investigate systematically and comprehensively, with experiments, what are their relationships, uncertainty, and performance; and how restricted perspectives in the growing volume of crowdsourced datasets collected by citizens affect the reliability of analyses typically conducted on panoramic imagery, which are paramount for urban research and planning.

The broader significance of this work is that its findings may lead to the standardisation and awareness of different approaches commonly used in research nowadays, and to giving recognition to crowdsourced street-level imagery and understanding its potential across many domains. Commercial services such as Google Street View and Baidu Maps, while with impressive volume of data, are not available everywhere, and this research may contribute to motivating using crowdsourced counterparts of this data to fill data gaps around the world, and lead to developing best practices (e.g. determining the optimal direction of imaging). Further, giving adequate attention to the theory and foundations, the paper doubles as an overview of these metrics for readers who are not acquainted with this topic and the workflows that underpin numerous studies in multiple application domains.

The outcome of this research is important also for researchers who use panoramic (commercial) imagery, not only to understand the interaction among different approaches to make use of panoramas, but also because if limited views turn out to be sufficient, it may mean that researchers now use more data than necessary. That is, it may not be essential to obtain and process the entire panorama, lowering costs of research as the use cases tend to be data- and power-hungry. Further, due to the increasing barriers, commercial sources should not be taken for granted, so researchers in the future may have to switch to crowdsourced imagery.

2. Background and related work

2.1. Street view imagery

Commercial services such as Google Street View, Baidu Maps, and Tencent offer billions of panoramic images of numerous urban areas around the world, enabling immersive insights in the streetscape and rivalling oblique and satellite imagery traditionally used in many research domains (Helbich et al., 2021; Gaw et al., 2022; Luo et al., 2022a). Companies provide standardised approaches to collect data (multiple cameras mounted on cars), resulting in relatively homogeneous data characteristics around the world.

The imagery has been used for estimating house prices (Qiu et al., 2022; Kang et al., 2021a), assessing walkability and bikeability (Ki and Lee, 2021; Ito and Biljecki, 2021), characterising the urban form (Middel et al., 2019; Hu et al., 2020), revealing the spatio-temporal evolution of urban visual environments (Liang et al., 2023), supporting epidemiology (Zhang et al., 2021a), estimating solar radiation (Deng et al., 2021; Lan et al., 2021), mapping street furniture and infrastructure (Zhong et al., 2021; Zhou et al., 2022b), generating 3D city models (Pang and Biljecki, 2022), understanding urban agriculture (Hawes et al., 2022), mapping land cover (Ning et al., 2022a), socio-economic studies (Suel et al., 2021; Byun and Kim, 2022), extracting building information (Ning et al., 2022b; Yan and Huang, 2022; Sun et al., 2022), inferring streetscape quality (Li et al., 2022b) and their vibrancy (Li et al., 2022a) and disorder (Chen et al., 2022b), and predicting perception (Guan et al., 2022; Verma et al., 2020). For the full list of applications, see Biljecki and Ito (2021). The process how these panoramic images are used will be expounded in the next section, together with the different approaches.

The rise of Volunteered Geographic Information (VGI) (Feng et al., 2022; Naghavi et al., 2022; So and Duarte, 2020; Yan et al., 2020),

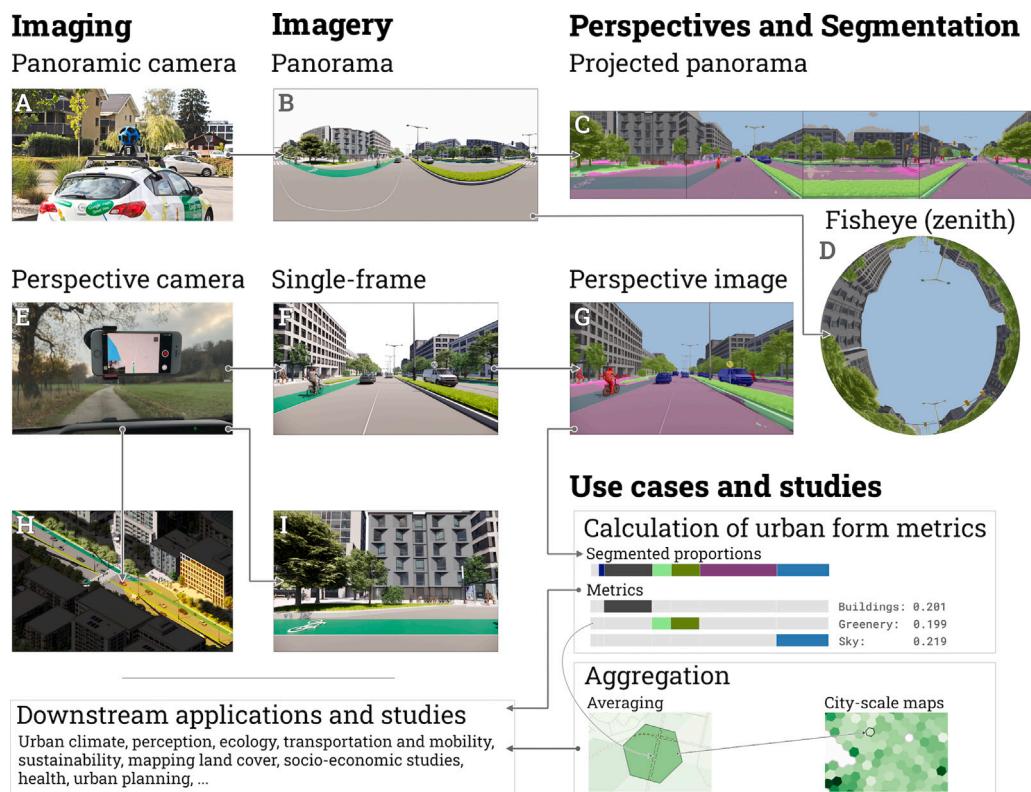


Fig. 1. An overview of typical practices and motivation for our multivariate work. Please refer to the text for the explanations. Source of the photos A and E: Wikimedia Commons.



(a) Longitudinal view from a perspective camera in a smartphone.

(b) Lateral view from an ultrawide angle action camera and with a different aspect ratio.

Fig. 2. Examples of crowdsourced images acquired from cars at the same location by different contributors, having different directions, fields of view, and aspect ratios. The increasingly heterogeneous image data on streetscapes necessitates investigation of the influence of myriads of parameters used. While having some advantages, in both cases, their insight into the streetscape tends to be limited in comparison with panoramic imagery typically found in commercial data. Source of data: Mapillary.

where citizens are producers of geospatial data that can be used freely, gave impetus to also volunteered street-level imagery and development of services such as Mapillary and KartaView, which have gained attention and now host a large volume of data from all over the world (Juhász and Hochmair, 2016; Leon and Quinn, 2019; Zhang et al., 2021b; Mahabir et al., 2020; Ma et al., 2019). VGI efforts have been contributed by companies as well (Anderson et al., 2019; Sarkar and Anderson, 2022), which is also the case for street-level imagery (Huang et al., 2019). Unlike commercial counterparts, the collection of these datasets is not standardised — anyone, anywhere, can collect and share sequential images using any kind of camera and mode of transportation (Hou and Biljecki, 2022; Luo et al., 2022b), which results in diverse imagery (Fig. 2). While there are different approaches and while panoramic imagery is supported, most of the images collected in practice are perspective (non-panoramic, single) images from dashcams and smartphones mounted inside cars, resulting in heterogeneous characteristics.

Now that billions of images are available for free and without restrictions, they have been used for some of the same research purposes as commercial panoramic imagery (Karasov et al., 2018; d'Andrimont et al., 2018; Stowell et al., 2020; Lummitz et al., 2021; Fan et al., 2021; Ding et al., 2021; Raviscioni et al., 2022; Bianconi et al., 2022; Inoue et al., 2022b; Nievias et al., 2022; Esch et al., 2022; Yap et al., 2022a; Yao et al., 2022a; León-Sánchez et al., 2022). Besides accessibility, their advantage is that it is often the only SVI data source at a location, e.g. in places where commercial services do not have any coverage yet (notwithstanding the significant accomplishment of commercial platforms offering data for about half of the world's jurisdictions, there remains an extensive portion of the globe they are yet to cover (Cinnamon and Gaffney, 2021)).

Despite such advantages and the several papers cited in the previous paragraph, user-generated street-level imagery remains rarely used in research, and the number of studies relying on it is dwarfed by those using Google Street View and Baidu Maps (Biljecki and Ito, 2021). Ostensibly, their disadvantage is the field of view limited by low-cost

consumer equipment (Fig. 2), in a stark contrast from professional panoramic imaging systems developed by commercial providers, thus researchers may have no motivation to use them if an easily available commercial source of panoramic imagery is available in the same area (Ali-bey et al., 2022). It is obvious that perspective photos (Fig. 2) may have shortcomings when sensing elements such as buildings, greenery, and sky, possibly being limited measuring the built environment. But is that really the case? Is panoramic imagery truly essential in such studies? In our work we investigate whether such limited view is actually sufficient to reliably sense commonly measured environmental elements of the streetscape for which researchers overwhelmingly rely on panoramas. Along with that, we investigate also whether the value of panoramic imagery is unparalleled or it has been overestimated for frequently employed analyses, and single-frame imagery from crowdsourced platforms such as Mapillary has been overlooked undeservedly.

These issues are also important considering the increasing number of studies that collect SVI on their own (Peng et al., 2018; Verma et al., 2019; Ao et al., 2019; Bochkarev and Smirnov, 2019; Gorgul et al., 2019; He et al., 2020; Wang et al., 2021; Oki and Ogawa, 2022; Kim and Lee, 2022; Ogawa et al., 2021), many of which are similar to crowdsourced imagery, e.g. collecting lateral imagery with perspective cameras (Raviscioni et al., 2022).

2.2. Characterising common elements: sky, buildings, and greenery

Literature review reveals that the most common urban form elements that are characterised using SVI are buildings (i.e. built form), greenery (or trees, vegetation), and sky (i.e. open canyon space). Thus, we focus on them. Because of the similarities of approaches and logic, and because many studies examine more than one of these elements in conjunction (Fu et al., 2019; Zhou et al., 2022a; Basu and Sevtsuk, 2022; Kawshalya et al., 2022), they are covered together. The explanations in this section are accompanied by Fig. 3 (the bottom portion of the image is related to the methodology of this study and it will be introduced in Section 3).

In general, studies transform the data into a fisheye image pointing up or into multiple perspective images that form a 360° panorama (but often not a vertical panorama, with the top and bottom portions trimmed) (Seiferling et al., 2017; Li et al., 2015; Yao et al., 2021). The former approach (fisheye) has been influenced by the traditional way to compute the Sky View Factor — using cameras with fisheye lenses pointing to the zenith and computing the proportion of the sky hemisphere visible from a ground-level point (Li et al., 2018; Gong et al., 2018; Miao et al., 2020; Lan et al., 2021). The latter one (set of perspective images forming a panorama) is done to mitigate distortions and facilitate labelling and processing with computer vision approaches (Zhou et al., 2022a; Beaucamp et al., 2022). It has been used mostly for parameterising greenery and buildings. For example, a plethora of studies has computed the Green View Index (GVI), which indicates the degree to which a person standing in a certain position can view vegetation (Yang et al., 2009; Ki and Lee, 2021).

Looking at Fig. 3, these general practices appear to make sense — a fisheye image pointing to the zenith gives justice to the complete parameterisation of sky and open space. On the other hand, it will not fully capture vegetation, as some street-side greenery (e.g. grass) will be left out, thus, panoramas without a full overview of the sky but with the horizon in their focal point may be more suitable for such purpose. The same goes for studies that are focused on mapping ground-level features such as sidewalks (Kang et al., 2021a; Hosseini et al., 2022; Ning et al., 2022c).

However, there are many studies that do not follow the same rationale. For example, there are those that characterise the proportion of buildings with fisheye imagery (Gong et al., 2018). It is rarely explained why that is the case. It is assumed that researchers do so to avoid including multiple approaches and to keep the derived metrics

in the same context, e.g. when computing both sky and greenery, they focus on either of the two approaches.

The first issue in this domain are terminological disparities. There are no standardised names. In fact, in some instances there are no particular terms used at all — especially in those in which environmental elements of the streetscapes that are analysed are many (Wu et al., 2020b). Further, researchers tend to use different terms for the same metric, e.g. the Green View Index is also known as *Visible Green Index* (Zhu et al., 2021), an issue that in general pervades lines of research that parameterise the urban form (Biljecki and Chow, 2022; Xu and Gao, 2022).

In this paper, influenced by the two most prominent metrics — SVF and GVI, and following the names of metrics in some established publications (e.g. Building View Factor in Gong et al., 2018), we define that metrics should suffix elements with *View Factor* if they are computed with fisheye imagery pointing to the zenith, while those that are computed from horizontal panoramas should use the term *View Index*, which suggest how much object pedestrians observe (see Fig. 3 for reference). Now that the terminology has been clarified, the application of each of these metrics is overviewed.

The SVF has a long history in domains ranging from satellite positioning to urban climate studies as it has a direct effect on the amount of satellite signals and solar radiation received at that point (Chapman et al., 2002; Chen et al., 2022d; Svensson, 2004; Demuzere et al., 2019; Wang et al., 2018; Ignatius et al., 2022). It has been also used for predicting ground shade provided by trees (Li et al., 2018; Sun et al., 2021a), explaining the urban heat island effect (Martin et al., 2022), informing urban planning interventions (Ye et al., 2018), forecasting dengue (Liu et al., 2021), explaining health-related factors and physical activity (Lu et al., 2019; Yao et al., 2022b), and understanding pedestrian volume (Chen et al., 2022c). The counterpart of SVF computed from panoramas — Sky View Index — has been used in a variety of studies (Chen et al., 2022a; Jing et al., 2021; Liao et al., 2021; Zhou et al., 2022c; Gao et al., 2022) (we purposefully do not abbreviate this metric to avoid the conflict with SVI that is reserved for street view imagery). Both metrics are encountered with alternative names such as level of open sky (Chen et al., 2022c), percentage of sky (Verma et al., 2020), and proportion of sky (Yin and Wang, 2016; Gong et al., 2019a).

The SVF can be computed in many ways, such as 3D city models, point clouds, building footprints, and most accurately — with field surveys with cameras equipped with fisheye lenses. But SVI, mimicking this approach by transforming the cylindrical panoramas into azimuthal fisheye images, provides a convenient and scalable approach retaining approximately the same level of quality (Dirksen et al., 2019; Chen et al., 2020b; Heo et al., 2020; Yuan and Chen, 2011; Li et al., 2015; Seiferling et al., 2017; Middel et al., 2017, 2018; Gong et al., 2018; Biljecki and Chow, 2022; Zeng et al., 2018; He et al., 2022b; Lipson et al., 2022), and it provides some advantages over other datasets. For example, as 3D city models often contain only buildings (Labetski et al., 2022; Peters et al., 2022; Wysocki et al., 2022), SVI has gained popularity over them thanks to the inclusion of all urban features, most importantly vegetation (Gong et al., 2018).

Moving on to greenery and buildings, both factors and indexes have been used extensively, e.g. Green View Factor (Peng et al., 2022; Hua et al., 2022), Green View Index (Chen et al., 2019; Wu et al., 2019; Chen et al., 2020c), Building View Factor (Gong et al., 2018), and Building View Index (Chen et al., 2020a). Greenery is commonly computed conflating trees and grass in a scene (Koo et al., 2022; Yu et al., 2021; Zhang and Hu, 2022), but there have been more specific instances such as Tree View Index (Qiu et al., 2022) and Tree View Factor (Gong et al., 2018), and differentiating different classes of greenery in cities (Sun et al., 2021b). These metrics have been so common and successful that they have transcended the street-level, e.g. some have been adopted at other platforms such as window-level in a building (Li et al., 2022c).

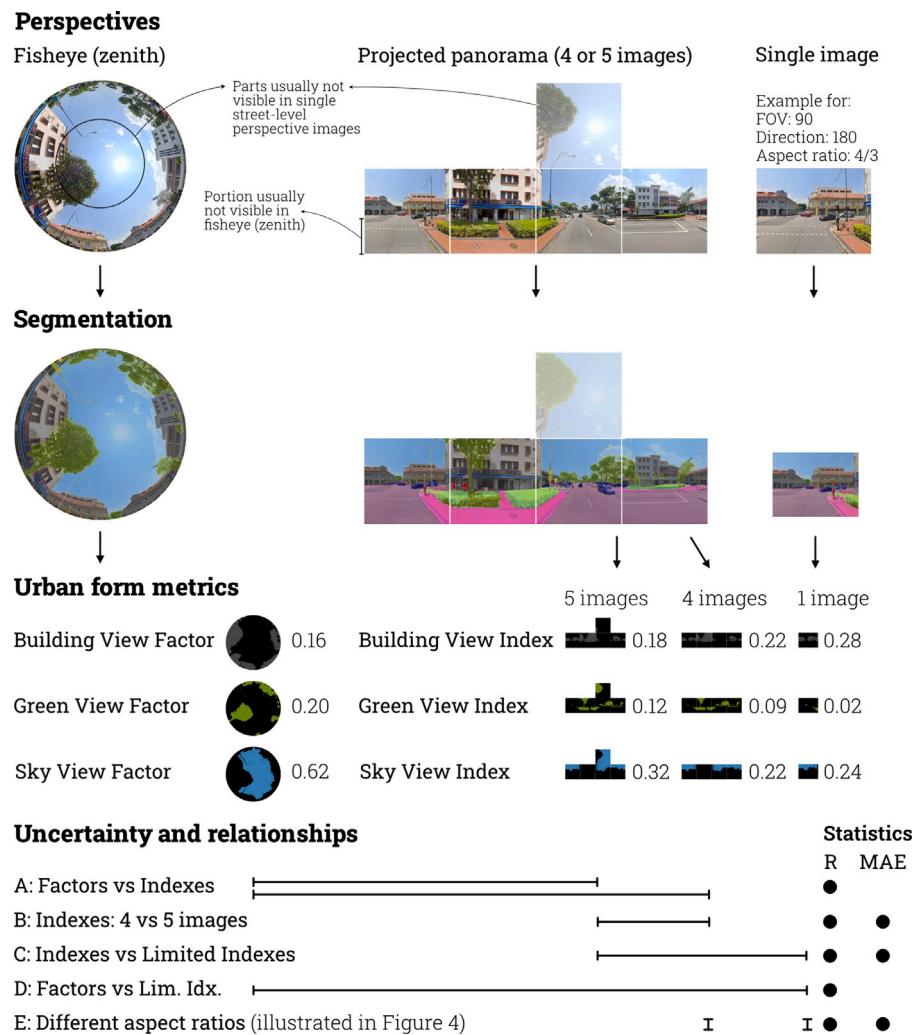


Fig. 3. Different approaches to compute the indicators. For the combination of single image parameters, refer to Fig. 4. Source of the imagery: Google Street View.

The proportion of buildings and vegetation, either alone or in combination, in a street scene have been used for estimating the real estate premium (Yang et al., 2021b; Chen et al., 2020a), monitoring street-side changes (Yu et al., 2022), understanding rideshare accessibility (Wang et al., 2022), crime occurrence (He et al., 2022a), perception (Zhang et al., 2020), and air pollution (Wu et al., 2020a).

These dimensionless metrics are not relevant only in urban canyons and densely built areas, but rather for a variety of urban configurations. Thus, in our work, we investigate diverse urban configurations around the world.

2.3. Related work

In general, little is known about the impact of parameters of imagery and the topic of errors, uncertainty, and quality in SVI. To the extent of our knowledge, the only work related to ours is the one by Kim et al. (2021). It has a dual purpose: understanding the influence of the sampling rate of SVI (e.g. the difference between retrieving imagery every 20 or 100 m along a road) and the influence of considering different combinations of perspective images derived from a panorama (e.g. perspective image directed at the right side of the car vs a panorama with a 360 view). The latter part, referred to as directional setting in their paper, is more related to ours. The study primarily computes indexes such as SVF for each perspective and its correlation with the one derived from the panorama at the same location. Their results suggest that: (i) information is lost when limited views are used;

(ii) the results obtained from an image directed at one lateral side (e.g. camera pointing left; cf. Fig. 11) can be significantly different from those aiming to the other side; and (iii) the correlation (i.e. measure for error or information loss) depends on both the index computed and the view. For example, having only the front view results in correlations between 0.652 and 0.96 for different indexes, while focusing on one index (sky view) computed from different directions reveals that the correlation varies between 0.822 (right view) and 0.961 (back view).

We bring several novelties that have not been focus in research yet, and several substantial advancements with respect to the state of the art. First, we systematically and comprehensively investigate a large number of scenarios occurring in practice and their relationships (e.g. panoramas obtained from imagery with different aspect ratios, and panoramas obtained from a different number of images, to name a few among many). Second, their work is concentrated on the topic of splitting imagery from panoramas (resulting in fixed fields of view in both horizontal and vertical dimensions), rather than considering also the context of crowdsourced imagery in which the field of view (FOV) is not standard due to myriads of equipment and the fact that the horizontal and vertical FOV are rarely equal (resulting in a square image). In our work, we zero in on the topic of field of view for the first time, and consider multiple options, reflecting different equipment used in practice. We will demonstrate with experiments that this aspect is quite important. Third, their work focuses on a single and limited study area (a town in California in USA), while we include multiple cities around the world to increase diversity of urban forms. Our decision

is affirmed by the fact that we obtained different results by city. In general, SVI studies involving more than one city are scant, thus, our work contributes as a cross-city comparative study in understanding the urban form and intra-city distribution of the aforementioned metrics. Fourth, their results include only one metric (correlation), while our work adds another metric (mean absolute error) that may be more suitable for interpreting the results.

3. Methodology

In a nutshell, we generate imagery according to different scenarios and approaches, and compare the values of the generated metrics based on a custom-built and highly-adaptable workflow.

3.1. Imagery and perspectives

We consider common practices, and regard the following types of perspectives (Figs. 3 and 4):

- Fisheye (zenith). These are used to compute the Building/Green/Sky View Factors.
- Panoramas, which are used to compute the indexes:
 - 4-image panoramas that are obtained from four perspective images extracted according to cardinal directions (0, 90, 180, and 270 degrees). We noticed in literature that they use different aspect ratios. So we compute three variants of these 4-image panoramas, using aspect ratios 1/1 (square, most common), 4/3, and 16/9.
 - 5-image panoramas — these are obtained from the 4-images listed above with an additional one pointing upwards. They are computed only using square images (1/1 aspect ratio). One may wonder why simply not focus on this kind of panorama, which gives an entire overview from a location and as it should be always possible to compute with commercial services. The reasons depend on the use case — e.g. in some studies, panoramas tend to mimic the view of pedestrians, so overhead views are not included.
- Perspective images (limited views), which represent those taken from consumer cameras and which are found in crowdsourced services, but are also used to construct panoramas when using commercial services (previous point). We regard 8 different directions: the same 4 as above (0, 90, 180, and 270 degrees) and 4 additional (diagonal) ones between (45, 135, 225, and 315 degrees), which are not common in practice, except in few publications (Chu et al., 2022) and some crowdsourced circumstances where e.g. dashcams are misaligned, but may be interesting to investigate as they combine longitudinal and lateral imagery (e.g. 45 — one that includes a bit of the front view and glimpse on the right). In parallel, to consider the heterogeneous equipment used in practice, we introduce multiple fields of view — 70, 80, and 90 to approximate typical ranges found in wide-angle cameras in consumer smartphones, and 120 to include those that are typical for dashcams and ultrawide-angle cameras in smartphones. Another motivation for the different FOVs is that nowadays, smartphones tend to have multiple cameras (e.g. wide- and ultrawide-angle), thus, given the choice, one may wonder which of these may be better to use for urban sensing. In our exploration of the metadata of a crowdsourced platform,¹ we find that most commonly images have aspect ratios of 4/3 and 16/9, but we include also 1/1 for consistency. Permuting all these

parameters, in total, we have 96 such images (Fig. 4), which we analyse separately. Because of the limited perspective, to distinguish them from panoramas, we call the indexes computed from such single perspectives as *limited indexes*.

3.2. Investigated metrics and statistics

We investigate the following five relationships (Fig. 3), and for notation we assign a letter to each. For each relationship, we compute one or two of the following metrics: Pearson correlation coefficient (R) and Mean Absolute Error (MAE). For all relationships, we compute the correlation coefficient to indicate the statistical relationship between two variables. However, we compute the MAE only for those metrics that are comparable, as MAE is used as measure of errors between observations expressing the same phenomenon (e.g. we do not do that when comparing indexes to factors, as it is meaningless to talk about errors in that sense).

Statistic A — Factors and indexes. We start by investigating what is the relationship between the factors (obtained from fisheye) with the indexes obtained from panoramas (both those computed from 4 and 5 images). Comparing these two key means to measure the urban form (Section 2.2) will shed light on their correspondence and potential differences.

Statistic B — Indexes from different variants of panoramas. While panoramas are used frequently, there are multiple ways to generate them: comprising 4 and 5 images. Thus, we analyse their difference across the three indexes. Besides R, for this relationship, we compute also MAE.

Statistic C — Indexes and limited indexes. This is the central part of the work, analysing the value of single (limited) views in comparison with panoramas, helping to answer whether crowdsourced SVI, which tends to have non-panoramic imagery is sufficient for sensing the urban form. Using this relationship, we analyse a large number of perspectives with different fields of view, directions, and aspect ratios. Thus, it helps us to understand and compare a variety of questions and scenarios, such as whether having a wide-angle camera pointing longitudinally is better than a ultrawide-angle camera pointing laterally; or if we had the choice of collecting data with a typical smartphone camera, in which direction it would be most useful to point it when measuring e.g. the Green View Index. This part is also of relevance to researchers using panoramic imagery, as it will help understanding does it make sense to use the entire panorama if one of its parts is sufficient, i.e. can panoramic imagery be reduced to a limited perspective, achieving similar results but decreasing the complexity of studies.

Statistic D — Factors and limited indexes. Similarly to C, this statistic also caters to understanding the usability of perspective cameras, but in comparison to metrics derived from fisheye imagery. That is, the results obtained from single-frame views are compared to those derived from fisheye imagery.

Statistic E — Different aspect ratios. This relationship regards the different aspect ratios found in practice, and it has two related parts. First, it focuses on 4-image panoramas, it compares the different aspect ratios when generating them, e.g. 4/3 vs 16/9. Second, it seeks to understand the role of consumer cameras with different aspect ratios (e.g. smartphones and dashcams, both used in crowdsourced imagery, tend to have different aspect ratios; see Fig. 2).

3.3. Study area, data, and image segmentation

Departing from the vast majority of studies focusing on a single city (Section 2.2), we choose five cities, with diverse morphological, geographical, and planning characteristics around the world: Amsterdam (the Netherlands), Hong Kong (China), Kampala (Uganda), San

¹ We examined metadata of 100k randomly sampled images from Mapillary from around the world collected and uploaded to the platform in the past decade.

Generating perspectives

Original panorama



8 directions
0
45
90
135
180
...

4 FOVs x 8 directions
32 combinations

FOV 70
FOV 80
FOV 90
FOV 120

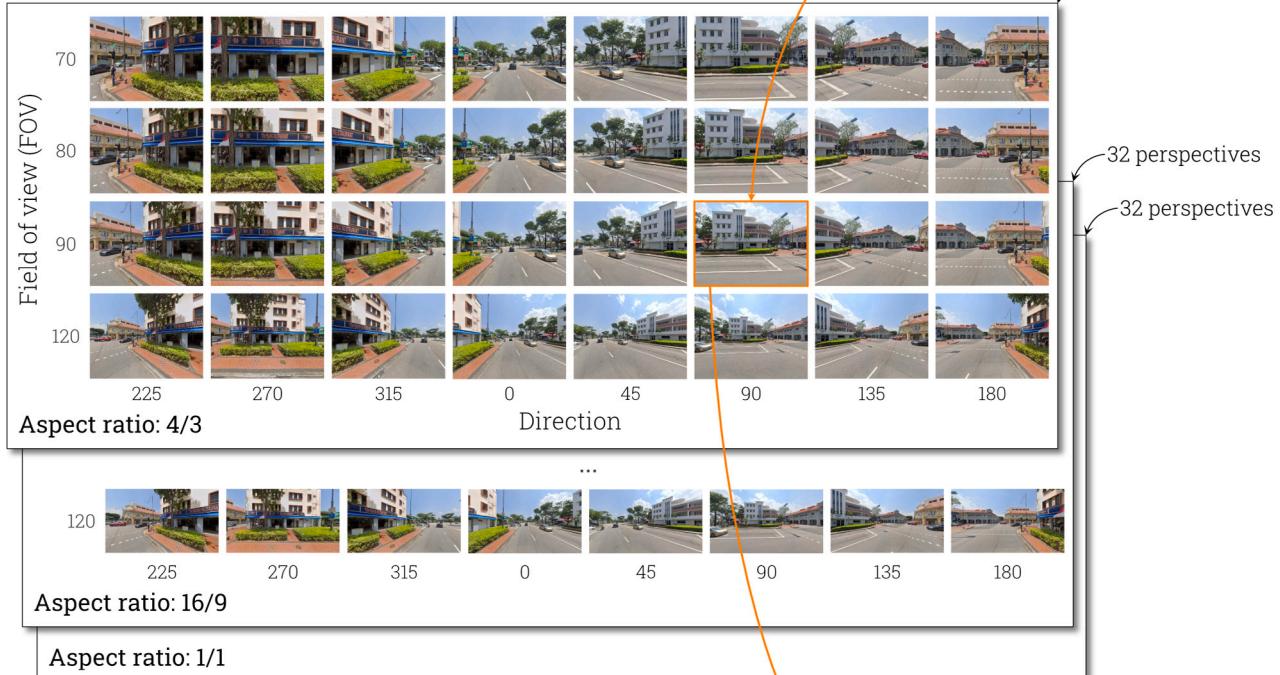
Examples for
direction = 90
with all 4 FOVs

32 FOV & direction combinations
x 3 aspect ratios = 96 perspectives

1/1
4/3
16/9

Examples for
direction = 90
FOV = 90
for all 3 aspect ratios

Limited perspectives (set of 96 single images)



Analysis

Reference values

Panorama	Fisheye
BVI = 0.18	BVF = 0.16
GVI = 0.12	GVF = 0.20
SVI = 0.32	SVF = 0.62

Computation of uncertainty
and relationships
(set of statistics C and D)

Computation of
limited indexes

Building = 0.13
Greenery = 0.09
Sky = 0.32

Segmentation of limited
perspectives



Fig. 4. The single image scenarios we simulated, as permutations of common aspect ratios, fields of view, and directions. We compute the three indexes of the urban form (which we dub as *limited indexes* to distinguish them from those generated from 360 imagery) for each scenario. Not all scenarios are shown for space constraints. The figure also expounds the metrics relevant for the limited perspectives. Source of the imagery: Google Street View.

Francisco (USA), and Singapore. The importance of measuring the urban form across diverse areas has been attested to identifying that all of these cities have been subject of past studies on the urban form (Mukisa et al., 2019; Rafiee et al., 2016; Liang et al., 2019; Sevtsuk et al., 2021; Wong and Kardinal Jusuf, 2010; Hosseini et al., 2022; Yang et al., 2021a; Gong et al., 2019b; Chen and Biljecki, 2023). Fig. 5 illustrates the diversity of our study areas, and considering such will reveal whether the relationships we investigate will be consistent across various morphologies. The data has been obtained from Google Street View. We have obtained imagery at 691,750 locations. For each location, we have obtained a fisheye image, a few panoramas (5-image, and multiple 4-image panoramas with different aspect ratios), and the 96 limited perspectives as illustrated in Fig. 4, resulting in nearly 70 million images (and more than 200 million metrics we compute from them). For the hemispherical transformation of the panorama into the fisheye photograph looking vertically (upwards with the zenith in the centre), we use the upper half of a panorama, and reproject it. The

factors have been computed using approaches found in the state of the art (Miao et al., 2020; Xia et al., 2021; Gong et al., 2018).

For the segmentation, we used Cityscapes² as the training dataset, which is a large-scale database that focuses on semantic understanding of urban street scenes. It provides semantic, instance-wise, and dense pixel annotations for 19 classes, including the ones in focus (c.f. Fig. 1). By comparing different semantic segmentation models, we adopted the state-of-the-art model, DeepLabV3, as the Mean Intersection over Union (mIoU) can exceed 80% in this task (Chen et al., 2018).

The statistical distributions (Fig. 5) reveal the diversity and data distribution in our study. For example, one might have expected that Hong Kong, well known for its urban jungle dotted by the world's largest number of skyscrapers, would yield a distribution with higher building indexes/factors and lower sky view indexes/factors than other cities. However, much of the imagery is taken all around the city,

² <https://www.cityscapes-dataset.com/>



Fig. 5. Distribution of values in the cities with example scenes from the study area. The values are derived from fisheye imagery looking up towards the zenith, but the visual samples of scenes are from a different perspective — standard panoramic imagery as they may be more suitable for a visual impression of the examples in question. Source of imagery: Google Street View.

including suburban villages, country parks and nature reserves, adding to the diversity of settings in our research.

4. Results

Because of the large number of results but limited space, this section features the most important results across the five statistics in focus. Most of the attention is given to limited perspectives, the central focus of this paper, which are dealt with by statistics C and D. The scatter plots contain a random subset of the observations to avoid clutter.

4.1. Factors and indexes (Statistic A)

Fig. 6 portrays the relationship between the indexes (computed from 4- and 5-image panoramas) as correlation coefficients and scatter plots. There are a few important observations. The correlation coefficients, while in general strong, are different for the three urban elements. The weakest (albeit still strong) relationship is for greenery. This could be explained by the fact that the Green View Factor (computed from fisheye) misses ground-level vegetation such as grass, which are included in Green View Indexes (both 4- and 5-image panoramas). On the other hand, 4-image panoramas have a limited peek into overhead greenery, which are well captured in fisheye imagery. Therefore, a key observation is that the different approaches may capture different aspects of urban greenery.

As expected, all elements in 4-image panoramas have a somewhat weaker correlation than 5-image panoramas (their relationship is investigated in the next section as a separate statistic). As it is the case with tall greenery, 4-image panoramas in high-rise settings may not be able to capture the full vertical extent of buildings (hints of these cases are visible in **Fig. 3** and some example images in **Fig. 5**), which affects the computation of the Sky View Index as well, resulting in over- and under-estimations, respectively. A 5-image panorama exhibits a very strong correlation with its counterparts (factors) derived from fisheye imagery; in fact, it suggests that in many cases there is not much difference between using a panoramic image and fisheye, but not

always and especially for greenery the correspondence may vary and be less strong.

Finally, in **Fig. 7**, we reveal major discrepancies between the two. The two perspectives yield almost the same results across three indexes in cases of low-rise architecture and in homogeneous streetscapes. Further – unsurprising but worth mentioning – the discrepancies tend to be quite small or none where there are no particular features present (e.g. the GVI and GVF are both zero in areas without any or with very little greenery, thus, their mismatch is also around zero; see some settings in **Fig. 5**). The major discrepancies where the fisheye perspective provides substantially different results from panoramas are often locations with high-rise buildings and grasslands. That is, the fisheye perspective will capture the full vertical extent of a building (and any other tall feature such as overhead vegetation), while panoramas will provide a complete view of the ground-level where the grass is (which in the fisheye perspective is below the horizon and cut off).

4.2. Indexes from different variants of panoramas (Statistic B)

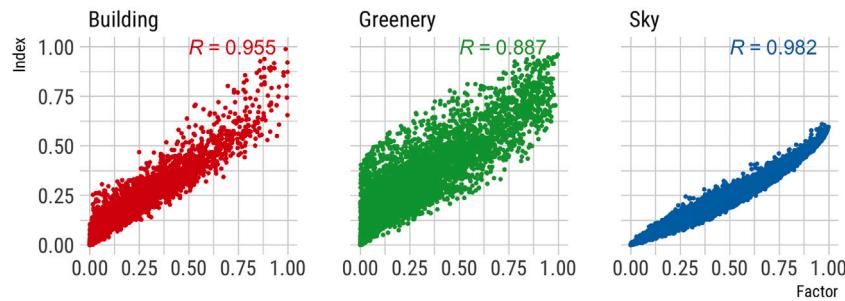
There is a strong relationship between the indexes derived from the 4- and 5-image panoramas (**Fig. 8**). If the goal of a study is to capture the entire streetscape including the overhead view, clearly a 5-image panorama will give more meaningful results. However, the results indicate that when one uses 4-image panoramas, the results are close to the 5-image approach.

4.3. Indexes and limited indexes (Statistic C)

In this statistic, we investigate the use of perspective images commonly found in crowdsourced imagery. To simplify the interpretation of the large number of combinations, we consider the most common ones in practice: images with 4/3 aspect ratio taken with wide and ultrawide cameras (FOVs of 70, 80, 90, and 120 degrees) and the ultrawide instance (FOV of 120 degrees) that has the aspect ratio of 16/9 to account for contributors that use dashcams and similar (denoted with an asterisk in the visuals).

Factor vs Index relationship

Panoramas with 5 images



Panoramas with 4 images

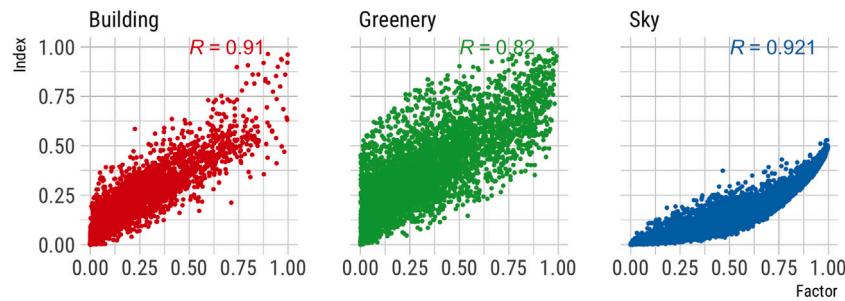


Fig. 6. Relationship between deriving the urban form from panoramas and fisheye imagery, with multiple scenarios on how panoramas are generated (Statistic A).

Low correspondence

Building



Greenery



Sky



High correspondence



Fig. 7. Examples of major (left) and minor (right) discrepancies between computing factors and indexes, from fisheye and panoramas, respectively (Statistic A). The images are sourced from Google Street View.

In Figs. 9 and 10 we give the results for all these combinations. The most important result is that in many scenarios, the measured indexes are very strongly correlated with metrics obtained from panoramas, suggesting that images with limited perspectives, which are considered to hamper their adoption in the research community, may actually be useful for inferring the urban form and come close to the metrics estimated from panoramic imagery. The strongest correlation is exhibited by ultrawide-angle longitudinal recordings, e.g. images with a FOV of 120 degrees taken in the direction of driving have correlation coefficients of 0.942, 0.950, and 0.927 with those obtained from panoramas, for the BVI, GVI, and SVI, respectively. Conveniently, this appears to be the most common provenance found in practice in crowdsourced street-level imagery.

The results are different for all the considered characteristics of equipment. As it was the case in the previous statistics, the results also differ for each element of the urban form, but they are in line with each other across the characteristics such as field of view and direction. In the continuation, we dissect them.

Starting with the direction — there is a significant influence in which direction the camera is pointing, confirming the findings of Kim et al. (2021). In the cases of all elements, the images tend to give more truthful results if they point front (0) or back (180) rather than to the sides (90 and 270). Moreover, while front and back views tend to achieve almost the same results, the left or right direction (90 vs 270 degrees) of the side views has a notable difference on the results. The diagonal views (45, 135, 225, and 315 degrees) do not bring much value, as they have only a marginal improvement over lateral

Comparison of indexes computed from different panoramas

Mean Absolute Error

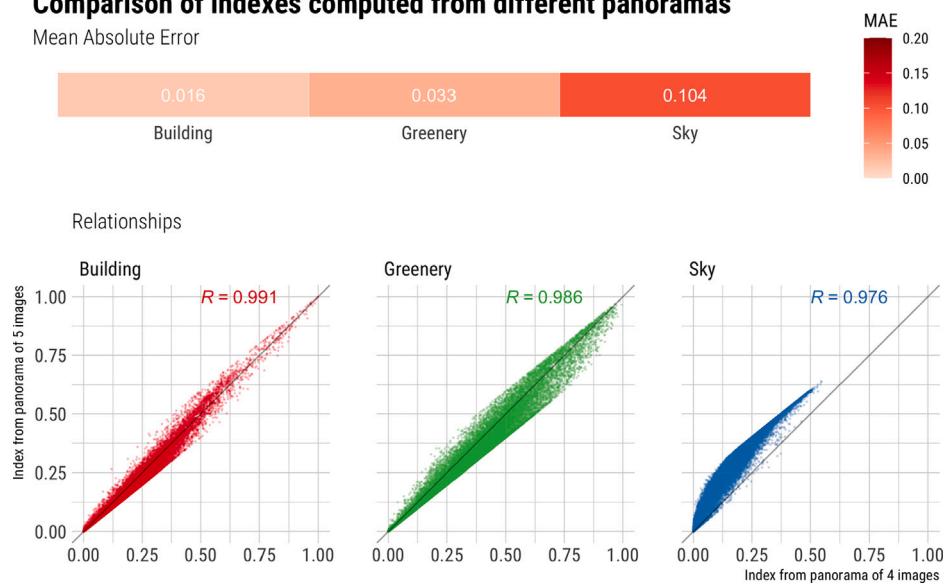


Fig. 8. Relationships between indexes derived from different panoramas (Statistic B).

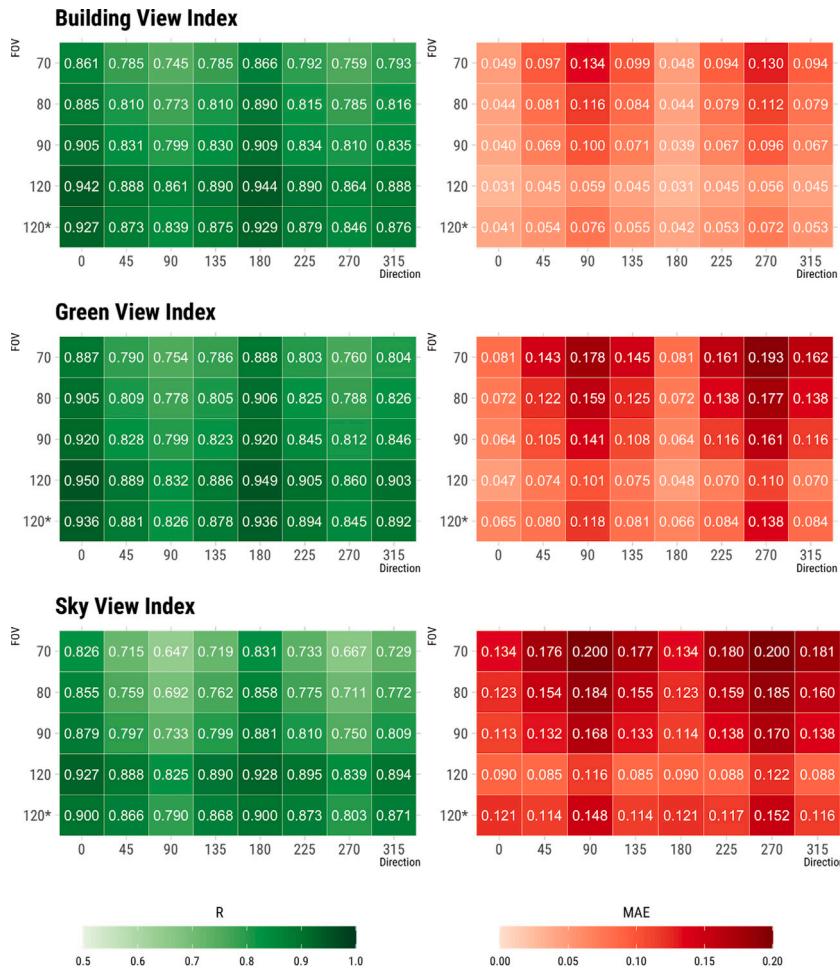


Fig. 9. Correspondence of a single street view with a full panorama (5 stitched images) for different urban form metrics (Statistic C).

views, and cannot compete with front and back views. Thus, starting to collect such imagery in practice would not be helpful, but at the same time such imagery, which is sometimes found in practice, would not compromise estimating these metrics.

The FOV has a notable influence on the results (see Fig. 10). However, while it is not a surprise that images obtained with ultrawide cameras will give better results than those with a narrower angle, the results suggest that it is not as important as the direction of view. For

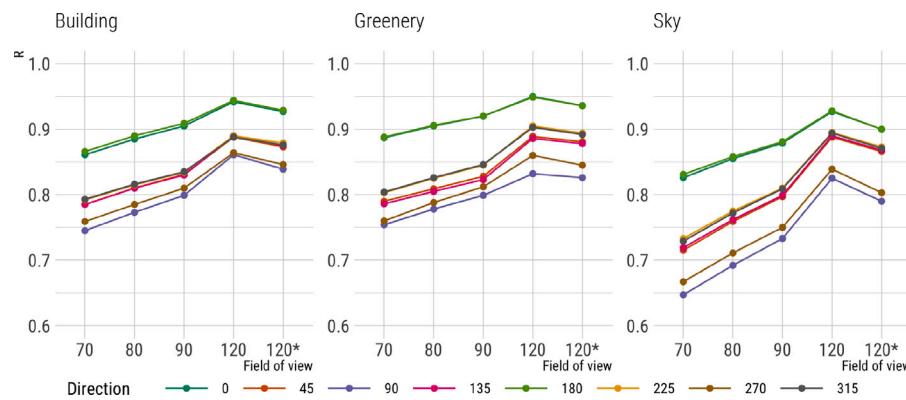


Fig. 10. The performance of using perspective (single) street view imagery for inferring the urban form depends on the camera type used and direction of the view (Statistic C). The asterisk indicates the images with the wider aspect ratio, which is frequent for ultrawide angle cameras.

example, a wide-angle camera (e.g. FOV = 80) pointing longitudinally (direction = 0 or 180) will give more truthful results ($R = 0.89$, for BVI) than an ultrawide-angle camera (e.g. FOV = 120) directed to the side (e.g. direction = 90) ($R = 0.86$), despite the much wider view offered by the latter.

As another outlook on the results, we computed the proportion of imagery that has an index within 0.05 of the one computed from a 4-image panorama, as the standard way to compute such metrics. For FOV = 120 and direction 0 (one of the most common combination of parameters found in practice), for building, greenery, and sky, in 78%, 62%, and 72% of cases, such images yield indexes that are within 0.05 in value from those derived from 4-image panoramas. Such conclusion uncovers the tremendous potential of freely available crowdsourced imagery collected with low-cost cameras. Further, the results suggest that in most studies researchers may be using much more data than they need. Instead of using several images from stitched panoramas in Google Street View, they may reduce data sourcing to single images, dramatically simplifying workflows and reducing computational complexity while maintaining comparable accuracy.

The results so far have been analysed for all five cities together. For further analysis, the discrepancies are grouped by city and urban form (Fig. 11). The middle plot illustrates the difference between the limited indexes from the panoramic indexes for each city (the top part of the figure shows the metrics together for reference). It suggests that there are some differences among the cities. While indexes related to buildings and greenery are comparable among cities, they are a bit more pronounced for the Sky View Index, with each city having its set of results. The differences are not substantial, thus, we conclude that our findings apply worldwide to most if not all cities.

Next, we focus on decomposing the results by the urban form, focusing on the sky, which exhibited larger differences than the other two elements in the previous paragraph. We bin the locations into 10 quantiles by the value of the SVF of each, and analyse the errors for each quantile (bottom plot in Fig. 11). The results are analysed in isolation for multiple FOVs and directions, and exhibit different behaviour. It can be concluded that the urban form has some influence on the results, and the exact statistics will depend for a particular city. Further, the results among the cities may not only be due to the different urban form, but also the different approaches to collect data in them. For example, in some cities, Google Street View collects data from parks and hinterlands (Chen and Biljecki, 2023), which have a significantly different array of indexes and factors, but such acquisition may not be conducted in some other cities, biasing the results.

Finally, in Fig. 12, we give examples for very accurate and least accurate estimations, to give an understanding of scenarios that are suitable and less suitable for imagery with limited views. In these examples, we focus on forward-facing perspective views (direction = 0) at the field of view of 90 degrees and aspect ratio 4:3. Analysing

the differences in the indexes computed from such perspective and panoramic imagery, we uncover that many of the major discrepancies in which the perspective images are not representative of the surrounding (panoramic) environment are those collected at the end of a street where a particular element (building, greenery) dominates the image. Further examples include images collected on underpasses (specifically for the Sky View Index), bridges over water (i.e. having an open sky around but not in front), hilly areas, and situations with large objects on the sides that are missed in front-facing images such as sideways trees outside the view (for similar scenarios see also the top left and middle left settings in the image as examples for greenery and building, respectively). In contrast, locations where a single image is representative of the entire surrounding are typically homogeneous and symmetric landscapes or those that are diverse but where by chance the metrics correspond in both perspectives (e.g. in the figure, the middle right example indicates a heterogeneous and asymmetric setting where the GVI is 0.254 both when computed from a single and a panoramic image).

4.4. Factors and limited indexes (Statistic D)

Similarly as the previous statistic, we analyse the performance of limited perspectives with the data that is usually obtained from commercial services, this time with the factors, which are derived from fisheye images. Fig. 13 suggests that albeit the correlations are weaker than it is the case with panoramas (Statistic C), they are still very strong for certain combinations, such as an ultrawide angle front-facing image. We hope that these results will motivate future studies to consider crowdsourced imagery in research that is traditionally engaging commercial (panoramic) data.

4.5. Different aspect ratios (Statistic E)

Finally, we examine the influence of the different aspect ratios. For limited perspectives, a hint of this influence is given in the previous two statistics (see Figs. 9 and 13), in which a certain scenario (FOV = 120) is compared between 4/3 and 16/9 perspectives, which are found across different cameras used in practice, and the results show an almost negligible influence. The influence of the aspect ratio is elaborated further here.

We start with the different aspect ratios (4/3 and 16/9, see Fig. 4 for their comparison) of perspective images. For BVI and GVI, in all cases of directions, the correlation coefficient is above 0.99 and the errors are small (MAE of up to 0.04), and in the shadow of the previous characteristics that have a more profound influence. Even in the least favourable combination (direction = 90 and FOV = 70) the correlation coefficient is still very strong ($R = 0.975$), not necessitating

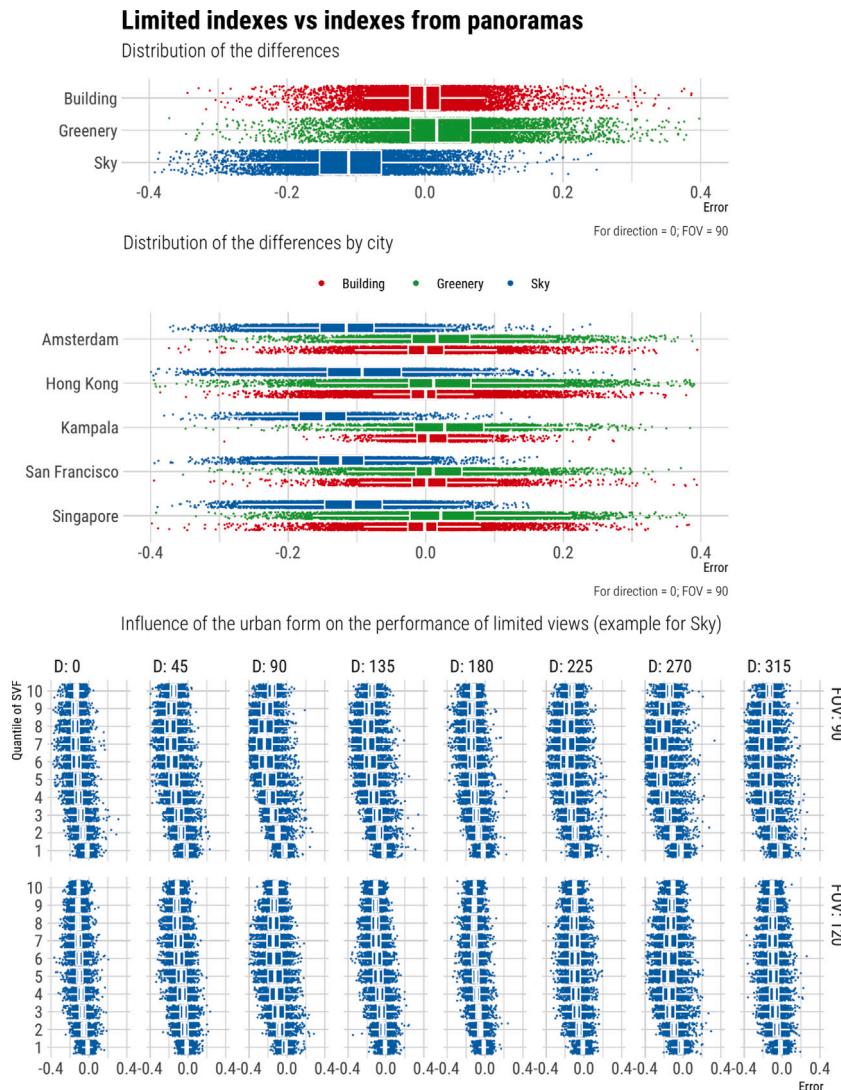


Fig. 11. The results (Statistic C in this case) are influenced not only by the different element of the urban form (i.e. building), but also city and its urban morphology.

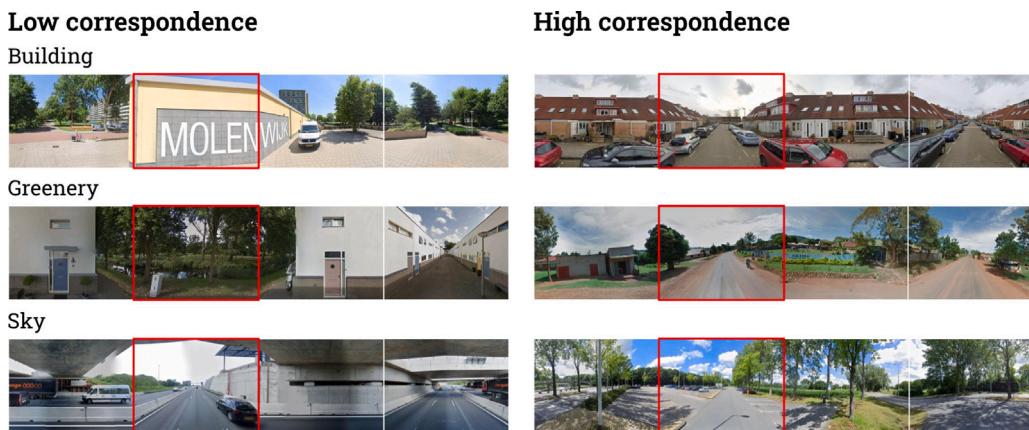


Fig. 12. Examples of large errors (left) and accurate instances (right) when computing indexes from panoramas and limited views (Statistic C). The forward-facing image, from which the limited indexes are computed, is the second one in each row. The images are sourced from Google Street View.

further research on this aspect that appears to have a consistently small influence.

Next, we investigate the relationship between regular 4-image panoramas (computed from perspective images with an aspect ratio



Fig. 13. Correspondence of a single street view with a fisheye for different urban form metrics (Statistic D).

of 1/1) with those computed with narrower aspect ratios such as 4/3 and 16/9, which are found in some instances. The correlations for 4/3 panoramas are 0.994, 0.994, and 0.990, for building, green, and sky view indexes, respectively. The results for 16/9 variants are: 0.978, 0.977, 0.961. Unless there is a strong reason (such as perception studies that would mimic the human vision, e.g. see Walker et al., 1990) to extract narrow panoramas, it is better to follow usual practices with 1/1 aspect ratios.

5. Aggregating the measurements: urban scale analysis of the sky view

The point-level results are often aggregated to understand the spatial distribution of the urban form in cities and the relationships with other variables (e.g. socio-economic indicators) (Demuzere et al., 2019; Gong et al., 2019a; Liao et al., 2021; Liu et al., 2021; Sun et al., 2021a,b; Verma et al., 2020; Wang et al., 2018; Wu et al., 2020b; Zhang et al., 2019) (Fig. 1 bottom right), e.g. at the scales of blocks (Yao et al., 2021), subdistricts (Lu et al., 2019), districts (Seiferling et al., 2017), and city-scale (Middel et al., 2018). Researchers also aggregate data at

regular grids (Chen et al., 2022c; Gong et al., 2019a; Dirksen et al., 2019; Verma et al., 2020; Ito and Biljecki, 2021). Such spatial aggregations are predominantly focused on the mean value of an indicator at all imaged points in a zone (e.g. cell of a grid), but some researchers compute also the variation therein, such as standard deviation (Middel et al., 2018).

As another angle of interpreting the results, we aggregated them in a city-scale analysis and compared the results to gauge the performance of single-frame images for this purpose. In this analysis, we focused on Singapore (a morphologically diverse city; see Fig. 5), the sky view, and the perspective image that is commonly found in crowdsourced imagery and when generating panoramas (direction of zero and FOV of 90 degrees), and compared the results to values obtained from fisheye and panorama (Statistics C and D; see Sections 4.3 and 4.4).

We aggregated the results according to a regular grid system (H3, a geospatial indexing system that partitions the world into hexagonal cells³) that is composed of hexagons, and it has been adopted in various disciplines (Bousquin, 2021; Woźniak and Szymański, 2021; Zhao et al., 2023; Liang et al., 2023). Each cell has the area of about 0.10 sq. km. and it contains 51 panoramas on average.

To put the results in context and enable straightforward comparison, for each approach, we compute the standard score (Z-score) of each cell — the number of standard deviations a value is above or below the mean value at the city scale. The resulting two maps (Fig. 14) are very similar to each other. The correlations between the two are very strong (0.951), with the one to a 4-image panorama even stronger — 0.981. Computations for building and greenery exhibit a similarly strong relationship. This finding indicates that non-panoramic SVI, typically found in platforms curating user-generated data, may be used for some use cases without a noticeable loss in accuracy. It also suggests that when the results are spatially aggregated, they tend to be more accurate, diminishing the image-level uncertainty of different perspectives (e.g. the R for that scenario at the image-level, i.e. Statistic D, is noticeably smaller — 0.84; see Fig. 13).

6. Discussion

6.1. Main insights and recommendations

The previous section has presented a comprehensive series of novel and insightful results, which will be elaborated further and untangled in this section.

When it comes to existing practices using panoramic imagery, it seems that there is a general correspondence in the results among the different approaches, with some exceptions in specific settings (cf. Fig. 12), and with varying magnitudes (e.g. for green view, the correlation is less strong than for sky view).

Because in certain cases single images come very close in deriving the same results, instead of fetching and processing numerous images, perhaps the process could be reduced considerably by deriving only the one with the most favourable combination of the direction and field of view, which may simplify data collection and reduce the computational complexity of their processing.

This result is also favourable for crowdsourced street-level imagery in which non-panoramic imagery still accounts for the vast majority of imagery found in mainstream platforms such as Mapillary and KartaView. Do we therefore suggest to replace commercial data sources with crowdsourced alternatives in future studies? Yes and no. It is certainly good news that in the majority of cases, a limited view may be sufficient to infer the urban form, potentially unlocking the value of billions of user-generated free images, some of which are available in locations where commercial services provide no coverage (León-Sánchez et al., 2022). But issues remain, and further investigations are

³ <https://h3geo.org/>

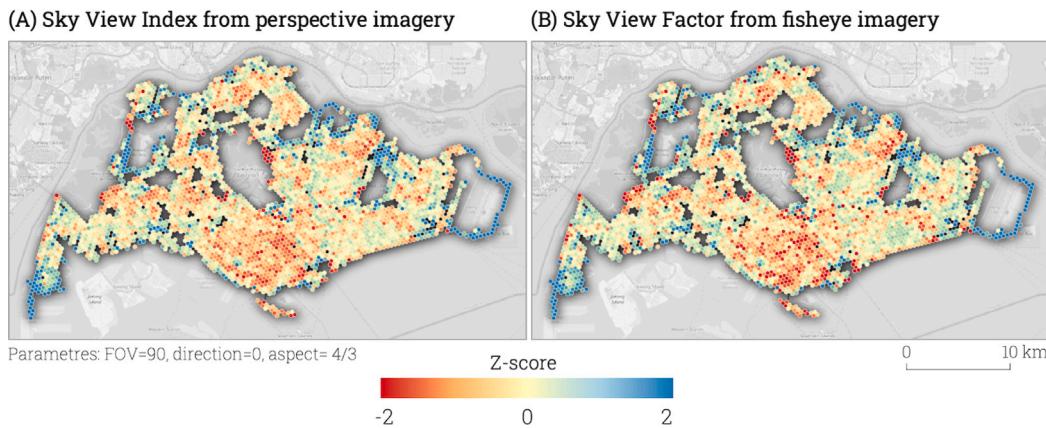


Fig. 14. Exposure to sky, at the urban scale and aggregated in a regular grid, computed from two entirely different perspectives (limited image akin to one taken from a smartphone vs panoramic fisheye image) have a very high degree of correspondence. Basemap: (c) OpenStreetMap contributors.

necessary. For example, quality of volunteered SVI is heterogeneous, primarily in terms of completeness and image quality (Quinn and León, 2019; Hou and Biljecki, 2022), aspects that are out of scope of this paper.

The results also depend on the urban form element that is measured. Estimating the sky view seems to be most sensitive one (cf. Fig. 10), thus caution should be exercised in morphologies with street canyons and researchers should be aware of the shortcomings of certain configurations.

As it is the case with other error propagation studies, the interpretation of the results much depends on what the data is used for. For example, an error of 0.05 in the computation of the Sky View Factor may not be a big deal in the general characterisation of the urban form, but may be influential in some other use cases. Experiments for specific use cases (e.g. using the Green View Index for walkability studies) need to be conducted to understand sensitivity to errors and establish application-specific data requirements.

6.2. Influence of the driving direction

The cities we focused on have different traffic rules. That means that in bidirectional traffic, a city that drives on the left hand side of the road, e.g. Singapore, the direction 90 degrees (on the right side of the car), may capture a wider spatial extent than the one that drives on the right (e.g. NYC), in which case the view in the same direction is closer to the street-side (see the top part of Fig. 15). We investigate this matter and group the results by this aspect — it appears that there is a notable influence of this aspect (see the middle of Fig. 15), which we raise for the first time in the SVI community, and which may warrant further investigations considering the growing volume of comparative studies involving multiple cities, which may involve contrasting traffic rules. The analysis focuses on the performance of the perspective images versus panoramas (Statistic C). For longitudinal views (directions of 0 and 180), there is no significant difference in the results, in contrast with the situation for lateral directions (90 and 270). However, in all cases, left-hand traffic cities have better results, suggesting that the influence may be biased by the physical difference among cities (e.g. see the previously discussed Fig. 11).

The different elements are affected differently — building and greenery exhibit mostly similar results, but they are different from sky. In the case of the Sky View Index, in almost all directions there is no difference for left- and right-hand traffic, including for the direction of the camera on the left (270), while a notable difference is found in the opposite side direction (90). The illustration and the examples from crowdsourced imagery in the figure affirm these results by demonstrating that the same direction from different sides of the road does not reveal sky and buildings entirely, and it is heavily influenced by the

physical configuration of the streetscape. Further aspects that influence such results may be the share of one-way or bidirectional roads and their width.

This aspect remains inconclusive and open for debate, but necessary to bring in this paper as it was not raised in literature so far and it may be relevant for studies involving multiple cities. It is a conjecture that will require further analysis, as it is not possible to isolate the effect of the side of the traffic with the influence of the urban form of the city and other influencing factors. For example, detailed geospatial analysis examining the influence and contribution of the e.g. the number of lanes of a road, directionality, and size of street/building setbacks and buffers, and accounting for the urban form (e.g. comparing only clusters of similar streets in different cities) will reveal the interaction of these aspects precisely.

6.3. Limitations and directions for future work

A part of the limitations of this study are generally caused by SVI and their analysis, and this work is driven by practices of typical studies in the field, mirroring their limitations. SVI is a boon and a bane for measuring the urban form metrics — its advantage is that it can be used to calculate a holistic (or ‘all-inclusive’) SVF, and at the same time its disadvantage is that it does not differentiate between the contributions of various types of features in the built environment (e.g. the SVF taking only buildings into account). Thus, it may be used in conjunction with other spatial data for a more insightful understanding in the SVF (e.g. a 3D building model may help to remove the contribution of buildings to isolate the SVF provided solely by vegetation) (Li et al., 2018). Further, imagery is often available only on roads (Chen and Biljecki, 2023), thus, the estimation of the indexes is possible only within their confines.

Following the practices of uncertainty propagation studies in GI-Science, we have relied on simulating perspective views from panoramas to establish a controlled environment, which enabled us to set a variety of parameters and isolate the uncertainty induced by a specific aspect such as the limited field of view. In this work, it might have been possible to compare panoramas from a service such as Google Street View with crowdsourced perspective images from Mapillary or KartaView in the same locations. However, that may not be a viable direction as in practice, images from different platforms will mismatch in many aspects, such as exact location of the panorama, time of acquisition, and quality of imagery. All these external factors will impact the results of the estimations, and the differences between two data sources will be a composite of all kinds of aspects and errors, e.g. different camera heights (Yan and Huang, 2022; Aikoh et al., 2023), heterogeneous completeness (Hou and Biljecki, 2022; Quinn and León, 2019; Seto and Nishimura, 2022), seasonality (Chiang et al., 2023; Han et al., 2023), and further inconsistencies such as varying light

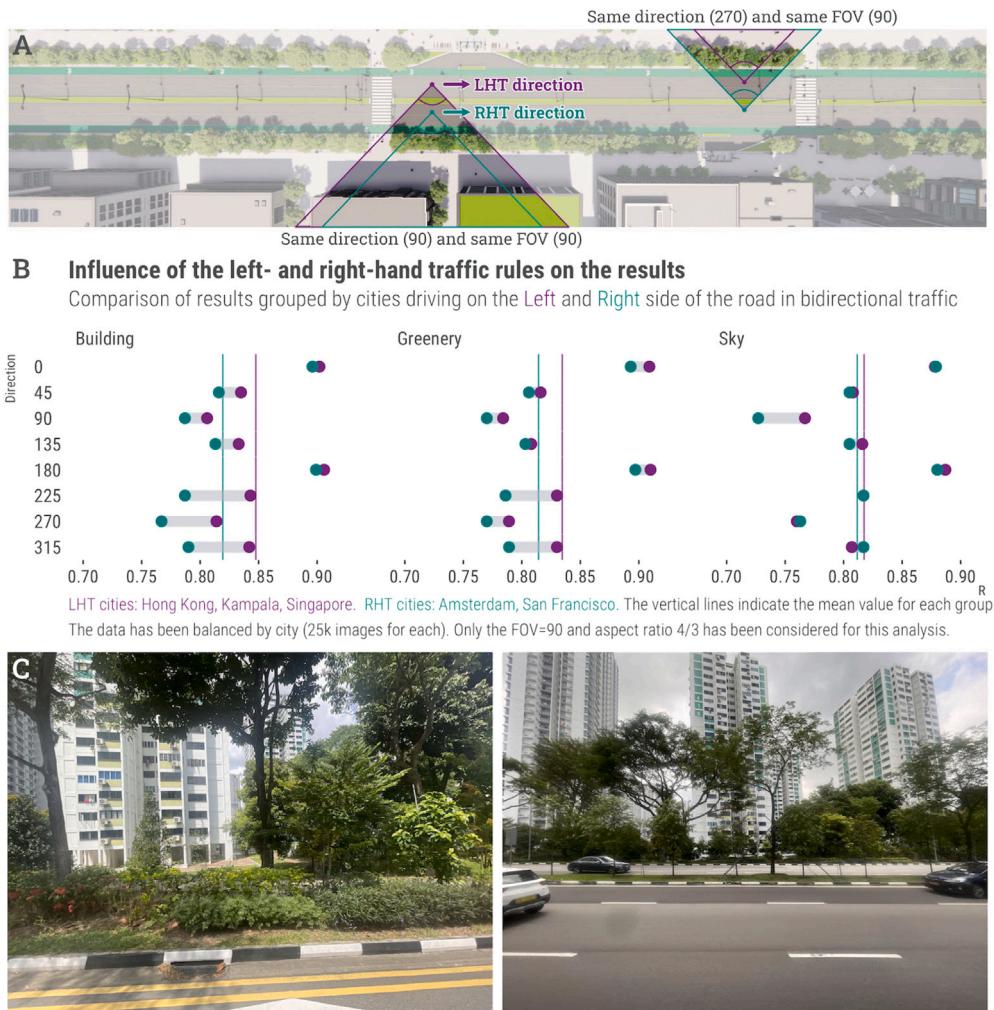


Fig. 15. Due to the different scope of view (A), the traffic rules may have an influence on the results (B) (demonstration on the Statistic C). The bottom row (C) shows an example from Mapillary in Singapore with two perspective images taken with the same equipment pointing in the same direction but different sides of roads. Precise conclusions are not possible without further analysis, and they depend on many factors such as road characteristics and the local urban form.

conditions and attribute errors (Zheng and Amemiya, 2023), making it difficult to isolate the influence of particular aspects such as field of view and direction, which we largely focused on this paper. Thus, both a strength and limitation of this work is that we use simulated data and focus on the aspect of camera properties and acquisition practices, and their influence in inferring the urban form. Future research will account for other aspects such as completeness and positional resolution of imagery, but likewise, this might be possible only with processed and controlled data.

As in all studies employing SVI, the segmentation that we used is not perfect (see Fig. 1 for examples of small typical misclassifications in the images), thus, the results are influenced to some degree by such uncertainty, which cannot be quantified due to cultural, architectural, and geographical differences (Inoue et al., 2022a). While we employ the same model everywhere, which may help minimise the influence of model errors, it remains unknown how much these imperfections affect the results. Further, we noticed that in some instances, certain houses are blurred for privacy reasons. Some of these are blurred with a soft filter (see an example in the upper sample for San Francisco in Fig. 5, at the right), resembling the sky and morphing the building into it, thus, the model has misclassified that part of the image, an error that propagates to the computation of the indexes. However, this is in line with all studies in the field, which have a degree of uncertainty in the segmentation process, and such errors should not have a major

influence on our findings, especially when dealing with millions of images.

When designing the method, we explored a crowdsourced platform to find the most common practices of capturing street-level imagery around the world (e.g. field of view, direction). The parameters we have established (Section 3) account for the vast majority of data, and they are expected to remain the same in the foreseeable future. However, it is not possible to capture all camera parameters and settings that are used in practice.

In our experiments, we did not take into account errors such as significant vertical tilt of a camera. For example, a dashcam may not always be pointing at the horizon. For future work, it would be beneficial to investigate whether a dashcam pointing upwards may give better results by encompassing the vertical extent of urban canyons.

Our work focuses on data from GSV, which is overwhelmingly collected with cars on driveable roads, which is also in line with the content of crowdsourced services such as Mapillary. However, an increasingly evident strength of crowdsourced services are off-road images and other transportation modes, e.g. those collected by pedestrians on sidewalks and in pedestrian zones, and cyclists on dedicated bike lanes, with its importance heralded by recent studies (Chen and Biljecki, 2023). It might be beneficial to take into account such modes, and perform a comparative study between imagery collected on different instances on the same street (e.g. sidewalk vs roadway).

Finally, in our work, we have focused on probably the most common use of SVI — inferring the urban form and greenery with image segmentation. However, the results may differ for other use cases, such as mapping street furniture and infrastructure (Zhong et al., 2021; Zhou et al., 2022b) and extracting building information (Ning et al., 2022b; Yan and Huang, 2022; Sun et al., 2022). For example, the latter may be significantly more sensitive to the lack of panoramic images because not all buildings are clearly visible from single-frame imagery pointing in the direction of driving (Biljecki et al., 2023).

7. Conclusion

Many research groups worldwide practising various urban disciplines rely on street view imagery, an emerging but widely available source of geospatial data and valuable means for urban sensing. Much of them use numerous images of streetscapes to infer the urban form and greenery at a high spatial resolution and accuracy. Examining the methodologies employed, we encounter disparate approaches among different research domains computing similar indicators, an aspect that may cause inconsistencies and one that warrants investigations. Further, street-level imagery became an important component of the landscape of crowdsourced geospatial data, as billions of images have been collected and shared by volunteers around the world, potentially competing with commercial sources, which have been used extensively. As the urban data research ecosystem is shifting towards openness in terms of software and data (Yap et al., 2022b; Fleischmann, 2019; Boeing, 2017; Rey, 2019; Liu et al., 2015; Benitez-Paez et al., 2018), it is important to intensify investigating their reliability and give them consideration for environmental and urban studies. Our work represents a significant initial stride towards achieving this goal by suggesting that the limited views of single-frame cameras and the use of consumer equipment may not represent a hindrance, but it recommends further investigations by incorporating other quality aspects that typically burden crowdsourced geographic information, such as heterogeneous completeness.

We think that this research is especially relevant given the increasing restrictions of commercial sources of SVI — urban data researchers may soon be impelled to divorce from their long dependency on Google Street View and Baidu Maps, in favour of crowdsourced imagery released at a free licence and involving citizens. However, that will come at the expense of forgoing data that is acquired in a standardised manner and one that gives an immersive view into streetscapes and the built environment thanks to panoramic imagery with a great regard to high quality, consistent weather conditions, and metadata. Thus, this research is timely to provide an understanding of the usability of limited imagery and set the scene for further research in understanding the usability of crowdsourced street-level imagery obtained from volunteers with consumer devices.

For the first time, this study gives a comprehensive set of insights into the influence of the parameters and practices for computing several urban form indicators such as the Sky View Factor, Green View Index, and Building View Index. There are many takeaways from our study, most importantly across three matters:

- *Panoramic imagery.* While bringing clarity to the metrics and investigating their interaction, a literature review suggests that not all researchers follow the same approaches, which is also aggravated by terminological inconsistencies. We establish that there is a difference in the results based on how the same dataset is used, but in most cases it is minor. However, researchers should be careful when exercising specific use cases, as some of them have associated limitations (e.g. using the fisheye image pointing upwards appears to be the way to capture the open sky in urban canyons, but such perspective may not be suitable for estimating the total amount of greenery as it tends to overlook grasslands). Further, the results suggest that studies may be using more images than necessary, and our findings may help simplify such studies.

- *Single imagery.* We quantify the performance of using single (limited) images in comparison with traditionally used panoramic imagery, and demonstrate that they can present a reasonable replacement for them, and may satisfy many use cases. In this aspect, we conclude that *narrow views are able to give broad perspectives*. We hope that this finding will simplify workflows by reducing the volume of imagery used, and even spur the community to consider the hitherto overlooked user-generated images in crowdsourced platforms, and take advantage of their free availability and flexibility, and the fact that they are available in various places where commercial counterparts are not present or have outdated photos. The results indicate that all the investigated camera characteristics play a role but with varying degrees — the direction of the image is more important than the field of view and aspect ratio of the camera. Further, when the data is spatially aggregated (e.g. at district scale), the uncertainty tends to be very low.

- *External factors.* The computation of indicators is somewhat affected by the urban form, city, perspective, and even traffic rules, but the results can be considered to apply globally. Understanding such relationships is complex and it is a viable direction for future work.

As autonomous vehicles, which rely on a variety of sensors such as cameras, become more common, it is possible that we may be able to use the collected data as a source of street-level imagery, and thus, contribute to the growing volume of the data shared openly and their quality.

The broader perspective of our work is the contribution to uncovering the potential of crowdsourced street-level imagery — the visual frontier of volunteered geographic information, and lead to further investigations, such as regarding the heterogeneous quality and completeness of data in crowdsourced data, and the development of benchmarks to identify the most optimal parameters of street view imagery while minimising computational complexity.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

We appreciate the constructive and encouraging comments by the editors and the reviewers, which have improved the paper. We thank the colleagues at the NUS Urban Analytics Lab for the discussions. Winston Yap, Koichi Ito, and Yangyang Li contributed with suggestions, while April Zhu is gratefully acknowledged for the help with the first illustration. This research is part of the projects (i) Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start Up Grant R-295-000-171-133; (ii) The Technical Key Project of Shenzhen Science and Technology Innovation Commission Grant JSGG20201103093401004; and (iii) Multi-scale Digital Twins for the Urban Environment: From Heartbeats to Cities, which is supported by the Singapore Ministry of Education Academic Research Fund Tier 1.

References

- Aikoh, T., Homma, R., Abe, Y., 2023. Comparing conventional manual measurement of the Green View Index with modern automatic methods using Google Street View and semantic segmentation. *Urban For. Urban Green.* 80, 127845. <http://dx.doi.org/10.1016/j.ufug.2023.127845>.
- Ali-bey, A., Chaib-draa, B., Giguère, P., 2022. GSV-cities: Toward appropriate supervised visual place recognition. *Neurocomputing* <http://dx.doi.org/10.1016/j.neucom.2022.09.127>.
- Anderson, J., Sarkar, D., Palen, L., 2019. Corporate editors in the evolving landscape of OpenStreetMap. *ISPRS Int. J. Geo-Inf.* 8 (5), 232. <http://dx.doi.org/10.3390/ijgi8050232>.
- Ao, Y., Wang, J., Zhou, M., Lindenbergh, R.C., Yang, M.Y., 2019. Fully convolutional networks for street furniture identification in panorama images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLII-2/W13, 13–20. <http://dx.doi.org/10.5194/isprs-archives-xlii-2-w13-13-2019>.
- Basu, R., Sevtuk, A., 2022. How do street attributes affect willingness-to-walk? City-wide pedestrian route choice analysis using big data from Boston and San Francisco. *Transp. Res.* A 163, 1–19. <http://dx.doi.org/10.1016/j.tra.2022.06.007>.
- Beaucamp, B., Leduc, T., Tourre, V., Servières, M., 2022. The whole is other than the sum of its parts: Sensibility analysis of 360 urban image splitting. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* V-4-2022, 33–40. <http://dx.doi.org/10.5194/isprs-annals-v-4-2022-33-2022>.
- Benitez-Paez, F., Comber, A., Trilles, S., Huerta, J., 2018. Creating a conceptual framework to improve the re-usability of open geographic data in cities. *Trans. GIS* 22 (3), 806–822. <http://dx.doi.org/10.1111/tgis.12449>.
- Bianconi, F., Filippucci, M., Ceccaroni, S., Seccaroni, M., 2022. The value of deep learning for landscape representation comparison between segmentation images maps and GIS. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLIII-B4-2022, 487–492. <http://dx.doi.org/10.5194/isprs-archives-xliii-b4-2022-487-2022>.
- Biljecki, F., Chow, Y.S., 2022. Global building morphology indicators. *Comput. Environ. Urban Syst.* 95, 101809. <http://dx.doi.org/10.1016/j.compenvurbsys.2022.101809>.
- Biljecki, F., Chow, Y.S., Lee, K., 2023. Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes. *Build. Environ.* 237, 110295. <http://dx.doi.org/10.1016/j.buildenv.2023.110295>.
- Biljecki, F., Ito, K., 2021. Street view imagery in urban analytics and GIS: A review. *Landscape Urban Plan.* 215, 104217. <http://dx.doi.org/10.1016/j.landurbplan.2021.104217>.
- Bochkarev, K., Smirnov, E., 2019. Detecting advertising on building façades with computer vision. *Procedia Comput. Sci.* 156, 338–346. <http://dx.doi.org/10.1016/j.procs.2019.08.210>.
- Boeing, G., 2017. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.* 65, 126–139. <http://dx.doi.org/10.1016/j.compenvurbsys.2017.05.004>.
- Bousquin, J., 2021. Discrete global grid systems as scalable geospatial frameworks for characterizing coastal environments.. *Environ. Model. Softw.* 146, 1–14. <http://dx.doi.org/10.1016/j.envsoft.2021.105210>.
- Byun, G., Kim, Y., 2022. A street-view-based method to detect urban growth and decline: A case study of Midtown in Detroit, Michigan, USA. *PLoS One* 17 (2), e0263775. <http://dx.doi.org/10.1371/journal.pone.0263775>.
- Chapman, L., Thornes, J.E., Bradley, A.V., 2002. Sky-view factor approximation using GPS receivers. *Int. J. Climatol.* 22 (5), 615–621. <http://dx.doi.org/10.1002/joc.649>.
- Chen, S., Bao, Z., Chen, J., Yang, L., Lou, V., 2022a. Sustainable built environment for facilitating public health of older adults: Evidence from Hong Kong. *Sustain. Dev.* 30, 1086–1098. <http://dx.doi.org/10.1002/sd.2303>.
- Chen, S., Biljecki, F., 2023. Automatic assessment of public open spaces using street view imagery. *Cities* 137, 104329. <http://dx.doi.org/10.1016/j.cities.2023.104329>.
- Chen, J., Chen, L., Li, Y., Zhang, W., Long, Y., 2022b. Measuring physical disorder in urban street spaces: A large-scale analysis using street view images and deep learning. *Ann. Am. Assoc. Geogr.* 1–19. <http://dx.doi.org/10.1080/24694452.2022.2114417>.
- Chen, L., Lu, Y., Ye, Y., Xiao, Y., Yang, L., 2022c. Examining the association between the built environment and pedestrian volume using street view images. *Cities* 103734. <http://dx.doi.org/10.1016/j.cities.2022.103734>.
- Chen, X., Meng, Q., Hu, D., Zhang, L., Yang, J., 2019. Evaluating greenery around streets using Baidu panoramic street view images and the panoramic Green view index. *Forests* 10 (12), 1109. <http://dx.doi.org/10.3390/f10121109>.
- Chen, S., Wong, N.H., Ignatius, M., Zhang, W., He, Y., Yu, Z., Hii, D.J.C., 2022d. ATLAS: Software for analysing the relationship between urban microclimate and urban morphology in a tropical city. *Build. Environ.* 208, 108591. <http://dx.doi.org/10.1016/j.buildenv.2021.108591>.
- Chen, L., Yao, X., Liu, Y., Zhu, Y., Chen, W., Zhao, X., Chi, T., 2020a. Measuring impacts of urban environmental elements on housing prices based on multisource data—A case study of Shanghai, China. *ISPRS Int. J. Geo-Inf.* 9 (2), 106. <http://dx.doi.org/10.3390/ijgi9020106>.
- Chen, S., Zhang, W., Wong, N.H., Ignatius, M., 2020b. Combining CityGML files and data-driven models for microclimate simulations in a tropical city. *Build. Environ.* 185, 107314. <http://dx.doi.org/10.1016/j.buildenv.2020.107314>.
- Chen, J., Zhou, C., Li, F., 2020c. Quantifying the green view indicator for assessing urban greening quality: An analysis based on internet-crawling street view data. *Ecol. Indic.* 113, 106192. <http://dx.doi.org/10.1016/j.ecolind.2020.106192>.
- Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European Conference on Computer Vision. ECCV*, pp. 801–818.
- Chiang, Y.C., Liu, H.H., Li, D., Ho, L.C., 2023. Quantification through deep learning of sky view factor and greenery on urban streets during hot and cool seasons. *Landscape Urban Plan.* 232, 104679. <http://dx.doi.org/10.1016/j.landurbplan.2022.104679>.
- Chu, T., Chen, Y., Su, H., Xu, Z., Chen, G., Zhou, A., 2022. A news picture geo-localization pipeline based on deep learning and street view images. *Int. J. Digit. Earth* 15 (1), 1485–1505. <http://dx.doi.org/10.1080/17538947.2022.2121437>.
- Cinnamon, J., Gaffney, A., 2021. Do-it-yourself street views and the urban imaginary of Google street view. *J. Urban Technol.* 1–22. <http://dx.doi.org/10.1080/10630732.2021.1910467>.
- d'Andrimont, R., Yordanov, M., Lemoine, G., Yoong, J., Nikel, K., Velde, M.v.d., 2018. Crowdsourced street-level imagery as a potential source of in-situ data for crop monitoring. *Land* 7 (4), 127. <http://dx.doi.org/10.3390/land7040127>.
- Demuzere, M., Bechtel, B., Middel, A., Mills, G., 2019. Mapping Europe into local climate zones. In: Mourshed, M. (Ed.), *PLoS One* 14 (4), e0214474. <http://dx.doi.org/10.1371/journal.pone.0214474>.
- Deng, M., Yang, W., Chen, C., Wu, Z., Liu, Y., Xiang, C., 2021. Street-level solar radiation mapping and patterns profiling using Baidu street view images. *Sustainable Cities Soc.* 75, 103289. <http://dx.doi.org/10.1016/j.scs.2021.103289>.
- Ding, X., Fan, H., Gong, J., 2021. Towards generating network of bikeways from Mapillary data. *Comput. Environ. Urban Syst.* 88, 101632. <http://dx.doi.org/10.1016/j.compenvurbsys.2021.101632>.
- Dirksen, M., Ronda, R.J., Theeuwes, N.E., Pagani, G.A., 2019. Sky view factor calculations and its application in urban heat island studies. *Urban Clim.* 30, 100498. <http://dx.doi.org/10.1016/j.ulclim.2019.100498>.
- Esch, T., Brzoska, E., Dech, S., Leutner, B., Palacios-Lopez, D., Metz-Marconcini, A., Marconcini, M., Roth, A., Zeidler, J., 2022. World settlement footprint 3D - A first three-dimensional survey of the global building stock. *Remote Sens. Environ.* 270, 112877. <http://dx.doi.org/10.1016/j.rse.2021.112877>.
- Fan, H., Kong, G., Zhang, C., 2021. An interactive platform for low-cost 3D building modeling from VGI data using convolutional neural network. *Big Earth Data* 5 (1), 49–65. <http://dx.doi.org/10.1080/20964471.2021.1886391>.
- Feng, Y., Huang, X., Sester, M., 2022. Extraction and analysis of natural disaster-related VGI from social media: Review, opportunities and challenges. *Int. J. Geogr. Inf. Sci.* 36 (7), 1275–1316. <http://dx.doi.org/10.1080/13658816.2022.2048835>.
- Fleischmann, M., 2019. Momepy: Urban morphology measuring toolkit. *J. Open Source Softw.* 4 (43), 1807. <http://dx.doi.org/10.21105/joss.01807>.
- Fu, X., Jia, T., Zhang, X., Li, S., Zhang, Y., 2019. Do street-level scene perceptions affect housing prices in Chinese megacities? An analysis using open access datasets and deep learning. In: Fu, S. (Ed.), *PLoS One* 14 (5), e0217505. <http://dx.doi.org/10.1371/journal.pone.0217505>.
- Gao, H., Xu, Z., Chen, Y., Lu, Y., Lin, J., 2022. Walking environment and obesity: A gender-specific association study in Shanghai. *Int. J. Environ. Res. Public Health* 19 (4), 2056. <http://dx.doi.org/10.3390/ijerph19042056>.
- Gaw, L.Y., Chen, S., Chow, Y.S., Lee, K., Biljecki, F., 2022. Comparing street view imagery and aerial perspectives in the built environment. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* X-4/W3-2022-49-2022. <http://dx.doi.org/10.5194/isprs-annals-X-4-W3-2022-49-2022>.
- Gong, Z., Ma, Q., Kan, C., Qi, Q., 2019a. Classifying street spaces with street view images for a spatial indicator of urban functions. *Sustainability* 11 (22), 6424. <http://dx.doi.org/10.3390/su11226424>.
- Gong, F.Y., Zeng, Z.C., Ng, E., Norford, L.K., 2019b. Spatiotemporal patterns of street-level solar radiation estimated using Google Street View in a high-density urban environment. *Build. Environ.* 148, 547–566. <http://dx.doi.org/10.1016/j.buildenv.2018.10.025>.
- Gong, F.Y., Zeng, Z.C., Zhang, F., Li, X., Ng, E., Norford, L.K., 2018. Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Build. Environ.* 134, 155–167. <http://dx.doi.org/10.1016/j.buildenv.2018.02.042>.
- Gorgul, E., Chen, C., Wu, K.K., Guo, Y., 2019. Measuring street enclosure and its influence to human physiology through wearable sensors. In: *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. ACM, <http://dx.doi.org/10.1145/3341162.3343794>.
- Guan, F., Fang, Z., Wang, L., Zhang, X., Zhong, H., Huang, H., 2022. Modelling people's perceived scene complexity of real-world environments using street-view panoramas and open geodata. *ISPRS J. Photogramm. Remote Sens.* 186, 315–331. <http://dx.doi.org/10.1016/j.isprsjprs.2022.02.012>.
- Han, Y., Zhong, T., Yeh, A.G.O., Zhong, X., Chen, M., Lü, G., 2023. Mapping seasonal changes of street greenery using multi-temporal street-view images. *Sustainable Cities Soc.* 92, 104498. <http://dx.doi.org/10.1016/j.scs.2023.104498>.
- Hawes, J.K., Gounaris, D., Newell, J.P., 2022. Does urban agriculture lead to gentrification? *Landscape Urban Plan.* 225, 104447. <http://dx.doi.org/10.1016/j.landurbplan.2022.104447>.
- He, H., Lin, X., Yang, Y., Lu, Y., 2020. Association of street greenery and physical activity in older adults: A novel study using pedestrian-centered photographs. *Urban For. Urban Green.* 55, 126789. <http://dx.doi.org/10.1016/j.ufug.2020.126789>.

- He, Z., Wang, Z., Xie, Z., Wu, L., Chen, Z., 2022a. Multiscale analysis of the influence of street built environment on crime occurrence using street-view images. *Comput. Environ. Urban Syst.* 97, 101865. <http://dx.doi.org/10.1016/j.compenvurbsys.2022.101865>.
- He, W., Zhang, L., Yuan, C., 2022b. Future air temperature projection in high-density tropical cities based on global climate change and urbanization – A study in Singapore. *Urban Clim.* 42, 101115. <http://dx.doi.org/10.1016/j.ulclim.2022.101115>.
- Helbich, M., Poppe, R., Oberski, D., van Emmichoven, M.Z., Schram, R., 2021. Can't see the wood for the trees? An assessment of street view- and satellite-derived greenness measures in relation to mental health. *Landscape. Urban Plan.* 214, 104181. <http://dx.doi.org/10.1016/j.landurbplan.2021.104181>.
- Heo, H.K., Lee, D.K., Park, C.Y., Kim, H.G., 2020. Sky view factor calculation in complex urban geometry with terrestrial LiDAR. *Phys. Geogr.* 42 (4), 1–21. <http://dx.doi.org/10.1080/02723646.2020.1778156>.
- Hosseini, M., Miranda, F., Lin, J., Silva, C.T., 2022. CitySurfaces: City-scale semantic segmentation of sidewalk materials. *Sustainable Cities Soc.* 79, 103630. <http://dx.doi.org/10.1016/j.scs.2021.103630>.
- Hou, Y., Biljecki, F., 2022. A comprehensive framework for evaluating the quality of street view imagery. *Int. J. Appl. Earth Obs. Geoinf.* 115, 103094. <http://dx.doi.org/10.1016/j.jag.2022.103094>.
- Hu, C.B., Zhang, F., Gong, F.Y., Ratti, C., Li, X., 2020. Classification and mapping of urban canyon geometry using Google street view images and deep multitask learning. *Build. Environ.* 167, 106424. <http://dx.doi.org/10.1016/j.buildenv.2019.106424>.
- Hua, J., Cai, M., Shi, Y., Ren, C., Xie, J., Chung, L.C.H., Lu, Y., Chen, L., Yu, Z., Webster, C., 2022. Investigating pedestrian-level greenery in urban forms in a high-density city for urban planning. *Sustainable Cities Soc.* 80, 103755. <http://dx.doi.org/10.1016/j.scs.2022.103755>.
- Huang, J., Yin, Y., Zimmermann, R., Hu, S., Wang, Y., Zhao, C., Wang, G., He, X., Huang, X., Zheng, S., Galloway, T., 2019. GrabView: A scalable street view system for images taken from different devices. In: 2019 IEEE Fifth International Conference on Multimedia Big Data. BigMM, IEEE, <http://dx.doi.org/10.1109/bigmm.2019.9000-6>.
- Ignatius, M., Xu, R., Hou, Y., Liang, X., Zhao, T., Chen, S., Wong, N.H., Biljecki, F., 2022. Local climate zones: lessons from Singapore and potential improvement with street view imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. X-4/W2-2022*, 121–128. <http://dx.doi.org/10.5194/isprs-annals-X-4-W2-2022-121-2022>.
- Inoue, T., Manabe, R., Murayama, A., Koizumi, H., 2022a. The effect of culture-specific differences in urban streetscapes on the inference accuracy of deep learning models. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. X-4/W3-2022*, 73–80. <http://dx.doi.org/10.5194/isprs-annals-X-4-w3-2022-73-2022>.
- Inoue, T., Manabe, R., Murayama, A., Koizumi, H., 2022b. Landscape value in urban neighborhoods: A pilot analysis using street-level images. *Landscape. Urban Plan.* 221, 104357. <http://dx.doi.org/10.1016/j.landurbplan.2022.104357>.
- Ito, K., Biljecki, F., 2021. Assessing bikeability with street view imagery and computer vision. *Transp. Res. C* 132, 103371. <http://dx.doi.org/10.1016/j.trc.2021.103371>.
- Jing, F., Liu, L., Zhou, S., Song, J., Wang, L., Zhou, H., Wang, Y., Ma, R., 2021. Assessing the impact of street-view greenery on fear of neighborhood crime in Guangzhou, China. *Int. J. Environ. Res. Public Health* 18 (1), 311. <http://dx.doi.org/10.3390/ijerph18010311>.
- Juhász, L., Hochmair, H.H., 2016. User contribution patterns and completeness evaluation of Mapillary, a crowdsourced street level photo service. *Trans. GIS* 20 (6), 925–947. <http://dx.doi.org/10.1111/tgis.12190>.
- Kang, B., Lee, S., Zou, S., 2021a. Developing sidewalk inventory data using street view images. *Sensors* 21 (9), 3300. <http://dx.doi.org/10.3390/s21093300>.
- Kang, Y., Zhang, F., Peng, W., Gao, S., Rao, J., Duarte, F., Ratti, C., 2021b. Understanding house price appreciation using multi-source big geo-data and machine learning. *Land Use Policy* 111, 104919. <http://dx.doi.org/10.1016/j.landusepol.2020.104919>.
- Karasov, O., Külvik, M., Chervanyov, I., Priadka, K., 2018. Mapping the extent of land cover colour harmony based on satellite Earth observation data. *GeoJournal* 84 (4), 1057–1072. <http://dx.doi.org/10.1007/s10708-018-9908-x>.
- Kawshalya, L.W.G., Weerasinghe, U.G.D., Chandrasekara, D.P., 2022. The impact of visual complexity on perceived safety and comfort of the users: A study on urban streetscape of Sri Lanka. In: Singh, K.K. (Ed.), *PLoS One* 17 (8), e0272074. <http://dx.doi.org/10.1371/journal.pone.0272074>.
- Ki, D., Lee, S., 2021. Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning. *Landscape. Urban Plan.* 205, 103920. <http://dx.doi.org/10.1016/j.landurbplan.2020.103920>.
- Kim, S.N., Lee, H., 2022. Capturing reality: Validation of omnidirectional video-based immersive virtual reality as a streetscape quality auditing method. *Landscape. Urban Plan.* 218, 104290. <http://dx.doi.org/10.1016/j.landurbplan.2021.104290>.
- Kim, J.H., Lee, S., Hipp, J.R., Ki, D., 2021. Decoding urban landscapes: Google street view and measurement sensitivity. *Comput. Environ. Urban Syst.* 88, 101626. <http://dx.doi.org/10.1016/j.compenvurbsys.2021.101626>.
- Koo, B.W., Guhathakurta, S., Botchwey, N., 2022. How are neighborhood and street-level walkability factors associated with walking behaviors? A big data approach using street view images. *Environ. Behav.* 54 (1), 211–241. <http://dx.doi.org/10.1177/00139165211014609>.
- Labetski, A., Vitalis, S., Biljecki, F., Arroyo Ohori, K., Stoter, J., 2022. 3D building metrics for urban morphology. *Int. J. Geogr. Inf. Sci.* 1–32. <http://dx.doi.org/10.1080/13658816.2022.2103818>.
- Lan, H., Gou, Z., Xie, X., 2021. A simplified evaluation method of rooftop solar energy potential based on image semantic segmentation of urban streetscapes. *Sol. Energy* 230, 912–924. <http://dx.doi.org/10.1016/j.solener.2021.10.085>.
- Leon, L.F.A., Quinn, S., 2019. The value of crowdsourced street-level imagery: Examining the shifting property regimes of OpenStreetCam and Mapillary. *GeoJournal* 84 (2), 395–414. <http://dx.doi.org/10.1007/s10708-018-9865-4>.
- León-Sánchez, C., Aguiar, G., Stoter, J., 2022. Creation of a citygml-based 3D city model testbed for energy-related applications. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLVIII-4/W5-2022*, 97–103. <http://dx.doi.org/10.5194/isprs-archives-xlviii-4-w5-2022-97-2022>.
- Li, X., Li, Y., Jia, T., Zhou, L., Hijazi, I.H., 2022a. The six dimensions of built environment on urban visibility: Fusion evidence from multi-source data. *Cities* 121, 103482. <http://dx.doi.org/10.1016/j.cities.2021.103482>.
- Li, S., Ma, S., Tong, D., Jia, Z., Li, P., Long, Y., 2022b. Associations between the quality of street space and the attributes of the built environment using large volumes of street view pictures. *Environ. Plan. B Urban Anal. City Sci.* 49 (4), 1197–1211. <http://dx.doi.org/10.1177/23998083211056341>.
- Li, X., Ratti, C., Seiferling, I., 2018. Quantifying the shade provision of street trees in urban landscape: A case study in Boston, USA, using Google Street View. *Landscape. Urban Plan.* 169, 81–91. <http://dx.doi.org/10.1016/j.landurbplan.2017.08.011>.
- Li, M., Xue, F., Wu, Y., Yeh, A.G.O., 2022c. A room with a view: Automatic assessment of window views for high-rise high-density areas using city information models and deep transfer learning. *Landscape. Urban Plan.* 226, 104505. <http://dx.doi.org/10.1016/j.landurbplan.2022.104505>.
- Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015. Assessing street-level urban greenery using Google Street View and a modified Green View Index. *Urban For. Urban Green.* 14 (3), 675–685. <http://dx.doi.org/10.1016/j.ufug.2015.06.006>.
- Liang, J., Gong, J., Zhang, J., Li, Y., Wu, D., Zhang, G., 2019. GSV2svf-an interactive GIS tool for sky, tree and building view factor estimation from street view photographs. *Build. Environ.* 168, 106475. <http://dx.doi.org/10.1016/j.buildenv.2019.106475>.
- Liang, X., Zhao, T., Biljecki, F., 2023. Revealing spatio-temporal evolution of urban visual environments with street view imagery. *Landscape. Urban Plan.* 237, 104802. <http://dx.doi.org/10.1016/j.landurbplan.2023.104802>.
- Liao, X., Deng, M., Huang, H., 2021. Analyzing multiscale spatial relationships between the house price and visual environment factors. *Appl. Sci.* 12 (1), 213. <http://dx.doi.org/10.3390/app1201213>.
- Lipson, M.J., Nazarian, N., Hart, M.A., Nice, K.A., Conroy, B., 2022. A transformation in city-descriptive input data for urban climate models. *Front. Environ. Sci.* 10, 866398. <http://dx.doi.org/10.3389/fenvs.2022.866398>.
- Liu, X., Song, Y., Wu, K., Wang, J., Li, D., Long, Y., 2015. Understanding urban China with open data. *Cities* 47, 53–61. <http://dx.doi.org/10.1016/j.cities.2015.03.006>.
- Liu, K., Yin, L., Zhang, M., Kang, M., Deng, A.P., Li, Q.L., Song, T., 2021. Facilitating fine-grained intra-urban dengue forecasting by integrating urban environments measured from street-view images. *Infect. Dis. Poverty* 10 (1), 40. <http://dx.doi.org/10.1186/s40249-021-00824-5>.
- Lu, Y., Ferranti, E.J.S., Chapman, L., Pfraun, C., 2023. Assessing urban greenery by harvesting street view data: A review. *Urban For. Urban Green.* 83, 127917. <http://dx.doi.org/10.1016/j.ufug.2023.127917>.
- Lu, Y., Yang, Y., Sun, G., Gou, Z., 2019. Associations between overhead-view and eye-level urban greenness and cycling behaviors. *Cities* 88, 10–18. <http://dx.doi.org/10.1016/j.cities.2019.01.003>.
- Lumnitz, S., Devisscher, T., Mayaud, J.R., Radic, V., Coops, N.C., Griess, V.C., 2021. Mapping trees along urban street networks with deep learning and street-level imagery. *ISPRS J. Photogramm. Remote Sens.* 175, 144–157. <http://dx.doi.org/10.1016/j.isprsjprs.2021.01.016>.
- Luo, J., Zhao, T., Cao, L., Biljecki, F., 2022a. Semantic riverscapes: Perception and evaluation of linear landscapes from oblique imagery using computer vision. *Landscape. Urban Plan.* 228, 104569. <http://dx.doi.org/10.1016/j.landurbplan.2022.104569>.
- Luo, J., Zhao, T., Cao, L., Biljecki, F., 2022b. Water view imagery: Perception and evaluation of urban waterscapes worldwide. *Ecol. Indic.* 145, 109615. <http://dx.doi.org/10.1016/j.ecolind.2022.109615>.
- Ma, D., Fan, H., Li, W., Ding, X., 2019. The state of Mapillary: An exploratory analysis. *ISPRS Int. J. Geo-Inf.* 9 (1), 10. <http://dx.doi.org/10.3390/ijgi9010010>.
- Mahabir, R., Schuchard, R., Crooks, A., Croitoru, A., Stefanidis, A., 2020. Crowdsourcing street view imagery: A comparison of Mapillary and OpenStreetCam. *ISPRS Int. J. Geo-Inf.* 9 (6), 341. <http://dx.doi.org/10.3390/ijgi9060341>.
- Martin, M., Chong, A., Biljecki, F., Miller, C., 2022. Infrared thermography in the built environment: A multi-scale review. *Renew. Sustain. Energy Rev.* 165, 112540. <http://dx.doi.org/10.1016/j.rser.2022.112540>.
- Meier, F., Fenner, D., Grassmann, T., Otto, M., Scherer, D., 2017. Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Clim.* 19, 170–191. <http://dx.doi.org/10.1016/j.ulclim.2017.01.006>.
- Miao, C., Yu, S., Hu, Y., Zhang, H., He, X., Chen, W., 2020. Review of methods used to estimate the sky view factor in urban street canyons. *Build. Environ.* 168, 106497. <http://dx.doi.org/10.1016/j.buildenv.2019.106497>.

- Middel, A., Lukasczyk, J., Maciejewski, R., 2017. Sky view factors from synthetic fisheye photos for thermal comfort routing—A case study in Phoenix, Arizona. *Urban Plan.* 2 (1), 19–30. <http://dx.doi.org/10.17645/up.v2i1.855>.
- Middel, A., Lukasczyk, J., Maciejewski, R., Demuzere, M., Roth, M., 2018. Sky view factor footprints for urban climate modeling. *Urban Clim.* 25, 120–134. <http://dx.doi.org/10.1016/j.uclim.2018.05.004>.
- Middel, A., Lukasczyk, J., Zakrzewski, S., Arnold, M., Maciejewski, R., 2019. Urban form and composition of street canyons: A human-centric big data and deep learning approach. *Landsc. Urban Plan.* 183, 122–132. <http://dx.doi.org/10.1016/j.landurbplan.2018.12.001>.
- Mukisa, N., Zamora, R., Lie, T.T., 2019. Analysis of solar cell temperature models used in solar photovoltaic simulating softwares. In: 2019 IEEE PES GTD Grand International Conference and Exposition Asia. GTD Asia, IEEE, <http://dx.doi.org/10.1109/gtdasia.2019.8715916>.
- Naghavi, M., Alesheikh, A.A., Hakimpour, F., Vahidnia, M.H., Vafaeinejad, A., 2022. VGI-based spatial data infrastructure for land administration. *Land Use Policy* 114, 105969. <http://dx.doi.org/10.1016/j.landusepol.2021.105969>.
- Nievas, C.I., Pilz, M., Prehn, K., Schorlemmer, D., Weatherill, G., Cotton, F., 2022. Calculating earthquake damage building by building: The case of the city of Cologne, Germany. *Bull. Earthq. Eng.* 1–47. <http://dx.doi.org/10.1007/s10518-021-01303-w>.
- Ning, H., Li, Z., Wang, C., Hodgson, M.E., Huang, X., Li, X., 2022a. Converting street view images to land cover maps for metric mapping: A case study on sidewalk network extraction for the wheelchair users. *Comput. Environ. Urban Syst.* 95, 101808. <http://dx.doi.org/10.1016/j.compenvurbsys.2022.101808>.
- Ning, H., Li, Z., Ye, X., Wang, S., Wang, W., Huang, X., 2022b. Exploring the vertical dimension of street view image based on deep learning: A case study on lowest floor elevation estimation. *Int. J. Geogr. Inf. Sci.* 36 (7), 1317–1342. <http://dx.doi.org/10.1080/13658816.2021.1981334>.
- Ning, H., Ye, X., Chen, Z., Liu, T., Cao, T., 2022c. Sidewalk extraction using aerial and street view images. *Environ. Plan. B Urban Anal. City Sci.* 49 (1), 7–22. <http://dx.doi.org/10.1177/2399808321995817>.
- Ogawa, Y., Oki, T., Chen, S., Sekimoto, Y., 2021. Joining Street-View Images and Building Footprint GIS Data. In: Proceedings of the 1st ACM SIGSPATIAL International Workshop on Searching and Mining Large Collections of Geospatial Data. ACM, Beijing China, pp. 18–24. <http://dx.doi.org/10.1145/3486640.3491395>.
- Oki, T., Ogawa, Y., 2022. A method for regional analysis using deep learning based on big data of omnidirectional images of streets. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLIII-B4-2022, 545–552. <http://dx.doi.org/10.5194/isprs-archives-xliiib4-2022-545-2022>.
- Pang, H.E., Biljecki, F., 2022. 3D building reconstruction from single street view images using deep learning. *Int. J. Appl. Earth Obs. Geoinf.* 112, 102859. <http://dx.doi.org/10.1016/j.jag.2022.102859>.
- Peng, Z., Bardhan, R., Ellard, C., Steemers, K., 2022. Urban climate walk: A stop-and-go assessment of the dynamic thermal sensation and perception in two waterfront districts in Rome, Italy. *Build. Environ.* 221, 109267. <http://dx.doi.org/10.1016/j.buildenv.2022.109267>.
- Peng, Z., Gao, S., Xiao, B., Guo, S., Yang, Y., 2018. CrowdGIS: Updating digital maps via mobile crowdsensing. *IEEE Trans. Autom. Sci. Eng.* 15 (1), 369–380. <http://dx.doi.org/10.1109/taes.2017.2761793>.
- Peters, R., Dukai, B., Vitalis, S., van Liempt, J., Stoter, J., 2022. Automated 3D reconstruction of LoD2 and LoD1 models for all 10 million buildings of the Netherlands. *Photogramm. Eng. Remote Sens.* 88 (3), 165–170. <http://dx.doi.org/10.14358/pers.21-00032r2>.
- Qiu, W., Zhang, Z., Liu, X., Li, W., Li, X., Xu, X., Huang, X., 2022. Subjective or objective measures of street environment, which are more effective in explaining housing prices? *Landsc. Urban Plan.* 221, 104358. <http://dx.doi.org/10.1016/j.landurbplan.2022.104358>.
- Quinn, S., León, L.A., 2019. Every single street? Rethinking full coverage across street-level imagery platforms. *Trans. GIS* 23 (6), 1251–1272. <http://dx.doi.org/10.1111/tgis.12571>.
- Rafiee, A., Dias, E., Koomen, E., 2016. Local impact of tree volume on nocturnal urban heat island: A case study in Amsterdam. *Urban For. Urban Green.* 16, 50–61. <http://dx.doi.org/10.1016/j.ufug.2016.01.008>.
- Raviscioni, S., Bolanos, M.A.R., Gaspari, F., Ponti, C., Biraghi, C.A., Tadi, M., 2022. Collaborative mapping strategy in support of slum upgrading: The case of Bogotá. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLIII-B4-2022, 273–280. <http://dx.doi.org/10.5194/isprs-archives-xliiib4-2022-273-2022>.
- Rey, S.J., 2019. PySAL: The first 10 years. *Spat. Econ. Anal.* 14 (3), 273–282. <http://dx.doi.org/10.1080/17421772.2019.1593495>.
- Sarkar, D., Anderson, J.T., 2022. Corporate editors in OpenStreetMap: Investigating co-editing patterns. *Trans. GIS* 26, 1879–1897. <http://dx.doi.org/10.1111/tgis.12910>.
- Seiferling, I., Naik, N., Ratti, C., Proulx, R., 2017. Green streets - Quantifying and mapping urban trees with street-level imagery and computer vision. *Landsc. Urban Plan.* 165, 93–101. <http://dx.doi.org/10.1016/j.landurbplan.2017.05.010>.
- Seto, T., Nishimura, Y., 2022. Analysis of the spatiotemporal accumulation process of Mapillary data and its relationship with OSM road data: A case study in Japan. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLVIII-4/W1-2022, 403–410. <http://dx.doi.org/10.5194/isprs-archives-xlviii-4-w1-2022-403-2022>.
- Sevtuk, A., Basu, R., Li, X., Kalvo, R., 2021. A big data approach to understanding pedestrian route choice preferences: Evidence from San Francisco. *Travel Behav. Soc.* 25, 41–51. <http://dx.doi.org/10.1016/j.tbs.2021.05.010>.
- So, W., Duarte, F., 2020. Cartographers of North Korea: Who are they and what are the technical, political, and social issues involved in mapping North Korea. *Geoforum* 110, 147–156. <http://dx.doi.org/10.1016/j.geoforum.2020.02.008>.
- Stowell, D., Kelly, J., Tanner, D., Taylor, J., Jones, E., Geddes, J., Chalstrey, E., 2020. A harmonised, high-coverage, open dataset of solar photovoltaic installations in the UK. *Sci. Data* 7 (1), 394. <http://dx.doi.org/10.1038/s41597-020-00739-0>.
- Suel, E., Bhatt, S., Brauer, M., Flaxman, S., Ezzati, M., 2021. Multimodal deep learning from satellite and street-level imagery for measuring income, overcrowding, and environmental deprivation in urban areas. *Remote Sens. Environ.* 257, 112339. <http://dx.doi.org/10.1016/j.rse.2021.112339>.
- Sun, M., Han, C., Nie, Q., Xu, J., Zhang, F., Zhao, Q., 2022. Understanding building energy efficiency with administrative and emerging urban big data by deep learning in Glasgow. *Energy Build.* 112331. <http://dx.doi.org/10.1016/j.enbuild.2022.112331>.
- Sun, Q.C., Macleod, T., Both, A., Hurley, J., Butt, A., Amati, M., 2021a. A human-centred assessment framework to prioritise heat mitigation efforts for active travel at city scale. *Sci. Total Environ.* 763, 143033. <http://dx.doi.org/10.1016/j.scitotenv.2020.143033>.
- Sun, Y., Wang, X., Zhu, J., Chen, L., Jia, Y., Lawrence, J.M., Jiang, L.H., Xie, X., Wu, J., 2021b. Using machine learning to examine street green space types at a high spatial resolution: Application in Los Angeles County on socioeconomic disparities in exposure. *Sci. Total Environ.* 787, 147653. <http://dx.doi.org/10.1016/j.scitotenv.2021.147653>.
- Svensson, M.K., 2004. Sky view factor analysis – implications for urban air temperature differences. *Meteorol. Appl.* 11 (3), 201–211. <http://dx.doi.org/10.1017/s1350482704001288>.
- Verma, D., Jana, A., Ramamritham, K., 2019. Machine-based understanding of manually collected visual and auditory datasets for urban perception studies. *Landsc. Urban Plan.* 190, 103604. <http://dx.doi.org/10.1016/j.landurbplan.2019.103604>.
- Verma, D., Jana, A., Ramamritham, K., 2020. Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. *Build. Environ.* 186, 107340. <http://dx.doi.org/10.1016/j.buildenv.2020.107340>.
- Walker, H.K., Hall, W.D., Hurst, J.W., 1990. *Clinical Methods: The History, Physical, and Laboratory Examinations*. Butterworths, Boston.
- Wang, C., Antos, S.E., Triveno, L.M., 2021. Automatic detection of unreinforced masonry buildings from street view images using deep learning-based image segmentation. *Autom. Constr.* 132, 103968. <http://dx.doi.org/10.1016/j.autcon.2021.103968>.
- Wang, M., Chen, Z., Rong, H.H., Mu, L., Zhu, P., Shi, Z., 2022. Ridesharing accessibility from the human eye: Spatial modeling of built environment with street-level images. *Comput. Environ. Urban Syst.* 97, 101858. <http://dx.doi.org/10.1016/j.compenvurbsys.2022.101858>.
- Wang, C., Middel, A., Myint, S.W., Kaplan, S., Brazel, A.J., Lukasczyk, J., 2018. Assessing local climate zones in arid cities: The case of Phoenix, Arizona and Las Vegas, Nevada. *ISPRS J. Photogramm. Remote Sens.* 141, 59–71. <http://dx.doi.org/10.1016/j.isprsjprs.2018.04.009>.
- Wong, N.H., Kardinal Jusuf, S., 2010. Air temperature distribution and the influence of sky view factor in a Green Singapore estate. *J. Urban Plann. Dev.* 136 (3), 261–272.
- Woźniak, S., Szmyański, P., 2021. Hex2vec - context-aware embedding H3 hexagons with OpenStreetMap tags. In: Proceedings of the 4th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery. ACM, <http://dx.doi.org/10.1145/3486635.3491076>.
- Wu, J., Cheng, L., Chu, S., Xia, N., Li, M., 2019. A Green View Index for urban transportation: How much greenery do we view while moving around in cities? *Int. J. Sustain. Transp.* 14 (12), 972–989. <http://dx.doi.org/10.1080/15568318.2019.1672001>.
- Wu, D., Gong, J., Liang, J., Sun, J., Zhang, G., 2020a. Analyzing the influence of urban street greening and street buildings on summertime air pollution based on street view image data. *ISPRS Int. J. Geo-Inf.* 9 (9), 500. <http://dx.doi.org/10.3390/ijgi9090500>.
- Wu, C., Peng, N., Ma, X., Li, S., Rao, J., 2020b. Assessing multiscale visual appearance characteristics of neighbourhoods using geographically weighted principal component analysis in Shenzhen, China. *Comput. Environ. Urban Syst.* 84, 101547. <http://dx.doi.org/10.1016/j.compenvurbsys.2020.101547>.
- Wysocki, O., Hoegner, L., Still, U., 2022. Refinement of semantic 3D building models by reconstructing underpasses from MLS point clouds. *Int. J. Appl. Earth Obs. Geoinf.* 111, 102841. <http://dx.doi.org/10.1016/j.jag.2022.102841>.
- Xia, Y., Yabuki, N., Fukuda, T., 2021. Sky view factor estimation from street view images based on semantic segmentation. *Urban Clim.* 40, 100999. <http://dx.doi.org/10.1016/j.uclim.2021.100999>.
- Xu, F., Gao, Z., 2022. Frontal area index: A review of calculation methods and application in the urban environment. *Build. Environ.* 109588. <http://dx.doi.org/10.1016/j.buildenv.2022.109588>.

- Yan, Y., Feng, C.C., Huang, W., Fan, H., Wang, Y.C., Zipf, A., 2020. Volunteered geographic information research in the first decade: A narrative review of selected journal articles in GIScience. *Int. J. Geogr. Inf. Sci.* 34 (9), 1–27. <http://dx.doi.org/10.1080/13658816.2020.1730848>.
- Yan, Y., Huang, B., 2022. Estimation of building height using a single street view image via deep neural networks. *ISPRS J. Photogramm. Remote Sens.* 192, 83–98. <http://dx.doi.org/10.1016/j.isprsjprs.2022.08.006>.
- Yang, L., Liu, J., Liang, Y., Lu, Y., Yang, H., 2021a. Spatially varying effects of street greenery on walking time of older adults. *ISPRS Int. J. Geo-Inf.* 10 (9), 596. <http://dx.doi.org/10.3390/ijgi10090596>.
- Yang, J., Rong, H., Kang, Y., Zhang, F., Chegut, A., 2021b. The financial impact of street-level greenery on new york commercial buildings. *Landscape Urban Plan.* 214, 104162. <http://dx.doi.org/10.1016/j.landurbplan.2021.104162>.
- Yang, J., Zhao, L., McBride, J., Gong, P., 2009. Can you see green? Assessing the visibility of urban forests in cities. *Landscape Urban Plan.* 91 (2), 97–104. <http://dx.doi.org/10.1016/j.landurbplan.2008.12.004>.
- Yao, S., Wang, F., Huang, B., 2022a. Measuring PM2.5 concentrations from a single smartphone photograph. *Remote Sens.* 14 (11), 2572. <http://dx.doi.org/10.3390/rs14112572>.
- Yao, Y., Xu, C., Yin, H., Shao, L., Wang, R., 2022b. More visible greenspace, stronger heart? Evidence from ischaemic heart disease emergency department visits by middle-aged and older adults in Hubei, China. *Landscape Urban Plan.* 224, 104444. <http://dx.doi.org/10.1016/j.landurbplan.2022.104444>.
- Yao, Y., Zhang, J., Qian, C., Wang, Y., Ren, S., Yuan, Z., Guan, Q., 2021. Delineating urban job-housing patterns at a parcel scale with street view imagery. *Int. J. Geogr. Inf. Sci.* 35 (10), 1–24. <http://dx.doi.org/10.1080/13658816.2021.1895170>.
- Yap, W., Chang, J.-H., Biljecki, F., 2022a. Incorporating networks in semantic understanding of streetscapes: Contextualising active mobility decisions. *Environ. Plan. B Urban Anal. City Sci.* <http://dx.doi.org/10.1177/23998083221138832>.
- Yap, W., Janssen, P., Biljecki, F., 2022b. Free and open source urbanism: Software for urban planning practice. *Comput. Environ. Urban Plan.* 96, 101825. <http://dx.doi.org/10.1016/j.compenvurbsys.2022.101825>.
- Ye, Y., Richards, D., Lu, Y., Song, X., Zhuang, Y., Zeng, W., Zhong, T., 2018. Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landscape Urban Plan.* 191, 103434. <http://dx.doi.org/10.1016/j.landurbplan.2018.08.028>.
- Yin, L., Wang, Z., 2016. Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Appl. Geogr.* 76, 147–153. <http://dx.doi.org/10.1016/j.apgeog.2016.09.024>.
- Yu, X., He, Y., Huo, W., Chen, G., Qi, W., 2022. Spatio-temporal monitoring of urban street-side vegetation greenery using Baidu street view images. *Urban For. Urban Green.* 73, 127617. <http://dx.doi.org/10.1016/j.ufug.2022.127617>.
- Yu, H., Zhou, Y., Wang, R., Qian, Z., Knibbs, L.D., Jalaludin, B., Schootman, M., McMillin, S.E., Howard, S.W., Lin, L.Z., Zhou, P., Hu, L.W., Liu, R.Q., Yang, B.Y., Chen, G., Zeng, X.W., Feng, W., Xiang, M., Dong, G.H., 2021. Associations between trees and grass presence with childhood asthma prevalence using deep learning image segmentation and a novel green view index. *Environ. Pollut.* 286, 117582. <http://dx.doi.org/10.1016/j.envpol.2021.117582>.
- Yuan, C., Chen, L., 2011. Mitigating urban heat island effects in high-density cities based on sky view factor and urban morphological understanding: A study of Hong Kong. *Archit. Sci. Rev.* 54 (4), 305–315. <http://dx.doi.org/10.1080/00038628.2011.613644>.
- Zeng, L., Lu, J., Li, W., Li, Y., 2018. A fast approach for large-scale sky view factor estimation using street view images. *Build. Environ.* 135, 74–84. <http://dx.doi.org/10.1016/j.buildenv.2018.03.009>.
- Zhang, Y., Chen, N., Du, W., Li, Y., Zheng, X., 2021a. Multi-source sensor based urban habitat and resident health sensing: A case study of Wuhan, China. *Build. Environ.* 198, 107883. <http://dx.doi.org/10.1016/j.buildenv.2021.107883>.
- Zhang, C., Fan, H., Kong, G., 2021b. VGI3D: An interactive and low-cost solution for 3D building modelling from street-level VGI images. *J. Geovisualization Spat. Anal.* 5 (2), 18. <http://dx.doi.org/10.1007/s41651-021-00086-7>.
- Zhang, J., Hu, A., 2022. Analyzing Green view index and Green view index best path using Google street view and deep learning. *J. Comput. Des. Eng.* 9, 2010–2023. <http://dx.doi.org/10.1093/jcdqe/qwac102>.
- Zhang, Y., Middel, A., Turner, B.L., 2019. Evaluating the effect of 3D urban form on neighborhood land surface temperature using Google Street View and geographically weighted regression. *Landscape Ecol.* 34 (3), 681–697. <http://dx.doi.org/10.1007/s10980-019-00794-y>.
- Zhang, L., Pei, T., Wang, X., Wu, M., Song, C., Guo, S., Chen, Y., 2020. Quantifying the urban visual perception of Chinese traditional-style building with street view images. *Appl. Sci.* 10 (17), 5963. <http://dx.doi.org/10.3390/app10175963>.
- Zhao, T., Liang, X., Tu, W., Huang, Z., Biljecki, F., 2023. Sensing urban soundscapes from street view imagery. *Comput. Environ. Urban Syst.* 99, 101915. <http://dx.doi.org/10.1016/j.compenvurbsys.2022.101915>.
- Zheng, X., Amemiya, M., 2023. Method for applying crowdsourced street-level imagery data to evaluate street-level greenness. *ISPRS Int. J. Geo-Inf.* 12 (3), 108. <http://dx.doi.org/10.3390/ijgi12030108>.
- Zhong, T., Zhang, K., Chen, M., Wang, Y., Zhu, R., Zhang, Z., Zhou, Z., Qian, Z., Lv, G., Yan, J., 2021. Assessment of solar photovoltaic potentials on urban noise barriers using street-view imagery. *Renew. Energy* 168, 181–194. <http://dx.doi.org/10.1016/j.renene.2020.12.044>.
- Zhou, H., Gu, J., Liu, Y., Wang, X., 2022a. The impact of the “skeleton” and “skin” for the streetscape on the walking behavior in 3D vertical cities. *Landscape Urban Plan.* 227, 104543. <http://dx.doi.org/10.1016/j.landurbplan.2022.104543>.
- Zhou, Y., Han, X., Peng, M., Li, H., Yang, B., Dong, Z., Yang, B., 2022b. Street-view imagery guided street furniture inventory from mobile laser scanning point clouds. *ISPRS J. Photogramm. Remote Sens.* 189, 63–77. <http://dx.doi.org/10.1016/j.isprsjprs.2022.04.023>.
- Zhou, H., Tao, G., Nie, Y., Yan, X., Sun, J., 2022c. Outdoor thermal environment on road and its influencing factors in hot, humid weather: A case study in Xuzhou City, China. *Build. Environ.* 207, 108460. <http://dx.doi.org/10.1016/j.buildenv.2021.108460>.
- Zhu, H., Yang, F., Bao, Z., Nan, X., 2021. A study on the impact of visible Green index and vegetation structures on brain wave change in residential landscape. *Urban For. Urban Green.* 64, 127299. <http://dx.doi.org/10.1016/j.ufug.2021.127299>.