ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-4/W3-2022
7th International Conference on Smart Data and Smart Cities (SDSC), 19–21 October 2022, Sydney, Australia

COMPARING STREET VIEW IMAGERY AND AERIAL PERSPECTIVES IN THE BUILT ENVIRONMENT

L. Y. Gaw[1], S. Chen[1], Y. S. Chow[1], K. Lee[2], F. Biljecki[1,3,*]

[1] Department of Architecture, National University of Singapore, Singapore - {leongaw, chenshuting, yoongshin, filip}@nus.edu.sg
[2] Yale-NUS College, Singapore - kay.lee@u.yale-nus.edu.sg
[3] Department of Real Estate, National University of Singapore, Singapore

Commission IV, WG IV/9

KEY WORDS: Buildings, Google Street View, Ground-level, Mobile Mapping Systems, Spatial Data Infrastructure

ABSTRACT:

Street view imagery (SVI) has gained prominence in the past decade, offering a new perspective to map and understand cities. It supports numerous studies in the built environment, by replacing or supplementing aerial and satellite imagery, where some studies have not yet been possible with traditional platforms and have now been enabled for the first time thanks to the increasing volume of SVI data. However, the two perspectives are often disconnected and there has not been an overarching paper to discuss the pros and cons of each. We provide an overview outlining and discussing the role of SVI in GIS and urban studies spanning six use cases. Our discourse is supported by a systematic literature review of more than 100 papers and our own experiments that reveal the added value and challenges of SVI in extracting information on buildings and other urban features, an increasingly important use case. We find that the key advantages of SVI over aerial imagery are that it represents more closely how streetscapes are perceived by people and that it enables extracting certain information that otherwise cannot be gathered from top-down perspectives. However, the spatial coverage of SVI tends to be limited to the vicinity of driveable roads, and its temporal coverage is comparatively sparse.

1. INTRODUCTION

Imagery obtained from aerial and satellite platforms has been a key source of spatial data supporting a variety of research in the built environment (Burke et al., 2021). In the past decade, their ground-level counterpart — street view imagery (SVI) — has been rapidly gaining attention (Biljecki and Ito, 2021), supporting mostly the same research disciplines that have relied on aerial/satellite imagery with some advantages, but also introducing new applications that have not been possible with satellite imagery (Verma et al., 2019; Zhang et al., 2018). Taking advantage of several benefits over aerial imagery such as having a different perspective (illustrated in Figure 1), SVI has been used in assessing walkability and bikeability, mapping street furniture, quantifying urban greenery, streetscape perception studies and so on (Hu et al., 2020; Guan et al., 2022; Verma et al., 2020; Ito and Biljecki, 2021; Hawes et al., 2022). In Figure 1, a – aerial perspectives provide an overview from the top, giving the means to numerous tasks, but it falls short in e.g. understanding the pedestrian perspective and extracting information about buildings other than rooftops; b – oblique perspectives show both the top and a half of buildings' side profile (facade), while its other half requires the opposite oblique perspective; c – street-level imagery enables a complementary perspective but it is limited by obstructions (e.g. vegetation), coverage and reach (e.g. not all features such as buildings may be visible because not all the roads have been covered by SVI, or being obscured by other buildings).

Both perspectives have positioned themselves in various fields, and recent studies in the built environment largely use either of the two, and sometimes their combination. While the advantages and disadvantages of each are somewhat intuitive, there has been no overview and comparison of the perspectives in the research literature nor has the topic been documented. Further, in many cases, they appear to be disconnected and it is seldom discussed whether it makes sense to combine them, as there is little research on complementing one with the other.

The goal of our paper is to investigate the value and challenges of SVI in comparison with aerial and satellite perspectives (from now on referred to only as *aerial* for brevity). The research encompasses a variety of aspects, both technical and non-technical. In the paper, we employ a two-fold method to shed light on such comparison: a review and an experiment on extracting information about the built environment. Our literature review is systematic, while the experiments we conducted are focused on buildings since they are the predominant feature of the built environment and a topic that is frequently the subject of both perspectives (Panagiotidou et al., 2021; Ogawa et al., 2021; Wang et al., 2021; Zhang et al., 2021, 2022).

The paper is organised as follows. In Section 2, we present the methodology of the two approaches. The results are elaborated in Section 3 and discussed in Section 4, while Section 5 concludes the paper with takeaways.

2. METHODS

2.1 Literature Review

To identify studies related to SVI and aerial imagery for our review, we searched for such terms in titles, keywords, and abstracts of articles in Scopus. While there is a plethora of publications that mentions either of the two, the general idea of our search was to include the articles that adopted both SVI and aerial imagery, as they could be include a comparative aspect and could be forthcoming in discussing the benefits and challenges

* Corresponding author

https://doi.org/10.5194/isprs-annals-X-4-W3-2022-49-2022

Figure 1. Illustration of some advantages and challenges of the perspective offered by SVI compared to traditionally used platforms.

of each, and thus, help us understand the two perspectives and draw conclusions.

We used 'street view' or 'street level imagery' or 'street level image' as the first search field, 'satellite imagery' or 'satellite image' or 'aerial imagery' or 'aerial image' or 'aerial' as the second search field in Scopus. The relationship between the two search fields is inter-sectional. That is, the literature we selected should investigate both SVI and aerial imagery, and the keywords that appear within each domain can be any of those listed above. The search was conducted in December 2021 and the time period searched was between then and the preceding five years, i.e. papers published from 2017 to 2021, when SVI gained most attention.

This initial search yielded 140 publications. Thereafter, literature screening was carried out to filter the papers that were relevant for our study. Articles were required to include both data sources to better understand the application of the two data sources in the same context. Further, we only focused on the applications in the urban context. For example, some articles focused on the generation or quality assessment of street view data, rather than the application (Majdik et al., 2017; Regmi and Borji, 2018), which we excluded. Also, applications in rural areas and articles focusing on image processing algorithm research were excluded.

Out of the initial 140 articles, 103 applied SVI or aerial imagery in urban contexts, of which 63 adopted both data sources. Next, we examined the 63 relevant articles to seek information relevant to our research and categorised them according to their application scenarios (Figure 2).

2.2 Experiments

To supplement our literature review and better understand the value and shortcomings of SVI in the built environment, we designed a comprehensive and global experiment that focused on collecting information on buildings from the ground-level perspective. We have examined SVI around the world and analysed how feasible is to extract information on buildings from it. By relying only on SVI in our interpretation, we were able to expose potential limitations of SVI in practice. Extracting information on buildings from SVI has been documented in literature (Kang et al., 2018; Fan et al., 2021; Zhou and Chang, 2021), and it is becoming a common use case of SVI. The implications of this experiment may be applicable to other use cases in mapping types of urban features other than buildings. However, studies usually focus on a specific study area and do not provide a critical overview of the platform. Thus, with this paper, we also contribute to the body of knowledge with a global mapping experiment, focusing on the advantages and disadvantages of SVI for this particular yet versatile use case.

In the experiment, we randomly selected a large number of buildings from OpenStreetMap (OSM), distributed worldwide. The buildings were first mapped to the aggregated WorldPop 2020 1 km raster (Lloyd et al., 2017). Then, we filtered out cells that have at least four buildings and are mapped to at least admin division level 2 based on the Database of Global Administrative Areas (GADM[1]). This process was important for two reasons. First, the administrative data is required to sample buildings in such a way to cover all countries in the world. Second, to account for local variations, we examined groups of buildings that are nearby (i.e. at the district scale). Balancing geographical distribution and a reasonable number of buildings to examine (a manual process, as explained in the continuation), we selected 6,578 buildings for the analysis, with all countries covered.

We used Google Street View (GSV) as the source of SVI, the most common dataset used in this research domain. The process to obtain the SVI that may cover each of the sampled buildings was manual. We located the location of the building in Google Maps, after which, Google Street View mode was activated. From there, the properties of the building such as the number of floors, function, roof shape, and material were collected visually. These attributes were commonly studied building attributes seen in our literature review and are also often mapped in OSM (Biljecki, 2020; Palliwal et al., 2021). For each building, we also denoted how difficult is to extract such information. The difficulties in extracting information from SVI of each building was rated on an ordinal scale of 1 to 4 — with 1 being the easiest and 4 being the most difficult or impossible because there was no GSV available covering the building. Each score was accompanied by a remark that provides more information about the rating.

3. RESULTS

3.1 State of the art

In our literature review, we endeavoured to understand the comparative value and challenges of aerial and street view perspectives. We focused on the applications in the built environment, and have grouped the use cases into six main applications. They are: (1) Buildings; (2) Property Prices; (3) Accessibility; (4) Land Use; (5) Greenery; and (6) Perception. Each of these applications is discussed in terms of the use of aerial imagery, SVI, or both where aerial and SVI complement one another.

3.1.1 Extracting building information According to the literature review, for extracting building information, overall, aerial imagery provides context of building usage, with higher

[1] https://gadm.org/index.html

https://doi.org/10.5194/isprs-annals-X-4-W3-2022-49-2022 |

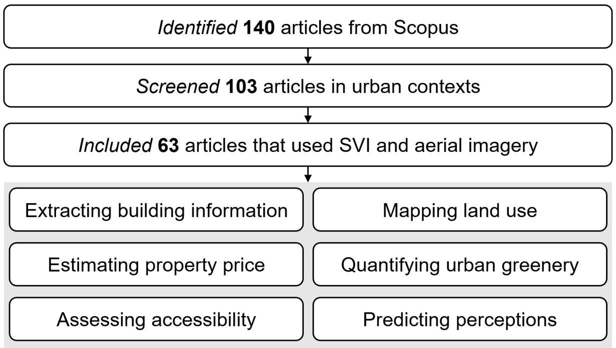

Figure 2. The result of our literature review yielded 63 papers that are of relevance to this paper and they can be categorised into six applications.

resolution aerial imagery achieved better model performance. SVI provides more details on building usage compared with aerial imagery (Hoffmann et al., 2019).

For industrial buildings, the same paper suggests that lower zoom levels (zoomed out) of aerial images achieved higher precision than higher zoom levels (zoomed in), most likely because some industrial buildings are very large and better represented at a lower zoom level. Hoffmann et al. (2019) also noted that it is easier to distinguish the structures of commercial and public buildings from SVI compared to aerial imagery. However, in the cited research, the model using aerial imagery classified higher proportion of commercial buildings than the one relying on SVI, and a higher proportion of public buildings are correctly classified by SVI compared to aerial images. By combining both types of imagery, Hoffmann et al. (2019) found that fusion models that contain the street view perspective performs better than any of the aerial models. This use case is closely related to our experiment, which is elaborated in Section 3.2.

In estimating energy consumption, aerial imagery captures parcel-level elements such as overall building geometry (e.g. footprint, perimeter and compactness), size of electricity-consuming constructions, and the coverage of landscape elements (e.g. green or tree coverage) (Rosenfelder et al., 2021). In comparison, SVI depicts the texture of the building walls and facades, providing information such as the number of storeys and size of the building, the year of construction and whether the building has been renovated (Rosenfelder et al., 2021).

3.1.2 Estimating property price

SVI has been used for estimating the value of real estate and other socio-economic parameters (Suel et al., 2021; Qiu et al., 2022). For property prices, Law et al. (2019) elaborate that the models with visual features derived from street view or aerial imagery perform more accurately than those without visual variables. In discerning the two, the model using aerial imagery performs better than with SVI, which may reflect the fact that buyers are more concerned about the environment of the whole neighbourhood rather than the streets (Law et al., 2019).

3.1.3 Assessing accessibility

For accessibility, specifically roads and sidewalks, Ning et al. (2022b) demonstrated that aerial imagery has a relatively small volume of data to cover the entire study area and can therefore be used as a primary data source. SVI as a supplementary data source provides reliable and ground-level obscured or missing pavements details in aerial imagery (Ning et al., 2022b). For further reading about extracting information of sidewalks with SVI, readers are referred to the studies by Kang et al. (2021) and Hosseini et al. (2022).

3.1.4 Mapping land use

For mapping land use, overall, aerial and SVI reflect different details of land use. For example, aerial imagery captures parcel-scale features; while SVI captures more ground-level details that aerial images lack and help improve the results, especially in ambiguous situations near roads (Srivastava et al., 2019; Cao et al., 2018; Ning et al., 2022a).

For the accuracy of land use classification models, Cao et al. (2018) suggested that pixel-based classification accuracy of the land use types using only aerial imagery is higher than using only SVI. The classification accuracy of land use in educational and transport categories is reported highest by Huang et al. (2020). Otherwise, the classification of commercial and civil land use is less accurate than average. The study reported that the model combining SVI, satellite (aerial) imagery, and auxiliary GIS data achieve highest accuracy. Whereas Cao et al. (2018) suggested classification results are not much improved when integrating SVI with aerial imagery. The exception is that SVI increases prediction accuracy when the aerial image resolution is low.

When it comes to the data coverage, aerial imagery provides a wider coverage than SVI (Cao et al., 2018), as not all buildings may be available in SVI (see Figure 1). The street view perspective may suffer from bias when classifying small scale parcels, and this is because SVI is captured sparsely and unevenly, which may lead to inaccurate results on machine learning (Feng et al., 2018; Cao et al., 2018; Qiao and Yuan, 2021). In SVI, only scenes near streets can be captured due to the limited coverage. Features obscured by large and tall roadside structures are hidden (Cao et al., 2018). Further, the available SVI is often biased towards prosperous areas of the city (Qiao and Yuan, 2021). Qiao and Yuan (2021) also caution that imaging the streets involves the collection and sharing of proximate sensing data that can lead to privacy and trust of anonyms.

3.1.5 Quantifying urban greenery

Mapping urban vegetation has been carried out extensively using a variety of data sources, including aerial and street view imagery, and it is the largest use case we have identified in our review.

To quantify greenery in urban contexts, in terms of spatial resolution, satellite imagery such as the Landsat constellation has relatively low spatial resolution for precise mapping of greenery in urban areas (Helbich et al., 2021; Baučić et al., 2020). However, SVI has comparatively high resolution images (Helbich et al., 2021; Tong et al., 2020; Wang et al., 2019; Kumakoshi et al., 2020). For example, compared to any existing global mapping products derived from aerial imagery, SVI provides a higher image resolution for calculating vegetation indices such as the Leaf Area Index (LAI) (Richards and Wang, 2020). For temporal resolution, imagery from satellites such as Sentinel 2, have a high temporal resolution (5 days at the equator) that allows monitoring changes at short time intervals (Baučić et al., 2020).

Meanwhile, the biggest limitation of SVI is its low frequency of updates (Barbierato et al., 2020). And not all the sample images are captured at the same time. There are instances where SVI in a neighbourhood are captured on different dates, which results in variations in foliage amounts and colours based on seasons (Richards and Wang, 2020).

Regarding data coverage, the data reach of SVI is limited to the immediate surrounds of the street space and public areas that are accessible by vehicles (Helbich et al., 2021; Tong et al., 2020; Wang et al., 2019; Kumakoshi et al., 2020; Barbierato et al., 2020). Whereas aircrafts and satellites are able to capture images of vegetation in areas that are off-road, which cannot be captured by SVI (Tong et al., 2020).

Aerial imagery provides a top-down perspective that can detect green roofs and shading of courtyards that are away from roads or are in private enclosures (Barbierato et al., 2020; Wu and Biljecki, 2021). Its shortcoming is that vegetation on building walls cannot be easily identified since aerial images are captured at near-nadir. In comparison, SVI is better in showcasing the shading of the ground and building facades provided by roadside trees (Barbierato et al., 2020). However, the limitation is when looking skywards at zenith with panoramic street view images, the canopies formed by tall trees can only be partially seen (Kumakoshi et al., 2020).

One key benefit of SVI is that it saves time from going down to a site as one can view a place virtually from a pedestrian's perspective. Thus, it is more easily interpreted compared to aerial imagery since SVI are closer to humans' perception (Aklibasinda, 2019; Kumakoshi et al., 2020).

Comparing the two, street level greenness was weakly to moderately positively correlated with greenness identified by satellite imagery (Helbich et al., 2021; Tong et al., 2020; Ye et al., 2019). Tong et al. (2020) demonstrated that the Green View Index (GVI) derived from SVI is moderately correlated to the Normalised Difference Vegetation Index (NDVI) derived from aerial imagery, which suggests that vegetation viewed from different angles produces different results of greenness. Ye et al. (2019) suggested that there was a positive and significant correlation between NDVI and visible street greenery. The Pearson's correlation coefficient gradually decreases with the increase of the distance away from the road where the SVIs were taken (Ye et al., 2019).

However, there will be a mismatch between top-down view versus on the ground view of *SVI* because the former views vegetation's canopy while the latter views not just the canopy but the sub-canopy structure of vegetation like roots, stem, and branches (Ye et al., 2019). The measured values of aerial imagery metrics (e.g. NDVI) is higher than the SVI metrics (e.g. GVI, sGVI) where vegetation is denser, whereas the opposite trend was observed in urban areas where more buildings are present (Tong et al., 2020; Kumakoshi et al., 2020). This may indicate buildings make it difficult for vegetation to be captured from a top-down perspective, and SVI captures urban vegetation in more details than aerial imagery (Kumakoshi et al., 2020). Hence, it is recommended that multiple indicators incorporating both SVI and aerial imagery be complemented (instead of substituted) to evaluate roadside greenery (Tong et al., 2020).

In terms cost of collection, SVI is acquired by vehicles and may be cheaper compared to aerial imagery from satellites or airborne sensors (Barbierato et al., 2020).

3.1.6 Predicting perceptions For perception studies, variables from SVI measure and explain more variation than aerial imagery (Larkin et al., 2021). Aerial imagery from satellites used to quantify built environment features and environmental exposures do not capture these perception differences (Larkin et al., 2021), and this is a use case where the value of SVI is convincing in comparison with aerial perspective, a finding consistent with the human-level perspective of SVI.

3.2 Experiments on extracting building information

The experiment revealed that for most buildings, it was not possible to extract building information from SVI, with significant differences by geography (Figure 3). On the other hand, where it was possible, SVI provided unparalleled benefits. Figure 4 provides visual examples of different buildings, and how the cases were scored.

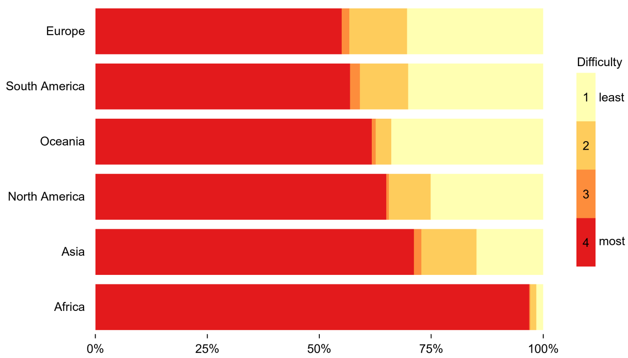

Figure 3. A summary of the difficulties in interpreting SVI by continents. Street view imagery from Europe are the easiest to interpret while those from Africa are the most difficult or they do not provide sufficient coverage.

In the process of collecting building information, a common limitation exhibited for both SVI and aerial imagery was the obstruction of buildings by greenery, including trees and shrubs, as well as neighbouring buildings (Table 1). Dissecting the two platforms, SVI was particularly affected: other structures that obstruct the buildings in SVI are fences, walls, advertisement boards, bus stops and moving objects (i.e. most commonly — vehicles on the road). Despite the seemingly advantageous vantage points offered by SVI, the building of interest might nonetheless be obstructed from all possible angles, with additional coverage lacking, a finding that resonates among related work (Pang and Biljecki, 2022). Thus, information such as the number of storeys and often the function of the building cannot be extracted. While we have focused on understanding the performance of SVI for extracting multiple key characteristics of buildings (see Section 2.2), most often the findings were the same, e.g. none of the characteristics could be acquired as the building was simply entirely obstructed by its neighbour.

While aerial imagery is available in most places, the quality of imagery is also not uniform and varies across areas, which is not the case with SVI sources such as GSV that tend to be consistent in quality.

Similarly, the coverage of SVI is not uniform. In places lacking extensive coverage, SVI is usually available only for main roads and highways. Thus, buildings located further from them cannot be seen. SVI can also be patchy, with breaks along a single road, leaving buildings in between inaccessible. Moreover, the terrain plays an important role in interpreting SVI data. When the building of interest is below the street level, the number of storeys of the building cannot be measured. In narrow streets and some other situations, it is also challenging to acquire information such as the number of storeys of the building if it

https://doi.org/10.5194/isprs-annals-X-4-W3-2022-49-2022

Figure 4. Examples of SVI studied and their interpretation difficulty levels. The OSM building footprints highlighted in orange (first column) are some of those that we randomly selected around the world. The orange point in the Google Maps' aerial imagery represents the street view point with its corresponding image to the right. Images are courtesy of © OpenStreetMap contributors, Google Maps, and Google Street View.

is tall, for example, more than 15 storeys, due to the angle of viewing.

4. DISCUSSION

Our research has revealed that there is a large overlap between aerial and SVI perspectives in practice. Many use cases are possible with either of the two. On the other hand, they often tend to be used in conjunction, taking the best of both worlds, but also allowing us to expose the pros and cons of each.

In this section, based on the results of the literature review and the experiments, we provide a comparative summary of the advantages and disadvantages of street view and aerial imagery that pertain to the built environment. The discussion is summarised in Table 2.

In general, based on both the literature review and experiment, SVI has an advantage that it provides a significantly higher resolution of details of objects in urban areas (Cao et al., 2018; Srivastava et al., 2019; Richards and Wang, 2020). The imagery represents closely the environment that people perceive and experience on the ground, usually following a route that can be traversed on foot or by vehicle (Kumakoshi et al., 2020). Hence, SVI can provide details seen on the ground that are not visible from aerial imagery (Hoffmann et al., 2019; Ning et al., 2022b; Cao et al., 2018; Rosenfelder et al., 2021; Barbierato et al., 2020). SVI is relatively cheaper to capture than aerial imagery thanks to free tiers of commercial companies and the emergence of crowdsourced SVI services such as Mapillary and KartaView (Barbierato et al., 2020; Inoue et al., 2022). What this means is that, even when mainstream services have no coverage of a particular area, SVI can be collected simply by mounting sensors on a backpack or a vehicle and moving around. Aerial imagery, in comparison, is more expensive to collect as sensors need to be mounted on aircrafts or spacecrafts that are very expensive to launch and operate. Startup capital aside, once spaceborne sensors are operational, they can continuously collect data for decades and be released openly for research and other uses. There are some websites where users can publish openly photogrammetric aerial images of the earth's surface captured by drones but coverage is minimal because of the need for expertise in photogrammetry, drone access, and large data storage requirements.

Disadvantages of SVI are just as well documented. Primarily, coverage is limited as only areas near routes that can be traversed by foot or vehicles can be imaged (Tong et al., 2020; Barbierato et al., 2020; Wang et al., 2019; Kumakoshi et al., 2020), potentially leaving some portions of urban areas out of

Difficulty	Most common reason	Second most common reason
1	Building not visible in satellite view but visible in GSV	Obstructed in GSV
2	Ambiguity in describing building that is visible	Obstructed in GSV
3	Ambiguity in describing building that is visible	Building under construction in GSV
4	No GSV available	Building not visible in GSV

Table 1. The predominant reasons by the four difficulty levels in interpreting SVI for extracting information on buildings.

	Advantages	Disadvantages
SVI	- High resolution and detail - Represent more closely how environments are perceived and experienced by people - Street view images can provide ground-level details that aerial images lack - Relatively low cost (i.e. freely available data thanks to both commercial and crowdsourced services) - Some objects in area with a dense SVI coverage may be observed more than once	- Coverage is limited, only area near streets are imaged - Imbalanced data coverage (both spatial and temporal) - Limited temporal coverage and revisit periods (i.e. sometimes once in a decade) - Difficulty in expanding research scale due to extensive processing time - Complicated processing workflow, advanced computer vision techniques are required to process the data - Tall objects (buildings, trees, etc.) may be only partially observed from SVI - The position and light condition of taking SVI may vary in the same point
Aerial	- Wider spatial coverage - Fine time granularity - Can be used to observe broad trends - Capture the overall information in a large scale - Data volume is relatively small	- Certain mismatch between satellite's top-down viewpoint and human-scale viewpoint - Cannot capture some details in building facade - Difficult for human's with limited experience to interpret - Openly available satellite imagery generally has too coarse spatial resolution

Table 2. Comparison of the general advantages and disadvantages of SVI and aerial imagery in studies of the built environment.

reach. Moreover, sometimes only major roads are imaged. As a consequence of this bias in the data collection, use cases may suffer. For example, measuring greenery on district scale (Section 3.1.5) may not be representative if only major and certain streets are considered, unlike it would be the case with aerial imagery, which has homogeneous coverage.

Other than the limited spatial coverage, the temporal coverage is neither extensive nor regular. Hence, the temporal resolution of images in between imagery periods is much coarser than aerial imagery (Baučić et al., 2020). Use cases are also burdened with long and complex processing workflows needed for advanced computer vision techniques and the large volume of data (often tens of thousands of images) (Ning et al., 2022b).

Some objects in areas with high densities of SVI such as dense road networks may have objects imaged more than once in a transect, which complicates data analysis (Kumakoshi et al., 2020). Tall objects such as buildings and trees may be only partly viewed in SVI, especially if the imagery is not panoramic (Kumakoshi et al., 2020). The position and lighting conditions when SVI is taken may vary in the same point (Wang et al., 2019; Richards and Wang, 2020). Thus, SVI in the day is much more common than those captured at night.

For the advantages of aerial imagery over SVI, they provide a wider, even global, coverage of urban areas and the earth (Cao et al., 2018; Ning et al., 2022b; Baučić et al., 2020). Specifically for aerial imagery captured by satellites, their temporal resolution is very high with revisit times of up to everyday (Baučić et al., 2020). This advantage is useful in monitoring changes to the urban landscape through time at a broad scale. The data volume in terms of spatial extents can be small (Ning et al., 2022b), but this increases greatly with the high temporal resolution.

For the advantage of aerial imagery, objects perceived in it are not as intuitive as SVI as humans are more familiar with viewing objects at ground level instead of from the sky (top-down) (Kumakoshi et al., 2020; Larkin et al., 2021). Most aerial imagery are captured with the camera pointing at nadir that makes imaging details of a building's facade difficult (Hoffmann et al., 2019; Rosenfelder et al., 2021). There are aerial imagery at very high resolutions from satellites and airborne missions but these are very expensive to operate, capture, and acquire. The alternative are to use open-sourced aerial imagery from satellites like Landsat from NASA and Sentinel from ESA (Barbierato et al., 2020). However, these imagery are too coarse to make out details of buildings. Rather, they are better suited for monitoring urban growth and areal information (Helbich et al., 2021; Baučić et al., 2020).

5. CONCLUSION

We examined literature that uses both SVI and aerial imagery in their research of the built environment and provided an evaluation across several use cases. We supplemented our review with empirical data of an examination of several thousands of SVI globally and evaluated their benefits and challenges in a use case. Given that each platform has advantages and disadvantages and that data availability is increasing, there should be more effort on how to integrate them, which we consider to be one of the key future directions. While for methodological reasons our research has examined papers that use both sources/perspectives, it should be noted that most studies nowadays use either of the two.

In this research, we have not taken into account unmanned aerial systems (UAS) and the perspectives they offer, which is as well gaining popularity in the built environment (Nex et al., 2022). While SVI offers some advantages over UAS such as not requiring a permit to fly in restricted urban spaces (Wang et al., 2021), the latter offers multiple perspectives (e.g. top and oblique view; cf. Figure 1(b)). For future work it may be beneficial to conduct a critical overview to include UAS that is an intermediary between SVI and aerial imagery.

There are also other comparisons that have not been thoroughly investigated in the scientific literature. Two gaps that were apparent to us were comparisons between data accessibility and quantifiable costs, and the issues that affect both perspectives, e.g. shadows in images that are problematic for both aerial and SVI but more so in SVI as Figure 1(c) depicts where part of the road is overcast by adjacent trees and buildings.

ACKNOWLEDGEMENTS

We thank the members of the NUS Urban Analytics Lab for the discussions, and the reviewers for their suggestions. April Zhu is acknowledged for the design of the 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; and (ii) 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

Aklibasinda, M., 2019. Determining the active green areas and their adequacy by using satellite images and GIS: The case of Nevsehir city (Turkey). *Fresenius Environmental Bulletin*, 28(10), 7274-7281.

Barbierato, E., Bernetti, I., Capecchi, I., Saragosa, C., 2020. Integrating remote sensing and street view images to quantify urban forest ecosystem services. *Remote Sensing*, 12(2).

Baučić, M., Gilić, F., Bačić, S., Duplančić-Leder, T., 2020. Open Geospatial Data for Urban Green Areas. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 6, 17–24.

Biljecki, F., 2020. Exploration of open data in Southeast Asia to generate 3D building models. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, VI-4/W1-2020, 37–44.

Biljecki, F., Ito, K., 2021. Street view imagery in urban analytics and GIS: A review. *Landscape and Urban Planning*, 215.

Burke, M., Driscoll, A., Lobell, D. B., Ermon, S., 2021. Using satellite imagery to understand and promote sustainable development. *Science*, 371(6535), eabe8628.

Cao, R., Zhu, J., Tu, W., Li, Q., Cao, J., Liu, B., Zhang, Q., Qiu, G., 2018. Integrating Aerial and Street View Images for Urban Land Use Classification. *Remote Sensing*, 10(10), 1553.

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.

Feng, T., Truong, Q.-T., Nguyen, D., Koh, J., Yu, L.-F., Binder, A., Yeung, S.-K., 2018. Urban zoning using higher-order markov random fields on multi-view imagery data. *Lecture Notes in Computer Science*, 11212 LNCS, 627-644.

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 Journal of Photogrammetry and Remote Sensing*, 186, 315–331.

Hawes, J. K., Gounaridis, D., Newell, J. P., 2022. Does urban agriculture lead to gentrification? *Landscape and Urban Planning*, 225, 104447.

Helbich, M., Poppe, R., Oberski, D., Zeylmans van Emmichoven, M., 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 and Urban Planning*, 214, 104181.

Hoffmann, E., Wang, Y., Werner, M., Kang, J., Zhu, X., 2019. Model fusion for building type classification from aerial and street view images. *Remote Sensing*, 11(11).

Hosseini, M., Miranda, F., Lin, J., Silva, C. T., 2022. CitySurfaces: City-scale semantic segmentation of sidewalk materials. *Sustainable Cities and Society*, 79, 103630.

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. *Building and Environment*, 167, 106424.

Huang, Z., Qi, H., Kang, C., Su, Y., Liu, Y., 2020. An ensemble learning approach for urban land use mapping based on remote sensing imagery and social sensing data. *Remote Sensing*, 12(19), 1-18.

Inoue, T., Manabe, R., Murayama, A., Koizumi, H., 2022. Landscape value in urban neighborhoods: A pilot analysis using street-level images. *Landscape and Urban Planning*, 221, 104357.

Ito, K., Biljecki, F., 2021. Assessing bikeability with street view imagery and computer vision. *Transportation Research Part C: Emerging Technologies*, 132.

Kang, B., Lee, S., Zou, S., 2021. Developing Sidewalk Inventory Data Using Street View Images. *Sensors*, 21(9), 3300.

Kang, J., Körner, M., Wang, Y., Taubenböck, H., Zhu, X. X., 2018. Building instance classification using street view images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 44–59.

Kumakoshi, Y., Chan, S., Koizumi, H., Li, X., Yoshimura, Y., 2020. Standardized green view index and quantification of different metrics of urban green vegetation. *Sustainability (Switzerland)*, 12(18).

Larkin, A., Gu, X., Chen, L., Hystad, P., 2021. Predicting perceptions of the built environment using GIS, satellite and street view image approaches. *Landscape and Urban Planning*, 216, 104257.

Law, S., Paige, B., Russell, C., 2019. Take a look around: Using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology*, 10(5).

Lloyd, C. T., Sorichetta, A., Tatem, A. J., 2017. High resolution global gridded data for use in population studies. *Scientific Data*, 4(1), e1298.

Majdik, A., Till, C., Scaramuzza, D., 2017. The Zurich urban micro aerial vehicle dataset. *International Journal of Robotics Research*, 36(3), 269-273. cited By 55.

Nex, F., Armenakis, C., Cramer, M., Cucci, D. A., Gerke, M., Honkavaara, E., Kukko, A., Persello, C., Skaloud, J., 2022. UAV in the advent of the twenties: Where we stand and what is next. *ISPRS Journal of Photogrammetry and Remote Sensing*, 184, 215–242.

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. *Computers, Environment and Urban Systems*, 95, 101808.

Ning, H., Ye, X., Chen, Z., Liu, T., Cao, T., 2022b. Sidewalk extraction using aerial and street view images. *Environment and Planning B: Urban Analytics and City Science*, 49(1), 7-22.

Ogawa, Y., Oki, T., Chen, S., Sekimoto, Y., 2021. Joining Street-View Images and Building Footprint GIS Data. *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Searching and Mining Large Collections of Geospatial Data*, 18–24.

Palliwal, A., Song, S., Tan, H. T. W., Biljecki, F., 2021. 3D city models for urban farming site identification in buildings. *Computers, Environment and Urban Systems*, 86, 101584.

Panagiotidou, M., Brito, M. C., Hamza, K., Jasieniak, J. J., Zhou, J., 2021. Prospects of photovoltaic rooftops, walls and windows at a city to building scale. *Solar Energy*, 230, 675–687.

Pang, H. E., Biljecki, F., 2022. 3D building reconstruction from single street view images using deep learning. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102859.

Qiao, Z., Yuan, X., 2021. Urban land-use analysis using proximate sensing imagery: a survey. *International Journal of Geographical Information Science*, 35(11), 2129-2148.

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? *Landscape and Urban Planning*, 221, 104358.

Regmi, K., Borji, A., 2018. Cross-view image synthesis using conditional gans. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3501–3510.

Richards, D., Wang, J., 2020. Fusing street level photographs and satellite remote sensing to map leaf area index. *Ecological Indicators*, 115.

Rosenfelder, M., Wussow, M., Gust, G., Cremades, R., Neumann, D., 2021. Predicting residential electricity consumption using aerial and street view images. *Applied Energy*, 301.

Srivastava, S., Vargas-Muñoz, J., Tuia, D., 2019. Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution. *Remote Sensing of Environment*, 228, 129-143.

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 Sensing of Environment*, 257, 112339.

Tong, M., She, J., Tan, J., Li, M., Ge, R., Gao, Y., 2020. Evaluating street greenery by multiple indicators using street-level imagery and satellite images: A case study in Nanjing, China. *Forests*, 11(12), 1-21.

Verma, D., Jana, A., Ramamritham, K., 2019. Machine-based understanding of manually collected visual and auditory datasets for urban perception studies. *Landscape and Urban Planning*, 190, 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. *Building and Environment*, 186, 107340.

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. *Automation in Construction*, 132, 103968.

Wang, R., Helbich, M., Yao, Y., Zhang, J., Liu, P., Yuan, Y., Liu, Y., 2019. Urban greenery and mental wellbeing in adults: Cross-sectional mediation analyses on multiple pathways across different greenery measures. *Environmental Research*, 176.

Wu, A. N., Biljecki, F., 2021. Roofpedia: Automatic mapping of green and solar roofs for an open roofscape registry and evaluation of urban sustainability. *Landscape and Urban Planning*, 214, 104167.

Ye, Y., Richards, D., Lu, Y., Song, X., Zhuang, Y., Zeng, W., Zhong, T., 2019. Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landscape and Urban Planning*, 191.

Zhang, C., Fan, H., Kong, G., 2021. VGI3D: an Interactive and Low-Cost Solution for 3D Building Modelling from Street-Level VGI Images. *Journal of Geovisualization and Spatial Analysis*, 5(2), 18.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., Ratti, C., 2018. Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180, 148–160.

Zhang, Z., Qian, Z., Zhong, T., Chen, M., Zhang, K., Yang, Y., Zhu, R., Zhang, F., Zhang, H., Zhou, F., Yu, J., Zhang, B., Lü, G., Yan, J., 2022. Vectorized rooftop area data for 90 cities in China. *Scientific Data*, 9(1), 66.

Zhou, P., Chang, Y., 2021. Automated classification of building structures for urban built environment identification using machine learning. *Journal of Building Engineering*, 43, 103008.