PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Augmented reality-based vision-aid indoor navigation system in GPS denied environment

Rajeev, Srijith, Wan, Qianwen, Yau, Kenny , Panetta, Karen, Agaian, Sos

Srijith Rajeev, Qianwen Wan, Kenny Yau, Karen Panetta, Sos Agaian, "Augmented reality-based vision-aid indoor navigation system in GPS denied environment," Proc. SPIE 10993, Mobile Multimedia/Image Processing, Security, and Applications 2019, 109930P (13 May 2019); doi: 10.1117/12.2519224

Event: SPIE Defense + Commercial Sensing, 2019, Baltimore, Maryland, United States

Augmented reality-based vision-aid indoor navigation system in GPS denied environment

Srijith Rajeev[a], Qianwen Wan[a], Kenny Yau[a], Karen Panetta[a], Sos Agaian[b]
[a]Dept of Electrical and Computer Engineering, Tufts University, Medford, MA, USA 02155;
[b]Dept of Computer Science, The City University of New York, New York City, NY, USA 10314;

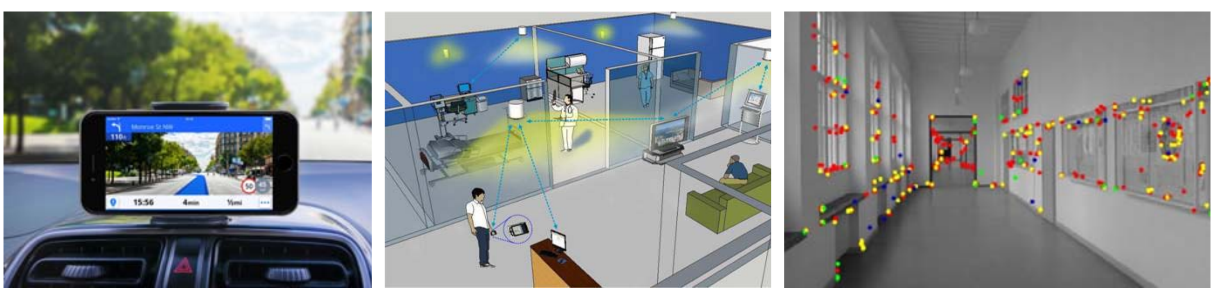
ABSTRACT

High accuracy localization and user positioning tracking is critical in improving the quality of augmented reality environments. The biggest challenge facing developers is localizing the user based on visible surroundings. Current solutions rely on the Global Positioning System (GPS) for tracking and orientation. However, GPS receivers have an accuracy of about 10 to 30 meters, which is not accurate enough for augmented reality, which needs precision measured in millimeters or smaller. This paper describes the development and demonstration of a head-worn augmented reality (AR) based vision-aid indoor navigation system, which localizes the user without relying on a GPS signal. Commercially available augmented reality head-set allows individuals to capture the field of vision using the front-facing camera in a real-time manner. Utilizing captured image features as navigation-related landmarks allow localizing the user in the absence of a GPS signal. The proposed method involves three steps: a detailed front-scene camera data is collected and generated for landmark recognition; detecting and locating an individual's current position using feature matching, and display arrows to indicate areas that require more data collects if needed. Computer simulations indicate that the proposed augmented reality-based vision-aid indoor navigation system can provide precise simultaneous localization and mapping in a GPS-denied environment.

Keywords: Augmented-reality, navigation, GPS, HoloLens, vision, positioning system, localization

1. INTRODUCTION

Localization can be defined as the process of estimating an individual's position relative to a reference frame [1]. The localization system performance is evaluated based on its accuracy. It is defined as the degree of conformance of an estimated or measured information at a given time to a defined reference value which is ideally the true value. The most common form of localization and navigation is to combine sensor readings and some form of closed-loop motion feedback to use a map [2]. Accurate outdoor and indoor location determination is an essential part of empowering various context-aware services and protocols. Figure 1 shows visual illustrations of common positioning systems.



GPS [3] | Wi-Fi, RFID, Bluetooth based positioning system [4] | Visual positioning system[4]

Figure 1: shows an illustrative example of the most widely used positioning systems.

The Global Positioning System (GPS) is a satellite-based navigation system developed by the U.S. Department of Defense (DoD) [5]. Although Global Position System (GPS) is one of the most widely used outdoor localization systems, its capabilities are limited inside an indoor environment. This is because GPS signals necessitate line-of-sight transmission

Mobile Multimedia/Image Processing, Security, and Applications 2019, edited by Sos S. Agaian, Vijayan K. Asari, Stephen P. DelMarco, Proc. of SPIE Vol. 10993, 109930P · © 2019 SPIE
CCC code: 0277-786X/19/$18 · doi: 10.1117/12.2519224

between the transmitter (satellites) and receivers. However, these signals generally lose signal intensity as they bounce off buildings, penetrate walls and other enclosures [5]. These effects degrade other known solutions for indoor localization which uses electromagnetic waves (EM). Other indoor positioning systems include Wi-Fi positioning system, Bluetooth, RFID tags, and visual positioning system.

Wi-Fi positioning system (WPS) is used for localization with wireless access points. WPS can be used to complement or substitute GPS systems. It can also be used along with GPS data to track users [6, 7]. Typical parameters useful to geolocate the Wi-Fi hotspot or wireless access point include the Media Access Control (MAC) address and the Service Set Identifier (SSID) of the access point. The possible signal fluctuations that may occur can increase errors and inaccuracies in the path of the user. Moreover, WPS is limited by Wi-Fi connectivity, and it does not work when out of range of Wi-Fi signals.

Bluetooth operates in the 2.4-GHz ISM band. Bluetooth is a “lighter” standard, highly ubiquitous, and supports several other networking services in addition to IP. Bluetooth positioning systems thrive on providing proximity details to the user, rather than exact location [8] as it was not designed to offer an exact location like GPS. Therefore, Bluetooth is used as a geo-fence or a micro-fence solution for indoor proximity solution.

RFID (radio-frequency identification) is a means of storing and retrieving information through the electromagnetic transmission to an radio-frequency compatible integrated circuit. However, its range is very limited. RFID is an emerging positioning technology that allows for mobility tracking of objects or people. It uses radio waves to wirelessly transmit identity (e.g., person identity or serial number) and other characteristics of an object. RFID is not suitable for exhaustive localization as it offers a limited range of less than a meter [9]. Moreover, the operability of such approaches requires some narrow passage to prevent from passing by out of range.

A visual positioning system used visual markers to determine the location of a camera-enabled mobile device. In such a system, markers or image targets are placed at specific locations. The system measures the visual angle from the device to the marker that enables it to estimate its location coordinates with respect to the marker. Table 1 summarizes the accuracy, coverage, cost of common positioning systems.

Table 1. Comparisons between popular navigation systems [10, 11]

Property	**GPS**	**Wi-Fi**	**RFID**	**Bluetooth**	**Vision**	**Beacons**
Accuracy	2-20 (m)	5-15 (m)	Passive: 4m Active: 100m	1-5 (m)	1/10-1 (m)	1-8 (m)
Coverage	Global	Depends on number of units	1/10 -100 (m)	1-30 (m)	1-10 (m)	30-75 (m)
Cost	High	Medium/High	High	Low	Medium/High	Medium
Environment	Outdoor	Indoor	Indoor	Indoor	Indoor	Indoor

Augmented reality (AR) is an interactive experience of a real-world environment where objects that exist in in the real-world are augmented or placed by computer-generated perceptual information. In AR, a plethora of ideas and concepts have been introduced that improve the depiction of virtual objects in a real scene. A significant advantage of AR systems is their depiction of digital information, where the user perceives virtual and real objects as coexisting in the same natural environment [12]. Moreover, AR system bridges the distance between computer vision, natural vision, and machine learning. Text recognition algorithms can be deployed on AR systems that can automatically detect text and convert them into digital information. At a glance, the user naturally recognizes the content of the information without needing to understand abstract metaphors that cumbersomely paraphrase the same information contained in conventional textual, graphical or even virtual reality systems [12]. Recently, augmented reality and computer vision-based navigation systems were introduced that have changed the scope of navigation[13]. However, this system is still applicable to augmented reality capable phone devices only. Moreover, another major concern in the indoor navigation community is the localization of the user/robot in a complex environment.

This article proposes an augmented reality-based indoor navigation system that can navigate to locations with and without GPS coordinates. The proposed augmented reality-based navigation system will use computer vision to identify and localize a user and Artificial Intelligence (AI) for path planning and AR for path visualization. The proposed system is designed to be non-invasive. It localizes the user's position based on the presence of a few distinctive image targets [14]. Therefore, it does not need expensive localization equipments like beacons or RFID tags. It does not depend on GPS, but the system can be extended to be used along with GPS coordinates. Furthermore, this system does not require QR codes for localizing its position. The rest of this paper is organized as follows: descriptions of related systems are presented in Section 2, the proposed system is described in Section 3, and computer simulations are provided in Section 4. Finally, conclusions and future works are drawn in Section 5.

2. RELATED WORK AND BACKGROUND

Vision-based navigation system has been a dominant approach in fields such as robotic navigation systems [19], autonomous driving products [15], UAVs (unmanned aerial vehicle)applications [16], capsule endoscopy (CE) procedures [31], medical applications [26], precision agriculture [32] and universal accessibility [25].

Since camera calibration process is becoming easier and the cost of imaging capture devices is dropping dramatically, the development of vision-based navigation methods has been an emerging topic in real-life human-machine interaction consumer products [24]. A wearable vision-based navigation system was prototyped in [18], which assists local navigation for blind and visually impaired people by providing online feedback on obstacles and useful objects. Localizing UAVs in GPS-denied conditions, such as indoor environments, through vision-based systems is explored and studied in [33]. A comprehensive survey on vision-based topological navigation system involves mapping and localization methods, image classification (global, local, bag-of-visual-words and combinations) techniques are investigated in [34].

Table 2. Vision-based Navigation systems, applications, and methods

Vision-based Navigation systems	Category	Related work
Commonly deployed devices	• UAVs • Mobile • Wearable device • Robots	[15, 16] [17] [18] [19]
Fields of application	• Cognitive psychology • Robot • Autonomous Driving • UAVs • Human-machine interaction • Universal Accessibility • Medical and clinical applications	[20] [21] [22] [23] [24] [25] [26]
Popular methodologies	• Visual Localization and Mapping, SLAM • Image Classification and • Object Detection and Tracking • 3D reconstruction	[27-30]

While Augmented Reality (AR) technology can create and blend virtual digital elements seamlessly into the real world, the virtual information can be displayed and used to assist navigation systems, especially in GPS denied or precise indoor situations [13]. AR technology platform can aid in supporting people with a more natural navigation experience [35]. A pilot work was found in the literature: And indoor positioning navigation system on mobile devices is actively researched in [17], which augmented the location-based virtual information overlapping on the real world video stream. Combination of AR platform and vision-based navigation system has demonstrated its advantages in real-life applications, such as getting route instructions in a shopping mall or an airport a person never go to. Considering an AR headset can provide human a truly immersive experience, this work introduces a vision-based AR navigation system. Table 2 presents the fields of applications and commonly used methods.

3. PROPOSED METHOD

The proposed Augmented Reality based Vision-Aid Indoor Navigation System localizes the user's position based on image targets, maps the features locally on a 3D map, and provides visual navigation aid on an augmented reality device. A user will utilize the components of the indoor navigation system in the order listed. First, the application is opened. Second, the user looks around the environment for an image target [14]. Then, the application localizes the user's position in a virtual 3D map. From here, the user can activate the Destination Input system using a voice command, in this case, "Set Destination." Once the destination has been specified, the Navigation System will perform the pathfinding and display a visual aid, guiding the user to the said destination. This project is deployed on Microsoft HoloLens [36]. Figure 2 shows the operation flow of the proposed system, and the following sections provide a more comprehensive description of this process.

2.1 User localization system

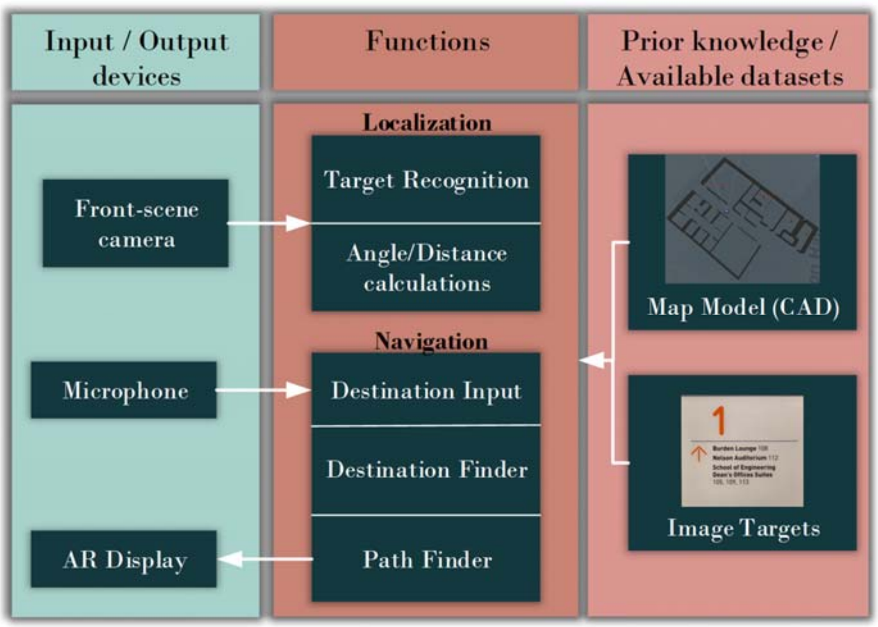


Figure 2: Block diagram of the proposed indoor navigation system.

Localization and mapping are the computational problems of updating a map of an known or unknown environment while simultaneously keeping track of an agent's location within it. In this system, image targets are used as markers to identify the location and orientation of the user in a 3D space. An image target is a visual landmark which the front-scene camera looks for to determine the user's position. Good image targets are rich in detail – busy patterns and/or images, have good contrast – edges in the image correspond to sharp color transitions, do not have repetitive patterns – corresponds to a unique image target., and contain many natural features in the image – these are sharp, spiked, chiseled details in the image. Additionally, the system should include only those image targets that will not change over time. Furthermore, the user's position in the model is updated automatically using the IMU device to keep world-space virtual objects in the same real-world position.

Every scene has a set of world axes and every object has its own set of axes. Positions are considered with refernce to the world axes. Rotations of an object are calculated by the rotational offset of the object's axes with the world axes, i.e., it is not calculated by the angle that a ray starting from the origin and passing through the position of the object makes with the world axes. An object's position and rotation information are contained in a transform. Thus, the word 'transform' is used to refer to an object's position and rotation information collectively.

Upon Target Recognition, the transform of the image target compared to the camera's transform is calculated. The location of the image target within the map is known and a permanent feature. A direct translation of the position offsets between the map and the camera cannot be performed because these entities have different coordinate systems. One set of information is taken with respect to the camera's axes (imagine three axes that stay fixed to the sagittal, coronal, and transverse planes of the head, no matter in which direction the user was looking). Another set of information is taken with respect to the map's axes. Therefore, it is necessary to perform a transformation on the data to align their reference axes.

On a high level: The goal of localization is to figure out how the camera is oriented relative to the map. In actuality, it is the process of finding how the map is oriented relative to the camera since it is not possible to shift the physical camera. Using the knowledge of the image target's orientation with respect to the map and that of the image target's orientation with respect to the camera, it is possible to determine how the map is oriented with reference to the camera. When the application loads, the camera and map Transforms are the same. Importantly, the image target's transform stays fixed relative to the map's transform because the image target is a child of the map – so any transformations applied to the map also apply to the image target.

To localize, first, translate the map by the image target-camera position offset, then rotate the map around its center by the image target-camera rotation offset. This effectively places the center of the map at the same transform relative to the camera as the image target. Next, determine the rotation offset between the image target and the map. The last step is to apply a negative translation of the image target-map position offset along the map's axes. More specifically: consider the reverse problem of localization, which is finding the coordinates and rotation of the map. Assuming apostrophes (or lack thereof) to indicate coordinates and rotations taken with respect to (w.r.t.) different sets of axes, i.e., x is the coordinates w.r.t. the world space axes, x' is coordinates w.r.t the map axes and let x'' be the coordinates and rotations w.r.t. the camera's axes. When the application is initialized, both the camera view and the map are set to (0,0) with orientation 0^o. This means that their axes are aligned with the world space axes. Since the camera is not being translated and rotated, the world space axes are the same as the camera. Thus, any measurements taken w.r.t the world space axes will also be the same if taken w.r.t. the camera's axes.

Assume that there is rotational invariance between different axes, i.e. $\Delta\theta^{tc} = \theta^t - \theta^c = \theta^{t'} - \theta^{c'} = \Delta\theta^{(tc)'}$. Image target (trackable) has map-relative coordinates $(x,y)^{t'}$ and rotation $\theta^{t'}$. Moreover, the map's map-relative coordinates $(x,y)^{m'}$ and rotation $\theta^{m'}$ are always (0,0) and 0^o. These two values stay locked because the trackable is a part of the map. The same trackable has camera-relative coordinates $(x,y)^{t''} = (x,y)^t$ with rotation $\theta^{t''} = \theta^t$. Finally, let the camera's (user view) position be $(x,y)^{c''} = (x,y)^c = (0,0)$ and $\theta^{c''} = \theta^c = 0^o$. This enables $(x,y)^{t''} = (x,y)^t$ and rotation $\theta^{t''} = \theta^t$. Finally, Algorithm 1 is used to shift the map by $(x,y)^{m''}$ and $\theta^{m''}$ → $(x,y)^m$ and θ^m in order to localize the user.

Algorithm I: User localization

Input: $(x,y,\theta)^t$, $(x,y,\theta)^{t'}$, $(x,y,\theta)^c = (0,0,0^o)$, $(x,y,\theta)^m = (0,0,0^o)$

Equivalencies: Expressions that are based on the input

Trackable-Camera Position Offset: $\Delta(x,y)^{tc} = (x,y)^t - (x,y)^c = (x,y)^t$ (given)

Trackable-Camera Rotation Offset: $\Delta\theta^{tc} = \theta^t - \theta^c = \theta^t$ (given)

Trackable-Map Rotation Offset: $\Delta\theta^{(tm)'} = \theta^{t'} - \theta^{m'} = \theta^{t'}$ (given)

Map-Camera Rotation Offset: $\Delta\theta^{mc} = \boldsymbol{\theta^m} = \Delta\theta^{tc} - \Delta\theta^{tm} = \Delta\theta^{tc} - \Delta\theta^{(tm)'} = \theta^t - \theta^{t'}$ (can calculate θ^m)

Algorithm: Transformations that can be applied to the map to get it to $(x,y)^m$ and θ^m

Map transforms starts at position (0,0) and rotation 0^o

Translate Map by $(x,y)^t$.

Rotate the Map around by θ^m.

Translate Map *along its own axes* by $-(x,y)^{t'}$

Output: Camera or (user view) is localized in the map

2.2 Map model and features

A navigation mesh is a bundle of two-dimensional (2D) convex polygons[37]. This mesh defines the area of an environment that is traversable by the user. Furthermore, navigational mesh provides a reference system for an area's image targets and features. Successful localization and navigation of the proposed indoor navigation system are highly dependent on an accurate map model and placement of features.

Model creation: A 3D model of the building is used to relate distances between points of interests in the building and represent walkable surfaces. As pathfinding is the goal of this application, high-resolution detail of the structure is not needed; a model that demonstrates the viable traversable paths is good enough. This would include floors to indicate walkable surfaces and walls to indicate barriers. In this paper, two methods are used to create the model[38]. The first is using geometric Game object prefabs in Unity to hand-create the walls and floors of a building. This method is feasible as long as one knew the key dimensions of the building (room widths, hallway lengths, etc. However, this process was time-consuming and can be erroneous. The second method and the preferred method is to create building features that correspond to a given map[39]. The floorplan of the building is first obtained. Next, it is geo-referenced using the open-source software QGIS[40].

Image target placement: One can create image targets and place them in the correct location manually. With many image targets, this may become infeasible, but it is possible to automate the creation and placement of image targets as long as there is a list of image targets and their locations. This list may be created by creating point features and then giving each point feature the name of the image target and its location.

2.3 Destination Input (Voice and lookup table based)

The user specifies the destination using a voice command. A prior list of locations and coordinates are required to add this feature into the system. This system uses voice input to set the destination, as voice is a natural way to communicate information and typing on a headset is very clunky. The voice event handler is triggered when it hears the phrase 'Set Destination.' Next, the system will listen and convert the user's speech to text. Then, the text is matched with an online dictionary of locations. This process can also be done in the proposed system using a lookup table.

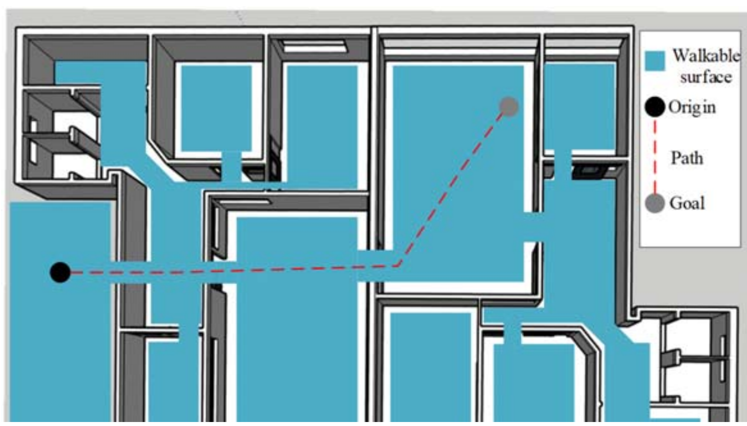


Figure 3: Shows a visual illustration of the navigation scenario in an indoor environment.

2.4 Navigation system

Consider the scenario shown in figure 3. Once the destination string has been obtained, the application searches the list of destinations, finds the correct one, and retrieves the coordinates of that destination [41]. The initial and goal locations have to be mapped to their nearest polygons to determine the path between them. Next, the search algorithm starts searching (visiting) all the neighboring polygons from the starting point, until it reaches the destination. This system uses the A* (A star) algorithm for navigation. A* uses a combination of searching based on the shortest path and heuristic searching [42]. In this algorithm, each cell or polygon is evaluated by the value:

$$f = g + h \quad (1)$$

Where h is the heuristic distance (Manhattan, Euclidean or Chebyshev) of the polygon to the goal state and g is the length of the path from the initial state to the goal state through the selected sequence of polygons. Using the path co-ordinates generated, the path is visualized on the augmented reality headset to provide navigational aid for the user.

4. COMPUTER SIMULATION

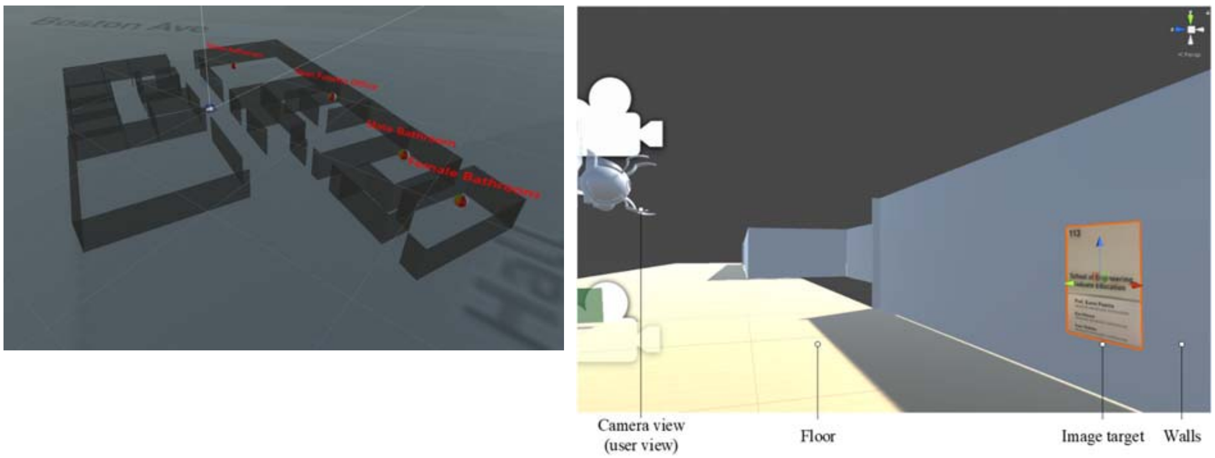
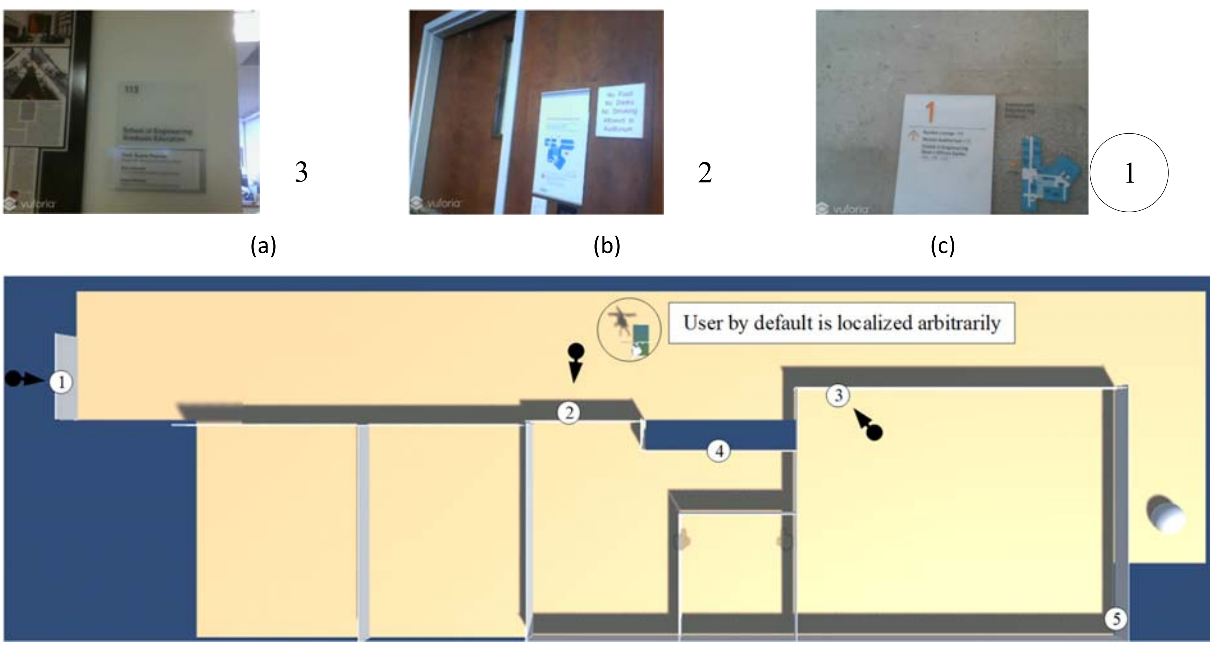


Figure 4: shows different views of the generated georeferenced 3D indoor map of a portion of Anderson Hall



(d)

Figure 5: (a) shows a section of the Anderson building with 5 image targets; [b-d] show the images of the targets that were used for localization in the computer simulation

4.1 Generate indoor 3D models and identify walkable areas

To test the proposed visual indoor navigation system, the floor plan of Anderson Hall, Tufts University, Medford was first obtained. Its features were extracted and georeferenced on a map (shown in figure 4). Importantly, the proposed indoor visual positioning system does not require virtual maps of the environment,. Furthermore, image targets were carefully selected and placed in the map.

4.2 Localizing user position

User localization has to be performed only once to localize the correct user position. When the program is initialized, the user or camera is set to be in the position (0,0) with 0º orientation (as shown in figure 5 d). The map is not oriented and positioned correctly with respect to the user. Once the front-view camera captures an image target, the application will run the user localization algorithm defined in the previous section and obtain the correct user-map orientation and positions. After applying and obtaining the correct transform, the application was shown the next image target. This process was repeated for all image targets in the scene. Figure 5 [a-c] show the successful transformation of the positions and orientation of the user with respect to the image targets. Table 3 shows the user-map transformation.

Table 3. shows the original and transformed position and orientation of the map with respect to the user (camera view)

w.r.t	Image Target (Figure 5 (a))		Image Target (Figure 5 (b))		Image Target (Figure 5 (c))	
	Original	Transformed	Original	Transformed	Original	Transformed
X	0	-2.9	-2.9	-4.0	-4.0	5.5
Y	0	0.6	0.6	-0.9	-0.9	-4.7
θ	0°	204.3°	204.3°	44.3°	44.3°	283.6°

5. CONCLUSION

This paper describes the development of a head-worn augmented reality (AR) based vision-aid indoor navigation system, which localizes the user without relying on a GPS signal. The proposed method determines a user's location via computer vision and provides relevant augmented reality content based on that location. The proposed localization algorithm can be used to localize a user based on the knowledge of image target locations and its relationship with the map. Moreover, users can be localized in offline systems using stored image features as navigation-related landmarks. The technology enables users to navigate indoor environments by issuing navigational aids and information directly on the augmented reality headset. More importantly, this indoor visual positioning system can be generated from floorplans or architectural CAD models. Computer simulations indicate that the proposed augmented reality-based vision-aid indoor navigation system can successfully provide precise simultaneous localization and mapping in a GPS-denied environment. Future work includes testing the application on larger datasets and incorporating eye-tracking capabilities to enhance navigation capabilities.

REFERENCES

[1] F. Zafari, A. Gkelias, and K. Leung, "A survey of indoor localization systems and technologies," arXiv preprint arXiv:1709.01015, (2017).

[2] D. Solanki, and V. Khare, "Optical navigation system for robotics application." 268-272 (2012).

[3] Sygic incorporates augmented reality into GPS navigation app : GPS World, https://www.gpsworld.com/sygic-incorporates-augmented-reality-into-gps-navigation-app/ (2019).

[4] RECENTLY POPULAR TECHNIQUES AND TECHNOLOGIES USED FOR INDOOR LOCATIONING SYSTEMS @Medium, https://medium.com/@Fizmon/recently-popular-techniques-and-technologies-used-for-indoor-locationing-systems-13afdb6a6adb (2016).

[5] A. M. Hossain, and W.-S. Soh, "A survey of calibration-free indoor positioning systems," Computer Communications, 66, 1-13 (2015).
[6] P. Bahl, and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system." 2, 775-784 (2000).
[7] Wikipedia, Indoor positioning system, https://en.wikipedia.org/wiki/Indoor_positioning_system.
[8] G. Sterling, and S. Hegenderfer, Everything You Always Wanted To Know about Beacons, https://www.brighttalk.com/webcast/635/107617 (2019).
[9] Indoor Localization Using RFID, https://www.infsoft.com/technology/sensors/rfid (2019).
[10] Indoor Navigation Using Beacons - Accuracy & More, https://www.infsoft.com/technology/sensors/bluetooth-low-energy-beacons (2019).
[11] Provides a comprehensive overview of indoor location technologies, including GPS, WiFi, iBeacon and RFID - comparing the advantages and disadvantages of each, http://lighthouse.io/indoor-location-technologies-compared/rfid (2019).
[12] W. Narzt, G. Pomberger, A. Ferscha et al., "Augmented reality navigation systems," Universal Access in the Information Society, 4(3), 177-187 (2006).
[13] Blippar, Augmented Reality (AR) & Computer Vision Company | Blippar, https://www.blippar.com/ (2019).
[14] Vuforia | Augmented Reality for the Industrial Enterprise, https://www.vuforia.com/ (2019).
[15] D. A. de Lima, and A. C. Victorino, "A hybrid controller for vision-based navigation of autonomous vehicles in urban environments," IEEE Transactions on Intelligent Transportation Systems, 17(8), 2310-2323 (2016).
[16] M. Faessler, F. Fontana, C. Forster et al., "Autonomous, vision-based flight and live dense 3D mapping with a quadrotor micro aerial vehicle," Journal of Field Robotics, 33(4), 431-450 (2016).
[17] G. Gerstweiler, E. Vonach, and H. Kaufmann, "HyMoTrack: A mobile AR navigation system for complex indoor environments," Sensors, 16(1), 17 (2016).
[18] H.-C. Wang, R. K. Katzschmann, S. Teng et al., "Enabling independent navigation for visually impaired people through a wearable vision-based feedback system." 6533-6540 (2017).
[19] J. S. Smith, J.-H. Hwang, F.-J. Chu et al., "Learning to Navigate: Exploiting Deep Networks to Inform Sample-Based Planning During Vision-Based Navigation," arXiv preprint arXiv:1801.05132, (2018).
[20] E. M. Barhorst-Cates, K. M. Rand, and S. H. Creem-Regehr, "The Effects of restricted peripheral field-of-view on spatial learning while navigating," PloS one, 11(10), e0163785 (2016).
[21] D. Ball, B. Upcroft, G. Wyeth et al., "Vision-based obstacle detection and navigation for an agricultural robot," Journal of field robotics, 33(8), 1107-1130 (2016).
[22] X. Liang, T. Wang, L. Yang et al., "Cirl: Controllable imitative reinforcement learning for vision-based self-driving." 584-599 (2018).
[23] W. G. Aguilar, V. S. Salcedo, D. S. Sandoval et al., "Developing of a video-based model for UAV autonomous navigation." 94-105 (2017).
[24] S. S. Rautaray, and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," Artificial Intelligence Review, 43(1), 1-54 (2015).
[25] J. Bai, D. Liu, G. Su et al., "A cloud and vision-based navigation system used for blind people." 22 (2017).
[26] H. Suenaga, H. H. Tran, H. Liao et al., "Vision-based markerless registration using stereo vision and an augmented reality surgical navigation system: a pilot study," BMC Medical Imaging, 15(1), 51 (2015).
[27] C. Kanellakis, and G. Nikolakopoulos, "Survey on Computer Vision for UAVs: Current Developments and Trends," Journal of Intelligent & Robotic Systems, 87(1), 141-168 (2017).
[28] F. Bonin-Font, A. Ortiz, and G. Oliver, "Visual navigation for mobile robots: A survey," Journal of intelligent and robotic systems, 53(3), 263-296 (2008).
[29] A. Mcfadyen, and L. Mejias, "A survey of autonomous vision-based see and avoid for unmanned aircraft systems," Progress in Aerospace Sciences, 80, 1-17 (2016).
[30] S. U. Kamat, and K. Rasane, "A Survey on Autonomous Navigation Techniques." 1-6 (2018).
[31] M. Mura, Y. Abu-Kheil, G. Ciuti et al., "Vision-based haptic feedback for capsule endoscopy navigation: a proof of concept," Journal of Micro-Bio Robotics, 11(1), 35-45 (2016).
[32] B. H. Y. Alsalam, K. Morton, D. Campbell et al., "Autonomous UAV with vision based on-board decision making for remote sensing and precision agriculture." 1-12 (2017).
[33] F. Valenti, D. Giaquinto, L. Musto et al., "Enabling Computer Vision-Based Autonomous Navigation for Unmanned Aerial Vehicles in Cluttered GPS-Denied Environments." 3886-3891 (2018).

[34] E. Garcia-Fidalgo, and A. Ortiz, "Vision-based topological mapping and localization methods: A survey," Robotics and Autonomous Systems, 64, 1-20 (2015).
[35] M. Neges, C. Koch, M. König et al., "Combining visual natural markers and IMU for improved AR based indoor navigation," Advanced Engineering Informatics, 31, 18-31 (2017).
[36] Microsoft HoloLens | Mixed Reality Technology for Business @hololens, https://www.microsoft.com/en-us/hololens (2019).
[37] W. G. Van Toll, A. F. Cook IV, and R. Geraerts, "A navigation mesh for dynamic environments," Computer Animation and Virtual Worlds, 23(6), 535-546 (2012).
[38] J. Whyte, N. Bouchlaghem, A. Thorpe et al., "From CAD to virtual reality: modelling approaches, data exchange and interactive 3D building design tools," Automation in construction, 10(1), 43-55 (2000).
[39] Mapbox @Mapbox, https://www.mapbox.com/ (2019).
[40] Welcome to the QGIS project!, https://www.qgis.org/en/site/ (2019).
[41] U. Technologies, Unity - Manual: Inner Workings of the Navigation System, https://docs.unity3d.com/Manual/nav-InnerWorkings.html (2019).
[42] F. Duchoň, A. Babinec, M. Kajan et al., "Path planning with modified a star algorithm for a mobile robot," Procedia Engineering, 96, 59-69 (2014).