Experiment Scheme

Multi-user Augmented Reality Ranging

Zhang Yuyang

3rd July 2023

Catalogue

1 Introduction

In this section we introduce the background and motivation of the project, then we discuss the objectives that will be achieved.

1.1 Background

Multi-user Augmented Reality (MuAR) [30] is essential to implementing the vision effect of Metaverse for its capability to provide immersive and interactive experiences .

The boom in the capabilities and features of mobile devices, like smartphones [3] and wearable headsets [14], combined with the ubiquitous and affordable Internet access and the advances in the areas of cooperative networking and computer vision transformed MuAR from science fiction to a reality [5].

Furthermore, the arrival of Augmented Reality Software Development Kits (SDK), Apple's ARKit for iOS, Google's ARCore [8] for Android allow developers to easily create Augmented Reality applications, the interest of these platforms is mainly to considerably reduce the difficulty of access to MuAR for developers [20]. With the help of ARKit, it is convenient for developers to keep track of mobile devices with its surroundings, such as the visual features, inertial data and environmental changes [6].

In MuAR experiences, **spatial awareness** [25,30]are critical to understand each other' s intentions and actions so as to guarantee the smooth cooperation among users. Meanwhile, having a rational **context awareness** [6, 25] can play a significant role in improving the overall experience and immersion for users in MuAR scenarios.

1.2 Motivation

Indoor Positioning [17] is a crucial element in enhancing the user experience of MuAR spatial awareness, and it forms the central focus of this experiment. Indoor positioning holds a promising future and extensive practicality in the field of engineering. For example, it shows potential in indoor navigation [23, 28], fire evacuation [19], object tracking [31] and indoor mobile gaming [15].

Apple ARKit allows developers to track the smartphone's position in the real world in real time, which combines the Visual Inertial Odometry (VIO) with data from the camera and the motion sensor of the device [20] to calculate pose of mobile device. In the mean time, Ultra Wide-Band (UWB) is an emerging technology in the field of indoor positioning that has shown better performance compared to others [2].

Furthermore, the accuracy of indoor positioning by UWB and VIO technique is influenced by various factors, including context awareness elements like lighting, visual textures and object occlusion.

1.3 Objective

- This work aims to combine ultra-wideband (UWB) distance measurement with visual-inertial-odometry (VIO) pose estimation on iOS device in MuAR context. Given the changing environmental conditions, it is possible to assess the pros and cons of two distance measurement methods from an engineering perspective in experiment.

- This work aims to design an app which enables users to switch between UWB and VIO position measurement method in real time and provides a more user-friendly positioning experience for MuAR scenarios.

2 Literature Review

In this section we review existing work on Multiuser Augment Reality and Ultra-Wideband ranging, with their performance under different circumstances.

2.1 Related Work

Miller et al. [18] present Cappella, an infrastructure-free 6-degrees-of-freedom (6DOF) positioning system for multi-user AR applications that uses motion estimates and range measurements between users to establish an accurate relative coordinate system. Choi et al. [7] proposed a Lightweight UWB-VIO(LUVI) relative positioning method for indoor localization on android devices, whose work was fully implemented and tested in several indoor environments, showing robustness to NLOS while significantly reducing computational complexity, and up to 30% lower average error.

Wang et al. [30] proposed a Coordinate System Alignment(CSA) method on iOS devices in multi-user AR context, which shows the potential to reduce position errors by 58.3% on average under both short-term and long-term error duration.

Scargill et al. [25] proposed the IoT-based Sensing for AR, which comprehensively outline the challenges and opportunities associated with several important research directions of next-generation AR, utilizing its results to emphasize the importance of spatial awareness and context awareness to next-generation AR.

2.2 Problem Formulation

This experiment can be divided into two sections: spatial awareness and context awareness.

2.2.1 Spatial Awareness

This part mainly contains the mathematical aspects of indoor distance measurement methods of UWB and VIO-CSA technique.

- **Visual Inertial Odometry:** VIO performs well for measuring device displacement [7] because it combines camera and inertial measurement unit (IMU) for pose estimation [29]. Camera uses computer vision to extract feature points from an RGB image sequence [32] to solve SfM problem. IMU usually acquire data at a much higher rate than the camera for state propogation [22].

  In single-user AR context, ARKit provides pose of camera relative to the world coordinate system [3], suppose there exists a stationary world coordinate system C_W, when user creates a virtual anchor(VA), its position under world coordinate system should be fixed, labelled as $P_{VA} = (x_{va}, y_{va}, z_{va})$, the euclidean distance between user, whose position is $P_U = (x_u, y_u, z_u)$, and VA is an absolute distance:

  $$d = ||P_{VA} - P_U|| = \sqrt{(x_{va} - x_u)^2 + (y_{va} - y_u)^2 + (z_{va} - z_u)^2} \quad (1)$$

  In MuAR context, VIO also has the capability to retrieve the pose of each user's camera in their world coordinate system.

- **Coordinate System Alignment.** In order to calculate the relative distance between different users in MuAR scenario, we have to align the coordinate system of users. Thus, it is necessary to calculate the relative affine transformation [13] between users, whose form shown as equation 2:

  $$T_r = \begin{bmatrix} R_r & t_r \\ 0 & 1 \end{bmatrix} \quad (2)$$

  R_r refers to relative rotation factor between two users' world coordinate origin and t_r refers to relative translational factor between two users' world coordinate.

Suppose now we have two coordinate systems, C_{w1} and C_{w2}, which is fixed by ARKit. ARKit provide us with camera pose T_1 and T_2 under C_{w1} and C_{w2}.

ARKit provides us relative pose within its API between peers, which is T_r, then we could achieve coordinate system alignment:

$$T_2^1 = T_r * T_2 \tag{3}$$

in which T_2^1 refers to the pose of device 2 in the coordinate of C_{w1}. By retrieving translation factor t_1^2 and t_1 from T_2^1 and T_1, we get:

$$t_2^1 = \left[x_2^1, y_2^1, z_2^1\right]^T \tag{4}$$

and:

$$t_1 = [x_1, y_1, z_1]^T \tag{5}$$

Applying equation 4 and 6 into equation 1, we can get distance between users, which is:

$$d_{12} = \left|\left|t_2^1 - t_1\right|\right| \tag{6}$$

- **UWB Ranging.** Impulse radio ultra-wideband (IR-UWB) communication technology [7] uses a very short pulse signal, which has excellent distance resolution and signal to noise ratio (SNR) compared with other communication technologies [2]. Nearby Interaction API [4] provides explicit peer position output, which is composed of two components: relative distance d_{NI} and relative direction (x_{NI}, y_{NI}, z_{NI}) [30].

2.2.2 Context Awareness

This part analyzes the potential impact of environmental factors on the sensors of the mobile devices employed in this experiment.

- **UWB Ranging.**

  Impacted by the obstruction of objects. Although UWB has excellent performance in ideal environments, it still has relatively unstable performance when there is an obstacle between the sender and receiver due to multi-path problems, reflection, diffraction, and energy loss through the medium [7]. From Juopper et al. [12], nowadays advanced UWB chip could make its signal penetrates well through most building materials such as wood and brick, but not metal.

  U1 chip limitation. Apple's U1 chip is restricted by certain environment circumstances. First, the maximum distance between any two peers is 9 meters [4, 30]; second, the screens of peer devices should be kept in the portrait mode [4]; third, peer devices should appear within the line of sight of each other and there is no presence of obstacles between them [4, 30].

- **RGB Camera.**

  Motion Blur. The visual data obtained from cameras onboard AR devices frequently suffers from distortions such as noise or motion blur [25].

  No Depth Data. Modern AR devices are increasingly relying on time-of-flight depth sensorsto aid with spatial understanding [25]. However, most iPhone available in the market equips with a monocular camera that does not have a lidar sensor. As a result, their ability to perceive distances is significantly restricted.

  Environmental Factors. Vision-based approaches are often dependent on environmental factors (e.g., lighting conditions or image resolution) [25]. The amount of light incident on environment surfaces per unit area, determines the accuracy with which environment surfaces can be mapped or tracked [25] by the camera. The edge strength and complexity of a texture, as well as how a texture is impacted by motion blur, affect pose estimate error magnitude and distance measurement error in MuAR [24, 25].

- **Inertial Measurement Unit.**

  Gimbal Lock. Gyroscope in IMU presents a problem for the overall solution and it is represented by the gimbal lock [10]. The problem is that within a small interval (about 20 degrees) around the vertical plane (+/- 10 degrees), the axes are reversed once on the negative side and another time on the positive side [21] effect.

如果使用图进行说明，Latex 中插入图并进行索引的方式如下：图1.

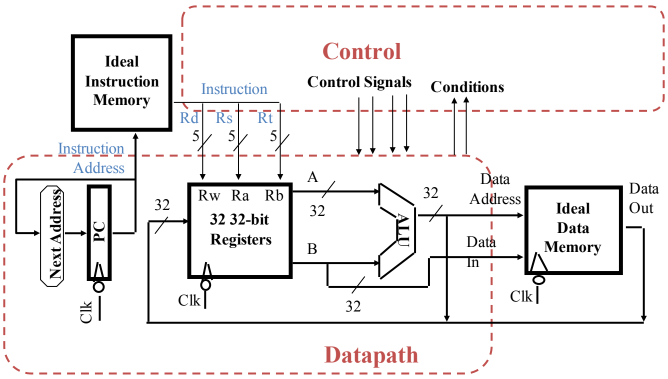

图 1: 单周期处理器结构图

3 Experiment Scheme

In this section, we mainly discuss the experimental method and process.

3.1 Experiment Method

The experiment utilizes the pose data gathered by VICON [1, 16], a motion camera capture system, as ground truth.

Based on the movement pattern of mobile devices, this experiment can be divided into two parts: static distance measurement and dynamic distance measurement.

3.1.1 Static Measurement

It is necessary to keep two iOS devices completely still during static measurement. This can be achieved by using tripod mounting technique.

The focal point of static measurement lies in the initialization process of the app, the change of surrounding environment settings and data post processing process, which will be further elucidated in the following parts 3.2.2 and 3.2.3.

3.1.2 Dynamic Measurement

The experiment shows great challenge in the form of dynamic measurement with the change of multiple variables like motion state, lighting and background texture because it's difficult to assess VIO and UWB distance measurement under these changing circumstances. Hence for this experiment, we employ the method of controlling variables. We strive to maintain the stability of all other environmental factors while manipulating a single condition. Through this approach, we can accurately assess the performance of two ranging methods.

- **Challenge 1: How to move.** Assume that the surroundings of the two mobile phones are similar. We aim to investigate the effect of device movement on VIO and UWB distance measurement method [27]. By drawing inspiration from this paper [27], we define five actions user may perform in multi-user AR context as shown below.
  - **Approaching** users approach each other in MuAR context.
  - **Normally** user moves around with smooth displacement and rotation.
  - **Walk away** users move that the distance between them or Field of View (FoV) exceeds the limitations of the U1 chip in MuAR context.
  - **Place Down** either of the users place down iOS device in MuAR context.
  - **Unstable movement** user is in an unstable state of motion, such as vigorous rotation, sudden acceleration or deceleration and in a high speed.

  In app, we analyze the numerical fluctuations of different inertial measurement components (like accelerator, gyroscope) in five scenarios using the Inertial Measurement Unit (IMU) [11] built-in to the iPhone.

- **Challenge 2: How to control environment condition.**

  This part discusses how to regulate the environmental conditions, for example, illumination and visual contexts effectively.

  Illumination To evaluate the impact of illumination on UWB and VIO techniques, we set four different lighting ranges shown as below.

  - **Neutral lighting** 900-1100 lumen. This condition can be achieved under normal indoor lighting.
  - **Over lighting** > 1400 lumen. If device is too close to the light source, for example, a desk lamp or flashlight.
  - **Dim lighting** 500-750 lumen, This condition can be achieved by turning off the indoor lights.
  - **Darkness** <100 lumen. When encountering such a scenario, the camera on the mobile device is obscured, and the sensor barely registers any incoming light.

  Textures. This part mainly tries to figure out the relationship between the performance of two distance measurement methods and the background of textures:

  We plan to utilize four different materials as the textures of MuAR experience, including a plain white wall, a paper, floor tiles and a check-board as 2 shows.

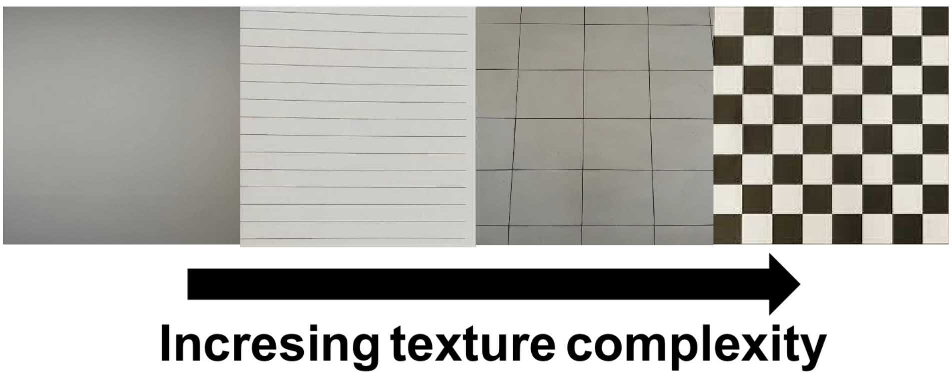

图 2: Textures

Under neutral lighting, the mean feature points number of these materials captured by ARKit are shown in table 1.

表 1: Certain Textures with Features' Number Captured by ARKit

Texture	**Avg number of features**
Wall	<10
Paper	10-20
Floor tiles	20-30
Chessboard	40-50
Wallpaper	>50

- **Challenge 3: How to detect environmental factors.**
  - **RGB Image.** At the lowest level, we can extract information from the overall distribution of pixel values about light level and the distinguishable elements of a scene (Img. Brightness, Img. Brightness RSD (Relative Standard Deviation), Contrast, Entropy) [26].
  - **Illumination.** AR platform itself (e.g., ARCore, ARKit) may provide useful data on light estimation and detected planes [3, 8, 26].
  - **Point Cloud Properties.** we can extract information fro the 3D point cloud generated by the AR platform while mapping the scene. From this we calculate Feature Point Count, Feature Point Density, Mean Depth, Feature Point Proximity, and how evenly distributed the points are across the mapped space (Spatial Heterogeneity) [26].

Moreover, further discussion in terms of virtual object position error (a function of pose estimate error magnitude) is valuable which considers both visual textures and illumination [25].

In conclusion, dynamic measurement serves to assess robustness of two distance measurement methods in a dynamic state, enabling the summarization of information, which is crucial basis for ranging selection algorithms.

3.1.3 Experiment Supplement

This context-awareness is key to realizing the full potential of AR; to deliver virtual content that provides a high-quality user experience, and enables optimal task performance, we require information on the specific circumstances in which it is presented [9].

- **Light** Illuminance, the amount of light incident on environment surfaces per unit area, determines the accuracy with which environment surfaces can be mapped or tracked, because it determines the extent to which visual features are detectable for tracking [25]. The following table 2 presents the specific outcomes of the illumination detection feature of ARKit, categorized into three distinct lighting scenarios.

表 2: ARKit Light Detection Module (Testing on iPhone 13)

Lighting	**Effect on UWB**	**Effect on VIO**
Neutral Lighting	nil	nil
Insufficient Lighting	nil	nil
Over Lighting	nil	nil

- **Texture** Both the edge strength and complexity of a texture will affect ranging magnitude and motion blur [26].

表 3: Possible Texture in Experiment

Texture	**Effect for ranging(VIO)**	**Effect for ranging(UWB)**
White wall	Nil	Nil
Wallpaper	Nil	Nil
Brick	nil	nil
Chessboard	nil	nil

- **Motion**

  During testing, we illustrate the varying impacts of different motion modes on VIO ranging, UWB ranging, and AR object drift in indoor testing environments, as shown in Table 4.

表 4: Possible Movements in Experiment

Movement	**Effect for drift**	**Effect for ranging(VIO)**	**Effect for ranging(UWB)**
Approaching	Nil	Nil	Nil
Walk away	Nil	Nil	Nil
Place down	nil	nil	nil
Unstable movement	nil	nil	nil

3.2 Experiment Process

This section mainly elucidates the intricate details pertaining to the experiment.

3.2.1 Preparation

- **App Design.** The application is created using the Xcode software and developed with the Swift programming language. The design of the app primarily involves two aspects: human-computer interaction and software programming. However, since it is not within the main scope of this experiment, it will be skipped. For a detailed view of the actual user interface, please refer to the image below.

- **Evaluation Model.** We aim to evaluate measurement spread based on root mean square error (RMSE) for the most recently measured window with N samples of ranging records. Therefore,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (d - d_i)^2}{N}} \tag{7}$$

  where i is the index of window and d is the ground truth of relative distance.

  We normalize RMSE (NRMSE) as a metric for NLOS identification using the difference between maximum and minimum ranging values in the window,

$$NRMSE = \frac{RMSE}{d_{max} - d_{min}} \tag{8}$$

  The software utilized for writing the code to evaluate the experimental data is Visual Studio Code, with Python serving as the programming language.

- **Experiment Tools.**
  - **Necessity.** Two iPhone with version higher than iOS 11.0 [3, 4], labelled as device A and B; Two Tripods; Few Thin wooden pieces or thin fabric materials; Some Wallpapers (wood, brick, stone, chessboard)
  - **Optionals.** Flashlights; Occlusion materials (piece of metal)

3.2.2 Experiment Setting

This section outlines the anticipated steps to be taken during the experimental process, including potential issues that may arise and the corresponding solutions.

- **App Initialization.**

  Step 1: At the beginning of the experiment, the app will commence initialization, creating a multi-user AR Collaboration session and matching the devices accordingly.

  Step 2: Users are required to constantly move their phones around to scan the surrounding environment in order to match the AR world map [3].

- **Static Measurement.**

  Assumption 1: Static measurements must be taken only after the app has been completely initialized.

  Assumption 2: The relative pose between the two devices have been adjusted within the limit of U1 chip.

  Step 1: Attaching both mobile devices cautiously to the tripod.

  Step 2: Pressing the data collection button in the app to gather environmental and ranging data, utilizing VICON [1] to catch ground truth.

  Step 3: After a certain period of time has elapsed, proceed to press the button again in order to end the distance measurement experiment temporarily. Make adjustments to the position of mobile devices and repeat the aforementioned steps multiple times to ensure the reliability of the gathered data.

- **Motion Control.**

  Assumption 1: Motion measurements must be taken only after the app has been completely initialized.

  Step 1: Pressing the data collection button in the app to gather data.

  Step 2: Users move around in certain motion states, which could be detected in app.

  Step 3: After a certain period of time, users press button once more to end the collection.

- **Lighting Control.**

  Assumption 1: Lighting measurements must be taken only after the app has been completely initialized.

  Assumption 2: Suppose devices are in the state of static measurement.

  Step 1: Turn on the indoor lights to achieve neutral light conditions. Pressing the data collection button in the app to gather data.

  Step 2: Then conduct the experiment repetitively under different lighting conditions.

- **Visual Textures Control.**

  Assumption 1: Texture measurements must be taken only after the app has been completely initialized.

  Assumption 2: Suppose devices are in the state of static measurement.

  Step 1: Change the background of experiment scene with different wallpapers to achieve different texture requirements.

  Step 2: Then conduct the experiment repetitively by collecting data in the app in various scenarios.

3.2.3 Data Collection and Post-Processing

Assuming that the coordinate system for VICON is denoted as C_W. Static measurement simply requires two iOS devices to remain stationary while our VICON system captures their 6 degree-of-freedom(6 DOF) poses in C_w, labeled as T_{WA} and T_{WB}.

The two iOS devices possess their own individual world coordinate systems, respectively labeled as C_{WA} and C_{WB}. Device A collects the location of Device B as P_{AB} in coordinate system C_{WA}. Conversely, Device B collects the location of Device A as P_{BA} in coordinate system Cwb.

In general, we can assess the UWB and VIO peer ranging data in MuAR context for the following scenarios.

- **Static Measurement.**
  - different lighting
  - different textures
  - occlusion or not
- **Dynamic Measurement.**
  - different lighitng
  - different textures
  - different motions

The data collected by the app is saved in either txt or csv format. Each line of data includes a timestamp and other information, such as distance, acceleration, angular velocity, and light intensity.

The post-processing phase contains three distinct parts: the **preprocessing** of experimental data, the **processing** of experimental data, and the **visualization** of the processed experimental data.

Preprocessing involves the categorization of experimental data and the utilization of ground truth to validate its reliability.

During processing stage, the average error of the experimental data is calculated using mentioned evaluation model. For instance, after conducting experiments on static measurement of ambient light changes, separate evaluation functions were used to calculate the RMSE of the distance data sets measured by UWB under different lighting conditions and VIO under different lighting conditions. The results were obtained to evaluate the reliability of the two ranging methods under different lighting conditions, which would be used for subsequent algorithm design selection.

Visualization provides a more straight-forward way to show the relative error than evaluated data itself.

参考文献

[1] Shaima Al Habsi, Maha Shehada, Marwah Abdoon, Ahmed Mashood, and Hassan Noura. Integration of a vicon camera system for indoor flight of a parrot ar drone. In *2015 10th International Symposium on Mechatronics and its Applications (ISMA)*, pages 1–6. IEEE, 2015.

[2] Abdulrahman Alarifi, AbdulMalik Al-Salman, Mansour Alsaleh, Ahmad Alnafessah, Suheer Al-Hadhrami, Mai A Al-Ammar, and Hend S Al-Khalifa. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors*, 16(5):707, 2016.

[3] Apple. ARKit. https://developer.apple.com/documentation/arkit, 2022.

[4] Apple. Nearby interaction. https://developer.apple.com/documentation/nearbyinteraction, 2022.

[5] Dimitris Chatzopoulos, Carlos Bermejo, Zhanpeng Huang, and Pan Hui. Mobile augmented reality survey: From where we are to where we go. *IEEE Access*, 5:6917–6950, 2017.

[6] Yunqiang Chen, Qing Wang, Hong Chen, Xiaoyu Song, Hui Tang, and Mengxiao Tian. An overview of augmented reality technology. In *Journal of Physics: Conference Series*, volume 1237, page 022082. IOP Publishing, 2019.

[7] Hong-Beom Choi, Keun-Woo Lim, and Young-Bae Ko. Luvi: Lightweight uwb-vio based relative positioning for ar-iot applications. *Ad Hoc Networks*, 145:103132, 2023.

[8] Google. ARCore. https://developers.google.com/ar/reference, 2022.

[9] Jens Grubert, Tobias Langlotz, Stefanie Zollmann, and Holger Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1706–1724, 2017.

[10] Evan G Hemingway and Oliver M O' Reilly. Perspectives on euler angle singularities, gimbal lock, and the orthogonality of applied forces and applied moments. *Multibody System Dynamics*, 44:31–56, 2018.

[11] Jianzhu Huai, Yujia Zhang, and Alper Yilmaz. The mobile ar sensor logger for android and ios devices. In *2019 IEEE SENSORS*, pages 1–4, 2019.

[12] Niko Juopperi. Uwb implementation and utilization in mpos device. 2022.

[13] Valtteri Kaatrasalo. Computer vision methods for augmented reality. 2022.

[14] S Karthika, P Praveena, and M GokilaMani. Hololens. *International Journal of Computer Science and Mobile Computing*, 6(2):41–50, 2017.

[15] Wenbin Luo et al. Building an ar treasure hunt game using indoor localization technology. 2018.

[16] Ahmed Mashood, Ahmed Dirir, Mousa Hussein, Hassan Noura, and Falah Awwad. Quadrotor object tracking using real-time motion sensing. In *2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, pages 1–4. IEEE, 2016.

[17] Rainer Mautz. Indoor positioning technologies. 2012.

[18] John Miller, Elahe Soltanaghai, Raewyn Duvall, Jeff Chen, Vikram Bhat, Nuno Pereira, and Anthony Rowe. Cappella: Establishing multi-user augmented reality sessions using inertial estimates and peer-to-peer ranging. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 428–440. IEEE, 2022.

[19] Gi Ho Nam, Han Sol Seo, Min Su Kim, Yun Kyoung Gwon, Cheol Min Lee, and Dong Myung Lee. Ar-based evacuation route guidance system in indoor fire environment. In *2019 25th Asia-Pacific Conference on Communications (APCC)*, pages 316–319, 2019.

[20] Zainab Oufqir, Abdellatif El Abderrahmani, and Khalid Satori. Arkit and arcore in serve to augmented reality. In *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, pages 1–7. IEEE, 2020.

[21] Alexandru Predescu, Mariana Mocanu, and Ciprian Lupu. Armax: A mobile geospatial augmented reality platform for serious gaming.

In *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, pages 313–318, 2019.

[22] Tong Qin, Peiliang Li, and Shaojie Shen. Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4):1004–1020, 2018.

[23] M S Ramesh, J Naveena Ramesh Vardhini, S Murugan, and J Albert Mayan. Indoor navigation using augmented reality for mobile application. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 1049–1052, 2023.

[24] Tim Scargill, Ying Chen, Nathan Marzen, and Maria Gorlatova. Integrated design of augmented reality spaces using virtual environments. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 297–306. IEEE, 2022.

[25] Tim Scargill, Sangjun Eom, Ying Chen, and Maria Gorlatova. Ambient intelligence for next-generation ar. *arXiv preprint arXiv:2303.12968*, 2023.

[26] Tim Scargill, Shreya Hurli, Jiasi Chen, and Maria Gorlatova. Will it move? indoor scene characterization for hologram stability in mobile ar. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, pages 174–176, 2021.

[27] Tim Scargill, Gopika Premsankar, Jiasi Chen, and Maria Gorlatova. Here to stay: A quantitative comparison of virtual object stability in markerless mobile ar. In *2022 2nd International Workshop on Cyber-Physical-Human System Design and Implementation (CPHS)*, pages 24–29. IEEE, 2022.

[28] Janis Tiemann, Andrew Ramsey, and Christian Wietfeld. Enhanced uav indoor navigation through slam-augmented uwb localization. In *2018 IEEE international conference on communications workshops (ICC workshops)*, pages 1–6. IEEE, 2018.

[29] Nguyen Hoang Khoi Tran and Vinh-Hao Nguyen. An ekf-based fusion of visual-inertial odometry and gps for global robot pose estimation. In *2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, volume 6, pages 1–5, 2021.

[30] Na Wang, Haoliang Wang, Stefano Petrangeli, Viswanathan Swaminathan, Fei Li, and Songqing Chen. Towards accurate positioning in multiuser augmented reality on mobile devices. In *2022 IEEE International Symposium on Multimedia (ISM)*, pages 214–217. IEEE, 2022.

[31] Ziming Wu, Jiabin Guo, Shuangli Zhang, Chen Zhao, and Xiaojuan Ma. An ar benchmark system for indoor planar object tracking. In *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pages 302–307, 2019.

[32] Harry Zhang, Jeffrey Ichnowski, Yahav Avigal, Joseph Gonzales, Ion Stoica, and Ken Goldberg. Dex-net ar: Distributed deep grasp planning using a commodity cellphone and augmented reality app. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 552–558, 2020.