

# A Review on Viewpoints and Path Planning for UAV-Based 3-D Reconstruction

Mehdi Maboudi , MohammadReza Homaei , Soohwan Song , Shirin Malihi , Mohammad Saadatseresht , and Markus Gerke 

**Abstract**—Unmanned aerial vehicles (UAVs) are widely used platforms to carry data capturing sensors for various applications. The reason for this success can be found in many aspects: the high maneuverability of the UAVs, the capability of performing autonomous data acquisition, flying at different heights, and the possibility to reach almost any vantage point. The selection of appropriate viewpoints and planning the optimum trajectories of UAVs is an emerging topic that aims at increasing the automation, efficiency, and reliability of the data capturing process to achieve a dataset with desired quality. On the other hand, 3-D reconstruction using the data captured by UAVs is also attracting attention in research and industry. This article investigates a wide range of model-free and model-based algorithms for viewpoints and path planning for 3-D reconstruction of large-scale objects. It presents a bibliography of more than 200 references to cover different aspects of the topic. The analyzed approaches are limited to those that employ a single-UAV as a data capturing platform for outdoor 3-D reconstruction purposes. In addition to discussing the evaluation strategies, this article also highlights the innovations and limitations of the investigated approaches. It concludes with a critical analysis of the existing challenges and future research perspectives.

**Index Terms**—Aircraft navigation path planning, autonomous aerial vehicles, image reconstruction, motion planning, remotely guided vehicles, surface reconstruction, 3-D displays, viewpoints planning.

## I. INTRODUCTION

A COMPLETE dataset which meets the predefined quality measures is the fundamental input for various applications like 3-D reconstruction of large-scale complex objects with a high level of details (for example cultural heritage documentation), inspection of construction sites, and building information modeling (BIM) just to name a few. Many sensors and platforms

Manuscript received 28 November 2022; revised 15 March 2023 and 25 April 2023; accepted 9 May 2023. Date of publication 16 May 2023; date of current version 9 June 2023. This work was supported by the German Research Foundation (DFG)—TRR 277/1 2020 under Project 414265976. (*Corresponding author: Mehdi Maboudi.*)

Mehdi Maboudi and Markus Gerke are with the Institute of Geodesy and Photogrammetry, Technische Universität Braunschweig, 38106 Braunschweig, Germany (e-mail: m\_maboudi@yahoo.com; m.gerke@tu-braunschweig.de).

MohammadReza Homaei and Mohammad Saadatseresht are with the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran (e-mail: mohammadreza-homaei@gmail.com; msaadat@ut.ac.ir).

Soohwan Song is with the Intelligent Robotics Research Division, ETRI, Daejeon 34129, South Korea (e-mail: soohwansong@etri.re.kr).

Shirin Malihi is with the School of Engineering, Institute for Infrastructure and Environment, University of Edinburgh, EH8 9YL Edinburgh, U.K. (e-mail: sh.malihi@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2023.3276427

for capturing high-quality images and point clouds have been introduced in the last decades. Especially, different types of unmanned aerial vehicles (UAVs) are widely used platforms that are agile, capable of performing autonomous data acquisition, can fly at different heights, and can reach almost any vantage point [1], [2]. Hence, UAV-based 3-D reconstruction is still attracting attention as a promising approach for generating semantic and geometric information about man-made and natural objects.

UAVs are utilized in various monitoring and inspection projects [3], [4], [5], [6], [7], [8], [9], [10], [11], [12]. Especially the demand for high-quality 3-D models in many applications such as construction, urban modelling, and emergency rescue is still increasing [13], [14], [15], [16], [17], [18]. Nevertheless, restricted access to the objects and practical considerations like limited flight time, safety protocols, battery constraints, and legal regulations should also be considered. Furthermore, 3-D reconstruction of complex objects with high level of details necessitate a thorough viewpoints and flight path planning. The quality of 3-D reconstructions highly depends on the acquisition process and camera network configuration [19], [20], [21].

Although 3-D reconstruction with impressive quality using state-of-the-art algorithms is attainable, the success of these reconstruction methods depends to a large degree on the input images with predefined specifications [22], [23], [24], [25], [26], [27]. Furthermore, adding more images cannot always deliver better reconstruction, and even diminishing returns at some points are reported in the literature [27], [28], [29]. Moreover, due to the limited battery capacity of the current UAVs and possible restrictions in accessing the area of interest, data capturing should be performed in a limited time [30]. Hence, efficient viewpoints and path planning algorithms are required to capture the data for the successful reconstruction of 3-D objects with various sizes, shapes, and complexities [31], [32], [33], [34].

The selection of appropriate viewpoints and trajectory for 3-D reconstruction has been a research topic for some decades [35], [36], [37], [38], [39], [40], [41]. In robotics, view and path planning is a research focus of many studies on 3-D reconstruction to make the robots explore and reconstruct unknown environments more competently [42], [43]. This task is known as the active vision/reconstruction or view-path-planning problem in robotics [1], [44], [45]. In Photogrammetry, the problem of determining the optimal viewpoints for accurate measurement of 3-D scenes or objects is called camera network design [21], [23], [46], [47] where mainly the geometric aspects of the problem

are investigated in controlled environments for highly accurate measurement of artificial targets using convergent imaging.

### A. Scope of the Paper

The review includes peer-reviewed publications and theses, focusing on a wide range of model-free and model-based algorithms on viewpoints and path planning for 3-D reconstruction of large-scale objects. In addition to discussing the evaluations strategies, this paper also highlights the innovations and limitations of the investigated approaches. It concludes with a critical analysis of the existing challenges and future research perspectives. Since various platforms and sensors are employed for different 3-D reconstruction applications, we exclude the studies which are not within the scope of this article.

1) *Platform*: Many researchers in robotics and other communities investigate view and path planning for UGVs (Unmanned Ground Vehicles) and wheeled robots [48], [49], [50], autonomous underwater vehicles [51], robotic arms [52], [53] as well as multi-UAVs [54], [55], [56], [57] for optimal data capturing. Although there are some common issues between path planning for these platforms and UAVs, the assumptions, challenges, trajectories, platforms abilities, and environmental conditions could be very different. For example, in contrast to UGVs, UAVs should consider a 6-D (in practice 5-D) search space for the sensor poses within the reconstruction workflow [2], [30]. Therefore, this article focuses on viewpoints and path planning for a single UAV.

2) *Sensor and 3-D Representation*: Various sensors are installed on UAVs for data capturing in different applications [58]. Monocular RGB camera [22], RGB-D sensor [59], depth cameras [60], [61], structured light scanners [62], multispectral camera [63], stereo cameras and LiDAR [64], [65], [66], [67] are the most commonly used sensors [68]. Compared to single RGB cameras, other sensors entail special hardware and setup, making them very efficient for some applications. However, it usually comes at the cost of more technical requirements, bounds applications and makes them more expensive. Different 3-D representation methods have been developed depending on the type of scanning sensors. Images acquired from a monocular camera are generally processed by structure from motion (SfM) and multiview stereo (MVS) algorithms [69] to reconstruct 3-D models. After computing the camera intrinsics and extrinsics using SfM, MVS can generate a dense 3-D reconstruction offline by matching the stereo correspondences of all images. This method can estimate a wide range of depths based on various baseline distances of images; therefore, MVS is appropriate for large-scale modeling. RGB-D sensors and stereo cameras can estimate accurate and dense depth maps in real-time. The estimated depth maps can be integrated into a volumetric model or a surface model by the volumetric mapping [70] or truncated signed distance field (TSDF) mapping [71], respectively. Volumetric models represent a coarse 3-D shape based on occupancy probability, whereas surface models represent the precise surfaces of an object. LiDAR can acquire very accurate point cloud data in real-time. The point clouds are relatively sparse compared to the depth maps acquired from depth sensors

and cannot represent dense surfaces. Therefore, LiDAR-based methods [66], [72] generally employ the volumetric mapping method for scanning.

3) *Application*: Geometric criteria like completeness and accuracy of the 3-D models acquired from UAV missions are the most interesting properties discussed in the literature [45]. However, researchers also aimed at other properties like semantic information, which could be extracted from the UAV data [73], [74], [75]. Apart from many other applications like precision agriculture [76], [77], [78], indoor exploration and reconstruction [79], [80], [81], [82], [83] to name a few, we focus on viewpoints and path planning for 3-D reconstruction of complex outdoor man-made objects like buildings and bridges which is a topic of interest for many applications [84], [85], [86]. However, the studies that are implemented for other applications, but are also applicable to large-scale scenes may be mentioned. A closely related topic is coverage path planning [87], [88], [89], [90], [91], [92] which deals with optimizing the trajectories that cover a known 2-D or 3-D environment. This topic could be considered as a set-cover problem [93], [94] or an art gallery problem followed by the shortest path optimization [95], [96], [97]. However, since these approaches do not consider MVS heuristics and 3-D reconstruction requirements, they are out of the scope of this article.

### B. SfM and MVS Heuristics in Viewpoints Planning

Since many of the 3-D reconstruction approaches rely on SfM and MVS, we shortly list the most important heuristics as follows [see also Fig. 1(a)].

- 1) *Distance (D)*: The distance between camera viewpoints and object surface defines the resulting model resolution and depends on the desired ground sampling distance and the camera intrinsics [19], [36].
- 2) *Multiple views*: Every part of the scene has to be observed from multiple views (theoretically at least two views and practically more for increasing the reliability and accuracy of 3-D reconstruction) from different perspectives with sufficient overlap between the views. Matching the corresponding points in overlapping images is the prerequisite for robust camera poses estimation and for triangulating 3-D object points [23], [24], [45], [98].
- 3) *Observation angle ( $\beta$ ) and parallax angle ( $\theta$ )*: Narrow observation angles between the optical-axis of the cameras and surface normals are preferred in MVS. Moreover, large parallax angles (a.k.a angular separation) between cameras (also known as the B/H ratio in photogrammetry) are favored to increase the triangulation quality. However, it makes it more difficult to find correspondences (matching) between the images, especially for complex 3-D structures [30], [85], [98], [99].

Various models are used in the literature to control these parameters [19], [31], [36]. Three examples are illustrated in Fig. 1(b)–(d).

The parameters defined above are visualized in Fig. 1(a). The strategy which is used in [99] is illustrated in Fig. 1(b), where a hemisphere is considered around the sample point. While

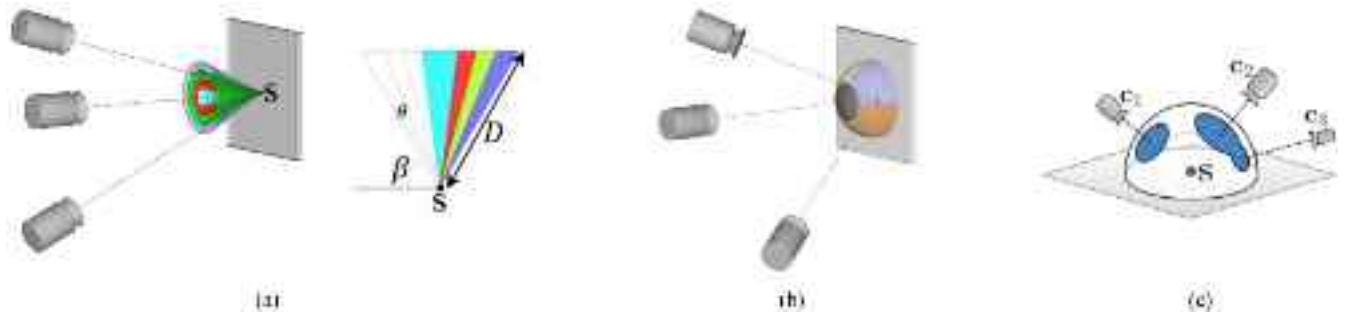


Fig. 1. Different strategies for considering SfM and MVS heuristics in viewpoints planning where S is a point on the surface of the object of interest. (a) Left: Co-apex cones for forcing the camera poses to adhere to the MVS and SfM heuristics. (a) Right: Side-view of four co-apex cones for describing the parameters: Distance ( $D$ ) of a camera to the object's point, Observation angle ( $\beta$ ) of a camera and parallax angle ( $\theta$ ) of two overlapping cameras. (b) Considering the segments of a hemisphere for viewpoints planning adopted from [99]. (c) MVS-aware coverage model in [31].

the hemisphere is divided into four segments, the ideal camera constellations intersect the hemisphere in different segments. The coverage model introduced in [31] is depicted in Fig. 1(c). Since each camera covers a disk area on the hemisphere around the sample point, the usefulness of the cameras is measured by total solid angle covered by all the disks.

### C. Problem Formulation

Since different terms, objectives, priorities, and approaches are utilized for viewpoints and path planning in various communities, we formulate the problem again to provide a holistic definition that covers different strategies and nomenclature, consistently.

Determining “good,” “ideal,” or “reasonable” imaging positions is a common research topic in various communities such as robotics, computer graphics, computer vision, and photogrammetry [23]. In UAV-based research, some viewpoints and path planning approaches (e.g., [23], [24]) solve the planning problem in two steps: The first step solves the optimum viewpoints while ignoring the travel budget and path quality and the second step, minimizes the cost of visiting all viewpoints which are computed in the first step. However, some other approaches solve both viewpoints and path planning problems in a unified approach [26], [45]. This is mainly a domain specific concern and depends on the application and requirements/priorities. Giving priority to the reconstruction quality of the object is useful (usually leads to) first solving the camera poses (sensor placement, a.k.a viewpoints selection) and then finding the shortest/fastest trajectory to visit all those points. The constraints are defined on the object and quality has the highest priority. In this case, after computing the vantage viewpoints, travelling salesman problem (TSP) as shortest Hamiltonian path between the viewpoints is a commonly used approach in the literature. On the other hand, giving priority to resources (flight time, trajectory length, or the number of the images) the constraints are defined on the sensor (also constrained by the obstacles [100], [101]). The aim would be to get as much as possible but hardly constrained by the resources. For maximum gain, the classical knapsack problem could be used. However, costs are not fixed in view-path planning (depending on the distance to the previous viewpoint).

Therefore, this is a combination of knapsack and TSP which is called the orienteering problem [102], [103]. The orienteering problem (reward-collecting graph optimization problem) was first introduced by Golden et al. [104]. The objective is maximizing the total gain collected from visited nodes. Because of the limited time budget, all available nodes may not be selected. In orienteering problem marginal rewards are additive, but in MVS the usability of each camera is not additive and depends on some heuristics which could be defined considering the poses of other cameras. Different strategies are employed to tackle this problem [26], [45]. Generally, we formulate the viewpoints and path planning as follows.

Considering a camera  $K$  with fixed intrinsics (at least during the data capturing) mounted on a drone, the objective of viewpoints and path planning is to find a feasible trajectory  $T$  of the drone which meets some constraints  $\Omega$  and passes through optimal camera positions  $P = \{p_1, p_2, \dots, p_N\}$  for taking high-quality pictures at optimal orientations  $\Phi = \{\varphi_1, \varphi_2, \dots, \varphi_N\}$  to achieve a predefined quality  $Q$  as much as possible [2], [105].

The camera locations and orientations  $\mathcal{C} = \{\{p_1, \varphi_1\}, \{p_2, \varphi_2\}, \dots, \{p_N, \varphi_N\}\}$  constitute a 6-D search space. Generally, roll angle could be neglected in practice, resulting in a 5-D search space for camera poses [2], [30]. Some researchers also add some more constraints (e.g., restricting the search space to a sphere around the object [106], some adaptive planes [105], or height-wise data capturing [107]). They further reduce the computations at the cost of having less chance to find the optimal solution. Various heuristics  $\mathcal{H}$  (cf. Section I-B) could be utilized to encourage the pose optimization algorithms to converge to a solution that meets the requirements of each specific application.

There are also various and different constraints  $\Omega$  which could be considered during path planning. Battery life, the shortest path between viewpoints, detecting and avoiding permanent and moving obstacles, and excluding no-go areas in the scene are the main constraints that are investigated. Continuity of the path or path quality is also another important aspect which is investigated in the path-oriented trajectory planning of [108].

Redefined quality  $Q$  of 3-D reconstruction depends on the objective of the project and could be different in specific domains and applications. For example, in some applications wherein the

full coverage of 3-D reconstruction is very important,  $Q$  could be formulated as the completeness of the results and other metrics like the accuracy of the reconstructed object could be relaxed. On the other hand, in some domains in which the accuracy is very important, other constraints have less priority, and accuracy has the highest weight in the planning. We will address these metrics in Section III.

## II. VIEWPOINTS AND PATH PLANNING STRATEGIES

Given that a maneuverable platform (e.g., UAV in our scope) is available, an initial tempting idea could be to capture so many images to achieve the best 3-D reconstruction. Considering the limited endurance of the UAVs and other practical flight-time limitations, diminishing returns and even destructive effect of adding views after some points [27], [28], [29], [108] and high computation cost of processing all images, this method is neither efficient nor applicable in real scenarios. Another idea could be to select/filter a subset of already densely captured images for 3-D reconstruction to decrease the significant computational overheads and avoid the destructive effect of some images on the final product. These techniques require a complete set of already captured images to filter some of them to achieve high quality 3-D reconstruction. Besides data capturing time which still is a problem, the total computational cost of viewpoints selection and processing the remaining images could be close to the processing time of all images. Furthermore, camera locations are densely sampled and the other three parameters of camera pose namely rotations are not optimized. Some other approaches try to improve the output of regular pattern or first planning of viewpoints by estimating the missing parts and low-quality 3-D reconstructed areas and iteratively adding some amending viewpoints to improve the final 3-D product [109], [110]. However, the best solution could be to have a holistic planning approach before starting data capturing which is the topic of a rich body of viewpoints and path planning literature in various research communities.

Automated flight planning methods can be classified either as model-free or model-based [32], [68], [99], [109], [111]. Model-based methods use an initial coarse model (geometric proxy, scene proxy, or simply proxy) for generating optimal viewpoints and path planning. But, model-free methods tackle the problem given that there is no such initial coarse model of the objects and environment and iteratively generate and update the model with new measurements. In the following parts of this section, we briefly discuss off-the-shelf mission planners. Afterward, we delve into the model-free and model-based methods.

### A. Off-the-Shelf Flight Planners

The manual piloting is inevitable in some applications [6]. However, the most common method for UAV-based data capturing in an automated fashion is to use an off-the-shelf flight planner, such as Pix4D capture,<sup>1</sup> Altizure,<sup>2</sup> Litchi,<sup>3</sup> DJI Terra,<sup>4</sup> DJI

flight planner,<sup>5</sup> 3dr robotics,<sup>6</sup> Precisionhawk,<sup>7</sup> PixHawk Ardu mission Planner,<sup>8</sup> DroneDeploy,<sup>9</sup> precision flight, UgCS.<sup>10</sup> The interested reader can refer to papers [112], [113], [114] for more information about the commercial flight planners. Furthermore, some open-source software package are listed in [41], Table 2.

Usually offering limited parameter settings, commercial flight planners generate conservative trajectories (e.g., an orbit, a zigzag, or a lawnmower pattern at a safe flight height) to cover the scene [19], [33], [45], [105]. However, these trajectories are generated regardless of the scene geometry and structure and distribution of the objects. Hence, these tools tend to oversample some regions (e.g., rooftops) while undersampling others (e.g., facades, convex areas, overhangs, and fine details), especially in dense areas and complex objects, and therefore sacrifice reconstruction quality [2], [26], [31], [115]. Furthermore, most of such planning tools do not directly choose viewpoints adhering to all SfM and MVS constraints and do not directly account for complete coverage of the scene. Recently, Agisoft introduced a tool for designing an optimum path-planning for image capturing and creating mission plans based on the rough model [109], [116]. In this tool, a set of camera positions are generated to cover the object's surface with sufficient overlap. However, it cannot promise the completeness of the dataset in real-world situations, and the accuracy of the final point is not explicitly considered in this approach. Although the Agisoft mission planning tool has a power-line detection algorithm, most off-the-shelf flight mission planners may cause an accident with adjacent obstacles in the environment and should be used cautiously and at the safe fixed distance to the object.

### B. Model-Free Methods

Model-free methods do not have any prior information for target structures or scenes. Since there is no prior information, it is challenging to compute an optimal scanning trajectory. Therefore, model-free methods have to find the best scanning trajectory in an online manner from a partially constructed model. This problem is the same as the exploration planning problem, which determines the scanning paths online to explore an unknown and spatially bounded space. In this problem, a scanning platform has to estimate its location and construct 3-D models in real-time. The location can be estimated using the simultaneous localization and mapping (SLAM) method [117], and sparse 3-D models can be constructed using an environmental mapping method [118]. SLAM systems are used to define the poses of the platforms, capture data from the scene, and direct the platform to the defined positions. Still, before directing the platform, the pose of the next station must be determined. In the last two decades, many SLAM approaches used a probabilistic method for reducing the impact of the inaccurate sensor on the map, using millimeter waves for building relative maps [119],

<sup>5</sup>[Online]. Available: <https://www.djiflightplanner.com/>

<sup>6</sup>[Online]. Available: <https://3drobotics.com/>

<sup>7</sup>[Online]. Available: <https://www.precisionhawk.com/>

<sup>8</sup>[Online]. Available: <https://ardupilot.org/planner/index.html>

<sup>9</sup>[Online]. Available: <https://www.dronedeploy.com/product/mobile/>

<sup>10</sup>[Online]. Available: <https://www.ugcs.com/>

<sup>1</sup>[Online]. Available: <https://www.pix4d.com/product/pix4dcapture>

<sup>2</sup>[Online]. Available: <https://www.altizure.com/>

<sup>3</sup>[Online]. Available: <https://flylitchi.com/>

<sup>4</sup>[Online]. Available: <https://www.dji.com/de/dji-terra>

integrating particle filter and extended Kalman filter [120], or introducing square root smoothing and mapping [121]. Another SLAM system is visual SLAM which can produce a fast 3-D reconstruction on-the-fly [122].

Most model-free methods employ the next-best-view (NBV) approach. This approach is a greedy method that finds a local solution from partial information. It consistently determines the best viewpoint that obtains the largest unknown information from a current map. The unknown information is defined according to the purpose of various scanning scenarios. Some methods [123], [124] try to explore a volumetric map within a minimum time thoroughly. They evaluate the unknown volumes to determine an NBV in the volumetric map. Some methods [22], [125] are trying to inspect all surfaces of an unknown structure. These methods find a viewpoint that observes the largest uncovered surfaces. The other methods [59], [126], [127] try to reconstruct precise 3-D models of target structures or environments. These methods analyze the reconstruction quality of 3-D models, such as completeness and accuracy [128]. This section classified the model-free methods into three categories: *frontier-based planning*, *volumetric-based planning*, and *surface-based planning*. The following sections present literature reviews on each category of model-free planning methods summarized in Table I.

**1) Frontier-Based Methods:** Frontier-based planning is one of the most widely used mobile robot exploration methods. The main idea was originally proposed by Yamauchi [129], who defined the frontier as a boundary between an explored and unexplored area in a 2-D occupancy grid map [130]. The occupancy grid map represents a 2-D environment as a set of grid cells composed of three states: free, occupied, and unknown. The frontier cells are estimated by collecting free cells adjacent to unknown cells. The frontier-based method starts at an initial location with an initially scanned map. The method continuously moves toward the nearest accessible and unvisited frontier cell until entire unknown cells are explored. This method generally shows acceptable exploration performances while intuitive and straightforward. Umari and Mukhopadhyay [131] extended the frontier-based method to detect the frontiers in the occupancy grid map efficiently. They found frontiers by expanding multiple rapidly exploring random trees (RRTs) instead of classical image processing methods like edge detection. The RRTs expanded towards unexplored regions based on the Voronoi diagram; they quickly extracted the frontiers with navigation paths.

The frontier-based method has been extended to 3-D exploration tasks of a UAV by utilizing a 3-D volumetric map [70] instead of a 2-D occupancy grid map. Frontier volumes are defined as the free volumes adjacent to the unknown volume in volumetric maps. The octree structures can rapidly access the 6- or 18-adjacent volumes. Most methods iteratively determine the NBVs based on the frontier information. Obwald et al. [132] improved the exploration performance of the general frontier-based method by providing additional guide information on global routes. They utilized a user-provided topological graph representing the topology information of the environment. A global route is computed by finding a TSP solution on the topological graph. A mobile robot sequentially visits every node based on the TSP tour while exploring local regions using the frontier-based method. This method can produce more effective

exploration paths that prevent revisiting a region that has already been passed. Cieslewski et al. [123] extended the frontier-based method to maintain the maximum speed of a mobile robot. They determined an NBV by selecting target frontier volumes inside the current field of view. This method can prevent a mobile robot from slowing down due to heading rotation; therefore, the robot can explore unknown regions at high speed. This method reduces the exploration time while it may increase the trajectory length. Several studies [124], [133] have focused on reducing the computation time of frontier extraction and clustering. Dai et al. [133] proposed an implicit frontier clustering method. They updated the frontiers only in the volumes inside the camera frustum at every depth integration step. Batinovic et al. [124] tried to filter out many frontier points in multi-resolution volumes on Octomap. They also applied the mean-shift clustering for frontier clustering. These approaches significantly reduce the number of frontiers, which reduces the computation time of candidate viewpoint evaluation.

Some others [134], [135] have tried to determine the next-best trajectories instead of the NBVs. They evaluated view sequences and found the most informative view paths. Shade and Newman [134] proposed a 3-D exploration method for stereo-camera-based scanning. They first extracted frontier volumes in a volumetric map and composed 3-D vector fields toward the frontier volumes. They then computed the steepest descent path in the vector fields to get the next-best trajectory for 3-D scanning. Meng et al. [135] proposed a two-stage planning method for 3-D exploration. The method first generates frontier sets by clustering adjacent frontiers and then computes the coverage path to cover all frontier sets. The coverage path can be computed by solving a fixed start open TSP. A mobile robot moves to the first edge of the coverage path and continuously recomputes the coverage path. Since this method considers global coverage instead of a single viewpoint, the number of revisits of the same area can be reduced.

**2) Volumetric-Based Planning:** Volumetric-based planning focuses on reconstructing a complete volumetric model instead of a fast exploration of unknown areas. Unlike frontier-based planning, volumetric-based planning analyzes not only the frontier but also the entire spatial information of the volumetric map. Estimated 3-D information is sequentially integrated into an octree structure [70]. Each volume has one of three states (free, occupied, or unknown volume), updated based on the occupancy probability. The volumetric map is very suitable for view planning since it can easily access entire spatial information by multiresolution representation of Octomap. Furthermore, it can efficiently perform a ray-casting for visibility check that is essential for viewpoint evaluation.

Volumetric-based planning method has also been widely used in object modeling. Vasquez-Gomez et al. [136] proposed a volumetric-based planning method for single object modeling. They evaluated viewpoints by analyzing unknown volumes with overlapped volumes for point cloud registration. They also applied a hierarchical ray-tracing method for fast visibility check. This method starts from the rough-resolution map and applies high-resolution ray-tracing only around occupied volumes. Delmerico et al. [137] proposed several information gain functions for viewpoint evaluation on the volumetric

TABLE I  
CATEGORIZATION AND MAIN FEATURES OF MODEL-FREE VIEW-PATH PLANNING METHODS

Category	Approach	References	Modeling Method	Modeling Target	Description
Frontier-based planning	NBV planning	Yamauchi [130], Umari & Mukhopadhyay [132]	2D occupancy grid mapping	Indoor/outdoor	Continuously move toward the nearest frontier in 2D grid map
		Zhu et al. [83]	2D occupancy grid mapping	Indoor	Deep reinforcement learning-based exploration planning
		Obwald et al. [133], Cieslewski et al. [124]	Volumetric mapping	Indoor/outdoor	Continuously move toward the nearest frontier in 3D space
		Dai et al. [134], Batinovic et al. [125]	Volumetric mapping	Indoor/outdoor	Reduce the computation time of frontier extraction and clustering
	Next-best trajectory	Shade & Newman [135], Meng et al. [136]	Volumetric mapping	Indoor/outdoor	Evaluate view sequences and determine the best trajectory
Volumetric-based planning	NBV planning	Vasquez-Gomez et al. [137], Delmerico et al. [138], Daudelin [139]	Volumetric mapping	Small-scale object	Provide various evaluation metrics for NBV planning on volumetric model
		Hepp et al. [156]	Volumetric mapping	Indoor/outdoor	3D CNN-based viewpoint evaluation for NBV planning
	Next-best trajectory	Charrow et al. [157], Charrow et al. [158], Wang et al. [159]	Volumetric mapping	Indoor/outdoor	Evaluate or optimize view paths by using an information-theoretic measure
	Receding horizon planning	Bircher et al. [140], Papachristos et al. [141], Dharmadhikari et al. [142], Batinovic et al. [143]	Volumetric mapping	Indoor/outdoor	Iteratively computes a next-best trajectory and only execute the first step of the trajectory
	Inspection planning	Heng et al. [97], Song & S. J [144], Song et al. [22]	Volumetric mapping	Indoor/outdoor	Online inspection path planning for complete modeling
Surface-based planning	NBV planning	Chen & Li [148], Wu et al. [52], Lee et al. [149]	Surface reconstruction	Small-scale object	Surface-based NBV planning for modeling a small-scale object
		Hardouin et al. [60], Kompis et al. [128]	TSDF mapping	Indoor/outdoor	Evaluate reconstruction quality for NBV planning
		Peralta et al. [160], Zeng et al. [161]	Point cloud mapping	Single object	Machine-learning-based NBV planning for point-cloud mapping
	Receding horizon planning	Schmid et al. [151]	TSDF mapping	Indoor/outdoor	Iteratively rewire nodes of RRT for coverage planning
	Inspection planning	Song & S. Jo [127]	TSDF mapping	Outdoor	Online inspection planning for dense surface reconstruction
		Song et al. [34], Song et al. [1]	MVS	Outdoor	View path planning via online multi-view stereo reconstruction

model. To quantify the visible information, they just counted unobserved voxels or calculated a weighting sum of voxel entropy. The authors also considered the object's rear side voxels with their entropy. They evaluated the performances of the proposed information gain functions through various modeling scenarios. Daudelin and Campbell [138] proposed a view planning method for modeling an object without any prior information such as size or bounding box. They dynamically extended search spaces for the sampling of viewpoints based on a partial reconstruction. The search spaces are not restricted to known free space. They estimated all reachable configurations of a mobile robot's sensor and generated viewpoint samples from the reachable configurations. Similar to the work [137], the total amount of information gain for each candidate viewpoint is computed.

Several studies also proposed a volumetric-based planning method for exploration tasks. Bircher et al. [139] employed a receding horizon planning strategy for exploration planning. The strategy iteratively computes an optimal exploration path and only executes the first step of the optimal path. This method generates a set of viewpoints by expanding an RRT and selects the best branch that provides the maximum information gain. The information gain is defined as the total volume of visible unknown cells penalized by the distance costs of nodes. The method then determines the NBV to the first node of the branch. This approach can efficiently find exploration paths with a low-computation complexity even in large-size environments. Papachristos et al. [140] extended the receding horizon planning method [139] to additionally consider the localization uncertainty in exploration tasks. The proposed method comprises two planning stages: volumetric exploration planning and uncertainty-optimization planning. Like the receding horizon planning, the exploration planning expands branches of a random tree and determines an NBV from the first node of the most informative branch. To compute an information gain, they consider not only unknown volumes but also the occupancy probability of occupied volumes. The uncertainty-optimization planning then finds an optimal trajectory to the NBV, which provides the minimum localization uncertainty for the SLAM module. Dharmadhikari et al. [141] also extended the receding horizon planning to provide fast and agile paths for a micro aerial vehicle (MAV). They directly generated view configurations and admissible paths based on motion primitives. They then determined a future-safe path that provides the maximum information gain and guarantees continuous fast motions. Batinovic et al. [142] proposed an exploration planning method for LiDAR scanning data. They applied a recursive shadow-casting algorithm for fast information gain computation of large-scale point clouds. They also proposed a cuboid-based path evaluation method that estimates the information gain of each RRT edge instead of a node.

Song and Jo [143] applied an inspection approach to a model-free planning method. Given a prior model of a target structure, the inspection approach precomputes a coverage path that provides full visual coverage of the whole surfaces of the prior model. On the other hand, Song and Jo [143] addressed an online inspection planning on an incrementally updated and partially known model. The inspection planning method first

determines an NBV that explores the largest unknown area in a volumetric map. Similar to the work [139], it expands branches of RRT and evaluates the information gain of each branch to determine the NBV. The method then plans a local inspection path to the NBV, providing complete visual coverage of near-frontiers. The local inspection path is continuously replanned according to the updated near-frontiers until a robot reaches the NBV. This method improves the completeness of volumetric modeling because it can thoroughly scan all small unreconstructed regions. Song et al. [22] also extended the online inspection planning to consider global coverage planning. They proposed an online map partitioning method, which decomposes entire space into a set of sectors by clustering free and unknown volumes. The decomposed sectors have a compact and convex shape; therefore, the sectors represent a topological map. The method plans a global coverage path of unexplored sectors and determines an NBV to move toward the next sector. It then plans a local inspection path that fully covers local frontiers. This method significantly reduces total exploration time and path length by reducing the number of revisits of the sectors.

*3) Surface-Based Planning:* Surface-based planning concentrates on reconstructing a precise 3-D surface model, which is represented as surface meshes [144], surfel points [145], or TSDF [146]. Volumetric models are difficult to sufficiently express complex surface information; therefore, the surface-based planning methods analyze the shape and trend of reconstructed surfaces instead of volume information. Most surface-based planning methods have been used to reconstruct small-scale objects. Chen and Li [147] predicted the shape and trend of simple and smooth surfaces from the curvature tendency and determined an NBV based on the predicted surface information. Wu et al. [52] also predicted a tentative surface model from a partial reconstruction by using the Poisson surface reconstruction algorithm [144]. They then evaluated the completeness and smoothness of Poisson iso-surfaces to determine an NBV. Lee et al. [148] detected surface shapes with surface primitives such as planes, cylinders, and spheres and utilized them for NBV determination.

Recently, several studies have attempted to apply the surface-based planning method to large-scale modeling tasks. Hardouin et al. [59] analyzed the reconstructed surfaces to determine an NBV for MAV exploration. They estimated 3-D data using a stereo camera or RGB-D sensor and integrated them into a surface model based on the TSDF [146]. TSDF can reconstruct detailed 3-D surfaces in real-time from a volumetric distance field, which contains sign distances to the closest surface. TSDF is also useful for identifying missing model parts during an online reconstruction [149]. To determine an NBV, Hardouin et al. [59] first detected incomplete surface elements (ISEs) from TSDF and then generated viewpoints along with normal directions of ISEs at a specified distance. They clustered the neighboring viewpoints and selected a cluster covering the most ISEs. They incrementally updated the TSDF model by scanning from an NBV until no ISE was detected.

Song and Jo [126] proposed a surface-based exploration, which is an extended method of online inspection planning [143]

for surface model reconstruction. The surface-based exploration analyzes both the volumetric map and surface model in TSDF for fast exploration and complete reconstruction. The method determined an NBV that explores the largest unknown areas in the volumetric map. The method then plans an inspection path covering the surrounding low-confidence surfaces and frontiers. To cover the low-confidence surfaces, it generates a set of viewpoints for each surface by inversely composing a view frustum from a surface to its normal direction. An inspection path is computed by finding the minimum distance trajectory that visits at least one viewpoint from each set of viewpoints. A generalized TSP algorithm can compute the minimum distance trajectory. The method then refines the inspection path to cover the near-frontiers by applying the inspection planning method of Song and Jo [143]. Their method has a longer path length and more completion time, but the percentage of the coverage and the quality of the model are improved in their method. However, using a better objective function could improve the quality of the 3-D reconstruction.

Schmid et al. [150] proposed an exploration planning method for reconstructing a surface model. They implemented a new RRT method, which continuously expands a single tree to get global coverage with the maximum utility. Unlike the original RRT, it rewrites tree nodes according to their utility to maintain nonexecuted nodes and subtrees. The method then refines intermediate paths and computes a global coverage path that maximizes the utility. They also proposed an information gain function for TSDF-based surface models. It considers not only a TSDF weight but also a sensing error. The sensing error of the depth sensor is modeled as the quadratic weight of the depth range. Their proposed method can reconstruct more accurate 3-D models than the other surface-based planning methods such as [66] and [126].

A new informed sampling method for exploration path planning is introduced in paper [127]. The informed sampling method ranks the sampled viewpoints based on a utility function and finds the viewpoints candidates that are likely to have high information gain in advance. They used a stereo camera for 3-D mapping the environment and Voxblox [151] as the map representation of the scene. They generated viewpoints based on surface normal and frontier voxels' normal and computed their ranks according to an artificial potential field. The artificial potential field considers MAV position and viewpoint's distinction and repetitiveness. The method then evaluated only a subset of viewpoints according to the ranks. This approach significantly reduces the computation time of viewpoint evaluation.

Song et al. [34] applied a surface-based planning method to online MVS reconstruction for the first time. On the other hand, Song et al. [1] implemented an online MVS system that reconstructs a large-scale scene in real-time. The system computes camera poses from the key-frame-based SLAM [152] and estimates depth maps of key-frames by performing the stereo matching of neighboring key-frames. For the depth estimation, it utilizes a monocular mapping algorithm called REMODE [153]. The estimated depth maps are integrated into a single 3-D model based on the surfel mapping method [145]. They utilized the surface-based exploration method to plan a local

inspection path for scanning low-confidence surfaces. Trajectory optimization is applied to maximize the MVS performance. The authors considered several multiview stereo heuristics for the trajectory optimization, such as parallax, relative distance, and focus angle. Furthermore, Song et al. [1] extended this work for a more precise 3-D model reconstruction. They utilized the deep-learning-based MVS, CasMVSNet [154], instead of REMODE for depth computation. CasMVSNet estimates a depth map by using multiple small cost volumes. It progressively reduces the depth hypothesis range in a coarse-to-fine manner, making it possible to process a high-resolution image in real-time. CasMVSNet provides better reconstruction performances with respect to accuracy and completeness than REMODE.

Surface-based planning analyzes detailed reconstructed surfaces in a surface model; however, the surface model may require a lot of memory. In this case, it can be limited to implementing the 3-D scanning system onboard. To address this problem, a coarse surface model can be utilized for path planning by lowering the resolution of the reconstructed model. However, the advantage of surface-based planning that analyzes the surface shape and trend of a precise 3-D model is lost. Finding a way to efficiently handle large-scale 3-D data for an onboard system would be an interesting direction for future work. TABLE I summarizes the state-of-the-art model-free approaches.

In this section, we addressed three kinds of model-free methods: frontier-based method, volumetric-based planning, and surface-based planning. The frontier-based method simply plans paths toward a boundary of known and unknown regions. This approach is the most efficient method for quickly exploring the largest unknown area. The volumetric-based planning considers the spatial information of 3-D volumetric maps to determine view paths. This method can produce more accurate and complete volumetric models than the frontier-based method. However, the volumetric-based planning is unsuitable for reconstructing precise 3-D models because a volumetric map cannot represent the complex 3-D surfaces exactly. The surface-based planning analyzes the quality and trend of reconstructed surfaces for 3-D scanning. This method can completely reconstruct high-quality 3-D surface models. However, it requires a lot of memory and computation to analyze 3-D data. In conclusion, an appropriate method should be applied according to the scanning purpose (exploration, volumetric mapping, dense surface reconstruction) and available computation power.

### C. Model-Based Methods

Model-based methods leverage a coarse representation of the scene geometry (geometric proxy) for generating optimal viewpoints and path planning. Recently, Agisoft Metashape 1.5 introduced functionality for designing an optimum path-planning for image capturing, which creates mission plans based on the rough model [109], [116]. In this approach, a set of viewpoints with sufficient overlap is generated to cover the surface of the object. However, the computed static network of camera poses cannot promise the completeness of the dataset in real-world situations, and the accuracy of the final point cloud is not considered explicitly. Recently explore-then-exploit (ETE)

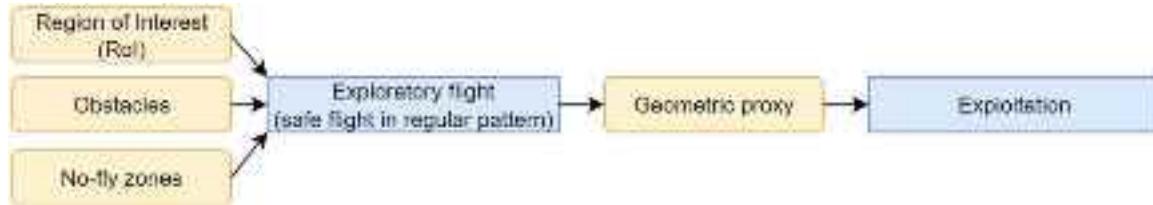


Fig. 2. Overall structure of explore-then-exploit methods.

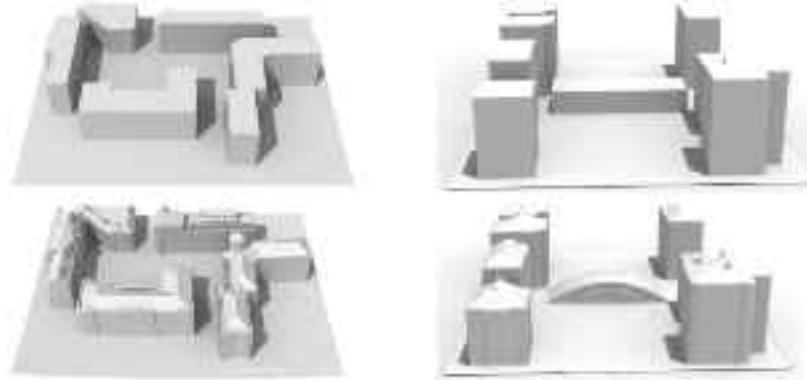


Fig. 3. Comparison of proxies. First row: coarse 2-D proxies generated in [20]; Second row: 3-D proxies generated using an exploratory flight. Images are adopted from [20].

methods (see Fig. 2) which consist of two phases are utilized in UAV mission planning [45], [68], [105], [109]. The first phase is the *exploration* that generates the initial viewpoints and flight path similar to off-the-shelf flight planners or manual flight [19]. The second phase is called the *exploitation* phase, where the geometric proxy is employed to design optimum viewpoints and trajectories to acquire proper images for 3-D reconstruction [55], [161].

The main advantage of the ETE strategy is that prior knowledge of the scene geometry facilitates the global optimization of 3-D coverage and accuracy in the secondary exploitation visit [105].

1) *Exploration*: Although most researchers use a nadir flight for this phase, employing orbit patterns is also reported in some literature [45]. The outcome of exploration phase approximates the scene geometry which is called “geometric proxy.” The geometric proxy is usually a mesh [23], [107], point cloud [68], or supported by voxel representation [26]. The initial model or geometric proxy obtained by exploration, may include many gaps or inaccurate areas [19], [26]. Therefore, this rough 3-D representation of the scene is mostly used for optimal viewpoints planning (to generate a high-fidelity reconstruction) as well as computing the 3-D safe navigable zone for the platform [26], [30], [45]. However, Yan et al. [68], and Zhang et al. [109] used this geometric proxy to locate incomplete or low-quality areas to guide the camera placement in the exploration phase.

Different from most other approaches, which use an exploratory flight for generating geometric proxy, in [20], a very coarse 2.5D coarse model is generated [see Fig. 3(a)] from the Google map (providing 2-D footprints of buildings) and a single

satellite image (for height estimation based on shadows). Using this approach, the planning can be done before visiting the target sites. However, it needs an up-to-date satellite image captured during a sunny day and assumes that the ground is relatively flat. Moreover, extrusion of the 2-D footprint of irregularly shaped architectural objects leads to nonoptimal viewpoints planning in the exploration phase.

For large-scale 3-D reconstruction, Liu et al. [2] utilized the 2-D site boundary and employed Ground-Aware 3-D Object Detector [162] to extract the bounding boxes of individual buildings, on-the-fly. Since these bounding boxes are coarser than geometric proxy, the authors addressed the tasks of explorations and reconstruction path planning through alternating steps.

A semantic-aware exploit-and-explore flight planning approach is introduced in [19]. The main contribution of this research lies at employing semantic information of the exploratory flight for extracting the target object and generating a semantically-enriched coarse proxy for defining inadmissible airspaces for safe trajectories.

A fully convolutional network [163] is applied on 3069 images (from Semantic Drone Dataset<sup>11</sup> and ISPRS 2D Semantic Labelling Benchmark [164] and also some manually annotated UAV images). The open street map information is used to improve the results. The object of interest is extracted using a region growing seeded by the user input, which is applied on the segmentation result. The semantics of the environment is very important (both practically and legally) in defining the flight trajectory, especially in dense urban areas.

<sup>11</sup>[Online]. Available: <https://www.tugraz.at/index.php?id=22387>

It is worth mentioning that with the recent advancements in the concepts and realization of BIM, digital twin, and smart cities, the 3-D rough models of large structures like buildings and bridges are going to be more available. Although this 3-D spatial information can alleviate the exploration in UAV mission planning, semantic and geometric information of the whole scene should be updated for a safe UAV mission and high-fidelity 3-D reconstruction.

2) *Exploitation*: After generating the geometric proxy, exploitation aims at designing optimum viewpoints and trajectories to provide the data for high-quality 3-D reconstruction [55], [161] constrained by platform and environment conditions.

Hoppe et al. [23] proposed an uncertainty-aware and heuristic-based viewpoint planning approach which aims at accurate 3-D coverage of the scene. Their approach belongs to the generate-and-test approaches which generates potential camera poses by assigning one fronto-parallel view to each triangle of the proxy-mesh. Then, an angular histogram (four bins in [0–40]) is computed for each triangle. Another histogram which is the accumulation of cameras of all triangles for each angular bin is used to score the cameras. Based on camera scores (which are updated after the selection of each viewpoint), a greedy selection of the best (i.e., observes the most triangles at a novel angle) overlapping neighboring viewpoints is employed in order to select a subset of the viewpoints to reach one of the two stopping conditions for each (in practice 95%) of the triangles: 1) estimated accuracy, measured by the largest semi-axis of the covariance ellipsoid, is acceptable, or 2) at least one camera of each bin is selected. Although the efficiency of this approach is not reported, we believe that initializing the camera positions by considering one candidate for each triangle of the proxy mesh may generate a very dense search space and is not applicable in real-world large-scale projects. Since this is an off-line approach, no feedback about texture or illumination conditions is considered during data capturing.

Jing et al. [111] subdivided the 3-D rough model using the bubble mesh method. The initial model is divided into small patches. Afterward, randomly sampled viewpoints are generated around the target building. After investigating the visibility (between each surface patch and each viewpoint) and adjacency (between each pair of viewpoints), the selection of viewpoints is performed as a modified set covering with constraints using a neighborhood greedy search algorithm. SfM and MVS heuristics are not involved in the viewpoints selection and computing the orientation of the cameras and quantitative evaluation of the results are missing. Although the proposed method for designing image orientations seems feasible, generating random positions is too risky, and it might cause some gaps or low-quality areas on complex objects.

In the seminal research of Roberts et al. [45], selecting the viewpoints and trajectory planning are jointly considered in a unified global optimization problem as a submodular orienteering problem. This innovative approach finds a path that maximizes scene coverage while adhering to the SfM and MVS heuristics and respecting the limited travel-budget of the drone. After introducing a coverage model, the approximated optimal orientations of all candidate cameras are computed. Submodular

orienteering is then transformed into an additive orienteering problem which is solved as an integer linear problem. However, the coverage model of this approach is monotone which means selecting more cameras never reduces the coverage score, which is not in line with the findings of Seitz et al. [27], and Hepp [28]. Moreover, there is no stopping condition except travel-budget, which is independent of the complexity of the object and desired quality. Lacking online processing also prevents reaction to low-texture areas and moving objects during flight. Lastly, uniform sampling of the object points avoids putting more attention on the edges.

Hepp et al. [26] proposed a system for 3-D reconstruction of building-scale scenes which is similar to [45] in spirit. Based on the coarse mesh (generated from images with overhead pattern), a volumetric occupancy map (fixed voxel size = 0.2 m) containing occupied, free-space, and unobserved voxels is generated where each voxel is also attributed by a measure of observation quality. The authors introduced an approximated camera model (to make the optimization problem suitable for a submodular formulation) to measure expected information gain from an individual viewpoint, independently from other viewpoints. After generating viewpoints candidates, ray-casting is used to compute the visible voxels for each viewpoint to evaluate the corresponding contributed information. Since this model entirely ignores the MVS heuristics, the proposed method encourages fronto-parallel views during the computation of information contribution of voxels. To solve the submodular optimization with travel-budget constraint, an adaption of the recursive strategy which is introduced in paper [165] is used. This strategy recursively splits the trajectory and travel-budget into two parts and selects the reachable (with the current travel budget) viewpoint with the highest information gain. The authors employed voxels' attributes (free, occupied, unknown) to find the free-space motion paths. Paths between viewpoints are preferably a straight line, or in case of an obstacle they are piecewise linear computed using RRT\* algorithm [166]. In spite of very promising results, this approach uses images from the exploratory flight in reconstruction and also in uncertainty computation. This would lead to a problem when the geometric proxy is not generated from the exploratory flight (cf. Section II-B1). Even if a good geometric proxy is available, the exploration should start after an initial flight. Another issue of these methods is that their optimization needs the minimum number of viewpoints as input and is not guaranteed to converge. The user should decide how long to run the optimization. Furthermore, using a recursive greedy algorithm to maximize the optimization objective may lead to some viewpoints which cannot make a single adjustable block. Hence, some other images are added to make sure that all the images are connected. Hoppe et al. [23], and Mostegel et al. [98] solve this problem by applying an overlapping constraint in the recursive selection process.

An ETE approach is introduced in [105] which after exploration and generating the geometric proxy, builds a set of adaptive viewing planes for viewpoint selection. This is an iterative approach in the exploration phase (usually one exploratory flight and two exploitation flights) which repeats data capturing after identifying the low-quality regions until quality converges or a

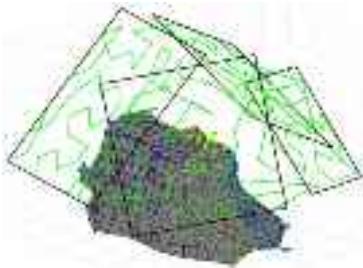


Fig. 4. Adaptive viewing rectangles for different clusters of the scene patch [105].

desired level of quality is achieved. The main novelty of this approach which is an extension of their previous study [36] is considering a set of adaptive viewing planes (see Fig. 4) to limit the search space for camera placement leading to similar resolution for the images in each patch. However, generating viewing planes based on the slope of regions may cause gaps because viewing planes are created based on the average of patch points normal. Hence, in the complex regions the viewing plane might not be able to cover all regions, properly.

In [30] a dilation operator with a ball structure element (whose radius is the desired minimum distance) is applied on the geometric proxy to estimate a closed surface boundary between safe and potentially dangerous airspace. The authors employed some reconstructability heuristics for predicting reconstruction quality. The contradicting effect of parallax on triangulation error and matchability are estimated and considered. Moreover, the distance of the camera to the object and binary visibility of the points based on the geometric scene proxy and the deviation from normal of the surface are also elaborated in the heuristic function. The authors use the geometric proxy to initialize the camera network with a fixed number of uniformly distributed cameras. Utilizing a set of uniformly distributed samples on the geometric proxy, pairwise view parameters are improved incrementally using the simplex method to optimize the reconstructability of the samples. Different aspects of 3-D reconstruction heuristics are formulated and embedded very well in the objective function of this approach. However, since fixed and uniformly distributed sample points are used, depth discontinuities posed some problems in the experiments. Moreover, the inability of this offline approach to adapt the number of views could be improved to avoid oversampling and undersampling of the viewpoints. Furthermore, like almost all other approaches, varying surface texture and lighting are not considered in their approach.

Sharma et al. [107] introduced a floor-wise viewpoints and path planning for optimal building surface coverage. The geometric proxy is a mesh which could be obtained either from building blueprints or cross-sections of a rough model generated by a reconnaissance (exploratory) flight. For each height a contour is generated and minimum 60% overlap between image strips is considered to cover the contours in 3-D. Offsetting the contours by a fixed distance, a 2-D flight map is generated. This initial flight path could be updated to avoid obstacles which are computed with MVS reconstruction from the reconnaissance

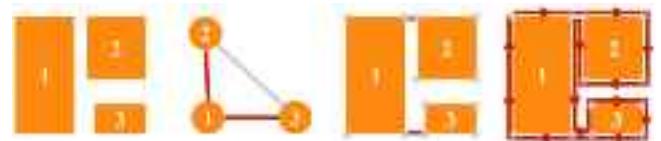


Fig. 5. Bi-level graph-based path planning in [33].

flight. The camera positions are initialized by sampling the flight path every 1 m and camera poses (position and orientations) are then optimized using genetic algorithm to maximize the total coverage such that at least  $N (= 3)$  cameras see each patch of the surfaces. Since this approach does not consider most MVS heuristics, the generated 3-D model is suitable for the applications like virtual reality which the details and completeness of the dataset are more important than the accuracy of 3-D reconstruction.

In [68], by evaluating the quality of the geometric proxy (point cloud), two types of new viewpoints are selected in a constrained view sampling space. The new viewpoints of first type are considered for improving incomplete or low-quality areas and the second type viewpoints are added to fully cover the entire scene. Finally, various optimization algorithms are employed to generate a smooth 3-D trajectory. For this step, ant colony optimization (ACO), particle swarm optimization (PSO), and branch and bound (BnB) methods are compared and the best results are reported using ACO. Similar to paper [30], the sample points on the geometric proxy are distributed uniformly. Therefore, most sample points with lower quality scores were located at the edges of buildings or structures with more geometric details.

Two main objectives in their method are maximizing the coverage and minimizing the path length, so the quality of the 3-D reconstruction does not play any role in their method for viewpoints planning.

Kuang et al. [33] proposed an online path planning method for large-scale urban scene reconstruction. In this research, the geometric proxy of the buildings is generated using an oblique initial flight around the buildings. Then, instances in the scene are segmented using the well-known Mask-RCNN [167] approach and simple geometric shapes are used to represent the layout of the scene. Flight trajectory is minimized in two levels, namely interbuildings and intrabuilding levels. In the first level of optimization each building is considered as a node of a graph (see Fig. 5) and the shortest path is computed using the classical Dijkstra algorithm. In the second level, the task is formulated as the Chinese postman problem.

For on-the-fly predicting the coverage of the whole scene a lightweight SLAM framework is utilized. Then, unobserved spaces are covered by adding new viewpoints. If the endurance of UAV allows, this online approach also has a mechanism which tries to capture close-up images of architectural details. The main limitation of this approach is that SfM and MSV heuristics are not considered in the planning, directly. Moreover, since prior information of the coarse scene model is not completely used during flight, more viewpoints are required.

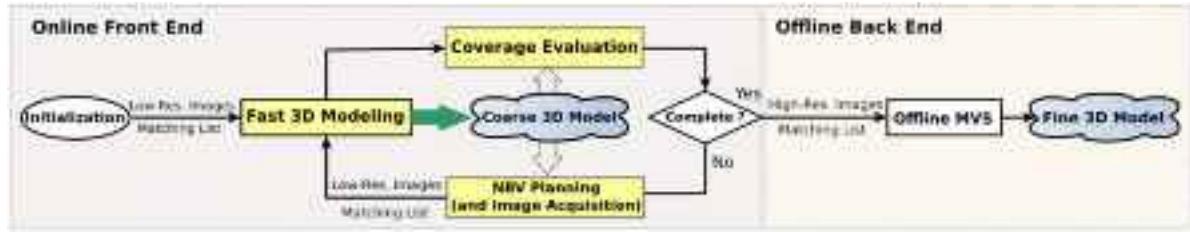


Fig. 6. Pipeline of the proposed approach in [168]. Image courtesy: [168].

Huang et al. [168] proposed an efficient viewpoints and path planning method based on a fast MVS and NBV algorithm. This approach consists of an online front-end for viewpoints and path planning to control the image capturing process and the offline back-end for generating the final dense 3-D model (see Fig. 6). The online front-end incorporates a closed-loop for iterative updating the initial model, coverage evaluation and NBV planning to achieve satisfactory coverage.

The proposed approach uses the initial model generated from an enclosing rectangular flight around the object. A confidence score is computed for each triangle of the mesh. In this confidence score position consistency, normal consistency, and front parallelism are considered. Searching for NBVs is done plane by plane, from high to low altitude to speed up the process by coarsening the search space from 5-D to 4-D. In each iteration, an obstacle-aware A\* algorithm is used for optimal path planning, which connects the NBVs starting from UAV's current position. The main novelty of this approach is using on-the-fly incremental SfM for local and per-vertex updating the mesh model.

Koch et al. [19] proposed a semantic-aware exploit-and-explore flight planning approach to define inadmissible airspaces for safe trajectories. The main contribution of this research lies in employing semantic information of the exploratory flight for extracting the target object and generating a semantically enriched coarse proxy for defining free and occupied airspace and avoiding prohibited flight zones. After sampling a large number of viewpoints on the bounding box of the extracted object of interest, the normalized distance-weighted mean of all visible 3-D points is utilized to compute the orientation of each viewpoint candidate. Distance-based and observation angle-based gains coupled with camera constellation hemisphere slices on each object point are used as heuristics to predict the suitability of the viewpoints for the reconstruction. Inspired by [26] and [45], greedy submodularity optimization is utilized to add the viewpoint with the highest additive information reward to the output set.

Since a prior model of the object is available, model-based methods are able to compute the reasonable poses that could be used to create a complete and accurate 3-D reconstruction. Employing various objective functions which include SfM and MVS heuristics, accuracy and completeness are involved explicitly in selecting optimum viewpoints. Most of these approaches use monocular images and tackle the problem of viewpoints and path planning separately, i.e., after computing the effective viewpoints, path optimization will start to compute an efficient solution to approach the computed poses. As Table II shows,

most model-based methods are offline. An interesting direction could be using a fast online SfM and MVS system to enable on-the-fly computations.

To the best of our knowledge, all model-based approaches use uniform sampling of the proxy for further inspections or use a uniform discretization of the environment as 3-D representation of the object of interest. Investigation of the results of the current approaches reveals that most of the large errors lie at highly curved parts and also sharp corners of the objects. We suggest considering an adaptive sampling of the objects where the overall density depends on the desired levels of details (LoD). However, the density should be also adaptive to local curvature which leads to more samples on the high curvature areas and geometric boundaries.

The Static environment is the prerequisite of the most MVS algorithms [98]. Therefore, detecting and considering moving objects during data capturing is very important. However, this issue is widely ignored by most planning and image acquisition approaches. Another limitation is low-texture areas and building parts with different materials. In some cases, especially for buildings, there might be some low-texture images that can't reconstruct the 3-D model of the object. None of the above-mentioned model-based methods consider neither moving objects nor low-texture images for viewpoints planning.

Although some researchers have already employed Google Map and satellite images for generating geometric proxy, with the recent advancements in the concepts and realization of BIM, digital twin and smart cities, the 3-D rough model of the large structures like buildings and bridges are going to be more available. Even LoD2 information is publicly available in some countries nowadays<sup>12,13</sup> that could be used as a-priori information in viewpoints and path planning algorithms.

Table II summarizes the main characteristics of discussed model-based approach.

### III. EVALUATION STRATEGIES

Evaluation of viewpoints and path planning methods for UAV-based 3-D reconstruction is an important task which poses many significant challenges [19], [26], [169], [170]. First, the important criteria in various communities and projects are different. For example, while visual fidelity and completeness are important in some cases, accuracy is a key in some other

<sup>12</sup>[Online]. Available: <https://3d.bk.tudelft.nl/opendata/opencities/>

<sup>13</sup>[Online]. Available: <https://opengeodata.lgl.niedersachsen.de/#lod2>

TABLE II  
MAIN CHARACTERISTICS OF INVESTIGATED MODEL-BASED APPROACHES

References	Exploration	Exploitation			Path planning	On-line processing		
		pose candidate generation		pose computation				
		position	angles					
Hoppe et al. [23]	Nadir flight	Centers of mesh triangles	Fronto-parallel to the mesh normals	Cluster then select	Greedy shortest path	No		
Huang et al. [169]	Flight on an enclosing rectangle	Plane-wise 2D grid	Regular sampling of pitch and yaw	NBV based on confidence score (function of position consistency, normal consistency, and front parallelism.)	A*	Yes		
Jing et al. [112]	2D map and estimated building height	Random	Surface patch attraction forces	coverage ratio and fully connected viewpoints constraints	---	No		
Peng & Isler [106]	Nadir flight	viewing grid on each adaptive viewing plane	Normals of the adaptive viewing planes	SfM and MVS heuristics	patch-wise TSP	No		
Roberts et al. [45]	Orbit flight	uniform sampling on 3D bounding box	Constrained uniform sampling	Additive submodular orienteering		No		
Smith et al. [30]	Nadir flight	Uniformly distributed views	Randomly oriented simplices around each view candidate	Simplex optimization spherical visibility histograms	TSP	No		
Yan et al. [69]	Nadir flight	centers of view sampling voxels	Normal angle of sample points	Local (confidence score (visibility, distance, sample points normal)) and Global (coverage)	TSP-ACO, PSO, BnB	No		
Kuang et al. [33]	Oblique images	Looking at center of unobserved cells cluster	encourages fronto-parallel views	sorted confidence score (function of distance, number of views, and front parallelism)	Dijkstra	Yes		
Hepp et al. [26]	Nadir flight	regular 3D sampling Add viewpoint to exploration queue	Random sampling with bias towards the interesting area	Information gain maximization and Sparse matching	RRT*	No		
Zhou et al. [20]	GoogleMap + satellite image	one candidate for each sampling point	Fronto-parallel to the sampling point	Max-Min optimization: minimizing the redundancy and maximizing the reconstructability	TSP	No		
Koch et al. [19]	Nadir flight	uniform sampling on 3D bounding box	Normalized distance-weighted mean of 3D points	Submodularity (distance-based and observation angle-based reward)	orienteering	No		
Sharma et al. [108]	Nadir flight	uniform sampling on height-wise boundaries	encourages fronto-parallel views	Genetic algorithm	Fast marching	No		

applications. Moreover, in some projects, overall flight time is critical. Second, environmental conditions such as lighting, weather conditions, and surrounding objects may change or be very challenging [171]. Moreover, moving objects and modifications of the object itself that can occur over the course of time can affect the evaluations. Furthermore, ground truth information is typically not available [19], [26], [169]. Hence, in most studies, the evaluations of the real datasets and scenes are restricted to the qualitative verification. One solution could be using a higher quality sensor, i.e., a terrestrial laser scanner (TLS) which provides accurate point clouds [172]. However, in order to cover the roof and other upper parts of the objects that cannot be captured using TLS, another high-quality low altitude UAV block could be also used. Some common targets can be used to co-register these datasets. This way, almost all parts of the object could be covered and a reliable ground truth could be generated. However, due to lack of resources (time, cost, and hardware) these exhaustive measurements may not be possible in practice. Hence, quantitative evaluations are usually possible and performed on synthetic datasets or specific dataset is used which make the fair and standard comparison of different approaches almost impossible. Last but not least, in most evaluations the quality of the generated 3-D data (e.g., point cloud) is investigated. This indicates the performance of the whole pipeline of planning, data capturing, and 3-D reconstruction (see Fig. 7), which is the nonseparable mixture of these three steps. The interested

reader can refer to paper [27] and the leader-board of [128] for a comparison of the reconstruction approaches. It should be accentuated that even in the case of perfect planning, the effect of hardware, 3-D reconstruction algorithm, and environmental condition could affect the quality of the acquired images. An intermediate solution (even where the ground truth data is not available) could be to employ a reconstructability measure for validation [2]. It leads to decoupling the planning evaluation from the performance of 3-D reconstruction algorithm. We will further elaborate on this topic in Section III-B.

In this section, we first mention the quality measures for the whole pipeline and then investigate measures that indicate the quality of viewpoints and path planning separately.

#### A. Evaluation of the Whole Pipeline

Denoting the ground truth as  $G$  and reconstruction result as  $R$ , we can compute the following quality measures:

- 1) *precision* (how much of  $R$  is supported by  $G$ );
- 2) *recall* (how much of  $G$  is correctly modeled by  $R$ );
- 3) *F-score* (Harmonic mean of *recall* and *precision*);
- 4) *accuracy* (how close  $R$  is to  $G$ ).

Generally, both  $R$  and  $G$  could be a mesh or point cloud. However, considering  $R$  as point cloud is better (avoiding surface reconstruction and its errors) and more convenient (avoiding sampling mesh at its vertices [27]). If  $G$  is delivered as a mesh

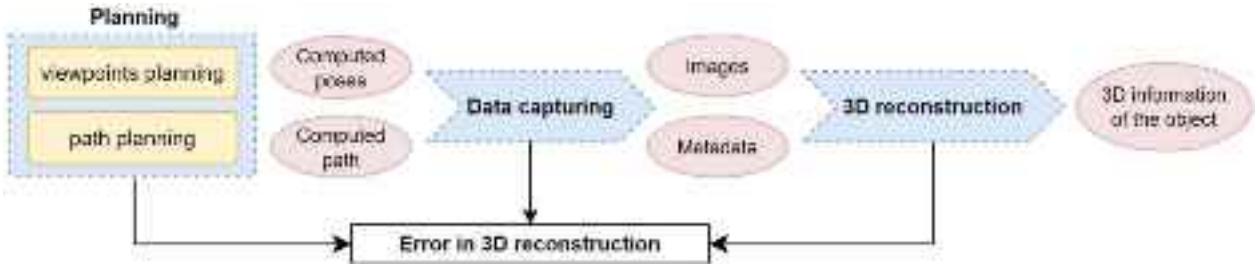


Fig. 7. General pipeline of planning, data capturing and 3-D reconstruction. For the sake of visualization, planning and data capturing are considered as two separate steps (which is not the case for some methods).

or CAD model, it is also better to use it directly to avoid the approximations in resampling it to a point cloud. Nevertheless, if  $G$  is itself a point cloud (e.g., captured by laser scanners) we can directly use it in the point-based evaluations. Following the similar notations as in [27] and [128], the quality measures could be expressed as follows.

*Precision* (Correctness): If the distance between a reconstructed point  $r \in R$  and the ground truth is

$e_{r \rightarrow G} = \min_{g \in G} \|r - g\|$ , the *precision* of the reconstruction for any cut-out distance  $d$  is defined as follows:

$$P(d) = \frac{\sum_{r \in R} [e_{r \rightarrow G} \leq d]}{|R|} \quad (1)$$

where  $[\cdot] \in \{01\}$  is the Iverson bracket and  $|\cdot|$  is the cardinality of the set and  $P(d) \in [01]$ . Selecting the proper value for  $d$  depends on the application and affects the results. For example, [26] considered  $d = 10$  cm which may not be acceptable for some high-quality reconstruction purposes.

*Recall* (Completeness): Similarly, for a ground-truth point  $g \in G$ , its distance to  $R$  is defined as

$e_{g \rightarrow R} = \min_{r \in R} \|g - r\|$  and the *recall* of the reconstruction  $R$  for any cut-out distance  $d$  is  $R(d) \in [01]$  and is defined as follows:

$$R(d) = \frac{\sum_{g \in G} [e_{g \rightarrow R} \leq d]}{|G|}. \quad (2)$$

*F-score*: Summarizes the *precision* and *recall* in a single score as follows:

$$F(d) = \frac{2P(d)R(d)}{P(d) + R(d)}. \quad (3)$$

$F(d) \in [01]$  has a very important property that if either  $P(d) \rightarrow 0$  or  $R(d) \rightarrow 0$  then  $F(d) \rightarrow 0$ , which is not always valid for arithmetic mean.

*Accuracy*: Various metrics using unsigned and signed histograms of distances from  $R$  to  $G$  could be used to measure the accuracy of the reconstruction. While mean and standard deviation of distances could be used [168], the median of the distances is more robust to outliers. Another statistic which is suggested in the literature [27], [108] is the distance  $D$  such that  $X\%$  of the points on  $R$  is within distance  $D$  of  $G$ . Median is a special case of this measure when  $X = 50$ . In [27]  $X = 90$  is used for computing the accuracy.

The main benefit of using signed distances compared to unsigned distances is that it gives a better sense of where and in which directions the reconstructed surface deviates from the ground truth [27], [173]. To this end, one can visualize the distribution and inspect the trend of the signed deviations. Moreover, using unsigned distances changes the distribution of errors. In the case of normally distributed signed errors, the unsigned errors follow the folded normal distribution [174], whose parameters are different from the normal distribution's parameters.

To the best of our knowledge, there is no established framework for the evaluation of viewpoints planning, directly. The most common way is to report one of the metrics mentioned in the previous section as a function of the number of images for the 3-D reconstruction of the object [30], [33], [168]. However, since these metrics are a function of both viewpoints planning quality and 3-D reconstruction quality, these strategies cannot evaluate the viewpoints planning separately. The topic of establishing an evaluation pipeline and proper quality criteria is discussed in paper [169]. Based on the proposed pipeline, the authors evaluated some 3-D UAS flight path planning algorithms. This pipeline focuses on pre-planned flight paths that have been computed based on available rough 3-D models and is not transferable to planning in unseen and unknown environments that need to be explored.

Path planning quality which aims at minimizing the mission duration is also highly correlated with the number and distribution of viewpoints. Therefore, reporting mission duration as a function of measures like recall and precision [33] depends also on view planning strategy. However, using the same viewpoints, path planning quality could be evaluated by measuring the flight path length and data capturing time. UAV battery duration and the speed at which the images can be captured are the main constraints that should be considered in path planning [30]. The impact of different image capturing strategies are investigated in [2] where one of the ablation studies compares Stop-then-Shoot and Shoot-While Flying strategies. Their experiments reveal that Shoot-While-Flying at 6 m/s speed reduces the battery consumption and flight time by more than 60% compared to Stop-then-Shoot at 8 m/s flight speed. Path quality in terms of path length, movement time, and turning time is also analyzed in [108]. We summarized the most common evaluation measures reported in the literature in Table III. For the sake of brevity, just model-based approaches are listed.

TABLE III  
EVALUATION MEASURES REPORTED IN THE MODEL-BASED STUDIES

	<b>Geometric measure</b>	<b>Efficiency measure</b>
Hoppe et al. [23]	Precision: C2M <sup>14</sup> (on synthetic dataset)	Number of images
Huang et al. [169]	Recall: C2C <sup>15</sup> Accuracy: C2C (mean and RMSE)	Flight time, and number of images
Peng & Isler [106]	Recall: based on depth images Accuracy: Mean and standard deviation of the depth	Number of images
Roberts et al. [45]	Recall: C2C Accuracy: C2C	Travel budget (fixed)
Smith et al. [30]	Recall: C2M (on synthetic dataset) Accuracy: C2M (on synthetic dataset)	Flight time, and Flight path length
Yan et al. [69]	Recall: Voxel occupancy Accuracy: C2M (on synthetic dataset)	Flight time, and Flight path length
Kuang et al. [33]	Recall: C2M (on synthetic dataset) Accuracy: C2M (RMSE)	Flight time, and Number of images
Hepp et al. [26]	Precision: C2C Recall: C2C	Flight path length and Number of images
Zhou et al. [20]	Recall: C2M Accuracy: C2M	Flight time, path length and Number of images
Koch et al. [19]	Precision: C2C Recall: C2C	Flight path length and Number of images
Sharma et al. [108]	Accuracy: Hausdorff distance (mean and RMS)	Flight time

<sup>14</sup>Cloud to Mesh.

<sup>15</sup>Cloud to Cloud.

There are also some attempts to combine different quality metrics to provide a unified score for both viewpoints and path planning steps to make different approaches comparable. In [169] an evaluation pipeline is introduced which employs a normalized score for weighted combination of viewpoints and path planning quality as follows:

$$\hat{S} = \Psi \frac{A}{L \cdot d^2} \quad (4)$$

where  $\Psi$  is the quality factor of the flight path,  $L$  is the cost of the flight path length,  $A$  is the surface area of the 3-D model, and  $d$  is the distance between camera and object. However, there is no optimal value for this normalized score and it could be used just for a relative comparison of the planning approaches. Moreover, some hyperparameters should be added to (4) to tune the importance of viewpoints and path planning quality measures for different applications.

#### B. Outdoor 3-D Reconstruction Benchmarks

Preparation of reliable ground truth for the whole scene is not a straight-forward task. Usually, more than one sensor should be utilized [175] to capture the entire object with the accuracy of at least one level of magnitude better than the data under investigation. Therefore, many researchers perform the qualitative evaluation on real datasets, while quantitative evaluation is

performed on synthetic datasets [26], [68]. Before discussing the synthetic datasets and development engines, it worth mentioning that some researchers employed the reconstructability measure to evaluate the planning algorithms.

While the reconstructability measure is usually used for optimizing the viewpoints and path planning algorithms (e.g., [30], [86]), it is utilized in [2] for evaluation. The authors also investigated the relation between the accuracy (w.r.t ground truth) and the reconstructability measure. The benefit of analyzing the reconstructability for evaluation is twofold. This not only provides the possibility to validate the planning algorithm in absence of ground truth data, but also decouples the planning and data capturing quality from 3-D reconstruction (cf. Fig. 7). However, it should be mentioned that this approach highly depends on the efficacy of the reconstructability measure and its underlying assumptions. For example, while most methods consider view pairs to compute the reconstruction error, Liu et al. [86] factored in multiple viewpoints' contribution in their uncertainty-aware learning-based approach.

Unreal Engine (UE<sup>16</sup>) is a game development engine that produces photorealistic scenes. UE is by far the most common simulation environment which is used in many research works in viewpoints and path planning field [26], [33], [36], [45]. An

<sup>16</sup>[Online]. Available: <https://www.unrealengine.com/en-US/>

open-source Python library called UnrealCV [176] could also be used to set the camera parameters and control a virtual camera in the simulated environments [45], [105]. Since there is full control over the camera parameters, objects (like buildings) and also lighting and texture it is possible to create many different scenes and reconfigure their attributes that is a very useful tool for evaluation purposes. Smith et al. [30] created a benchmark which consists of 34 uniquely different buildings in five large urban scenes and packaged them within a release project of UE, which is publicly accessible via their project website.<sup>17</sup> The authors developed a new synthetic benchmark dataset and simulation environment for UAV path planning problem in urban environments making it possible to quantitatively evaluate other viewpoints and path planning algorithms. A small part of one of synthetic samples from this benchmark is illustrated in Fig. 8(a). Microsoft AirSim [177], an open-source UAV simulator built on Unreal Engine, is also employed in the literature to simulate the flight process and plan the UAV flight path [33], [178].

Unity game engine<sup>18</sup> is another environment that is employed in [107] to create a 3-D scene [see Fig. 8(b)]. Scripting capabilities of the game engine provides the possibility to capture the images using a virtual camera. An experimental Unity release of AirSim is also announced.<sup>19</sup> In some studies like [19] Koch et al. generated their synthetic dataset [see Fig. 8(c)] using other environments, e.g., open-source computer graphics software Blender [179] which is a well-known 3-D modeling and rendering package. 3-D architecture models from the 3-D Ware-house of SketchUp<sup>20</sup> are also used [168]. The authors used the Gazebo robot simulator<sup>21</sup> for simulation and controlling the platform.

Terragen,<sup>22</sup> a scenery generator program, is another environment employed in the literature [180], [181]. Recently, the concept of an autonomous environment generator for UAV-based simulation based on machine learning algorithms was introduced in [182]. Based on satellite images, this approach introduced the concept of procedurally generating, scaling, and placing 3-D models to create a realistic environment. The interested reader can refer to papers [182] and [183] for more details on simulation environments.

#### IV. DISCUSSION AND FUTURE TRENDS

Although much work and progress are observable in the field of viewpoints and path planning, existing algorithms are far from the last word on this topic, and there is plenty of exciting work left to do. In this section, we discuss the current challenges in viewpoints and path planning approaches which show the direction for future research on this topic.

##### A. Model-Free versus Explore-Then-Exploit Methods

Most of the model-free approaches rely on NBV and rapidly traverse along the direction that decreases the model uncertainty

without using any prior information about the scene and object. Although this is very efficient for fast exploration of the objects, their local search strategy within a nonlinear objective function *may get stuck in local minima*. It cannot guarantee a complete coverage of the object, especially its details. Furthermore, most model-free methods use stereo cameras or depth cameras to capture data for 3-D reconstruction. It means the platform should get close enough to the object. Although it is feasible for indoor applications and small objects, it could be very time-consuming or even impossible for large-scale 3-D reconstruction purposes. On the other hand, having a priori knowledge of the general geometry enables the explore-then-exploit methods to optimize the coverage and accuracy of the results globally leading to high-quality results and smoother trajectories. However, *most of the existing explore-then-exploit approaches are offline* and do not update the computed poses based on the online feedback from the data capturing unit. Therefore, considering the current computational power of on-board processing units, refining the global design of the poses and updating them locally based on the feedback from the acquired images could be a very interesting direction for future research. Nowadays, this idea could also be more realistic with the availability of very fast data transfer technologies like 5G and the popularity of cloud-based processing. Furthermore, finding a way to efficiently handling the large-scale 3-D data on an onboard system would be an interesting future work. This also could close the gap between two flight missions (exploration and exploitation) which requires two visits from the site, which is not always possible in practice.

##### B. Practical Considerations

Like almost all other projects in engineering, one interesting direction could be *as-planned versus as-is* investigation. Even in the case of a perfect planning, many other issues could affect the final outcome. Most viewpoints and path planning methods assume that the estimated 3-D data and localization information are very accurate. They also assume that an MAV could accurately follow the planned paths. However, in practice, these assumptions are not usually satisfied. Hardware limitations, meteorological conditions, objects complexities, and sensor-object relative conditions are the main factors that can further increase the uncertainty of the motion and state estimation of the MAV or the quality of the final product. GNSS hardware and setup is an example in this regard which affects the data capturing quality. Although GNSS provides valuable information for UAV-based data capturing and processing, when the platforms fly at a low distance to the object of interest or under the structures like bridges, the GNSS signal can be strongly disturbed. Therefore, integrating visual SLAM and GNSS-based navigation and data capturing could be an interesting topic to have the best of both worlds. Drift on the gimbal and the magnetometer compass affects capturing the data with the already planned angles. Furthermore, due to safety issues, e.g., minimum controllable flight height, some lower parts of the structures cannot be well captured and reconstructed [33]. To the best of our knowledge, most of the current approaches are designed for ideal conditions and neglect the effect of wind and other meteorological conditions. For example, wind affects the platform's ability to reach the planned

<sup>17</sup>[Online]. Available: <https://vccimaging.org/Publications/Smith2018UAVPathPlanning/>

<sup>18</sup>[Online]. Available: <https://unity.com/>

<sup>19</sup>[Online]. Available: <https://microsoft.github.io/AirSim/>

<sup>20</sup>[Online]. Available: <https://3dwarehouse.sketchup.com/>

<sup>21</sup>[Online]. Available: <http://gazebosim.org>

<sup>22</sup>[Online]. Available: <https://planetside.co.uk/>

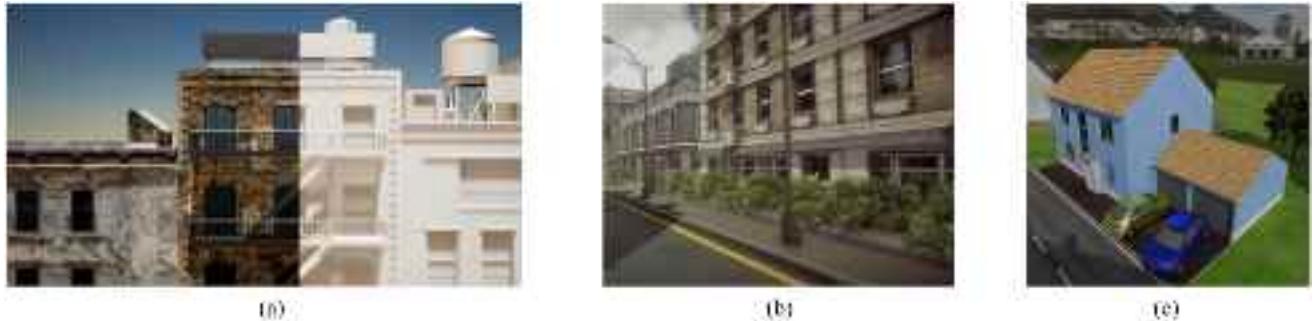


Fig. 8. Three examples of synthetic environments. (a) Part of one of the synthetic samples from [30] left part of the scene is shown after texturing, while the right side is shown before texturing. (b) Series of building in Unity engine.<sup>23</sup> (c) Synthetic dataset created in [19] using Blender.

viewpoints. Some studies [140], [184] addressed the problems occurring by fast motion of the platform by considering the localization uncertainty of SLAM for view planning. They integrated exploration or coverage planning with an active SLAM approach, which aims to minimize the localization uncertainty of SLAM. They produced navigation paths to track many SLAM features to reduce localization uncertainty. However, although their methods improved the localization accuracy, MAVs are still unable to navigate quickly because the generated paths do not satisfy MAV's dynamic properties. To address this issue, some studies [1], [22], [139] restricted the maximum limit of the platform's motion speed to be small and then performs a scanning task. However, this approach cannot be the ultimate solution, because the limited speed comes at the cost of longer data capturing time. Therefore, *considering localization uncertainty in view-path-planning* using the approaches like active SLAM [140], [185], [186] and the fast motion generation methods [123], [141] could be another direction for future work. Since battery limitation (flight time) is a challenge for almost all systems, multi-agent systems like [187] seem to be very effective for large-scale 3-D reconstruction purposes. However, current issues like efficient and reliable tasks distribution between the agent are still challenging.

Besides all deviations caused by hardware imperfection and environmental conditions, the *appearance of the objects with different materials* can also affect the quality of 3-D reconstruction. Shadow and illumination are also very important factors that can hinder reaching a high-fidelity 3-D model. Reflective or low-texture surfaces like glass should also be considered. Moreover, since the texture resolution is a function of image scale, it can pose another constraint on viewpoints planning which affects the acceptable distance between the sensor and low-texture parts of the objects. Hence, deviation of as-is poses of the captured images from as-planned ones necessitates an iterative closed-loop and online feedback between data capturing and planning modules that could be another direction for future research.

Since the perception of the surrounding environment is a key for collision-free UAV flights, commercial UAVs are equipped

with some algorithms/sensors to perform a safe flight. However, when *dealing with dynamic scenes* (which is the case in many real scenarios), other considerations like moving objects detection also comes into play and should be considered in data capturing. Recently, [188] employed a YoLO object detector [189] to find obstacles and pedestrians in UAV's navigation environment. This would also be important to avoid any occlusion caused by moving objects between the sensor and the object of interest.

While some algorithms prioritize reconstruction quality, others emphasize the flight time (equivalently trajectory length or the number of images). Future research could also concentrate on finding reasonable and flexible solutions that can switch between these two possibilities or even add some hyperparameters for *tuning the importance of 3-D reconstruction quality and the trajectory length* in order to fulfill the requirements of different applications.

Decoupling the planning quality from efficacy of 3-D reconstruction algorithms should be emphasized. Therefore, establishing *standard frameworks and strategies for evaluation* of viewpoints planning, path planning, and combination of both is crucial. Improving reconstructability predictors could be a relevant research direction in this regard. Moreover, availability of publicly available benchmark datasets like UrbanScene3D [190], and open-source implementation of the methods like paper [30] will pave the way for future research in this field.

### C. Artificial Intelligence (AI) and Machine Learning-Based Approaches

Some studies recently applied AI and machine learning techniques for viewpoints and path planning. Hepp et al. [35] proposed an NBV planning method where the viewpoint's utility is computed from a 3-D convolutional neural network (CNN). The method used a multiscale voxel representation of a partially explored scene as an input of the CNN. It trains utility scores of viewpoints from an oracle with access to ground truth information. Peralta et al. [159] also proposed a learning-based NBV planning method for scanning houses. They provide a dataset of 3-D house models for training NBV policies. The authors trained a deep Q-network [191] and deep deterministic policy gradient [191] based on the house dataset. Zeng et al. [160] proposed a deep-learning network for NBV planning, which directly uses

<sup>23</sup>[Online]. Available: <https://unityassetsfreedom.club/download/town-constructor-pack/>

raw point cloud data instead of a volumetric model. The network extracts the feature of partially reconstructed point cloud and predicts the information gain of viewpoints from the extracted feature. This network can estimate an information gain more efficiently than the conventional ray-casting approaches. An exploration planning method based on deep reinforcement learning is proposed in [82]. This method trains topological information of an office-like environment, which provides a guide to efficiently compute the visiting sequence for unexplored regions. Martin et al. [181] proposed using genetic algorithms to optimize viewpoints planning for accurate 3-D reconstruction using small UAVs. The coverage and accuracy of the generated model are formulated in the objective function of the genetic algorithms. The objective function is defined based on the number of visible terrain points in the images of a solution set and the need to capture terrain points from multiple angles. The evaluation results show the improvement in the completeness and standard deviations of case studies in comparison with the basic grid survey. A hierarchical framework for path generation, coverage path planning, and dynamic obstacle avoidance is presented in [192]. Satellite images and maps of farms were used via a deep learning method for path planning. A faster R-CNN network [193] localizes and identifies objects and obstacles of humans and vehicles. Lie et al. [194] introduced VGF-NET, a visual-geometric fusion learning method consisting convolutional layers and SLAM, followed by a directional attention model for navigating the drone and generating a 2.5D height map, simultaneously. Xie et al. [195] proposed a deep reinforcement learning approach for path planning in a dynamic environment. They formulated this issue as a partially observable markov decision process and solved it by utilizing the local information and relative distance without global information. Historical state-action sequences and multisensors are used by the proposed method to achieve more reasonable decision-making. Recently, a model-based and data-driven approach was proposed that employs a neural network to predict the reconstructability as a function of the proxy geometry, a set of viewpoints, and optionally on-the-flight images [86]. This learning model consists of an attention-based view fusion network to predict the reconstructability and another attention-based image fusion network for uncertainty modelling. The authors also integrated a novel view planning framework to avoid sticking in the local minima. Although there is a rich body of literature on AI-based path and coverage planning [196], [197], [198], [199], [200], [201], [202], viewpoints planning for 3-D reconstruction based on AI and machine/deep learning approaches is still in its infancy and there is a large room to deepen the insight into this topic.

#### ACKNOWLEDGMENT

The authors would like to thank all the reviewers and editors for their valuable and constructive comments. M. Maboudi and M. Gerke would like to thank the DFG for the support within the SFB/Transregio 277- Additive manufacturing in construction. (Subproject C06). The authors would also like to acknowledge support by the Open Access Publication Funds of the Technische Universität Braunschweig.

#### REFERENCES

- [1] S. Song, D. Kim, and S. Choi, "View path planning via online multiview stereo for 3-D modeling of large-scale structures," *IEEE Trans. Robot.*, vol. 38, no. 1, pp. 372–390, Feb. 2022.
- [2] Y. Liu, R. Cui, K. Xie, M. Gong, and H. Huang, "Aerial path planning for online real-time exploration and offline high-quality reconstruction of large-scale urban scenes," *ACM Trans. Graph.*, vol. 40, no. 6, pp. 1–16, Dec. 2021, doi: [10.1145/3478513.3480491](https://doi.org/10.1145/3478513.3480491).
- [3] P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas, and I. Moscholios, "A compilation of UAV applications for precision agriculture," *Comput. Netw.*, vol. 172, 2020, Art. no. 107148, doi: [10.1016/J.COMNET.2020.107148](https://doi.org/10.1016/J.COMNET.2020.107148).
- [4] C. Torresan et al., "Forestry applications of UAVs in Europe: A review," *Int. J. Remote Sens.*, vol. 38, no. 8–10, pp. 2427–2447, 2016, doi: [10.1080/01431161.2016.1252477](https://doi.org/10.1080/01431161.2016.1252477).
- [5] R. Eskandari, M. Mahdianpari, F. Mohammadimansh, B. Salehi, B. Brisco, and S. Homayouni, "Meta-analysis of unmanned aerial vehicle (UAV) imagery for agro-environmental monitoring using machine learning and statistical models," *Remote Sens.*, vol. 12, no. 21, Oct. 2020, Art. no. 3511, doi: [10.3390/RS12213511](https://doi.org/10.3390/RS12213511).
- [6] M. Maboudi, A. Alamouri, V. López De Arriba, M. S. Bajauri, C. Berger, and M. Gerke, "Drone-based container crane inspection: Concept, challenges and preliminary results," in *Proc. ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, 2021, pp. 121–128, doi: [10.5194/ISPRS-ANNALS-V-1-2021-121-2021](https://doi.org/10.5194/ISPRS-ANNALS-V-1-2021-121-2021).
- [7] J. Seo, L. Duque, and J. Wacker, "Drone-enabled bridge inspection methodology and application," *Automat. Construction*, vol. 94, pp. 112–126, Oct. 2018, doi: [10.1016/J.AUTCON.2018.06.006](https://doi.org/10.1016/J.AUTCON.2018.06.006).
- [8] J. I. Vasquez-Gomez, D. E. Troncoso Romero, M. Antonio-Cruz, and E. Zamora, "Spiral trajectories for building inspection with quadrotors," in *Proc. Int. Conf. Unmanned Aircr. Syst.*, 2022, pp. 891–896, doi: [10.1109/ICUAS54217.2022.9836133](https://doi.org/10.1109/ICUAS54217.2022.9836133).
- [9] P. Gospodnetić, D. Mosbach, M. Rauhut, and H. Hagen, "Viewpoint placement for inspection planning," *Mach. Vis. Appl.*, vol. 33, no. 1, pp. 1–21, Jan. 2022, doi: [10.1007/S00138-021-01252-Z/TABLES7](https://doi.org/10.1007/S00138-021-01252-Z/TABLES7).
- [10] T. Rakha and A. Gorodetsky, "Review of unmanned aerial system (UAS) applications in the built environment: Towards automated building inspection procedures using drones," *Automat. Construction*, vol. 93, pp. 252–264, Sep. 2018, doi: [10.1016/J.AUTCON.2018.05.002](https://doi.org/10.1016/J.AUTCON.2018.05.002).
- [11] N. Jacob-Loyola, F. Muñoz-La Rivera, R. F. Herrera, and E. Atencio, "Unmanned aerial vehicles (Uavs) for physical progress monitoring of construction," *Sensors*, vol. 21, no. 12, Jun. 2021, Art. no. 4227, doi: [10.3390/S21124227](https://doi.org/10.3390/S21124227).
- [12] Y. Xu and Y. Turkan, "BrIM and UAS for bridge inspections and management," *Eng. Construction Architectural Manage.*, vol. 27, no. 3, pp. 785–807, Mar. 2020, doi: [10.1108/ECAM-12-2018-0556/FULL/XML](https://doi.org/10.1108/ECAM-12-2018-0556/FULL/XML).
- [13] Y. Li and C. Liu, "Applications of multirotor drone technologies in construction management," *Int. J. Construction Manage.*, vol. 19, no. 5, pp. 401–412, Sep. 2018, doi: [10.1080/15623599.2018.1452101](https://doi.org/10.1080/15623599.2018.1452101).
- [14] S. Malihí, M. J. V. Zoej, M. Hahn, and M. Mokhtarzade, "Window detection from UAS-derived photogrammetric point cloud employing density-based filtering and perceptual organization," *Remote Sens.*, vol. 10, no. 8, Aug. 2018, Art. no. 1320, doi: [10.3390/RS10081320](https://doi.org/10.3390/RS10081320).
- [15] S. Sahebdivan, H. Arefi, and M. Maboudi, "Rail track detection and projection-based 3D modeling from UAV point cloud," *Sensors*, vol. 20, no. 18, 2020, Art. no. 5220, doi: [10.3390/s20185220](https://doi.org/10.3390/s20185220).
- [16] J. M. de Luis-Ruiz, J. Sedano-Cibrián, R. Pereda-García, R. Pérez-álvarez, and B. Malagón-Picón, "Optimization of photogrammetric flights with UAVs for the metric virtualization of archaeological sites. application to juliobriga (Cantabria, Spain)," *Appl. Sci.*, vol. 11, no. 3, Jan. 2021, Art. no. 1204, doi: [10.3390/APPI1031204](https://doi.org/10.3390/APPI1031204).
- [17] S. Karam et al., "Micro and macro quadcopter drones for indoor mapping to support disaster management," in *Proc. XXIV ISPRS Congr.*, 2022, pp. 203–210, doi: [10.5194/ISPRS-ANNALS-V-1-2022-203-2022](https://doi.org/10.5194/ISPRS-ANNALS-V-1-2022-203-2022).
- [18] N. Kerle, F. Nex, M. Gerke, D. Duarte, and A. Vetrivel, "UAV-based structural damage mapping: A review," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 1, Dec. 2019, Art. no. 14, doi: [10.3390/ijgi9010014](https://doi.org/10.3390/ijgi9010014).
- [19] T. Koch, M. Körner, and F. Fraundorfer, "Automatic and semantically-aware 3D UAV flight planning for image-based 3D reconstruction," *Remote Sens.*, vol. 11, no. 13, 2019, Art. no. 1550, doi: [10.3390/rs11131550](https://doi.org/10.3390/rs11131550).
- [20] X. Zhou et al., "Offsite aerial path planning for efficient urban scene reconstruction," *ACM Trans. Graph.*, vol. 39, no. 6, pp. 1–16, 2020, doi: [10.1145/3414685.3417791](https://doi.org/10.1145/3414685.3417791).

- [21] C. Fraser, "Network design considerations for non-topographic photogrammetry," *Photogrammetric Eng. Remote Sens.*, vol. 50, no. 8, pp. 1115–1126, 1984.
- [22] S. Song, D. Kim, and S. Jo, "Online coverage and inspection planning for 3D modeling," *Auton. Robots*, vol. 44, no. 8, pp. 1431–1450, 2020, doi: [10.1007/s10514-020-09936-7](https://doi.org/10.1007/s10514-020-09936-7).
- [23] C. Hoppe et al., "Photogrammetric camera network design for micro aerial vehicles photogrammetric camera network design for micro aerial vehicles," in *Proc. 17th Comput. Vis. Winter Workshop*, vol. 8, 2012, pp. 1–8.
- [24] K. Schmid, H. Hirschmüller, A. Dömel, I. Grix, M. Suppa, and G. Hirzinger, "View planning for multi-view stereo 3D reconstruction using an autonomous multicopter," *J. Intell. Robot. Syst. Theory Appl.*, vol. 65, no. 1–4, pp. 309–323, Jan. 2012, doi: [10.1007/S10846-011-9576-2](https://doi.org/10.1007/S10846-011-9576-2).
- [25] M. Saponaro, A. Capolupo, E. Tarantino, and U. Fratino, "Comparative analysis of different UAV-based photogrammetric processes to improve product accuracies," in *Proc. Int. Conf. Comput. Sci. Appl.*, 2019, pp. 225–238, doi: [10.1007/978-3-030-24305-0\\_18](https://doi.org/10.1007/978-3-030-24305-0_18).
- [26] B. Hepp, M. Niebner, and O. Hilliges, "Plan3D: Viewpoint and trajectory optimization for aerial multi-view stereo reconstruction," *ACM Trans. Graph.*, vol. 38, no. 1, pp. 1–17, Dec. 2018, doi: [10.1145/3233794](https://doi.org/10.1145/3233794).
- [27] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2006, vol. 1, pp. 519–526, doi: [10.1109/CVPR.2006.19](https://doi.org/10.1109/CVPR.2006.19).
- [28] B. Hepp, *Planning for Autonomous Micro-Aerial Vehicles with Applications to Filming and 3D Modeling*. Zürich, Switzerland: ETH Zürich, 2018.
- [29] A. Hornung, B. Zeng, and L. Kobbelt, "Image selection for improved multi-view stereo," in *Proc. IEEE 26th Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8, doi: [10.1109/CVPR.2008.4587688](https://doi.org/10.1109/CVPR.2008.4587688).
- [30] S. Neil, M. Moehrle, M. Goesele, and W. Heidrich, "Aerial path planning for urban scene reconstruction," *ACM Trans. Graph.*, vol. 37, Dec. 2018, Art. no. 16, doi: [10.1145/3272127.3275010](https://doi.org/10.1145/3272127.3275010).
- [31] M. Roberts, "Trajectory optimization methods for drone cameras," Ph.D. thesis, Stanford Univ., Stanford, CA, USA, 2019.
- [32] M. D. Kaba, M. G. Uzunbas, and S. N. Lim, "A reinforcement learning approach to the view planning problem," in *Proc. IEEE 30th Conf. Comput. Vis. Pattern Recognit.*, 2017, vol. 2017, pp. 5094–5102, doi: [10.1109/CVPR.2017.541](https://doi.org/10.1109/CVPR.2017.541).
- [33] Q. Kuang, J. Wu, J. Pan, and B. Zhou, "Real-time UAV path planning for autonomous urban scene reconstruction," in *Proc. Int. Conf. Robot. Automat.*, 2020, pp. 1156–1162, doi: [10.1109/ICRA40945.2020.9196558](https://doi.org/10.1109/ICRA40945.2020.9196558).
- [34] S. Song, D. Kim, and S. Jo, "Active 3D modeling via online multi-view stereo," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 5284–5291, doi: [10.1109/ICRA40945.2020.9197089](https://doi.org/10.1109/ICRA40945.2020.9197089).
- [35] C. I. Connolly, "The determination of next best views," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1985, pp. 432–435, doi: [10.1109/ROBOT.1985.1087372](https://doi.org/10.1109/ROBOT.1985.1087372).
- [36] C. Peng and V. Isler, "View selection with geometric uncertainty modeling," 2017, *arXiv:1704.00085*.
- [37] A. H. Ahmadabadian, S. Robson, J. Boehm, and M. Shortis, "Stereovision network design for precise and dense 3D reconstruction," *Photogrammetric Rec.*, vol. 29, no. 147, pp. 317–336, 2014, doi: [10.1111/phor.12076](https://doi.org/10.1111/phor.12076).
- [38] W. R. Scott, G. Roth, and J. F. Rivest, "View planning for automated three-dimensional object reconstruction and inspection," *ACM Comput. Surv.*, vol. 35, no. 1, pp. 64–96, Mar. 2003, doi: [10.1145/641865.641868](https://doi.org/10.1145/641865.641868).
- [39] J. Aloimonos, I. Weiss, and A. Bandyopadhyay, "Active vision," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 333–356, Jan. 1988, doi: [10.1007/BF00133571](https://doi.org/10.1007/BF00133571).
- [40] J. Zeng et al., "Efficient view path planning for autonomous implicit reconstruction," Sep. 2022, *arXiv:2209.13159*.
- [41] S. Jiang, W. Jiang, and L. Wang, "Unmanned aerial vehicle-based photogrammetric 3D mapping: A survey of techniques, applications, and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 135–171, Jun. 2022.
- [42] D. Scaramuzza et al., "Vision-controlled micro flying robots: From system design to autonomous navigation and mapping in GPS-denied environments," *IEEE Robot. Automat. Mag.*, vol. 21, no. 3, pp. 26–40, Sep. 2014.
- [43] S. Chen, Y. Li, and N. M. Kwok, "Active vision in robotic systems: A survey of recent developments," *Int. J. Robot. Res.*, vol. 30, no. 11, pp. 1343–1377, Aug. 2011, doi: [10.1177/0278364911410755](https://doi.org/10.1177/0278364911410755).
- [44] Y. Kong, F. Zhu, H. Sun, L. Zhiyuan, Q. Wang, and J. Wang, "A view planning method for 3D reconstruction with unknown feature prediction," in *Proc. 8th Int. Conf. Mechatron. Robot. Eng.*, 2022, pp. 182–187, doi: [10.1109/ICMRE54455.2022.9734106](https://doi.org/10.1109/ICMRE54455.2022.9734106).
- [45] M. Roberts et al., "Submodular trajectory optimization for aerial 3D scanning," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 5334–5343, doi: [10.1109/ICCV.2017.569](https://doi.org/10.1109/ICCV.2017.569).
- [46] M. Saadatseresht, F. Samadzadegan, and A. Azizi, "Automatic camera placement in vision metrology based on a fuzzy inference system," *Photogrammetric Eng. Remote Sens.*, vol. 71, no. 12, pp. 1375–1385, 2005, doi: [10.14358/PERS.71.12.1375](https://doi.org/10.14358/PERS.71.12.1375).
- [47] S. Mason, "Heuristic reasoning strategy for automated sensor placement," *Photogrammetric Eng. Remote Sens.*, vol. 63, no. 9, pp. 1093–1102, 1997.
- [48] A. Hosseiniaveneh and F. Remondino, "An imaging network design for UGV-based 3D reconstruction of buildings," *Remote Sens.*, vol. 13, no. 10, pp. 1–28, 2021, doi: [10.3390/rs13101923](https://doi.org/10.3390/rs13101923).
- [49] P. Huang, L. Lin, K. Xu, and H. Huang, "Autonomous outdoor scanning via online topological and geometric path optimization," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 4, pp. 3682–3695, Apr. 2022.
- [50] C. Cao, H. Zhu, H. Choset, and J. Zhang, "TARE: A hierarchical framework for efficiently exploring complex 3D environments," in *Robot.: Sci. Syst. Conf.*, 2021, doi: [10.15607/rss.2021.xvii.018](https://doi.org/10.15607/rss.2021.xvii.018).
- [51] M. Panda, B. Das, B. Subudhi, and B. B. Pati, "A comprehensive review of path planning algorithms for autonomous underwater vehicles," *Int. J. Automat. Comput.*, vol. 17, no. 3, pp. 321–352, Jan. 2020, doi: [10.1007/S11633-019-1204-9](https://doi.org/10.1007/S11633-019-1204-9).
- [52] S. Wu et al., "Quality-driven poisson-guided autoscaning," *ACM Trans. Graph.*, vol. 33, no. 6, Nov. 2014, Art. no. 203, doi: [10.1145/2661229.2661242](https://doi.org/10.1145/2661229.2661242).
- [53] B. Bogaerts, S. Sels, S. Vanlanduit, and R. Penne, "Near-optimal path planning for complex robotic inspection tasks," 2019, pp. 1–16, [Online]. Available: <http://arxiv.org/abs/1905.05528>
- [54] W. Jing, D. Deng, Y. Wu, and K. Shimada, "Multi-UAV coverage path planning for the inspection of large and complex structures," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2020, pp. 1480–1486, doi: [10.1109/IROS45743.2020.9341089](https://doi.org/10.1109/IROS45743.2020.9341089).
- [55] A. Sadeghi, A. B. Asghar, and S. L. Smith, "On minimum time multi-robot planning with guarantees on the total collected reward," in *Proc. Int. Symp. Multi-Robot Multi-Agent Syst.*, 2019, pp. 16–22, doi: [10.1109/MRS.2019.8901079](https://doi.org/10.1109/MRS.2019.8901079).
- [56] R. Nagasawa, E. Mas, L. Moya, and S. Koshimura, "Model-based analysis of multi-UAV path planning for surveying postdisaster building damage," *Sci. Rep.*, vol. 11, no. 1, pp. 1–14, Sep. 2021, doi: [10.1038/s41598-021-97804-4](https://doi.org/10.1038/s41598-021-97804-4).
- [57] S. Ivić, B. Crnković, L. Grbčić, and L. Matleković, "Multi-UAV trajectory planning for 3D visual inspection of complex structures," *Automat. Construction*, vol. 147, 2023, Art. no. 104709.
- [58] F. Nex et al., "UAV in the advent of the twenties: Where we stand and what is next," *ISPRS J. Photogrammetry Remote Sens.*, vol. 184, pp. 215–242, Feb. 2022, doi: [10.1016/J.ISPRSJPTRS.2021.12.006](https://doi.org/10.1016/J.ISPRSJPTRS.2021.12.006).
- [59] G. Hardouin, F. Morbidi, J. Moras, J. Marzat, and E. M. Mouaddib, "Surface-driven next-best-view planning for exploration of large-scale 3D environments," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 15501–15507, 2020, doi: [10.1016/j.ifacol.2020.12.2376](https://doi.org/10.1016/j.ifacol.2020.12.2376).
- [60] A. Deris, I. Trigonis, A. Aravanis, and E. K. Stathopoulou, "Depth cameras on UAVs: A first approach," *Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 42, no. 2W3, pp. 231–236, Feb. 2017, doi: [10.5194/ISPRS-ARCHIVES-XLII-2-W3-231-2017](https://doi.org/10.5194/ISPRS-ARCHIVES-XLII-2-W3-231-2017).
- [61] L. Xu, W. Cheng, K. Guo, L. Han, Y. Liu, and L. Fang, "FlyFusion: Realtime dynamic scene reconstruction using a flying depth camera," *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 1, pp. 68–82, Jan. 2021.
- [62] L. Ufanowicz and R. Sabak, "Unmanned aerial vehicles supporting imagery intelligence using the structured light technology," *Arch. Transp.*, vol. 58, no. 2, pp. 35–45, 2021, doi: [10.5604/01.3001.0014.8796](https://doi.org/10.5604/01.3001.0014.8796).
- [63] F. DadrasJavan, F. Samadzadegan, S. H. Seyed Pourazar, and H. Fazeli, "UAV-based multispectral imagery for fast citrus greening detection," *J. Plant Dis. Protection*, vol. 126, no. 4, pp. 307–318, Jun. 2019, doi: [10.1007/S41348-019-00234-8](https://doi.org/10.1007/S41348-019-00234-8).
- [64] N. Bolourian and A. Hammad, "LiDAR-equipped UAV path planning considering potential locations of defects for bridge inspection," *Automat. Construction*, vol. 117, Sep. 2020, Art. no. 103250, doi: [10.1016/J.AUTCON.2020.103250](https://doi.org/10.1016/J.AUTCON.2020.103250).
- [65] B. Alsadik and F. Remondino, "Flight planning for LiDAR-based UAS mapping applications," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 6, Jun. 2020, Art. no. 378, doi: [10.3390/IJGI9060378](https://doi.org/10.3390/IJGI9060378).

- [66] L. Yoder and S. Scherer, "Autonomous exploration for infrastructure modeling with a micro aerial vehicle," in *Field and Service Robotics*, D. S. Wettergreen and T. D. Barfoot, Eds. Cham, Switzerland: Springer, 2016, pp. 427–440.
- [67] C. Song, Z. Chen, K. Wang, H. Luo, and J. C. P. Cheng, "BIM-supported scan and flight planning for fully autonomous LiDAR-carrying UAVs," *Autom. Construction*, vol. 142, Oct. 2022, Art. no. 104533, doi: [10.1016/J.AUTCON.2022.104533](https://doi.org/10.1016/J.AUTCON.2022.104533).
- [68] F. Yan, E. Xia, Z. Li, and Z. Zhou, "Sampling-based path planning for high-quality aerial 3D reconstruction of urban scenes," *Remote Sens.*, vol. 13, no. 5, pp. 1–23, 2021, doi: [10.3390/rs13050989](https://doi.org/10.3390/rs13050989).
- [69] J. L. Schöberle, E. Zheng, J. M. Frahm, and M. Pollefeys, "Pixelwise view selection for unstructured multi-view stereo," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 501–518, doi: [10.1007/978-3-319-46487-9\\_31](https://doi.org/10.1007/978-3-319-46487-9_31).
- [70] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Auton. Robots*, vol. 34, no. 3, pp. 189–206, 2013, doi: [10.1007/s10514-012-9321-0](https://doi.org/10.1007/s10514-012-9321-0).
- [71] T. Whelan, M. Kaess, H. Johannsson, M. Fallon, J. J. Leonard, and J. McDonald, "Real-time large-scale dense RGB-D SLAM with volumetric fusion," *Int. J. Robot. Res.*, vol. 34, no. 4/5, pp. 598–626, Dec. 2014, doi: [10.1177/0278364914551008](https://doi.org/10.1177/0278364914551008).
- [72] H. Qin et al., "Autonomous exploration and mapping system using heterogeneous UAVs and UGVs in GPS-denied environments," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1339–1350, Feb. 2019.
- [73] M. Popovic, G. Hitz, J. Nieto, I. Sa, R. Siegwart, and E. Galceran, "Online informative path planning for active classification using UAVs," in *Proc.-IEEE Int. Conf. Robot. Automat.*, 2017, pp. 5753–5758, doi: [10.1109/ICRA.2017.7989676](https://doi.org/10.1109/ICRA.2017.7989676).
- [74] J. Valente, J. Del Cerro, A. Barrientos, and D. Sanz, "Aerial coverage optimization in precision agriculture management: A musical harmony inspired approach," *Comput. Electron. Agriculture*, vol. 99, pp. 153–159, Nov. 2013, doi: [10.1016/J.COMPAG.2013.09.008](https://doi.org/10.1016/J.COMPAG.2013.09.008).
- [75] F. Stache, J. Westheider, F. Magistri, M. Popovic, and C. Stachniss, "Adaptive path planning for UAV-based multi-resolution semantic segmentation," in *Proc. Eur. Conf. Mobile Robots*, 2021, pp. 1–6, doi: [10.1109/ECMR50962.2021.9568788](https://doi.org/10.1109/ECMR50962.2021.9568788).
- [76] A. Basiri, V. Mariani, G. Silano, M. Aatif, L. Iannelli, and L. Glielmo, "A survey on the application of path-planning algorithms for multi-rotor UAVs in precision agriculture," *J. Navigation*, vol. 75, pp. 364–383, 2022, doi: [10.1017/S0373463321000825](https://doi.org/10.1017/S0373463321000825).
- [77] G. E. Just, M. E. Pellenz, L. A. De Paula Lima, B. S. Chang, R. D. Souza, and S. Montejo-Sánchez, "UAV path optimization for precision agriculture wireless sensor networks," *Sensors*, vol. 20, no. 21, Oct. 2020, Art. no. 6098, doi: [10.3390/S20216098](https://doi.org/10.3390/S20216098).
- [78] T. Zaenker, C. Smitt, C. McCool, and M. Bennewitz, "Viewpoint planning for fruit size and position estimation," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2021, pp. 3271–3277, doi: [10.1109/IROS51168.2021.9636701](https://doi.org/10.1109/IROS51168.2021.9636701).
- [79] C. Gao et al., "A UAV-based explore-then-exploit system for autonomous indoor facility inspection and scene reconstruction," *Automat. Construction*, vol. 148, Apr. 2023, Art. no. 104753, doi: [10.1016/J.AUTCON.2023.104753](https://doi.org/10.1016/J.AUTCON.2023.104753).
- [80] S. Karam, F. Nex, B. T. Chidura, and N. Kerle, "Microdrone-based indoor mapping with graph SLAM," *Drones*, vol. 6, no. 11, Nov. 2022, Art. no. 352, doi: [10.3390/DRONES6110352](https://doi.org/10.3390/DRONES6110352).
- [81] S. Shen, N. Michael, and V. Kumar, "Autonomous indoor 3D exploration with a micro-aerial vehicle," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2012, pp. 9–15, doi: [10.1109/ICRA.2012.6225146](https://doi.org/10.1109/ICRA.2012.6225146).
- [82] D. Zhu, T. Li, D. Ho, C. Wang, and M. Q. H. Meng, "Deep reinforcement learning supervised autonomous exploration in office environments," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 7548–7555, doi: [10.1109/ICRA.2018.8463213](https://doi.org/10.1109/ICRA.2018.8463213).
- [83] Y. Sun, Q. Huang, D.-Y. Hsiao, L. Guan, and G. Hua, "Learning view selection for 3D scenes," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 14464–14473, [Online]. Available: [https://openaccess.thecvf.com/content/CVPR2021/html/Sun\\_Learning\\_View\\_Selection\\_for\\_3D\\_Scenes\\_CVPR\\_2021\\_paper.html](https://openaccess.thecvf.com/content/CVPR2021/html/Sun_Learning_View_Selection_for_3D_Scenes_CVPR_2021_paper.html)
- [84] R. Almadhoun, T. Taha, L. Seneviratne, and Y. Zweiri, "A survey on multi-robot coverage path planning for model reconstruction and mapping," *SN Appl. Sci.*, vol. 1, no. 8, pp. 1–24, Jul. 2019, doi: [10.1007/S42452-019-0872-Y](https://doi.org/10.1007/S42452-019-0872-Y).
- [85] Z. Shang and Z. Shen, "Flight planning for survey-grade 3D reconstruction of truss bridges," *Remote Sens.*, vol. 14, no. 13, Jul. 2022, Art. no. 3200, doi: [10.3390/RS14133200](https://doi.org/10.3390/RS14133200).
- [86] Y. Liu et al., "Learning reconstructability for drone aerial path planning," Sep. 2022, *arXiv:2209.10174*.
- [87] C. Peng and V. Isler, "Visual coverage path planning for urban environments," *IEEE Robot. Automat. Lett.*, vol. 5, no. 4, pp. 5961–5968, Oct. 2020.
- [88] A. Bircher et al., "Structural inspection path planning via iterative viewpoint resampling with application to aerial robotics," in *Proc. Int. Conf. Robot. Automat.*, 2015, pp. 6423–6430, doi: [10.1109/ICRA.2015.7140101](https://doi.org/10.1109/ICRA.2015.7140101).
- [89] C. S. Tan, R. Mohd-Mokhtar, and M. R. Arshad, "A comprehensive review of coverage path planning in robotics using classical and heuristic algorithms," *IEEE Access*, vol. 9, pp. 119310–119342, 2021.
- [90] E. Galceran and M. Carreras, "A survey on coverage path planning for robotics," *Robot. Auton. Syst.*, vol. 61, no. 12, pp. 1258–1276, Dec. 2013, doi: [10.1016/J.ROBOT.2013.09.004](https://doi.org/10.1016/J.ROBOT.2013.09.004).
- [91] Z. Shang, J. Bradley, and Z. Shen, "A co-optimal coverage path planning method for aerial scanning of complex structures," *Expert Syst. Appl.*, vol. 158, Nov. 2020, Art. no. 113535, doi: [10.1016/J.ESWA.2020.113535](https://doi.org/10.1016/J.ESWA.2020.113535).
- [92] M. Lodel, B. Brito, A. Serra-Gomez, L. Ferranti, R. Babuska, and J. Alonso-Mora, "Where to look next: Learning viewpoint recommendations for informative trajectory planning," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2022, pp. 4466–4472, doi: [10.1109/ICRA46639.2022.9812190](https://doi.org/10.1109/ICRA46639.2022.9812190).
- [93] W. Jing, D. Deng, Z. Xiao, Y. Liu, and K. Shimada, "Coverage path planning using path primitive sampling and primitive coverage graph for visual inspection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2019, pp. 1472–1479, doi: [10.1109/IROS40897.2019.8967969](https://doi.org/10.1109/IROS40897.2019.8967969).
- [94] S. Pan, H. Hu, and H. Wei, "SCVP: Learning one-shot view planning via set covering for unknown object reconstruction," *IEEE Robot. Automat. Lett.*, vol. 7, no. 2, pp. 1463–1470, Apr. 2022.
- [95] C. Papachristos et al., "Autonomous exploration and inspection path planning for aerial robots using the robot operating system," *Stud. Comput. Intell.*, vol. 778, pp. 67–111, 2019, doi: [10.1007/978-3-319-91590-6\\_3](https://doi.org/10.1007/978-3-319-91590-6_3).
- [96] L. Heng, A. Gotovos, A. Krause, and M. Pollefeys, "Efficient visual exploration and coverage with a micro aerial vehicle in unknown environments," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2015, pp. 1071–1078, doi: [10.1109/ICRA.2015.7139309](https://doi.org/10.1109/ICRA.2015.7139309).
- [97] R. Almadhoun, T. Taha, L. Seneviratne, J. Dias, and G. Cai, "A survey on inspecting structures using robotic systems," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 6, pp. 1–18, Nov. 2016, doi: [10.1177/1729881416663664](https://doi.org/10.1177/1729881416663664).
- [98] C. Mostegel, M. Rumpler, F. Fraundorfer, and H. Bischof, "UAV-based autonomous image acquisition with multi-view stereo quality assurance by confidence prediction," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, 2016, pp. 1–10, doi: [10.1109/CVPRW.2016.8](https://doi.org/10.1109/CVPRW.2016.8).
- [99] T. Koch, "Automated and precise 3D building reconstruction using UAVs," Ph.D. thesis, TU Munich, Munich, Germany, 2020.
- [100] A. Batinovic, J. Goricanec, L. Markovic, and S. Bogdan, "Path planning with potential field-based obstacle avoidance in a 3D environment by an unmanned aerial vehicle," in *Proc. Int. Conf. Unmanned Aircr. Syst.*, 2022, pp. 394–401, doi: [10.1109/ICUAS54217.2022.9836159](https://doi.org/10.1109/ICUAS54217.2022.9836159).
- [101] X. Wang, M. C. Gursoy, T. Erpek, and Y. E. Sagduyu, "Learning-based UAV path planning for data collection with integrated collision avoidance," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 16663–16676, Sep. 2022.
- [102] A. Gunawan, H. C. Lau, and P. Vansteenwegen, "Orienteering problem: A survey of recent variants, solution approaches and applications," *Eur. J. Oper. Res.*, vol. 255, no. 2, pp. 315–332, Dec. 2016, doi: [10.1016/J.EJOR.2016.04.059](https://doi.org/10.1016/J.EJOR.2016.04.059).
- [103] P. Vansteenwegen, W. Souffriau, and D. Van Oudheusden, "The orienteering problem: A survey," *Eur. J. Oper. Res.*, vol. 209, no. 1, pp. 1–10, Feb. 2011, doi: [10.1016/J.EJOR.2010.03.045](https://doi.org/10.1016/J.EJOR.2010.03.045).
- [104] B. L. Golden, L. Levy, and R. Vohra, "The orienteering problem," *Nav. Res. Logistics*, vol. 34, no. 3, pp. 307–318, Jun. 1987, doi: [10.1002/1520-6750\(198706\)34:3<307::aid-nav.3240030305>2.0.co;2-1](https://doi.org/10.1002/1520-6750(198706)34:3<307::aid-nav.3240030305>2.0.co;2-1).
- [105] C. Peng and V. Isler, "Adaptive view planning for aerial 3D reconstruction," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 2981–2987, doi: [10.1109/ICRA.2018.8793532](https://doi.org/10.1109/ICRA.2018.8793532).
- [106] J. I. Vasquez-Gomez, E. Lopez-Damian, and L. E. Sucar, "View planning for 3D object reconstruction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 4015–4020, 2009, doi: [10.1109/IROS.2009.5354383](https://doi.org/10.1109/IROS.2009.5354383).
- [107] O. Sharma, N. Arora, and H. Sagar, "Image acquisition for high quality architectural reconstruction," in *Proc. 45th Graph. Interface Conf. Proc. Graph. Interface*, 2019, pp. 1–9, doi: [10.20380/GI2019.18](https://doi.org/10.20380/GI2019.18).

- [108] H. Zhang, Y. Yao, K. Xie, C. W. Fu, H. Zhang, and H. Huang, "Continuous aerial path planning for 3D urban scene reconstruction," *ACM Trans. Graph.*, vol. 40, no. 6, pp. 1–15, Dec. 2021, doi: [10.1145/3478513.3480483](https://doi.org/10.1145/3478513.3480483).
- [109] S. Zhang, C. Liu, and N. Haala, "Three-dimensional path planning of uavs imaging for complete photogrammetric reconstruction," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 5, no. 1, pp. 325–331, 2020, doi: [10.5194/isprs-annals-V-1-2020-325-2020](https://doi.org/10.5194/isprs-annals-V-1-2020-325-2020).
- [110] Z. Shang and Z. Shen, "Topology-based UAV path planning for multi-view stereo 3D reconstruction of complex structures," *Complex Intell. Syst.*, vol. 9, pp. 909–926, 2023, doi: [10.1007/s40747-022-00831-5](https://doi.org/10.1007/s40747-022-00831-5).
- [111] W. Jing, J. Polden, P. Y. Tao, W. Lin, and K. Shimada, "View planning for 3D shape reconstruction of buildings with unmanned aerial vehicles," in *Proc. 14th Int. Conf. Control, Automat., Robot. Vis.*, 2016, pp. 1–6, doi: [10.1109/ICARCV.2016.7838774](https://doi.org/10.1109/ICARCV.2016.7838774).
- [112] G. Tmušić et al., "Current practices in UAS-based environmental monitoring," *Remote Sens.*, vol. 12, no. 6, Mar. 2020, Art. no. 1001, doi: [10.3390/RS12061001](https://doi.org/10.3390/RS12061001).
- [113] I. Singh, "DJI's commercial drone market share dropped dramatically in 2021: Report," 2021, Accessed: Oct. 20, 2022. [Online]. Available: <https://dronedj.com/2021/09/14/droneanalyst-dji-market-share-2021/>
- [114] H. J. Przybilla, "Wer wird denn gleich in die Luft gehen? APPs zur Bildflugplanung für UAV," in *Proc. Schriftenreihe des DVW*, 2022, pp. 17–38.
- [115] Z. Kang, J. Yang, Z. Yang, and S. Cheng, "A review of techniques for 3D reconstruction of indoor environments," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 5, May 2020, Art. no. 330, doi: [10.3390/ijgi9050330](https://doi.org/10.3390/ijgi9050330).
- [116] Agisoft, "Agisoft: Mission planning for complex structures," 2022, Accessed: Feb. 20, 2022. [Online]. Available: <https://agisoft.freshdesk.com/support/solutions/articles/3100157953-mission-planning-for-complex-structures>
- [117] J. J. Leonard and H. F. Durrant-Whyte, "Simultaneous map building and localization for an autonomous mobile robot," in *Proc. IEEE/RSJ Int. Workshop Intell. Robots Syst.*, pp. 1442–1447, Dec. 1991, doi: [10.1109/IROS.1991.174711](https://doi.org/10.1109/IROS.1991.174711).
- [118] C. Estrada, J. Neira, and J. D. Tardós, "Hierarchical SLAM: Real-time accurate mapping of large environments," *IEEE Trans. Robot.*, vol. 21, no. 4, pp. 588–596, Aug. 2005.
- [119] G. Dissanayake, S. Sukkarieh, E. Nebot, and H. Durrant-Whyte, "The aiding of a low-cost strapdown inertial measurement unit using vehicle model constraints for land vehicle applications," *IEEE Trans. Robot. Automat.*, vol. 17, no. 5, pp. 731–747, Oct. 2001.
- [120] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," in *Proc. Nat. Conf. Artif. Intell.*, 2002, pp. 593–598.
- [121] F. Dellaert and M. Kaess, "Square root SAM: Simultaneous localization and mapping via square root information smoothing," *Int. J. Robot. Res.*, vol. 25, no. 12, 2006, pp. 1181–1203, doi: [10.1177/0278364906072768](https://doi.org/10.1177/0278364906072768).
- [122] B. Fang and Z. Zhan, "A visual SLAM method based on point-line fusion in weak-matching scene," *Int. J. Adv. Robot. Syst.*, vol. 17, Mar. 2020, Art. no. 2, doi: [10.1177/1729881420904193](https://doi.org/10.1177/1729881420904193).
- [123] T. Cieslewski, E. Kaufmann, and D. Scaramuzza, "Rapid exploration with multi-rotors: A frontier selection method for high speed flight," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2017, pp. 2135–2142, doi: [10.1109/IROS.2017.8206030](https://doi.org/10.1109/IROS.2017.8206030).
- [124] A. Batinovic, T. Petrovic, A. Ivanovic, F. Petric, and S. Bogdan, "A multi-resolution frontier-based planner for autonomous 3D exploration," *IEEE Robot. Automat. Lett.*, vol. 6, no. 3, pp. 4528–4535, Jul. 2021.
- [125] A. Bircher, M. Kamel, K. Alexis, H. Oleynikova, and R. Siegwart, "Receding horizon path planning for 3D exploration and surface inspection," *Auton. Robots*, vol. 42, no. 2, pp. 291–306, 2018, doi: [10.1007/s10514-016-9610-0](https://doi.org/10.1007/s10514-016-9610-0).
- [126] S. Song and S. Jo, "Surface-based exploration for autonomous 3D modeling," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 4319–4326, doi: [10.1109/ICRA.2018.8460862](https://doi.org/10.1109/ICRA.2018.8460862).
- [127] Y. Kompis, L. Bartolomei, R. Mascaro, L. Teixeira, and M. Chli, "Informed sampling exploration path planner for 3D reconstruction of large scenes," *IEEE Robot. Automat. Lett.*, vol. 6, no. 4, pp. 7893–7900, Oct. 2021.
- [128] A. Knapitsch, J. Park, Q.-Y. Zhou, and V. Koltun, "Tanks and temples: Benchmarking large-scale scene reconstruction," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–13, Jul. 2017.
- [129] B. Yamauchi, "Frontier-based approach for autonomous exploration," in *Proc. IEEE Int. Symp. Comput. Intell. Robot. Automat.*, 1997, pp. 146–151, doi: [10.1109/cira.1997.613851](https://doi.org/10.1109/cira.1997.613851).
- [130] H. P. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1985, pp. 116–121, doi: [10.1109/ROBOT.1985.1087316](https://doi.org/10.1109/ROBOT.1985.1087316).
- [131] H. Umari and S. Mukhopadhyay, "Autonomous robotic exploration based on multiple rapidly-exploring randomized trees," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2017, pp. 1396–1402, doi: [10.1109/IROS.2017.8202319](https://doi.org/10.1109/IROS.2017.8202319).
- [132] S. Obwald, M. Bennewitz, W. Burgard, and C. Stachniss, "Speeding-up robot exploration by exploiting background information," *IEEE Robot. Automat. Lett.*, vol. 1, no. 2, pp. 716–723, Jul. 2016.
- [133] A. Dai, S. Papathodorou, N. Funk, D. Tzoumanikas, and S. Leutenegger, "Fast frontier-based information-driven autonomous exploration with an MAV," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 9570–9576, doi: [10.1109/ICRA40945.2020.9196707](https://doi.org/10.1109/ICRA40945.2020.9196707).
- [134] R. Shade and P. Newman, "Choosing where to go: Complete 3D exploration with Stereo," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 2806–2811, doi: [10.1109/ICRA.2011.5980121](https://doi.org/10.1109/ICRA.2011.5980121).
- [135] Z. Meng et al., "A two-stage optimized next-view planning framework for 3-D unknown environment exploration, and structural reconstruction," *IEEE Robot. Automat. Lett.*, vol. 2, no. 3, pp. 1680–1687, Jul. 2017.
- [136] J. I. Vasquez-Gomez, L. E. Sucar, and R. Murrieta-Cid, "View/state planning for three-dimensional object reconstruction under uncertainty," *Auton. Robots*, vol. 41, no. 1, pp. 89–109, 2017, doi: [10.1007/s10514-015-9531-3](https://doi.org/10.1007/s10514-015-9531-3).
- [137] J. Delmerico, S. Isler, R. Sabzevari, and D. Scaramuzza, "A comparison of volumetric information gain metrics for active 3D object reconstruction," *Auton. Robots*, vol. 42, no. 2, pp. 197–208, 2018, doi: [10.1007/s10514-017-9634-0](https://doi.org/10.1007/s10514-017-9634-0).
- [138] J. Daudelin and M. Campbell, "An adaptable, probabilistic, next-best view algorithm for reconstruction of unknown 3-D objects," *IEEE Robot. Automat. Lett.*, vol. 2, no. 3, pp. 1540–1547, Jul. 2017.
- [139] A. Bircher, M. Kamel, K. Alexis, H. Oleynikova, and R. Siegwart, "Receding horizon next-best-view planner for 3D exploration," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2016, pp. 1462–1468, doi: [10.1109/ICRA.2016.7487281](https://doi.org/10.1109/ICRA.2016.7487281).
- [140] C. Papachristos, F. Mascarich, S. Khattak, T. Dang, and K. Alexis, "Localization uncertainty-aware autonomous exploration and mapping with aerial robots using receding horizon path-planning," *Auton. Robots*, vol. 43, no. 8, pp. 2131–2161, 2019, doi: [10.1007/s10514-019-09864-1](https://doi.org/10.1007/s10514-019-09864-1).
- [141] M. Dharmadhikari et al., "Motion primitives-based path planning for fast and agile exploration using aerial robots," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 179–185, doi: [10.1109/ICRA40945.2020.9196964](https://doi.org/10.1109/ICRA40945.2020.9196964).
- [142] A. Batinovic, A. Ivanovic, T. Petrovic, and S. Bogdan, "A shadowcasting-based next-best-view planner for autonomous 3D exploration," *IEEE Robot. Automat. Lett.*, vol. 7, no. 2, pp. 2969–2976, Apr. 2022.
- [143] S. Song and S. Jo, "Online inspection path planning for autonomous 3D modeling using a micro-aerial vehicle," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2017, pp. 6217–6224, doi: [10.1109/ICRA.2017.7989737](https://doi.org/10.1109/ICRA.2017.7989737).
- [144] M. Kazhdan and H. Hoppe, "Screened Poisson surface reconstruction," *ACM Trans. Graph.*, vol. 32, no. 3, pp. 1–13, 2013, doi: [10.1145/248728.2487237](https://doi.org/10.1145/248728.2487237).
- [145] T. Whelan, R. F. Salas-Moreno, B. Glocker, A. J. Davison, and S. Leutenegger, "ElasticFusion: Real-time dense SLAM and light source estimation," *Int. J. Robot. Res.*, vol. 35, no. 14, pp. 1697–1716, 2016, doi: [10.1177/0278364916669237](https://doi.org/10.1177/0278364916669237).
- [146] R. A. Newcombe et al., "KinectFusion: Real-time dense surface mapping and tracking," in *Proc. 10th IEEE Int. Symp. Mixed Augmented Reality*, 2011, pp. 127–136, doi: [10.1109/ISMAR.2011.6092378](https://doi.org/10.1109/ISMAR.2011.6092378).
- [147] S. Y. Chen and Y. F. Li, "Vision sensor planning for 3-D model acquisition," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 35, no. 5, pp. 894–904, Oct. 2005.
- [148] I. D. Lee, J. H. Seo, Y. M. Kim, J. Choi, S. Han, and B. Yoo, "Automatic pose generation for robotic 3-D scanning of mechanical parts," *IEEE Trans. Robot.*, vol. 36, no. 4, pp. 1219–1238, Aug. 2020.
- [149] R. Monica and J. Aleotti, "Contour-based next-best view planning from point cloud segmentation of unknown objects," *Auton. Robots*, vol. 42, no. 2, pp. 443–458, 2018, doi: [10.1007/s10514-017-9618-0](https://doi.org/10.1007/s10514-017-9618-0).
- [150] L. Schmid, M. Pantic, R. Khanna, L. Ott, R. Siegwart, and J. Nieto, "An efficient sampling-based method for online informative path planning in unknown environments," *IEEE Robot. Automat. Lett.*, vol. 5, no. 2, pp. 1500–1507, Apr. 2020.
- [151] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, and J. Nieto, "VoxBlox: Incremental 3D Euclidean signed distance fields for on-board MAV planning," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2017, pp. 1366–1373, doi: [10.1109/IROS.2017.8202319](https://doi.org/10.1109/IROS.2017.8202319).

- [152] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Trans. Robot.*, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.
- [153] M. Pizzoli, C. Forster, and D. Scaramuzza, "REMODE: Probabilistic, monocular dense reconstruction in real time," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2014, pp. 2609–2616, doi: [10.1109/ICRA.2014.6907233](https://doi.org/10.1109/ICRA.2014.6907233).
- [154] X. Gu, Z. Fan, S. Zhu, Z. Dai, F. Tan, and P. Tan, "Cascade cost volume for high-resolution multi-view stereo and stereo matching," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 2492–2501, doi: [10.1109/CVPR42600.2020.00257](https://doi.org/10.1109/CVPR42600.2020.00257).
- [155] B. Hepp, D. Dey, S. N. Sinha, A. Kapoor, N. Joshi, and O. Hilliges, "Learn-to-score: Efficient 3D scene exploration by predicting view utility," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 455–472, doi: [10.1007/978-3-030-01267-0\\_27](https://doi.org/10.1007/978-3-030-01267-0_27).
- [156] B. Charrow et al., "Information-theoretic planning with trajectory optimization for dense 3D mapping," *Robot. Sci. Syst.*, vol. 11, 2015, doi: [10.15607/RSS.2015.XI.003](https://doi.org/10.15607/RSS.2015.XI.003).
- [157] B. Charrow, S. Liu, V. Kumar, and N. Michael, "Information-theoretic mapping using Cauchy-Schwarz quadratic mutual information," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2015, pp. 4791–4798, doi: [10.1109/ICRA.2015.7139865](https://doi.org/10.1109/ICRA.2015.7139865).
- [158] C. Wang, H. Ma, W. Chen, L. Liu, and M. Q. H. Meng, "Efficient autonomous exploration with incrementally built topological map in 3-D environments," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9853–9865, Dec. 2020, doi: [10.1109/TIM.2020.3001816](https://doi.org/10.1109/TIM.2020.3001816).
- [159] D. Peralta, J. Casamiro, A. M. Nilles, J. A. Aguilar, R. Atienza, and R. Cajote, "Next-best view policy for 3D reconstruction," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 558–573, doi: [10.1007/978-3-030-66823-5\\_33](https://doi.org/10.1007/978-3-030-66823-5_33).
- [160] R. Zeng, W. Zhao, and Y. J. Liu, "PC-NBV: A point cloud based deep network for efficient next best view planning," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2020, pp. 7050–7057, doi: [10.1109/IROS5743.2020.9340916](https://doi.org/10.1109/IROS5743.2020.9340916).
- [161] X. Zhou, Z. Yi, Y. Liu, K. Huang, and H. Huang, "Survey on path and view planning for UAVs," *Virtual Reality Intell. Hardware*, vol. 2, no. 1, pp. 56–69, Feb. 2020, doi: [10.1016/j.vrih.2019.12.004](https://doi.org/10.1016/j.vrih.2019.12.004).
- [162] Y. Liu, Y. Yixuan, and M. Liu, "Ground-aware monocular 3D object detection for autonomous driving," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 919–926, Apr. 2021.
- [163] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 431–440, doi: [10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965).
- [164] F. Rottensteiner, G. Sohn, M. Gerke, J. D. Wegner, U. Breitkopf, and J. Jung, "Results of the ISPRS benchmark on urban object detection and 3D building reconstruction," *ISPRS J. Photogrammetry Remote Sens.*, vol. 93, pp. 256–271, Jul. 2014, doi: [10.1016/J.IJSPRSJPRS.2013.10.004](https://doi.org/10.1016/J.IJSPRSJPRS.2013.10.004).
- [165] C. Chekuri and M. Pál, "A recursive greedy algorithm for walks in directed graphs," in *Proc. 46th Annu. IEEE Symp. Foundations Comput. Sci.*, 2005, vol. 2005, pp. 245–253, doi: [10.1109/SFCS.2005.9](https://doi.org/10.1109/SFCS.2005.9).
- [166] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *Int. J. Robot. Res.*, vol. 30, no. 7, pp. 846–894, Jun. 2011, doi: [10.1177/0278364911406761](https://doi.org/10.1177/0278364911406761).
- [167] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2980–2988, doi: [10.1109/ICCV.2017.322](https://doi.org/10.1109/ICCV.2017.322).
- [168] R. Huang, D. Zou, R. Vaughan, and P. Tan, "Active image-based modeling with a toy drone," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 6124–6131, doi: [10.1109/ICRA.2018.8460673](https://doi.org/10.1109/ICRA.2018.8460673).
- [169] P. Debus and V. Rodehorst, "Evaluation of 3D uas flight path planning algorithms," in *Proc. 24th ISPRS Congr.*, 2021, pp. 157–164, doi: [10.5194/isprs-archives-xliii-b1-2021-157-2021](https://doi.org/10.5194/isprs-archives-xliii-b1-2021-157-2021).
- [170] M. Maboudi, A. Elbillehy, Y. Ghassoun, and M. Gerke, "Improving the internal accuracy of UAV-image blocks using local low-altitude flights and scale-bars," in *Proc. 24th ISPRS Congr.*, 2021, pp. 183–188, doi: [10.5194/isprs-archives-XLIII-B1-2021-183-2021](https://doi.org/10.5194/isprs-archives-XLIII-B1-2021-183-2021).
- [171] P. Burdziakowski, K. Bobrowska, E. Meroño De Larriba, F. Javier, and M. Carrascosa, "UAV photogrammetry under poor lighting conditions—Accuracy considerations," *Sensors*, vol. 21, no. 10, May 2021, Art. no. 3531, doi: [10.3390/S21103531](https://doi.org/10.3390/S21103531).
- [172] Q. Li, H. Huang, W. Yu, and S. Jiang, "Optimized views photogrammetry: Precision analysis and a large-scale case study in Qingdao," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 1144–1159, Jan. 2023.
- [173] M. Maboudi, D. Bánádi, and M. Gerke, "Investigation of geometric performance of an indoor mobile mapping system," in *Proc. 24th ISPRS Congr.*, 2018, pp. 637–642, doi: [10.5194/isprs-archives-XLII-2-637-2018](https://doi.org/10.5194/isprs-archives-XLII-2-637-2018).
- [174] M. Tsagris, C. Beneki, and H. Hassani, "On the folded normal distribution," *Mathematics*, vol. 2, no. 1, pp. 12–28, Feb. 2014, doi: [10.3390/MATH2010012](https://doi.org/10.3390/MATH2010012).
- [175] M. Bouziani, H. Chaaba, and M. Ettarid, "Evaluation of 3D building model using terrestrial laser scanning and drone photogrammetry," in *Proc. 24th ISPRS Congr.*, 2021, pp. 39–42, doi: [10.5194/isprs-archives-XLVI-4-W4-2021-39-2021](https://doi.org/10.5194/isprs-archives-XLVI-4-W4-2021-39-2021).
- [176] W. Qiu et al., "UnrealCV: Virtual worlds for computer vision," in *Proc. ACM Multimedia Conf.*, 2017, pp. 1221–1224, doi: [10.1145/3123266.3129396](https://doi.org/10.1145/3123266.3129396).
- [177] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "AirSim: High-fidelity visual and physical simulation for autonomous vehicles," in *Proc. Field Service Robot. Conf.*, 2017, vol. 5, pp. 621–635, doi: [10.48550/arxiv.1705.05065](https://doi.org/10.48550/arxiv.1705.05065).
- [178] S. A. Munoz, "Optimized 3D reconstruction for infrastructure inspection with automated structure from motion and machine learning methods learning methods," MS thesis, Brigham Young Univ., Provo, UT, USA, 2020.
- [179] Blender\_Online\_Community, "Blender - A 3D modelling and rendering package," Stichting Blender Foundation, Amsterdam, The Netherlands, 2018. [Online]. Available: <http://www.blender.org>
- [180] C. Liu, S. Zhang, and A. Akbar, "Ground feature oriented path planning for unmanned aerial vehicle mapping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 4, pp. 1175–1187, Apr. 2019.
- [181] R. A. Martin, I. Rojas, K. Franke, and J. D. Hedengren, "Evolutionary view planning for optimized UAV terrain modeling in a simulated environment," *Remote Sens.*, vol. 8, no. 1, Dec. 2015, Art. no. 26, doi: [10.3390/RS8010026](https://doi.org/10.3390/RS8010026).
- [182] J. Nakama, R. Parada, J. P. Matos-Carvalho, F. Azevedo, D. Pedro, and L. Campos, "Autonomous environment generator for UAV-based simulation," *Appl. Sci.*, vol. 11, no. 5, Mar. 2021, Art. no. 2185, doi: [10.3390/app11052185](https://doi.org/10.3390/app11052185).
- [183] A. I. Hentati, L. Krichen, M. Fourati, and L. C. Fourati, "Simulation tools, environments and frameworks for UAV systems performance analysis," in *Proc. 14th Int. Wirel. Commun. Mob. Comput. Conf.*, 2018, pp. 1495–1500, doi: [10.1109/IWCMC.2018.8450505](https://doi.org/10.1109/IWCMC.2018.8450505).
- [184] A. Kim and R. M. Eustice, "Active visual SLAM for robotic area coverage: Theory and experiment," *Int. J. Robot. Res.*, vol. 34, no. 4/5, 2015, doi: [10.1177/0278364914547893](https://doi.org/10.1177/0278364914547893).
- [185] E. M. Lee, J. Choi, H. Lim, and H. Myung, "REAL: Rapid exploration with active loop-closing toward large-scale 3D mapping using UAVs," in *Proc. IEEE Int. Conf. Intell. Robot. Syst.*, 2021, pp. 4194–4198, doi: [10.1109/IROS51168.2021.9636611](https://doi.org/10.1109/IROS51168.2021.9636611).
- [186] Y. Zhang, B. Zhou, L. Wang, and S. Shen, "Exploration with global consistency using real-time re-integration and active loop closure," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2022, pp. 9682–9688, doi: [10.1109/ICRA46639.2022.9811892](https://doi.org/10.1109/ICRA46639.2022.9811892).
- [187] W. He, X. Qi, and L. Liu, "A novel hybrid particle swarm optimization for multi-UAV cooperate path planning," *Appl. Intell.*, vol. 51, no. 10, pp. 7350–7364, Oct. 2021, doi: [10.1007/S10489-020-02082-8](https://doi.org/10.1007/S10489-020-02082-8).
- [188] A. Tullu, B. Endale, A. Wondesen, and H.-Y. Hwang, "Machine learning approach to real-time 3D path planning for autonomous navigation of unmanned aerial vehicle," *Appl. Sci.*, vol. 11, no. 10, May 2021, Art. no. 4706, doi: [10.3390/APP1104706](https://doi.org/10.3390/APP1104706).
- [189] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," Apr. 2018, doi: [10.48550/arxiv.1804.02767](https://doi.org/10.48550/arxiv.1804.02767).
- [190] L. Lin, Y. Liu, Y. Hu, X. Yan, K. Xie, and H. Huang, "Capturing, reconstructing, and simulating: The UrbanScene3D dataset," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 93–109, doi: [10.1007/978-3-031-20074-8\\_6/FIGURES/9](https://doi.org/10.1007/978-3-031-20074-8_6/FIGURES/9).
- [191] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," in *Proc. 4th Int. Conf. Learn. Representations*, 2016, pp. 168–174.
- [192] T. Lei, C. Luo, G. E. Jan, and Z. Bi, "Deep learning-based complete coverage path planning with re-join and obstacle fusion paradigm," *Front. Robot. AI*, vol. 9, pp. 1–14, 2022, doi: [10.3389/frobt.2022.843816](https://doi.org/10.3389/frobt.2022.843816).
- [193] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
- [194] Y. Liu, K. Xie, and H. Huang, "VGF-Net: Visual-geometric fusion learning for simultaneous drone navigation and height mapping," *Graph. Models.*, vol. 116, Jul. 2021, Art. no. 101108, doi: [10.1016/J.GMOD.2021.101108](https://doi.org/10.1016/J.GMOD.2021.101108).
- [195] R. Xie, Z. Meng, L. Wang, H. Li, K. Wang, and Z. Wu, "Unmanned aerial vehicle path planning algorithm based on deep reinforcement learning in large-scale and dynamic environments," *IEEE Access*, vol. 9, pp. 24884–24900, 2021.

- [196] W. Zhang, S. Zhang, F. Wu, and Y. Wang, "Path planning of UAV based on improved adaptive grey wolf optimization algorithm," *IEEE Access*, vol. 9, pp. 89400–89411, 2021.
- [197] Y. V. Pehlivanoglu and P. Pehlivanoglu, "An enhanced genetic algorithm for path planning of autonomous UAV in target coverage problems," *Appl. Soft Comput.*, vol. 112, 2021, Art. no. 107796, doi: <https://doi.org/10.1016/j.asoc.2021.107796>.
- [198] A. Sonmez, E. Kocyigit, and E. Kugu, "Optimal path planning for UAVs using genetic algorithm," in *Proc. Int. Conf. Unmanned Aircr. Syst.*, 2015, pp. 50–55, doi: [10.1109/ICUAS.2015.7152274](https://doi.org/10.1109/ICUAS.2015.7152274).
- [199] X. Zhou, F. Gao, X. Fang, and Z. Lan, "Improved bat algorithm for UAV path planning in three-dimensional space," *IEEE Access*, vol. 9, pp. 20100–20116, 2021.
- [200] Y. Zhao, Z. Zheng, and Y. Liu, "Survey on computational-intelligence-based UAV path planning," *Knowl.-Based Syst.*, vol. 158, pp. 54–64, Oct. 2018, doi: [10.1016/J.KNOSYS.2018.05.033](https://doi.org/10.1016/J.KNOSYS.2018.05.033).
- [201] S. Aggarwal and N. Kumar, "Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges," *Comput. Commun.*, vol. 149, pp. 270–299, Jan. 2020, doi: [10.1016/J.COMCOM.2019.10.014](https://doi.org/10.1016/J.COMCOM.2019.10.014).
- [202] R. Zhang, S. Li, Y. Ding, X. Qin, and Q. Xia, "UAV path planning algorithm based on improved harris hawks optimization," *Sensors*, vol. 22, no. 14, Jul. 2022, Art. no. 5232, doi: [10.3390/S22145232](https://doi.org/10.3390/S22145232).



**Mehdi Maboudi** received the Ph.D. degree in photogrammetry from the University of Tehran, Tehran, Iran, in 2017.

He is a Lecturer and Senior Researcher with the Technical University of Braunschweig, Braunschweig, Germany. Since 2022, he has been the Chair of ISPRS ICWG II/Ib on "Digital Construction: Reality Capture, Automated Inspection, and Integration to BIM." His research interests include close-range photogrammetry, UAV-based structural health monitoring, reality capture, inspection and object tracking in digital construction, indoor mobile mapping, and 3-D indoor modelling.



**MohammadReza Homaei** received the B.Sc. degree in surveying and geomatics engineering, and the M.Sc. degree in photogrammetry, both from the University of Tehran, Tehran, Iran, in 2015 and 2018, respectively. His M.Sc. thesis was about selecting optimum images from a dataset for increasing the speed of the 3-D reconstruction under the supervision of Dr. Mohammad Saadatseresht.

His research interests include close-range photogrammetry, photogrammetric network design, UAV view-point planning, indoor mobile mapping, and 3-D indoor modelling.



**Soohwan Song** received the B.S. degree in information and communication engineering from Dongguk University, Seoul, South Korea, in 2013, and the M.S. and Ph.D. degrees in computer science from KAIST, Daejeon, South Korea, in 2015 and 2020, respectively.

He worked as a Postdoctoral Researcher with Neuro-Machine Augmented Intelligence Laboratory, KAIST. Since March 2021, he has been with the Mobility Robot Research Division, ETRI, as a Senior Researcher. His research interests include robotics, motion planning, and computer vision.



**Shirin Malih** received the B.Sc. degree in surveying engineering from the University of Isfahan, Isfahan, Iran, and the master's degree in photogrammetry engineering from K.N.Toosi University of Technology, Tehran, Iran, in 2008, and the Ph.D. degree in Photogrammetry, from K.N.Toosi University of Technology and HFT Stuttgart, in 2019.

Her Ph.D. on large-scale building reconstruction using UAV Photogrammetry and big point clouds processing was funded by young scientist award of DAAD. As a Postdoc Research Assistant and Associate she has collaborated with Universities of Tehran, New York, and Edinburgh. Her current research interests include digital twining of construction and BIM.



**Mohammad Saadatseresht** received the B.S. degree in surveying and geomatics engineering, the M.Sc. degree in photogrammetry, and the Ph.D. degree in industrial photogrammetry, all from the University of Tehran, Tehran, Iran, in 1996, 1998, and 2005, respectively. His M.Sc. thesis was about building extraction from large scale aerial digital images and his Ph.D. thesis was about automation of vision metrology network design which has been done in cooperation with supervision of Prof. Clive Fraser in Melbourne University in 2005.

Since 2006, he has been working as an Assistant/Associate Professor with the School of Surveying and Geospatial Engineering, University of Tehran, and has published more than 300 journal and conference papers in close range and aerial photogrammetry, remote sensing, mobile mapping systems and computer vision subjects. He has been a Supervisor of several dozen graduated students these years.



**Markus Gerke** received the diploma degree in surveying and the Ph.D. degree in photogrammetry from Leibniz University, Hannover, Germany, in 2000 and 2006, respectively. The topic of his thesis was automated road data verification from remote sensing images.

From 2007 to 2016, he was an Assistant Professor with the University of Twente, Enschede, The Netherlands, faculty ITC, where he worked on image geometry and semantics. Since 2017 he has been a Full Professor with the Technical University of Braunschweig and heading the Institute of Geodesy and Photogrammetry. His research interests include UAV-based photogrammetry, close-range sensing for quality inspection and reality capture in the construction engineering science and remote sensing.