Year: 2020

Automatic 3D Reconstruction of Structured Indoor Environments

Pintore, Giovanni ; Mura, Claudio ; Ganovelli, Fabio ; Fuentes-Perez, Lizeth ; Pajarola, R ; Gobbetti, Enrico

Abstract: Creating high-level structured 3D models of real-world indoor scenes from captured data is a fundamental task which has important applications in many fields. Given the complexity and variability of interior environments and the need to cope with noisy and partial captured data, many open research problems remain, despite the substantial progress made in the past decade. In this tutorial, we provide an up-to-date integrative view of the field, bridging complementary views coming from computer graphics and computer vision. After providing a characterization of input sources, we define the structure of output models and the priors exploited to bridge the gap between imperfect sources and desired output. We then identify and discuss the main components of a structured reconstruction pipeline, and review how they are combined in scalable solutions working at the building level. We finally point out relevant research issues and analyze research trends.

DOI: https://doi.org/10.1145/3388769.3407469

Originally published at:
Pintore, Giovanni; Mura, Claudio; Ganovelli, Fabio; Fuentes-Perez, Lizeth; Pajarola, R; Gobbetti, Enrico (2020). Automatic 3D Reconstruction of Structured Indoor Environments. In: ACM SIGGRAPH Courses, Los Angeles, 17 August 2020 - 28 August 2020, ACM Digital Library.
DOI: https://doi.org/10.1145/3388769.3407469

Automatic 3D Reconstruction of Structured Indoor Environments

Tutorial Notes

Giovanni Pintore
CRS4, Italy

Claudio Mura
University of Zurich, Switzerland

Fabio Ganovelli
ISTI-CNR, Italy

Lizeth Fuentes-Perez
University of Zurich, Switzerland

Renato Pajarola
University of Zurich, Switzerland

Enrico Gobbetti
CRS4, Italy

ABSTRACT

Creating high-level structured 3D models of real-world indoor scenes from captured data is a fundamental task which has important applications in many fields. Given the complexity and variability of interior environments and the need to cope with noisy and partial captured data, many open research problems remain, despite the substantial progress made in the past decade. In this tutorial, we provide an up-to-date integrative view of the field, bridging complementary views coming from computer graphics and computer vision. After providing a characterization of input sources, we define the structure of output models and the priors exploited to bridge the gap between imperfect sources and desired output. We then identify and discuss the main components of a structured reconstruction pipeline, and review how they are combined in scalable solutions working at the building level. We finally point out relevant research issues and analyze research trends.

CCS CONCEPTS

• **Computing methodologies** → **Computer graphics**; *Shape modeling*; **Computer vision**; *Computer vision problems*; *Shape inference*; *Reconstruction*; • **Applied computing** → *Computer-aided design*.

KEYWORDS

indoor reconstruction, indoor scanning, structured reconstruction

ACM Reference Format:
Giovanni Pintore, Claudio Mura, Fabio Ganovelli, Lizeth Fuentes-Perez, Renato Pajarola, and Enrico Gobbetti. 2020. Automatic 3D Reconstruction of Structured Indoor Environments: Tutorial Notes. In *Proceedings of SIGGRAPH 2020 Courses (SIGGRAPH2020 Courses).* ACM, New York, NY, USA, 10 pages.

1 FORMAT AND PRE-REQUISITES

Format. Long (3 hours).
Necessary background. The tutorial is at the intermediate level. Basic computer-vision and graphics background is a pre-requisite.
Intended audience. The target audience includes researchers in geometric modeling, as well as practitioners in the relevant application fields. Researchers will find a structured overview of the field, which organizes the various problems and existing solutions, classifies the existing literature, and indicates challenging open problems. Domain experts will, in turn, find a presentation of the areas where automated methods are already mature enough to be ported into practice, as well as an analysis of the kind of indoor environments that still pose major challenges.
Previous presentations. This tutorial builds on an extensive state-of-the-art survey that has been presented at Eurographics 2020 [Pintore et al. 2019b]. The Eurographics presentation version was a condensed STAR aimed at experts, and focused on the presentation of the literature survey. This course significantly extends it with tutorial-style presentations to accommodate a much more varied audience and to make the content more self-contained.

SIGGRAPH2020 Courses, July 2020, Washington, DC, USA
2020.

2 COURSE DESCRIPTION

The automated reconstruction of 3D models from acquired data, be it images or 3D point clouds, has been one of the central topics in computer graphics and computer vision for decades. This field is now thriving, as a result of complementing scientific, technological and market trends. In particular, in recent years, the widespread availability and proliferation of high-fidelity visual/3D sensors (e.g., smartphones, commodity and professional stereo cameras and depth sensors, panoramic cameras, low-cost and high-throughput scanners) has been matched with increasingly cost-effective options for large data processing (e.g., cloud and GPU-accelerated computation), as well as with novel means of visual exploration, from mobile phones to immersive personal displays.

In this context, one of the rapidly emerging sub-fields is concerned with the automatic reconstruction of indoor environments. That is, a 3D representation of an interior scene must be inferred from a collection of measurements that sample its shape and/or appearance, exploiting and/or combining sensing technologies ranging from passive methods, such as single- and multi-view image capturing, to active methods, such as infrared or time-of-flight cameras, optical laser-based range scanners, structured-light scanners, and LiDAR scanners [Berger et al. 2017]. Based on the raw data acquired by these devices, many *general* surface reconstruction methods focus on producing accurate and dense 3D models that faithfully replicate even the smallest geometry and appearance details. In this sense, their main goal is to provide the most accurate representation possible of all the surfaces that compose the input scene, disregarding its structure and semantics or possibly only exploiting them to maximize the fidelity of the output surface model. A number of more *specialized* indoor reconstruction solutions focus, instead, on abstracting simplified high-level structured models that optimize certain application-dependent characteristics [Ikehata et al. 2015].

The focus on high-level structured models is motivated by several reasons. First of all, their availability is necessary in many fields. For example, applications such as the generation or revision of building information models (BIM) require, at least, the determination of the bare architectural structure [Mura et al. 2014b; Turner et al. 2015]. On the other hand, information on the interior clutter, in terms of 3D footprint of major indoor objects, is necessary in many other use cases, such as guidance, energy management, security, evacuation planning, location awareness or routing [Ikehata et al. 2015]. Even when the goal is solely for visualization, structured simplified models need to be extracted as a fundamental component of a renderable model. This is because narrow spaces, windows, non-cooperative materials, and abundant clutter make the transition from the acquisition of indoor scenes to their modeling and rendering a very difficult problem. Thus, applying standard dense surface reconstruction approaches, which optimize for completeness, resolution and accuracy, leads to unsatisfactory results.

Automatic 3D reconstruction and modeling of indoor scenes, has thus attracted a lot of research in recent years, making it an emerging well-defined topic. In particular, the focus has been on developing specialized techniques for very common and very structured multi-room environments, such as residential, office, or public buildings, which have a substantial impact on architecture, civil engineering, digital mapping, urban geography, real estate, and more [Ikehata et al. 2015]. In this context, the fundamental tasks are the discovery of structural elements, such as rooms, walls, doors, and indoor objects, and their combination in a consistent structured 3D shape and visual representation. The research community working on these problems appears, however, fragmented, and many different vertical solutions have been proposed for the various motivating applications. In this course, we provide an up-to-date integrative view of the field, bridging complementary views coming from computer graphics and computer vision.

3 COURSE RATIONALE

Reconstruction of visual and geometric models from images or point clouds is a very broad topic in computer graphics and computer vision. This course focuses on the specific problems and solutions relating to the reconstruction of *structured 3D indoor models*, that is rapidly emerging as a very important and challenging problem, with specific solutions and very important applications. Thus, we complement existing courses and surveys focusing on reconstructing detailed surfaces from dense high-quality data or on assigning semantic to existing geometry, by covering the extraction of an *approximate structured geometry* connected to a *visual representation* from sparse and incomplete measurements.

The tutorial content is based on a recent survey of the state-of-the-art that we have published in Computer Graphics Forum [Pintore et al. 2019b], and presented at the 2020 Eurographics conference. We refer the audience to that STAR for an in-depth presentation of the concept and a detailed reasoned bibliography.

A general coverage of methods for 3D surface reconstruction and primitive identification is available in recent surveys [Berger et al. 2017; Kaiser et al. 2019], and we will build on them for the definition of general problems and solutions. In the same spirit, we do not specifically cover interactive or online approaches; those interested in online reconstruction can find more detail on the topic in the survey by Zollhöfer et al. [Zollhöfer et al. 2018]. We also will refer the audience to an established state-of-the-art report on urban reconstruction [Musialski et al. 2013] for an overview of the companion problem of reconstructing (from the outside) 3D geometric models of urban areas, individual buildings, façades, and further architectural details.

The techniques surveyed in this course also have an overlap with the domains of Scan-to-BIM or Inverse-CAD, where the goal is the automatic reconstruction of full (volumetric) information models from measurement data. However, the overlap is only partial, since we do not cover the assignment of full semantic information and/or the satisfaction of engineering construction rules, and Scan-to-BIM generally does not cover the generation of visual representations, which is necessary for rendering. Moreover, most Scan-to-BIM solutions are currently targeting (dense) point cloud data, while we cover solutions starting from a variety of input sources. It should be noted that, obviously, relations do exist, and many of the solutions surveyed here can serve as good building blocks to tackle the full Scan-to-BIM problem. We will refer the audience to established surveys in the Scan-to-BIM area for a review of related techniques based on point-cloud data [Pătrăucean et al. 2015; Tang et al. 2010; Volk et al. 2014], general computer vision [Fathi et al. 2015], and RGB-D data [Chen et al. 2015a].

In addition, commodity mobile platforms are emerging as a very common solutions both for capture and for exploration of mobile environments. On this specific topics, we refer the audience to two recent tutorials on the subject, which also contain sections devoted to indoor environments [Agus et al. 2017a,b].

4 DETAILED OUTLINE

The course will be organizes in two sessions of 1.5 hours. After providing a general overview of the subject (Session 1.1), we will discuss shape and color sources generated by indoor mapping devices and describe several open datasets available for research purposes (Session 1.2). We will then provide an abstract characterization of the typical structured indoor models, and of the main problems that need to be solved to create such models from imperfect input data, identifying the specialized priors exploited to address significantly challenging imperfections in visual and geometric input (Session 1.3). The various solutions proposed in the literature, and their combination into global reconstruction pipelines will be then analyzed by providing a general overview, pointing out the various solutions proposed in the literature, and discussing their pros and cons. Session 1.4 will be dedicated to room segmentation, while Session 1.5 will cover boundary surface reconstruction from dense 3D data. After a break, we will continue with a presentation of boundary surface reconstruction from images and/or sparse 3D data (Session 2.1), object detection and reconstruction (Session 2.2), final model assembly (Session 2.3), and visual representation generation (Session 2.4). We will finally point out relevant research issues and analyze research trends (Session 2.5).

SESSION 1.1: Opening and introduction

In the introductory session, we will define the topic of structured indoor reconstruction and point out to the many applications of it. We will then provide an outline of the rest of the presentation.

SESSION 1.2: Data capture and representation

Indoor reconstruction starts from measured data obtained by surveying the indoor environment. Many options exist for performing capture, ranging from very low-cost commodity solutions to professional devices and systems. In this session, we first provide a characterization of the various input sources and then provide a link to the main public domain datasets available for research purposes.

Input data sources. Indoor mapping is required for a wide variety of applications, and an enormous range of 3D acquisition devices have been proposed over the last decades. From LiDAR to portable mobile mappers, these sensors gather shape and/or color information in an effective, often domain-specific, way [Lehtola et al. 2017; Xiong et al. 2013]. In addition, many general-purpose commodity solutions, e.g., based on smartphones and cameras, have also been exploited for that purpose [Pintore et al. 2014; Sankar and Seitz 2012]. However, a survey of acquisition methods is out of the scope of this survey. We rather provide a classification in terms of the characteristics of the acquired information that have an impact on the processing pipeline. Our classification will differentiate *Purely visual input sources*, *Purely geometric input sources*, and *Multimodal colorimetric and geometric input sources*.

Open research data. A notable number of freely available datasets containing indoor scenes have been released in recent years for the purposes of benchmarking and/or training learning-based solutions. However, most of them are more focused on scene understanding [University of Zurich 2016] than reconstruction, and often only cover portions of rooms [Cornell University 2012; New York University 2012; Princeton University 2015; Stanford University 2016b; Technical University of Munich 2015; Washington University 2014]. Many of them have been acquired with RGB-D scanners, due to the flexibility and low-cost of this solution (see an established survey [Firman 2016] for a detailed list of them). We will summarize the major open datasets that have been used in general 3D indoor reconstruction research, detailing their characteristics and possible usage. These will include *SUN360 Database* [Massachussets Institute of Technology 2012; Pintore et al. 2018a,b; Xiao et al. 2012; Yang and Zhang 2016; Zhang et al. 2014], *SUN3D Database* [Chang et al. 2017; Choi et al. 2015; Dai et al. 2017c; Princeton University 2013; Xiao et al. 2013], *UZH 3D Dataset* [Mattausch et al. 2014; Mura et al. 2014b, 2016; University of Zurich 2014], *SUNCG Dataset* [Armeni et al. 2017; Chang et al. 2017; Liu et al. 2018b; Princeton University 2016; Song et al. 2017], *BundleFusion Dataset* [Dai et al. 2017c; Fu et al. 2017; Huang et al. 2017; Stanford University 2016a], *ScanNet Data* [Chang et al. 2017; Dai et al. 2017a,b], *Matterport3D Dataset* [Chang et al. 2017; Matterport 2017], *2D-3D-S Dataset* [Armeni et al. 2017; Stanford University 2017], *FloorNet Dataset* [Chen et al. 2019; Liu et al. 2018b,c], *CRS4/ViC Research Datasets* [CRS4 Visual Computing 2018; Pintore et al. 2019a, 2018a,b], *Replica Dataset* [Straub et al. 2019], and *Structured3D Dataset* [Sun et al. 2019; Zheng et al. 2019a].

SESSION 1.3: Targeted structured 3D model

The goal of structured 3D indoor reconstruction is to transform an input source containing a sampling of a real-world interior environment into a compact structured model containing both geometric and visual abstractions. Each distinct input source tends to produce only partial coverage and imperfect sampling, making reconstruction difficult and ambiguous. For this reason, research has concentrated on defining priors in order to combat imperfections and focus reconstruction on very specific expected indoor structures, shapes, and visual representations. In this session, we first characterize the artifacts typical of indoor model measurement, before defining the structure and priors commonly used in structured 3D indoor reconstruction research, and the sub-problems connected to its generation.

Artifacts. In this session, we will introduce the characterization provided by Berger et al. [Berger et al. 2017] for point clouds, which characterized sampled sources according to the properties that have the most impact on reconstruction algorithms, identifying them into *sampling density*, *noise*, *outliers*, *misalignment*, and *missing data*. We will then show how this characterization extends to visual and mixed data. We will then discuss how the artifacts associated with each one of these characteristics have some specific forms for indoor environments.

Reconstruction priors. We will show how, without prior assumptions, the reconstruction problem for indoor environments is ill-posed, since an infinite number of solutions may exist that fit under-sampled or partially missing data. We will discuss how structured indoor reconstruction has focused its efforts on formally or implicitly restricting the target output model, in order to cover a large variety of interesting use-cases while making reconstruction tractable, introducing in particular the separation between permanent structures and movable objects, and the organization of permanent structures into a graph of rooms connected by passages. We will then survey very specific geometric priors for structural recovery that have been introduced in the indoor reconstruction literature, including *floor-wall* [Delage et al. 2006], *cuboid* [Hedau et al. 2009], *Manhattan world* [Coughlan and Yuille 1999], *Atlanta world* (a.k.a. *Augmented Manhattan World*) [Schindler and Dellaert 2004], *Indoor World Model* [Lee et al. 2009], *Vertical Walls* [Pintore et al. 2018a], and *Piece-wise planarity* [Furukawa et al. 2009].

Main problems. Starting from the above definitions, we identify a core set of basic problems that need to be solved to construct the model from observed data, which are then discussed in the following sessions: *room segmentation*, *bounding surfaces reconstruction*, *indoor object detection and reconstruction*, *integrated model computation*, and *visual representation generation*.

SESSION 1.4: Room segmentation

While a number of early methods focused on reconstructing the bounding surface of the environment as a single entity, without considering the problem of recognizing individual sub-spaces within it, structuring the 3D model of an indoor environment according to its subdivision into different rooms has gradually become a fundamental step in all modern indoor modeling pipelines, regardless of the type of input they consider (e.g. visual vs. 3D data) or of their main intended goal (e.g. virtual exploration vs. as-built BIM) [Ikehata et al. 2015]. In this session we will discuss approaches that segment the *input* before the application of the reconstruction pipeline, as well as approaches that structure the *output* 3D model according to its subdivision into different rooms.

SESSION 1.5: Bounding surfaces reconstruction - part 1

While room segmentation deals with the problem of decomposing an indoor space into disjoint spaces (e.g., hallways, rooms), the goal of bounding surface reconstruction is to further parse those spaces into the structural elements that bound their geometry (e.g. floor, ceiling, walls, etc.). This task is one of the major challenges in indoor reconstruction, since building interiors are typically cluttered with furniture and other objects. Not only are these elements not relevant to the structural shape of a building, and should therefore considered as outliers for this task, but they also generate viewpoint occlusions resulting in large amounts of missed sampling of the permanent structures. Larger amounts of missed 3D samplings are also present in visual input sources. Thus, generic surface reconstruction approaches are doomed to fail. In this session, we will discuss an array of specific state-of-the-art approaches, focusing primarily on the extraction of walls, ceilings, and floors. Given the complexity of the topic, the session is subdivided in two parts. In this first session, we will introduce the topic and discuss methods for reconstruction *with dense geometric measures*, acquired either by stereo or by direct measurement of depth.

SESSION 2.1: Bounding surfaces reconstruction - part 2

The second part of the bounding surface reconstruction session will be devoted to techniques that perform reconstruction *without geometric measures as input sources* and *with sparse geometric measures*. As we will see, these technique exploit mostly visual input data (single- and multi-view).

SESSION 2.2: Object detection and reconstruction

Modeling objects that occur in indoor scenes is a recurrent problem in computer graphics and computer vision research. In this context, the term *object* refers to a part of the environment that is movable (typically, furniture) and thus does not belong to the architectural structure. In this session, we will survey those aspects of indoor object modeling that are integrated in the reconstruction of the entire indoor scene. In particular, we will present approaches where object detection is exploited for clutter removal, methods where 3D indoor objects are approximately reconstructed, and specialized techniques targeting the detection and modeling of flat objects attached to walls and ceilings.

SESSION 2.3: Integrated model computation

The structured reconstruction of a complex environment requires not only the analysis of isolated structures, permanent or not, but also to ensure their integration into a coherent structured model. In this session, we will first discuss how the boundary models of the different rooms are made geometrically and structurally consistent, ensuring for instance that the separating wall boundaries between adjacent rooms are correctly modeled based on the specific output representation of choice. Secondly, we will show methods that find connections among rooms, so that adjacent rooms are connected by doors or large passages that directly reflect the intended functionality of the environment and that can therefore be integrated in its structured representation in the form of graph edges. Moreover, the structure of a multi-room environment goes beyond the plain geometric description of its rooms and is strongly related to the way such rooms are connected. For this reason, we will also present approaches for the extraction of a graph that encodes the room interconnections in multi-room and multi-floor environments.

SESSION 2.4: Visual representation generation

The geometric and topological description coming out of the previous steps may not be enough for the applications that should ultimately visualize the reconstructed model. It is therefore necessary to enrich the structured representation with information geared towards visual representation. In this session, we will discuss how generating visual representations translates into two different problems: the improvement of appearance of reconstructed models with additional geometric and visual data, and the generation of structures to support exploration and navigation. We will then discuss techniques to improve the appearance of reconstructed models by refining the color or by refining the geometry. We will finally show how providing support for visualizing/exploring the dataset has especially been tackled in the context of applications that link the structured reconstruction to the original data, and will present current approaches.

SESSION 2.5: Wrap-up and discussion

In this concluding session, we will summarize the main result coming out of the literature survey and provide examples of applications in which the techniques are exploiting, focusing especially on emerging software-as-a-service approaches. We will then provide a view on open problems and current and future works. We will particularly mention work that exploits less constraining priors, performing data fusion to combine visual and depth cues into multi-modal feature descriptors to help reconstruction, improving reconstruction from visual input from commodity cameras and smartphones, as well as exploiting data-driven priors to learn hidden relations from the available data.

5 TUTORIAL NOTES CONTENTS

At the end of this tutorial, we include a full bibliography, as well as commented slides for all the tutorial sessions.

6 SCHEDULE

Duration	Lecturer	Topic	Sub-topics
10'	Gobbetti	Opening and introduction	Topic definition; Main applications; Course outline
10'	Gobbetti	Data capture and representation	Input data sources; Capture setups; Open research data
15'	Gobbetti	Targeted structured 3D model	Artifacts; Reconstruction priors; Main problems
25'	Mura	Room segmentation	Segmentation of input; Segmentation of output
25'	Pajarola	Bounding surfaces reconstruction - part 1	With dense geometric measures
BREAK			
25'	Pintore	Bounding surfaces reconstruction - part 2	Without geometric measures as input sources; With sparse geometric measures
20'	Pintore	Indoor object detection and reconstruction	Object detection for clutter removal; 3D indoor objects detection and reconstruction; Flat indoor objects detection and reconstruction
15'	Ganovelli	Integrated model computation	Ensuring consistency of multi-room models; Finding and modeling connections; Multi-room and multi-floor graphs
15'	Ganovelli	Visual representation generation	Geometry refinement; Texture refinement; Visual exploration
15'	Gobbetti	Wrap-up and discussion	Summary of techniques and assessment of capabilities; Open problems; Q&A

7 AUTHORS AND LECTURERS

- **Giovanni Pintore** is a senior research engineer at the Visual Computing (ViC) group at the Center for Advanced Studies, Research, and Development in Sardinia (CRS4). He holds a Laurea (M. Sc.) degree (2002) in Electronics Engineering from the University of Cagliari. His research interests include methods for 3D reconstruction of structured indoor scenes from images, multi-resolution representations of large and complex 3D models, as well as visual computing applications of mobile graphics. He has published a number of works in the field of both interactive and automatic reconstruction of indoor structures and has given several courses in international conferences, such as Eurographics, SIGGRAPH Asia, and 3DV, focusing on mobile capture and metric reconstruction of architectural scenes. He has contributed as key developer and manager in international industrial and research projects in the areas of security, space exploration and smart cities. He served as program chair, editor and reviewer in international conferences and journals.
- **Claudio Mura** is a postdoctoral researcher and lecturer at the Visualization and MultiMedia Lab of the University of Zurich, from which he obtained a Ph.D. in Informatics in 2017 while working as an Early-Stage Researcher in the EU FP7 MSCA-ITN project DIVA. Before that, he received a M.Sc. degree in Computer Science from the University of Cagliari, Italy. His research, for which he has obtained direct funding from several public and private institutions, has been awarded with the Best Student Paper Award at the 2016 Pacific Graphics Conference and the 2nd Best Paper Award at the 2018 Computer Graphics International Conference. He has also collaborated with industry partners in R&D and technology transfer projects. His current research interests include 3D modeling and semantic understanding of interiors, point-based shape analysis and point cloud processing.
- **Lizeth Fuentes** is a doctoral candidate at the Visualization and MultiMedia Lab of the University of Zurich, working as an Early-Stage Researcher in the H2020 MSCA-ITN project EVOCATION. She obtained a B.Sc. degree in Computer Science from the National University of Saint Augustine, Peru, and a M.Sc. degree (2017) in Computer Science from the Federal Fluminense University, Rio de Janeiro, Brazil. Her research interests are geometry processing, computer vision, shape analysis and machine learning.
- **Fabio Ganovelli** is a research scientist at the Istituto di Scienza e Tecnologie dell'Informazione (ISTI) of the National Research Council (CNR) in Pisa, Italy. He received his PhD from University of Pisa in 2001. Since then, he published in the fields of deformable objects, geometry processing, out-of-core rendering and manipulation of massive models, photorealistic rendering, image-to-geometry registration, indoor reconstruction, and education. He is a core developer of the Visualization and Computer Graphics Library and served as reviewer and/or chair for all the main journals and conferences in Computer Graphics.
- **Renato Pajarola** is a full Professor in the Department of Informatics at the University of Zürich (UZH). He received a Dipl. Inf-Ing ETH as well as a Dr. sc. techn. degree in computer science from the Swiss Federal Institute of Technology (ETH) Zurich in 1994 and 1998 respectively. Subsequently he was a post-doctoral researcher and lecturer in the Graphics, Visualization and Usability Center at Georgia Tech. In 1999 he joined the University of California Irvine as an Assistant Professor where he established the Computer Graphics Lab. Since 2005 he has been leading the Visualization and MultiMedia Lab at UZH. He is a Senior Member of ACM and IEEE as well as a Fellow of the Eurographics Association. Dr. Pajarola's research interests include interactive large-scale data visualization, real-time 3D graphics, 3D scanning and reconstruction, geometry processing, as well as remote and parallel rendering. He has published a wide range of internationally peer-reviewed research articles in top journals and conferences. Prof. Pajarola regularly serves on program committees, such as for example the IEEE Visualization Conference, Eurographics, EuroVis Conference, IEEE Pacific Visualization or ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games. He organized and co-chaired the Eurographics Conference in 2015, chaired the 2010 EG Symposium on Parallel Graphics and Visualization and was papers co-chair in 2011, and also of the 2007 and 2008 IEEE/EG Symposium on Point-Based Computer Graphics. His recent co-authored papers received a SPIE Best Paper Award in 2013, a Best Student Paper at the Pacific Graphics Conference and an Honorable Mention Award at the ACM SIGGRAPH Symposium on Visualization both in 2016, as well as a (2nd) Best Paper Award at the Computer Graphics International Conference in 2018.

- **Enrico Gobbetti** is the director of Visual Computing (ViC) and Data-Intensive Computing (DiC) at the Center for Advanced Studies, Research, and Development in Sardinia (CRS4), Italy. He holds an Engineering degree (1989) and a Ph.D. degree (1993) in Computer Science from the Swiss Federal Institute of Technology in Lausanne (EPFL). Prior to joining CRS4, he held research and/or teaching positions at EPFL, University of Maryland, and NASA. His main research interests span many areas of visual and distributed computing, with emphasis on scalable technology for acquisition, storage, processing, distribution, and interactive exploration of complex objects and environments. Systems based on these technologies have been used in as diverse real-world applications as internet geoviewing, scientific data analysis, surgical training, and cultural heritage study and dissemination. Enrico has (co-)authored over 200 papers, eight of which received best paper awards. He regularly serves the scientific community through participation in editorial boards, conference committes, and working groups, as well as through the organization and chairing of conferences. He is a Fellow of Eurographics.

ACKNOWLEDGMENTS

This work has received funding from Sardinian Regional Authorities under projects VIGECLAB, AMAC, and TDM (POR FESR 2014-2020 Action 1.2.2). We also acknowledge the contribution of the European Union's H2020 research and innovation programme under grant agreements 813170 (EVOCATION) and 820434 (ENCORE).

REFERENCES

3DVista. 1999. 3DVista:Professional Virtual Tour software. https://www.3dvista.com.
Antonio Adan and Daniel Huber. 2011. 3D reconstruction of interior wall surfaces under occlusion and clutter. In *Proc. 3DIMPVT*. 275–281.
A. Agarwala, A. Colburn, A. Hertzmann, B. Curless, and M. F. Cohen. 2013. Image-Based Remodeling. *IEEE TVCG* 19, 01 (2013), 56–66.
Marco Agus, Enrico Gobbetti, Fabio Marton, Giovanni Pintore, and Pere-Pau Vázquez. 2017a. Mobile Graphics. In *SIGGRAPH Asia 2017 Courses*.
Marco Agus, Enrico Gobbetti, Fabio Marton, Giovanni Pintore, and Pere-Pau Vázquez. 2017b. Mobile Graphics. In *Proc. EUROGRAPHICS Tutorials*, Adrien Bousseau and Diego Gutierrez (Eds.).
Mohamed Aly and Jean-Yves Bouguet. 2012. Street view goes indoors: Automatic pose estimation from uncalibrated unordered spherical panoramas. In *Proc. WACV*. 1–8.
Rareş Ambruş, Sebastian Claici, and Axel Wendt. 2017. Automatic Room Segmentation From Unstructured 3-D Data of Indoor Environments. *IEEE Robotics and Automation Letters* 2, 2 (2017), 749–756.
Abhishek Anand, Hema Swetha Koppula, Thorsten Joachims, and Ashutosh Saxena. 2013. Contextually guided semantic labeling and search for three-dimensional point clouds. *The International Journal of Robotics Research* 32, 1 (2013), 19–34.
Iro Armeni, Zhi-Yang He, JunYoung Gwak, Amir R. Zamir, Martin Fischer, Jitendra Malik, and Silvio Savarese. 2019. 3D Scene Graph: A Structure for Unified Semantics, 3D Space, and Camera. In *Proc. ICCV*.
I. Armeni, A. Sax, A. R. Zamir, and S. Savarese. 2017. Joint 2D-3D-Semantic Data for Indoor Scene Understanding. *ArXiv e-prints* (Feb. 2017). arXiv:1702.01105
Iro Armeni, Ozan Sener, Amir R. Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 2016. 3D Semantic Parsing of Large-scale Indoor Spaces. In *Proc. CVPR*. 1534–1543.
S. Y. Bao, A. Furlan, L. Fei-Fei, and S. Savarese. 2014. Understanding the 3D layout of a cluttered room from multiple images. In *Proc. IEEE WACV*. 690–697.
Matthew Berger, Andrea Tagliasacchi, Lee M. Seversky, Pierre Alliez, Gaël Guennebaud, Joshua A. Levine, Andrei Sharf, and Claudio T. Silva. 2017. A Survey of Surface Reconstruction from Point Clouds. *Computer Graphics Forum* 36, 1 (2017), 301–329.
Dmytro Bobkov, Martin Kiechle, Sebastian Hilsenbeck, and Eckehard Steinbach. 2017. Room Segmentation in 3D Point Clouds using Anisotropic Potential Fields. In *Proc. ICME*. 727–732.
András Bódis-Szomorú, Hayko Riemenschneider, and Luc Van Gool. 2014. Fast, approximate piecewise-planar modeling based on sparse structure-from-motion and superpixels. In *Proc. CVPR*. 469–476.
Alexandre Boulch, Martin de La Gorce, and Renaud Marlet. 2014. Piecewise-Planar 3D Reconstruction with Edge and Corner Regularization. *Computer Graphics Forum* 33, 5 (2014), 55–64.
Alexandre Boulch, Simon Houllier, Renaud Marlet, and Olivier Tournaire. 2013. Semantizing Complex 3D Scenes using Constrained Attribute Grammars. *Computer Graphics Forum* 32, 5 (2013), 33–42.
Yuri Boykov, Olga Veksler, and Ramin Zabih. 2001. Fast Approximate Energy Minimization via Graph Cuts. *IEEE TPAMI* 23, 11 (November 2001), 1222–1239.
Gabriel J. Brostow, Jamie Shotton, Julien Fauqueur, and Roberto Cipolla. 2008. Segmentation and Recognition Using Structure from Motion Point Clouds. In *Proc, ECCV*, David Forsyth, Philip Torr, and Andrew Zisserman (Eds.). 44–57.
Angela Budroni and Jan Böhm. 2010. Automated 3D Reconstruction of Interiors from Point Clouds. *International Journal of Architectural Computing* 8, 1 (2010), 55–73.
R. Cabral and Y. Furukawa. 2014. Piecewise Planar and Compact Floorplan Reconstruction from Images. In *Proc. CVPR*. 628–635.
Angel Chang, Angela Dai, Thomas Funkhouser, Maciej Halber, Matthias Niessner, Manolis Savva, Shuran Song, Andy Zeng, and Yinda Zhang. 2017. Matterport3D: Learning from RGB-D Data in Indoor Environments. In *Proc. 3DV*. 667–676.
Anne-Laure Chauve, Patrick Labatut, and Jean-Philippe Pons. 2010. Robust Piecewise-Planar 3D Reconstruction and Completion from Large-scale Unstructured Point Data. In *Proc. CVPR*. 1261–1268.
Jiacheng Chen, Chen Liu, Jiaye Wu, and Yasutaka Furukawa. 2019. Floor-SP: Inverse CAD for Floorplans by Sequential Room-wise Shortest Path. *Proc. ICCV* (2019).
Kang Chen, Yu-Kun Lai, and Shi-Min Hu. 2015a. 3D indoor scene modeling from RGB-D data: a survey. *Computational Visual Media* 1, 4 (2015), 267–278.
Kang Chen, Yu-Kun Lai, Yu-Xin Wu, Ralph Martin, and Shi-Min Hu. 2014. Automatic Semantic Modeling of Indoor Scenes from Low-quality RGB-D Data Using Contextual Information. *ACM TOG* 33, 6 (Nov. 2014), 208:1–208:12.
Kang Chen, Kun Xu, Yizhou Yu, Tian-Yi Wang, and Shi-Min Hu. 2015b. Magic Decorator: Automatic Material Suggestion for Indoor Digital Scenes. *ACM TOG* 34, 6 (Oct. 2015), 232:1–232:11.
Sungjoon Choi, Qian-Yi Zhou, and Vladlen Koltun. 2015. Robust Reconstruction of Indoor Scenes. In *Proc. CVPR*. 5828–5839.
Marc Christie, Patrick Olivier, and Jean-Marie Normand. 2008. Camera control in computer graphics. *Computer Graphics Forum* 27, 8 (2008), 2197–2218.
Ibrahim Cinaroglu and Yalin Bastanlar. 2016. A direct approach for object detection with catadioptric omnidirectional cameras. *Signal, Image and Video Processing* 10, 2 (2016), 413–420.

Daniel Cohen-Or, Yiorgos L Chrysanthou, Claudio T. Silva, and Frédo Durand. 2003. A survey of visibility for walkthrough applications. *IEEE TVCG* 9, 3 (2003), 412–431.

Cornell University. 2012. Cornell RGBD dataset. http://pr.cs.cornell.edu/sceneunderstanding/data/data.php. [Accessed: 2019-09-25].

James M Coughlan and Alan L Yuille. 1999. Manhattan world: Compass direction from a single image by bayesian inference. In *Proc. ICCV*, Vol. 2. 941–947.

CRS4 Visual Computing. 2018. CRS4 ViC Research Datasets. http://vic.crs4.it/download/datasets/. [Accessed: 2019-09-25].

Y. Cui, Q. Li, B. Yang, W. Xiao, C. Chen, and Z. Dong. 2019. Automatic 3-D Reconstruction of Indoor Environment With Mobile Laser Scanning Point Clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2019), 1–14.

Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Niessner. 2017a. ScanNet Data. http://www.scan-net.org/. [Accessed: 2019-09-25].

Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Niessner. 2017b. ScanNet: Richly-annotated 3D Reconstructions of Indoor Scenes. In *Proc. CVPR*.

Angela Dai, Matthias Niessner, Michael Zollhofer, Shahram Izadi, and Christian Theobalt. 2017c. BundleFusion: Real-Time Globally Consistent 3D Reconstruction Using On-the-Fly Surface Reintegration. *ACM TOG* 36, 4 (2017), 24:1–24:18.

Angela Dai, Daniel Ritchie, Martin Bokeloh, Scott Reed, Jürgen Sturm, and Matthias Niessner. 2018. ScanComplete: Large-Scale Scene Completion and Semantic Segmentation for 3D Scans. In *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*.

Soheil Darabi, Eli Shechtman, Connelly Barnes, Dan B Goldman, and Pradeep Sen. 2012. Image melding: Combining inconsistent images using patch-based synthesis. *ACM TOG* 31, 4 (2012), 82–1.

L. Del Pero, J. Bowdish, D. Fried, B. Kermgard, E. Hartley, and K. Barnard. 2012. Bayesian geometric modeling of indoor scenes. In *Proc. CVPR*. 2719–2726.

L. Del Pero, J. Bowdish, B. Kermgard, E. Hartley, and K. Barnard. 2013. Understanding Bayesian Rooms Using Composite 3D Object Models. In *Proc. CVPR*. 153–160.

E. Delage, Honglak Lee, and A. Y. Ng. 2006. A Dynamic Bayesian Network Model for Autonomous 3D Reconstruction from a Single Indoor Image. In *Proc. CVPR*, Vol. 2. 2418–2428.

M. Di Benedetto, F. Ganovelli, M. Balsa Rodriguez, A. Jaspe Villanueva, R. Scopigno, and E. Gobbetti. 2014. ExploreMaps: Efficient Construction and Ubiquitous Exploration of Panoramic View Graphs of Complex 3D Environments. *Computer Graphics Forum* 33, 2 (2014), 459–468.

Youli Ding, Xianwei Zheng, Yan Zhou, Hanjiang Xiong, et al. 2019. Low-Cost and Efficient Indoor 3D Reconstruction through Annotated Hierarchical Structure-from-Motion. *Remote Sensing* 11, 1 (2019), 58.

Herbert Edelsbrunner, Joseph O'Rourke, and Raimund Seidel. 1986. Constructing Arrangements of Lines and Hyperplanes with Applications. *SIAM J. Comput.* 15, 2 (May 1986), 341–363.

ETH Zurich. 2017. ETH3D Dataset. https://www.eth3d.net/datasets. [Accessed: 2019-09-25].

Habib Fathi, Fei Dai, and Manolis Lourakis. 2015. Automated as-built 3D reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenges. *Advanced Engineering Informatics* 29, 2 (2015), 149–161.

Michael Firman. 2016. RGBD Datasets: Past, Present and Future. In *Proc. CVPR Workshop on Large Scale 3D Data: Acquisition, Modelling and Analysis*.

Michael Firman, Oisin Mac Aodha, Simon Julier, and Gabriel J Brostow. 2016. Structured prediction of unobserved voxels from a single depth image. In *Proc. CVPR*. 5431–5440.

Matthew Fisher, Manolis Savva, Yangyan Li, Pat Hanrahan, and Matthias Niessner. 2015. Activity-centric Scene Synthesis for Functional 3D Scene Modeling. *ACM TOG* 34, 6 (2015), 170:1–179:13.

Alex Flint, Christopher Mei, David Murray, and Ian Reid. 2010. A Dynamic Programming Approach to Reconstructing Building Interiors. In *Proc ECCV*, Kostas Daniilidis, Petros Maragos, and Nikos Paragios (Eds.). 394–407.

A. Flint, D. Murray, and I. Reid. 2011. Manhattan scene understanding using monocular, stereo, and 3D features. In *Proc. ICCV*. 2228–2235.

Stephen Friedman, Hanna Pasula, and Dieter Fox. 2007. Voronoi Random Fields: Extracting Topological Structure of Indoor Environments via Place Labeling. In *IJCAI*, Vol. 7. 2109–2114.

Qiang Fu, Xiaowu Chen, Xiaotian Wang, Sijia Wen, Bin Zhou, and Hongbo Fu. 2017. Adaptive Synthesis of Indoor Scenes via Activity-associated Object Relation Graphs. *ACM Trans. Graph.* 36, 6 (Nov. 2017), 201:1–201:13.

Y. Furukawa, B. Curless, S. M. Seitz, and R. Szeliski. 2009. Manhattan-world stereo. In *Proc. CVPR*. 1422–1429.

Yasutaka Furukawa, Brian Curless, Steven M. Seitz, and Richard Szeliski. 2009. Reconstructing building interiors from images. In *Proc. ICCV*. 80–87.

David Gallup, Jan-Michael Frahm, and Marc Pollefeys. 2010. Piecewise planar and non-planar stereo for urban scene reconstruction. In *Proc. CVPR*. 1418–1425.

Christopher Geyer and Kostas Daniilidis. 2000. A Unifying Theory for Central Panoramic Systems and Practical Implications. In *Proc. ECCV*. 445–461.

Enrico Gobbetti. 2019. Creation and Exploration of Reality-based Models. *Computers Graphics Forum* 38, 2 (2019), xvii.

Enrico Gobbetti, Dave Kasik, and Sung-eui Yoon. 2008. Technical strategies for massive model visualization. In *Proc. ACM Symp. on Solid and physical modeling*. 405–415.

Mani Golparvar Fard, Feniosky Pea-Mora, Carlos A. Arboleda, and Sanghyun Lee. 2009. Visualization of construction progress monitoring with 4D simulation model overlaid on time-lapsed photographs. *Journal of Computing in Civil Engineering* 23, 6 (2009), 391–404.

Ruiqi Guo and Derek Hoiem. 2013. Support Surface Prediction in Indoor Scenes. In *Proc. ICCV*. 2144–2151.

A. Gupta, S. Satkin, A. A. Efros, and M. Hebert. 2011. From 3D scene geometry to human workspace. In *Proc. CVPR*. 1961–1968.

A. Handa, T. Whelan, J.B. McDonald, and A.J. Davison. 2014. A Benchmark for RGB-D Visual Odometry, 3D Reconstruction and SLAM. In *Proc. ICRA*.

David Harel and Yehuda Koren. 2001. On Clustering Using Random Walks. In *Proc. FST TCS*. 18–41.

Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask R-CNN. In *Proc. ICCV*. 2961–2969.

V. Hedau, D. Hoiem, and D. Forsyth. 2009. Recovering the spatial layout of cluttered rooms. In *Proc. ICCV*. 1849–1856.

Varsha Hedau, Derek Hoiem, and David Forsyth. 2010. Thinking Inside the Box: Using Appearance Models and Context Based on Room Geometry. In *Proc. ECCV*. 224–237.

V. Hedau, D. Hoiem, and D. Forsyth. 2012. Recovering free space of indoor scenes from a single image. In *Proc. CVPR*. 2807–2814.

Derek Hoiem, Alexei A. Efros, and Martial Hebert. 2007. Recovering Surface Layout from an Image. *International Journal of Computer Vision* 75, 1 (01 Oct 2007), 151–172.

Binh-Son Hua, Quang-Hieu Pham, Duc Thanh Nguyen, Minh-Khoi Tran, Lap-Fai Yu, and Sai-Kit Yeung. 2016. SceneNN: A Scene Meshes Dataset with aNNotations. In *Proc. 3DV*.

Jingwei Huang, Angela Dai, Leonidas Guibas, and Matthias Niessner. 2017. 3Dlite: Towards Commodity 3D Scanning for Content Creation. *ACM TOG* 36, 6 (2017), 203:1–203:14.

ICL. 2017. ICL-NUIM RGB-D Dataset. https://www.doc.ic.ac.uk/~ahanda/VaFRIC/iclnuim.html. [accessed: 2019-09-24].

Satoshi Ikehata, Hang Yang, and Yasutaka Furukawa. 2015. Structured Indoor Modeling. In *Proc. ICCV*. 1323–1331.

A. Iraqui, Y. Dupuis, R. Boutteau, J. Y. Ertaud, and X. Savatier. 2010. Fusion of Omnidirectional and PTZ Cameras for Face Detection and Tracking. In *Proc. Int. Conf. on Emerging Security Technologies*. 18–23.

Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon. 2011. KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera. In *Proc. UIST*. 559–568.

Philipp Jenke, Benjamin Huhle, and Wolfgang Strasser. 2009. Statistical Reconstruction of Indoor Scenes. In *Proc. WSCG*. 17–24.

Hou Ji, Angela Dai, and Matthias Niessner. 2019. 3D-SIS: 3D Semantic Instance Segmentation of RGB-D Scans. In *Proc. CVPR*.

Z. Jia, A. Gallagher, A. Saxena, and T. Chen. 2013. 3D-Based Reasoning with Blocks, Support, and Stability. In *Proc. CVPR*. 1–8.

Adrien Kaiser, Jose Alonso Ybanez Zepeda, and Tamy Boubekeur. 2019. A Survey of Simple Geometric Primitives Detection Methods for Captured 3D Data. *Computer Graphics Forum* 38, 1 (2019), 167–196.

Sungil Kang, Annah Roh, Bodam Nam, and Hyunki Hong. 2011. People detection method using graphics processing units for a mobile robot with an omnidirectional camera. *Optical Engineering* 50 (2011), 50:1–50:9.

Z. Sadeghipour Kermani, Z. Liao, P. Tan, and H. Zhang. 2016. Learning 3D Scene Synthesis from Annotated RGB-D Images. In *Proc. SGP*. 197–206.

Young Min Kim, Niloy J Mitra, Dong-Ming Yan, and Leonidas Guibas. 2012. Acquiring 3D indoor environments with variability and repetition. *ACM TOG* 31, 6 (2012), 138:1–138:10.

Kujiale.com. 2019. Structured3D Data. https://structured3d-dataset.org/. [Accessed: 2019-09-25].

Avanish Kushal, Ben Self, Yasutaka Furukawa, David Gallup, Carlos Hernandez, Brian Curless, and Steven M Seitz. 2012. Photo tours. In *Proc. 3DIMPVT*. 57–64.

K. Lai, L. Bo, and D. Fox. 2014. Unsupervised feature learning for 3D scene labeling. In *Proc. ICRA*. 3050–3057.

David C. Lee, Abhinav Gupta, Martial Hebert, and Takeo Kanade. 2010. Estimating Spatial Layout of Rooms Using Volumetric Reasoning About Objects and Surfaces. In *Proc. NIPS*. 1288–1296.

David C Lee, Martial Hebert, and Takeo Kanade. 2009. Geometric reasoning for single image structure recovery. In *Proc. CVPR*. 2136–2143.

K. Lee, S. Ryu, S. Yeon, H. Cho, C. Jun, J. Kang, H. Choi, J. Hyeon, I. Baek, W. Jung, H. Kim, and N. L. Doh. 2016. Accurate Continuous Sweeping Framework in Indoor Spaces With Backpack Sensor System for Applications to 3-D Mapping. *IEEE Robotics and Automation Letters* 1, 1 (2016), 316–323.

Ville Lehtola, Harri Kaartinen, Andreas Nüchter, Risto Kaijaluoto, Antero Kukko, Paula Litkey, Eija Honkavaara, Tomi Rosnell, Matti Vaaja, Juho-Pekka Virtanen, et al.

2017. Comparison of the selected state-of-the-art 3D indoor scanning and point cloud generation methods. *Remote sensing* 9, 8 (2017), 796.

Yangyan Li, Angela Dai, Leonidas Guibas, and Matthias Niessner. 2015. Database-Assisted Object Retrieval for Real-Time 3D Reconstruction. *Computer Graphics Forum* 34, 2 (May 2015), 435–446.

Chen Liu, Kihwan Kim, Jinwei Gu, Yasutaka Furukawa, and Jan Kautz. 2019. Planercnn: 3D Plane Detection and Reconstruction from a Single Image. In *Proc. CVPR*. 4450–4459.

C. Liu, P. Kohli, and Y. Furukawa. 2016. Layered Scene Decomposition via the Occlusion-CRF. In *Proc. CVPR*. 165–173.

Chen Liu, Jiaye Wu, and Yasutaka Furukawa. 2018a. Data. https://github.com/art-programmer/. [Accessed: 2019-09-25].

Chen Liu, Jiaye Wu, and Yasutaka Furukawa. 2018b. FloorNet: A Unified Framework for Floorplan Reconstruction from 3D Scans. In *Proc. ECCV*, Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss (Eds.). 203–219.

Chen Liu, Jiaye Wu, and Yasutaka Furukawa. 2018c. FloorNet Data. https://github.com/art-programmer/FloorNet. [Accessed: 2018-10-24].

Chen Liu, Jiajun Wu, Pushmeet Kohli, and Yasutaka Furukawa. 2017. Raster-to-vector: Revisiting Floorplan Transformation. In *Proc. ICCV*. 2195–2203.

Chen Liu, Jimei Yang, Duygu Ceylan, Ersin Yumer, and Yasutaka Furukawa. 2018d. Planenet: Piece-wise Planar Reconstruction from a Single RGB Image. In *Proc. CVPR*. 2579–2588.

Andelo Martinovic and Luc Van Gool. 2013. Bayesian grammar learning for inverse procedural modeling. In *Proc. CVPR*. 201–208.

Eleonora Maset, Federica Arrigoni, and Andrea Fusiello. 2017. Practical and efficient multi-view matching. In *Proc. ICCV*. 4568–4576.

Massachussets Institute of Technology. 2012. SUN360 Database. http://people.csail.mit.edu/jxiao/SUN360/. [Accessed: 2019-09-25].

Oliver Mattausch, Daniele Panozzo, Claudio Mura, Olga Sorkine-Hornung, and Renato Pajarola. 2014. Object Detection and Classification from Large-Scale Cluttered Indoor Scans. *Computer Graphics Forum* 33, 2 (2014), 11–21.

Matterport. 2017. Matterport3D. https://github.com/niessner/Matterport. [Accessed: 2019-09-25].

Kevin Matzen, Michael F. Cohen, Bryce Evans, Johannes Kopf, and Richard Szeliski. 2017. Low-cost 360 Stereo Photography and Video Capture. *ACM TOG* 36, 4 (2017), 148:1–148:12.

Paul Merrell, Eric Schkufza, Zeyang Li, Maneesh Agrawala, and Vladlen Koltun. 2011. Interactive Furniture Layout Using Interior Design Guidelines. *ACM TOG* 30, 4 (July 2011), 87:1–87:10.

Aron Monszpart, Nicolas Mellado, Gabriel J. Brostow, and Niloy J. Mitra. 2015. RAPter: Rebuilding Man-Made Scenes with Regular Arrangements of Planes. *ACM TOG* 34, 4 (2015), 103:1–103:12.

Claudio Mura, Alberto Jaspe Villanueva, Oliver Mattausch, Enrico Gobbetti, and Renato Pajarola. 2014a. Reconstructing Complex Indoor Environments with Arbitrary Walls Orientations. In *Eurographics Posters*.

Claudio Mura, Oliver Mattausch, Alberto Jaspe Villanueva, Enrico Gobbetti, and Renato Pajarola. 2014b. Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts. *Computers & Graphics* 44 (2014), 20–32.

Claudio Mura, Oliver Mattausch, and Renato Pajarola. 2016. Piecewise-planar Reconstruction of Multi-room Interiors with Arbitrary Wall Arrangements. *Computer Graphics Forum* 35, 7 (2016), 179–188.

Claudio Mura and Renato Pajarola. 2017. Exploiting the Room Structure of Buildings for Scalable Architectural Modeling of Interiors. In *ACM SIGGRAPH Posters*. 4:1–4:2.

Srivathsan Murali, Pablo Speciale, Martin R. Oswald, and Marc Pollefeys. 2017. Indoor Scan2BIM: Building Information Models of House Interiors. In *Proc. IROS*. 6126–6133.

Przemyslaw Musialski, Peter Wonka, Daniel G Aliaga, Michael Wimmer, Luc Van Gool, and Werner Purgathofer. 2013. A survey of urban reconstruction. *Computer graphics forum* 32, 6 (2013), 146–177.

Liangliang Nan, Ke Xie, and Andrei Sharf. 2012. A Search-classify Approach for Cluttered Indoor Scene Understanding. *ACM TOG* 31, 6 (2012), 137:1–137:10.

Pushmeet Kohli Nathan Silberman, Derek Hoiem and Rob Fergus. 2012. Indoor Segmentation and Support Inference from RGBD Images. In *Proc. ECCV*.

NavVis. 2012. TUMViewer. https://www.navvis.lmt.ei.tum.de/view/. [Accessed: 2019-09-25].

New York University. 2012. NYU-Depth V2. https://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html. [Accessed: 2019-09-25].

Sebastian Ochmann, Richard Vock, , and Reinhard Klein. 2019. Automatic Reconstruction of Fully Volumetric 3D Building Models from Oriented Point Clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 151 (2019), 251 – 262.

Sebastian Ochmann, Richard Vock, Raoul Wessel, and Reinhard Klein. 2016. Automatic Reconstruction of Parametric Building Models from Indoor Point Clouds. *Computers & Graphics* 54 (February 2016), 94–103.

Sebastian Ochmann, Richard Vock, Raoul Wessel, Martin Tamke, and Reinhard Klein. 2014. Automatic generation of structural building descriptions from 3D point cloud scans. In *Proc. GRAPP*. 1–8.

Sven Oesau, Florent Lafarge, and Pierre Alliez. 2014. Indoor Scene Reconstruction using Feature Sensitive Primitive Extraction and Graph-cut. *ISPRS Journal of Photogrammetry and Remote Sensing* 90 (2014), 68–82.

Sven Oesau, Florent Lafarge, and Pierre Alliez. 2016. Planar Shape Detection and Regularization in Tandem. *Computer Graphics Forum* 35, 1 (2016), 203–215.

Viorica Pătrăucean, Iro Armeni, Mohammad Nahangi, Jamie Yeung, Ioannis Brilakis, and Carl Haas. 2015. State of Research in Automatic As-Built Modelling. *Advanced Engineering Informatics* 29, 2 (2015), 162–171.

Mark Pauly, Niloy J. Mitra, Joachim Giesen, Markus Gross, and Leonidas J. Guibas. 2005. Example-based 3D Scan Completion. In *Proc. SGP*. 23:1–23:10.

Giovanni Pintore, Marco Agus, and Enrico Gobbetti. 2014. Interactive mapping of indoor building structures through mobile devices. In *Proc. 3DV*, Vol. 2. 103–110.

Giovanni Pintore, Fabio Ganovelli, Enrico Gobbetti, and Roberto Scopigno. 2016a. Mobile Mapping and Visualization of Indoor Structures to Simplify Scene Understanding and Location Awareness. In *Proc. ECCV Workshops*. 130–145.

Giovanni Pintore, Fabio Ganovelli, Enrico Gobbetti, and Roberto Scopigno. 2016b. Mobile reconstruction and exploration of indoor structures exploiting omnidirectional images. In *Proc. SIGGRAPH Asia Symposium on Mobile Graphics and Interactive Applications*.

Giovanni Pintore, Fabio Ganovelli, Alberto Jaspe Villanueva, and Enrico Gobbetti. 2019a. Automatic modeling of cluttered floorplans from panoramic images. *Computer Graphics Forum* 38, 7 (2019), 347–358.

Giovanni Pintore, Fabio Ganovelli, Ruggero Pintus, Roberto Scopigno, and Enrico Gobbetti. 2018a. 3D floor plan recovery from overlapping spherical images. *Computational Visual Media* 4, 4 (2018), 367–383.

Giovanni Pintore, Valeria Garro, Fabio Ganovelli, Marco Agus, and Enrico Gobbetti. 2016c. Omnidirectional image capture on mobile devices for fast automatic generation of 2.5D indoor maps. In *Proc. IEEE WACV*. 1–9.

Giovanni Pintore, Claudio Mura, Fabio Ganovelli, Lizeth Fuentes-Perez, Renato Pajarola, and Enrico Gobbetti. 2019b. State-of-the-art in Automatic 3D Reconstruction of Structured Indoor Environments. *Computer Graphics Forum* 39, 2 (2019), 667–699.

Giovanni Pintore, Ruggero Pintus, Fabio Ganovelli, Roberto Scopigno, and Enrico Gobbetti. 2018b. Recovering 3D existing-conditions of indoor structures from spherical images. *Computers & Graphics* 77 (2018), 16–29.

Ruggero Pintus, Enrico Gobbetti, Marco Callieri, and Matteo Dellepiane. 2017. *Techniques for seamless color registration and mapping on dense 3D models*. Springer, 355–376.

Princeton University. 2013. SUN3D Database. https://sun3d.cs.princeton.edu/. [Accessed: 2019-09-25].

Princeton University. 2015. SUNRGBD Database. http://3dvision.princeton.edu/projects/2015/SUNrgbd/. [Accessed: 2019-09-25].

Princeton University. 2016. SceneCG Dataset. https://sscnet.cs.princeton.edu/. [Accessed: 2019-09-25].

Andrzej Pronobis, Barbara Caputo, Patric Jensfelt, and Henrik I Christensen. 2010. A realistic benchmark for visual indoor place recognition. *Robotics and autonomous systems* 58, 1 (2010), 81–96.

K. Pulli, H. Abi-Rached, T. Duchamp, L. G. Shapiro, and W. Stuetzle. 1998. Acquisition and visualization of colored 3D objects. In *Proc. Pattern Recognition*, Vol. 1. 11–15.

Reconstruct Inc. 2016. Reconstruct: A Visual Command Center. https://www.reconstructinc.com/.

Joseph Redmon and Ali Farhadi. 2017. YOLO9000: Better, Faster, Stronger. In *Proc. CVPR*. 7263–7271.

Kensaku Saitoh, Takashi Machida, Kiyoshi Kiyokawa, and Haruo Takemura. 2006. A 2D-3D integrated interface for mobile robot control using omnidirectional images and 3D geometric models. In *Proc. ACM/IEEE Int. Symp. on Mixed and Augmented Reality*. 173–176.

Victor Sanchez and Avideh Zakhor. 2012. Planar 3D Modeling of Building Interiors from Point Cloud Data. In *Proc. ICIP*. 1777–1780.

Aditya Sankar and Steven Seitz. 2012. Capturing Indoor Scenes with Smartphones. In *Proc. UIST*. 403–412.

Scott Satkin, Maheen Rashid, Jason Lin, and Martial Hebert. 2015. 3DNN: 3D Nearest Neighbor. Data-Driven Geometric Scene Understanding Using 3D Models. *International Journal of Computer Vision* 111 (2015), 69–97.

Manolis Savva, Angel X. Chang, Pat Hanrahan, Matthew Fisher, and Matthias Niessner. 2016. PiGraphs: Learning Interaction Snapshots from Observations. *ACM TOG* 35, 4 (2016).

G. Schindler and F. Dellaert. 2004. Atlanta world: an expectation maximization framework for simultaneous low-level edge grouping and camera calibration in complex man-made environments. In *Proc. CVPR*, Vol. 1. I–I.

Ruwen Schnabel, Roland Wahl, and Reinhard Klein. 2007. Efficient RANSAC for Point-Cloud Shape Detection. *Computer Graphics Forum* 26, 2 (2007), 214–226.

D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stueckler, and D. Cremers. 2018. The TUM VI Benchmark for Evaluating Visual-Inertial Odometry. In *Proc. IROS*.

A. G. Schwing, S. Fidler, M. Pollefeys, and R. Urtasun. 2013. Box in the Box: Joint 3D Layout and Object Reasoning from Single Images. In *Proc. ICCV*. 353–360.

T. Schöps, J. L. Schönberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, and A. Geiger. 2017. A Multi-view Stereo Benchmark with High-Resolution Images and Multi-camera Videos. In *Proc. CVPR*. 2538–2547.

Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms.

In *Proc. CVPR*, Vol. 1. 519–528.

Tianjia Shao, Weiwei Xu, Kun Zhou, Jingdong Wang, Dongping Li, and Baining Guo. 2012. An interactive approach to semantic modeling of indoor scenes with an RGBD camera. *ACM TOG* 31, 6 (2012), 136:1–136:10.

Chao-Hui Shen, Hongbo Fu, Kang Chen, and Shi-Min Hu. 2012. Structure recovery by part assembly. *ACM TOG* 31, 6 (2012), 180:1–180:10.

H. Shin, Y. Chon, and H. Cha. 2012. Unsupervised Construction of an Indoor Floor Plan Using a Smartphone. *IEEE TPAMI* 42, 6 (2012), 889–898.

S. N. Sinha, D. Steedly, and R. Szeliski. 2009. Piecewise planar stereo for image-based rendering. In *Proc. ICCV*. 1881–1888.

Sudipta N. Sinha, Drew Steedly, Richard Szeliski, Maneesh Agrawala, and Marc Pollefeys. 2008. Interactive 3D Architectural Modeling from Unordered Photo Collections. *ACM TOG* 27, 5 (Dec. 2008), 159:1–159:10.

Noah Snavely, Steven M. Seitz, and Richard Szeliski. 2008. Modeling the World from Internet Photo Collections. *International Journal of Computer Vision* 80, 2 (2008), 189–210.

S. Song, S. P. Lichtenberg, and J. Xiao. 2015. SUN RGB-D: A RGB-D scene understanding benchmark suite. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 567–576.

Shuran Song, Fisher Yu, Andy Zeng, Angel X Chang, Manolis Savva, and Thomas Funkhouser. 2017. Semantic Scene Completion from a Single Depth Image. *Proc. CVPR* (2017).

Stanford University. 2016a. Bundle Fusion Dataset. https://graphics.stanford.edu/projects/bundlefusion. [Accessed: 2019-09-25].

Stanford University. 2016b. PiGraphs Dataset. https://graphics.stanford.edu/projects/pigraphs/. [Accessed: 2019-09-25].

Stanford University. 2017. BuildingParser Dataset. http://buildingparser.stanford.edu/dataset.html. [Accessed: 2019-09-25].

Julian Straub, Thomas Whelan, Lingni Ma, Yufan Chen, Erik Wijmans, Simon Green, Jakob J. Engel, Raul Mur-Artal, Carl Ren, Shobhit Verma, Anton Clarkson, Mingfei Yan, Brian Budge, Yajie Yan, Xiaqing Pan, June Yon, Yuyang Zou, Kimberly Leon, Nigel Carter, Jesus Briales, Tyler Gillingham, Elias Mueggler, Luis Pesqueira, Manolis Savva, Dhruv Batra, Hauke M. Strasdat, Renzo De Nardi, Michael Goesele, Steven Lovegrove, and Richard Newcombe. 2019. The Replica Dataset: A Digital Replica of Indoor Spaces. arXiv:cs.CV/1906.05797

M. Stroila, A. Yalcin, J. Mays, and N. Alwar. 2012. Route Visualization in Indoor Panoramic Imagery with Open Area Maps. In *Proc. ICME Workshops*. 499–504.

StructionSite. 2016. VideoWalk. https://www.structionsite.com/products/videowalk/.

Hao Su, Charles R Qi, Yangyan Li, and Leonidas J Guibas. 2015. Render for CNN: Viewpoint estimation in images using CNNs trained with rendered 3D model views. In *Proc. ICCV*. 2686–2694.

Cheng Sun, Chi-Wei Hsiao, Min Sun, and Hwann-Tzong Chen. 2019. HorizonNet: Learning Room Layout With 1D Representation and Pano Stretch Data Augmentation. In *Proc. CVPR*.

Pingbo Tang, Daniel Huber, Burcu Akinci, Robert Lipman, and Alan Lytle. 2010. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction* 19, 7 (2010), 829–843.

Technical University of Munich. 2015. TUM LSI Dataset. https://hazirbas.com/datasets/tum-lsi/. [Accessed: 2019-09-25].

G. Tsai, Changhai Xu, Jingen Liu, and B. Kuipers. 2011. Real-time indoor scene understanding using Bayesian filtering with motion cues. In *Proc. ICCV*. 121–128.

Shubham Tulsiani, Abhishek Kar, Joao Carreira, and Jitendra Malik. 2016. Learning category-specific deformable 3D models for object reconstruction. *IEEE TPAMI* 39, 4 (2016), 719–731.

E. Turner, P. Cheng, and A. Zakhor. 2015. Fast, Automated, Scalable Generation of Textured 3D Models of Indoor Environments. *IEEE Journal of Selected Topics in Signal Processing* 9, 3 (2015), 409–421.

Eric Turner and Avideh Zakhor. 2012. Watertight as-Built Architectural Floor Plans Generated from Laser Range Data. In *Proc. 3DIMPVT*. 316–323.

Eric Turner and Avideh Zakhor. 2013. Watertight Planar Surface Meshing of Indoor Point-Clouds with Voxel Carving. In *Proc. 3DV*. 41–48.

Eric Turner and Avideh Zakhor. 2014. Floor Plan Generation and Room Labeling of Indoor Environments from Laser Range Data. In *Proc. Int. Conf. on Computer Graphics Theory and Applications*. 22–33.

University of Zurich. 2014. UZH 3D Dataset. https://www.ifi.uzh.ch/en/vmml/research/datasets.html. [Accessed: 2019-09-25].

University of Zurich. 2016. SceneNN Dataset. https://www.ifi.uzh.ch/en/vmml/research/datasets.html. [Accessed: 2019-09-25].

Andrea Vedaldi and Andrew Zisserman. 2018. Object instance recognition. http://www.robots.ox.ac.uk/~vgg/practicals/instance-recognition/index.html. [Accessed: 2018-10-24].

Rebekka Volk, Julian Stengel, and Frank Schultmann. 2014. Building Information Modeling (BIM) for existing buildings — Literature review and future needs. *Automation in Construction* 38 (2014), 109 – 127.

F. Walch, C. Hazirbas, L. Leal-Taixé, T. Sattler, S. Hilsenbeck, and D. Cremers. 2017. Image-based localization using LSTMs for structured feature correlation. In *Proc. ICCV*.

Huayan Wang, Stephen Gould, and Daphne Koller. 2010. Discriminative Learning with Latent Variables for Cluttered Indoor Scene Understanding. In *Proc. ECCV*, Kostas Daniilidis, Petros Maragos, and Nikos Paragios (Eds.). 435–449.

M. L. Wang and H. Y. Lin. 2009. Object recognition from omnidirectional visual sensing for mobile robot applications. In *Proc. IEEE Int. Conf. on Systems, Man and Cybernetics*. 1941–1946.

Washington University. 2014. Washington RGBD dataset. https://rgbd-dataset.cs.washington.edu/dataset.html. [Accessed: 2019-09-25].

R. T. Whitaker, J. Gregor, and P. F. Chen. 1999. Indoor scene reconstruction from sets of noisy range images. In *Proc. 3-D Digital Imaging and Modeling*. 348–357.

Erik Wijmans and Yasutaka Furukawa. 2017. WUSTL Indoor RGBD Dataset. https://cvpr17.wijmans.xyz/data/. [Accessed: 2019-09-25].

T. Wu, J. Liu, M. Li, R. Chen, and J. Hyyppä. 2018. Automated large scale indoor reconstruction using vehicle survey data. In *Proc. UPINLBS*. 1–5.

Yu Xiang, Roozbeh Mottaghi, and Silvio Savarese. 2014. Beyond PASCAL: A benchmark for 3D object detection in the wild. In *Proc. WACV*. 75–82.

J. Xiao, K. A. Ehinger, A. Oliva, and A. Torralba. 2012. Recognizing scene viewpoint using panoramic place representation. In *Proc. CVPR*. 2695–2702.

Jianxiong Xiao and Yasutaka Furukawa. 2014. Reconstructing the World's Museums. *International Journal of Computer Vision* 110, 3 (Dec 2014), 243–258.

J. Xiao, A. Owens, and A. Torralba. 2013. SUN3D: A Database of Big Spaces Reconstructed Using SfM and Object Labels. In *2013 IEEE International Conference on Computer Vision*. 1625–1632.

Xuehan Xiong, Antonio Adan, Burcu Akinci, and Daniel Huber. 2013. Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in Construction* 31 (2013), 325–337.

Xuehan Xiong and Daniel Huber. 2010. Using Context to Create Semantic 3D Models of Indoor Environments. In *Proc. BMVC*. BMVA Press, 45.1–45.11.

J. Xu, B. Stenger, T. Kerola, and T. Tung. 2017. Pano2CAD: Room Layout from a Single Panorama Image. In *Proc. WACV*. 354–362.

Kun Xu, Kang Chen, Hongbo Fu, Wei-Lun Sun, and Shi-Min Hu. 2013. Sketch2Scene: Sketch-based Co-retrieval and Co-placement of 3D Models. *ACM TOG* 32, 4 (July 2013), 123:1–123:15.

Kai Xu, Hui Huang, Yifei Shi, Hao Li, Pinxin Long, Jianong Caichen, Wei Sun, and Baoquan Chen. 2015. Autoscanning for Coupled Scene Reconstruction and Proactive Object Analysis. *ACM TOG* 34, 6 (2015), 177:1–177:14.

Bo Yang, Stefano Rosa, Andrew Markham, Niki Trigoni, and Hongkai Wen. 2018b. Dense 3D object reconstruction from a single depth view. *IEEE TPAMI* (2018).

Fan Yang, Gang Zhou, Fei Su, Xinkai Zuo, Lei Tang, Yifan Liang, Haihong Zhu, and Lin Li. 2019c. Automatic Indoor Reconstruction from Point Clouds in Multi-room Environments with Curved Walls. *Sensors* 19, 17 (Sep 2019), 3798.

Fengting Yang and Zihan Zhou. 2018. Recovering 3D Planes from a Single Image via Convolutional Neural Networks. In *Proc. ECCV*. 85–100.

H. Yang and H. Zhang. 2016. Efficient 3D Room Shape Recovery from a Single Panorama. In *Proc. CVPR*. 5422–5430.

Jingyu Yang, Ji Xu, Kun Li, Yu-Kun Lai, Huanjing Yue, Jianzhi Lu, Hao Wu, and Yebin Liu. 2019b. Learning to Reconstruct and Understand Indoor Scenes from Sparse Views. *CoRR* (2019). http://arxiv.org/abs/1906.07892

Shang-Ta Yang, Fu-En Wang, Chi-Han Peng, Peter Wonka, Min Sun, and Hung-Kuo Chu. 2019a. DuLa-Net: A Dual-Projection Network for Estimating Room Layouts from a Single RGB Panorama. In *Proc. CVPR*.

Yang Yang, Shi Jin, Ruiyang Liu, , and Jingyi Yu. 2018a. Automatic 3D Indoor Scene Modeling From Single Panorama. In *Proc. CVPR*. 3926–3934.

Yao Yao, Zixin Luo, Shiwei Li, Tianwei Shen, Tian Fang, and Long Quan. 2019. Recurrent MVSNet for High-Resolution Multi-View Stereo Depth Inference. In *Proc. CVPR*.

Edward Zhang, Michael F. Cohen, and Brian Curless. 2016. Emptying, Refurnishing, and Relighting Indoor Spaces. *ACM TOG* 35, 6 (2016), 174:1–174:14.

Jian Zhang, Chen Kan, Alexander G Schwing, and Raquel Urtasun. 2013. Estimating the 3D layout of indoor scenes and its clutter from depth sensors. In *Proc. ICCV*. 1273–1280.

Jianming Zhang, Stan Sclaroff, Zhe Lin, Xiaohui Shen, Brian Price, and Radomir Mech. 2015a. Minimum Barrier Salient Object Detection at 80 FPS. In *Proc. ICCV*. 1404–1412.

Yinda Zhang, Shuran Song, Ping Tan, and Jianxiong Xiao. 2014. PanoContext: A Whole-Room 3D Context Model for Panoramic Scene Understanding. In *Proc. ECCV*. 668–686.

Yizhong Zhang, Weiwei Xu, Yiying Tong, and Kun Zhou. 2015b. Online structure analysis for real-time indoor scene reconstruction. *ACM TOG* 34, 5 (2015), 159:1–159:13.

Jia Zheng, Junfei Zhang, Jing Li, Rui Tang, Shenghua Gao, and Zihan Zhou. 2019a. Structured3D: A Large Photo-realistic Dataset for Structured 3D Modeling. arXiv:cs.CV/1908.00222

Liang Zheng, Yi Yang, and Qi Tian. 2017. SIFT meets CNN: A decade survey of instance retrieval. *IEEE TPAMI* 40, 5 (2017), 1224–1244.

Lintao Zheng, Chenyang Zhu, Jiazhao Zhang, Hang Zhao, Hui Huang, Matthias Niessner, and Kai Xu. 2019b. Active Scene Understanding via Online Semantic Reconstruction. *Computer Graphics Forum* 38, 7 (2019), 103–114.

J. Zhu, Y. Guo, and H. Ma. 2018. A Data-Driven Approach for Furniture and Indoor Scene Colorization. *IEEE TVCG* 24, 9 (2018), 2473–2486.

S. Zingg, D. Scaramuzza, S. Weiss, and R. Siegwart. 2010. MAV navigation through indoor corridors using optical flow. In *Proc. IEEE IROS*. 3361–3368.

Michael Zollhöfer, Patrick Stotko, Andreas Görlitz, Christian Theobalt, Matthias Niessner, Reinhard Klein, and Andreas Kolb. 2018. State of the Art on 3D Reconstruction with RGB-D Cameras. *Computer Graphics Forum* 37, 2 (2018), 625–652.

Chuhang Zou, Alex Colburn, Qi Shan, and Derek Hoiem. 2018. LayoutNet: Reconstructing the 3D Room Layout from a Single RGB Image. In *Proc. CVPR*. 2051–2059.

Chuhang Zou, Jheng-Wei Su, Chi-Han Peng, Alex Colburn, Qi Shan, Peter Wonka, Hung-Kuo Chu, and Derek Hoiem. 2019. 3D Manhattan Room Layout Reconstruction from a Single 360 Image. arXiv:cs.CV/1910.04099