

# 5G Edge Computing Enhanced Mobile Augmented Reality

Jacky Cao,  
supervised by Xiang Su.  
*University of Oulu*  
Oulu, Finland  
jacky.cao@oulu.fi

**Abstract**—Physical world enhancements through virtually drawn annotations on mobile devices are a core component of mobile augmented reality (MAR) experiences. However, resource intensive computer vision tasks typically limit the full deployment of MAR pipelines on-device due to a lack of required computation resources. As an alternative, a distributed scheme between client devices and edge servers enables offloading computation intensive tasks to powerful machines. This introduces challenges, such as maintaining low communication latency and high bandwidth channels between client and edge server. Current LTE and WiFi infrastructure do not fulfil such requirements. 5G networks are an ideal candidate to provide seamless MAR experiences. This research investigates the enhancement of MAR with both 5G and edge computing and aims to address two particular challenges: 1) the creation of a new 5G-based edge computing infrastructure to support reliable MAR, and 2) investigating which particular transport protocols allow for optimised quality of experience (QoE) for MAR applications within 5G infrastructure.

**Index Terms**—mobile augmented reality, edge computing, 5g

## I. INTRODUCTION AND MOTIVATION

Mobile augmented reality (MAR) combines physical world views with virtually drawn annotations on mobile devices. The core enabling technology for MAR applications are computationally intensive computer vision algorithms, e.g., feature extraction, objection detection, and object recognition. These algorithms are often not deployed on user devices, i.e., smartphones and head-mounted display devices, as they lack the on-board processing capabilities to successfully analyse images and render results within an acceptable latency. Smaller resolution images could be analysed on-device and require less processing time, however, this comes at the detriment of the object detection and recognition accuracy.

This trade-off between latency and accuracy can be partially solved with offloading the computation intensive image processing tasks to external servers, where substantially more processing resources can be used to analyse larger images at higher accuracy values. Conventional cloud computing can provide the required computational power, but required communication resources to cloud servers are high. Edge computing is an emerging paradigm, allowing access to computation resources in the proximity to data sources and users. Edge computing is therefore a more suitable approach for computation offloading in latency-constrained MAR [1], where computation is placed at network edges, e.g., co-located with cellular base stations. In addition to low and predictable

latency, edge computing introduces benefits, such as lower bandwidth costs, improved user battery life, privacy- and security-preserving services, and system scalability to meet demand [2].

Edge computing enables a significant reduction in the overall latency. As more applications and devices are connected together, current LTE and WiFi networks will reach their capacity and inevitably become saturated, leading to general performance degradation for all network users, including MAR applications. 5G provides promising solutions to enable: 1) reduced communication latency, 2) higher bandwidth channels, 3) more reliable networks, and 4) the ability to support a considerably larger amount of devices.

Applying 5G edge computing technologies in MAR greatly enhances the perceived user experiences through reduced overall latency, and would introduce additional enhancements to MAR. For example, situational awareness through the interaction of MAR applications with the large availability of IoT devices, this would allow for context- and situation-aware environments which would adaptively respond in real-time to provide users with the best possible experiences. For 5G edge infrastructure to be successfully taken advantage of in MAR, protocols and algorithms that exploit the full capability of 5G need to be developed and tested.

## II. OUR APPROACH

This PhD project will investigate two significant scientific challenges related to 5G, edge computing, and MAR. The combination of which will lead to an MAR system which is capable of providing the best possible QoE. Addressing each challenge independently is not a viable option and a holistic approach towards developing a successful system is required. The services used by MAR clients are supported by underlying infrastructure, i.e., edge nodes and a 5G mobile network.

The validation of the proof of concept infrastructure will be made on the 5G Test Network (5GTN). 5GTN is an open mobile network which is deployed at the University of Oulu, Finland and provides a carrier grade testing platform. Key network features include non-standalone 5G and LTE connectivity through respective base stations across the university campus. The network contains additional capabilities such as MEC servers and an IoT network, the former is of particular interest as they are employable as edge nodes. The



Fig. 1. Front-end user interface of the context-aware MAR air quality system.

IoT network is additionally important as context-awareness could be introduced through the inclusion of such data sources.

#### A. 5G Edge Computing Infrastructure in MAR

Reliability is a key concern for MAR systems, i.e., operation regardless of network issues or interruptions. Consequently, a 5G edge computing infrastructure will be investigated where intensive computations are re-located from centralised cloud servers to 5G network edges. A particular focus for this work will be the ability of the novel edge architectures to handle vast data quantities and provide high performance and quality to MAR experiences, especially with the most latency-sensitive operations, such as mapping, tracking, and rendering. This infrastructure will also enhance the MAR pipeline with context metadata, handle knowledge synchronisation between cloud and edge nodes, and the management of virtualisation technologies in edge nodes.

We further investigate how to enable context-awareness for MAR within this infrastructure through a context-aware head-mounted display MAR system which leverages context data and user profile context for holographic user navigation (Fig. 1). This system is split into a HoloLens and Android smartphone application because HoloLens 2 cannot provide the necessary GPS information and cellular data connection. We evaluate the prediction accuracy and latency of the system, and conclude that holographic rendering is one of the largest bottlenecks. Other latency sources, such as data transfer and model prediction time are minimised, for example, compressing and reducing the size of the data packets between smartphone and HoloLens, and converting machine learning models to be suitable for mobile devices.

#### B. Protocols for Optimised MAR within 5G Infrastructure

User MAR experiences should remain seamless while highly unstable networks are utilised to transfer data. Most research to date has focused on application-level approaches and algorithmic optimisation rather than maximising the link utilisation. To minimise the network impact on overall latency perceived by the user, novel transport protocols and methods to mitigate network instability will be studied to accommodate the needs of MAR applications on wireless networks.

We evaluate the suitability of several transport protocols (UDP, TCP, MPEG-TS, RTP, and QUIC) through quantifying quality of service (QoS) metrics, i.e., return-trip time (data latency between client and edge server), jitter (variation in packet delay from client to server), and throughput (data transferred per second from client to server) [3]. In addition to varying the transport protocol, we also change the image resolution to investigate the impact of data size, and the network connection type to compare the different available QoS. Cellular 5G and LTE, and WiFi connections are used during data collection. One key finding is that TCP is the most suitable transport protocol when using a 5G connection, owing to the protocol having the lowest RTT value for seven out of the nine tested image resolutions.

### III. CONCLUSIONS AND FUTURE WORK

For both good QoS and user QoE, MAR requires a low-latency pipeline which can be achieved using computation offloading. To achieve an acceptable offloading scheme, edge computing and 5G connectivity are necessary to enhance MAR. This work aims to: 1) create a new 5G and edge computing-based MAR infrastructure, and 2) to investigate which transport protocols provide the best QoE. We have investigated a context-aware MAR system which leverages context data for holographic user navigation and analysed the performance of current transport protocols on different network connections.

In future work, we will investigate the integration of more real-time personal context information, and perform extensive system level analytics to quantitatively define which areas of the context-aware system could be further optimised, and how scalability would impact QoS and QoE. We also plan to quantify transport protocol performance in different scenarios (i.e., varying signal quality, frequencies, and scalability), and investigate multipath offloading to mitigate network instability and guarantee acceptable QoE. Building upon these works, we will study privacy and security aspects for MAR, and the best mechanisms to integrate into our MAR system to guarantee that user data is protected, while also maintaining high QoE. This will ultimately lead to the development of the proposed final 5G edge computing infrastructure for MAR, where we will investigate system scalability, and adapt the system to accommodate several scenarios with high constraints and requirements such as for smart cities, vehicles, and industry. All previous work will prove an invaluable asset in the development, testing, and deployment of this system.

### REFERENCES

- [1] P. Mach and Z. Beevar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1628–1656, 2017.
- [2] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE internet of things journal*, vol. 3, no. 5, pp. 637–646, 2016.
- [3] J. Cao, X. Su, B. Finley, P. Zhou, and P. Hui, "Evaluating transport protocols on 5g for mobile augmented reality," *arXiv preprint arXiv:2006.02859*, 2020.