TableCast

E8: Caroline Crooks, Tahaseen Shaik, Sumayya Syeda
18-500 Capstone Design, Spring 2024
Electrical and Computer Engineering Department
Carnegie Mellon University

Product Pitch

TableCast is a unique product that brings technology to kitchen spaces to create an interactive culinary experience for cooks of all levels. Users will be guided through recipes with widgets displayed on to the kitchen table by a projector. TableCast is a completely hands-free interactive user experience through gesture commands, voice commands and computer vision techniques. These are our most key requirements and will guide our priorities for the development timeline.

Based on our user study and quantitative tests, we have determined that the system is usable with our reliance on symbols and images to convey information. The projected content is interactable and accurate to the table configuration as shown in Figure 3. The voice commands and gesture recognition have been shown to work with different user demographics as required. Using TableCast, we hope more users will feel confident in a kitchen setting.

System Architecture

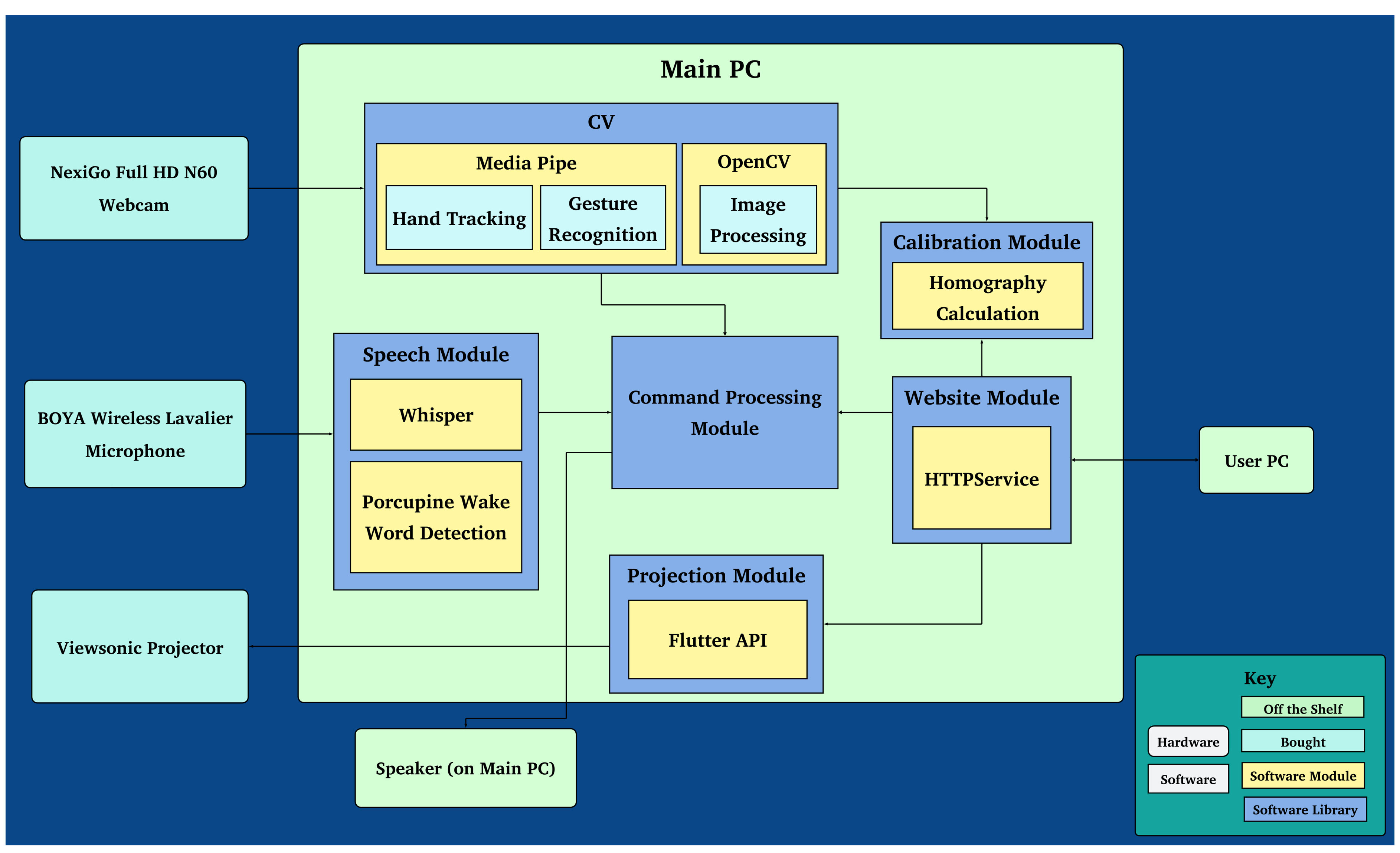

Figure 1: Hardware & Software System Components

Our main system is hosted on a Legion Pro 7i Gen 8 Intel computer. To accommodate our physical components, the host computer communicates with several off the shelf hardware components: a USB camera, a wireless microphone, and a 4500 ANSI Lumens projector. These hardware components allow the user to interact with recipes using a hands-free approach by utilizing speech and gestures. The chosen projector allows our system to work in a dim kitchen environment with a sufficient brightness. The microphone enables the user to use TableCast is a noisier kitchen environment.

We applied a module-based structure for our software system integrations. This makes it easy to unit test and validate each subsystem. Our calibration and CV modules extensively utilize NumPy and OpenCV to process the kitchen setup and projected content. We further use Media Pipe to track hands for the gesture tracking component of TableCast. The voice command use Whisper and Porcupine to accurately detect and process user commands. All modules communicate over multiple serial data buses using a Publisher/Subscriber interface. These modules are all managed via Flutter which is then used to display dynamic content onto the table and in our web user interface.

Conclusions & Additional Information

Our overall system assessment was guided by our user study and our V&V process that held the project against our aspirations. Users were generally satisfied with gesture and voice commands and the overall user interface.

We learned a lot regarding the limitations of various projectors and of a warping algorithm. Our project can be extended by using 2 higher quality projectors to eliminate issues (shadow, light dispersion, etc.) we faced.

http://course.ece.cmu.edu/~ece500/projects/s24-teame8/

System Description

The user starts TableCast through the webapp available on their local computer. The webapp will walk the user through the projector calibration process and then guide the user to selecting a recipe. In the projector calibration step, TableCast uses the table coordinates indicated by red markers to compute the necessary homography to warp the such that it appears straightened. The homography matrix and detected table outline is communicated to Flutter to be applied to the projection UI. The user should only need to complete the calibration once after the setup has been completed.

Next, the user can use gesture and voice commands processed by the CV & Speech Module to navigate their cooking experience. Media Pipe was heavily utilized to recognize gestures for button presses and track hand landmarks for button swipes (to indicate previous/next step). OpenAI's speech recognition tool Whisper was used to interpret voice commands with our wake word 'TableCast' so user's can pause, play, and navigate between steps.

With our dynamic display, the projections are updated and displayed on the kitchen counter as the commands are processed so that the user can enjoy guidance with reduced latency. At the conclusion of the recipe, the user is able to give feedback.

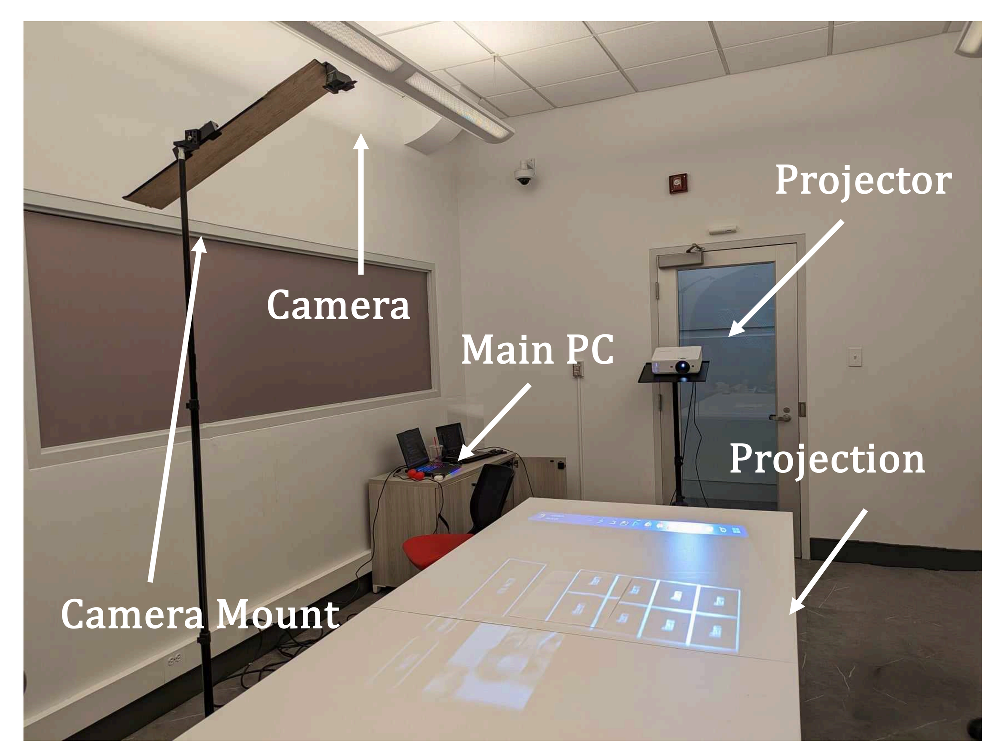

Figure 2: System Set up

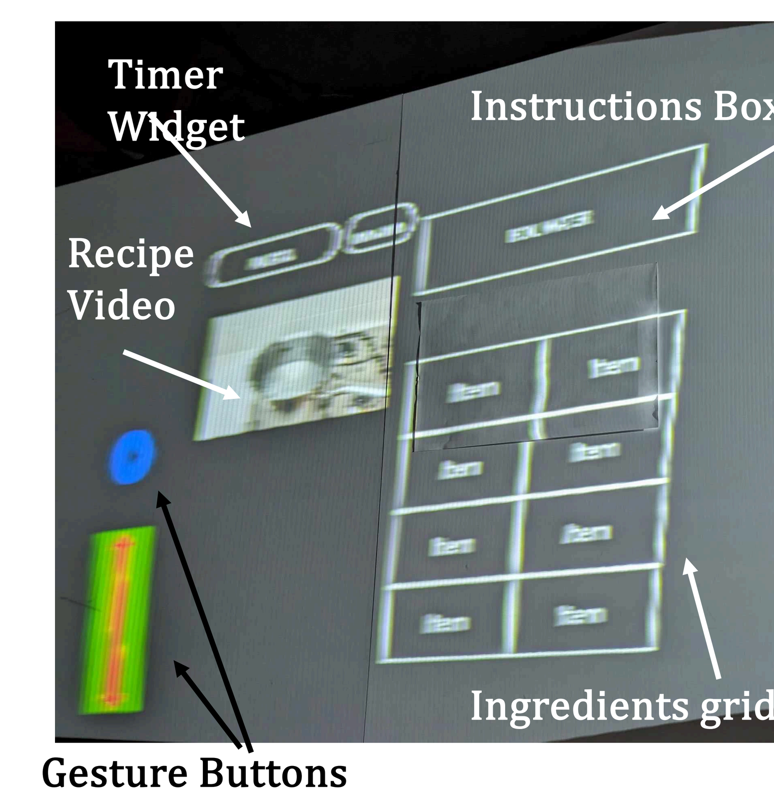

Figure 3: Projection UI

System Evaluation

Below are some of the key tests to evaluate our system's functionality.

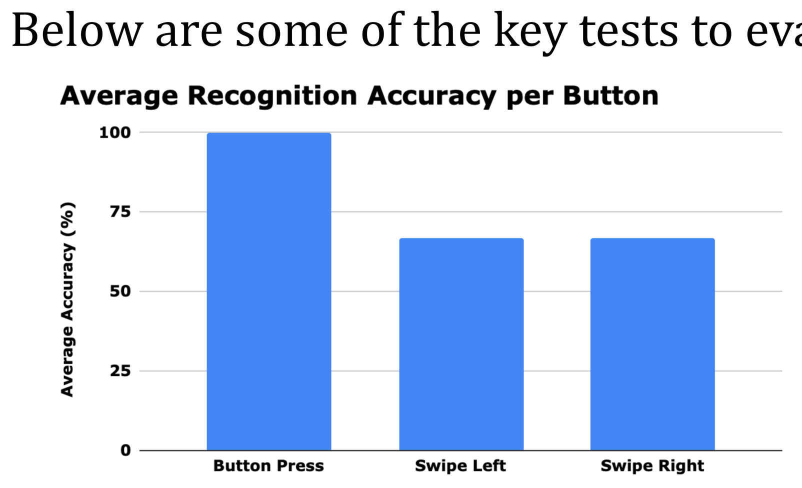

Figure 4: Accuracy Test for Gesture Commands

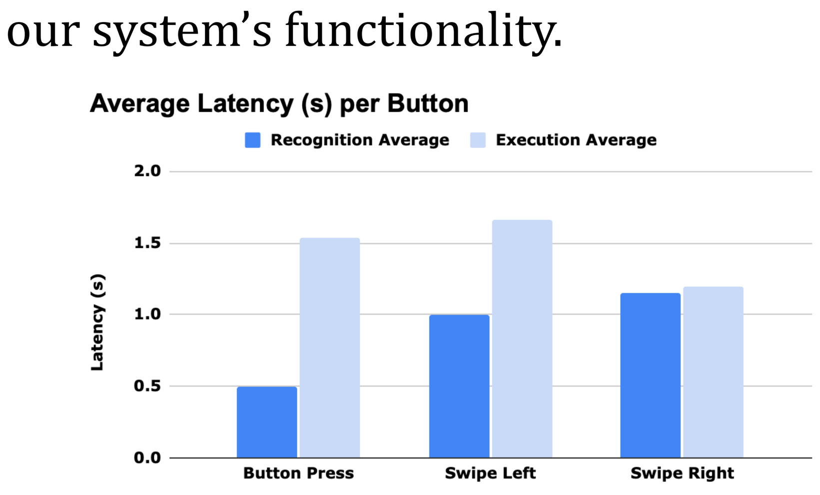

Figure 5: Latency Test for Gesture Commands

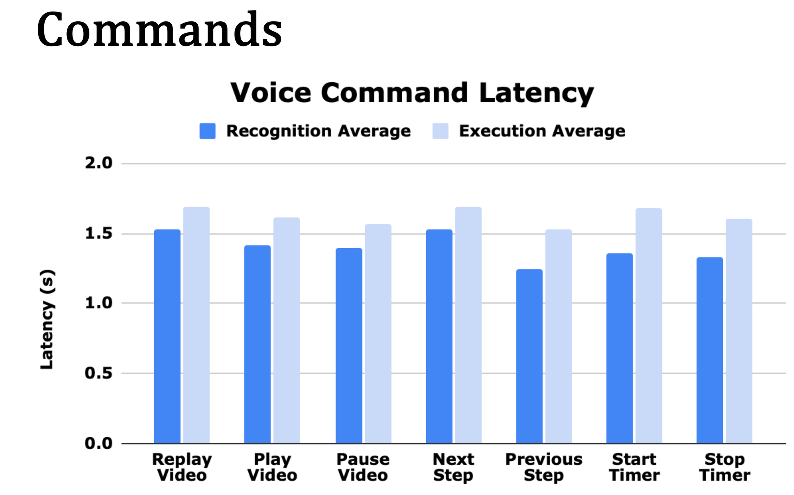

Figure 7: Accuracy Test for Voice Commands

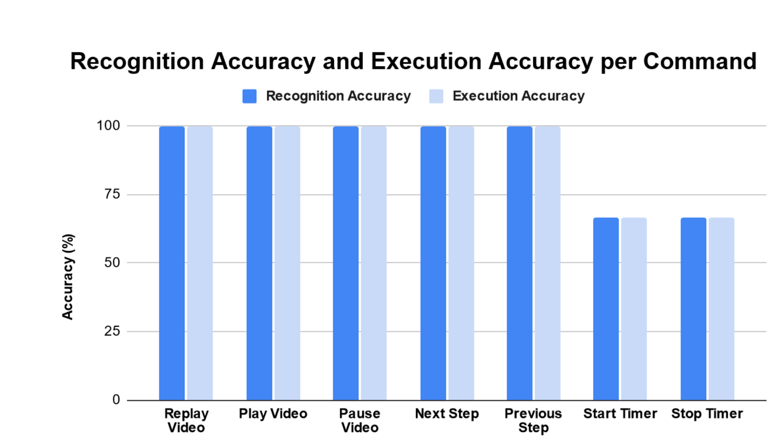

Figure 8: Latency Test for Voice Commands

Design Trade Offs	Reasoning
Moved Projector to side of user rather than placing it across the user	This change decreases the impact of a shadow on the countertop as the user cooks and increases the projector throw to maximize table space.
Choosing a PC over a Nvidia Jetson AGX Xavier for the main processing unit	AGX Xavier was incompatible with the voice command module and other necessary software packages.

Electrical & Computer ENGINEERING
Carnegie Mellon