Improving the utilization of close-range photogrammetry and terrestrial laser scanning for photorealistic urban 3D modeling

Arttu Julin
Improving the utilization of close-range photogrammetry and terrestrial laser scanning for photorealistic urban 3D modeling

Arttu Julin

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Engineering, at a public examination held at the lecture hall 216 (Otakaari 4) of the school on 19 November 2021 at 12 noon.

Aalto University
School of Engineering
Department of Built Environment
Research Institute of Measuring and Modeling for the Built Environment (MeMo)
Supervising professor
Professor Matti T. Vaaja, Aalto University

Thesis advisors
Professor Hannu Hyyppä, Aalto University
Professor Juha Hyyppä, Finnish Geospatial Research Institute FGI
Juho-Pekka Virtanen, Aalto University

Preliminary examiners
Professor Fabio Remondino, FBK Trento, Italy
Professor Stephan Nebiker, FHNW, Switzerland

Opponent
Professor Fabio Remondino, FBK Trento, Italy
Abstract

Photorealistic urban 3D models have become valuable and much-requested tools for visualizing, managing, developing, and understanding growing and densifying urban environments in diverse use cases ranging from interactive digital twins to engaging XR experiences. Close-range photogrammetry and terrestrial laser scanning (TLS) can produce accurate, detailed, and photorealistic 3D models from complex urban environments. These close-range 3D measuring techniques are well acknowledged for their complementary benefits and high level of geometric accuracy. However, the quality of the model appearance has been rarely assessed for photorealism. Furthermore, the 3D model itself does not guarantee its utilization and usefulness, though the application platform plays a significant role in putting it into beneficial use.

The dissertation developed ways to improve the utilization of close-range photogrammetry and TLS for photorealistic urban 3D modeling. First, relevant application platforms were identified by studying real-life urban 3D modelling activities. Second, automated multi-sensor urban 3D modeling was experimented for a photorealistic web-based application and the quality of the produced models (close-range photogrammetry, TLS, and a combination of them) was evaluated from the perspective of both model geometry and textures. Finally, the quality of TLS point cloud colorization was evaluated by analyzing the TLS instrument’s capability to reproduce color and details in the scene.

The results identified real-time 3D platforms (i.e., game engines and virtual globes) as the most relevant application platforms for utilizing photorealistic urban 3D models. Integrating close-range photogrammetry and TLS proved a good compromise for efficient model production between the superior texture quality of photogrammetry and the better geometric quality of TLS. Finally, a new method was developed for evaluating the quality of TLS point cloud colorization, which revealed quality differences among all tested commercial TLS instruments and settings.

The dissertation research improved the utilization of and offered a deeper understanding of the efficiency and quality of close-range photogrammetry and TLS for photorealistic urban 3D modeling. For photorealistic use cases, it is crucial to understand the quality of a 3D model as a combination of geometry and appearance. Together, the close-range 3D measuring methods and real-time 3D platforms enable the efficient and beneficial creation and utilization of highly detailed, photorealistic, urban 3D models in both new and well-established fields.

Keywords laser scanning, photogrammetry, photorealism, urban 3D modeling, quality evaluation, model appearance, point cloud, mesh model, game engine, virtual globe

ISBN (printed) 978-952-64-0406-6
ISBN (pdf) 978-952-64-0407-3
ISSN (printed) 1799-4934
ISSN (pdf) 1799-4942
Location of publisher Helsinki
Location of printing Helsinki
Year 2021
Pages 180

# Tiivistelmä


Väitöskirjatyössä kehitettiin tapoja parantaa lähifotoogrammetria ja maalaserkeilausen hyötykäyttöä fotorealistisessa 3D-kaupunkimallintamisessa. Ensinnäkin tunnistettiin olennaiset sovelluslukat tutkimalla käynnissä olevia 3D-kaupunkimallinnushankkeita. Sitten tehokasta mallintuotantoa testattiin kaupunkiympäristön fotorealistentta verkkosovellusta varten ja eri mallinnusmenetelmiä (lähifotoogrammetria, maalaserkeilausta ja niiden yhdistelmää) verrattiin toisinsa arvioimalla mallien geometrican ja tekstuuroinnin laatua. Lopuksi maalaserkeilaudenten tuottaman pisteipilven varjaysen laatua analysoitiin värien ja yksityiskohtien toistokyyyn näkökulmasta.

Tulokset osottivat 3D-grafiikan reaalikaikaiseen renderöintiin perustuvien alustojen (pelimoottorit ja virtual globe -alustat) olevan olennaisimpa sovellusalustoja fotorealistenten 3D-kaupunkimallien hyödyntämiseen. Lähifotoogrammetrija ja maalaserkeilausen välisen integroinnin havaittiin olevan hyvä kompromissi, jossa yhdistyvät lähifotoogrammetrijan parempia teksturooinnin laatua sekä maalaserkeilausen parempi geometrisinen laatua. Lopuksi maalaserkeilaudenten tuottaman pisteipilven varjaysen laadun arviointiin kehitettiin uusi menetelmä, joka paljasti fotorealistenten sovellusten kannalta merkittäviä laatueroja kaikkien menetelmällä testattujen kaupallisten laserkeilaudenten ja asetusten välillä.

Tutkimustyö paransi testattujen 3D-mittausmenetelmien hyödyntämismahdollisuuksia fotorealisten 3D-kaupunkimallinnuksen tarpeisiin sekä tarjosi uutta tietoa niiden tehokkuudesta ja laadusta. Fotorealistentisissa sovelluksissa on tärkeää ymmärtää 3D-mallin laatua kokonaisvaltaisesti sekä mallin geometrican että ulkonäön kannalta. Työssä tutkittiin 3D-mittausmenetelmät ja reaalikaikaiseen renderöintiin perustuvat alustat yhdessä mahdollistavat fotorealistenten ja yksityiskohtaisten 3D-kaupunkimallien tehokkaan tuotannon sekä hyötykäytön niin uusilla kuin jo vakiintuneilla sovellusaloilla.

## Avainsanat
- laserkeilaus, fotogrammetria, fotorealismi, 3D-kaupunkimallinnus, laadun arviointi, 3D-mallin ulkonäkö, pisteipilvi, kolmivierakomalli, pelimoottori

## ISBN
- painettu: 978-952-64-0406-6
- pdf: 978-952-64-0407-3

## ISSN
- 1799-4934
- 1799-4942

## Julkaisupaikka
- Helsinki

## Painopaikka
- Helsinki

## Vuosi
- 2021

## Sivumäärä
- 180

## URL
Acknowledgements

Over the years as a doctoral student, I have had a unique chance to meet and collaborate with awesome people and to work with numerous inspiring topics and cutting-edge technology. Finally, the time has come to acknowledge all of those who have been part of this adventure.

I would first like to thank Prof. Hannu Hyyppä who attracted me into this eventful and memorable academic journey that has now resulted in this dissertation. None of this would have been possible without his support, advice, and encouragement. I also wish to thank my supervisor Prof. Matti Vaaja for his time and support, especially with concluding this dissertation process. I would also like to extend my gratitude to Prof. Henrik Haggrén, who was my supervisor at the beginning of this research work, for his inspiring and longlasting support that started already in the early years of my career when I was a young tech student. Special thanks to my advisors Prof. Juha Hyyppä and Dr. Juho-Pekka Virtanen for their support and excellent scientific mentorship throughout all these years. I also want to acknowledge my pre-examiners Prof. Fabio Remondino and Prof. Stephan Nebiker for their contribution to this dissertation. I am thankful to Prof. Remondino for agreeing to act as my opponent.

I must also thank all my colleagues who have helped me towards this achievement and who I have had an honor to navigate with in the sometimes-meandering maze of academic projects. Eventhough I am the sole author of this dissertation, this work is in part also their achievement and for that I am extremely grateful. I thank all my fellow co-authors: Kaisa Jaalama, Matti Kurkela, Toni Rantanen, Dr. Mikko Maksimainen, Prof. Harri Kaartinen, Prof. Antero Kukko, Dr. Johanna Ylipulli, and Dr. Matti Pouke. Also, I wish to thank especially Mattias Ingman, Aino Keitaanenniemi, Hannu Handolin, Juhani Talvela, Milka Nuikka, and Marika Ahlavuo who have helped with daily work activities, organizing field trips, and all sorts of problem solving over the years. It has been a great pleasure to work with you all.

The research work leading to this dissertation has been accomplished in a variety of projects. This work has been supported by the Academy of Finland, the Centre of Excellence in Laser Scanning Research (CoE-LaSR) and the Strategic Research Council project “Competence-Based Growth Through Integrated Disruptive Technologies of 3D Digitalization, Robotics, Geospatial Information and Image Processing/Computing – Point Cloud Ecosystem” (COMBAT), Business Finland project “VARPU”, the Aalto Energy Efficiency Research Programme, the Energizing Urban Ecosystems (EUE) -programme, the European Regional

Related to these projects I have had a chance of working with numerous public organizations, private companies, and individuals that deserve to be thanked. Particularly, I wish to thank Tommy Mård from the Finnish Broadcasting Company Yle, Henrik and Johan Gullmets, Prof. Roope Raisamo from Tampere University, and Jarmo Suomisto from the City of Helsinki for providing valuable assistance for my research. I have also been fortunate to receive personal support from the Jenny and Antti Wihuri Foundation, the Finnish Foundation for Technology Promotion, and the Foundation for Promoting the Field of Surveying.

It would be impossible to imagine this achievement without the invaluable support from all my friends who have stood by me in this endeavour. This achievement would have been immensely more difficult without all the counter-balancing activities and happy distractions along the way. Thank you Joonas, Joni, Anna, Aleksi, Jukka, Petteri, Jesse, Laura, Olli S., Lauri, Juho L., Henna, Juho A., Okko, Niko, Teemu, Matti, Mikael, Simo, Janne, Olli K., Panu, Hanna, Esko, Anssi, Jori, Felix, Juha, and all the rest of you who I anyway forgot to mention.

I am profoundly grateful to my family, my mother Marjaana, my father Kai, and my brother Otto for their unconditional support through my life in all its twists and turns. They have always believed in me and taught me curiosity and resiliency that has helped me incredibly in this journey.

Most importantly, I am forever thankful to my dear wife Elisa for your undying support and for being an everyday source of energy and sanity, and to our cheerful little Oskari who has really turned my life around and made me a person that I am today. Hence, I dedicate this work to both of you.

Espoo, 30 September 2021
Arttu Julin
Contents

Acknowledgements ........................................................................................................... i
List of Abbreviations ....................................................................................................... v
List of Publications ......................................................................................................... vii
Author’s Contribution ..................................................................................................... viii

1. Introduction ................................................................................................................. 1
   1.1 Background and motivation ............................................................................... 1
   1.2 Research gap ......................................................................................................... 5
   1.3 Research objectives, questions, and scope ....................................................... 7
   1.4 Research approach and dissertation structure ................................................. 10

2. Literature review ......................................................................................................... 13
   2.1 Photorealistic 3D modeling ............................................................................. 13
   2.2 Close-range 3D measuring ............................................................................. 19
   2.3 Platforms for photorealistic urban 3D models .............................................. 24

3. Materials and methods ............................................................................................. 29
   3.1 Multiple-case study .......................................................................................... 29
   3.2 Test sites ............................................................................................................ 30
   3.3 Instruments and datasets ................................................................................... 32
   3.4 Methods for data processing, evaluation, and analysis ................................ 33
   3.5 Supplementary qualitative methods ................................................................ 35

4. Results ......................................................................................................................... 37
   4.1 Platforms for photorealistic urban 3D models (Publication 1) .................... 37
   4.2 Efficient production of photorealistic urban 3D models (Publication 2) ........ 39
   4.3 Quality of photorealistic close-range 3D measuring (Publication 3) ............ 45

5. Discussion ..................................................................................................................... 51
   5.1 Platforms for photorealistic urban 3D models (Publication 1) .................... 51
5.2 Efficient production of photorealistic urban 3D models
(Publication 2) .................................................................................. 54

5.3 Quality of photorealistic close-range 3D measuring
(Publication 3) .................................................................................. 57

5.4 Future outlook ........................................................................... 62

6. Summary and conclusions .......................................................... 65

References ....................................................................................... 69
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D</td>
<td>Three-dimensional</td>
</tr>
<tr>
<td>ALS</td>
<td>Airborne laser scanning</td>
</tr>
<tr>
<td>API</td>
<td>Application programming interface</td>
</tr>
<tr>
<td>AR</td>
<td>Augmented reality</td>
</tr>
<tr>
<td>ASCII</td>
<td>American standard code for information interchange</td>
</tr>
<tr>
<td>BIM</td>
<td>Building information modeling</td>
</tr>
<tr>
<td>BRDF</td>
<td>Bidirectional reflectance distribution function</td>
</tr>
<tr>
<td>CAD</td>
<td>Computer-aided design</td>
</tr>
<tr>
<td>CSG</td>
<td>Constructive solid geometry</td>
</tr>
<tr>
<td>DSLR</td>
<td>Digital single-lens reflex camera</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic information system</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global navigation satellite system</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphics processing unit</td>
</tr>
<tr>
<td>HDR</td>
<td>High dynamic range</td>
</tr>
<tr>
<td>IFC</td>
<td>Industry foundation classes, a BIM model format</td>
</tr>
<tr>
<td>INS</td>
<td>Inertial navigation system</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>JPEG</td>
<td>Joint Photographic Experts Group, an image format</td>
</tr>
<tr>
<td>LDR</td>
<td>Low dynamic range</td>
</tr>
<tr>
<td>LiDAR</td>
<td>Light detection and ranging</td>
</tr>
<tr>
<td>MTF</td>
<td>Modular transfer function</td>
</tr>
<tr>
<td>NURBS</td>
<td>Non-uniform rational B-spline</td>
</tr>
<tr>
<td>OGC</td>
<td>Open Geospatial Consortium</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>PBR</td>
<td>Physically based rendering</td>
</tr>
<tr>
<td>RGB</td>
<td>A color model based on red, green, and blue</td>
</tr>
<tr>
<td>SLAM</td>
<td>Simultaneous localization and mapping</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>TLS</td>
<td>Terrestrial laser scanning</td>
</tr>
<tr>
<td>UAV</td>
<td>Unmanned aerial vehicle</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual reality</td>
</tr>
<tr>
<td>XR</td>
<td>Extended reality</td>
</tr>
</tbody>
</table>
List of Publications

This doctoral dissertation consists of a summary and of the following publications which are referred to in the text by their numerals:


Additional publications by the author that are relevant for the dissertation topic include Virtanen et al., 2018a; Virtanen et al., 2018b; Kurkela et al., 2020; Virtanen et al., 2020; Jaalama et al., 2021; Kurkela et al., 2021; Virtanen et al., 2021. They are cited where appropriate in this dissertation.
Author’s Contribution

**Publication 1:** Characterizing 3D City Modeling Projects: Towards a Harmonized Interoperable System

The author envisioned and wrote the manuscript with support from K. Jaalama and J.-P. Virtanen. M. Pouke, J. Ylipulli, and J.-P. Virtanen organized and participated in the interviews together with the author. J.-P Virtanen assisted in the multiple-case study and K. Jaalama and J. Ylipulli contributed to the interview analysis together with the author. M.T. Vaaja, J. Hyvppää, and H. Hyvppää participated in finalizing the manuscript with other authors.

**Publication 2:** Automated Multi-Sensor 3D Reconstruction for the Web

The author was mainly responsible for organizing the field campaign and conceptualizing and writing the article with help from K. Jaalama, J.-P. Virtanen, M. Maksimainen, and M. Kurkela. J.-P. Virtanen, K. Jaalama, M. Kurkela, and M. Maksimainen assisted with the research methodology and data analysis. J. Hyvppää and H. Hyvppää participated in finalizing the manuscript with other authors.

**Publication 3:** Evaluating the Quality of TLS Point Cloud Colorization

The manuscript was largely outlined and written by the author. Data collection and processing was done by the author with help from T. Rantanen. Conceptualization and data analysis was assisted by M. Kurkela, T. Rantanen, M. Maksimainen, and J.-P. Virtanen, who also helped with the study methodology and finalizing the manuscript together with M.T. Vaaja, H. Kaartinen, A. Kukko, J. Hyvppää, and H. Hyvppää.
1. Introduction

1.1 Background and motivation

3D models have become important and desired tools for managing, developing, and understanding globally expanding and densifying urban environments. These urban 3D models have been envisioned as platforms for collaboration and services, and as digital and virtual counterparts to real functioning cities. No strict definition exists regarding just what constitutes an urban 3D model, and many arguably overlapping and similar concepts, such as 3D city models (e.g., Kolbe et al., 2005; Gröger & Plumer, 2012; Biljecki et al., 2015), virtual cities (e.g., Batty et al., 2005; Hudson-Smith et al., 2007; Morton et al., 2012; Neuvillé et al., 2018; Saran et al., 2018), or more recently, digital twins (e.g., Ruohomäki et al., 2018; Ketzler et al., 2020; Schrotter & Hürzeler, 2020), essentially explaining the same phenomena have been widely and frequently used. In this dissertation, the term urban 3D model is used broadly to refer to a three-dimensional and digital model that represents a selected area in the urban environment in its entirety with elements such as buildings, vegetation, and infrastructure. Therefore, the focus of the dissertation is not limited to citywide information modeling of predefined sets of elements (most typically buildings), an approach commonly used in the 3D city modeling domain.

Urban 3D models are utilized by city organizations, citizens, academics, and businesses in a growing number of cities and nations all over the world. The increasing and diversifying number of urban use cases with advanced analyzes, visualizations, and simulations require realistic, detailed, and accurate depictions of the complex urban environments in their entirety. More efficient management and communication of urban information could yield better insights into and improve decisions that result in better performance, services, and financial savings (Bolton et al., 2018). Figure 1 shows examples of two different urban 3D modeling approaches.
3D measuring technologies enable the efficient production of photorealistic 3D models of real-world environments with high levels of detail and accuracy. Photogrammetry and laser scanning (also referred to as LiDAR) especially have become more efficient, accessible, and accurate techniques thanks to achievements in the development of computation and algorithms and the increasing affordability of sensors (e.g., cameras and laser scanners) and platforms (e.g., drones and backpack-style mobile platforms). The rapid technological development (e.g., Lévy et al., 2002; Fröhlich & Mettenleiter, 2004; Nistér, 2004; Hirschmüller, 2007; Kolb et al., 2010; Jancosek & Pajdla, 2011; Zhang, 2018; Qin & Gruen, 2021) has made it possible to create and present virtual 3D environments on a significantly more realistic, immersive, and efficient level than could be achieved previously. In the urban context, 3D measuring techniques have frequently been used to produce photorealistic 3D models at a wide range of scales, qualities, and levels of detail, ranging from massive citywide 3D datasets of simple block models to highly detailed, photorealistic-textured, 3D mesh models or colorized 3D point clouds with high verisimilitude, but typically limited regional coverage. Photographs and LiDAR scans have been frequently and variously used as an input data for a diverse set of automated (e.g., Musialski et al., 2013) and manual (e.g., Yin et al., 2008) urban modelling approaches. Over the years, modeling automation has been an avid research topic and a persisting challenge. An extensive overview of urban 3D modeling approaches can be found, for example, in Musialski et al. (2013).

The present dissertation focuses on close-range photogrammetry and terrestrial laser scanning (TLS) techniques, which can achieve millimeter-level accuracies and detail compared to mobile or airborne mapping methods. Mobile mapping methods that rely on GNSS-, INS-, and SLAM-based positioning can typically reach centimeter-level accuracies and detail, whereas airborne methods are capable of producing decimeter-level 3D data. A comparison of different 3D measuring techniques and scales at which they are typically applied in urban 3D modeling is illustrated in Figure 2.
Introduction

Figure 2. Typical scales of various 3D measuring techniques in the context of modeling urban environments.

Utilizing close-range 3D measuring techniques can result in significant savings in time and money and enable the use of detailed photorealistic 3D models in completely new use cases, especially when compared to traditional and often dissimilar workflows, where the detailed 3D content is produced more or less manually by 3D artists, designers, or other professionals. 3D modeling of complex environments like cities for demanding use cases, such as for web applications, has generally been viewed as inefficient, expensive, and even a barrier to entry for developers and end users.

Photorealistic 3D models of real-world environments are becoming increasingly important elements in applications, services, digital twins, and virtual and augmented reality (VR/AR) experiences created for numerous fields, including urban planning, architecture, archaeology, construction, tourism, digital games, education, cultural heritage, the real estate business, and various industry sectors. A realistic visual model appearance is a key aspect in many fields, such as cultural heritage (e.g., Gaiani et al., 2017), archaeology (e.g., Pepe et al., 2016), architecture, tourism, and entertainment. Real-world 3D data is increasingly being applied in visually extremely demanding use cases, such as in digital games (Statham, 2018) or in VR/AR applications (Çöltekin et al., 2020). Most of these diverse use cases for 3D models are centered around their ability to powerfully and realistically visualize the real world with a high level of detail and accuracy, for example to communicate information, analyses, and designs and to demonstrate the impact of decisions. In this dissertation, photorealistic 3D models are used to refer to 3D models that aim to be as indistinguishable from the real world as possible. This is analogous to the definition of perceptual realism (McMahan, 2003). The production of these 3D models requires highly detailed real-world 3D data that is both visually and dimensionally accurate. Thus, these models must have a high degree of detail and accuracy from
both the perspective of 3D geometry and model appearance. In the dissertation, the 3D model appearance is defined as non-geometric information that is captured by close-range 3D measuring instruments and applied to the dimensionally accurate 3D geometry. More specifically, the model appearance can be represented by color information in a colorized 3D point cloud and as texture images in a 3D mesh model. Aspects of photorealism and their main representations in relation to a detailed urban 3D model are presented in Figure 3.

The 3D model itself does not solely guarantee utilization and usefulness, but the **application platform** plays a significant role in putting the urban 3D model into beneficial use. Currently, utilization is complicated by a great number of various types of application platforms, ranging from professional GIS and CAD software to real-time 3D platforms like game engines and web-based virtual globes. In this dissertation, these software solutions are primarily understood as application platforms, which are essential to the development and creation of a 3D model application and for defining its functionalities and characteristics. The term real-time 3D platform is used to refer to both game engines and browser-based systems like virtual globes that utilize real-time rendering of 3D graphics and appear similar from the perspective of 3D model production, such as with their wide support for textured 3D mesh models.

The key concepts related to the dissertation are illustrated in Figure 4.
1.2 Research gap

Utilization of 3D models remains one of the key challenges in urban 3D modeling, a challenge that has often been acknowledged (e.g., Neuville et al., 2018). 3D data collection and modeling processes and suitable application platforms together establish a technical foundation for utilizing urban 3D models. Currently, the application of urban 3D models is complicated by the great variety of different software solutions, ranging from GIS and CAD software (e.g., Zlatanova et al., 2002) to real-time 3D platforms like web-based virtual globes (e.g., Blaschke et al., 2012; Loesch et al., 2012; Krämer & Gutbell, 2015) and game engines (e.g., Fritsch & Kada, 2004; Greenwood et al., 2009). Few studies have focused on forming an overview of the technical solutions used in real-life projects beyond reporting individual experiments and demonstrations. Furthermore, no recent study exists that adequately analyzes the functionality and usability aspects of these urban 3D modeling activities, including the application platforms that determine or limit the features of a 3D model application from the perspectives of, for instance, the supported 3D data types and qualities, level of accessibility, interaction, interoperability, user experience, and visualization.

The utilization of urban 3D data is often viewed and developed from the perspective of a city or other public organization, and very few scholars have acknowledged the social aspect of urban 3D modeling and explained the motivations behind developing these applications. Over the years, urban 3D models have been made openly and freely available for public use (e.g., Ruohomäki et al., 2018; Schrotter & Hürzeler, 2020) in part to encourage the development of more specialized urban applications and utilization outside city organizations and to expand the economic benefits of urban 3D modeling. The general movement towards open distribution of urban 3D model data underlines the need to ensure the applicability of the 3D data beyond internal use cases within city organizations and also the need to motivate and encourage larger developer communities to create new applications.
A significant number of the benefits of urban 3D models are achieved through efficient visualization of the often-complex urban scenery to communicate information, analyses, designs, and decisions. Yet, a great share of city modeling literature over the last decade has focused on semantic city modeling, which has especially been driven by the development and application of the open CityGML standard (e.g., Kolbe et al., 2005; Stadler et al., 2009; Gröger & Plumer, 2012). Therefore, the focus has not typically been on photorealistic model appearance but on often-oversimplified semantic driven or non-photorealistic visualization (e.g., Döllner & Buchholz, 2005; Glander & Döllner, 2009; Neuville et al., 2018) of a selected set of urban elements most typically limited to buildings. A more human-scale perspective, one containing indoor spaces or streetscapes, the complex city environments most commonly experienced by people, have rarely been included in common urban 3D modeling approaches. Moreover, using CityGML for efficient visualization purposes has been seen as overly complex for developers, which has led to the introduction of more developer-friendly encoding specifications, perhaps most notably CityJSON (Ledoux et al., 2019).

Photorealistic applications have been considered difficult to efficiently produce, but the recent advances in 3D measuring, modeling, and computer graphics are lowering the barrier for production and use of increasingly realistic urban 3D models. The increased level of detail and realism could enable new, and enhance existing, use cases for photorealistic 3D models. For example, a more accurate and detailed geometry and model appearance could improve the accuracy and reliability of spatial analyses (similar to that reported by Willenborg et al., 2018) and the effectiveness of urban 3D model visualizations (e.g., Jaalama et al., 2021). The significance of high levels of detail for immersion into and the realism of the virtual environment and the challenge of modeling automation have been acknowledged (e.g., Walmsley & Kersten, 2019). To conclude, there exists a need for further research on how to apply close-range 3D measuring techniques for the flexible and efficient production of photorealistic 3D models.

The more detailed and accurate the model’s geometry and appearance, the more indistinguishable it can be from reality (e.g., Shiode, 2000). The focus on using 3D measuring technologies for city applications has typically been on applying airborne laser scanning (ALS) (e.g., Lohani & Ghosh, 2017) and often existing building footprint data to produce large-scale citywide 3D models of buildings (e.g., Tomljenovic et al., 2015; Nys et al., 2020) and terrain. Another popular approach has been to utilize aerial photography and photogrammetric modeling techniques (e.g., Rothermel et al., 2012; Rupnik et al., 2014) to produce more realistic citywide, textured 3D mesh models, similarly to the popular 3D cities offered in map services by Google (Google, 2021a), Apple (Apple, 2021), and Microsoft (Microsoft, 2021a). These large-scale airborne LiDAR- and image-based techniques have been also used increasingly in combination (e.g., Mandlburger et al., 2017; Zhang & Lin, 2017; Toschi et al., 2018) and on a smaller scale using unmanned platforms (e.g., Haala et al., 2020). Both typically citywide LiDAR or image-based approaches result in extreme data volumes with...
Spatial accuracies ranging from decimeters to meters and do not aim to reconstruct the city on a street level, where city environments are typically experienced by people. Despite being efficient and highly automated data collection methods, data collection on a citywide scale is usually only possible at longer, annual-level time cycles, compared to more localized and flexible close-range 3D data collection approaches. The move to understand complex and dynamic city environments cannot rely solely on bird’s-eye views, and realistically 3D modeling those environments can benefit from more extreme levels of detail and flexible, more localized data collection approaches. Close-range photogrammetry and TLS have the potential to efficiently produce 3D data at a high level of accuracy and detail, even on a 1:1 scale, and to enable new benefits for urban 3D models. The complementary benefits of integrating close-range photogrammetry and TLS have frequently been noted in the existing literature (e.g., Ramos & Remondino, 2015), and multi-sensor integration has been realized using commercially available (and potentially more easily applicable) solutions. However, there is a need to gather more knowledge on the quality that these integrated modeling approaches achieve compared with the results obtained using individual methods.

Being well-established 3D measuring methods, the quality of photogrammetry and laser scanning has mainly been studied from the perspective of reconstructing 3D geometry. However, surprisingly little attention has been paid to assessing the quality of 3D measuring techniques from the perspective of the texturing and point colorization that reflects the appearance of the 3D model, and thus, the achievable degree of photorealism. This gap in the literature is especially notable in relation to modern TLS instruments that utilize integrated cameras to produce photorealistic 3D point cloud data. Thus, more comprehensive knowledge is needed about the quality of the image data provided by the 3D measuring techniques when data is collected for photorealistic purposes.

1.3 Research objectives, questions, and scope

The objective of this dissertation was to develop ways to improve the utilization of close-range photogrammetry and TLS for photorealistic urban 3D modeling. To fulfill its objective, the dissertation addresses the following three research questions:

RQ1: What are the relevant application platforms for utilizing photorealistic urban 3D models?

RQ1 approaches the dissertation topic on an application level and focuses on reviewing real-life urban 3D modeling activities via a multiple-case study and identifying the relevant application platforms for utilizing photorealistic 3D models. The activities are studied in the six largest cities in Finland: Helsinki, Espoo, Tampere, Vantaa, Oulu, and Turku. Furthermore, as a supplementary approach, semi-structured interviews with city authorities and experts focus on their internal drivers and factors to better understand the expectations behind
applying urban 3D models. This approach increases understanding of the impacts of software platforms as an integral part of utilizing urban 3D models. The application platforms are important for putting the 3D models to use and also for determining and limiting the features (e.g., support for 3D data types, user experience, visualization capabilities, accessibility) of the application.

**RQ2:** How can close-range 3D measuring be utilized in efficient model production for photorealistic urban 3D models?

RQ2 approaches the dissertation topic on a modeling level and makes it possible to experiment with the automated production of photorealistic and dimensionally accurate urban 3D models using close-range photogrammetry, TLS, and a combination of them in a real-life project setting as well as evaluate the modeling results from the perspective of geometric and texturing quality. Furthermore, the experiment is based on the real-life joint project “Puhos 3D” by Aalto University and the Finnish national public broadcasting company Yle (Yle, 2021). More specifically, the focus is on exploring the production of a photorealistic 3D model for a web-based, real-time 3D Sketchfab platform (Sketchfab, 2021a) with strict performance requirements. In addition to the quantitative geometry and texture quality evaluation methods, a supplementary online survey was organized to compare the modeling approaches.

**RQ3:** How can the quality of a close-range 3D measuring technique be evaluated for photorealistic applications?

RQ3 approaches the dissertation topic on a sensor level and focuses on how best to develop a method for evaluating the quality of TLS point cloud colorization. The method applies and combines established image quality metrics and analyzes the TLS instrument’s capability for reproducing color and details in the scene. In addition, the method is tested and the results compared for benchmarking purposes with four different commercial TLS instruments.

The research scope focuses on photorealistic urban 3D models and applications that utilize and visualize these models. From the perspective of urban 3D data, the focus is on the model’s geometry and appearance. Other aspects of the model, such as its semantics and topology, do not fall within the scope of the dissertation. With respect to the model’s appearance, the focus is on attaining non-geometric information from the whole scene using camera sensors, while defining the physical and material properties of the scene, a common task in the computer graphics field related to realistic rendering, is not within the scope of the dissertation. In relation to 3D model geometry, the focus is on polygonal mesh models and 3D point clouds. Therefore, the main representations of the model’s appearance are texture images and color values collected with camera sensors.

When producing and utilizing urban 3D models, it is important to take into account not only the 3D data, but also the application platforms that enable the use of this data. As an application platform, we considered software solutions

---

**References:**


that present the urban 3D model to its users and enable its utilization. These include various GIS and CAD solutions and real-time 3D platforms, such as game engines and web-based virtual globes. Furthermore, the study focuses more on analyzing platform categories than on singling out specific software solutions since the software field is evolving rapidly and undergoing constant change. Related to the other software solutions, the various data management systems (e.g., 3DCityDB, Oracle Spatial, PostGIS) that specialize in the aspects of managing and hosting urban 3D model data are outside the scope of the dissertation. Due to the great number of urban 3D modeling activities all around the globe, the multiple-case study and the supplementary interviews were limited to the six largest cities in Finland.

With respect to data collection, the focus is on close-range photogrammetry and terrestrial laser scanning (TLS), both of which enable data acquisition in urban environments, including even indoor spaces, at a very high level of detail and accuracy (see Figure 2) from measuring ranges up to a few hundred meters. The level of detail and accuracy achieved with them is potentially higher than with mobile mapping and SLAM-based technologies, which often rely on similar cameras and LiDAR sensors, or airborne data collection. For example, TLS is often used to produce reference datasets when the geometric quality of other 3D measuring techniques is evaluated. Apart from the model's geometry, both close-range photogrammetry and TLS commonly utilize camera sensors and are thus well suited for collecting information about the appearance of the target. Close-range photogrammetry and TLS can also be used to complement each other and can also be significantly more cost efficient and practical to experiment with compared to citywide data collection methods involving ALS and aerial photography. Despite being beyond the scope of this dissertation, current use of unmanned aerial vehicle (UAV-)based photogrammetry can be seen in many ways as analogous (e.g., similar sensors and data processing workflows) and often complementary to close-range photogrammetry with the main difference typically having to do with perspective and longer measuring ranges. From the modeling perspective, emphasis is placed on the photorealistic modeling of existing (or as-built) information in the urban environment, while as-planned information, such as IFC models from building information modeling (BIM) processes, are not included within the scope of this dissertation. Also, since the focus is on efficient model production, the diverse and dissimilar manual modeling workflows based on 3D modeling suites often related to CAD and 3D GIS are excluded. Other aspects often linked to 3D measuring and modeling, such as metric sensor calibration or georeferencing the model data, are also not within the scope of this dissertation.

Overall, the dissertation focuses on utilizing real-life cases, approaches, and methods, and it evaluates their practical implications. To underline their reproducibility and practical applicability, the dissertation focuses on using software and hardware solutions that are commercially or openly available. All the instruments and software are used as is in the experiments and without modifications.
1.4 Research approach and dissertation structure

The research aspect of this dissertation focuses mainly on conducting experiments and on a quantitative analysis of the acquired empirical results. In addition, a set of supplementary qualitative methods (i.e., semi-structured interviews and an online survey) are used to collect complementary and explanatory data about phenomena that are not solely technical in nature. For example, cities can be considered socio-technical systems (e.g., Hillier, 2009); to understand the processes taking place and affecting their virtual and digital counterparts, it is relevant to also examine the societal aspects related to developing and using those models. Additionally, a supplementary online survey is used to assess the perceived quality of the compared models, which benefits from applying qualitative methods (McMahan, 2003) as well as using solely quantitative approaches. Furthermore, this work emphasizes the practical applicability of the 3D model by focusing on real-life use cases and environments and using commercially or openly available instruments and solutions. Therefore, the philosophical underpinnings of this dissertation are grounded in pragmatism, where the role of research is primarily underlined as a practical problem-solving instrument.

The research process in this dissertation is designed to spotlight the multidisciplinary topic sequentially in three publications (illustrated in Figure 5), with the corresponding research questions discussed in section 1.3 and summarized in Table 1.

![Figure 5. An illustration of the applied sequential research process.](image)

The dissertation consists of a summary and three original publications. To fulfill the research objective, the three research questions, RQ1, RQ2, and RQ3 are answered in turn in Publications 1, 2, and 3. The original publications included in the dissertation follow a sequential structure (see Figure 5), where the scope and objectives of each publication is narrowed down and follows from the previous publication. The relationship between the research questions, the publications, and the objectives for each part of the dissertation are presented in the following Table 1.
Table 1. Relationship between the research questions, publications, and research approaches in the dissertation.

<table>
<thead>
<tr>
<th>Research question</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1: What are the relevant application platforms for utilizing photorealistic urban 3D models? (Publication 1)</td>
<td>Reviewing real-life urban 3D modeling activities in a multiple-case study to identify relevant application platforms. Conducting supplementary semi-structured interviews with city authorities and practitioners to understand the internal motivators and barriers to applying photorealistic urban 3D models.</td>
</tr>
<tr>
<td>RQ2: How can close-range 3D measuring be utilized in efficient model production for photorealistic urban 3D models? (Publication 2)</td>
<td>Experimenting with automated multi-sensor modeling for a web-based, urban 3D model application in a complex urban environment. Evaluating the quality of the model’s geometry and appearance using quantitative approach and a supplementary online survey.</td>
</tr>
<tr>
<td>RQ3: How can the quality of a close-range 3D measuring technique be evaluated for photorealistic applications? (Publication 3)</td>
<td>Developing a method for evaluating the quality of TLS point cloud colorization and experimenting with its use in benchmarking commercial TLS instruments in a laboratory test environment.</td>
</tr>
</tbody>
</table>

This summary consists of six sections. This introductory section is followed by section 2, which reviews the relevant literature related to the dissertation topic. The materials and methods used in this research study are covered in section 3. The results of this dissertation are presented in section 4. The results, implications, and future outlooks are discussed in section 5. Finally, a summary and conclusions are presented in section 6.
2. Literature review

In this section, the relevant background theory for the dissertation is presented. It covers the aspects of photorealistic 3D modeling and introduces the topics of TLS, close-range photogrammetry and their integration, including the topic of application platforms for photorealistic urban 3D models.

2.1 Photorealistic 3D modeling

No single definition exists for what makes a 3D model photorealistic. For example, in the computer graphics domain the concept of photorealism has frequently been referred to in rendering (or image synthesis) where the aim is to create (render) images that aim to be indistinguishable from photographs (e.g., Ferwerda, 2003). In practice, computer graphics offer many techniques for visualizing the real world based on 3D data produced with photogrammetry and laser scanning. Generally, photorealism can be understood as how closely the objects or environments match those that exist in the real world. The degree of reality depends on the number of details conveyed by the 3D model (e.g., Shiodo, 2000). Compared to other photorealistic digitization methods, such as traditional or 360-degree photography, the 3D model-based approaches allow for a full six degrees of freedom (6DoP) of movement for the observer in a 3D space. Conversely, exploring image-based models in applications, such as Google Street View (Anguelov et al., 2010), typically require limiting transition techniques (e.g., image warping or fading) when moving between discrete photographs (Nebiker et al., 2010).

2.1.1 Model geometry

The 3D geometry of a 3D model can be digitally represented using surfaces, points, and volumes or by applying other descriptions. Points are commonly represented using 3D point clouds, surface representations include polygonal meshes (Baumgart, 1972; Smith, 2006), and parametric representations contain such approaches as non-uniform rational B-splines (NURBS) (Dimitrov & Golparvar-Fard, 2014). Volumetric representations include voxels (Nourian et al., 2016) and constructive solid geometry (CSG) (Requicha & Voelcker, 1983). Voxels are cuboid volume elements that can be generated from 3D point clouds and mesh models with voxelization methods (Nourian et al., 2016).
**Polygonal meshes** are among the most popular representations of 3D geometry, especially due to their wide support by modern graphics hardware, namely the graphics processing units (GPU) used in desktop and mobile computing platforms. The mesh model consists of vertices, edges, and faces. The faces are most typically triangles (triangle mesh), but quadrilateral (quad mesh) faces are often also manually created or automatically generated from triangles using remeshing approaches because of their potential advantages in, for example, the polygonal modeling of objects and characters, as well as in texturing (Bommes et al., 2013). In addition to the 3D coordinates, other information, such as vertex colors and normals, can be stored into the model’s vertices. Numerous mesh generation algorithms exist, such as marching cubes (Lorensen & Cline, 1987) and Poisson reconstruction algorithms (Kazhdan et al., 2006), and they have been implemented in numerous commercial and open software solutions, such as Sequoia (AWS Thinkbox, 2021), CloudCompare (CloudCompare, 2021), and Meshlab (Cignoni et al., 2008). The production of textured mesh models is commonly integrated into modern photogrammetric 3D modeling pipelines (e.g., AliceVision, 2021), but the implemented algorithms are commonly proprietary. Nocerino et al. (2020) have presented examples of how to evaluate the geometric quality of mesh-based surface reconstruction and provided an overview of surface assessment criteria. In practice, the level of detail of the 3D geometry is dependent on the density of the vertices (or polygons) in the model. The mesh models generated from 3D measured data can be quite complex and heavy to visualize and often require simplification (decimation) (Kobbelt et al., 1998) to be useful for an application. With simplification, the vertex (or polygon) count of a 3D model is reduced to meet a certain vertex (or polygon) budget depending on the performance requirements of the use case. Common file formats for polygonal meshes include OBJ (Wavefront OBJ), FBX (Filmbox), PLY (Polygon File Format or Stanford Triangle Format), and glTF (GL Transmission Format). Mesh models are often textured to enhance their detail and visual appearance, especially in photorealistic applications (Figure 6).
3D point clouds have become an increasingly relevant form of representing 3D geometry due to their simplicity and production efficiency. Point clouds consist of a large number of X, Y, and Z data points with numerous possible additional attributes, such as RGB color information derived from the imagery or information about point classification. Point clouds can be produced using various methodologies, such as measuring directly from a target with laser scanning or deriving the 3D coordinates from images using photogrammetric dense image matching. The level of detail of the 3D geometry of a point cloud is highly dependent on its point density. 3D measuring techniques enable the efficient collection of point clouds but processing them can be a challenging task due to such typical problems as heterogeneous point density and possible holes, noise, outliers, and misalignments in the data (Berger et al., 2017). 3D point clouds are often used directly and many of its aspects from visualization (e.g., Nebiker et al., 2010; Virtanen et al., 2020) to data management (e.g., van Oosterom et al., 2015) have been frequently studied. Direct application of 3D point clouds is still a challenge, and the data is often processed into other representations such as 3D mesh models using meshing (or triangulation) techniques to improve their applicability. Common 3D point cloud file formats include ASCII or binary formats like LAS (ASPRS, 2011) or E57 (Huber, 2011). One example of a photorealistic colored 3D point cloud is shown in Figure 7.

Figure 6. An example of a textured 3D mesh model with its 3D geometry represented as triangle mesh and its appearance represented as automatically generated texture images. The model was produced using close-range photogrammetry.
2.1.2 Model appearance

The appearance of a photorealistic 3D model is commonly represented with color values or texture images. Color values can be connected to the 3D points of a point cloud or to the vertices of a 3D mesh model and are most commonly described as 8-bit RGB values in a sRGB color space (IEC, 2021).

Texture mapping, or texturing, is a process of defining a texture image on the surface of a 3D model (Yuksel et al., 2019), typically to improve its appearance by incorporating surface colors, details, and materials. Texture images can be based on real photographs or created artificially with various procedures, such as digital painting. UV mapping is a process where each 3D coordinate of a mesh vertex is assigned a 2D coordinate in an UV texture space. This can be done either automatically (e.g., Lévy et al., 2002) or manually using various software tools. UV mapping requires unwrapping, which is a process where the 3D geometry of a mesh model is projected onto a planar surface. This can be done using mesh parts, and to reduce texture size, the texture images that represent parts of the mesh can be aggregated into texture atlases (see Figure 6) that contain information about all the faces in the mesh (Pagés et al., 2010). Another strategy for reducing the storage and memory requirements is to exploit the symmetries or tiling so that the same texture data can be used over multiple parts of the model (Yuksel et al., 2019). An overview of alternative texturing methods can be found in a study by Yuksel et al. (2019). The RGB color values of a texture are stored into texels (texture elements), which are essentially pixels in the texture images (or texture atlases) that are represented by standard image file formats, such as JPEG or TIFF.

Apart from color, textures can be used to store other shading parameters over 3D surfaces (Yuksel et al., 2019). With physically based rendering (PBR), mul-
titexturing-based approaches are used where more than one texture is used simultaneously to represent material properties, such as surface albedo (i.e., illumination independent surface color), roughness, and metalness. The real-time PBR approaches applied in modern game engines commonly aim to achieve photorealism by using simplified approximations of bidirectional reflectance distribution functions (BRDF) and render equations to model the extremely complex topic of how light flows in the scene and behaves with material surfaces (Pharr et al., 2016). Related to texturing, normal mapping is a popular technique that can be used to reduce the number of polygons needed to convey the same level of detail in the model (Apollonio et al., 2021). It is commonly used to enhance the appearance and details of a simplified low polygon model by utilizing a normal map generated from a high polygon model.

**Appearance modeling** is a well-established topic that addresses a fundamental problem in computer graphics, and it is often related to how different materials in the scene are rendered or simulated under different lighting and viewing conditions (Müller, 2008; Dorsey et al., 2010; Dong, 2019). This goal is related to physical realism (Ferwerda, 2003) and has been pursued with state-of-the-art rendering techniques, such as real-time PBR and real-time ray tracing (NVIDIA, 2021), and with specialized devices that are capable of measuring reflection properties (e.g., Wang et al., 2013). Related to 3D measuring, a similar distinction has been made between geometric and appearance modeling in the field of cultural heritage (e.g., Remondino & Rizzi, 2010), where experts have great appreciation for realistic visual appearances. The examples found in the 3D documentation of cultural heritage have typically focused on individual objects, such as producing digital replicas of museum artifacts (e.g., Berrier et al., 2015; Nöll et al., 2015; Apollonio et al., 2021). Many of these approaches rely on controlled lighting and utilizing highly customized hardware and software systems (e.g., Nöll et al., 2015). From the perspective of quality, the concept of model appearance can be ambiguous and has been often referred to as the radiometric quality of a 3D model (e.g., Gaiani et al., 2016), especially in the remote sensing domain.

In the context of 3D measuring, the model’s appearance is formed from the radiance that is captured by the sensor instrument. In the case of photorealistic 3D models, this is achieved with a camera sensor that is integral to (i.e., in photogrammetry) or integrated with (i.e., in laser scanning) the methodology. This radiance is affected by a complex combination of factors, such as the illumination, geometry, and specular and diffuse reflectivity of the target. Regarding photorealistic 3D models, the model’s appearance can be understood as color and tone (lightness) information that is then connected to a 3D point in a point cloud or to a pixel in a texture image, and it is commonly described as red, green, and blue values of the RGB color model. Furthermore, the sRGB is by far the most widely used and supported color space standard due to its consistency among various devices (e.g., computers, monitors, mobile devices, and cameras), the Internet, and common 3D graphics programming interfaces, such as WebGL (Khronos Group, 2021a), OpenGL (Khronos Group, 2021b), and Direct3D (Microsoft, 2021b). On their own, the RGB color values that represent
the appearance of a 3D model do not describe only the colorfulness but also the differences in luminance and in color (i.e., contrast), which can be visually perceived as details in the scene. Thus, from the perspective of the model’s appearance, the quality of a photorealistic 3D model can be understood as the measuring instruments’ capability of reproducing color and detail from the scene. Since this reproduction is enabled, de facto, by camera sensors, the quality of the model’s photorealistic appearance can be approached from the perspective of image quality assessment.

2.1.3 Image quality assessment

A great number of image quality studies have tested digital cameras (e.g., Loebich & Wueller, 2001; Wueller, 2006) or more specific types of devices and systems, such as mobile phone cameras (e.g., Jin, 2008; Peltoketo, 2014), digital aerial imaging sensors (e.g., Honkavaara et al., 2008; Dąbrowski & Jenerowicz, 2015; Orych, 2015), and 360-degree cameras (e.g., Peltoketo, 2016a; Yang et al., 2017). On the whole, image quality is the result of a complex combination of quality factors inherited from the lens, imaging sensor, and image processing pipeline. Image quality can be assessed subjectively via various perceptual methods that assess human subjects (Wang & Bovik, 2006) or by analyzing the image data using various automated image-quality measurement techniques. Many studies and numerous standards address the issue of objective quality assessment, such as ISO 12233 (ISO, 2021a) for measuring resolution and spatial frequency response, ISO 15739 (ISO, 2021b) for measuring the amount of noise, and ISO/CIE 11664-6:2014 (ISO/CIE, 2021) for assessing color reproduction, all of which define a wide range of quality metrics that are typically calculated from a variety of test charts (see, e.g., Figure 13) using such software solutions as Imatest Master (Imatest, 2021a) or iQ-Analyser (Image Engineering, 2021).

The quality of a camera system cannot be depicted with a single metric. However, the color reproduction, sharpness, and noise of a system have been considered the most important metrics of image quality (e.g., Peltoketo, 2014). First, color reproduction indicates the capability of the imaging sensor to reproduce colors in the scene, which can be understood as an accurate color description consisting of a combination of luminance and chromatic components. Color reproduction can be quantified with color accuracy measurements, such as by calculating total color difference (e.g., Sharma et al., 2015) from a color reference chart like the ColorChecker. Second, sharpness describes the amount of detail the imaging sensor is able to capture. It can be quantified using resolution measurements, most notably the modular transfer function (MTF), which defines the capability of an imaging sensor to maintain the optical contrast of increasingly finer details in a targeted scene (Peltoketo, 2016b). MTF can be measured using test charts like Siemens stars, and it has been interchangeably referred to as a spatial frequency response (SFR) in the literature. In addition, other metrics such as Shannon information capacity (Koren, 2020a) can be used to assess the imaging quality. It is derived from Shannon’s original information capacity theory (Shannon, 1948) and relies on the hypothesis that the image
sharpness (i.e., MTF) and noise correlate with the information capacity and that it corresponds with the perceived image quality. Finally, noise can be understood as an unwanted random spatial variation in an image that obscures details. The degree of noise can be assessed by measuring an image’s signal-to-noise ratio (SNR).

2.2 Close-range 3D measuring

2.2.1 Terrestrial laser scanning

TLS is a ground-based and close-range 3D measuring technique that is based on laser scanning (also called LiDAR), where the coordinates are measured directly from a target in 3D space using the laser ranging principle (Shan & Toth, 2018). Laser scanning is an active 3D measuring technique that sends and receives its own energy in the form of a laser beam and allows for measuring 3D information independently of scene illumination and texture. TLS scanners measure the distance to the target using two main principles: time-of-flight (ToF) or phase-shift (phase comparison) methods. The ToF method is based on measuring the time taken by the laser pulse to reach the target surface. The phase-shift method is based on determining the distance to the target using the phase shift between the transmitted and received modulated continuous waves. TLS scanners use mirrors to be able to scan the surrounding environment. Modern TLS scanners are typically panoramic and designed to cover their surroundings to a full 360-degree dome and at distances typically up to a few hundred meters. Also, camera-type scanner instruments exist that are designed to scan much more limited angular areas within a specific field of view (FOV) (Shan & Toth, 2018). Modern TLS instruments are capable of recording millions of 3D observations per second that together form a 3D point cloud of the target environment. TLS is typically performed from a static platform, that is, a tripod. However, apart from their typical intended use, TLS instruments have been applied for mobile mapping purposes (e.g., Kukko et al., 2012). Over the years, TLS instruments have become more capable 3D measuring devices via such advancements as increasing measuring rates, better portability, affordability, and user-friendliness.

To record more comprehensive information about the surrounding environments, many modern TLS instruments use integrated cameras to colorize the 3D point cloud (see an example in Figure 8). To potentially increase the flexibility and the imaging quality (especially in challenging lighting conditions), many scanner manufacturers, such as Faro, Leica Geosystems, Trimble, and Zoller + Fröhlich, have added high dynamic range (HDR) imaging capabilities to their TLS systems (e.g., Faro, 2021; Leica Geosystems, 2021a; Trimble, 2021a; Z+F, 2021). HDR is a widely used photographic imaging technique where several images at different exposure times are combined to produce an image of a greater dynamic range (e.g., Pourreza-Shahri & Nasser Kehtarnavaz, 2015). Further, to increase the speed of the imaging process during data acquisition, some scanner
manufacturers have added more camera components to the scanner frame instead of relying on the traditional approach of using a single camera sensor mounted coaxially with the laser to minimize the parallax effects in point cloud colorization. The detailed imaging specifications of these commercial instruments and the related processing pipelines and colorization algorithms have rarely been disclosed.

Figure 8. (a) Laser scanning urban environment with a TLS instrument; and (b) an example of a colored 3D point cloud acquired from an urban environment.

TLS is a well-established and versatile 3D measuring technique that can produce highly detailed and colored 3D point clouds of real-world environments for numerous applications. For example, TLS has been applied to agriculture (Lumme et al., 2008), archaeology (Lerma et al., 2010), as-built modeling for architecture (Sternberg, 2004), building information modeling (Wang et al., 2019), city modeling (Dold & Brenner, 2004; Pu & Vosselman, 2009; Akmaliaa et al., 2014), construction and facility management (Yuan et al., 2020), cultural heritage (Abmayr et al., 2005; Guarnieri et al., 2006), geology and geomorphology (Buckley et al., 2008; Abellán et al., 2014), forestry (Bienert et al., 2006; Liang et al., 2008; Yrttimaa et al., 2019), structural inspection and the monitoring of bridges (Riveiro et al., 2011), dams (Alba et al., 2006), hydropower stations (Schäfer et al., 2004), tunnels (Nuttens et al., 2010), roads (Hassan et al., 2020), industry (Sternberg & Kersten, 2007; Kawashima et al., 2014; Stenz et al., 2020), and land surveying (Pinkerton, 2011). Also, TLS-derived point clouds have commonly been used as reference datasets for various quality verification purposes (Vaaja et al., 2011; Lehtola et al., 2017).

Quality of TLS instruments has frequently been studied throughout their existence. In general, the focus has been on studying the geometric aspects of the quality of the resulting point cloud data. From the perspective of geometry, the quality of a point cloud is affected by the precision and accuracy of the laser scanner (e.g., Gordon et al., 2000; Lichti et al., 2000; Boehler & Marbs, 2003; Staiger, 2003; Mechelke et al., 2007; Pfeifer & Briese, 2007; Wunderlich et al., 2013; Schmitz et al., 2019). Regarding 3D point clouds, the level of identifiable
detail is influenced by scan resolution and disrupted by undesirable errors induced by, for example, edge effects (e.g., Boehler & Marbs, 2003; Wunderlich et al., 2013; Lichti & Jamtsho, 2006; Ling et al., 2008; Pesci et al., 2011). The target surface geometry and its material properties, such as the color, texture, and reflectivity, impacts the geometric quality of the measured 3D point cloud (e.g., Gordon et al., 2000; Boehler & Marbs, 2003; Mechelke et al., 2007; Clark & Robson, 2004; Kersten et al., 2005; Soudarissanane et al., 2007; Voegtle et al., 2008). Moreover, environmental conditions (Borah & Voelz, 2007) and the scanning geometry (Soudarissanane et al., 2011; Kawashima et al., 2014) influence the geometric quality of the 3D scanned data. TLS is generally considered a 3D measuring technique with a potentially high degree of accuracy and precision. It has often been used to produce reference data for evaluating other 3D modeling approaches, such as photogrammetry (e.g., Remondino et al., 2014).

From the perspective of radiometric quality, the measuring, calibrating, and utilizing of TLS intensity information has been investigated with notable frequency (e.g., Bucksch et al., 2007; Pfeifer et al., 2007; Kukko et al., 2008; Kassalainen et al., 2009; Krooks et al., 2013; Tan et al., 2015). Furthermore, the relationship between the TLS intensity and colorimetric data of a color chart has been studied (Balaguer-Puig et al., 2017), and the influences of different materials and colors on the intensity values of TLS has been investigated (Voegtle et al., 2008). Besides focusing on TLS intensity, Hassan et al. (2017) investigated the ability of the RGB value distribution to assist with identifying construction materials when using TLS. TLS has been widely used as a data source for colorized 3D point clouds that are either applied directly or further processed into textured 3D models. Previous investigations have addressed point cloud colorization from the perspective of deriving color information from external image sources (e.g., Abdelhafiz et al., 2005; Stal et al., 2011; Moussa et al., 2012; Forkuo & King, 2014; Crombez et al., 2015; Pepe et al., 2016; Pleskacz & Rzonca, 2016), finding that the utilization of external cameras improves the visual quality (e.g., Gašparović & Malarić, 2012; Valero et al., 2016). However, since integrated camera sensors are common in modern TLS instruments, the use of an external camera can also be seen as an extra step in the data collection and processing pipeline that requires additional software, hardware, and manual labor.

### 2.2.2 Close-range photogrammetry

Photogrammetry is a technique for deriving reliable and accurate 3D measurements from images (e.g., Granshaw, 2020). The photogrammetric pipeline has evolved significantly over the years thanks to advances made in integrating photogrammetric and computer vision algorithms (e.g., Remondino et al., 2014; Remondino et al., 2017) and general developments in computation power. In addition to the previously mentioned meshing and texturing, significant progress has been achieved towards creating an automated pipeline that includes all modern core components, including structure-from-motion and dense-image matching.

Structure-from-motion (SfM) is a procedure for simultaneously estimating 3D coordinates (structure) and camera positions (motion) (Ullman, 1979), as well
as the camera calibration parameters. First, the 2D features (tie points) are detected and matched among images via feature extraction (Lowe, 2004) and image feature (tie-point) matching (Hartmann et al., 2016). Second, a highly redundant bundle adjustment (Triggs et al., 1999) procedure is applied to estimate and optimize all (interior and exterior) camera parameters and 3D coordinates for these 2D features. SfM can be implemented with a variety of strategies, ranging from incremental (Agarwal et al., 2011; Schönberger & Frahm, 2016), to hierarchical (Toldo et al., 2015; Cefalu et al., 2017), to global (Moulon et al., 2013; Sweeney et al., 2015) approaches.

Dense-image matching is a procedure where the obtained camera poses and a sparse 3D point cloud representing the scene geometry are used as a starting point to derive a dense colorized 3D point cloud of the scene. Dense-image matching is used to estimate a depth map by retrieving a depth value for each pixel for all cameras that have been resolved by SfM. Semi-global matching (Hirschmuller, 2007) is perhaps the most notable example of such approaches. Furthermore, a dense surface representation of the 3D geometry can be created from the depth maps using 3D meshing techniques, such as 3D Delaunay tetrahedralization (Jancosek & Pajdla, 2011). Unnecessary vertices of the 3D mesh geometry can be reduced using simplification (or decimation) approaches (Kobbelt et al., 1998). After creating the 3D mesh model, it can be textured with automated UV mapping techniques (Lévy et al., 2002, Allène et al., 2008). In contrast to TLS, one benefit of photogrammetry is that the 3D geometry can be derived from the same set of images as the color or texture information. Figure 9 illustrates an example of a photogrammetric 3D reconstruction result.

These aforementioned advances in the photogrammetric pipeline have spawned many open source and commercial software solutions for producing textured 3D mesh models from photogrammetric imagery, such as 3DF Zephyr (3Dflow, 2021), COLMAP (COLMAP, 2021), Meshroom (AliceVision, 2021), Metashape (Agisoft, 2021), nFrames (nFrames, 2021), Pix4D (Pix4D, 2021),
and RealityCapture (Capturing Reality, 2021a), for further use in 3D modeling software suites, game engines, or web-based 3D model publishing platforms. The rise of these solutions has also made the creation of photorealistic 3D models increasingly accessible for non-expert users. The quality and performance of various photogrammetric pipeline implementations have been reviewed and benchmarked by researchers over the years (e.g., Remondino et al., 2014; Nikolov & Madsen, 2016; Stathopoulou et al., 2019; Nocerino et al., 2020).

With **close-range photogrammetry**, the target object or environment is photographed by hand or using a tripod from relatively short distances of up to a few hundred meters, primarily depending on the camera sensor and lens combination being used. Nowadays, a great variety of camera sensors can be used for such a task, ranging from professional-level, digital single-lens reflex (DSLR) or mirrorless digital cameras to diverse camera sensors applied in mobile devices. Yet, applications that demand the most accuracy still require calibrated metric cameras (Luhmann et al., 2016). Depending on the use case, airborne platforms like UAVs can be used to gain an aerial perspective of the target. The increasing popularity, quality, and affordability of digital cameras is undoubtedly lowering the entry barrier for people to apply photogrammetric techniques.

In practice, photogrammetric 3D reconstruction techniques are limited by the capability of the data collection instruments, that is, the digital cameras. Apart from the instrument, the quality of the resulting photogrammetric 3D model can be understood as being the result of a diverse set of possible approaches related to the way the target is photographed and linked to the selected software solutions and how they implement the photogrammetric pipeline. In contrast to TLS, as a direct 3D measuring instrument the quality of TLS is more dependent on the capability of the sensor itself.

Close-range photogrammetry has been applied in a great number of applications, such as in archeology (Samaan et al., 2013; Balletti et al., 2015), architecture (Werner et al., 2002; Nex & Rinaudo, 2009), city modeling (Portalés et al., 2010), construction and civil engineering (Hampel & Maas, 2003), cultural heritage documentation (Grussenmeyer et al., 2012; Apollonio et al., 2021), digital gaming (Statham, 2018), structural dynamics (Baqersad et al., 2017), reverse engineering (Görski et al., 2010), industrial measurements (Luhmann, 2010), forensics (Gonzalez-Aguilera & Gomez-Lahoz, 2009), and underground mining (Benton et al., 2016).

### 2.2.3 Integration of TLS and close-range photogrammetry

The integration of TLS and close-range photogrammetry has frequently been studied and reported in the existing literature, with a great number of discussions pertaining to the differences between and complementary benefits of the methods (e.g., Velios & Harris, 2001; Habib et al., 2004; Guarnieri et al., 2006; Remondino & Rizzi, 2010; Gašparović & Malarić, 2012; Grussenmeyer et al., 2012). The resulting models have typically been evaluated by focusing on their geometric quality, whereas the texturing quality (e.g., Gašparović & Malarić, 2012) has received surprisingly little attention. Furthermore, the quality of the
integrated model has rarely been compared to the results produced using individual methods only.

Use cases that integrate TLS and close-range photogrammetry have included improving building extraction (Becker & Haala, 2007), reconstructing the details of building facades (Li et al., 2011), and improving the extraction of building outlines (Nex & Rinaudo, 2011). In many cases, the desire to use integrated approaches has been motivated by their potential to enable high-quality texture mapping and improve the geometry and visual quality of the 3D model (e.g., Alshawabkeh, 2006). However, many of the reported integration approaches have involved manual labor and have been limited to specific use cases and datasets. In practice, integrating the two methods has often been seen as colorizing 3D point cloud data (Wang et al., 2019), texturing laser-scanned 3D models with external images (Haala & Alshawabkeh, 2006), or generally merging separately produced point cloud, image, or model data (Guarnieri et al., 2006; Balletti et al., 2015; Jo & Hong, 2019) at the end of the 3D modeling pipeline, where the weaknesses of individual data sources become more difficult to fix (Rönnholm et al., 2007; Remondino, 2011). One approach used quite often has been to merge separately collected 3D point cloud datasets, often for the purpose of ensuring a more comprehensive dataset by combining datasets from multiple perspectives (e.g., Jo & Hong, 2019; Valenti & Paternò, 2019).

With respect to the openly or commercially available software solutions, RealityCapture can utilize 3D geometry and the color values from TLS scans as input data in addition to the photogrammetric images. Integration is done in the early stage of the 3D reconstruction pipeline, and thus, it can benefit from the inherited dimensional accuracy of laser scanning (Capturing Reality, 2021b). Many comparable solutions, such as the implemented pipeline in 3DF Zephyr, are based on reconstructing laser scans and images separately and interactively registering them together. Valenti and Paternò (2019) have presented one example of this type of approach.

When comparing the advantages and disadvantages of both close-range 3D measuring approaches, the photogrammetric techniques typically struggle with measuring featureless surfaces, whereas their capability to detect edges and corners is more reliable than with TLS. Close-range photogrammetry is dependent on scene illumination and scale determination. However, to produce good photorealistic results TLS must rely on camera integration to colorize the point cloud data. (e.g., Ramos & Remondino, 2015)

### 2.3 Platforms for photorealistic urban 3D models

An increasing number of software solutions exist that utilize urban 3D models and enable the development of photorealistic applications (e.g., Virtanen et al., 2018a; Julin et al., 2020). The diverse field of application platforms ranges from simple tools and plugins to complex systems, including professional GIS and CAD software, 3D game engines, and web-based 3D model platforms such as virtual globes. That alone poses a challenge for anyone wishing to develop applications or utilize urban 3D models efficiently. Platforms enable the creation
of urban 3D model applications and define their characteristics and functionalities. The technical features and capabilities of the platforms greatly define and influence the way in which and what types of 3D model data can be visualized, interacted with, and accessed. From the perspective of photorealism, platforms can support various rendering qualities and support such aspects of urban 3D models as geographic coordinate systems and various representation types of model geometry and appearance.

The professional GIS and CAD software suites are typically designed for and used by professionals, such as city planners, GIS engineers, architects, and surveyors, and they are not meant for public use. They are typically complex software systems, such as ArcGIS (Esri, 2021a), Locus (Trimble, 2021b), AutoCAD (Autodesk, 2021), and Microstation (Bentley, 2021a), which include a wide variety of tools and functionalities for handling, analyzing, and maintaining model (typically vector or raster) data and support a wide range of spatial data formats and geographical coordinate systems.

Game engines have gained popularity as platforms for developing engaging and interactive applications for visually demanding and usually more specialized use cases. Game engines rely on the real-time rendering of 3D graphics and often support advanced state-of-the-art rendering techniques, such as PBR. This real-time 3D technology makes it possible to create interactive game-like applications and experiences with freedom of movement (6DoF) for the user, realistic real-time simulations of physics and lighting, animation, advanced visual effects, and complex data integrations, and to support multiple simultaneous users or immersive VR and AR systems. Related to applying spatial 3D data, the limited support for global coordinate systems has been a persistent issue when applying geospatial data to game engines (Virtanen et al., 2018a). Game engines such as Unreal Engine (Epic Games, 2021a) and Unity (Unity, 2021a) have traditionally been used to create digital games, but the rapidly advancing technology has increasingly been adopted as application development platforms in architecture, archaeology, the automotive industry, construction, education, sports, various cultural fields, medicine, the real estate business, and various other industry sectors. Some examples of game engine-based urban 3D modeling applications can be found in studies by, for example, Fritsch and Kada (2004), Greenwood et al. (2009), Laksono and Aditya (2019), Buyukdemircioglu and Kocaman (2020), and Wahbeh et al. (2021). As an immersive application example, Schmohl et al. (2020) present a photorealistic VR application based on aerial photography derived photogrammetric 3D mesh model data in a local, but otherwise similar approach to popular Google Earth VR application (Käser et al., 2017). To visually improve the model data, Schmohl et al. (2020) acknowledged the requirement of visually appealing model façades. Figure 10 presents an example of a photorealistic urban 3D model in a game engine environment with highly detailed façade textures.
Figure 10. (a) An example of a photorealistic urban 3D model in a game engine environment (Unreal Engine); and (b) immersive exploration of the same virtual environment using a VR headset.

The Internet has evolved into a major dissemination and sharing platform for 3D models, and web experiences are increasingly based on 3D graphics. The browser-based rendering technology for real-time 3D graphics has been under development since the mid-nineties. The emergence of web-based interactive 3D graphics has enabled real-time rendering natively in a web browser without requiring any separate plugins or installations. The most notable example of this development is WebGL (Khronos Group, 2021a), which is a JavaScript application programming interface (API) that is based on OpenGL ES (Embedded Systems) (Khronos Group, 2021c) and supported natively by most modern mobile and desktop web browsers. Many so-called high-level JavaScript libraries that currently exist, such as Three.js (Three.js, 2021), are designed to make WebGL more accessible and friendlier for application developers. No standard file format exists for web-based 3D models; however, the glTF (GL Transmission Format) format has been popularly utilized (e.g., Schilling et al., 2016; Miao et al., 2017). Furthermore, 3D Tiles (OGC, 2021) is based on glTF and has emerged as a tileable format designed for streaming and rendering massive 3D content, including textured 3D meshes and point clouds. 3D Tiles has been adopted as a standard by the Open Geospatial Consortium (OGC). The performance of web-based, real-time 3D graphics applications is constrained by the memory limitations of web browsers, also file sizes must be limited to ensure tolerable download times, and the potentially diverse user hardware might vary greatly in terms of performance. WebGL has been implemented in numerous web-based platforms from general 3D model publishing platforms such as Sketchfab (Sketchfab, 2021a) to virtual globes like CesiumJS (Cesium, 2021a) and OpenCities Planner (Bentley, 2021b).

Related to the utilization of photorealistic urban models available on the web, perhaps the most relevant platform category is virtual globes, which are platforms designed for viewing geospatial information that have been applied as generic 3D model visualization platforms (e.g., Nebiker et al., 2010; Krämer & Gutbell, 2015; Prandi et al., 2015; Buyukdemircioglu & Kocaman, 2020), or in
more specialized use cases, such as for near-real time semantic view analysis (Virtanen et al., 2021), visualizing energy consumption and production potentials (Helsinki, 2021), and mapping electromagnetic field exposure (Rennes, 2021). The early evolution of virtual globes and related research activities in web-based interactive 3D technologies are discussed, for example, in Christen (2008), Christen and Nebiker (2011), and Loesch et al. (2012). Notably, Nebiker et al. (2010) proposed direct streaming of massive, and colorized point clouds in virtual globes as an alternative modeling approach for numerous urban application scenarios. Currently perhaps the most notable platform in the category, Cesium, has been applied by many cities such as Berlin (Berlin, 2021), Helsinki (Helsinki, 2021), and Rotterdam (Rotterdam, 2021) to host their urban 3D models (Figure 11). Also, Cesium has evolved into many open source (e.g., TerriaJS, 2021) and commercial platforms (e.g., Virtual City Systems, 2021). Other well-known virtual globes include Google Earth (Google, 2021b) and NASA World Wind (NASA, 2021), but arguably the most popular web-based map services, with photorealistic 3D visualizations from Google (Google, 2021a), Apple (Apple, 2021), and Microsoft (Microsoft, 2021a), can also be considered virtual globes.

![Figure 11. Examples of urban 3D models on web-based virtual globe platforms (based on Cesium): (a) Berlin; (b) Helsinki; and (c) Rotterdam.](image)

The ability to create realistic and believable virtual environments has been one of the key premises of utilizing real-time 3D platforms. Notably, many of these platforms have vast global developer communities outside the geospatial domain, and the platforms themselves can even be free to use and monetize within certain limitations. Even completely open source-based workflows, from open 3D data to real-time 3D-based applications, have been studied (Edler et al., 2018). One hindrance to utilizing these real-time 3D platforms is that a great share of the compatible 3D content for these platforms has traditionally been created manually by 3D artists, designers, or other professionals. This task has been seen as difficult to automate and complicated by the diversity of 3D modeling solutions, and often by artistic workflows and practices as well. Content creation can be costly and inefficient and serve as a barrier to entry for both
developers and end users (Evans et al., 2014). However, 3D measuring technology has been widely considered promising for improving the efficiency of these visually demanding content creation pipelines, especially with respect to real-world environments and objects. This potential has also been recognized and promoted within the 3D graphics and gaming communities (e.g., Statham, 2018; Metanaut, 2021).
3. Materials and methods

This section presents the materials and methods used in the dissertation. An overview of the research methods and materials used is provided in Table 2.

Table 2. Overview methods and materials used in the dissertation.

<table>
<thead>
<tr>
<th>Publ.</th>
<th>Research methods</th>
<th>Materials</th>
<th>Study / test sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Multiple-case study</td>
<td>19 selected urban 3D modeling cases</td>
<td>Cities of Helsinki, Espoo, Tampere, Vantaa, Oulu, and Turku</td>
</tr>
<tr>
<td></td>
<td>Qualitative thematic analysis of semi-structured interviews (supplementary method - see section 3.5.)</td>
<td>Six semi-structured interviews with city authorities and experts</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Experimenting with automated, photorealistic, multi-sensor, 3D modeling for the web</td>
<td>TLS and close-range photogrammetry</td>
<td>Complex urban test environment</td>
</tr>
<tr>
<td></td>
<td>Survey analysis (supplementary method - see section 3.5.)</td>
<td>Online survey of 33 experts</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Evaluating the quality of TLS point cloud colorization</td>
<td>TLS and photography</td>
<td>Laboratory test environment</td>
</tr>
</tbody>
</table>

3.1 Multiple-case study

A multiple-case study was conducted to review real-life urban 3D modeling activities with the goal of identifying relevant platforms for utilizing photorealistic 3D models. To collect the data, publicly available information on the selected activities in the six largest cities in Finland was extensively screened. These activities included the 3D city datasets maintained by practitioners in the cities as well as notable individual projects accomplished by private companies. See Table 3 for an overview of the selected activities (cases) and the studied application platforms. More detailed descriptions are presented in Publication 1.

Table 3. A list of the selected urban 3D modeling cases and application platforms investigated as part of the multiple-case study.

<table>
<thead>
<tr>
<th>City</th>
<th>Case</th>
<th>Application platform(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Espoo</td>
<td>3D city model data</td>
<td>Locus</td>
</tr>
<tr>
<td></td>
<td>Mission Leppävaara</td>
<td>CityPlanner</td>
</tr>
<tr>
<td></td>
<td>Otaniemi lighting simulator</td>
<td>Unity</td>
</tr>
<tr>
<td></td>
<td>Tapiola</td>
<td>Unity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helsinki</td>
<td>Helsinki 3D+ information model</td>
<td>Cesium</td>
</tr>
<tr>
<td></td>
<td>Helsinki 3D+ mesh model</td>
<td>Microstation</td>
</tr>
<tr>
<td></td>
<td>3D city model data</td>
<td>Unity</td>
</tr>
<tr>
<td></td>
<td>Oulunkylä 2030</td>
<td></td>
</tr>
<tr>
<td>Oulu</td>
<td>3D city model data</td>
<td>Locus</td>
</tr>
<tr>
<td></td>
<td>VirtualOulu</td>
<td>Unity, Unreal Engine</td>
</tr>
<tr>
<td></td>
<td>Hiukkanavaara 3D model</td>
<td>Unity</td>
</tr>
</tbody>
</table>
This selected sample of activities was analyzed by focusing on a set of indicators: 1) the application platform being used, 2) data accessibility, 3) regional data coverage, and 4) the utilization of as-planned data. These indicators were recurring and unambiguous in the studied sample cases and could be considered essential for the functionality and usability of the urban 3D models.

### 3.2 Test sites

The experiments done as part of this dissertation were performed both in a real-world urban environment with real-life circumstances and in a specifically designed laboratory environment.

#### 3.2.1 Complex urban test environment

Close-range photogrammetry, TLS, and a combination of the two were experimented with in the Puhos shopping mall in Helsinki, Finland (Figure 12). The Puhos shopping mall was designed by Finnish architect Erkki Karvinen and opened in 1964, and it was at the time the largest shopping mall in Finland (Laitinen, 2012). The selected test site contained a partially open, two-story space around an oval-shaped courtyard and a combination of indoor and outdoor spaces with challenging materials (e.g., prominent metal and glass surfaces), complex geometries (e.g., staircases, escalators, curved structures), and difficult lighting conditions. The data acquisition campaign at the test site was part of an interdisciplinary project between Aalto University and the Finnish broadcasting company Yle, with the goal of studying the use of a reality-based and photorealistic 3D model in a journalistic web story: “Puhos: Take a look around a corner of multicultural Finland under threat” (Yle, 2021). Therefore, the data collection was constrained by uncontrollable real-life circumstances of a fixed schedule, changing weather conditions, and a consistently large number of people on site.
Materials and methods

3.2.2 Laboratory test environment

A specifically prepared indoor laboratory test environment was used to evaluate the 3D point cloud colorization quality of commercial TLS instruments. The environment at Aalto University's premises consisted of a blacked-out room illuminated with uniform, standard illuminant D65 (for sRGB and noon daylight), as specified in the ISO 11664-2:2007 standard (ISO, 2021c). The scanner mounting position was fixed, and a set of standardized image quality test charts were horizontally mounted on the wall at a fixed distance of two meters (Figure 13).

The sinusoidally modulated Siemens start (Loebich et al., 2007; ISO, 2021a) chart (Figure 13a) was selected for measuring quality factors related to detail reproduction. The 50.0 × 66.7 cm chart contains 144 pattern bands and is based on the ISO 12233:2017 standard (annex E). The Siemens star chart is considered...
a reliable test chart for measuring system sharpness from multiple angles for all types of cameras, and it is less affected by image processing (e.g., artificial sharpening) than other comparable methods that rely on determining system sharpness based on high contrast edges (Loebich et al., 2007). The X-Rite ColorChecker Classic (based on McCamy et al., 1976) was selected as a reference target (Figure 13b) for assessing quality factors related to color reproduction. The 21.6 × 27.9 cm chart contains 24 reference patches representing primary, chromatic, grayscale, and natural colors. The simplified ISO 15739 digital camera noise test chart (Figure 13c) was selected for measuring noise. The 30.5 × 45.7 cm chart is based on the ISO 15739 standard (ISO, 2021b) and contains 15 uniform greyscale patches for measuring noise. The applied image quality analysis is described further in subsection 3.4.1.

3.3 Instruments and datasets

The studied instruments and datasets collected for the dissertation are listed in Table 4 below. More detailed instrument specifications are presented in the corresponding publications.

Table 4. The instruments and datasets used in the research.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Publ.</th>
<th>Type</th>
<th>Acquired dataset(s)</th>
<th>Test site</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faro Focus 3D S120</td>
<td>2</td>
<td>TLS</td>
<td>RGB-colored 3D point clouds from 22 scan stations</td>
<td>Complex urban test environment</td>
<td>Experimenting with TLS for photorealistic multi-sensor 3D modeling for the web</td>
</tr>
<tr>
<td>Trimble TX5</td>
<td></td>
<td></td>
<td>RGB-colored 3D point clouds from 21 scan stations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leica ScanStation P40</td>
<td>3</td>
<td></td>
<td>Four RGB-colored 3D point clouds with two alternative scan resolution (3.1 mm @ 10 m and 6.3 mm @ 10 m) and dynamic range (HDR$^1$ and LDR$^2$) settings</td>
<td>Laboratory test environment</td>
<td>Evaluating point cloud colorization quality</td>
</tr>
<tr>
<td>Faro Focus S 350</td>
<td></td>
<td></td>
<td>Four RGB-colored 3D point clouds with two alternative scan resolution (3.1 mm @ 10 m and 6.1 mm @ 10 m) and dynamic range (HDR and LDR) settings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leica RTC360</td>
<td></td>
<td></td>
<td>Two RGB-colored 3D point clouds with two alternative scan resolution (3.0 mm @ 10 m and 6.0 mm @ 10 m) settings with an HDR imaging mode.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leica BLK360</td>
<td></td>
<td></td>
<td>Two RGB-colored 3D point clouds with two alternative dynamic range (HDR and LDR) settings and a scan resolution of 5.0 mm @ 10 m.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nikon D800E with Nikkor AF-S 14-24 mm f/2.8G lens</td>
<td>2</td>
<td>Digital camera (DSLR)</td>
<td>A photogrammetric dataset of 433 images.</td>
<td>Complex urban test environment</td>
<td>Experimenting with photogrammetry for photorealistic multi-sensor 3D modeling for the web</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>A photographic reference dataset.</td>
<td>Laboratory test environment</td>
<td>Verifying the test conditions and testing an alternative data source for evaluating point cloud colorization quality</td>
</tr>
</tbody>
</table>

$^1$ HDR (high dynamic range) is an imaging technique where an image with a greater dynamic range is produced by combining several images at different exposure times.

$^2$ LDR (low dynamic range) is a non-HDR imaging technique where the dynamic range of an image consists of an exposure time of a single image.
3.4 Methods for data processing, evaluation, and analysis

The methods used for data processing, evaluation, and analysis are presented in Table 5 and Table 6 below. More detailed descriptions are presented in the corresponding publications.

Table 5. Summary of the data processing methods used in the research.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Publ.</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw scan data processing</td>
<td>2</td>
<td>Pre-processing of TLS data collected with Faro Focus 3D S120 and Trimble TX5 scanners, including data filtering, registration, point cloud colorization, and exporting individual scans as PTX files using Faro SCENE (version 6.0.2.23).</td>
</tr>
<tr>
<td>Raw image data processing</td>
<td>3</td>
<td>Pre-processing of TLS data collected with Leica ScanStation P40, Faro Focus 3D 360, Leica RTC360, and Leica BLK360 scanners with respective original software from the scanner manufacturers (Leica Cyclone REGISTER 360 version 1.6.2. and Faro SCENE version 7.5.2.3361). Colorizing the point cloud data via alternative strategies for using default and linear tone mapping operators to test the impact of dynamic range settings and exporting individual scans as E57 files.</td>
</tr>
<tr>
<td>Point cloud preparation</td>
<td>2</td>
<td>Pre-processing of the equirectangular, panoramic, 32-bit HDR photographs from TLS scanners, including manual adjustment of exposure and white balance levels when experimenting with the use of a linear tone mapping operator as an alternative HDR tone mapping strategy.</td>
</tr>
<tr>
<td>Photorealistic, multi-sensor, 3D modeling for the web</td>
<td>2</td>
<td>Processing of a web-applicable, textured, 3D mesh model and colorized point cloud datasets using three comparable approaches: photogrammetry, TLS, and hybrid model.</td>
</tr>
</tbody>
</table>

Table 6. Summary of the evaluation and analysis methods used in the research.

<table>
<thead>
<tr>
<th>Evaluation and analysis approach</th>
<th>Publ.</th>
<th>Methods</th>
<th>Statistical values and metrics</th>
<th>Available references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geometric quality analysis</td>
<td>2</td>
<td>Using a multiscale model-to-model cloud comparison (M3C2) method for analyzing floor surface deviations between the compared modeling approaches: photogrammetry, TLS, and hybrid model.</td>
<td>Mean distance and standard deviation, as implemented in CloudCompare (version 2.10-alpha).</td>
<td>Lague et al., 2013.</td>
</tr>
<tr>
<td>Texture quality analysis</td>
<td></td>
<td>Analyzing histograms for evaluating the quality of 3D model texture atlases between the compared modeling approaches: photogrammetry, TLS, and hybrid model.</td>
<td>Mean, standard deviation, mode, and the number and percentage of black and white pixels in the texture atlases, as implemented in ImageJ2.</td>
<td>Rueden et al., 2016.</td>
</tr>
<tr>
<td>Computing time calculation</td>
<td></td>
<td>Calculating computing times for model production steps between the compared modeling approaches: photogrammetry, TLS, and hybrid model.</td>
<td>Total time for 3D reconstruction and computing times for meshing and texturing. Computations were done with a PC workstation: AMD Ryzen 7 2700X CPU, 32 GB RAM, Nvidia GeForce 1070 GTX GPU.</td>
<td>Imatest, 2021b; Sharma et al., 2015; ISO/IEC, 2021.</td>
</tr>
<tr>
<td>Image quality analysis for color reproduction</td>
<td>3</td>
<td>Evaluating the instrument’s ability to reproduce colors in the scene. The X-Rite ColorChecker Classic was used as a color reference target.</td>
<td>Color difference (ΔE00), chroma difference (ΔC00), exposure error, and white balance error, as implemented in Imatest Master (version 2020.1.0.45711 Alpha).</td>
<td>Imatest, 2020a; Imatest, 2021c; Imatest, 2021d; Imatest, 2021e; ISO.</td>
</tr>
<tr>
<td>Image quality analysis for detail reproduction</td>
<td></td>
<td>Evaluating the instrument’s ability to reproduce details in the scene using selected quality metrics related to: sharpness and information capacity using a sinusoidally modulated Siemens star chart (based on ISO 12233)</td>
<td>MTF50P and MTF10P (ISO 12233), information capacity, and SNR (ISO 15739), as implemented in Imatest Master (version 2020.1.0.45711 Alpha).</td>
<td>Koren, 2020a; Imatest, 2021c; Imatest, 2021d; Imatest, 2021e; ISO.</td>
</tr>
</tbody>
</table>
3.4.1 Image quality analysis

The dissertation proposes a new method for evaluating the quality of TLS point clouds based on established and standardized image quality assessment techniques. The method is based on measuring selected objective image quality factors (related to color and detail reproduction) from 2D images that represent the 3D scanned image quality test charts in a laboratory test environment.

The color reproduction was assessed by calculating the metrics related to color accuracy based on 2D renders of a ColorChecker chart for each tested scan. CIEDE2000 color difference formulas (Sharma et al., 2015), which are part of the ISO/CIE 11664-6:2014 standard for colorimetry (ISO/CIE, 2021), were used to measure the mean color difference \( \Delta E_{oo} \) and chroma difference \( \Delta C_{oo} \) metrics. The color difference describes the total color error, whereas the chroma difference gives a better indication of the error in colorfulness by describing the color accuracy with a minimal effect from exposure error, indicated by the error in the luminance component (Imatest, 2021b). Additionally, the exposure error and white balance error were calculated as additional image quality metrics related to color reproduction using the grayscale patches of the ColorChecker chart with calculation methods described by Imatest (2021b).

The detail reproduction was assessed by calculating sharpness-related metrics, a Shannon information capacity, and an SNR using the 2D renders from the sinusoidally modulated Siemens star chart and the simplified ISO 15739 noise chart for each tested scan. A modulation transfer function (MTF) was used as a metric for system sharpness, which can determine the maximum level of detail that an imaging system can capture. The MTF was measured as a mean from the star pattern along the radii of a circle in eight segments. The spatial frequencies of MTF10P and MTF50P were selected to summarize the calculated MTF curves for more straightforward comparison purposes. The MTF10P describes a spatial frequency where the image contrast drops to 10% of its peak value and represents a limiting resolution below which all information is considered useless. Compared to similarly used metrics, the MTF50P is considered a robust indicator (e.g., versus artificial image sharpening) of a system’s performance and describes a spatial frequency where the image contrast drops to 50% of its peak value (Koren, 2020b). The algorithms used for calculating the MTF from a Siemens star were based on Loebich et al. (2007), Imatest (2021c), and the ISO 12233 (ISO, 2021a) standard. In addition to the MTF measurements, the Shannon information capacity (Imatest, 2021d) was studied as a novel image quality metric that potentially accounted for both the MTF and the level of noise in the tested datasets and that could calculate them from the same location using the Siemens star chart. The method is described in detail by Koren.
(2020a). Furthermore, the SNR was calculated to assess the amount of noise in the data for each tested scan, as described by Imatest (2021e) and based on the ISO 15739 (ISO, 2021b) standard, derived from the simplified ISO 15739 gray-scale noise test chart.

### 3.5 Supplementary qualitative methods

#### 3.5.1 Semi-structured interviews

As a supplementary method presented in the first publication of the dissertation, semi-structured interviews were organized for a total of 34 informants in the six largest cities in Finland (Helsinki, Espoo, Tampere, Vantaa, Oulu, and Turku) to further understand and explain the observations made during the multiple-case study. The informants were chosen by the cities themselves, and their backgrounds included GIS management, city surveying, city planning, urban development, and supervisory control of building. The durations of the voice-recorded and transcribed interviews were between 60 and 143 minutes, and they were conducted using a semi-structured interview plan:

1. Description of the current situation in the reference city. What is the current situation of 3D city modeling in the city you are representing?
2. Expectations regarding 3D city models. Which tasks should the future 3D city models serve?
3. Expected users. Who are the expected users of the 3D city model?
4. Key factors in the development of 3D city modeling. What can hinder or enhance the development?
5. Data. What kind of data is needed to implement a 3D city model?
6. Visualization. How should 3D city models be visualized?

The qualitative thematic analysis of the interviews was based on the transcriptions. The data was divided into two themes: 1) the internal drivers and potential of 3D city modeling, and 2) the internal challenges and barriers to developing 3D city modeling. The background of an informant did not affect the interview design, mainly due to the variety of ways that 3D city modeling had been organized in the targeted cities.

#### 3.5.2 Online Survey

As a supplementary method in the second publication, an online survey was organized to evaluate the perceived visual quality of the compared 3D models produced using three close-range 3D modeling approaches: photogrammetry, TLS, and a hybrid approach. A total of 33 experts from the fields of 3D measuring and modeling, geoinformatics, computer graphics, and digital gaming participated anonymously in the survey and were asked to choose, without any pre-existing knowledge about any of the models, which of the compared models they liked the best. The compared models were uploaded into, and easily viewed by the participants in Sketchfab, a browser-based 3D model publishing platform.
The online survey questions (from Appendix A of Publication 2) were as follows:

1a. From the perspective of photorealism and visual appeal, which model did you like the best? (multiple-choice question)
1b. Why? Explain briefly (open question with a text box)
2a. Which model has the best geometric quality? (multiple-choice question)
2b. Why? Explain briefly (open question with a text box)
3a. Which model has the best texturing quality? (multiple-choice question)
3b. Why? Explain briefly (open question with a text box)
4. Results

4.1 Platforms for photorealistic urban 3D models (Publication 1)

4.1.1 A multiple-case study to identify the relevant platforms for utilizing photorealistic urban 3D models

The first publication included in the dissertation focused on identifying the relevant platforms for utilizing photorealistic urban 3D models. To achieve this goal, a multiple-case study was conducted that analyzed the observed 3D modeling activities in the six largest cities in Finland. A summary of the results is presented in Table 7 below.

Table 7. A summary of the multiple-case study results (adapted from Publication 1).

<table>
<thead>
<tr>
<th>City</th>
<th>Case</th>
<th>Application platform(s)</th>
<th>Platform type</th>
<th>Data publicly viewable</th>
<th>Data publicly downloadable</th>
<th>Full city data coverage</th>
<th>Utilization of as-planned data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Espoo</td>
<td>3D city model data Mission</td>
<td>Locus CityPlanner</td>
<td>GIS/CAD Virtual</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Leppävaara</td>
<td>Unity</td>
<td>game engine</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Otaniemi lighting simulator</td>
<td>Unity</td>
<td>Game engine</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tapiola</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helsinki</td>
<td>Helsinki 3D+ information model</td>
<td>Cesium</td>
<td>Virtual globe</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Helsinki 3D+ mesh model</td>
<td>Cesium</td>
<td>Virtual globe</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3D city model data</td>
<td>Microstation Unity</td>
<td>GIS/CAD Game engine</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Oulunkylä 2030</td>
<td></td>
<td>Game engine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oulu</td>
<td>3D city model data</td>
<td>Locus Unity, Unreal</td>
<td>GIS/CAD Game engine</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>VirtualOulu</td>
<td>Engine</td>
<td>Game engine</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Hiukkavaara 3D model</td>
<td>MAPGETS</td>
<td>Virtual globe</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>SmartOulu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tampere</td>
<td>3D city model data</td>
<td>Novapoint, Quadri, Auto-</td>
<td>GIS/CAD Virtual</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>campus and science park area</td>
<td>CAD, Viasys VDC</td>
<td>globe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turku</td>
<td>3D city model data</td>
<td>Locus Sova 3D</td>
<td>GIS/CAD Virtual</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3D model of Turku campus and science park area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vantaa</td>
<td>3D city model data</td>
<td>Microstation Unity</td>
<td>GIS/CAD Game engine</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Kivistö</td>
<td>Minecraft</td>
<td>Game engine</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Minecraft</td>
<td>Minecraft</td>
<td>Game engine</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Myymälä</td>
<td>Unity</td>
<td>Game engine</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>13</td>
<td>3</td>
<td>13</td>
<td>10</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>
In the six target cities, the study found a fragmented landscape of a total of 19 different urban model realizations that utilized 13 different platforms. Most of the platforms (a total of 68%) were real-time 3D platforms, and more specifically, game engines (42%). All the investigated photorealistic urban 3D models utilized a real-time 3D platform (i.e., game engines or web-based virtual globes). Among the studied cases, the real-time 3D platforms were used to create specialized applications (e.g., Otaniemi lighting simulator) and applications that targeted broader audiences via participatory urban planning (e.g., Mission Leppävaara). The urban models that applied game engines had the highest verisimilitude and emphasized game-like interaction mechanics and immersion with the goal of user engagement and high-quality visualization, even on a pedestrian level (e.g., VirtualOulu). Due to the performance constraints of web browsers, the urban 3D models that applied web-based, real-time 3D platforms (i.e., virtual globes) appeared to fall visually in between the GIS/CAD platforms and 3D game engines. The observed GIS- and CAD-based platforms appeared to be more for internal use within the city organizations and were more focused on aspects of data management and non-photorealistic visualization.

The study found that a total of 63% of the project cases were viewable on the web, and almost all (92%) of those publicly accessible cases utilized real-time 3D platforms. This was accomplished with separate web-based plugins or natively by the platform itself (especially in the case of virtual globes). A total of 53% of the studied models were publicly downloadable, and those datasets typically consisted of citywide 3D building models and digital terrain models produced using ALS. The 3D models were updated periodically and maintained by the cities together with their statutory base mapping processes, and each city utilized internally at least one GIS/CAD solution for such operations. In all cases, the downloadable urban 3D models had undergone at least some level of processing and no original raw data was distributed to, for example, application developers.

The models used on real-time 3D platforms were more typically limited in their regional scope (with the exception of the citywide Helsinki 3D+ mesh model). In addition, from the data integration perspective, all the studied urban 3D models that focused on integrating as-planned data (i.e., BIM models or architectural plans) utilized real-time 3D platforms.

4.1.2 A supplementary semi-structured interview study

The goal of organizing the supplementary semi-structured interviews was to further understand and explain the observations from the multiple-case study. Both the perceived potential and the barriers hindering the adoption and application of urban 3D models were studied.

Better visuality and the ability to demonstrate the impact of decisions were mentioned as a key element in enhancing the desired interaction and trust between different stakeholders both inside and outside the city organizations. Compared to 2D presentation, a 3D model was generally seen as unambiguous and easier to understand. The level of detail of a 3D model was seen as potentially adaptable based on the use case. The urban 3D model was imagined as a
Results

A forum for integrating city services and information, even as a virtual place where everyone could create their own vision of the city. From the perspective of accessibility, the participants envisioned urban 3D models as open services that enhance communication and collaboration and enable public participation and even crowdsourcing. So, the potential was seen in not merely accessing the models but in activities that result from user engagement and such interactions as co-design and decision-making. Furthermore, the urban 3D models were seen as potentially scalable and flexible data integration platforms serving multiple purposes.

One of the acknowledged barriers was that the group of users is heterogeneous, with different needs, expectations, and views. The future uses for urban 3D models were seen as difficult to predict, leading to them being neither easy to modify nor scalable. The lack of established revenue models has prevented fruitful cooperation between the public and private sector. Furthermore, the participants complained about the lack of expertise, coordination, and leadership in the city organizations and view the role of cities more as providers of 3D data than as application developers.

4.2 Efficient production of photorealistic urban 3D models (Publication 2)

To follow the sequential research process (see section 1.4), the second publication included in this dissertation was motivated by the results of the first publication, which highlighted the relevance of real-time 3D platforms (i.e., game engines and web-based virtual globes) in developing photorealistic urban 3D model applications. A further motivation was to experiment with automated model production approaches and evaluate the resulting quality from the perspectives of the model’s geometry and appearance to lower the observed barriers related to content creation and application development. The first publication also showed that especially photorealistic 3D model applications with the highest level of visual realism on a pedestrian level appeared to rely on highly manual modeling workflows and did not utilize close-range 3D measuring technology to its fullest potential.

Automated multi-sensor urban 3D modeling was experimented for a photorealistic web-based application, and the quality of the compared modeling approaches of close-range photogrammetry, TLS, and a combination of them was evaluated from the perspective of both the model’s geometry and textures. Additionally, a supplementary online survey was organized to gather insights about the perceived visual quality of the compared modeling approaches. To ensure practical applicability, the 3D models were produced using commercial instruments and software solutions in a complex, real-life urban test environment.
4.2.1 Automated photorealistic multi-sensor 3D modeling for the web

The 3D content creation process was specified and experimented with for a web-based, real-time 3D Sketchfab platform, with stricter performance requirements compared to creating 3D content for game engines. The resulting three comparable textured 3D mesh models (photogrammetry, TLS, and hybrid) were simplified and reduced to a budget of 500,000 polygons and a maximum of ten 4k (4096 x 4096 pixel) texture image files, according to the recommendations by Sketchfab (Sketchfab, 2021b).

The photographs and 3D measurements used as an input dataset for the 3D modeling were collected from the urban test environment using close-range photogrammetry and terrestrial laser scanning. To mitigate the effects of changing weather and lighting conditions, the techniques were used simultaneously in the challenging multi-story, indoor-outdoor public space of Puhos shopping mall in Helsinki, Finland. After the data pre-processing phase, the compared photorealistic textured 3D mesh models were produced using RealityCapture (Capturing Reality, 2021a), a commercial photogrammetry software package capable of integrating photogrammetric images and TLS data. An overview of the modeling approaches and the computing times for different data processing phases is presented in Table 8, and the compared models are illustrated in Figure 14. To match the specified polygon count budget, all produced 3D models had to be significantly simplified from their original high polygon count.

Table 8. An overview of the processing of the three compared photorealistic 3D modeling approaches (adapted from Publication 2).

<table>
<thead>
<tr>
<th>Alignment</th>
<th>Photogrammetry</th>
<th>TLS</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total input data file size</td>
<td>6.5 GB</td>
<td>21.1 GB</td>
<td>27.6 GB</td>
</tr>
<tr>
<td>Number of automatically registered images</td>
<td>306/433</td>
<td>-</td>
<td>363/433</td>
</tr>
<tr>
<td>Number of automatically registered laser scans</td>
<td>-</td>
<td>43/43</td>
<td>43/43</td>
</tr>
<tr>
<td>Metric scale ¹</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Reconstruction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of polygons</td>
<td>386,145,064 (full size)</td>
<td>318,875,950 (full size)</td>
<td>693,603,980 (full size)</td>
</tr>
<tr>
<td>Final web-compatible models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of polygons</td>
<td>500,000 (as specified)</td>
<td>500,000 (as specified)</td>
<td>500,000 (as specified)</td>
</tr>
<tr>
<td>Number of 4k textures</td>
<td>10 (as specified)</td>
<td>10 (as specified)</td>
<td>10 (as specified)</td>
</tr>
<tr>
<td>Computing times ²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meshing time</td>
<td>07 h: 07 min: 50 s</td>
<td>00 h: 43 min: 06 s</td>
<td>19 h: 51 min: 23 s</td>
</tr>
<tr>
<td>Texturing time</td>
<td>00 h: 22 min: 15 s</td>
<td>00 h: 01 min: 51 s</td>
<td>00 h: 34 min: 15 s</td>
</tr>
<tr>
<td>Total time</td>
<td>07 h: 30 min: 05 s</td>
<td>00 h: 44 min: 57 s</td>
<td>20 h: 25 min: 38 s</td>
</tr>
</tbody>
</table>

¹ As a technique, laser scanning makes it possible to determine the metric scale of the model data directly, and it can also be used in the hybrid approach.

² The calculated computing times did not account for the data pre-processing stage.
Figure 14. The automatically produced, web-compatible 3D models: (a) photogrammetry; (b) TLS, and (c) hybrid.

4.2.2 Geometric quality

The geometric quality of the compared modeling approaches was evaluated by analyzing the ground floor surface deviations between the produced 3D models and the manually processed, full-resolution, TLS-derived, reference dataset using a multiscale model-to-model cloud comparison (M3C2) method (Lague et al., 2013). The results of this analysis are presented in Table 9 and further visualized in Figures 15 and 16.
Table 9. Summary of the ground floor surface deviation analysis to assess the quality of the 3D model geometry (adapted from Publication 2)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Photogrammetry</th>
<th>TLS</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean distance (signed)</td>
<td>0.41 mm</td>
<td>-0.15 mm</td>
<td>-0.05 mm</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>6.20 mm</td>
<td>2.72 mm</td>
<td>3.18 mm</td>
</tr>
</tbody>
</table>

Figure 15. Distance values of the compared modeling approaches vs. the reference (photogrammetry (green), TLS (red), and hybrid (blue)).

Figure 16. The ground floor surface deviations between all modeling approaches and the reference: (a) the photogrammetry approach; (b) the terrestrial laser scanning approach; and (c) the hybrid approach.

4.2.3 Texture quality

To quantitatively assess the model’s appearance, the texture quality of the compared modeling approaches was evaluated by analyzing the tonal distributions of all the 4k (4096 x 4096) texture atlases of the produced 3D models using a histogram analysis. The results of this analysis are presented in Table 10 and further visualized in Figure 17. The TLS model suffered from clear overexposure and underexposure.
Table 10. Summary of the histogram analysis to assess the quality of the model texturing (adapted from Publication 2).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Photogrammetry</th>
<th>TLS</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (8-bit)</td>
<td>92</td>
<td>126</td>
<td>100</td>
</tr>
<tr>
<td>Standard deviation (8-bit)</td>
<td>43</td>
<td>59</td>
<td>49</td>
</tr>
<tr>
<td>Mode (8-bit)</td>
<td>79</td>
<td>254(^1)</td>
<td>95</td>
</tr>
<tr>
<td>Number of black pixels</td>
<td>1533</td>
<td>1,768,527</td>
<td>4921</td>
</tr>
<tr>
<td>Number of white pixels</td>
<td>1609</td>
<td>3,864,622</td>
<td>909,088</td>
</tr>
<tr>
<td>Percentage of black pixels</td>
<td>0.00091%</td>
<td>1.05%</td>
<td>0.0029%</td>
</tr>
<tr>
<td>Percentage of white pixels</td>
<td>0.00096%</td>
<td>2.30%</td>
<td>0.54%</td>
</tr>
</tbody>
</table>

\(^1\) The 8-bit value of 254 appeared as the maximum pixel value in the resulting texture images generated in RealityCapture.

Figure 17. A histogram analysis of the 8-bit pixel values of all texture atlases in the three compared modeling approaches: photogrammetry (green), TLS (red), and hybrid (blue). The peak in the hybrid model was caused by the empty space between the texture islands on the texture atlases and had no perceivable impact on the visual quality of the model (adapted from Publication 2).

4.2.4 A supplementary online survey

As a supplementary evaluation method, an online survey was conducted to assess the perceived quality of the compared models. The hybrid approach was clearly favored for having the best overall visual appearance (91%), geometry (82%), and texturing (79%). The photogrammetry-based approach had the worst performance in terms of model geometry (0%) and TLS-based model with texturing quality (6%). A summary of the online survey results is presented in Figure 18 and an example of the visually perceivable differences in the compared models is shown in Figure 19.
Results

Figure 18. Results of the online survey on visual quality for the three modeling approaches: photogrammetry (green), TLS (red), and hybrid (blue) (adapted from Publication 2).

Figure 19. Comparison of the visually perceivable differences in the model appearances between the three approaches: (a) the photogrammetry-based model suffers from holes in the data on shiny and non-textured surfaces, such as taped windows; (b) the TLS-based model suffers from illumination differences, which cause visible seam lines (distortions) between the textured areas, and a lack of data underneath the scanning stations, which causes circular patterns in the texture; (c) many of the visible problems were fixed in the hybrid model (adapted from Publication 2).

Overall, the research results supported the view that creating photorealistic 3D models by integrating photogrammetry and TLS is a good compromise for both geometric and texture quality. Compared to approaches that used a single method it was slower and more computationally expensive but combined many complementary advantages of both approaches, for example the superior image quality of photogrammetry and the direct scale determination and higher geometric quality of TLS. In part, the observed lower quality of TLS in producing model textures provided the motivation to focus on the colorization quality of TLS-derived point clouds in Publication 3 (as further explained in section 1.4) of the dissertation.
4.3 Quality of photorealistic close-range 3D measuring (Publication 3)

The third publication included in the dissertation proposed a new method for evaluating the quality of TLS point cloud colorization and experimented with its use in benchmarking the colorization quality of four commercial TLS instruments. This research was motivated by the observed lack of quantitative methods for assessing the quality of model appearance and by the results from the second publication, which highlighted the weaknesses of TLS-based 3D modeling in terms of texturing quality compared to close-range photogrammetry.

4.3.1 A new method for evaluating the quality of TLS point cloud colorization

As a result, a method (outlined in Figure 20) for evaluating the colorization quality of TLS-derived 3D point clouds was developed. It was based on utilizing well-established and standardized image quality metrics related to detail and color reproduction: color accuracy (ISO/CIE 11664-6), sharpness (ISO 12233), information capacity, and SNR (ISO 15739) using corresponding standardized test charts (X-Rite ColorChecker, sinusoidal Siemens star (ISO 12233) and a simplified noise test chart (ISO 15739), as previously described in Figure 13).

Figure 20. An overview of the proposed method for evaluating the colorization quality of TLS-derived 3D point cloud data (adapted from Publication 3).

The proposed method was experimented with using four different commercial TLS instruments (Leica ScanStation P40, Faro Focus S150, Leica RTC360, and Leica BLK360) in a specially designed laboratory test environment. The tested scanners represent diverse design choices, for instance in terms of their portability, affordability, operating range, and camera configurations and specifications. In general, the results revealed quality differences between the tested scanners and settings that can be considered significant in use cases where high-quality color information is required, such as in applications that rely on visual appearance or photorealism. The key results are summarized in the following two sections, whereas the remaining results are described in detail in the original Publication 3.
4.3.2 Color reproduction

To better evaluate the color reproduction of the tested commercial TLS instruments, the color accuracy was measured based on colorized 3D point cloud data representing the 3D scanned ColorChecker chart. As additional quality metrics, an exposure error and a mean white balance error were also calculated. A summary of the results related to color reproduction is presented in Table 11 below.

Table 11. A summary of the color reproduction metrics (adapted from Publication 3; higher numbers indicate a higher degree of error).

<table>
<thead>
<tr>
<th>Scan (TLS instrument, dynamic range, scan resolution, HDR tone mapping)</th>
<th>Color Difference (mean ΔE&lt;sub&gt;∞&lt;/sub&gt;)</th>
<th>Chroma Difference (mean ΔE&lt;sub&gt;∞&lt;/sub&gt;)</th>
<th>Exposure Error (f-stops)</th>
<th>White Balance Error (saturation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leica P40, HDR, 3.3 mm, Default</td>
<td>3.48</td>
<td>2.58</td>
<td>-0.69</td>
<td>0.15</td>
</tr>
<tr>
<td>Leica P40, HDR, 3.3 mm, Linear</td>
<td>3.39</td>
<td>2.59</td>
<td>-0.65</td>
<td>0.15</td>
</tr>
<tr>
<td>Leica P40, LDR, 3.3 mm</td>
<td>3.16</td>
<td>2.72</td>
<td>0.65</td>
<td>0.14</td>
</tr>
<tr>
<td>Leica P40, HDR, 6.3 mm, Default</td>
<td>3.47</td>
<td>2.93</td>
<td>-0.64</td>
<td>0.16</td>
</tr>
<tr>
<td>Leica P40, HDR, 6.3 mm, Linear</td>
<td>3.40</td>
<td>2.92</td>
<td>-0.44</td>
<td>0.18</td>
</tr>
<tr>
<td>Leica P40, LDR, 6.3 mm</td>
<td>3.14</td>
<td>2.70</td>
<td>0.85</td>
<td>0.14</td>
</tr>
<tr>
<td>Faro S 350, HDR, 3.3 mm, Default</td>
<td>3.94</td>
<td>3.37</td>
<td>-0.23</td>
<td>0.05</td>
</tr>
<tr>
<td>Faro S 350, HDR, 3.3 mm, Linear</td>
<td>3.92</td>
<td>3.35</td>
<td>-0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>Faro S 350, LDR, 3.3 mm</td>
<td>3.68</td>
<td>3.15</td>
<td>0.41</td>
<td>0.05</td>
</tr>
<tr>
<td>Faro S 350, LDR, 6.1 mm</td>
<td>3.16</td>
<td>2.55</td>
<td>1.03</td>
<td>0.05</td>
</tr>
<tr>
<td>Faro S 350, HDR, 6.1 mm, Linear</td>
<td>3.13</td>
<td>2.54</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>Faro S 350, LDR, 6.1 mm</td>
<td>3.92</td>
<td>3.15</td>
<td>0.60</td>
<td>0.04</td>
</tr>
<tr>
<td>Leica RTC360, HDR, 3.0 mm, Default</td>
<td>3.44</td>
<td>2.68</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Leica RTC360, HDR, 3.0 mm, Linear</td>
<td>3.59</td>
<td>2.93</td>
<td>-0.10</td>
<td>0.06</td>
</tr>
<tr>
<td>Leica RTC360, HDR, 6.0 mm, Default</td>
<td>4.66</td>
<td>3.73</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Leica BLK360, HDR, 5.0 mm, Default</td>
<td>3.98</td>
<td>3.15</td>
<td>0.16</td>
<td>0.08</td>
</tr>
<tr>
<td>Leica BLK360, HDR, 6.0 mm, Linear</td>
<td>3.69</td>
<td>2.54</td>
<td>0.71</td>
<td>0.32</td>
</tr>
<tr>
<td>Leica BLK360, LDR, 5.0 mm, Linear</td>
<td>3.72</td>
<td>3.56</td>
<td>-0.55</td>
<td>0.17</td>
</tr>
<tr>
<td>Nikon D900E photographic reference</td>
<td>4.77</td>
<td>3.80</td>
<td>-0.11</td>
<td>0.05</td>
</tr>
</tbody>
</table>

To represent the total color accuracy, the color difference (ΔE<sub>∞</sub>) for each tested scan is presented in Figure 21. For all tested TLS scanners, the absolute total color error was the highest when HDR imaging was used with default tone mapping settings and without any manual adjustments. When using the default tone mapping settings, the HDR scans collected with the Leica RTC360 produced the most accurate colors. When the manually adjusted linear tone mapping settings were used, or when the HDR mode was not used at all, the results from the Faro S 350 were closest to the ColorChecker reference. The scan resolution did not have any significant effect on the color difference. The measured mean white balance errors are presented in Figure 22. The Faro S 360 had the most accurate white balance, while the Leica BLK360 produced the largest white balance error of all the compared scans.

Figure 21. The total color accuracy measured as color difference (ΔE<sub>∞</sub>) between each tested scan and the ColorChecker reference values (adapted from Publication 3).
Figure 22. The mean white balance error between each tested scan and the reference values (adapted from Publication 3).

4.3.3 Detail reproduction

To assess the detail reproduction of the tested scanner instruments, the selected quality metrics related to system sharpness and information capacity were measured based on a 3D scanned Siemens star chart. Furthermore, the SNR (based on ISO 15739) was measured using a 3D scanned, simplified ISO 15739 noise test chart to estimate the level of noise in the colorized 3D point clouds. A summary of the results related to the detail reproduction is presented in Table 12 below.

Table 12. A summary of the detail reproduction metrics (adapted from Publication 3).

<table>
<thead>
<tr>
<th>Scan (TLS instrument, dynamic range, scan resolution, HDR tone mapping)</th>
<th>MTF50P (cycles/pixel)</th>
<th>MTF10P (cycles/pixel)</th>
<th>Shannon Information Capacity (bits/pixel)</th>
<th>ISO 15739 SNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leica P40, HDR, 3.1 mm, Default</td>
<td>0.199</td>
<td>0.400</td>
<td>1.32</td>
<td>27.4</td>
</tr>
<tr>
<td>Leica P40, HDR, 3.1 mm, Linear</td>
<td>0.191</td>
<td>0.357</td>
<td>0.83</td>
<td>46.4</td>
</tr>
<tr>
<td>Leica P40, LDR, 3.1 mm</td>
<td>0.255</td>
<td>0.698</td>
<td>3.12</td>
<td>35.6</td>
</tr>
<tr>
<td>Leica P40, HDR, 6.3 mm, Default</td>
<td>0.167</td>
<td>0.316</td>
<td>1.07</td>
<td>27.2</td>
</tr>
<tr>
<td>Leica P40, HDR, 6.3 mm, Linear</td>
<td>0.173</td>
<td>0.296</td>
<td>0.81</td>
<td>47.5</td>
</tr>
<tr>
<td>Leica P40, LDR, 6.3 mm</td>
<td>0.201</td>
<td>0.353</td>
<td>1.04</td>
<td>35.6</td>
</tr>
<tr>
<td>Faro 5300, HDR, 3.1 mm, Default</td>
<td>0.237</td>
<td>0.419</td>
<td>2.21</td>
<td>26.8</td>
</tr>
<tr>
<td>Faro 5300, HDR, 3.1 mm, Linear</td>
<td>0.237</td>
<td>0.417</td>
<td>2.31</td>
<td>38.5</td>
</tr>
<tr>
<td>Faro 5300, LDR, 3.1 mm</td>
<td>0.225</td>
<td>0.406</td>
<td>2.41</td>
<td>28.5</td>
</tr>
<tr>
<td>Faro 5300, HDR, 6.3 mm, Default</td>
<td>0.191</td>
<td>0.336</td>
<td>1.34</td>
<td>27.8</td>
</tr>
<tr>
<td>Faro 5300, HDR, 6.3 mm, Linear</td>
<td>0.196</td>
<td>0.339</td>
<td>1.38</td>
<td>40.1</td>
</tr>
<tr>
<td>Faro 5300, LDR, 6.3 mm</td>
<td>0.190</td>
<td>0.328</td>
<td>1.45</td>
<td>36.4</td>
</tr>
<tr>
<td>Leica RTC360, HDR, 3.0 mm, Default</td>
<td>0.227</td>
<td>0.377</td>
<td>1.16</td>
<td>32.5</td>
</tr>
<tr>
<td>Leica RTC360, HDR, 6.0 mm, Default</td>
<td>0.188</td>
<td>0.312</td>
<td>0.97</td>
<td>41.0</td>
</tr>
<tr>
<td>Leica BLK360, HDR, 4.0 mm, Linear</td>
<td>0.182</td>
<td>0.310</td>
<td>0.89</td>
<td>36.0</td>
</tr>
<tr>
<td>Leica BLK360, HDR, 5.0 mm, Default</td>
<td>0.135</td>
<td>0.226</td>
<td>1.08</td>
<td>33.3</td>
</tr>
<tr>
<td>Leica BLK360, HDR, 5.0 mm, Linear</td>
<td>0.136</td>
<td>0.227</td>
<td>0.76</td>
<td>39.5</td>
</tr>
<tr>
<td>Nikon D600E photographic reference</td>
<td>0.212</td>
<td>0.357</td>
<td>2.89</td>
<td>34.0</td>
</tr>
</tbody>
</table>

The MTF curves (according to ISO 12233) were measured using the sinusoidal Siemens star chart as a metric for system sharpness. The curves were summarized as the spatial frequencies corresponding to MTF50P and MTF10P values and are presented in Figure 23. As expected, the system’s sharpness values increased with the scan resolution, whereas the differences between the tested dynamic range settings were not so clear. Overall, the Leica BLK360 produced the least sharp results of all the tested settings.
Results

Figure 23. The sharpness values, corresponding to MTF50P and MTF10P values, between each tested scan (adapted from Publication 3; a larger number indicates sharper results).

As an experimental metric, a Shannon information capacity was calculated per scan using the sinusoidal Siemens star chart, and the results are described in Figure 24. The Faro S 350 produced the highest information capacity, whereas the results from the Leica BLK360 were the lowest of all the tested scans.

Figure 24. Measured Shannon information capacity for all tested 3D scans measured using a sinusoidal Siemens star chart (adapted from Publication 3).

To assess the level of noise, SNR (according to ISO 15739) was measured using the simplified ISO 15739 noise test chart for each 3D scan, and the results are described in Figure 25. When HDR imaging was applied, using the linear tone mapping approach reduced the amount of noise in all scans, except for the ones acquired with the Leica RTC360.
In addition, for benchmarking purposes selected measured quality metrics (color difference, white balance error, sharpness, information capacity, and noise) were summarized under one combined quality score metric (Figure 26). When the HDR imaging mode was used with linear tone mapping settings (with manual adjustments), or when the HDR mode was not used at all (LDR), the Faro S 350 produced the best overall colorization quality, whereas when the default HDR settings were used (with minimal processing steps), the Leica RTC360 appeared to produce the best colorization quality. The Leica RTC360 was the only scanner that did not indicate any improvement in quality when the HDR imaging mode was used with a linear tone mapping approach. Among all the tested settings, the Leica BLK360 produced the lowest quality score. As expected, the increase in scan resolution improved the colorization quality for all tested scanners. Likewise, the tested photographic reference with a Nikon D800E produced a better quality score than the tested TLS scans.
The dissertation developed the applicability and contributed to a deeper understanding of the efficiency and quality of close-range 3D measuring techniques for use cases that require photorealistic urban 3D models. To fulfill this general objective, the research was divided into three research questions: RQ1, RQ2, and RQ3. The key findings, their implications, and limitations of this dissertation are discussed below.

5.1 Platforms for photorealistic urban 3D models (Publication 1)

The first publication included in the dissertation reviewed real-life urban 3D modeling activities to identify relevant application platforms and assessed interviews with city authorities and practitioners to understand the internal motivators and barriers to applying photorealistic urban 3D models. The goal was to answer research question RQ1.

The results showed that real-time 3D platforms (i.e., game engines and web-based virtual globes) clearly constitute a relevant technology for utilizing photorealistic urban 3D models. In the multiple-case study, all identified urban 3D models that relied on photorealistic visualization utilized real-time 3D-based application platforms (see Figure 27 for examples). Also, most of the studied projects that were openly accessible for users outside city organizations used real-time 3D platforms.

Figure 27. Examples of the studied photorealistic urban 3D models in Publication 1: (a) game-engine-based VirtualOulu, and (b) citywide Helsinki 3D+ mesh model (image courtesy of City of Helsinki).
Generally, the most detailed and visually appealing photorealistic urban 3D models (similar to, e.g., VirtualOulu) appeared to be based heavily on manual modeling techniques, whereas the aerial photogrammetry-based Helsinki 3D+ mesh model was the only photorealistic 3D model that at the same time had a high degree of automation and broader citywide coverage. But due to the large scale of the model, the quality was not feasible for street-level or human-scale visualization purposes. Since the time of publication, a similar modeling approach has been completed in, for example, Tampere (Tampere, 2021). These findings partially provided the motivation for experimenting with the automated production of highly detailed and photorealistic urban 3D models reported in Publication 2.

The supplementary semi-structured interview study highlighted many favorable aspects that would motivate cities to utilize real-time 3D platforms in the future, such as the stated desire for visuality, interactivity, accessibility, flexibility, and availability for users and developers outside the city organization. The interview results underlined the need to develop workflows that enable the use of real-world 3D model data in application platforms that are accessible for larger developer communities outside the typical professional fields of, for example, 3D GIS, surveying, urban planning, or architecture. However, while the real-time 3D platforms clearly appeared the most relevant types of platforms for utilizing photorealistic urban 3D models, the results indicated that no single platform exists that would meet all the desired requirements related to the production, upkeep, maintenance, and utilization of urban 3D models.

Many of the stated barriers, such as heterogeneous user groups with varying needs and inharmonious expectations and views regarding urban 3D models, or the difficulty in predicting use cases, make it difficult to want to adopt one single multi-purpose model. Rather, they encourage the flexible application development of specialized applications. The stated lack of expertise in city organizations calls for capacity building or the need to encourage a larger group of developers outside cities. The cities also saw their role more as 3D model data providers than as application developers, which also underscores the idea that application development should be efficiently possible outside city organizations. To some degree, this had already been achieved in many of the observed urban 3D modeling cases that relied on real-time 3D platforms, since they had often been produced by private contractors. Despite the generally good accessibility for at least viewing the urban 3D data, none of the studied cities offered any raw data for developers, which to some degree hinders the ability to freely develop applications and thus add value to the existing 3D data. For example, a similar 3D modeling approach to the hybrid modeling approach described in Publication 2 would not be feasible without access to the raw sensor data. On the other hand, this would probably raise more issues and concerns regarding privacy and security, which many participants had already mentioned as barriers during the interviews.

The multiple-case study revealed a large degree of variance in all the studied aspects. For example, among the six targeted cities the study found a total of 19 urban 3D modeling cases that utilized a total of 13 different platforms. In reality,
modeling projects are accomplished by people from different fields using different tools and targeting different outcomes. To further explain the identified level of fragmentation among the studied urban 3D modeling cases, the different aspects can be divided into three partially overlapping operational cultures in the urban 3D modeling domain: 3D GIS, BIM, and computer graphics (Figure 28). Each of these identifiable operational cultures has its own perspective with respect to urban 3D modeling, one that results in different realizations.

![Figure 28](image.png)

**Figure 28.** The three proposed operational cultures for urban 3D modeling. Most of the studied photorealistic urban 3D modeling cases could be classified under the visually oriented computer graphics culture heading (adapted from Publication 1).

The interview study indicated an aspiration to achieve the ideal of a single and completely harmonized urban 3D model. To realize this goal and fix the observed fragmentation in urban 3D modeling, the communication and interplay between people and technologies across these three operational cultures should be improved and all unnecessary obstacles removed. A proposed concept for harmonizing urban 3D modeling between the three operational cultures is discussed in more detail in Publication 1.

The research presented in Publication 1 reviews public urban 3D models, meaning all undisclosed modeling efforts accomplished both inside and outside city organizations were inevitably left outside the scope of the research. Also, the scope of the study was limited to Finland, and therefore, some of the identified platforms were possibly applied only in Finland and not more widely elsewhere. Furthermore, the research was affected by the typical limitations of interview studies, such as the possible influence of the interviewees’ subjective views, conforming to popular opinion, or social desirability bias.

To date, not many studies have aimed at providing a larger picture of the ongoing real-life activities in the urban 3D modeling scene. Furthermore, there have not been many reported efforts to understand the internal motivations and barriers related to these activities. This underlines the often-understated issue
that realizing the potential benefits of urban 3D modeling is more than just a technological challenge, as it also involves social and economic aspects. Moreover, compared to other related research, the role of platforms as an integral part of utilizing urban 3D data has not often been discussed. To maximize the benefits and fulfill the application potential of urban 3D modeling, the challenge and significance of putting the 3D data into use, especially beyond internal use by city organizations, has been understated.

The research reported in Publication 1 has gained the attention of many researchers (e.g., Murshed et al., 2018; Buyukdemircioglu & Kocaman, 2020; Sun et al., 2020), and it can even be argued that the findings presented in Publication 1 have, at least to some degree, influenced the proposed national building standard in Sweden (e.g., Eriksson et al., 2020). Motivated partially by Publication 1, the technical capabilities of real-time 3D platforms from the perspective of developing applications have been further investigated and discussed by Julin et al. (2020). Since Publication 1, the significance of real-time 3D platforms has arguably been increasing and the use of real-time graphics technologies in the urban 3D modeling context has been increasingly reported in academic literature (e.g., Laksono & Aditya, 2019; Buyukdemircioglu & Kocaman, 2020; Schmohl et al., 2020; Keil et al., 2021; Wahbeh et al., 2021) and elsewhere (e.g., Esri, 2021b; Epic Games, 2021b; Virtual Helsinki, 2021). Over the last couple of years, the urban 3D modeling domain has been dominated by discussion on digital twins, ranging from local-level modeling efforts (Ruohomäki et al., 2018; Dembski et al., 2020; Schrotter & Hürzeler, 2020) to national-level ecosystems (Bolton et al., 2018).

5.2 Efficient production of photorealistic urban 3D models (Publication 2)

The second publication included in the dissertation experimented with automated multi-sensor urban 3D modeling for an interactive and photorealistic web-based application and evaluated the quality of the model’s geometry and textures. The goal was to answer research question RQ2.

The results showed that close-range photogrammetry and TLS can be utilized for a highly automated and integrated model production process to create photorealistic urban 3D models. When comparing the experimented techniques, TLS produced the best geometric quality, whereas the texturing quality of photogrammetry was superior. A hybrid approach combining photogrammetry and TLS proved a good compromise for the quality of the model’s geometry and textures but came with added costs and manual labor as a result of essentially using two methods for data collection and pre-processing. The experiments further demonstrated that the integration of these two complimentary close-range 3D measuring methods could be directly reproduced in a commercially available software package, RealityCapture. In addition, 3D modeling was applied in a real-life project case with a complex urban indoor-outdoor environment with uncontrollable lighting and the strict performance requirements of an online real-time 3D Sketchfab platform. Furthermore, the study highlighted problems
that the TLS has in delivering texture quality comparable to the quality of photogrammetry, thus underlining the need for further investigation, which was achieved later in Publication 3.

Results of the supplementary online survey were even more favorable for the hybrid model than the numeric quality analyses. The quality of the model's geometry and texturing went hand in hand, such that a good geometric quality improved the visually perceived quality of texturing, while good texturing quality had the same positive effect on the visually perceived quality of the model's geometry. It was also evident that informants tended to focus on coarse errors and visual artifacts, such as holes or texture artifacts (see examples in Figure 19), in the models. These types of errors are often inherited from the selected 3D measuring approach and how well it can be made to accommodate to the existing circumstances: the errors included circle-shaped artifacts related to the scanning positions of TLS, problems in the data coverage caused by occlusions, and uncontrollable changes in the environment, such as weather, lighting, moving people, and objects. These types of problems can be extremely difficult and painstaking to correct in a later stage of the 3D modeling pipeline.

Even though the model production process was highly automated, the model's web compatibility remains a key challenge when creating real-world-based 3D models, especially due to the performance limitations of web browsers and WebGL-based platforms. All compared 3D models had to be heavily simplified from the original high polygon count to run on the targeted application platform. For example, to meet the target of 500,000 polygons, the hybrid model had to be simplified and reduced to 0.07% of its original polygon count of almost 694 million polygons. As designed, this simplification process inevitably reduces detail in the models, which, when using automated solutions, results in the loss of meaningful details.

Prioritizing photorealism sets a high barrier for the visual quality of the models. Both the appearance and geometry of the model are expected to be as free of errors and visible artifacts as possible. In practice, the higher the visual quality requirements, the more difficult it becomes to automate the modeling process. Many commonly executed manual model editing and optimization techniques exist to improve the visual quality of textured 3D models, such as cleaning and fixing errors (e.g., holes, double faces, non-manifold geometry, other artifacts), optimizing the model’s topology via retopologizing processes, manual UV mapping, normal mapping, texture de-lighting, and creating material maps for PBR. Similarly, various types of needs for 3D model optimization have been reported by others (e.g., Kersten et al., 2017; Virtanen et al., 2018b). These optimization techniques are often done by 3D modeling professionals, such as 3D artists, with diverse and specific workflows and tools that pose a great challenge for automation.

Unsurprisingly, the results suggest that no single 3D measuring technique exists that could fulfill all requirements for creating photorealistic 3D assets. In general, the observations made about the strengths and weaknesses of close-range photogrammetry and TLS are well in line with findings presented in pre-
Discussion

Previous studies touching on the topic (e.g., Ramos & Remondino, 2015). The research clarifies the complementary benefits of using a multi-sensor (hybrid) modeling approach compared to individual approaches, and, for example, the findings of this study motivated the use of a similar hybrid workflow to produce the indoor 3D point cloud used in a study by Virtanen et al. (2020). Publication 2 showed that multi-sensor approaches have become productionally available using fully commercial solutions. However, they still support 3D data in a certain form; for example, in the case of RealityCapture the laser scan data must be in a specific ordered file format where the data is ordered according to the individual scan stations. To date, very few studies have been published that consider the integration of TLS and close-range photogrammetry for photorealistic and web-compatible 3D models. Evaluations of geometric quality are a mainstay of studies focusing on 3D measuring, but assessing both the quality of a model’s geometry and its appearance based on comparisons of several modeling approaches has not been much reported. Furthermore, the approach of assessing the qualitative aspects of the compared 3D models in the same study with the same dataset as the quantitative quality evaluation is unique to this research.

The approaches used in Publication 2 are limited to the capabilities of the used hardware and software solutions. Since its publication, advances in TLS instruments have made it more efficient and capable of acquiring HDR images, which would have potentially mitigated part of the illumination-related problems regarding the texturing quality of TLS. In part, this provided the motivation for the research conducted in Publication 3, which assessed the performance of TLS-based HDR imaging and manual adjustment of tonal scales or the white balance of the TLS-based image data prior to point cloud colorization. Furthermore, the approach of Publication 2 was limited to a commercially available software package, RealityCapture, which uses proprietary algorithms in the 3D reconstruction pipeline. Also, the polygon count limitations of the Sketchfab platform might no longer apply and other alternative approaches, such as taking advantage of the tileable 3DTiles format (OGC, 2021), have since become potentially feasible. Furthermore, the scope of the research involved adapting a real-life project (Yle, 2021) with real-life limitations, such as changing weather, lighting, and moving people in the modeled scene. It is also noteworthy that the emphasis of the 3D modeling was photorealistic visualization, not achieving the highest level of dimensional accuracy and precision. The results from the online survey are limited to presenting subjective views about the visual quality of the compared models.

Moreover, the regional scope of the produced 3D models in Publication 2 could be scaled from a single building site to cover a larger area. Extending the urban 3D model could be achieved by using complementary airborne and/or ground-based techniques (see Figure 2 for an overview). To gain an aerial perspective, additional UAV- or other aerial-based image data (similar to e.g., Wen et al., 2019) could be used, at least to some extent, directly within the same modeling workflow that was reported in Publication 2. Additionally, using ALS, oblique imaging, or hybrid mapping sensors (e.g., Toschi et al., 2018; Haala et al., 2020) could enable photorealistic modelling even on a citywide scale, but
likely with significantly increased cost and labor. The scope of the 3D model could be extended at the ground-level using vehicle based or portable mobile mapping systems (e.g., Blaser 2020; Li et al., 2020). More information about mobile mapping of urban environments can be found, for example, in Wang et al. (2019) and Lehtomäki (2021). However, while being mobile and potentially efficient data collection methods, reconstructing larger urban areas or entire cities with similar visual quality and detail that was reported in Publication 2 would be laborious and challenging. The expected extreme data volumes would not only pose a challenge to data acquisition and 3D modeling but also to data management and utilization. For example, visualizing significantly larger areas with similar level of detail and quality to Publication 2 would likely benefit from out-of-core rendering techniques (e.g., Richter & Döllner, 2010; Schütz et al., 2020). Often in real-time applications, the amount of data (e.g., number of polygons and textures) is still limited by certain budget. Consequently, there still exists a trade-off between the level of detail and regional coverage of the 3D dataset. However, 3D modeling efforts will be easier to scale in the future as the data acquisition technology develops towards more mobile and integrated methods and the data processing advances closer to full automation.

5.3 Quality of photorealistic close-range 3D measuring (Publication 3)

The third publication included in the dissertation proposed and experimented with a method for evaluating the quality of TLS-based 3D point cloud colorization using commercial TLS instruments. The goal was to answer research question RQ3.

More specifically, the developed method evaluated the point cloud colorization quality of modern commercial TLS systems with integrated imaging sensors, but it can arguably be applied to evaluating the colorization quality of virtually any colored point cloud. The tested TLS instruments' capabilities to reproduce colors and details were investigated by experimenting with the use of image quality assessment methods. In addition, the proposed method was experimented with in benchmarking the colorization quality of commercial TLS instruments.

The results showed clear differences between the tested TLS instruments in all measured quality aspects related to color and detail reproduction. These numerical differences are well supported by visual observations (Figure 29). The established objective image quality metrics could be successfully and usefully implemented when evaluating the colorization quality of a TLS-derived 3D point cloud in cases where the quality of an imaging system involves a complex combination of an unknown number of unknown factors related to the quality of the imaging sensor, lens, image processing pipeline (e.g., sharpening, compression, denoising, tone mapping, panoramic stitching, and point cloud colorization), and test chart.
The perceived and measured quality inconsistencies limit the usefulness and reliability of the color information of the point cloud in use cases that rely on colored 3D point clouds. For example, producing uniform color information between different scan settings and scanners, and thus a higher quality model appearance, becomes more difficult to achieve. The identified quality inconsistencies in colorization could be more efficiently corrected if the scanner manufacturers would be more transparent about the imaging specifications and capabilities of their “black box” type of systems. This would also assist the users in making better and more proven decisions on instrument selection. Furthermore, either entirely or better controlled measurement parameters would be beneficial for many scientific applications. The research results will hopefully raise awareness about the importance of quality point cloud colorization and its potential implications for applications that rely on colored 3D point clouds.

The results further indicated that there is a need to develop better and more accessible colorization tools and workflows as well as automated image processing pipelines that would increase the efficiency and quality of 3D point cloud colorization. Also, investments in improving the quality of point cloud color information would increase the applicability of 3D point cloud in use cases that rely on visual analysis, reliable object interpretation, recognition, and photorealism. Furthermore, higher quality radiometric data could increase the level of automation and quality in various data processing, analysis, and modeling.
tasks. This would make 3D point clouds and the inherent 3D models more relevant and useful not only in traditional application areas such as surveying, engineering, and cultural heritage but would also introduce the 3D measuring technologies into new application areas.

The findings suggest that the measured and visually detectable problems in color reproduction (i.e., differences in color, white balance, and exposure) could be alleviated in data processing. The errors related to luminance (similar to errors that were reported in real-life conditions in Publication 2) could be corrected using adjustments similar to those commonly used when editing digital photographs, whereas correcting the errors in chroma would require color calibrating the instrument in the field, similar to the color corrections applied for a photogrammetric dataset by, for example, Gaiani et al. (2017) and Apollonio et al. (2021). Compared to color reproduction, the quality issues related to detail reproduction (i.e., measured sharpness and noise) seemed much harder to fix since they are more complex and beyond the control of the scanner user. Related to TLS, the processes and tools used to correct problems (e.g., in-field color calibration) have either not been implemented or else would be labor intensive to achieve in a real-life project with typically uncontrollable lighting and up to hundreds of scan stations. Nevertheless, correcting the errors in color or detail would directly improve the quality of the 3D point cloud or appearance of the inherent textured 3D model.

The data collection speed cannot be ignored when evaluating the colorization performance of a TLS instrument. The speed of the scanner depends on the selected scan and imaging settings (e.g., scan resolution and dynamic range settings) as well as the performance of the instrument. Long data acquisition times increase the duration and cost of real-life projects and can reduce the data quality by including more uncontrollable changes in the environment, such as changing weather, lighting, or moving objects in the scene, of the collected data. The scan times between the tested scanners were strikingly different, with possible practical implications. As an example, there was a tenfold difference between the fastest scanner, the Leica RTC360, and the slowest scanner, the Faro Focus S 350, when using equivalent settings to collect HDR data. Especially the decisions about the imaging settings (e.g., whether to use HDR or LDR imaging mode) had the greatest impact on the total data collection time spent in the field. The fastest scanners (the Leica RTC360 and Leica BLK360) clearly benefit from using three camera sensors mounted on the scanner body instead of just using a more traditional approach where the scene is imaged through a single camera mounted coaxially with a laser. The relationship between the point cloud colorization quality score and scanning speed is illustrated in Figure 30.
Figure 30. The relationship between the quality score and the scanner speed for each tested scanner with the highest tested scan resolution setting (adapted from Publication 3).

Lighting has a significant effect on the colorization quality. Using the HDR imaging mode results in few benefits if the scene is evenly and stably illuminated, such as in our laboratory test environment. When the colorization was done using the default settings without any scene-specific manual adjustments, the LDR imaging mode yielded better results than the HDR mode. However, we can expect the HDR imaging mode to be much more useful in challenging real-life conditions with changing and uneven lighting, such as in typical outdoor environments or in the urban test environment of Publication 2. It is also notable that the current point cloud processing workflows do not fully take advantage of the potential of HDR, since the luminance range (or radiometric resolution) is compressed from 32 bits to 8 bits per channel, causing inaccuracies in the color data. Storing and displaying color information at 8 bits per channel is the most widely supported approach with 3D point clouds. Furthermore, the HDR data was alternatively processed with linear settings for the purpose of assessing the colorization quality with fewer unknown image processing steps compared to automated tone mapping. The observed clear improvement in quality when the linear processing was experimented with suggests that there is the potential to achieve better colorization quality by optimizing and improving the image processing pipeline.

With respect to color, it is important to acknowledge the difference between accurate and visually pleasing color. The goal of automatic image processing is often to produce pleasing color by using artificial sharpening and boosting saturation. Nonetheless, in many applications, particularly in remote sensing, the
unmodified and truthful presentation of radiometric values is typically more im-
portant and useful than pleasing visual appearance and can be considered a ne-
cessity for such use cases as measuring imaging luminance (e.g., Kurkela et al.,
2020; Maksimainen et al., 2020; Kurkela et al., 2021) or more generally in many
applications related to remote sensing.

The TLS instruments are well acknowledged for their geometric accuracy,
which has been investigated and discussed by scholars for more than two dec-
aades. However, to date the research reported in Publication 3 remains one of
the few sensor-level attempts to evaluate the point cloud colorization quality of
TLS from the perspective of producing a high-quality model appearance. Simi-
lar approaches have been made related to photogrammetry by Gaiani et al.
(2017). Publication 3 has encouraged new research efforts in, for example, uti-
lizing TLS for 3D luminance measurements (Kurkela et al., 2021).

Publication 3 provides motivation and information for improving the colori-
zation quality of 3D point cloud data collected with TLS instruments. Further-
more, it reveals objective quality differences between various commercially
available TLS instruments and gives useful information about their producible
quality and its practical implications for “black box” types of systems with pro-
prietary colorization algorithms. Furthermore, the developed method should be
applicable for all colored 3D point clouds, regardless of the method being used
or how the color information was acquired, such as via photogrammetry. Also,
the reported methodology and the use of standardized image quality assessment
metrics could be of interest for instrument manufacturers or more generally in
the geospatial industry.

The developed and studied method assessed colorization quality by analyzing
the resulting colorized 3D point cloud data acquired from the TLS instruments,
but it did not analyze the raw panoramic pictures or the raw image data directly.
Thus, the methodology best reflects a situation where the colored 3D point cloud
is applied or visualized directly, which can often be the case in urban 3D mod-
eling. In the case of producing textured 3D models, the color information in the
3D point cloud can also be transferred to the pixel values of the textures or it
could be used directly as vertex colors. As further limitations, it was not feasible
to experiment with all possible image quality metrics, and the software and
scanner firmware versions used in the research may be subject to change.

Furthermore, Publication 3 did not examine the impact of co-registration er-
rors, or the parallax effects caused by different perspectives, between the LiDAR
sensor and the camera(s) integrated to the scanner. Co-registration errors and
parallax influence the colorization quality by causing offset when mapping the
color data to the point cloud data (e.g., Olsen et al., 2010). In practice this might
lead into discoloration or color mismatches in real-world environments that
contain poles, railings, branches of trees, other thin elements, or edges. Presum-
ably these co-registration errors can be mitigated with sensor calibration and
are more prevalent with scanner designs that have camera(s) mounted to the
scanner frame instead of mounting the camera coaxially with the laser. The sen-
tor calibration is often beyond the control of the scanner user but in this case
the texturing quality could be potentially improved by increasing the number of overlapping scan stations or with additional data processing.

5.4 Future outlook

The interest in utilizing 3D measuring techniques for creating photorealistic 3D models is rapidly growing as photogrammetry and laser scanning are becoming increasingly popular outside such traditional application domains as surveying, engineering, and cultural heritage. As example, the vast number of developer communities in 3D graphics and gaming domains are being encouraged to apply these techniques and the game engine companies are bringing the production of photorealistic 3D assets and application development closer together and more accessible. Examples of such development have been the acquisitions of photogrammetric software and 3D scanning companies (Epic Games, 2019; Epic Games, 2021c). Furthermore, the development of real-time rendering and the streaming of 3D graphics (e.g., Epic Games, 2021d) is making it possible to utilize larger and larger datasets and keep reducing the need for manual data optimizations.

XR (i.e., VR and AR) technologies are still in their infancy, but they have been rapidly developing hand in hand with real-time 3D technology towards more serious applications, such as for virtual training, education, and social events beyond gaming and entertainment. As new generations of hardware are introduced, VR and AR will continue to become intermixed and change the way virtual environments are viewed, interacted with, and experienced. There will be an increasing need for producing compelling, highly detailed, and realistic 3D model content, which is essential in the creation of new XR applications.

The growing interest in creating real-world-based 3D models is being accelerated also by the increasing affordability of sensors, such as the Apple LiDAR (Apple, 2020), or the progressively affordable sensor technology made available by the autonomous driving and robotics industries. Also, sensor platforms such as UAVs are becoming widely available and can help to further automate data collection (e.g., Skydio, 2021). All told, this increase in the affordability and accessibility of 3D measuring is moving the bottleneck from 3D content creation towards utilization. Furthermore, the rapid development of SLAM technology and more seamless integration of LiDAR and photogrammetry are increasing the efficiency of data collection. Over the last few years, commercial mobile SLAM-based sensor systems have emerged that can produce photorealistic data (e.g., NavVis, 2021; Leica Geosystems, 2021b). Also, multispectral or hyperspectral laser scanning shows great potential for removing the need for additional camera sensors altogether. Data from multiple laser wavelengths could be used to produce accurate color information for 3D point clouds directly and independently of scene lighting. Despite being in its infancy, this technology could advance existing applications (e.g., automated recognition of objects or materials) and create completely new applications (e.g., illumination-independent texturing).
The applicability and significance of 3D point clouds is expected to increase. Thanks to advances in point-based rendering and data streaming, the 3D point cloud data is becoming easier to apply directly on real-time 3D platforms (e.g., Schütz, 2016; Epic Games, 2021e) without the need for complex meshing or vectorization workflows. Point clouds are becoming more widely supported in, for example, the new CityGML 3.0 standard via the PointCloud module (Kutzner et al., 2020). This makes the application of close-range 3D measuring techniques more direct and efficient.

Novel methods based on machine learning and deep learning techniques are increasingly being introduced and show great promise for improving the level of automation (Qin & Gruen, 2021), for example in automated object detection (e.g., Shi & Rajkumar, 2020) or in data segmentation (e.g., Qi et al., 2017; Tong et al., 2020). AI-based approaches are also already increasing the potential of using automation to create material maps for PBR or texture de-lighting (Unity, 2021b). In the future, novel deep learning-based approaches, such as neural radiance fields (NeRF), could be used to model and visualize complex real-world scenes much more efficiently (Mildenhall et al., 2020).

Urban 3D model data is becoming increasingly available and is even being distributed for free via online data portals and application programming interfaces (APIs). Not only the data, but also instructions on how to apply the data, are being shared more and more via the Internet in the form of tutorials and online courses. The recent development and introduction of CityJSON is another example of an encoding that is making it easier for developers to use urban 3D models for application development. Likewise, application platforms are constantly improving the utilization of urban 3D model data. Thanks to recent developments, perhaps most notably the collaboration between Epic Games and Cesium (Cesium, 2021b), game engines and virtual globes are converging and many stated obstacles, such as the limited support for geographical coordinates and geospatial data formats (e.g., 3DTiles), are being removed.

Over the last few years, the discussion related to urban 3D modeling has shifted strongly towards digital twins (Ketzler et al., 2020) that feature real-world-based 3D model data as the basis for dynamic data integration. This highlights the need to be able to update the virtual counterpart of a dynamic and functioning city as rapidly as possible and calls for efficient 3D modeling techniques. Furthermore, the requirement of near real-time interactivity between the virtual and physical worlds has increased the relevance of real-time 3D platforms as one of the technical foundations required for creating digital twins. Also in part, the digital twin paradigm and advances in the semantic segmentation of 3D mesh models (e.g., Rouhani et al., 2017) and 3D point clouds (e.g., Perez-Perez et al., 2021) are bringing the semantic and photorealistic modeling and visualization approaches closer together. Also, modeling indoor spaces and indoor-outdoor integration is becoming an increasingly relevant topic (Kang, et al., 2020), expanding the relevance of using close-range 3D measuring techniques. For example, in the city modeling domain this growing demand for indoor modeling has been noted by the inclusion of indoor spaces in the specifications for the upcoming 3.0 version of the CityGML standard (Kutzner et al.,
2020). In the end, as 3D models become more and more detailed and photorealistic and are updated closer to real time, the questions pertaining to privacy, security, and copyright issues will become more relevant and urgent.
6. Summary and conclusions

This dissertation developed ways to improve the utilization of close-range 3D photogrammetry and TLS for photorealistic urban 3D modeling. To fulfill this goal, real-life urban 3D modeling activities were first reviewed to identify relevant application platforms, and city authorities and practitioners were interviewed to understand the internal motivators and barriers to applying photorealistic 3D models. Second, automated multi-sensor urban 3D modeling was studied for a photorealistic web-based 3D application, and the quality of the model’s geometry and texturing was evaluated. Finally, a method was proposed for evaluating the quality of TLS-based 3D point cloud colorization using commercial TLS instruments. Research questions RQ1, RQ2, and RQ3 were answered in publications 1, 2, and 3 using quantitative experiments, a multiple-case study, and supplementary qualitative methods.

As a result, Publication 1 demonstrated that the real-time 3D platforms (i.e., game engines and web-based virtual globes) are the most relevant technology for utilizing photorealistic urban 3D models. These application platforms have an integral role in putting 3D data into beneficial use in cases that require engaging visuality, advanced interactivity, open access, data integration, and versatility for both developers and users. The big picture offered in Publication 1 has been of interest for many researchers (e.g., Murshed et al., 2018; Buyukdemircioglu & Kocaman, 2020; Sun et al., 2020), and its findings have, at least to some degree, influenced the proposed national building standard in Sweden (e.g., Eriksson et al., 2020).

Publication 2 showed that close-range 3D measuring techniques (i.e., close-range photogrammetry, TLS, and a combination of them) can be utilized to produce highly automated and integrated 3D models for applications that require photorealistic urban 3D models. The results supported the view that creating photorealistic 3D models using both photogrammetry and TLS is a good compromise for the quality of the model’s geometry and its appearance. It combined many complementary advantages of both approaches, like the typically superior image quality of photogrammetry and the typically higher geometric quality and direct scale determination of TLS. Prioritizing photorealism sets a high requirement for the visual quality of the models. Automation remains a key challenge to efficiently producing models, and the higher the visual requirements, the more difficult it becomes to automate a task.

A new method for evaluating the quality of a 3D measuring technique for photorealistic applications was proposed and implemented in Publication 3. The
method was developed for evaluating the point cloud colorization quality of modern commercial TLS systems and was studied by benchmarking four commercial panoramic TLS instruments with HDR imaging capabilities. The results showed that it is possible to evaluate the colorization quality even in the case of essentially “black box” types of instruments with proprietary algorithms. Furthermore, the method can be applied to evaluate the colorization quality of virtually any colored 3D point cloud. The quality of the model’s appearance was apparent through the measuring instrument’s capability to reproduce color and detail in the scene. Furthermore, the experiments revealed quality differences between the tested instruments and settings, which can be considered significant especially in use cases that value photorealism and benefit from high-quality and reliable color information. The findings suggested that the measured and visually detectable problems in color reproduction (i.e., differences in color, white balance, and exposure) could be alleviated during the data processing phase, whereas the quality issues related to detail reproduction (i.e., sharpness and noise) were more difficult to fix. Improving the colorization quality of 3D point clouds and the inherent 3D models could make them more useful for traditional application fields, such as surveying, engineering, and cultural heritage, but also open up possibilities in new and emerging application areas, such as in 3D content creation for video games and immersive experiences or in virtual production in the film industry.

As a synthesis, the dissertation proposes that together, close-range 3D measuring methods (i.e., photogrammetry and TLS) and real-time 3D platforms (i.e., game engines and web-based virtual globes) enable the efficient and beneficial creation and utilization of photorealistic urban 3D models. The results highlight the role and significance of real-time 3D technology in developing urban applications and introduces new knowledge and methods for their utilization together with close-range 3D measuring. Clear complementary benefits exist for utilizing 3D measuring technologies and real-time 3D platforms together. For photorealistic use cases, it is important to understand the quality of a 3D model as a combination of geometry and appearance. The model’s appearance, represented by texture images or color values, is a crucial element in the photorealistic 3D modeling of real-life environments and its significance has often been overlooked. Investing more effort in improving the quality of a model’s appearance could also enable new use cases in many remote sensing-related applications where the truthful and unmodified presentation of radiometric values can be considered more important than visual appearance.

Detailed and photorealistic 3D modeling of real-world environments still often relies on manual modeling approaches, which can be expensive, labor intensive, and even a barrier to entry for developers, businesses, and organizations. This dissertation improves knowledge on how both dimensionally and visually accurate real-world-based 3D data can be efficiently created and utilized. However, realizing the potential benefits of urban 3D modeling is not only a technological challenge; it also often involves economic and social aspects that are use-
ful to acknowledge. The results could both help and motivate developers, experts, and practitioners to create and apply photorealistic urban 3D models in both new and well-established fields.

Despite the fact that close-range photogrammetry and TLS are increasingly efficient and accessible measuring techniques, there is a need to develop better, more accessible, and more automated tools and workflows to increase the quality, production efficiency, and applicability of photorealistic 3D models even further. The demand for visually and dimensionally realistic environments is increasingly driven by global trends, such as digital transformation, robotization, XR, and digital twins. This is stimulating rapid development in multiple areas, enabling larger audiences to create and utilize 3D models for a wider range of use cases in growing and partly unforeseen application fields with diverse requirements. This development will ensure that there is an increasing need for robust quality control and for skilled and well-informed people to create and use photorealistic 3D models now and in the future.
References


Baumgart, B.G. (1972). Winged edge polyhedron representation. Stanford University, Department of Computer Science, 1-44.


References


References


