



Review

# Detailed Three-Dimensional Building Façade Reconstruction: A Review on Applications, Data and Technologies

Anna Klimkowska<sup>1,\*</sup>, Stefano Cavazzi<sup>2</sup> , Richard Leach<sup>3</sup> and Stephen Grebby<sup>1</sup>

<sup>1</sup> Nottingham Geospatial Institute, University of Nottingham, Jubilee Campus, Nottingham NG7 2TU, UK; stephen.grebby@nottingham.ac.uk

<sup>2</sup> Ordnance Survey, Adanac Drive, Southampton SO16 0AS, UK; stefano.cavazzi@os.uk

<sup>3</sup> Department of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, University Park, Nottingham NG7 2RD, UK; richard.leach@nottingham.ac.uk

\* Correspondence: anna.klimkowska@nottingham.ac.uk

**Abstract:** Urban environments are regions of complex and diverse architecture. Their reconstruction and representation as three-dimensional city models have attracted the attention of many researchers and industry specialists, as they increasingly recognise the potential for new applications requiring detailed building models. Nevertheless, despite being investigated for a few decades, the comprehensive reconstruction of buildings remains a challenging task. While there is a considerable body of literature on this topic, including several systematic reviews summarising ways of acquiring and reconstructing coarse building structures, there is a paucity of in-depth research on the detection and reconstruction of façade openings (i.e., windows and doors). In this review, we provide an overview of emerging applications, data acquisition and processing techniques for building façade reconstruction, emphasising building opening detection. The use of traditional technologies from terrestrial and aerial platforms, along with emerging approaches, such as mobile phones and volunteered geography information, is discussed. The current status of approaches for opening detection is then examined in detail, separated into methods for three-dimensional and two-dimensional data. Based on the review, it is clear that a key limitation associated with façade reconstruction is process automation and the need for user intervention. Another limitation is the incompleteness of the data due to occlusion, which can be reduced by data fusion. In addition, the lack of available diverse benchmark datasets and further investigation into deep-learning methods for façade openings extraction present crucial opportunities for future research.

**Keywords:** façade parsing; building openings; object detection; images; point cloud; platforms; sensors



**Citation:** Klimkowska, A.; Cavazzi, S.; Leach, R.; Grebby, S. Detailed Three-Dimensional Building Façade Reconstruction: A Review on Applications, Data and Technologies. *Remote Sens.* **2022**, *14*, 2579. <https://doi.org/10.3390/rs14112579>

Academic Editor: Csaba Benedek

Received: 31 March 2022

Accepted: 15 May 2022

Published: 27 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

For decades, three-dimensional (3D) city models have been primarily used for visualisation, with an increasing number of stakeholders and practitioners recognising their advantages in decision-making processes. A key reason for this is that the world around us, presented in 3D form, as opposed to two-dimensional (2D) maps and drawings, is typically more comprehensible and easier to perceive and can readily serve as a tool for communicating and sharing information [1].

Therefore, 3D city models play an essential role when analysing and managing urban data, and this is evidenced by the large number of applications that utilise this information [2]. The integration of 3D city models with non-geometrical data, such as social, economic, acoustic or historical information has proven its utility in fields including 3D cadastre [3], emergency response [4], decision-making, urban planning [5], smart cities [6] and more recently, in the digital twin field where 3D city models serve as a fundamental source of information [7].

Among the many elements that make up city models, one of the most prominent, if not the most important, are the buildings. Nevertheless, despite their usefulness, the publicly available 3D building models are often presented as coarse, solid blocks in most cases. Occasionally, these block models are augmented with simplified roof shapes and image texture [8]. However, details of the building openings (windows and doors) are often missing.

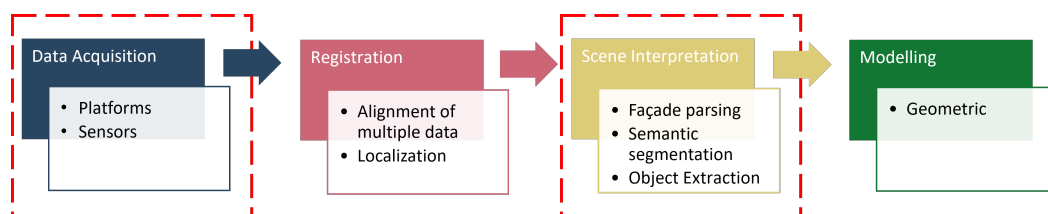
There is a significant body of literature focused on trying to automate building model reconstruction to make it low-cost, faster, accurate and easy to update. Nevertheless, many of the approaches proposed to solve the challenge of automatic generation of detailed models require a great deal of pre-and post-processing, therefore restricting their utilisation as standalone solutions [9].

This challenge is partly associated with the complex geometry of buildings, restricted access in performing measurements, or the presence of objects that obscure the façade, especially their lower parts [10]. An example illustrating the complexity of buildings based on their architectural style is presented in Figure 1. This diversity of styles poses problems to developing universal and robust methods for the reconstruction of building façades in the city environment.



**Figure 1.** Examples of architectural styles in United Kingdom. Top row from left: Tudor, Baroque, Georgian, Victorian and Edwardian; bottom row from left: Art Deco, 1930s, Modern (white-grey building) and Green ‘Living Wall’. The images were taken from multiple modalities : Modern: <https://architizer.com/blog/inspiration/collections/architecture-on-the-market-london-contemporary-homes/>, accessed on 17 January 2022, Living Wall: <https://tugc.co.uk/portfolio-item/maintaining-one-of-the-biggest-living-walls-in-london/>, accessed on 17 January 2022, Others: <https://www.bohaglass.co.uk/british-architectural-styles/>, accessed on 17 January 2022.

3D building reconstruction is relevant to a wide range of research fields, such as computer vision, computer graphics, photogrammetry, geodesy, architecture, civil engineering and construction. To avoid confusion, in this review, the concept of 3D building reconstruction encompasses the processing workflow consisting of several steps: data acquisition, data registration, scene interpretation and object extraction and 3D modelling (Figure 2).



**Figure 2.** The general workflow of the reconstruction process includes data acquisition and registration, scene interpretation and modelling. Red boxes indicate elements of the process covered in this review for detailed façade reconstruction.

Data acquisition is a crucial step in building reconstruction. A building’s architectural style, size, location (i.e., city centre versus suburbs) and the level of detail at which the

final product is to be presented can all determine the choice of the sensor technologies and platforms used to acquire the necessary data used as a basis for reconstructing a building. The next step in the 3D building reconstruction pipeline is data registration. In many cases, to obtain a complete picture of a building, it is necessary to combine data from different perspectives (e.g., terrestrial and airborne) and sensors. Multiple methods are used to register data from point clouds [11–16] or 2D images [17–21].

Scene interpretation is the next stage of the reconstruction pipeline. Many concepts relate to scene interpretation, including façade parsing, segmentation and object extraction. The primary purpose of these methods is to thoroughly extract and analyse the shape and position of walls and façade openings (windows, doors and balconies) from the acquired data. The final stage of the reconstruction process is the geometric modelling of the façade and its elements. These can be achieved using parametric [22], surface [23–25] or volumetric [26,27] modelling approaches.

To date, solutions for detecting the openings of the façade are scarce and generally limited in application to buildings with a relatively narrow architectural variety. This is due to, *inter alia*, a relatively monolithic type of publicly available data and from the generally small amount of available data that would enable more research to be conducted. Nevertheless, the number of applications that could take advantage of this knowledge of detailed 3D building models is increasing.

Therefore, this review initially focuses on highlighting the current research on data and key technologies used to enrich building façades, especially the methods used for façade parsing and building-opening detection. Accordingly, the registration of input data [28] and the geometric modelling of buildings [10,29] are beyond the scope of this review.

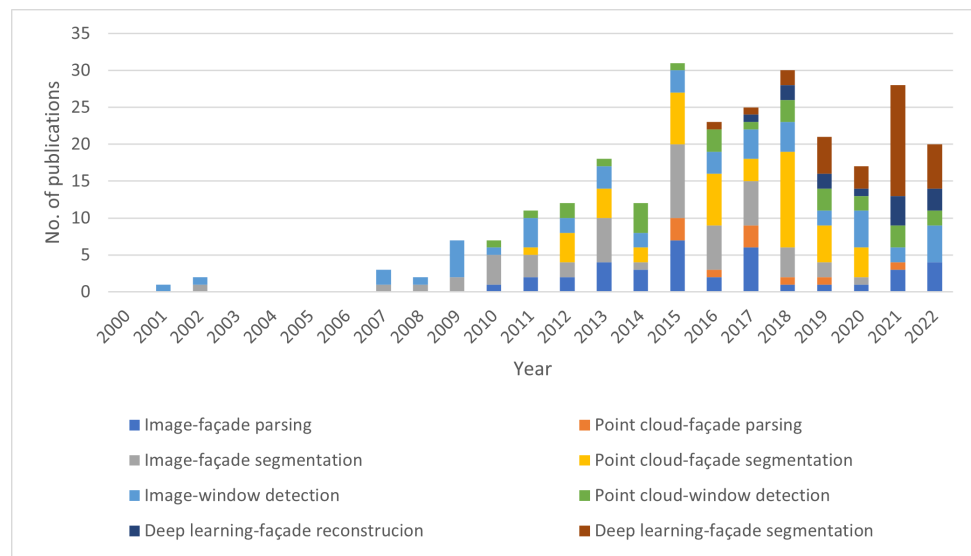
The first part of this review (Section 3) presents an overview of applications, such as building-information modelling and solar energy potential estimation, that could benefit from detailed 3D building models. Section 4 presents an overview of the platforms most commonly used to record building façades, in addition, we introduce emerging platforms, such as handheld scanners, the use of mobile phone images and volunteered geography information. Next, the advantages and disadvantages of range-based and image-based data used to identify façade openings are discussed (Section 5).

In Section 6, we discuss the methods developed to extract the façade elements from 3D and 2D data and highlight the challenges of the current approaches. Finally, we present the key limitations of the current methods and identify knowledge gaps in the current body of research, and we subsequently proposed key areas for future investigation.

## 2. Methodology

The approach taken to collate existing publications on detailed façade reconstruction for this review was based on an investigation of building reconstruction described in other literature reviews [10,30,31]. This enabled an initial understanding of what had been achieved within the building reconstruction field in the past and the identification of relevant keywords to search for related literature. To the best of our knowledge, there is no literature review dedicated to façade opening extraction, aside from that by Neuhausen et al. [32], which presented methods for window detection from images.

The increased technological development in data collection and associated processing methods has impacted the scientific community, as reflected in the increase in the number of journal articles, books and series on facade-related studies (Figure 3). The origins of research related to façade reconstruction involved the processing and analysing of images. About a decade later, in 2011, the possibilities of point clouds for building reconstruction began to be explored. The year 2016 represents the advent of research introducing deep-learning methods, which, in 2021, comprised about two-thirds of publications on the use of deep learning in façade opening extraction-related topics. To order the content of the collated literature into a manageable taxonomy, a set of themes and questions was developed to guide this review (Table 1).



**Figure 3.** Number of façade opening extraction-related publications per year in journals, series and books according to the Web of Science portal based on the following keywords (in quotes) and Boolean Operators (in uppercases): (1) ‘façade parsing’ AND ‘image’, (2) ‘façade parsing’ AND ‘point cloud’, (3) ‘façade segmentation’ AND ‘image’ NOT ‘indoor’, (4) ‘façade segmentation’ AND ‘point cloud’ NOT ‘indoor’, (5) ‘window detection’ AND ‘image’ AND ‘façade’ NOT ‘indoor’, (6) ‘window detection’ AND ‘point cloud’ AND ‘façade’ NOT ‘indoor’, (7) ‘façade reconstruction’ AND ‘deep learning’, and (8) ‘façade segmentation’ AND ‘deep learning’.

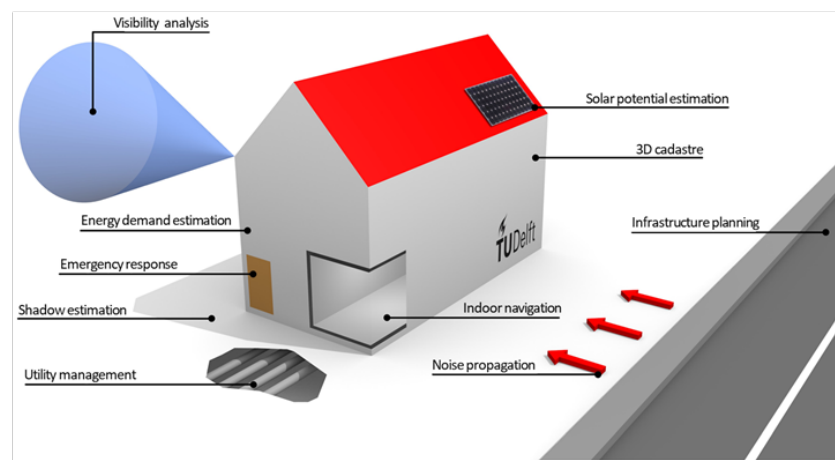
**Table 1.** Questions defining the main topics covered by the review.

Review Section	Questions and Example
Section 3: Applications	What are the applications that could benefit from more detailed 3D building model? For example, solar potential estimation, building information modelling and energy analysis.
Section 4: Platforms	What measurement technology is used? What strategy can deliver reliable and suitable data for façade opening extraction? For example, airborne, ground level, mobile and static platforms. Are there any emerging technologies? Handheld, mobile phones and volunteered geography information.
Section 5: Data types \representation	What are the advantages and disadvantages of data used for façade opening extraction? What are the challenges of existing methods? For example, what characteristics of range-based and image-based data should be considered when extracting façade openings.
Section 6: Scene interpreta- tions	How is the data processed? For example, analysis of point cloud density for façade opening extraction. Machine learning capabilities for façade opening extraction. What are the benefits of data fusion?

### 3. Applications

Simplified and coarse representations of buildings can serve as an effective source for analysis on either a city or national scale. However, a lack of detailed information on the façade elements poses difficulties for analysis at street or building scale [33]. Despite much research, the reconstruction of buildings with semantically rich façades remains a challenge due to the complexity of the task or lack of adequate data. Nevertheless, the realisation of an increasing number of applications that could benefit from detailed information about

the composition and appearance of existing buildings has sparked growing interest among academic and commercial researchers (Figure 4).



**Figure 4.** Applications of 3D building models for decision-making support. Reprinted with permission from Ref. [2].

### 3.1. Building Information Modelling

Among the applications requiring detailed façade information, building information modelling (BIM), solar potential analysis estimation and a broad spectrum of energy analyses are most common. BIMs are digital facility models containing rich semantic information. The potential of BIM is recognised among the Architecture, Construction, Engineering and Facility Management (ACE/FM) community, since BIMs facilitate communication between a range of stakeholders involved at different stages of a facility's life cycle.

BIM models, in most cases, are generated from blueprints at a level of detail that exceeds that which is typically obtained based on remotely sensed data [10]. However, despite the lower level of detail, creating 3D models of buildings directly from remotely sensed data for BIM is useful when blueprints are unavailable (e.g., historic buildings) or are old and outdated. In such cases, remote sensing methods allow the creation of a 3D representation of buildings in as-built/as-is conditions [34–37].

### 3.2. Solar Energy Potential

Solar energy potential estimation is another application for which detailed information about the façade and its elements is relevant. Roof surface analysis has been used for many years in solar energy estimation, often overlooking the possibilities of also exploiting walls. This is due to a tedious and time-consuming process if considering analyses on a city scale and not individual buildings [38] or the need for more sophisticated tools based on 3D geospatial data [39].

However, the growing interest in renewable energy policies has led to intensified research on the solar energy potential associated with vertical surfaces (e.g., walls) in the urban environment. This is of particular importance in the modern urban landscape because the area of a façade often exceeds that available on the roof. Another advantage of utilising vertical surfaces is that they are not typically affected by dust and snow cover during the winter months. In recent years, several studies have investigated the potential of solar energy from a building's façade (vertical surfaces) [39,40].

For instance, Desthieux et al. [39] assessed the solar energy potential for rooftops and façades in Geneva, Switzerland, by combining light detection and ranging (LiDAR) with 2D and 3D cadastre data (i.e., 3D building models, 2D roof layer, 2.5D digital surface model). Their experiment proved the practicality of 2D and 2.5D data for solar radiation analyses conducted on large urban scales. However, the obtained level of detail was insufficient when features, such as windows or balconies, were necessary for the analysis.

Catita et al. [40] investigated the relevance of roofs and façades for the photovoltaic potential of the University of Lisbon, Portugal, by combining aerial LiDAR data with 3D building models and solar radiation models. The authors presented a methodology addressing which façade is most suitable for solar panels or where on the façade a panel should be localised.

### 3.3. Energy Analysis

Detailed façade representations are also beneficial to applications under the broad spectrum of energy analysis. For instance, the size and orientation of windows can have a significant impact on the energy consumption of buildings [41]. To minimize energy wastage, the area of doors and windows should be less than or equal to 20% of the floor area [42]. Moreover, the number of floors (which can be estimated based on the number of windows) and the window-to-wall ratio, among other factors, can contribute to heat demand modelling [43].

Building height and the percentage of glaze were also among the many factors used by Nageler et al. [44] for dynamic urban building energy simulation. Other energy analysis applications that could make use of detailed façades include estimation of energy demand [45], energy-efficient retrofits planning [46], assessment of energy performance [47] and luminance mapping [48].

### 3.4. Civil Engineering

One of the important aspects considered by civil engineers is an assessment of the stiffness of buildings, the resistance of which can be weakened by ground movements (e.g., caused by underground construction or geohazards). In this case, information about building openings, and in particular windows, is considered since they affect the overall stiffness of the building structure [8,49,50]. Information on façade openings is also critical for evacuation planning and emergency response, where not only a geometrical representation of façade openings are essential, but also their semantic representations [4].

### 3.5. Other Applications

The applications highlighted above are domains within which the advantages of using more detailed building façades are clear. Other applications that could benefit from detailed façades include light pollution and shadow simulations [51], prerequisite for disaster management [52], movies and virtual reality [53], and potentially bird-window collision analysis [54–56] where the window angle of orientation plays an important role in assessing avian mortality. As outlined, the enrichment of volumetric 3D models with additional information about façade openings clearly has significant potential for opening new possibilities for more effective and accurate analyses and urban environment management.

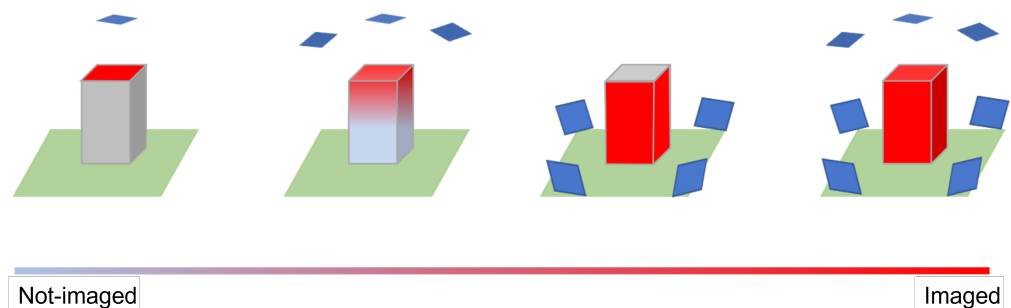
## 4. Platforms

Although research into the reconstruction of 3D buildings has increased in recent decades (see Figure 3), the majority of studies have sought to improve the accuracy of determining the external geometry of buildings and marginalizing the detection and reconstruction of building openings. In general, the data used for the reconstruction of the façade can be grouped according to the viewing perspective from which they were obtained or the type of data collected. Regarding the perspective, aerial and ground-based platforms are mostly used due to their viewing angle. As for the sensor, the most commonly used data are those acquired using ranging-based and imaging-based sensors (Figure 5).



**Figure 5.** Data commonly used for 3D building reconstruction. Reprinted with permission from Ref. [50].

The main factors that determine the choice of the platform from which the data for 3D building reconstruction are acquired are the complexity and height of the specific building feature/component of interest (e.g., roof, façade). Figure 6 shows what information about the building can be collected depending on the level of the platform. For instance, data acquired from aerial platforms are most suitable for reconstruction of the roof [57] or the top of a tall building. In contrast, data collected from ground level captures parts of buildings that are not visible from a plan view, such as the façade and its openings.



**Figure 6.** Building elements captured depending on the platform used. From left to right: nadir view allows rooftop detection, the oblique view extends nadir by mapping façades, terrestrial platforms capture façade information, and the combination of all allows mapping the whole building. Adapted with permission from Ref. [58].

Airborne platforms provide the most pertinent data for roof modelling, building footprint detection (oblique images), showcasing streets and mapping the locations of open spaces. However, the capabilities of airborne platforms for detailed 3D elevation modelling are limited. This is largely due to the downward-facing (nadir) viewing perspective offered by most sensors that are mounted to airborne platforms. Alternative solutions to mapping façades from airborne platforms include the use of sensors (e.g., cameras) that provide an oblique viewing perspective [59–61] or unmanned aerial vehicles [62].

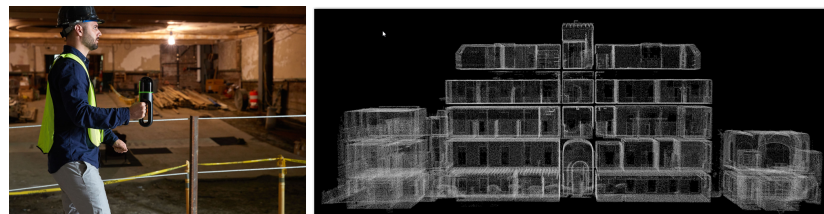
An airborne survey allows information for a large area to be readily acquired in a single flight but at a high cost. A solution to this is through the use of unmanned aerial vehicles (UAVs). Nevertheless, the use of UAVs for surveying complex urban environments poses several challenges of its own. First, a UAV survey requires knowledge and appropriate certificates on how to operate the drones. Moreover, the area and height at which UAVs

can operate in cities are limited due to the safety and security risks posed by skyscrapers and overhead power lines.

In addition to airborne platforms, data obtained from ground-based platforms are used to gather information for the 3D reconstruction of buildings. This group of platforms includes both static and mobile units (e.g., vehicle, backpack, robot). Static ground platforms consist of a sensor most often placed on a tripod, with data collected in stop-and-go mode. This means that, once all the required data are collected from one measuring point, the equipment is then moved to the next point to perform measurements from a different perspective [63]. Datasets acquired using static scanners are typically large, therefore, requiring considerable storage and time for subsequent processing. Moreover, collecting data for multiple buildings is time-consuming.

The availability of Global Positioning System (GPS) technology contributed to the creation of mobile mapping systems (MMS) [64] in which sensors are mounted on moving platforms (e.g., vehicle, rail cars, robot and backpack). Most commonly, mobile technologies are equipped with sensors, such as a scanner and camera, Global Navigation Satellite System (GNSS) and Inertial Measurement Unit (IMU), with data collected along the driving path. Mobile systems reduce the time and cost needed to undertake a survey, as they can be used to acquire data from large areas in a shorter amount of time.

In recent years, manual simultaneous localization and mapping (SLAM) based laser scanners have emerged, allowing the relatively seamless collection of 3D data (Figure 7). These handheld scanners are growing in popularity due to their measurement speed, low cost and ease of use. These devices combine different sensors (LiDAR sensors, RGB camera and IMUs), the data of which are used by SLAM technologies to track the position of the device while creating a mapped environment [65]. Handheld scanners are especially useful in the absence of a GPS signal which makes it difficult to determine the position of the sensor or platform (e.g., satellite signal blockage due to building, bridges, trees, indoor surveys and 'multipath' effects). Solutions based on the SLAM approach allow for faster and cheaper data acquisition at the expense of lower accuracy compared to traditional methods.

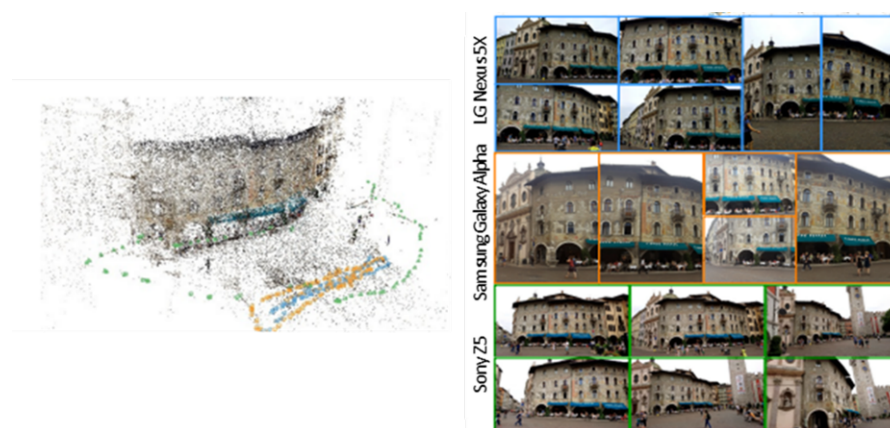


**Figure 7.** An example of handheld technology. Source: Leica.

Over the past decade, in addition to traditional solutions, there has been a growing interest in the geospatial community to use images acquired from smartphones. This data collection method is cheap and fast and enables the acquisition of images from different perspectives, making it possible to register façade elements not visible from other platforms. An additional advantage of mobile phones is the GNSS sensor with which the device is equipped.

The GNSS sensor allows for a relatively accurate determination of the position of the camera when taking an image, depending on the quality of the signal and the sensor [66]. The potential use of images acquired from mobile phones for building modelling has been investigated in several studies [67–69]. In these, phone images were used to create point clouds of buildings through the structure from motion (SfM) [70] algorithm. The presented results demonstrated that data acquired from mobile phones by multiple users could be successfully combined to derive a 3D point cloud [67] (Figure 8).





**Figure 8.** 3D building reconstruction from mobile phone images. Image boundary associate with the colour coded cameras in the point cloud. Reprinted with permission from Ref. [67].

Data collection from multiple users echoes the concept of ‘citizens as sensors’ of Goodchild [71], which opened up the world of ‘volunteered geography information’ (VGI) and ‘volunteered geography’ (VG) [72]. Due to the adoption of crowd-sourcing, many non-experts can now share geospatial data. For example, Fan et al. [73] designed an integrated platform to reconstruct 3D building models based on images collected by volunteers or non-expert users. The results appear promising for delivering more detailed building models, however, no information on the location accuracy was reported.

Despite the undisputed positives associated with citizens actively participating in data collection, this approach raises questions about the quality assurance [74]. As for using crowd-sourced data for a more detailed reconstruction of buildings, a set of well-designed data acquisition guidelines could ultimately help enhance façade openings detection.

For example, many images from different locations, collected under various conditions, could allow for more accurate recognition of city elements and thus the creation of more robust and flexible methods for openings detection and modelling. In addition, such data could be used to develop databases for training, validation and testing new strategies, where dedicated and diverse datasets could assist data-driven approaches.

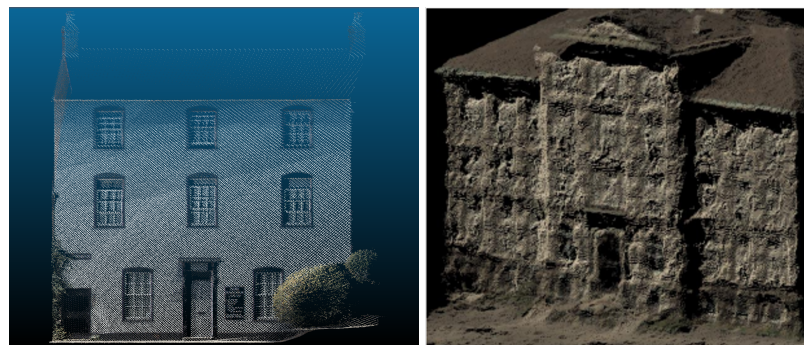
Data acquired from terrestrial platforms enable the collection of information on the lower parts of buildings, which allows subsequent identification of building openings. Nevertheless, data collected from this level poses several limitations. First, permission to enter private properties (e.g., residential neighbourhoods) may be restricted for legal reasons. Furthermore, objects on the ground can be obstacles that limit a sensor’s field of view. These include, among others, pedestrians, passing and standing cars or street furniture. These limitations pose difficulties in obtaining a complete dataset for a given area of interest and, consequently, lead to problems with recreating a complete and detailed 3D building model. Missing façade elements can be supplemented with data from aerial platforms obtained with an oblique camera. Although reconstructing detailed building façades is possible, it is still a challenging task. This is mainly due to constraints associated with the complexity of a building façade, occlusion and camera angle, radiometric changes in illumination and perspective in different images, sensitivity to sun glint and hot spots [75,76]. This leads to incomplete coverage and holes in the data that make it difficult to correctly identify objects or problems with image matching that can be used to replicate the point cloud from images (Figure 9).



**Figure 9.** Left: nadir, middle and right: oblique views of the same urban scene. The differences in radiometric resolution between the views, cause difficulties in image matching. Reprinted with permission from Ref. [77].

### 5. Data Types/Representation

The two most popular measurement techniques used in building reconstruction are laser scanning and photogrammetry. The basics for these techniques can be found in McManamon [78] and Mikhail [79] and Wolf et al. [80], respectively. Point cloud datasets obtained from laser scanning or multi-view stereo vision are commonly used for 3D building reconstruction and city mapping (Figure 10). Compared to data, such as 1D measured distance or projected 2D images, point clouds provide high-resolution data for accurately determining the shape, size and position of an object in 3D space, enabling object reconstruction and modelling [81].



**Figure 10.** An example of point cloud from terrestrial laser scanning (left, source OS © Crown copyright and database rights 2022 OS and dense multi-stereo matching from oblique aerial imagery (right, reprinted with permission from Ref. [82]).

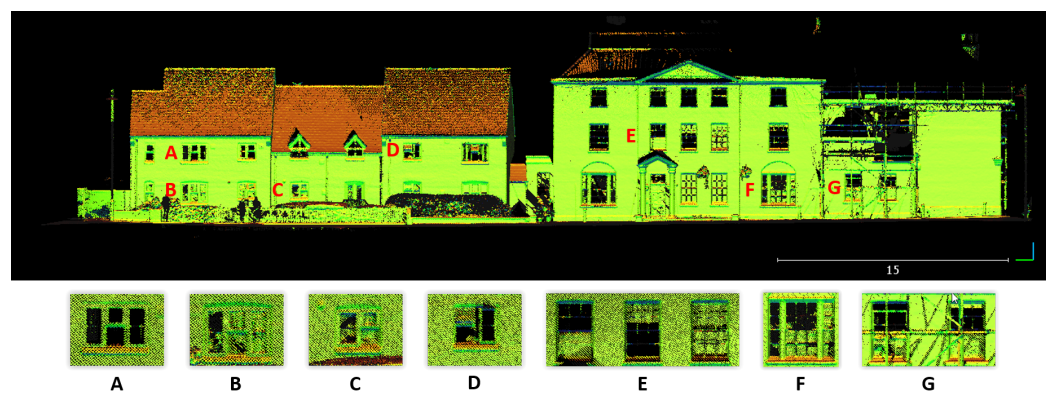
Range-based point clouds are quick to obtain, not affected by illumination and of high accuracy, which makes them a good source of data for 3D building reconstruction. However, the range-based point clouds have shortcomings which include lack of semantic information, high costs of the equipment and high technical requirements for operation [83].

Additionally, the distribution and density of points obtained is not always uniform and depends on the distance between the sensor and the object. A subsequent lack of information, high redundancy and low distribution at the edges of objects are all factors that can lead to difficulties in the reconstruction of complex buildings [84,85]. Moreover, differences in the accuracy and density of point clouds do not allow for a simple fusion of data obtained from different viewpoints and sensors [86].

Image-based point clouds are generated by applying triangulation to stereo-image pairs, such as through SfM. One of the main advantages of image-based point clouds is that they contain high-resolution textural and colour information, which enables the generation of photorealistic models [87]. Nevertheless, point clouds obtained from SfM can suffer from non-uniformities and may contain a higher noise level than range-based point clouds. Compared to the raw point cloud from range-based methods, in which the 3D information is obtained directly from the implicit scale factor, image-based point clouds require this information to be extracted from the images using camera parameters (i.e., intrinsic and extrinsic parameters).

One of the essential differences between the two types of point clouds is the accuracy, which, for range-based point clouds, is greater in the depth direction compared to image-based point clouds. On the other hand, the accuracy of the image-based product is greater in the direction perpendicular to the depth plane.

One of the point cloud characteristics used in detecting building openings is the presence of holes (i.e., data gaps) where glass objects are located, which is due to signal penetrating through the transparent surface (range-based) or difficulties with image matching for glass/transparent surfaces (image-based). This feature is helpful in the detection of elements, such as windows or glass doors; however, when the objects are covered (e.g., curtains and blinds), their detection based on the missing data is challenging. Figure 11 illustrates a point cloud with various examples of windows. As presented, the assumption that a hole represents windows in a point cloud can be misleading. Specifically, some of the windows are partially or entirely obscured, meaning that the holes in the point clouds do not match the actual shape of the window.



**Figure 11.** Mobile mapping system point cloud representing a building façade. (A)—window represented by holes without noise. (B–G)—samples of windows with varying degrees of obscuring. Source: OS © Crown copyright and database rights 2022 OS.

In addition to using images to create a point cloud, 2D image data itself provides a source of information that can be used for the semantic enrichment of building façades. For instance, the edges of objects are easier to detect in images, which allows the boundaries of objects to be easily defined. Furthermore, the colour information (presented in different colour spaces) can be used to segment the façade. Recky et al. [88] successfully adapted a k-means clustering in a CIE-Lab colour space to detect windows on complex historical buildings. Despite its advantages, the reconstruction of buildings based only on photogrammetric methods provides several challenges. Occlusions related to vegetation, pedestrians or vehicles, duplicate measurements of the same object, or insufficient data volume related to site accessibility are among the main factors that prevent an accurate reconstruction of objects.

## 6. Scene Interpretation

During the past decade, a large body of research has been conducted on 3D building reconstruction. Nevertheless, most of this deals with the manifold facets of the buildings. Specifically, there has been a primary focus on improving the accurate representation of the external shape of the building without consideration of its details. To the best of our knowledge, the work by Musialski et al. [30] is the most comprehensive review of the methods used for automatic and semi-automatic reconstruction of buildings from images and LiDAR data, which includes methods for window detection from image data.

The number of studies that addressed the subject of façade opening detection prior to 2013, when Musialski's work was published, was minimal. However, recent advances in

data processing approaches (e.g., computer vision and deep learning) and more accessible hardware has contributed to the growing interest in detecting façade openings.

Here, we present an update of more recent research on building reconstruction, with a focus on façade elements identification. With respect to scene interpretation methods, we assume that segmentation refers to the assignment of a label only to known objects, whereas scene parsing involves assigning annotations to all points. In contrast, object detection selects specific elements (e.g., windows and doors) from a given dataset.

### 6.1. 3D Scene Interpretation

Essentially, point clouds generated from laser scanning or multi-view stereo vision are commonly used for 3D reconstruction of buildings due to their ability to significantly enhance geometry reconstruction and surface modelling [81]. Interpreting scenes obtained from point clouds to extract information about the elevation of a building and its openings can be divided into two stages. The process begins with segmentation that involves grouping points belonging to one surface (e.g., a wall). These regions are further subdivided to obtain more detailed façade information (e.g., doors, windows and balconies) using feature extraction methods.

#### 6.1.1. Segmentation

Although the segmentation process does not directly translate into the delineation of the façade openings, it is a process that enables the distinction of those points within the point cloud that belong to the building (Figure 12). The general concept is to group points into homogeneous components based on the common features to exclude elements that are outside the area of interest, thus reducing the number of points and allowing for quicker processing. The advantages and disadvantages of common point cloud segmentation methods are presented in Table 2 and further information can be found in the work of Wang et al. [89] and Xu and Stilla [10].



**Figure 12.** Example of building segmentation. Raw point cloud collected from terrestrial laser scanner with RGB values. Source OS © Crown copyright and database rights 2022.

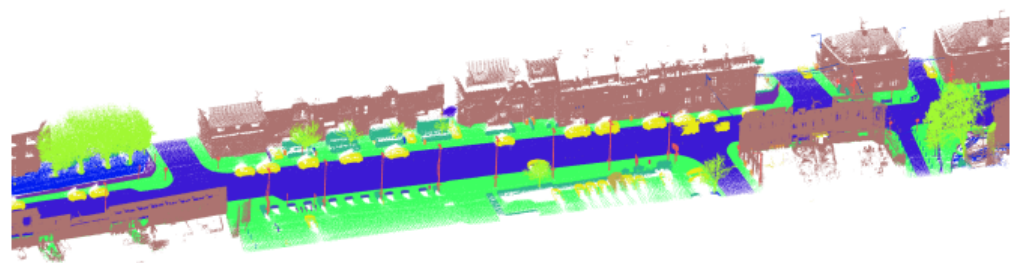
**Table 2.** Summary of different point cloud segmentation methods for building reconstruction

Method	Advantages	Disadvantages	Examples
Model-based	Resilience to outliers and noise	Computationally expensive, time-consuming, high memory consumption, difficulties with reconstruction of objects for which there are no mathematical expressions	[90]
Region-growing	Good preservation of edges and boundaries of surfaces and artifacts	Sensitive to outliers and noise, over segmentation	[91–93]
Clustering-based	Easy to implement, no need for setting of seeds	Over and under segmentation, high computational costs, points on the edges may meet the requirements of more than one cluster	[94,95]

### 6.1.2. Scene Parsing

Façade parsing is a crucial step in the 3D building reconstruction process that leads to the classification of input data into semantic regions, such as walls, doors and windows [96] (Figure 13). Façade parsing enables a better urban scene understanding and enables more effective information storage on building façades. However, most examples of point cloud semantic segmentation of the urban environment do not include the detection of openings in the building façade.

Li et al. [24] introduced a façade parsing approach for which a point cloud is first decomposed into depth planes. Next, façade elements are detected using a combination of machine-learning-based classification and prior semantic knowledge. The overall performance of the method presents good results. However, it can be challenging to apply if the occluded area varies from the visible façade elements, hindering accurate reconstruction of the occluded parts of façade. To improve the method, the authors suggested considering the use of weak architectural principles or fusing the point cloud with images or GIS data.



**Figure 13.** Part of Paris-Lille-3D dataset. Semantic labels for each object class, with unique colour for each class. Reprinted with permission from Ref. [97].

### 6.1.3. Façade Openings Extraction

The next step in the detailed 3D façade reconstruction is the extraction of the façade openings. Current façade openings detection strategies for point clouds can be categorized into hole-based and rule/symmetry-based (Table 3). Most building openings have glass

elements, and their characteristic ‘transparency’, which generates information gaps in the point cloud, can be exploited for the detection of openings [98,99]. A second characteristic often used to detect openings from point clouds is their typical rectangular shape represented by two vertical and horizontal edges.

Hole-based opening detection methods search for empty spaces in a point cloud and rely heavily on the point cloud density. The reason for this is twofold. On the one hand, a dense raw point cloud for walls and low point density for glass objects allows the detection of gaps in the point cloud. On the other hand, a low raw point density may be insufficient for the reliable generation of feature geometries [100].

Many studies on hole-based opening detection use terrestrial laser scanners (TLS) since their data density typically exceeds that of point clouds obtained from mobile mapping systems (MMS) or aerial laser systems (ALS). Pu and Vosselman [101] tested their method based on a Triangular Irregular Network of data from TLS (average point density 500 pts/m<sup>2</sup>) and MMS (around 100 pts/m<sup>2</sup>); while the method recovered the outlines of building openings from the TLS data, the point cloud density from MMS was too low to detect windows and doors.

High-density point clouds of the urban environment are not always available as TLS data collection, compared to MMS and ALS, is more costly and time-consuming when surveying large areas. The additional inconvenience of a highly-dense point cloud is the additional data storage and management required.

To address these issues and assess the effect of point cloud density on correct façade opening detection, Zolanvari et al. [102] tested their slicing method on three buildings of diverse size and complexity. They suggested that a density of 130 pts/m<sup>2</sup> should be sufficient to achieve at least 90% of accuracy. Although hole-based methods allow for façade opening detection, they present several shortcomings, such as sensitivity to outliers and noisy points, user intervention [101,103,104] and challenges with detecting windows covered by curtains [104–106].

Another façade opening detection strategy for 3D point clouds involves rule and symmetry-based methods. Rules play an important role in detailed façade reconstruction as openings follow architectural rules concerning their shape and spatial relationships (e.g., the locations of windows and doors relative to walls). Rules can be applied to openings extracted from corner detectors [9,107], a graph model [73], machine learning [108] or based on the assumption of symmetry in window distribution [109].

Applying rule and symmetry-based methods allow detection of openings from partially missing data. However, these methods are heavily depend on symmetry, and repetitive elements can affect their robustness [110]. Moreover, the approach is unsuitable for low-rise buildings, as vertically repeated features are unlikely to be present. This diversity, complexity and asymmetry of the façade structure does not allow for the creation of flexible rules describing the façades and their elements, posing problems in automating the reconstruction process of detailed 3D building models.

To address this problem, Fan et al. [73] proposed façade layout graph model method based on Gestalt [111] principles and the principle of architectural form [112]. Although the proposed method describes different types of façade structures, its applicability for complex structures and automation needs further improvements.

The methods outlined are often tested on a limited number of datasets that are lacking in the representation of buildings with diverse architectural styles. The lack of benchmark datasets with various styles makes it difficult to develop and test new methods (e.g., deep-learning based) for façade opening detection or point cloud parsing. Despite a recent increase in the number and availability point clouds of urban environments, their classification does not include façade openings.

**Table 3.** Summary of different methods for façade opening extraction from point cloud.

Method	Advantages	Disadvantages	Examples
Hole-based	The assumption that windows are an integral part of the façade reduces the search area	Depends on point cloud density, misclassification due to occlusion holes, points reflected from curtains may be similar to those reflected from wall causing difficulties in distinguishing façade openings, high-density point cloud require more data storage	[100–106,113,114]
Rule-symmetry-based	Successful with low resolution data, allows detection from partially missing data	Necessary user intervention to create rules, high reliability on symmetry and rules, difficulties with application for non-symmetrical façade types	[9,73,107,108,115]

## 6.2. 2D Scene Interpretation

While this review focuses on 3D building reconstruction, it is also crucial to outline methods used to detect façade elements from 2D data for two main reasons. First, there is increasing amount of data being collected by MMS that, in addition to the acquiring images, also provide the camera model parameters, GPS and IMU data [116], allowing for 2D to 3D projection of data. Secondly, the 2D images allow the isolation of features with linear components (e.g., windows and doors) that can improve the quality of the final detailed building model [117,118].

### 6.2.1. Scene Parsing

Image semantic segmentation is a process of adding a semantic class to every image pixel. Generally, segmentation-based methods for extracting façade elements from 2D images can be divided into two groups (Table 4). These are: (i) bottom-up methods [119–123]; (ii) grammar driven (top-down) methods [124–126].

Bottom-up methods solve the image segmentation problem at the pixel level. Although bottom-up methods deliver satisfactory pixel-wise results, a result of the segmentation is not suitable for every application due to divergence from basic architectural rules; for instance, irregular shape (pixel-wise segmentation may not result in straight and even lines) and incorrect size of windows or doors. These irregularities may be enough to estimate the glass-to-wall ratio, which might be suitable for thermal related applications.

However, their usability for constructions-related applications, such as BIM, is insufficient. The reason for pixel labelling challenges is that radiometric variation in images, dynamic or static occlusion, and shadows can hinder the precise recognition of objects. The above-mentioned issues can lead to a partial or complete gap in the segmented scene. To improve the quality of bottom-up methods, weak architectural constraints can be applied [127–129], thus, reducing occlusion issues and improving the overall accuracy; nevertheless, issues, such as the window’s vertical alignment, still persist.

Recent bottom-up methods rely on convolutional networks, however, the lack of datasets with diverse architectural styles poses difficulties for 2D data segmentation [130,131]. Specifically, Lotte et al. [130] trained a neural network on different available datasets and then

tested it on an entirely new dataset. The results suggest the need for accurate annotation of the images and the lack of an exact extraction of façade elements on an unknown data set, except where training and testing datasets share similar features. Moreover, Kong and Fan [131] note that existing benchmark datasets constitute rectified images without distortion, which are different from those acquired from the mobile mapping systems. Hence, at present, methods developed on specific datasets cannot be easily transferable to other scenarios.

Top-down methods segment images based on grammar rules. As opposed to bottom-up methods, the implementation of grammar rules supports the prediction of the façade element's location, which is particularly important in the case of occlusion or changes in illumination. Moreover, applying the grammar rules restrict the area in which the algorithm searches for elements, as it considers only these segments that follow the rules.

Nevertheless, grammatical rules involve significant human intervention and prior knowledge to construct the rules describing a given architectural style, making it difficult to automate and adapt the methods to different data or sites. To reduce user intervention automatic learning of rules with ground-truth image annotations have been proposed [124,132].

For example, Gadde et al. [124] experimented with applying a set of grammar rules learned on one architectural style to another style. Their results indicated that the location of the façade elements is not ground truthful, therefore demonstrating the problems with the interchangeable application of grammatical rules.

**Table 4.** Summary of different methods for façade parsing from images.

Method	Advantages	Disadvantages	Examples
Bottom-up	Better performance compared to rule-based methods [131], less prior-knowledge necessary	Susceptible to changes in lighting, irregular shape of the output (e.d., edges are not straight lines), occlusion hinders precise segmentation, lack of diverse training datasets	[53,100–102,104,106,113,114,119,130,131,133–135]
Top-down	Higher prediction of façade element's location, better structure of detected façade elements	Rely on strong prior knowledge, difficulties in applying rules for one architectural style to other, lack of flexibility when facade is changed	[124–126,132,136]

### 6.2.2. Façade Openings Extraction

The studies presented above focus on parsing an entire building façade, however, there is a considerable body of literature dedicated to window detection from 2D images. Neuheusen et al. [32] published the most recent overview of window detection from images. In that review, the methods are categorized into the following groups: grammar-based [137], image processing [138] and machine learning [122]. The same author has since focused on window detection from image data through a series of studies on window detection for risk assessment analysis [8,50,139].

These specific studies involved the assessment of a soft cascade classifier for window detection. Through a series of experiments, the authors proposed a final three-step pipeline. The process commences with image rectification, then the detection of objects is performed by a sliding window detector with a soft cascade classifier consisting of thresholded Haar-like features (Figure 14). The last step concerns post-processing to refine the detection step.



In future work, the authors plan to categorize windows by a type similar to that presented by Lee and Nevatia [140].



**Figure 14.** Example results of windows (presented as green rectangles) detected in image. Reprinted with permission from Ref. [50].

One of the main elements of the façades is entrance doors, the location of which could find applications in navigation [141], emergency response and flood analysis. Nevertheless, there has been little research dedicated to external front door detection. In contrast, there is a considerable body of research on door detection and localization in indoor environments. This has largely been driven by a growing demand for a variety of applications, such as navigational assistance [142] and 3D building reconstruction [143].

Many of these case studies are associated with mobility restrictions for disabled people or intelligence systems allowing safe navigation through indoor spaces. Although, in theory, the appearance of doors from inside and outside is identical, their characteristics differ significantly in terms of their surroundings, occlusion and lighting, which limits the interchangeability of the same methods for door detection in varying environments.

In the case of entrance door detection in outdoor scenes, Liu et al. [144], Kang et al. [145] and Talebi et al. [146] tackled this by implementing vision-based techniques, such as edge detection, line segments or local features to detect front doors from RGB images. The limitations of these studies are that they use small and relatively simple datasets to assess the results [147] or, in the case of Talebi et al. [146], have challenges with door detection from strongly tilted images and when the door and wall are almost identical in colour.

### 6.3. Data Fusion for Scene Interpretation

Many approaches to reconstructing façades utilise single-source datasets, where data from one source (e.g., camera and scanner) but multiple viewpoints are linked together. While this approach can cover a large area of interest, it does not mitigate or eliminate the limitations associated with a given technology. For example, point cloud data are unstructured, and the features determined directly from them are characterized by low geometric continuity. On the other hand, 2D images, which can be registered for real-world scale according to a point cloud, make it possible to describe the spatial distribution of elements in the image (e.g., linear features).

Therefore, the combination of the features extracted from both images and a point cloud allows for a more accurate reconstruction of building façades and their elements. Such a solution can be synergistic through supplementing the limitations of one method with the advantages of another (Figure 15), providing more reliable information [148].

	Camera	LiDAR	Camera + LiDAR
Object detection	●	●	●
Object classification	●	●	●
Distance measurement	●	●	●
Object edge precision	●	●	●
Range	●	●	●
Colour	●	●	●
3D reconstruction	●	●	●
Contrast, resolution	●	●	●
Poor weather performance	●	●	●
Low-light performance	●	●	●

**Figure 15.** Sensing comparison: LiDAR vs. Camera. Green refers to good capability, yellow to fair and red to poor capability.

Although LiDAR and cameras are the most popular data sources used for 3D building reconstruction, both methods have limitations, that can be overcome by fusing data of different modalities. Accordingly, multimodal sensor fusion can increase the level of detail and quality of a building model and enable the creation of more information-rich 3D city models. For instance, incorporating spectral information from hyperspectral imagery would allow identification of the material properties of a wall.

Furthermore, combining data from thermal imagery allows for thermal attribute mapping [149–151], while multispectral information can help assess the conservation state of a building [152]. An interesting example of multimodal data fusion is reported by Jarzabek-Rychard et al. [153]. In that study, 3D scene reconstruction is obtained from Red-Green-Blue (RGB) and thermal infrared (TIR) images. The authors investigated the applicability of spectral and geometric features for thermal analysis and the usability of a 3D thermal point cloud for façade opening extraction through a set of experiments.

The same type of data was used in Lin et al. [154] to detect windows through a combination of geometric features and thermal attributes. One of the fundamental assumptions of this approach was that there is a significant difference in temperature between windows and their surroundings (e.g., a wall). In some cases, due to the reflections from warm objects, this temperature difference was apparent and, therefore, caused misclassification. Additional grammar rules and topology analysis could help improve the performance of this method.

The fusion of 3D point clouds and 2D optical images for façade feature extraction has been the subject of several studies [98,118,148,155]. Pu and Vosselman [118] used point cloud data to establish the general structure of a building façade before applying computer vision edge detection on the images. Linear features in the images were then compared with model edges in order to improve the final result. The authors concluded that 3D point clouds enable straightforward extraction of areas and surface normals, which is challenging with 2D image data.

On the other hand, image data outperforms point clouds when detecting boundaries due to overall higher image resolution. Wang et al. [155] explored the fusion of a 3D point cloud and image data in three steps. First, the structural information about a building façade was retrieved from images. This was followed by the exploration of different methods for mapping between the image feature and the 3D point cloud. In the last step, the point cloud optimisation is performed by considering the structural information.

The potential limitation of the method is the difficulty with the feature extraction of complex façades and missing data due to occlusions. Becker and Haala [98] integrated information from a TLS and camera with a 3D city model. In the first stage of this approach, point cloud data was used to retrieve larger building parts, and then the images were processed to extract more detailed elements, such as window crossbars.

## 7. Limitations and Challenges

Musialski et al. [30] identified three challenges affecting the methods of that time: full automation, quality and scalability and acquisition constraints. Despite the passage of time and the development of new algorithms and more sophisticated hardware, these problems still persist nearly a decade later. Moreover, a few more factors can be added to the list of limitations highlighted by Musialski et al. [30].

- *Automation:* Automation of 3D building reconstruction has been an aim for researchers over the last couple of decades. However, the complexity of this task is high since we seek to reconstruct a well-defined underlying object; however, as we do not know the underlying object, it needs to be estimated from the data. Moreover, combining the above-mentioned data acquisition methods with the presented data processing techniques still relies upon many manual steps. Many new approaches require human intervention in setting up processing parameters, resulting in low levels of automation. Improving automation of the 3D building reconstruction by limiting the human input should involve designing concepts that create more intelligent and flexible workflows.
- *Quality:* The reconstruction of 3D buildings is directly related to level of detail (LoD) and level of accuracy (LoA), which are selected according to the application of the 3D model, method of data acquisition, processing and labour costs. The current low level of automation in the 3D reconstruction of buildings leads to the necessity of operator intervention. This intervention may concern data capture, data quality control, data processing settings, or the final result of the reconstruction assessment. Each operator has different knowledge and experience, which may affect the quality of the final 3D model. Therefore, it is imperative to develop good practice guidelines to eliminate the subjectivity between operators. Automation can be the solution to this problem. Hence, it is necessary to develop solutions for autonomous data processing that would minimize the need for intervention by an experienced operator.
- *Data acquisition:* Due to its complexity, it is challenging to obtain a complete dataset for an urban environment. For instance, differences in the dimensions of the buildings lead to differences in the characteristics of the acquired dataset, e.g., a terrestrial point cloud captured for a tall building will be less dense in the higher parts as the sensor's distance from the target will increase. Moreover, issues accessing buildings can result in gaps in the acquired point clouds and hence lower quality of the data used in 3D building reconstruction. Incomplete data are often the result of static objects (e.g., vegetation) or dynamic objects (e.g., vehicles and pedestrians) obscuring parts of buildings, while data acquisition at different times of the day may help limit dynamic occlusions [156], removing static occlusions is a more complex problem. In this regard, the fusion of data collected from various platforms offering different fields-of-view can help reduce the impact of static occlusions. For example, data from oblique cameras show blueprints of the building, which are hidden behind the stationary occluding object in the ground-level image. Another factor limiting data acquisition is the sensitivity of optical data to weather conditions (e.g., shadow and direct sunshine), resulting in a non-uniform image. One solution to this problem is incorporating data from laser scanning, which is more resistant to lightning and weather conditions.

## 8. Knowledge Gaps and Future Directions

- *Deep-learning implementation:* There has been a recent rapid increase in the development and use of deep-learning algorithms for processing point clouds and images. However, the full application of these methods in reconstructing 3D buildings is still in its infancy. One of the reasons for this is the lack of proven, state-of-the-art methods that could be practically applied in commercial projects. In addition, most of the developed algorithms focus on modelling indoor environments that differ from the outdoor environment characterized by lower quality data (higher noise and outliers).

The significant differences between outdoor and indoor environments therefore preclude the opportunity to interchange techniques used to reconstruct objects within each of these environments. Additionally, deep-learning approaches require a significant amount of test data to train the algorithm in order to recognize individual elements. The preparation of such data requires considerable time and manual work. Data collection itself is an expensive task, especially in the case of laser scanning over vast areas, and this is partly responsible for limiting the number of available datasets that could be used for training an algorithm.

A possible solution to overcoming the requirement of an extensive training dataset is the current data-centric deep-learning trend that emphasizes data quality rather than quantity. Focusing on well-prepared training data instead of the selection of the hyperparameters of a deep-learning model could help in the development of more robust workflows for the reconstruction of 3D building models.

- *Benchmark Datasets:* Publicly available benchmark datasets are essential for the development and evaluation of new methods and algorithms since they enable the direct comparison with the existing strategies [157]. Despite the advantages and demand for benchmark datasets, the number of suitable, readily-available datasets available for the detailed segmentation of building façades remains low. One of the reasons for this may be the relatively recent interest in building modelling beyond its traditional representation in the form of a volumetric model.

Additionally, let us assume that the façade openings are not only to be located on an image or a point cloud but also modelled (i.e., fitted into the body of the building). In this case, information is needed to reconstruct the object's location in space, while it should not be an issue with point clouds, 2D image-based benchmark datasets need to include ancillary information, such as camera calibration parameters and sensor locations.

Only such sets of comparative data permit an accurate validation of the results of 3D building reconstruction approaches. An additional requirement for creating a versatile and universal benchmark dataset is sufficient diversity of feature classes and their corresponding labels. Although currently low, the number of available benchmark datasets for complex urban environments is increasing with the advent of smart cars and autonomous vehicle applications (Table 5).

Even though these datasets, which often consist of both 3D and 2D data, have a specific category of 'building', they do not typically include sub-categories, such as windows or doors. This is, however, slowly changing and since 2020 there has been a growing interest in point cloud data that includes façade-level classes, such as windows, doors, balconies [158,159]. Benchmark datasets based only on images contain more detailed classes, which sometimes include windows and doors, however, they often lack the internal and external parameters, therefore making 3D reconstruction of building façades challenging.

Table 5 summarises benchmark datasets commonly used for detailed façade segmentation (with camera images) as well as data from multimodal sensors with the potential to be used for complex 3D building reconstruction. However, it is worth recognising the problems related to the use of benchmark datasets presented in Table 5, especially the fact that these datasets were not explicitly designed for the detailed reconstruction of buildings. Benchmark datasets are designed with a specific application in mind; however, in many cases, their use goes beyond the boundaries of the tasks for which these datasets were originally created [160].

In a recent study by Koch et al. [161], the problem of 'heavy borrowing' of datasets within the machine-learning community was noted, where data created to solve one problem are used to solve another task, which can lead to misalignment. Many of the available benchmark datasets are generated in a specific geographic location, and using them elsewhere may yield poor outcomes. For example, a deep-learning

model trained on data collected in the United States may not perform well on images from Asia.

Additionally, factors, such as the sun path, weather conditions, variety of backgrounds and different camera specifications, can affect the model's accuracy. At the moment, there are no available well-constructed benchmark datasets that can be used for detailed 3D building reconstruction. Despite the possibility of using the data presented in Table 5, it is necessary to approach them critically and verify that they are appropriate for the proposed application, or whether there is a need to create an entirely new dataset.

- *Computing power and data bottlenecks:* The variety of applications utilising 3D building models is leading to an increasing demand for data at different scales, from global to local. This, in turn, is associated with an increase in the amount of data that needs to be maintained, transmitted and processed. Although sophisticated methods for building reconstruction may produce more desirable results, the time associated with processing large volumes of data is typically greater than that of less complex techniques, subsequently reducing their cost-effectiveness.

The solution to the big data processing problem may lie with high-performance computing (HPC) technologies, a set of tools and systems that provide significantly higher performance than a typical desktop computer, laptop, or workstation. HPC can be used to overcome the issue of slow data processing on a local system or limitations of CPU capacity. However, since HPC is often chargeable, a common method is to run a small-scale test on a local device and then migrate the analysis to HPC [162]. Ultimately, coping with big data may contribute to the development of machine-learning algorithms and the search for efficient data storage of pipelines [163,164].

**Table 5.** List of the benchmark datasets from mobile (t.m.) and static (t.s.) terrestrial platforms with labels. For data from multimodal technology, the annotations do not specify building openings. Information about building openings is distinguished in the datasets from the single-sensor technology.

Multimodal sensor							
	Platform	Position (GPS,IMU)	LiDAR	Cameras	Building	Openings	Objects
A2D2	Geyer et al. [165]	t.m.	Yes	Yes	Yes	No	3D, 2D
KITTI	Geiger et al. [166]	t.m.	Yes	Yes	Yes	No	3D, 2D
Apollo Space	Huang et al. [167]	t.m.	Yes	Yes	Yes	No	3D, 2D
Mapillary	Neuhold et al. [168]	t.m.	Yes	Yes	Yes	No	3D, 2D
CityScape	Cordts et al. [169]	t.m.	Yes	Yes	Yes	No	3D, 2D
TUM-FACADE	[158]	t.m.	No	Yes	No	Yes	3D
Semantic3D	[170]	t.s.	Yes	Yes	No	Yes	3D
IDD: India Driving Dataset	Varma et al. [171]	t.m.	Yes	Yes	Yes	No	3D,2D
Single sensor							
eTRIMS	Korč and Förstner [172]	t.s.	No	No	Yes	Yes	2D
RueMonge	Riemenschneider et al. [173]	t.s.	No	No	Yes	Yes	2D
CMP	Tyleček and Šára [174]	t.s.	No	No	Yes	Yes	2D
BDD100K	Yu et al. [175]	t.m.	Yes	No	Yes	Yes	2D

- *Completeness:* 3D building reconstruction from images is a reverse-engineering task that requires methods to validate newly produced results. One of the metrics for estimating

the quality of the obtained 3D building model is completeness. Completeness has two kinds of errors: omission and commission. The first relates to omitted building elements, while the second concerns elements incorrectly classified as part of the building.

Assessing the completeness of 3D building models generated from terrestrial images is a challenging task due to elements that obscure the façade of the building (e.g., cars, people and vegetation) and a lack of reference data. Although blueprints and BIM can serve as a reference to assess the surface-based performance of developed algorithms, the available BIM and blueprints data are again limited. Therefore, it is essential to develop methods to readily determine the completeness of detailed building models.

## 9. Conclusions

Detailed façade segmentation, which includes the detection and location of building openings, is becoming an essential topic among researchers and industry professionals. However, despite the growing interest in more accurate building modelling, the detection of façade elements still needs to be investigated, and many of the challenges require further attention. In this review, we provided a comprehensive overview of the techniques and sensors used to acquire data to detect and reconstruct building façades and their openings.

The benefits, limitations and research gaps in the field of detailed 3D building reconstruction with a focus on opening extraction have been discussed. The main limitations of the current methods include the problems with the automation of processes that require user intervention and pre- and post-processing. User intervention throughout the entire workflow also influences the final product's quality, indicating the need for operator guidance.

Finally, the constraints associated with obtaining complete data representing buildings pose a challenge to the detailed reconstruction of buildings. Through assessment of the state of the art, the presented overview allows conclusions to be drawn about future research directions. The first is data fusion from multi-modal sources and different platforms to minimize the obscuring of façade elements by static and dynamic elements of the urban environment.

In addition, there is a need for more diverse sets of comparative data that would increase the possibility of creating robust and reproducible solutions for detailed 3D building reconstruction. Another potential direction for future research is the adaptation of computer vision, machine-learning and deep-learning methods to automate the process and limit user interference.

Despite the growing interest in using these strategies, well-established, proven and publicly available solutions are lacking. Finally, there is a need to establish quality matrices to help evaluate the proposed strategies and to develop pipelines for completeness, effectiveness and efficiency.

**Author Contributions:** Conceptualization, A.K., S.C., R.L. and S.G.; methodology, A.K.; formal analysis, A.K.; investigation, A.K.; writing—original draft preparation, A.K.; writing—review and editing, S.C., R.L. and S.G.; visualization, A.K.; supervision, S.C., R.L. and S.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Engineering and Physical Sciences Research Council Centre for Doctoral Training in Geospatial Systems (grant number EP/S023577/1) with the support of the Ordnance Survey.

**Data Availability Statement:** No new data were created during this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Billen, R.; Cutting-Decelle, A.; Marina, O.; De Almeida, J.; Cagliioni, M.; Falquet, G.; Leduc, T.; Métral, C.; Moreau, G.; Perret, J.; et al. *3D City Models and Urban Information: Current Issues*; EDP Sciences: Ulys, France, 2014.
2. Biljecki, F.; Stoter, J.; Ledoux, H.; Zlatanova, S.; Çöltekin, A. Applications of 3D city models: State of the art review. *SPRS Int. J. -Geo-Inf.* **2015**, *4*, 2842–2889. [[CrossRef](#)]

3. Stoter, A.; Ploeger, J.; Roes, H.; Van Der Riet, R.; Biljecki, E.; Ledoux, F. First 3D Cadastral Registration of Multi-Level Ownerships Rights in the Netherlands. In Proceedings of the 5th International FIG Workshop on 3D Cadastres, Athens, Greece, 18–20 October 2016; pp. 491–504.
4. Chen, L.C.; Wu, C.H.; Shen, T.S.; Chou, C.C. The application of geometric network models and building information models in geospatial environments for fire-fighting simulations. *Comput. Environ. Urban Syst.* **2014**, *45*, 1–12. [[CrossRef](#)]
5. Kaňuk, J.; Gallay, M.; Hofierka, J. Generating time series of virtual 3-D city models using a retrospective approach. *Landsc. Urban Plan.* **2015**, *139*, 40–53. [[CrossRef](#)]
6. Prandi, F.; Soave, M.; Devigili, F.; Andreolli, M.; De Amicis, R. Services Oriented Smart City Platform Based On 3d City Model Visualization. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *2*, 59. [[CrossRef](#)]
7. Ruohomaki, T.; Airaksinen, E.; Huuska, P.; Kesaniemi, O.; Martikka, M.; Suomisto, J. Smart City Platform Enabling Digital Twin. In Proceedings of the Ninth International Conference on Intelligent Systems 2018: Theory, Research and Innovation in Applications, IS 2018-Proceedings, Funchal, Portugal, 25–27 September 2018. [[CrossRef](#)]
8. Neuhausen, M.; Obel, M.; Martin, A.; Mark, P.; König, M. Window detection in facade images for risk assessment in tunneling. *Vis. Eng.* **2018**, *6*, 1–16. [[CrossRef](#)]
9. Gruen, A.; Schubiger, S.; Qin, R.; Schrotter, G.; Xiong, B.; Li, J.; Ling, X.; Xiao, C.; Yao, S.; Nuesch, F. Semantically enriched high resolution LoD 3 building model generation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 11–18. [[CrossRef](#)]
10. Xu, Y.; Stilla, U. Toward Building and Civil Infrastructure Reconstruction from Point Clouds: A Review on Data and Key Techniques; *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 2857–2885. [[CrossRef](#)]
11. Min, Z.; Wang, J.; Meng, M.Q. Joint Rigid Registration of Multiple Generalized Point Sets with Hybrid Mixture Models. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 334–347. [[CrossRef](#)]
12. Zhu, J.; Jiang, Z.; Evangelidis, G.D.; Zhang, C.; Pang, S.; Li, Z. Efficient registration of multi-view point sets by K-means clustering. *Inf. Sci.* **2019**, *488*, 205–218. [[CrossRef](#)]
13. Cheng, L.; Chen, S.; Liu, X.; Xu, H.; Wu, Y.; Li, M.; Chen, Y. Registration of Laser Scanning Point Clouds: A Review. *Sensors* **2018**, *18*, 1641. [[CrossRef](#)] [[PubMed](#)]
14. Elbaz, G.; Avraham, T.; Fischer, A. 3D point cloud registration for localization using a deep neural network auto-encoder. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017. [[CrossRef](#)]
15. Yang, B.; Zang, Y. Automated registration of dense terrestrial laser-scanning point clouds using curves. *ISPRS J. Photogramm. Remote Sens.* **2014**, *95*, 109–121. [[CrossRef](#)]
16. Xie, J.; Hsu, Y.F.; Feris, R.S.; Sun, M.T. Fine Registration of 3D Point Clouds with ICP Using an RGB-D Camera. *J. Vis. Commun. Image Represent.* **2013**, *6938*, 194–204.
17. Fond, A.; Berger, M.O.; Simon, G. Model-image registration of a building’s facade based on dense semantic segmentation. *Comput. Vis. Image Underst.* **2021**, *206*, 103185. [[CrossRef](#)]
18. Piasco, N.; Sidibé, D.; Demonceaux, C.; Gouet-Brunet, V. A survey on Visual-Based Localization: On the benefit of heterogeneous data. *Pattern Recognit.* **2018**, *74*, 90–109. [[CrossRef](#)]
19. Sattler, T.; Maddern, W.; Toft, C.; Torii, A.; Hammarstrand, L.; Stenborg, E.; Safari, D.; Okutomi, M.; Pollefeys, M.; Sivic, J.; et al. Benchmarking 6DOF Outdoor Visual Localization in Changing Conditions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018. [[CrossRef](#)]
20. Schonberger, J.L.; Pollefeys, M.; Geiger, A.; Sattler, T. Semantic Visual Localization. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018. [[CrossRef](#)]
21. Sattler, T.; Leibe, B.; Kobbelt, L. Efficient & Effective Prioritized Matching for Large-Scale Image-Based Localization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 1744–1756. [[CrossRef](#)]
22. Li, M.; Nan, L.; Smith, N.; Wonka, P. Reconstructing building mass models from UAV images. *Comput. Graph.* **2016**, *54*, 84–93. [[CrossRef](#)]
23. Zhang, L.; Li, Z.; Li, A.; Liu, F. Large-scale urban point cloud labeling and reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 86–100. [[CrossRef](#)]
24. Li, Z.; Zhang, L.; Mathiopoulos, P.T.; Liu, F.; Zhang, L.; Li, S.; Liu, H. A hierarchical methodology for urban facade parsing from TLS point clouds. *ISPRS J. Photogramm. Remote Sens.* **2017**, *123*, 75–93. [[CrossRef](#)]
25. Rusu, R.B.; Marton, Z.C.; Blodow, N.; Dolha, M.; Beetz, M. Towards 3D Point cloud based object maps for household environments. *Robot. Auton. Syst.* **2008**, *56*, 927–941. [[CrossRef](#)]
26. Ochmann, S.; Vock, R.; Klein, R. Automatic reconstruction of fully volumetric 3D building models from oriented point clouds. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 251–262. [[CrossRef](#)]
27. Lin, H.; Gao, J.; Zhou, Y.; Lu, G.; Ye, M.; Zhang, C.; Liu, L.; Yang, R. Semantic decomposition and reconstruction of residential scenes from LiDAR data. *ACM Trans. Graph.* **2013**, *32*, 1–10. [[CrossRef](#)]
28. Pomerleau, F.; Colas, F.; Siegwart, R. A Review of Point Cloud Registration Algorithms for Mobile Robotics. *Found. Trends Robot.* **2015**, *4*, 1–104. [[CrossRef](#)]
29. Gimenez, L.; Hippolyte, J.L.; Robert, S.; Suard, F.; Zreik, K. reconstruction of 3D building information models from 2D scanned plans. *J. Build. Eng.* **2015**, *2*, 24–35. [[CrossRef](#)]

30. Musialski, P.; Wonka, P.; Aliaga, D.G.; Wimmer, M.; Van Gool, L.; Purgathofer, W. *A Survey of Urban Reconstruction*; Wiley Online Library: Hoboken, NJ, USA, 2013; Volume 32, pp. 146–177. [[CrossRef](#)]
31. Haala, N.; Kada, M. An update on automatic 3D building reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 570–580. [[CrossRef](#)]
32. Neuhausen, M.; Koch, C.; König, M. Image-based Window Detection—An Overview. In Proceedings of the 23rd International Workshop of the European Group for Intelligent Computing in Engineering, Kraków, Poland, 29 June–1 July 2016.
33. Lee, D.; Pietrzyk, P.; Donkers, S.; Liem, V.; van Oostveen, J. Modelling and observation of heat losses from buildings: The impact of geometric detail on 3D heat flux modelling. In Proceedings of the 33rd European Association of Remote Sensing Laboratories (EARSeL) Symposium, Matera, Italy, 3–6 June 2013.
34. Deng, Y.; Cheng, J.C.; Anumba, C. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Autom. Constr.* **2016**, *67*, 1–21. [[CrossRef](#)]
35. Pătrăucean, V.; Armeni, I.; Nahangi, M.; Yeung, J.; Brilakis, I.; Haas, C. State of research in automatic as-built modelling. *Adv. Eng. Infor.* **2015**, *29*, 162–171. [[CrossRef](#)]
36. Wang, J.; Sun, W.; Shou, W.; Wang, X.; Wu, C.; Chong, H.Y.; Liu, Y.; Sun, C. Integrating BIM and LiDAR for Real-Time Construction Quality Control. *J. Intell. Robot. Syst. Theory Appl.* **2015**, *79*, 417–432. [[CrossRef](#)]
37. Tang, P.; Huber, D.; Akinci, B.; Lipman, R.; Lytle, A. Automatic Reconstruction of As-Built Building Information Models from Laser-Scanned Point Clouds: A Review of Related Techniques. *Autom. Constr.* **2010**, *19*, 829–843. [[CrossRef](#)]
38. Freitas, S.; Catita, C.; Redweik, P.; Brito, M.C. Modelling solar potential in the urban environment: State-of-the-art review. *Renew. Sustain. Energy Rev.* **2015**, *41*, 915–931. [[CrossRef](#)]
39. Desthieux, G.; Carneiro, C.; Camponovo, R.; Ineichen, P.; Morello, E.; Boulmier, A.; Abdennadher, N.; Dervev, S.; Ellert, C. Solar energy potential assessment on rooftops and facades in large built environments based on lidar data, image processing, and cloud computing. Methodological background, application, and validation in Geneva (solar cadaster). *Front. Built Environ.* **2018**, *4*, 1–22. [[CrossRef](#)]
40. Catita, C.; Redweik, P.; Pereira, J.; Brito, M.C. Extending solar potential analysis in buildings to vertical facades. *Comput. Geosci.* **2014**, *66*, 1–12. [[CrossRef](#)]
41. Gasparella, A.; Pernigotto, G.; Cappelletti, F.; Romagnoni, P.; Baggio, P. Analysis and modelling of window and glazing systems energy performance for a well insulated residential building. *Energy Build.* **2011**, *43*, 1030–1037. [[CrossRef](#)]
42. Perera, D.W.U.; Pfeiffer, C.F.; Skeie, N.O. Modelling the Heat Dynamics of a Residential Building Unit: Application to Norwegian Buildings. *MIC J. Nor.* **2014**, *35*, 43–57. [[CrossRef](#)]
43. Schwanebeck, M.; Krüger, M.; Duttmann, R. Improving GIS-Based Heat Demand Modelling and Mapping for Residential Buildings with Census Data Sets at Regional and Sub-Regional Scales. *Energies* **2021**, *14*, 1029. [[CrossRef](#)]
44. Nageler, P.; Zahrer, G.; Heimrath, R.; Mach, T.; Mauthner, F.; Leusbrock, I.; Schranzhofer, H.; Hochenauer, C. Novel validated method for GIS based automated dynamic urban building energy simulations. *Energy* **2017**, *139*, 142–154. [[CrossRef](#)]
45. Krüger, A.; Kolbe, T.H. Building analysis for urban energy planning using key indicators on virtual 3d city models—The energy atlas of Berlin. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *XXXIX-B2*, 145–150. [[CrossRef](#)]
46. Previtali, M.; Barazzetti, L.; Brumana, R.; Cuca, B.; Oreni, D.; Roncoroni, F.; Scaioni, M. Automatic façade modelling using point cloud data for energy-efficient retrofitting. *Appl. Geomat.* **2014**, *6*, 95–113. [[CrossRef](#)]
47. Schlueter, A.; Thesseling, F. Building information model based energy/exergy performance assessment in early design stages. *Autom. Constr.* **2009**, *18*, 153–163. [[CrossRef](#)]
48. Saran, S.; Wate, P.; Srivastav, S.K.; Krishna Murthy, Y.V. CityGML at semantic level for urban energy conservation strategies. *Ann. GIS* **2015**, *21*, 27–41. [[CrossRef](#)]
49. Obel, M.; Mark, P.; Neuhausen, M.; König, M.; Schindler, S. Settlement risk assessment by means of categorized surface infrastructure. Structural analysis of historical constructions: Anamnesis, diagnosis, therapy, Controls. In Proceedings of the Tenth International Conference on Structural Analysis of Historical Constructions, Leuven, Belgium, 13–15 September 2016. [[CrossRef](#)]
50. Neuhausen, M.; König, M. Automatic window detection in facade images. *Autom. Constr.* **2018**, *96*, 527–539. j.autcon.2018.10.007 [[CrossRef](#)]
51. Zhang, X.; Lippoldt, F.; Chen, K.; Johan, H.; Erdt, M.; Zhang, X.; Lippoldt, F.; Chen, K.; Johan, H.; Erdt, M. A Data-driven Approach for Adding Facade Details to Textured LoD2 CityGML Models. In Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Prague, Czech Republic, 25–27 June 2019; pp. 294–301.
52. Duarte, D.; Nex, F.; Kerle, N.; Vosselman, G. Damage detection on building façades using multi-temporal aerial oblique imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *IV-2/W5*, 29–36.
53. Kelly, T.; Guerrero, P.; Steed, A.; Wonka, P.; Mitra, N.J. FrankenGAN: guided detail synthesis for building mass-models using style-synchronized GANs. *arXiv* **2018**, arXiv:1806.07179.
54. Klem, D., Jr. Bird–window collisions: A critical animal welfare and conservation issue. *J. Appl. Anim. Welf. Sci.* **2015**, *18*, S11–S17. [[CrossRef](#)] [[PubMed](#)]
55. Klem, D., Jr.; Farmer, C.J.; Delacretaz, N.; Gelb, Y.; Saenger, P.G. Architectural and landscape risk factors associated with bird–glass collisions in an urban environment. *Wilson J. Ornithol.* **2009**, *121*, 126–134. [[CrossRef](#)]



56. Klem, D.; Keck, D.C.; Marty, K.L.; Ball, A.J.M.; Niciu, E.E.; Platt, C.T. Effects of window angling, feeder placement, and scavengers on avian mortality at plate glass. *Wilson Bull.* **2004**, *116*, 69–73. [[CrossRef](#)]
57. Wang, R.; Peethambaran, J.; Chen, D. LiDAR Point Clouds to 3-D Urban Models: A Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 606–627. [[CrossRef](#)]
58. Lotte, R.G. 3-Dimensional (3D) Urban Mapping: A Study of Detection and Reconstruction of Building's Facade Through Structure-From-Motion (SfM) and Convolutional Neural Network (CNN). Ph.D. Thesis, Instituto Nacional De Pesquisas Espaciais, State of Sao Paulo, Brazil, 2018.
59. Frommholz, D.; Linkiewicz, M.; Poznanska, A.M. Inlining 3d reconstruction, multi-source texture mapping and semantic analysis using oblique aerial imagery. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2016**, *XLI-B3*, 605–612. [[CrossRef](#)]
60. Toschi, I.; Ramos, M.M.; Nocerino, E.; Menna, F.; Remondino, F.; Moe, K.; Poli, D.; Legat, K.; Fassi, F. Oblique photogrammetry supporting 3d urban reconstruction of complex scenarios. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.-ISPRS Arch.* **2017**, *42*, 519–526. [[CrossRef](#)]
61. Nex, F.; Rupnik, E.; Remondino, F. Building Footprints Extraction from Oblique Imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*; **2013**, *2*, 61–66. [[CrossRef](#)]
62. Rakha, T.; Gorodetsky, A. Review of Unmanned Aerial System (UAS) Applications in the Built Environment: Towards Automated Building Inspection Procedures Using Drones. *Autom. Constr.* **2018**, *93*, 252–264. [[CrossRef](#)]
63. Lin, Y.; Hyyppä, J.; Kukko, A. Stop-and-go mode: Sensor manipulation as essential as sensor development in terrestrial laser scanning. *Sensors* **2013**, *13*, 8140–8154. [[CrossRef](#)]
64. Novak, K. Mobile mapping systems: new tools for the fast collection of GIS information. In *State-of-the-Art Mapping. International Society for Optics and Photonics*; SPIE: Bellingham WA USA, 1993; Volume 1943, pp. 188–198.
65. Newman, P.; Cole, D.; Ho, K. Outdoor SLAM using visual appearance and laser ranging. In Proceedings of the IEEE International Conference on Robotics and Automation, Orlando, FL, USA, 15–19 May 2006; Volume 2006. [[CrossRef](#)]
66. Huber, D. Background Positioning for Mobile devices-Android vs. iphone. In Proceedings of the Joint Conference of IEEE Computer & Communication Societies, Shanghai, China 10–15 April 2011; Volume 7.
67. Poiesi, F.; Locher, A.; Chippendale, P.; Nocerino, E.; Remondino, F.; Van Gool, L. Cloud-based collaborative 3D reconstruction using smartphones. In Proceedings of the 14th European Conference on Visual Media Production (CVMP 2017), London, UK, 11–13 December 2017; pp. 1–9.
68. Zhang, S.; Shan, J.; Zhang, Z.; Yan, J.; Hou, Y. Integrating smartphone images and airborne lidar data for complete urban building modelling. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 741–747. [[CrossRef](#)]
69. Jeong, J.; Kim, T. Semi-automatic building models and façade texture mapping from mobile phone images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 613. [[CrossRef](#)]
70. Westoby, M.J.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, J.M. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314. [[CrossRef](#)]
71. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [[CrossRef](#)]
72. Elwood, S. Volunteered geographic information: future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal* **2008**, *72*, 173–183. [[CrossRef](#)]
73. Fan, H.; Wang, Y.; Gong, J. Layout graph model for semantic façade reconstruction using laser point clouds. *Geo-Spat. Inf. Sci.* **2021**, *24*, 403–421. [[CrossRef](#)]
74. Daniel, F.; Kucherbaev, P.; Cappiello, C.; Benatallah, B.; Allahbakhsh, M. Quality control in crowdsourcing: A survey of quality attributes, assessment techniques, and assurance actions. *ACM Comput. Surv. (CSUR)* **2018**, *51*, 1–40. [[CrossRef](#)]
75. Verykokou, S.; Ioannidis, C. *Oblique Aerial Images: A Review Focusing on Georeferencing Procedures*; International Journal of Remote Sensing, Taylor & Francis: London, UK, 2018. [[CrossRef](#)]
76. Remondino, F.; Toschi, I.; Gerke, M.; Nex, F.; Holland, D.; McGill, A.; Lopez, J.T.; Magarinos, A. Oblique aerial imagery for rma—some best practices. *Off. Publ.-EuroSDR* **2017**, *2017*, *XLI-B4*, 639–645. [[CrossRef](#)]
77. Rupnik, E.; Nex, F.; Remondino, F. Oblique multi-camera systems-orientation and dense matching issues. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives*; ISPRS: Vienna, Austria, 2014; Volume 40. doi: [[CrossRef](#)]
78. McManamon, P.F. *LiDAR Technologies and Systems*; SPIE Press Book: Bellingham, WA, USA, 2019. [[CrossRef](#)]
79. Mikhail, E.M. *Introduction to Modern Photogrammetry*; Wiley Online Library: Hoboken, NJ, USA; Chichester, UK, 2001.
80. Wolf, P.R.; Dewitt, B.A.; Wilkinson, B.E. *Elements of Photogrammetry with Application in GIS*; McGraw-Hill Education: Berkshire, UK, 1983; Volume 110.
81. Che, E.; Jung, J.; Olsen, M.J. Object Recognition, Segmentation, and Classification of Mobile Laser Scanning Point Clouds: A State of the Art Review. *Sensors* **2019**, *19*, 810. [[CrossRef](#)]
82. Haala, N.; Rothermel, M.; Cavegn, S. Extracting 3D urban models from oblique aerial images. In Proceedings of the 2015 Joint Urban Remote Sensing Event (JURSE), Lausanne, Switzerland, 30 March–1 April 2015; pp. 1–4.
83. Bechtold, S.; Höfle, B. Helios: A multi-purpose lidar simulation framework for research, planning and training of laser scanning operations with airborne, ground-based mobile and stationary platforms. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 161–168. [[CrossRef](#)]

84. Zhu, Q.; Li, Y.; Hu, H.; Wu, B. Robust point cloud classification based on multi-level semantic relationships for urban scenes. *ISPRS J. Photogramm. Remote Sens.* **2017**, *129*, 86–102. [[CrossRef](#)]
85. Ni, H.; Lin, X.; Ning, X.; Zhang, J. Edge detection and feature line tracing in 3D-point clouds by analyzing geometric properties of neighborhoods. *Remote Sens.* **2016**, *8*, 710. [[CrossRef](#)]
86. Kedzierski, M.; Fryskowska, A. Methods of laser scanning point clouds integration in precise 3D building modelling. *Meas. J. Int. Meas. Confed.* **2015**, *74*, 221–232. [[CrossRef](#)]
87. Kersten, T.; Mechelke, K.; Maziull, L. 3D model of Al Zubarah fortress in qatar-Terrestrial laser scanning vs. dense image matching. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 1–8. [[CrossRef](#)]
88. Recky, M.; Leberl, F. Windows detection using K-means in CIE-Lab color space. In Proceedings of the International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010. doi: [[CrossRef](#)]
89. Wang, Q.; Tan, Y.; Mei, Z. Computational methods of acquisition and processing of 3D point cloud data for construction applications. *Arch. Comput. Methods Eng.* **2020**, *27*, 479–499. [[CrossRef](#)]
90. Chen, D.; Zhang, L.; Mathiopoulos, P.T.; Huang, X. A methodology for automated segmentation and reconstruction of urban 3-D buildings from ALS point clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4199–4217. [[CrossRef](#)]
91. Vo, A.V.; Truong-Hong, L.; Laefer, D.F.; Bertolotto, M. Octree-based region growing for point cloud segmentation. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 88–100. [[CrossRef](#)]
92. Xiao, J.; Adler, B.; Zhang, J.; Zhang, H. Planar segment based three-dimensional point cloud registration in outdoor environments. *J. Field Robot.* **2013**, *30*, 552–582. [[CrossRef](#)]
93. Nurunnabi, A.; Belton, D.; West, G. Robust segmentation in laser scanning 3D point cloud data. In Proceedings of the 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA), Fremantle, WA, Australia, 3–5 December 2012; pp. 1–8.
94. Lu, X.; Yao, J.; Tu, J.; Li, K.; Li, L.; Liu, Y. Pairwise linkage for point cloud segmentation. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 201–208. [[CrossRef](#)]
95. Xu, Y.; Yao, W.; Tuttas, S.; Hoegner, L.; Stilla, U. Unsupervised segmentation of point clouds from buildings using hierarchical clustering based on gestalt principles. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 4270–4286. [[CrossRef](#)]
96. Gadde, R.; Jampani, V.; Marlet, R.; Gehler, P.V. Efficient 2D and 3D Facade Segmentation Using Auto-Context. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 1273–1280. [[CrossRef](#)]
97. Roynard, X.; Deschaud, J.E.; Goulette, F. Paris-lille-3d: A point cloud dataset for urban scene segmentation and classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2027–2030.
98. Becker, S.; Haala, N. Integrated LIDAR and image processing for the modelling of building facades. *Photogramm. Fernerkund. Geoinf.* **2008**, *2*, 65–81.
99. Mesolongitis, A.; Stamos, I. Detection of windows in point clouds of urban scenes. In Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, RI, USA, 16–21 June 2012; pp. 17–24. [[CrossRef](#)]
100. Zolanvari, S.M.; Laefer, D.F.; Natanzi, A.S. Three-dimensional building façade segmentation and opening area detection from point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *143*, 134–149. [[CrossRef](#)]
101. Pu, S.; Vosselman, G. Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 575–584. [[CrossRef](#)]
102. Iman Zolanvari, S.M.; Laefer, D.F. Slicing Method for curved façade and window extraction from point clouds. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 334–346. [[CrossRef](#)]
103. Truong-Hong, L.; Laefer, D.F.; Hinks, T.; Carr, H. Combining an angle criterion with voxelization and the flying voxel method in reconstructing building models from LiDAR data. *Comput.-Aided Civil Infrastruct. Eng.* **2013**, *28*, 112–129. [[CrossRef](#)]
104. Zhou, M.; Ma, L.; Li, Y.; Li, J. Extraction of building windows from mobile laser scanning point clouds. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018. [[CrossRef](#)]
105. Pu, S.; Vosselman, G. Extracting windows from terrestrial laser scanning. *Intl Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2007**, *36*, 12–14.
106. Li, J.; Xiong, B.; Qin, R.; Gruen, A. A flexible inference machine for global alignment of wall openings. *Remote Sens.* **2020**, *12*, 1968. [[CrossRef](#)]
107. Li, J.; Xiong, B.; Biljecki, F.; Schrotter, G. A sliding window method for detecting corners of openings from terrestrial LiDAR data. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives*; ISPRS: Vienna, Austria, 2018; Volume 42. [[CrossRef](#)]
108. Dehbi, Y.; Hadiji, F.; Gröger, G.; Kersting, K.; Plümer, L. Statistical relational learning of grammar rules for 3D building reconstruction. *Trans. GIS* **2017**, *21*, 134–150. [[CrossRef](#)]
109. Aijazi, A.K.; Checchin, P.; Trassoudaine, L. Automatic detection and feature estimation of windows for refining building facades in 3D urban point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *2*, 1. [[CrossRef](#)]
110. Friedman, S.; Stamos, I. Online detection of repeated structures in point clouds of urban scenes for compression and registration. *Int. J. Comput. Vis.* **2013**, *102*, 112–128. [[CrossRef](#)]
111. Koffka, K. *Principles of Gestalt Psychology*; Routledge: Oxfordshire, UK, 2013.

112. Flemming, U. *Syntactic Structures in Architecture: Teaching Composition with Computer Assistance*; CUMINCAD: 1990.
113. Truong-Hong, L.; Laefer, D.F. Octree-based, automatic building facade generation from LiDAR data. *Comput.-Aided Des.* **2014**, *53*, 46–61. [[CrossRef](#)]
114. Yang, J.; Shi, Z.K.; Wu, Z.Y. Towards automatic generation of as-built BIM: 3D building facade modeling and material recognition from images. *Int. J. Autom. Comput.* **2016**, *13*, 338–349. [[CrossRef](#)]
115. Nguatam, W.; Drauschke, M.; Mayer, H. Localization of Windows and Doors in 3d Point Clouds of Facades. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *2*, 87. [[CrossRef](#)]
116. Maddern, W.; Pascoe, G.; Linegar, C.; Newman, P. 1 Year, 1000km: The Oxford RobotCar Dataset. *Int. J. Robot. Res.* **2017**, *36*, 3–15. [[CrossRef](#)]
117. Becker, S.; Haala, N. Combined feature extraction for façade reconstruction. In Proceedings of the ISPRS Workshop Laser Scanning, Espoo, Finland, 12–14 September 2007; pp. 241–247.
118. Pu, S.; Vosselman, G. Building facade reconstruction by fusing terrestrial laser points and images. *Sensors* **2009**, *9*, 4525–4542. [[CrossRef](#)]
119. Rahmani, K.; Mayer, H. High quality facade segmentation based on structured random forest, region proposal network and rectangular fitting. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; ISPRS: Vienna, Austria, 2018; Volume 4. [[CrossRef](#)]
120. Yan, L.; Fan, B.; Liu, H.; Huo, C.; Xiang, S.; Pan, C. Triplet Adversarial Domain Adaptation for Pixel-Level Classification of VHR Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3558–3573. [[CrossRef](#)]
121. Liu, H.; Xu, Y.; Zhang, J.; Zhu, J.; Li, Y.; Hoi, S.C. DeepFacade: A Deep Learning Approach to Facade Parsing with Symmetric Loss. *IEEE Trans. Multimed.* **2020**, *22*, 3153–3165. [[CrossRef](#)]
122. Fathalla, R.; Vogiatzis, G. A deep learning pipeline for semantic facade segmentation. In Proceedings of the British Machine Vision Conference 2017, London, UK, 4–7 September 2017. [[CrossRef](#)]
123. Mathias, M.; Martinović, A.; Van Gool, L. ATLAS: A Three-Layered Approach to Facade Parsing. *Int. J. Comput. Vis.* **2016**, *118*, 22–48. [[CrossRef](#)]
124. Gadde, R.; Marlet, R.; Paragios, N. Learning Grammars for Architecture-Specific Facade Parsing. *Int. J. Comput. Vis.* **2016**, *117*, 290–316. [[CrossRef](#)]
125. Teboul, O.; Kokkinos, I.; Simon, L.; Koutsourakis, P.; Paragios, N. Parsing facades with shape grammars and reinforcement learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1744–1756. [[CrossRef](#)]
126. Riemenschneider, H.; Krispel, U.; Thaller, W.; Donoser, M.; Havemann, S.; Fellner, D.; Bischof, H. Irregular lattices for complex shape grammar facade parsing. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012. [[CrossRef](#)]
127. Martinović, A.; Mathias, M.; Weissenberg, J.; Gool, L.V. A three-layered approach to facade parsing. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 416–429.
128. Cohen, A.; Schwing, A.G.; Pollefeys, M. Efficient structured parsing of facades using dynamic programming. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 3206–3213.
129. Ma, W.; Ma, W.; Xu, S.; Zha, H. Pyramid ALKNet for Semantic Parsing of Building Facade Image. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1009–1013. [[CrossRef](#)]
130. Lotte, R.G.; Haala, N.; Karpina, M.; de Aragão, L.E.O.e.; Shimabukuro, Y.E. 3D façade labeling over complex scenarios: A case study using convolutional neural network and structure-from-motion. *Remote Sens.* **2018**, *10*, 1435. [[CrossRef](#)]
131. Kong, G.; Fan, H. Enhanced facade parsing for street-level images using convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 10519–10531. [[CrossRef](#)]
132. Martinovic, A.; Van Gool, L. Bayesian grammar learning for inverse procedural modeling. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; Columbus, OH, USA, 25–27 June 2013. doi: [[CrossRef](#)]
133. Schmitz, M.; Mayer, H. A convolutional network for semantic facade segmentation and interpretation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 709. [[CrossRef](#)]
134. Ma, W.; Xu, S.; Ma, W.; Zha, H. Multiview Feature Aggregation for Facade Parsing. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 1–5. [[CrossRef](#)]
135. Nishida, G.; Bousseau, A.; Aliaga, D.G. Procedural modeling of a building from a single image. In *Computer Graphics Forum*; Wiley Online Library: Hoboken, NJ, USA, 2018; Volume 37, pp. 415–429.
136. Wu, F.; Yan, D.M.; Dong, W.; Zhang, X.; Wonka, P. Inverse procedural modeling of facade layouts. *arXiv* **2013**, arXiv:1308.0419.
137. Teboul, O.; Simon, L.; Koutsourakis, P.; Paragios, N. Segmentation of building facades using procedural shape priors. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010. [[CrossRef](#)]
138. Meixner, P.; Leberl, F.; Brédif, M. Interpretation of 2d and 3d building details on facades and roofs. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2011**, XXXVIII-3/W22, 137–142. [[CrossRef](#)]
139. Neuhausen, M.; Martin, A.; Obel, M.; Mark, P.; König, M. A cascaded classifier approach to window detection in facade images. In Proceedings of the 34th International Symposium on Automation and Robotics in Construction, Taipei, Taiwan, 28 June–1 July 2017. [[CrossRef](#)]

140. Lee, S.C.; Nevatia, R. Extraction and integration of window in a 3D building model from ground view images. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; Volume 2. [\[CrossRef\]](#)
141. Hu, X.; Noskov, A.; Fan, H.; Novack, T.; Li, H.; Gu, F.; Shang, J.; Zipf, A. Tagging the main entrances of public buildings based on OpenStreetMap and binary imbalanced learning. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 1773–1801. [\[CrossRef\]](#)
142. Tian, Y.; Yang, X.; Yi, C.; Arditi, A. Toward a computer vision-based wayfinding aid for blind persons to access unfamiliar indoor environments. *Mach. Vis. Appl.* **2013**, *24*, 521–535. [\[CrossRef\]](#)
143. Díaz-Vilariño, L.; Verbree, E.; Zlatanova, S.; Diakité, A. Indoor modelling from SLAM-based laser scanner: Door detection to envelope reconstruction. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives*; ISPRS: Vienna, Austria, 2017; Volume 42. doi: [\[CrossRef\]](#)
144. Liu, Y.; Liu, J.; Korah, T.; Hedau, V.; Parameswaran, V.; Grzeszczuk, R. Entrance Detection from Street-View Images. In Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 24–27 June 2014.
145. Kang, J.; Deng, F.; Li, X.; Wan, F. Automatic texture reconstruction of 3d city model from oblique images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 341–347. [\[CrossRef\]](#)
146. Talebi, M.; Vafaei, A.; Monadjemi, A. Vision-based entrance detection in outdoor scenes. *Multimed. Tools Appl.* **2018**, *77*. [\[CrossRef\]](#)
147. Kang, S.J.; Trinh, H.H.; Kim, D.N.; Jo, K.H. Entrance detection of buildings using multiple cues. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2010; Volume 5990 LNAI.
148. Yang, L.; Sheng, Y.; Wang, B. 3D reconstruction of building facade with fused data of terrestrial LiDAR data and optical image. *Optik* **2016**, *127*, 2165–2168. [\[CrossRef\]](#)
149. Lin, D.; Jarzabek-Rychard, M.; Tong, X.; Maas, H.G. Fusion of thermal imagery with point clouds for building façade thermal attribute mapping. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 162–175. [\[CrossRef\]](#)
150. Zhu, J.; Xu, Y.; Hoegner, L.; Stilla, U. Direct co-registration of tir images and mls point clouds by corresponding keypoints. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; ISPRS: Vienna, Austria, 2019; Volume 4. [\[CrossRef\]](#)
151. Fryskowska, A.; Stachelek, J. A no-reference method of geometric content quality analysis of 3D models generated from laser scanning point clouds for hBIM. *J. Cult. Herit.* **2018**, *34*, 95–108. [\[CrossRef\]](#)
152. Costanzo, A.; Minasi, M.; Casula, G.; Musacchio, M.; Buongiorno, M.F. Combined use of terrestrial laser scanning and IR Thermography applied to a historical building. *Sensors* **2014**, *15*, 194–213. [\[CrossRef\]](#)
153. Jarzabek-Rychard, M.; Lin, D.; Maas, H.G. Supervised detection of façade openings in 3D point clouds with thermal attributes. *Remote Sens.* **2020**, *12*, 543. [\[CrossRef\]](#)
154. Lin, D.; Dong, Z.; Zhang, X.; Maas, H.G. Unsupervised window extraction from photogrammetric point clouds with thermal attributes. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *4*, 45–51. [\[CrossRef\]](#)
155. Wang, Y.; Ma, Y.; Zhu, A.x.; Zhao, H.; Liao, L. Accurate facade feature extraction method for buildings from three-dimensional point cloud data considering structural information. *ISPRS J. Photogramm. Remote Sens.* **2018**, *139*, 146–153. [\[CrossRef\]](#)
156. Omar, H.; Mahdjoubi, L.; Kheder, G. Towards an automated photogrammetry-based approach for monitoring and controlling construction site activities. *Comput. Ind.* **2018**, *98*, 172–182. [\[CrossRef\]](#)
157. Xie, Y.; Tian, J.; Zhu, X.X. Linking Points with Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geosci. Remote. Sens. Mag.* **2020**, *8*, 38–59. [\[CrossRef\]](#)
158. Wysocki, O.; Hoegner, L.; Stilla, U. TUM-FAÇADE: Reviewing and Enriching Point Cloud Benchmarks for FAÇADE Segmentation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *46*, 529–536. [\[CrossRef\]](#)
159. Matrone, F.; Lingua, A.; Pierdicca, R.; Malinverni, E.; Paolanti, M.; Grilli, E.; Remondino, F.; Murtiyoso, A.; Landes, T. A benchmark for large-scale heritage point cloud semantic segmentation. *XXIV ISPRS Congr.* **2020**, *43*, 1419–1426. [\[CrossRef\]](#)
160. Raji, I.D.; Bender, E.M.; Paullada, A.; Denton, E.; Hanna, A. AI and the everything in the whole wide world benchmark. *arXiv* **2021**, arXiv:2111.15366.
161. Koch, B.; Denton, E.; Hanna, A.; Foster, J.G. Reduced, Reused and Recycled: The Life of a Dataset in Machine Learning Research. *arXiv* **2021**, arXiv:2112.01716.
162. Nowogrodzki, A. Eleven tips for working with large data sets. *Nature* **2020**, *577*, 439–441. [\[CrossRef\]](#) [\[PubMed\]](#)
163. Liu, K.; Boehm, J. Classification of big point cloud data using cloud computing. *ISPRS-Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **2015**, *40*, 553–557. [\[CrossRef\]](#)
164. Boehm, J.; Liu, K. NoSQL for storage and retrieval of large LiDAR data collections. *ISPRS-Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **2015**, *40*, 577–582. [\[CrossRef\]](#)
165. Geyer, J.; Kassahun, Y.; Mahmudi, M.; Ricou, X.; Durgesh, R.; Chung, A.S.; Hauswald, L.; Pham, V.H.; Mühlegg, M.; Dorn, S.; et al. A2D2: Audi Autonomous Driving Dataset *arXiv* **2020**, arXiv:2004.06320.
166. Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets Robotics: The KITTI Dataset. *Int. J. Robot. Res.* **2013**, *32*, 1231–1237. [\[CrossRef\]](#)
167. Huang, X.; Wang, P.; Cheng, X.; Zhou, D.; Geng, Q.; Yang, R. The ApolloScape Open Dataset for Autonomous Driving and Its Application. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*. [\[CrossRef\]](#)

168. Neuhold, G.; Ollmann, T.; Bulo, S.R.; Kotschieder, P. The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017. [[CrossRef](#)]
169. Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The Cityscapes Dataset for Semantic Urban Scene Understanding. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016. [[CrossRef](#)]
170. Hackel, T.; Savinov, N.; Ladicky, L.; Wegner, J.D.; Schindler, K.; Pollefeys, M. SEMANTIC3D.NET: A new large-scale point cloud classification benchmark. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-1-W1*, 91–98.
171. Varma, G.; Subramanian, A.; Namboodiri, A.; Chandraker, M.; Jawahar, C. IDD: A dataset for exploring problems of autonomous navigation in unconstrained environments. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 7–11 January 2019; pp. 1743–1751.
172. Korč, F.; Förstner, W. *eTRIMS Image Database for Interpreting Images of Man-Made Scenes*; Tech. Rep. TR-IGG-P-2009-01t; University of Bonn: Bonn, Germany, 2009.
173. Riemenschneider, H.; Bódis-Szomorú, A.; Weissenberg, J.; Van Gool, L. Learning where to classify in multi-view semantic segmentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 8436. [[CrossRef](#)]
174. Tyleček, R.; Šára, R. Spatial pattern templates for recognition of objects with regular structure. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8142, pp. 364–374.
175. Yu, F.; Chen, H.; Wang, X.; Xian, W.; Chen, Y.; Liu, F.; Madhavan, V.; Darrell, T. Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2636–2645.