

Project Analysis: Shortcomings, Future Work

This document outlines the limitations of the current Context-Aware Conversation Support with Augmented Reality and Generative AI system and proposed necessary future improvements.

Team Members:

- Sabrina Alvarado
- Keren Rivera
- Jonathan Martinez
- Peace Passos

Mentor: Dr. Chen Chen, FIU KFSCIS, Assistant Professor

Instructor: Seyedmasoud Sadjadi

Shortcomings and Limitations

The current system, while functional, faces several real-world constraints inherent to the initial development platform and architecture.

A. Hardware and Experience Limitations

- **Monochrome Passthrough AR:** The use of the Meta Quest 2 restricts the augmented reality overlays to a black and white, low resolution passthrough view of the real world. This significantly degrades the visual clarity and immersion of the Mixed Reality experience.
- **Hardware Constraints:** overall system performance (e.g., processing speed, graphical fidelity, and hand tracking quality) is limited by the older hardware of the Meta Quest 2.

B. Architectural and Performance Constraints

- **Network Dependency:** the core process – which relies on a chain of Speech-to-Text, AI processing, and Text-to-Speech – is fully reliant on network strength. Poor connectivity leads to increased latency and a degraded user experience.
- **AI Contextual Challenges:** The system can be slower for slang, niche domain-specific terms, or complex idioms due to the inherent difficulty in achieving real-time, context-aware translation and clarification.
- **Latency in Noisy Environment:** The initial microphone input and processing experience occasional latency in noisy environment.

Future Wishlist & Proposed Improvements (Future Work)

To enhance the system's robustness and utility, the following improvements are prioritized, aligning closely with the project's original vision:

- **Hardware Upgrade & Color Passthrough:** The most critical upgrade is moving to a device like the Meta Quest 3 (or equivalent) to enable full-color, high-resolution passthrough. This will create a much more realistic and seamless Mixed Reality experience.
- **Comprehensive Multilingual Expansion:** actively implement support for a wide range of global language beyond the initial demo language pairs by fully leveraging the extensive capabilities of the Google Gemini API and robust STT/TTS/SDKs.
- **Enhanced AI and Processing:**
 - Implement on-device AI models to handle primary language and concept processing locally, significantly reducing network dependency and latency.
 - Integrate contextual scene understanding (using new headset depth sensors) to ground AR augmentations to real-world objects.
- **User Interaction and Collaboration:**
 - Develop support for multi-user AR conversation environments to facilitate group collaboration.
 - Incorporate gesture interaction for a more natural and hands-free control interface.
- **Post-Conversation Utility:**
 - Add features for summarization + transcripts of the conversation for review and documentation.