IARD: Interconnected Augmented Reality Device for Disaster Rescue Missions

Abhik Kumar Dey
Arizona State University
akdey@asu.edu

Rohith Eppepalli
Arizona State University
reppepal@asu.edu

Sethu Manickam
Arizona State University
smanick4@asu.edu

Sree Vashini Ravichandran
Arizona State University
sravic17@asu.edu

***Abstract* — Fire rescue missions involve several issues pertaining to locating victims trapped in buildings and navigating to their location amidst poor visibility due to the ongoing fire. This topic is a hotspot for research using cutting edge technology to aid the first responders in attempting safe and efficient rescue operations. Existing methods include the use of an augmented reality enabled mask equipped with thermal and radar enabled vision to see through walls. While earlier methods did not attempt to connect and improve perception, our proposal aims to expand the vision coverage of each user wielding the Interconnected Augmented Reality Device (IARD) mask by identifying and locating victim's coordinates. The cloud server connects the data generated by other IARD masks by correlating environmental and contextual data while the AR system overlays the obtained victim's coordinates in the physical world in real-time. The precise working of a mask involves downloading the refreshed data every 't' seconds from the cloud, reconstructing the environment of the newly downloaded data of coordinates of various articles in and around the entire disaster environment, and finally overlaying it on the first responder's augmented reality screen. This being an idea paper, we have proposed architecture using IoT along with AR to alleviate this problem.**

***Keywords*— Augmented Reality, IoT, Computer Vision, Convolutional Neural Network, Cloud Computing**

I. Introduction

During disaster rescue missions, locating the individuals in trouble is one of the primary hurdles faced by the rescue team, especially during fires when people get trapped inside buildings. Augmented Reality (AR) glasses equipped with thermal imaging can help in spotting areas of higher thermal sensitivity but the range of each sensor is limited to a few meters. Another aspect of this problem is that the path traced out to reach the located target may have been blocked by the ongoing fire. There needs to be a unified solution to combat such situations.

In today's scenario, first responders gear up with a radio sender and a receiver, to get updates about the area of rescue. Together, they pinpoint a few probable spots where individuals could be stuck. But the information remains limited to the sender and the receiver involved, increasing chances of miscommunication or incorrect interpretation of instructions, both of which could delay the rescue process. Hence, to mitigate these problems, we propose an augmented reality system where first responders have interconnected vision and share critical information about the victims to have faster and safer rescue missions.

II. Related Work & Motivation

Existing interconnected AR systems are limited by computing power and communicating bandwidth. Researchers are looking to improve the interconnected augmented reality experience by improving memory management, having better response time by computing closer to end-nodes through advancements in fog and edge processing, computer vision, and object detection [1]. Interconnected AR systems for ad-hoc networks of autonomous vehicles where hazards and occlusion information can be shared between cars to have more visibility in order to have more time to respond are being pursued [2]. Current vision enhancement systems used by disaster management teams do not employ real-time shared vision, essentially making every person rely on their own to perceive the disaster environment.

Range-R is a device that uses a radar system that can see through walls [3]. The radar system's high sensitivity is capable of detecting a man's heartbeat deep inside the debris of a building, behind several walls. The sensor technology used in this device is called Through-The-Wall Sensors (TTWS). Though theoretically, it seems easy to implement, in practice the TTWS has to combine several technologies and advanced data processing units in one device. Most TTWS operate in the range of 1 to 10 GHz, whose radiation is good enough to penetrate through concrete, wood, and walls but not suitable for walls thicker than 12 inches. Also, lowering the frequency may not give accurate information.

Other more pronounced solutions make use of AR to overlay additional information. Several manufacturers such as Qwake Tech [4] have been working on developing AR masks that help in overlaying a simplified image of the environment along with some key information on their AR glasses. This device looks promising as it tends to eliminate the issues caused by smoke and debris inside the burning building, and shows the user only necessary imagery. However, the concept of a shared vision is yet to be built.

III. Proposed Solution

Thermal and radar enabled goggles [3] are capable of spotting human activity beyond walls up to a few meters. The thermal sensors are also capable of detecting fire in and around the first responder's location. Moreover, the AR device is equipped with a GPS and a Gyroscope. Thus, we can not only get the 2-dimensional coordinates of the target but also the height at which the target is located (using the Gyroscope), thereby adding a third dimension to the space. Owing to this, the 1st first responder, equipped with IARD, will have access to the raw coordinates of a target's location as identified by the thermal and radar enabled goggles. To explain, suppose persons A and B are on the 2nd and 4th floor of a 5 storey building wearing IARD. The sensors on each device can scan up to a maximum of 3 consecutive floors. This means person A can scan the 1st, 2nd, and 3rd floor, and person B can scan the 3rd, 4th, and 5th floor. Our proposed idea targets to combine the information collected (possibly victim location) from both the devices in such a way that both A and B will have a visual of the whole 5 storey building on their respective AR screen. Thus, after combining, A will also have a visual of any activity happening on the 5th floor and vice versa.

Our proposed method involves the use of the Gaussian edge detection technique [5] to eliminate occlusion and showing only the relevant edges. We are also proposing an AR overlay, which would help the first responder in quickly navigating through the smoke-laden building. What the current AR devices do not provide is a connected vision, we envisage a system where multiple devices when connected over IoT can be used to provide a more robust disaster scene by expanding the coverage of the firefighter's vision. Our proposed solution involves several steps which have been organized as follows:

A. ***Thermal Scan of the Environment***: Once the first responder switches on his device, the IARD will begin scanning thermal spots in and around his current location. We are also planning to extend X-Ray imagery using a toned-down and safe to use radar enabled goggles with adaptive frequency change that is capable of scanning through walls for spotting human activity. The data is sent to the cloud server for object identification using advanced Machine Learning algorithms [6]. The classified information is stored in a centralized cloud repository.

B. ***Offloaded Object Detection***: The machine learning model will be available on the cloud. The amount of data that is being transmitted will be of high volume as we are talking about high-quality thermal vision videos and corresponding radar information, and would require a wider bandwidth connection to be able to transfer, and compute quickly and classify the entities as a victim, pet, or other.

C. ***Downloading from Server and Rendering the Visual***: The AR device will begin fetching data from the centralized cloud repository every 't' seconds to see if there are any new entrants. Any new victim location is then rendered and augmented into the physical environment overlay onto the screen of the AR glass. All the first responders connected in the same network

Figure 1 *Visual through the eyes of a first responder*

will begin visualizing the victim on their respective AR glasses. An important point to note here is that the person wielding the device may be at building 1, but he will also be able to perceive the victim spotted at building 2 which may be 50 feet away from his current location.

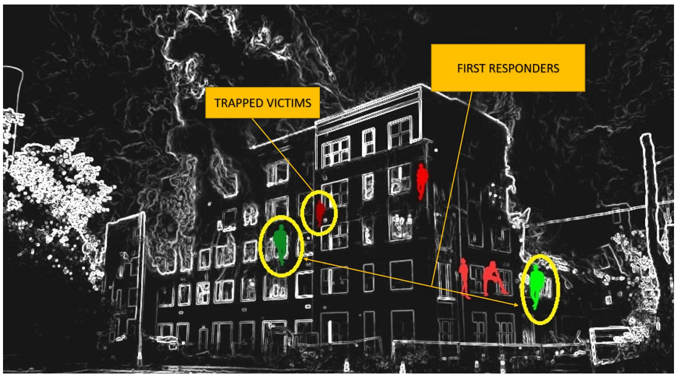

Figure 2 *Proposed visual from outside the building*

Upon rendering, the visual perceived by the first responder is shown in Figure 1. The proposed visual from outside the building is shown in Figure 2.

- The wearables will have sensors with which they will be connected to each other within a certain range of distance and another sensor through which they will transfer data to the server.
- Once the wearables are switched on and connected, they start sending data at the same time to the server.
- The server takes the data and combines in a way that the vision is expanded over a specific range.

A. ***System Architecture***:

Figure 3 shows the architecture of the proposed idea. Clients (P_i) are wearable devices connected to each other and the server. They start sending data to the server at the same time. The server on receiving the response from the wearables starts extracting videos at every *'t'* seconds. Then it pushes the *'t'* second video to a queue (Queue1 in the diagram) at instance say *t'* (t prime). Queue 1 contains queues of all the devices. Each cell contains the video of instance *t'*. Thus at *t'* = 0 we have the queues of all the wearables with *'t'* second video and like this data will be added to the queue.

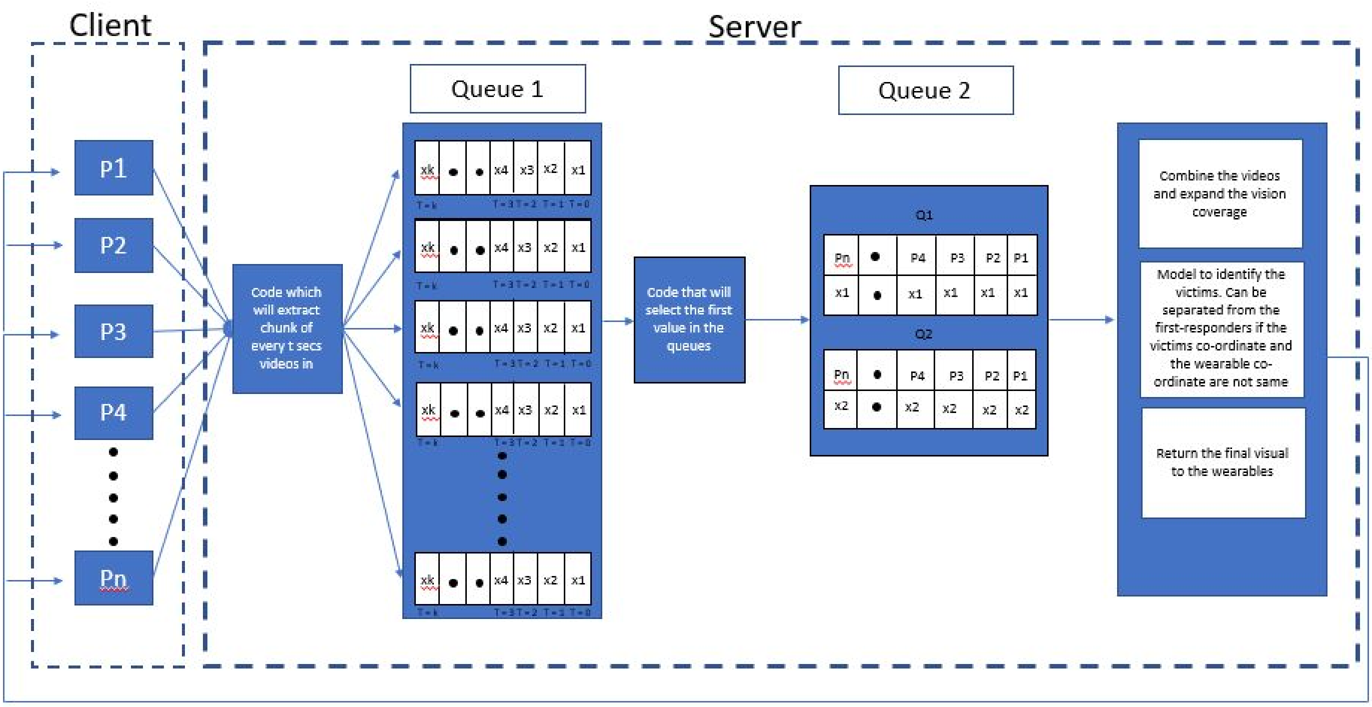

Figure 3 *Architecture Overview*

IV. Implementation Overview

To implement our proposal, the system overview is as follows:

In the next tier, the videos from Queue 1 are taken for all the wearables. We are maintaining a queue data structure. Hence, the first video in the queue will be taken for all the devices and placed in another queue (Queue 2). Now, this queue contains videos of all the devices. These videos will then be merged in such a way that the resultant video will be of length *t* seconds but will have a large dimension. The resultant video will be the expanded vision of the total area in which all the devices are present. Also, there are other operations that will be performed like identifying and marking victims, plotting the safest exit route through the building based on the structure, etc. In order to reduce latency in data processing, we have introduced another queue in the module Queue 2. Thus, whenever Q1 is not empty, the data will be pushed to Q2 and once the processing of Q1 is complete, the data will be taken from Q2.

Once all the data is gathered, the resultant video will be overlayed on the wearable devices by augmenting the victim location in the physical environment. There will be a round trip delay between sending the transmission and receiving this information, though. Suppose at 10.52 pm the data was sent to the server, all the people wearing the device will see the visual of time 10.52 at 10.53 pm. We should aim to minimize this delay as much as possible so that the users can get nearly real-time overlays.

V. Feasibility and Analysis

The idea we have proposed in the paper and the system architecture we explained above has not been implemented yet. Developing the prototype of the application was beyond the scope of our activity. However, our idea and the model is currently not feasible to implement because of the below points:

- Merging visions from multiple devices together so as to expand the vision coverage is not yet implemented but research is going on.
- Wide range of bandwidth is required to send data to the server, process them, and return it to the device so as to have a real-time view. With current technologies we have, it is difficult to implement the whole process rapidly and the target is to minimize the delay as much as possible. The latency between sending and receiving the data should be between a few milliseconds.

VI. Conclusion

Advanced scanners that are capable of adapting to the physical environment so as to switch to different frequencies and utilize a wide range of radio spectrum are under research. The tradeoff between the intervenable frequencies required for walls but not harmful to humans requires experimental investigation. Besides, the device has to undergo surplus training to adapt to the environment and requires advanced Machine Learning algorithms with better computational speed to interpret and classify the perceived objects, especially through a fast measurement process. Though experimental solutions look promising, the technology needed to implement it is still quite far fetched which requires a minimum of 3 to 4 years to be readily available for use. The voluminous data comprising high-quality videos of thermal and radar imaging would require 5G technology at the bare minimum. As per our estimation, with the currently ongoing research on merging the visions and increasing bandwidth to have fast movement of enormous amounts of data, IARD should be ready within five years

Acknowledgment

We would like to thank Dr. Ayan Banerjee for encouraging and giving us the opportunity to express our idea and taking the time to review, support, and consider our proposal.

References

[1] E. Bastug, M. Bennis, M. Medard and M. Debbah, "*Toward Interconnected Virtual Reality: Opportunities, Challenges, and Enablers*," in IEEE Communications Magazine, vol. 55, no. 6, pp. 110-117, June 2017.

[2] K. Naseer Qureshi, F. Bashir and S. Iqbal, "*Cloud Computing Model for Vehicular Ad hoc Networks*," 2018 IEEE 7th International Conference on Cloud Networking (CloudNet), Tokyo, 2018, pp. 1-3.

[3] https://www.sds.l3t.com/military-first-responders/RANGE-R.htm

[4] https://www.qwake.tech/

[5] .G. Deng and L. W. Cahill, "*An adaptive Gaussian filter for noise reduction and edge detection*," 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference, San Francisco, CA, USA, 1993, pp. 1615-1619 vol.3.

[6] Zhao, Mingmin, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina Katabi. "*Through-wall human pose estimation using radio signals*." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7356-7365. 2018.