5G-Enabled Mobile AR System for Location-aware Assistance in Buildings – System Overview and Preliminary Evaluation

Baris ÖZCAN, Jan GROTTKE, Zhenyu LIU, Christoph BLUT, Ralf BECKER and Jörg BLANKENBACH, Germany

Key words: 5G, Real-Time Streaming, Augmented Reality (AR), Technical Building Equipment (TBE), Digital Transformation

SUMMARY

The increasing complexity of technical building equipment (TBE) requires advanced digital solutions to support technicians in tasks such as planning, installation, maintenance, and inspection inside buildings. The N5GEH-LocI4AR project addresses this demand by integrating the 5G communication standard, Internet of Things (IoT), and Augmented Reality (AR) technologies into an intelligent location-aware mobile assistance system designed to provide real-time support during complex technical operations. This system is specifically intended to help technicians navigate large installations, identify components, access relevant technical data, and monitor sensor readings in real time.
The core components of the system include coarse indoor positioning, fine-grained localization and pose tracking, and AR streaming and visualization, which collectively enable context-aware and real-time support for technicians. Coarse positioning relies on sensor fusion, combining data from inertial measurement units (IMUs), WLAN, Bluetooth, and 5G fingerprinting to establish an initial location estimate independent of GNSS. This estimate is then refined through fine pose tracking, which ensures the precise spatial alignment of AR content—such as BIM models of the TBE—with the technician's physical surrounding. The AR streaming and visualization component utilizes 5G to offload computationally intensive tasks to a server or cloud, enabling the real-time delivery of high-resolution 3D content to mobile devices while maintaining minimal latency and optimized device performance.
In this paper, we present the design, development, and preliminary evaluation of the N5GEH-LocI4AR system, focusing on its core components for providing context-aware AR overlays in real-world environments. Initial results demonstrate the capability of the system to deliver stable performance across devices, with high frame rates and reliable indoor positioning accuracy. Future work will aim to further optimize the fine localization process, enhance AR content stability, and conduct comprehensive evaluations of workflow latency to ensure robust deployment in practical applications.

5G-Enabled AR Streaming System for Mobile Assistance in Technical Building Equipment (TBE) – System Overview and Preliminary Evaluation

Baris ÖZCAN, Jan GROTTKE, Zhenyu LIU, Christoph BLUT, Ralf BECKER and Jörg BLANKENBACH, Germany

1. INTRODUCTION

The energy sector is undergoing a transformative shift driven by the decentralization of energy supply systems, a necessity highlighted by Germany's "Energiewende" initiative. This shift introduces new demands on information and communication technology (ICT) infrastructures, which needs to be modernized to support the management and maintenance of increasingly complex energy systems. To address these challenges, the German government initiated the National 5G Energy Hub (N5GEH), a project leveraging 5G and advanced digital technologies to improve the efficiency and sustainability of energy systems.

As part of the N5GEH initiative, the N5GEH-LocI4AR subproject focuses on developing a 5G-enabled, location-based mobile assistance system. This system supports planners, technicians, and installers working in environments where conventional GNSS-based systems are unreliable. By integrating 5G, IoT, and AR technologies, it enables real-time access to essential building and sensor data, facilitating tasks such as navigation, component identification, and maintenance within complex energy installations.

This paper centers on the core elements of the N5GEH-LocI4AR system. The framework allows real-time access and precise alignment of digital content, such as BIM models, with the technician's physical environment. It comprises three key components: coarse indoor positioning, fine-grained indoor localization and pose tracking, and 5G-based real-time streaming. These components work together to ensure accurate positioning, efficient data delivery, and intuitive AR overlays, enhancing the efficiency of technical operations.

The remainder of this paper is structured as follows: Section 2 presents the foundational and state-of-the-art technologies integrated into the N5GEH-LocI4AR system. Section 3 provides an overview of the N5GEH-LocI4AR project, highlighting its objectives and key innovations. Section 4 focusses on the AR streaming and visualization framework, covering its core components on coarse indoor positioning, fine-graned indoor localization and pose tracking, and 5G-based real-time streaming. Section 5 presents preliminary testing results, and Section 6 concludes the paper with insights into future work.

2. TECHNOLOGICAL FOUNDATION AND STATE-OF-THE-ART

This section provides an overview of the foundational and state-of-the-art technologies integrated into or related to the N5GEH-LocI4AR project, including indoor positioning systems, 5G and technology, augmented reality (AR), and mobile applications for AR. These technologies are specifically utilized within the AR streaming and visualization framework to enable accurate indoor localization, rapid data transmission, and seamless AR-based visualization.

2.1 Technologies for Indoor Positioning Systems

Various technologies enable automated indoor localization (Blankenbach 2017), including WLAN, ultrasound, Ultra-Wideband (UWB) (Norrdine et al. 2013) and artificial magnetic fields (Blankenbach & Norrdine 2010a). IMU-based methods further contribute to localization by estimating position and orientation relative to a starting point, though high-quality IMUs are required to mitigate drift over time, limiting their use in large-scale, low-cost applications (Becker 2015).

Sensor fusion combines data from multiple sensors to improve localization accuracy and robustness. By integrating complementary data sources, such as inertial, magnetic, and environmental sensors, alongside radio-based methods like WLAN or Bluetooth Low Energy (BLE), sensor fusion can provide a more comprehensive estimate of position and orientation. This approach is particularly valuable in environments where individual sensors may face limitations, such as signal attenuation or drift. The flexibility of sensor fusion also allows for adaptability across diverse use cases and hardware configurations.

Smartphone-based sensor fusion has become a practical solution for pedestrian localization. Using low-cost micro-electromechanical systems (MEMS) sensors (IMU, magnetometer, barometer) in combination with WLAN, BLE, and digital building models, smartphones can achieve reliable indoor positioning (Grottke & Blankenbach 2021). Techniques such as sequential Monte Carlo (SMC) filters or particle filters further enhance data fusion accuracy (Real-Ehrlich 2019; Willemsen, Keller & Sternberg 2015).

While commercial solutions exist for specific use cases—such as visitor guidance using Ekahau for WLAN (Ekahau 2024), asset tracking with Ubisense for UWB (Ubisense 2024), and indoor navigation with Indoors for BLE (Indoors 2024)—these systems typically rely on a single technology without integrating multiple sensor inputs. In contrast, approaches that incorporate sensor fusion, combining data from diverse sources like IMUs, WLAN, and BLE, offer enhanced accuracy and robustness, particularly in complex environments. Adapting 5G for indoor positioning in energy technology remains largely unexplored, presenting opportunities for further innovation.

2.2 5G Technology

The 5G New Radio (5G NR) communication standard (Dahlman, Parkvall & Skold 2018) supports enhanced mobile broadband (eMBB), massive machine-type communication (mMTC), and ultra-reliable, low-latency communication (uRLLC). These advancements enable high data rates and latency below one millisecond. Enhanced bandwidth allows finer spatial resolution, while beamforming through advanced antenna arrays supports precise positioning using signal timing, angles of arrival (AoA), angles of departure (AoD), and signal propagation time (Time of Flight, ToF). Techniques such as Round-Trip Time (RTT) and Time Difference of Arrival (TDoA) leverage ToF to achieve centimeter-level localization accuracy, particularly in environments where traditional GNSS systems are unavailable.

2.3 Augmented Reality (AR)

AR overlays virtual data onto the real world in real-time, involving three core elements: 1) visualization of georeferenced data, 2) perception of the real environment, and 3) alignment of virtual and physical worlds. Game engines such as Unreal Engine (Unreal Engine 2024) or Unity (Unity 2024) facilitate virtual rendering, with mobile implementations using optical see-through for headsets or video see-through for mobile devices like smartphones.
Pose determination is critical for AR, requiring precise, real-time tracking of device position and orientation. While GNSS-IMU integration serves outdoor applications (Schall 2009), indoor systems often depend on WLAN, BLE, or infrared. Building information systems often leverage precise indoor positioning methods, such as those utilizing WLAN or infrared, to enable seamless integration of virtual and physical data in AR environments (Blankenbach & Norrdine 2010b). Recent advances in visual tracking, such as marker-based and natural feature tracking (NFT) approaches, minimize infrastructure reliance and improve accuracy. Systems like ARCore (Google AR 2024) and ARKit (Apple AR 2024) utilize visual-inertial odometry to align AR content, though achieving long-term stability remains a challenge. Integrating localization data with NFT is well-supported by 5G, allowing for precise and stable pose tracking essential for realistic AR applications.

2.4 Applications on mobile devices

Both mobile operating systems iOS and Android support robust AR development through ARKit and ARCore, respectively. These frameworks facilitate environment recognition, motion tracking, and light analysis, enabling accurate integration of 2D and 3D objects into live images. Cross-platform frameworks such as Unreal Engine and Unity extend these capabilities across devices by incorporating ARCore or ARKit functions.
AR data can be divided into two main types. Dynamic objects are processed locally on mobile devices, while more complex models, such as 3D scans or BIM models, often require significant storage and computational resources, making local processing impractical. These larger models are typically stored in Digital Asset Management (DAM) systems and accessed on demand.
The high-speed and low-latency connectivity of 5G fundamentally changes how these large data models are utilized. With 5G, 3D content can be seamlessly accessed and rendered in real time, reducing the reliance on device hardware capabilities. Computationally intensive tasks, such as rendering and data processing, can be offloaded to cloud platforms or servers. This approach minimizes the computational demand on mobile devices while ensuring the delivery of smooth, real-time AR experiences, even for data-intensive applications. However, it requires a stable online connection with 5G to ensure consistent performance and uninterrupted access to large data models.

3. PROJECT OVERVIEW

The project N5GEH-LocI4AR, supported by Germany's Federal Ministry for Economic Affairs and Climate Action (BMWK), aims to drive digitalization and efficiency improvements in the energy sector by integrating modern ICTs such as 5G, IoT, and AR. The project consortium includes industry experts, academic researchers, and small and medium enterprises (SME), which collaborate on developing a mobile assistance system designed to support technicians with planning, installation, commissioning, and maintenance of complex energy systems.

3.1 Project Objectives

The primary objective of the N5GEH-LocI4AR project is to create a location-based indoor mobile assistance system for devices like smartphones and tablets. This system combines accurate indoor positioning with AR-based visualization, enabling technicians to navigate complex technical environments and access essential, context-relevant information in real-time. Tailored specifically to meet the demands of industrial energy applications, the system is intended to overcome limitations of conventional GNSS-based solutions in indoor settings, where GNSS signal is unreliable or unavailable. By providing precise localization and intuitive AR overlays, the system enhances the spatial perception and operational efficiency of technicians.

3.2 Key Innovation Areas

The N5GEH-LocI4AR project integrates 5G with advanced technologies, including indoor positioning, AR, and IoT, to develop a mobile assistance system tailored for technicians in complex TBE environments. Its innovations focus on three key areas:

(1) Real-Time Localization with 5G and Low-Cost Sensors

The system performs flexible, real-time localization without requiring a fixed start point, combining data from IMUs (e.g., accelerometers, gyroscopes), Bluetooth, WLAN, and 5G. Coarse positioning combines data from these sources to estimate an initial location, providing a reliable starting point in complex indoor environments. Sensor fusion algorithms account for signal reflections and delays, maintaining continuous and reliable positioning. Building on this initial positioning, fine-grained localization integrates natural feature tracking (NFT) and a BIM model to establish a refined localization with centimeter-level accuracy, which ensures precise AR alignment. This multimodal approach facilitates for accurate navigation and AR functionality in environments where GNSS is unavailable or unreliable.

(2) 5G-Based AR Streaming for Real-Time Visualization

The system leverages the high bandwidth and low latency of 5G to offload intensive tasks such as AR rendering, and large-sized data such as BIM models, to a cloud platform. Processed AR content is then streamed back to the mobile device (e.g., smartphone), allowing for smooth, responsive overlays that align precisely with the physical surroundings of the user. Positioning data is further refined using visual tracking (NFT), supporting accurate AR alignment for tasks like installation and maintenance.

(3) Real-Time IoT Data Integration with 5G for On-Site Monitoring

The system integrates real-time IoT data from environmental sensors, such as temperature, humidity, and equipment status, transmitted over 5G. These IoT data points not only provide essential operational parameters for technicians but also contribute to the indoor positioning process, offering supplementary location information via Bluetooth and RFID beacons. The live data feed enhances situational awareness, allowing technicians to monitor environmental conditions in real time.

3.3 Practical Testing and Evaluation

The N5GEH-LocI4AR system is intended to undergo practical testing at the Technology and Innovation Park Nordheide (TIP) in Germany, where real-world energy infrastructure offers an ideal setting for validating the system's performance. These tests will evaluate accuracy, responsiveness, and ease of use in real-world conditions that demand precise integration of sensor data and dynamic AR content. The outcomes are anticipated to demonstrate the project's potential for enhancing productivity and accuracy in complex technical environments.
Through its innovative combination of 5G-based localization, AR streaming, and IoT data integration, the N5GEH-LocI4AR project supports Germany's climate objectives by promoting resource-efficient practices and improved energy infrastructure maintenance.

4. AR STREAMING AND VISUALIZATION FRAMEWORK

This section outlines the workflow of the AR streaming and visualization framework of the N5GEH-LocI4AR system, followed by details on its key components: coarse indoor positioning, fine-grained indoor localization and pose tracking, and 5G-based real-time streaming.

4.1 Workflow Overview

In the framework, real-time localization, pose tracking, and AR visualization are distributed between the mobile device (client) and the server (Fig. 1). The mobile device collects raw sensor data from IMUs, barometric sensors, magnetometers, and Bluetooth/WLAN signals to establish room-level indoor positioning. This data is processed locally on the device, with the calculated position sent to the server. The mobile device also transmits the live camera feed and user interactions, such as selecting elements or triggering pose corrections.
The server refines the initial coarse positioning by matching visual features from the camera feed with elements of the BIM model, like the corners of a door, achieving precise alignment between virtual and real-world components. Once the initial alignment is set, continuous pose tracking ensures ongoing accuracy, with corrections made only when significant drift occurs, by again matching visual features from the camera feed to the BIM model.
Finally, AR visualization is generated either on the server, where BIM content is combined with the camera feed and streamed back to the device, or directly on the mobile device, which overlays the BIM-rendered content onto the live camera feed.

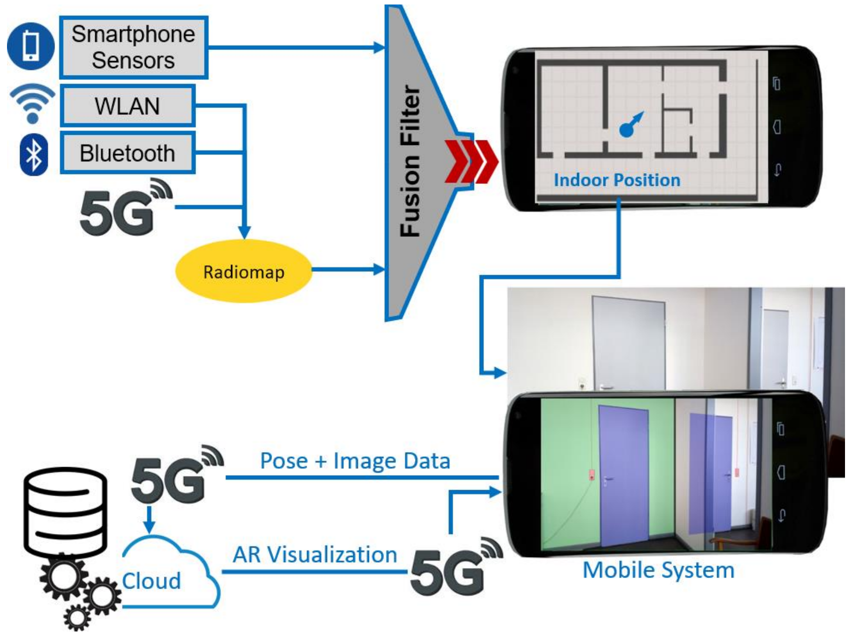

Fig. 1: System Overview of the AR Streaming and Visualization Framework.

4.2 Coarse Indoor Positioning

To enable accurate coarse positioning within indoor environments, a raster-based particle filter was developed to perform real-time pedestrian positioning by integrating multiple independent localization methods into a unified spatial probability map, represented as a 2D radio map. Here, absolute position information from different sources—such as 5G signals, magnetic field distortions, and WLAN signal strengths—is converted into distinct probability distributions over the building layout. A BIM model further refines this map, designating navigable and non-navigable areas to exclude physically unreachable positions, like walls (Fig. 2, left).

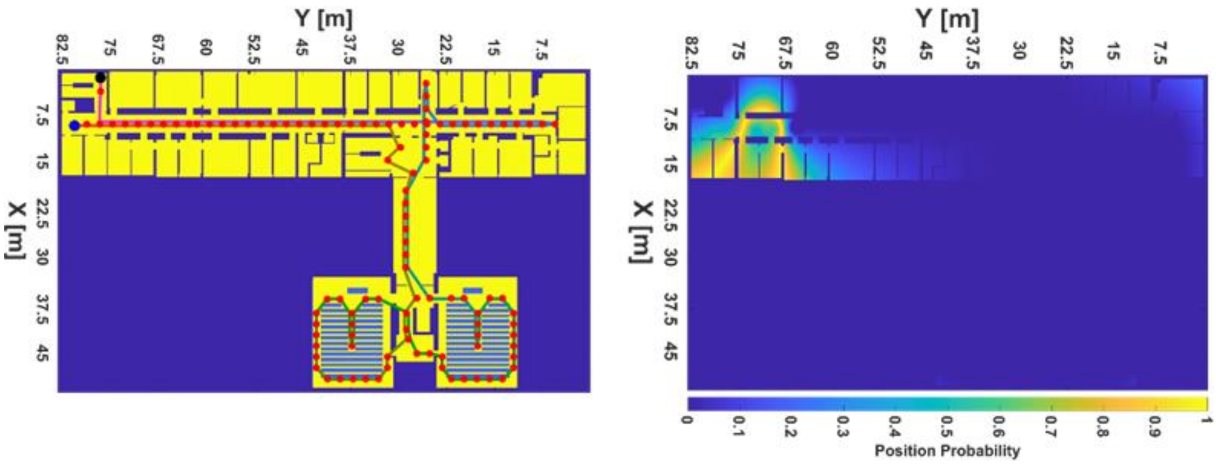

Fig. 2: Left: Rasterized model of a test building with a cell size of 15 cm and 555x336 cells; yellow for walkable areas, blue for non-walkable areas, and light blue for benches, right: Map of estimated positions based on WLAN fingerprinting.

Each data source contributes a spatial probability distribution indicating the likelihood of the device's location. Using Bayesian methods, these distributions are combined and normalized to create a composite probability map that highlights the most probable device location (Fig. 2, right). This map fusion is computed on a graphics processor using element-wise matrix multiplications, enabling real-time performance.
The filter dynamically weighs each data source based on its reliability, optimized through test trajectories and refined using genetic algorithms for adaptation to each building's unique signal environment. To track device movement, data from IMUs (accelerometer, gyroscope, magnetometer) capture positional changes relative to the building layout via convolution operations, also processed on the Graphics Processing Unit (GPU) for speed.
The step size and direction (including uncertainties) are represented as a two-dimensional displacement pattern, serving as the convolution kernel to refine the current probability distribution (Fig. 3). Singular Value Decomposition (SVD) optimizes the convolution process by focusing only on the most significant singular values, thereby reducing computation load. By simplifying data into matrix operations, the system delivers rapid, stable real-time positioning estimates that inform the fine localization stage for precise AR alignment.

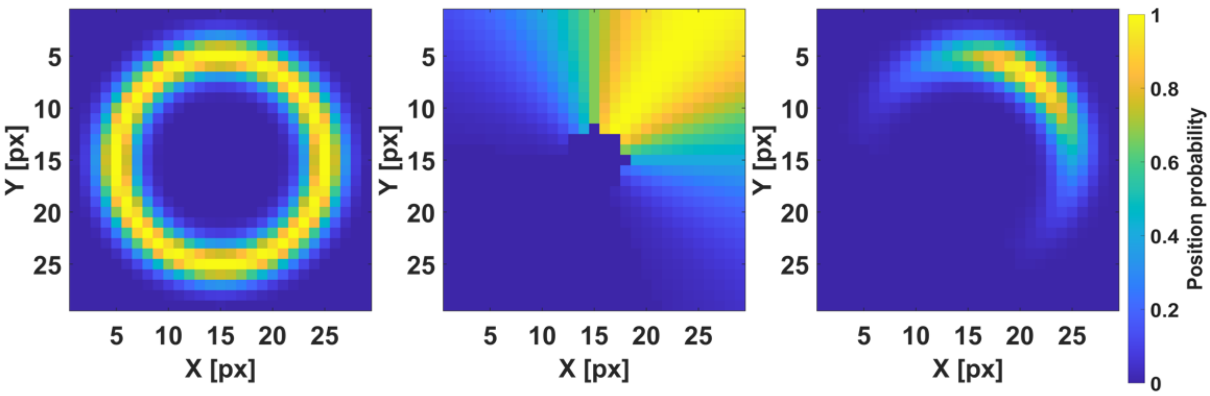

Fig. 3: Pedestrian Dead Reckoning (PDR) grids: left - step size distribution, center - heading distribution, right - step probability distribution.

4.3 Fine-grained Localization and Pose Tracking

The pose-tracking and orientation procedure depends on a BIM-based visual localization method for AR environments, with a smartphone serving as the AR mobile device. The core objective is to achieve precise alignment between virtual and real-world elements by leveraging natural visual features from objects like doors, including door frames, along with georeferenced Building Information Modeling (BIM) models.
The process starts with coarse positioning, which identifies the correct room in the BIM model. This is achieved by integrating localization data, ensuring the AR system selects the appropriate georeferenced environment. Since each room is associated with at least one door in the BIM model, the fine positioning begins with the selection of a reference door object in the AR environment.

Through user interaction, the desired door object is selected. Once selected, all vertices from the corresponding door object in the BIM model are extracted. To address potential inconsistencies in the BIM data, a minimum bounding box algorithm is applied to identify the eight corners of the door. From these, only the four corners of the front surface are retained for further processing.
Using a pre-trained YOLOv5 model (Jocher et al. 2020), doors are detected within images captured by the smartphone camera. YOLOv5 divides each input image into grids and predicts bounding boxes and class probabilities for objects within these grids. If a door is detected, it returns a region of interest (ROI) containing the door, which narrows down subsequent processing areas.
Inspired by AprilTag algorithms (Olson 2011), the 2D coordinates of door corners are extracted by detecting edges and contours within the ROI. The process includes bilateral filtering to enhance edges and reduce noise, followed by edge detection using Canny's algorithm. Mathematical morphology operations (Matheron & Serra 2002; Balado et al. 2020; Liu et al. 2022) connect broken edges to form complete contours from which convex quadrilateral contours are filtered out as potential doors (Fig. 4).
The final step uses Perspective-n-Point (PnP) methods (Terzakis & Lourakis 2020) to estimate the pose of the smartphone based on matching 2D-3D point pairs obtained from previous steps (Blut & Blankenbach 2020). The intrinsic camera parameters and distortion coefficients are considered during this phase to minimize reprojection errors. A state-of-the-art PnP solver then computes rotation matrices and translation vectors that update the camera pose within the BIM model coordinate system.

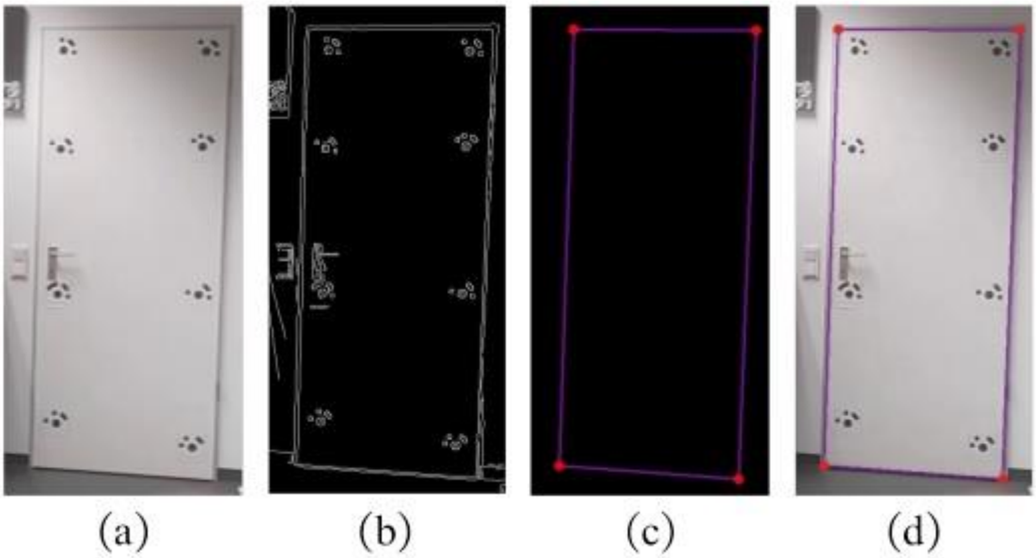

Fig. 4: Door corner detection in the image: (a) input image, (b) extract edges, (c) extract convex quadrilateral contour, (d) obtain the outline of the door and extract the four door corners.

4.4 5G-based Real-Time Streaming

The streaming system is built on Unreal Engine 5, utilizing its Pixel Streaming plugin, which relies on WebRTC (Web Real-Time Communication) and the RTP (Real-Time Protocol) for communication. A signaling and web server manages the connection and data exchange between the mobile device (client) and the server, establishing a peer-to-peer (P2P) connection. Once the connection is established, the mobile device transmits data, including coarse positioning (see section 4.2) and camera image (Fig. 5). Future plans include integrating user interactions, such as touch-based selections of BIM model elements, for example, to filter or highlight specific building components or retrieve additional information about them, or button-based trigger for pose corrections. The data is streamed as structured JSON objects to the server via the WebRTC connection through a fast, low-latency communication channel, with 5G planned for the final system.

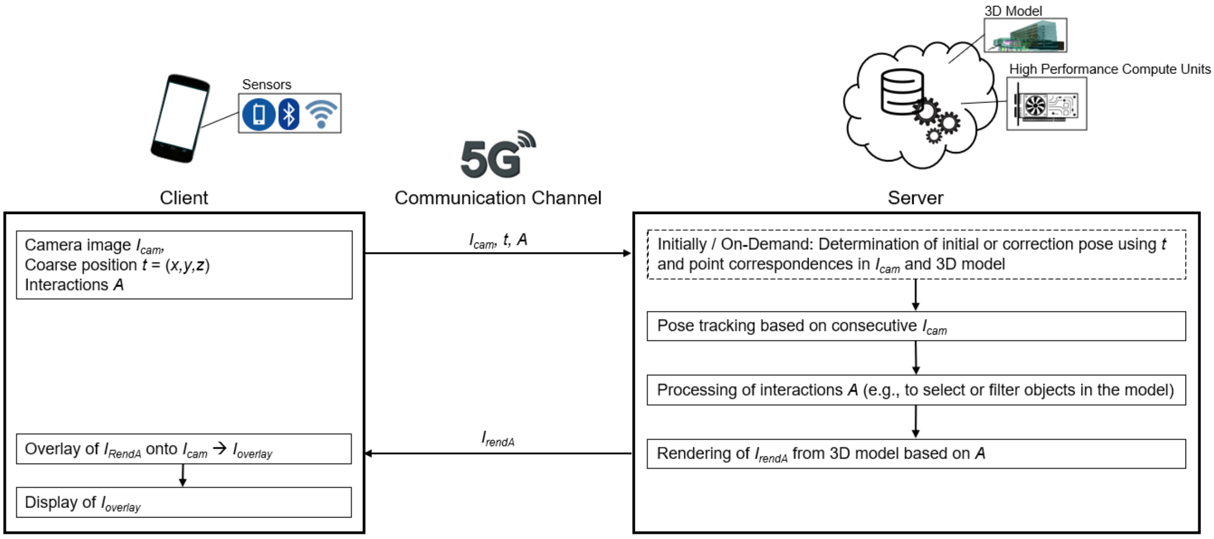

Fig. 5: Concept of the workflow for 5G-based Real-Time Streaming (Architecture R2).

On the server side, a dedicated workflow within Unreal Engine is implemented to process the incoming data streams. The data is initially parsed and separated into its components, such as coarse positioning and camera images, before being further processed. The coarse positioning data is crucial for the fine-grained localization and pose tracking (see section 4.3), as it establishes the initial position of the AR device. The camera image serves two purposes: first, it is likewise used for localization and pose tracking (see section 4.3), and second, it can be used for a server-side combination with the BIM-rendered content to generate the AR view (see following rendering architecture R1).
To create the AR view, consisting of BIM-rendered scene overlapping with the camera image, two rendering architectures were implemented and investigated (see Fig. 6):

- **R1 (Server-Side Overlay)**: After fine-grained localization and pose tracking (see section 4.3), the camera image is combined with BIM-rendered content on the server and then streamed back to the client. While this minimizes client-side processing, it introduces a minor delay as the entire AR image must be transmitted, leading to a slight lag between the real environment and the AR visualization. Additionally, the current widget-based overlay setup places the camera image over the BIM content, leading to an incorrect layering order, and superimposes all BIM content, including the background, without distinguishing between relevant and irrelevant sections.

- **R2 (Client-Side Overlay)**: In this architecture, only the BIM-rendered content is streamed to the client, which handles the overlay. Although this slightly increases client processing demand, it ensures the displayed camera feed is live. This may cause a shift or misalignment between the camera feed and the BIM content. However, the low-latency of 5G enables near real-time synchronization, minimizing such misalignments and ensuring a seamless AR experience. A chroma keying technique renders the BIM scene's background in a uniform color (e.g., green), which the client filters out, leaving only relevant BIM objects visible.

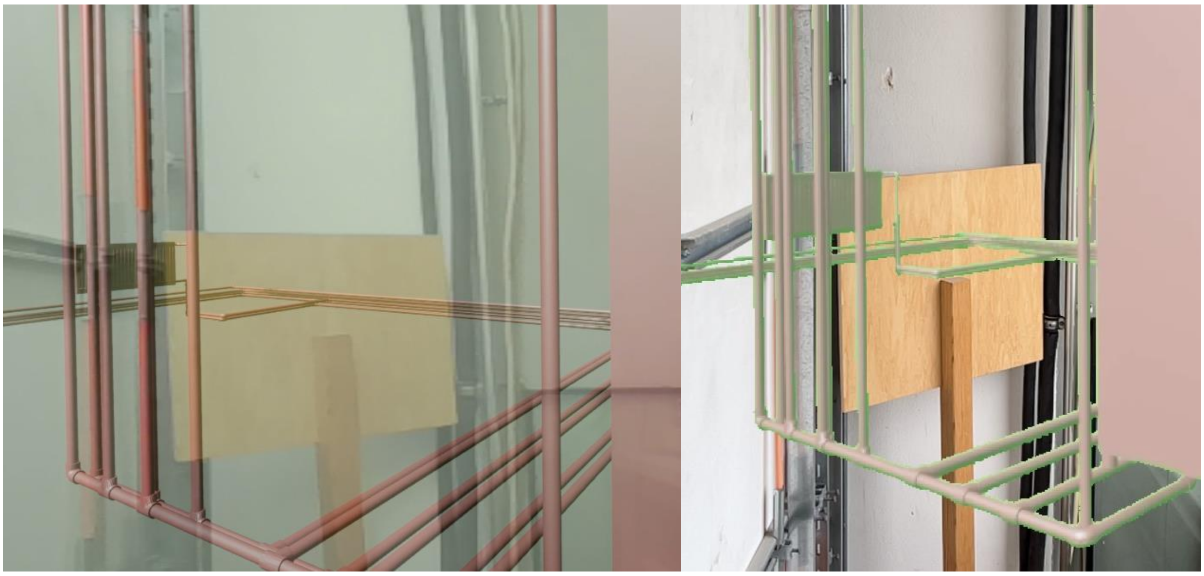

Fig. 6: Examples of both server-side overlay (R1) and client-side overlay (R2).

5. SYSTEM TESTING AND RESULTS

This chapter presents the testing and evaluation results for the AR streaming and visualization framework of the N5GEH-LocI4AR project, focusing on the performance of the coarse indoor positioning, fine-grained localization and pose tracking, and 5G-based real-time streaming under real-world conditions.

5.1 Evaluation of Coarse Indoor Positioning

To assess the accuracy and reliability of the coarse indoor positioning, data collection trials were conducted within a multi-level building, using a Samsung Galaxy S7, running Android Oreo (8.0.0) and equipped with an Exynos 8890 eight-core CPU.
As a reference, three BLE beacons (Bluetooth v4.3) were placed near stairways, and 96 WLAN access points (IEEE 802.11ac, 2.4/5 GHz) were utilized. The test environment spanned a single floor, marked with 107 ground truth points (GTPs) for accuracy validation (Fig. 7, red dots, the start in black, and the end in blue). Eight runs were completed—five for testing and three for

evaluation—with an average walking speed of 1.6 m/s over approximately 580 steps. The building layout was defined in a local coordinate system and is mapped onto a 15 x 15 cm raster grid, resulting in a resolution of 555 x 336 cells.

For data recording, an in-house application was used to control the sampling rate of each sensor signal, including accelerometer, gyroscope, magnetometer, Wi-Fi, and BLE, along with optional data from the barometer, proximity, and light sensors. Each sensor's data was saved as a separate CSV file with timestamps, and GTPs were noted by manual input. A total of 157 GTPs were used as control points, with some points sampled multiple times.

In post-processing, the IMU sensor data was interpolated at 100 Hz using spline interpolation to achieve equidistant data points, and a moving-median filter with a 0.31-second width was applied to remove outliers—this value, based on the median step duration, was optimized for consistent step frequency estimation.

The results in Fig. 7 highlight key positioning error metrics across different methods using empirical cumulative distribution function (ECDF) curves and violin charts. The infrastructure-free method (orange) achieved the lowest median error at 1.05 m and a mean deviation of 1.23 m. Its root mean square error (RMSE) of 1.51 m and third quartile (Q3) value of 1.63 m further demonstrate its consistent accuracy. The BLE-WLAN-LTE-Magnetic Field variant (yellow) performed comparably, with a median error of 1.23 m, mean error of 1.31 m, RMSE of 1.53 m, and Q3 value of 1.70 m. In contrast, the infrastructure-based method (blue) had slightly higher deviations, with a median of 1.33 m, mean of 1.44 m, RMSE of 1.72 m, and Q3 value of 1.91 m. The ninety-five-percentile (2σ) errors ranged from 2.74 m for the BLE-WLAN-LTE-Magnetic Field variant to 3.26 m for the infrastructure-based method.

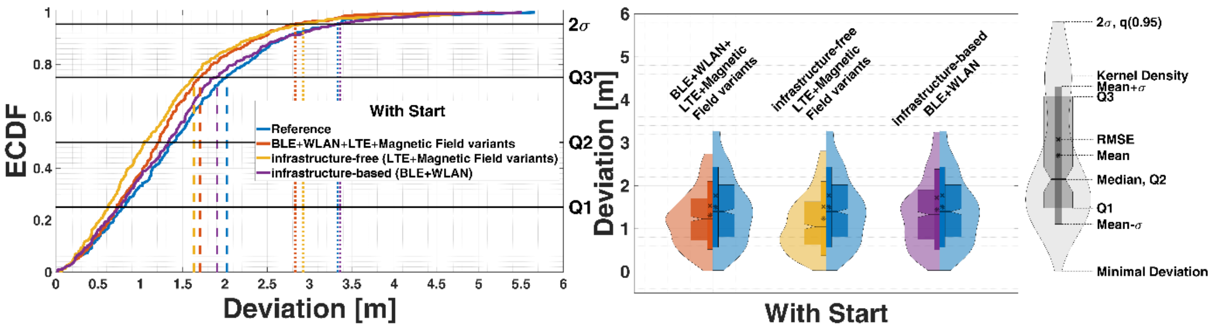

Fig. 7: Left: ECDF of positioning errors for different methods, showing key metrics (Q2, Q3, and 2σ). Right: Bar chart summarizing deviation metrics (median, mean, RMSE, Q3, and 2σ).

5.2 Evaluation of Fine-grained Localization and Pose Tracking

For the evaluation of the fine-grained localization and pose tracking two BIM models were used: one with a flat door frame and another with a protruding frame. Experiments were conducted at four orientation angles (0°, 10°, 20°, and 30°) and ground truth was obtained using a ring-marker localization system (Fig. 8). The method achieved an average translation error in the lower centimeter level and a rotation error of lower single-digit degree range (Fig. 9). The alignment process is performed instantaneously, with a typical execution time of under one second.

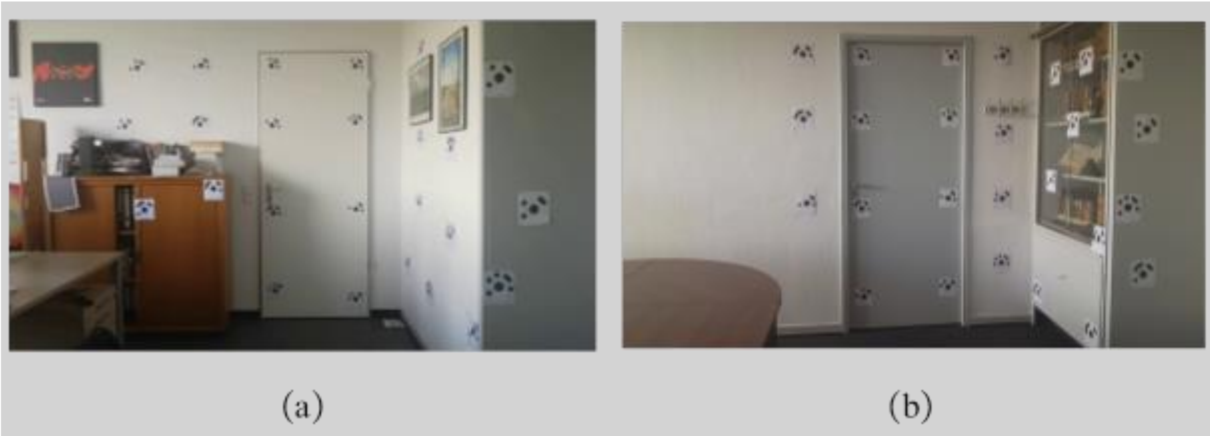

Fig. 8: Ring marker localization system.

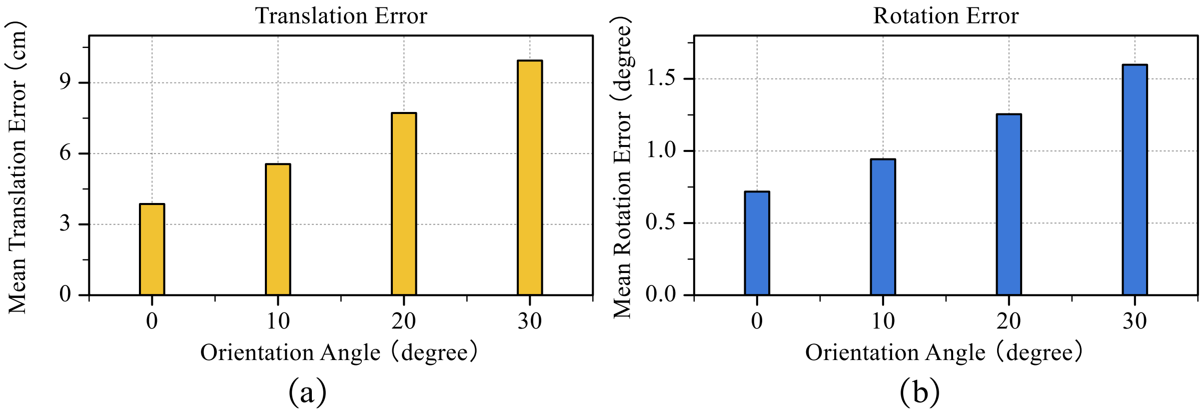

Fig. 9: Top: Mean translation (a) and rotation (b) errors of the door at different orientation angles (0°-30°).

5.3 Evaluation of 5G Real-Time Streaming

This section quantifies the performance of the AR streaming system, focusing on reliability, throughput, and visual quality for real-time AR applications. While the system is designed to utilize 5G for its low latency and high bandwidth capabilities, the evaluations presented here were conducted using a local WLAN network due to the unavailability of a suitable 5G setup during testing. Despite this limitation, the data transmission pipeline was fully operational, with data packages being transmitted without loss, ensuring reliable communication between the client and server.

In the initial evaluation of the R1 (Server-Side Overlay) architecture, a laptop served as the client device, using its webcam to simulate the smartphone camera. This setup achieved a frame rate of 30 fps, delivering a smooth AR experience with no noticeable delay or stuttering, indicating a robust real-time performance. Although no further quantitative analysis has been performed for this configuration, the high frame rate and seamless AR imagery highlighted the suitability of the R1 architecture for stable, real-time AR applications.

For the R2 (Client-Side Overlay) configuration, tests were conducted on an iPhone 13 and a Samsung Galaxy S20 using an 80% compression rate for JPG images. The iPhone 13 achieved frame rates of 20.3 fps at a resolution of 1280x720 and 21.1 fps at 640x480, indicating consistent performance across these resolutions. Quantitative results for the Samsung Galaxy S20 were not evaluated during this phase.

6. CONCLUSION

In this paper, we presented the core components of the AR streaming and visualization framework for the N5GEH-LocI4AR system. This framework is combines 5G, and AR technologies for coarse indoor positioning, fine-grained localization and pose tracking, and real-time streaming capabilities to support technical tasks within TBE systems. Initial evaluations demonstrated the potential of the system in achieving reliable indoor positioning, robust fine-grained localization and pose tracking, and effective AR streaming, highlighting the suitability of this integrated approach for real-world applications.

As development continues, future work will focus on further integrating IoT data for sensor visualization in real-time, enhancing on-site monitoring. The integration of 5G for streaming is crucial to ensure high-bandwidth, low-latency performance. Additionally, quantitative evaluations will assess workflow efficiency, with a focus on latency, particularly under 5G conditions.

7. ACKNOWLEDGEMENTS

The authors thank all partners namely Internet Marketing Services GmbH, IOX GmbH and Wirtschaftsförderung im Landkreis Harburg GmbH for trustful cooperation. We gratefully acknowledge the project funding from the German Federal Ministry for Economic Affairs and Climate Action under the EnOB: N5GEH-LocI4AR funding measure (grant number 03EN1059A-D).

REFERENCES

Apple AR. (2024). Retrieved November 25, 2024, from https://developer.apple.com/augmented-reality/.

Balado, J., van Oosterom, P., Díaz-Vilariño, L., & Meijers, M. (2020). Mathematical morphology directly applied to point cloud data. ISPRS Journal of Photogrammetry and Remote Sensing, 168, 208–220. https://doi.org/10.1016/j.isprsjprs.2020.08.011

Becker, D., & Becker, M. (2015). A study of non-stochastic IMU errors in strapdown airborne gravimetry. In: DGON Inertial Sensors and Systems Symposium (ISS), IEEE, pp. 1–18.

Blankenbach, J. (2017). Indoor-Positionierung & lokale Positionierungssysteme. In: Brand, M., Blankenbach, J., & Kolbe, T. H. (Eds.), Leitfaden - Mobile GIS: Von der GNSS-basierten Datenerfassung bis zu Mobile Mapping

(Vol. 6, No. 3). München, Germany: Selbstverlag, Runder Tisch GIS e.V. Retrieved from https://rundertischgis.de/publikationen/leitfaeden.html#a_mobilegis.

Blankenbach, J., & Norrdine, A. (2010a). Position estimation using artificially generated magnetic fields. In: Proceedings of the 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE. Catalog Number: CFPI009J-ART.

Blankenbach, J., & Norrdine, A. (2010b). Building information systems based on precise indoor positioning. Journal of Location Based Services, 5(1), 22–37. https://doi.org/10.1080/17489725.2010.538016

Blut, C., & Blankenbach, J. (2020). Three-dimensional CityGML building models in mobile augmented reality: A smartphone-based pose tracking system. International Journal of Digital Earth, 14(1), 32–51. https://doi.org/10.1080/17538947.2020.1733680

Dahlman, E., Parkvall, S., & Skold, J. (2018). 5G NR: The Next Generation Wireless Access Technology. Academic Press.

Ekahau. (2024). Retrieved November 25, 2024, from http://www.ekahau.com.

Google AR. (2024). Retrieved November 25, 2024, from https://developers.google.com/ar.

Grottke, J., & Blankenbach, J. (2021). Evolutionary optimization strategy for indoor position estimation using smartphones. Electronics, 10(5), 618. https://doi.org/10.3390/electronics10050618.

Indoors. (2024). Retrieved November 25, 2024, from https://indoo.rs/.

Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., Michael, K., TaoXie, Fang, J., imyhxy, Lorna, Yifu, Z., Wong, C., V, A., Montes, D., Wang, Z., Fati, C., Nadar, J., Laughing, … Jain, M. (2020). Ultralytics/yolov5: V7.0—YOLOv5 SOTA Realtime Instance Segmentation (Version v7.0) [Computer software]. Ultralytics Inc. https://doi.org/10.5281/zenodo.7347926

Liu, Z., van Oosterom, P., Balado, J., Swart, A., & Beers, B. (2022). Detection and reconstruction of static vehicle-related ground occlusions in point clouds from mobile laser scanning. Automation in Construction, 141, 104461. https://doi.org/10.1016/j.autcon.2022.104461

Matheron, G., & Serra, J. (2002). The birth of mathematical morphology. In Proceedings of the 6th International Symposium on Mathematical Morphology, Sydney, Australia, 2002, 1–16. https://smil.cmm.minesparis.psl.eu/publis/C-72.pdf (accessed March 29, 2022).

Norrdine, A., Kasmi, Z., Sieprath, A., & Blankenbach, J. (2013). Ranging by UWB Time-of-Flight Measurements. Proceedings of the 2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN), October 28–31, Montbéliard, France.

Olson, E. (2011). AprilTag: A robust and flexible visual fiducial system. 2011 IEEE International Conference on Robotics and Automation, 3400–3407. https://doi.org/10.1109/ICRA.2011.5979561

Real-Ehrlich, C. (2019). Echtzeit-Positionierung für Fußgänger innerhalb von Gebäuden auf Basis von Smartphone-Sensoren. Dissertation, RWTH Aachen University, Aachen, Germany.

Schall, G., Wagner, D., Reitmayr, G., Taichmann, E., Wieser, M., Schmalstieg, D., & Hofmann-Wellenhof, B. (2009). Global pose estimation using multi-sensor fusion for outdoor Augmented Reality. In: Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), October 19–22, 2009, Orlando, FL, USA, pp. 153–162. https://doi.org/10.1109/ISMAR.2009.5336489

Terzakis, G., & Lourakis, M. (2020). A consistently fast and globally optimal solution to the perspective-n-point problem. In European Conference on Computer Vision (pp. 478–494). Springer International Publishing.

Ubisense. (2024). Retrieved November 25, 2024, from https://ubisense.com/dimension4.

Unity. (2024). Retrieved November 25, 2024, from https://unity.com.

Unreal Engine. (2024). Retrieved November 25, 2024, from https://www.unrealengine.com.

Willemsen, T., Keller, F., & Sternberg, H. (2015). A topological approach with MEMS in smartphones based on a routing graph. In: Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE.

CONTACTS

Baris Özcan, M.Sc.,
Jan Grottke, Dipl.-Ing.,
Zhenyu Liu, M.Sc.,
Dr.-Ing. Christoph Blut,
Dr.-Ing. Ralf Becker,
Univ.-Prof. Dr.-Ing. Jörg Blankenbach
RWTH Aachen University,
Geodetic Institute and Chair for
Computing in Civil Engineering & GIS
Mies-van-der-Rohe-Str. 1, 52074 Aachen
GERMANY
Tel. +49 241 80 95300
Email: {oezcan, grottke, liu, blut, becker, blankenbach}@gia.rwth-aachen.de
Website: http://www.gia.rwth-aachen.de