CS3483

Multimodal Interface Design Project:

Smart Kitchen: Gesture-Driven Contactless Control System

Mid-term Design Report

College of Computing BSc Computer Science

5 NOV 2025

Student Name:	Student ID:
Wong Siu Lun	58383709
Tsoi Ho Chun	59380140

WFH Smart Kitchen: Gesture-Driven Contactless Control System

1. Project Objective

In the project, the design aims to provide a contactless kitchen control interface for the **working from home (WFH)** white-collar employees, and the housewives, allowing them to manage kitchen tasks (e.g., boiling water, making coffee, heating food) via computer camera with gesture recognition, and support of voice control.

WFH employees and the housewives usually face the challenges:

1. Cannot easily leave their workspace (especially during business meeting calls)
2. Kitchen tasks (e.g. boiling water, coffee, heating food) disrupt workflow
3. Safety issues: Wet hands, hot surfaces, or mobility impairments
4. Multi-tasking: doing housework/ taking care children while still handle kitchens work
5. Vision-impaired users may feel hard to use smart kitchen interface controls

Our system needs to minimize their distractions and prioritize safety. The target users are WFH white-collar employees and the housewives with computers, including those with mobility or vision impairments, ensuring accessibility. The interface should enhance work efficiency of the employees by reducing the time of handling kitchen works. Users should be able to control the interface with following devices: desktop/ laptop computer, with at least 720p front-facing computer camera (old MacBook camera spec.) for accurate gesture detection.

Core Task Overview

Users can perform the following tasks from their room via the multimodal interface:

1. **Boil Water**: Start the kettle -> monitor water temperature -> automatically/ manually stop boiling water.
2. **Make Coffee**: Activate the capsule coffee machine -> select size of coffee (espresso or lungo).
3. **Heat Food**: Control the oven to warm pre-prepared food (e.g., bread).
4. **Emergency Stop**: Quickly shut down all devices for safety.

User Activities

Users will:

1. Use webcam gestures in front of work desk to control kitchen appliances.
2. Receive real-time visual/ auditory feedback.
3. Use voice only after gesture trigger to avoid application unexpected activation.
4. Operation logs for records and emergency stop with simply one gesture.

2. Presentation of Designs

2.1 Micro Float Panel

The design of the micro float panel aims to show the information of current condition of the devices controlled by the application. It defaulted at the right upper corner of the screen. The condition of the devices connected with the system shown in the panel. For instance, the empty message would be indicated for the water kettle in the panel to alert users, preventing them to start the kettle without water, which may cause waste of electrical and the potential of fire hazard.

The application mode also indicated in our design hence the users would be alerted if they are using the wrong mode. For instance, toggle box of “meeting” should be checked under meeting mode, users can prevent device mis-activation with purely hand gesture.

The design should provide minimal features with minimal screen space consumption ,to minimize distraction to users’ working float and focus on working tasks.

Design 1:

Figure 1. Listing design of micro float panel

The design 1 of the micro float panel aims to show the information of current condition of the devices controlled by the application in list format. The condition of the devices connected with the system shown in the panel. Any mode selected would also be shown in the panel.

For example, in figure1, the water kettle indicated empty, with deep grey as bar background color 

, indicating that the water kettle is now empty. And the oven indicating a 64s with an unfinished loading bar 

, the 50% full progress bar showing heating process is about 50%, with the time 64s left till finish. The row of coffee makes shows which size of coffee was done/ or is there any coffee capsule in the machine ready for used. 

shows an espresso ready to serve in figure 1. Undoubtfully, 

would be shown if no coffee capsule stalled in the machine.

The vertical design of window control  , prevent users from mis-clicking the “X” button close window button by not setting it at the easiest reaching region (most up-right). Also, the vertical design can reduce the size consumed with the showing application name “smart kitchen”.

shows recent mode chosen. With the toggle choice, users can switch between modes (i.e. it is in meeting mode in figure 1.)

Design 2:

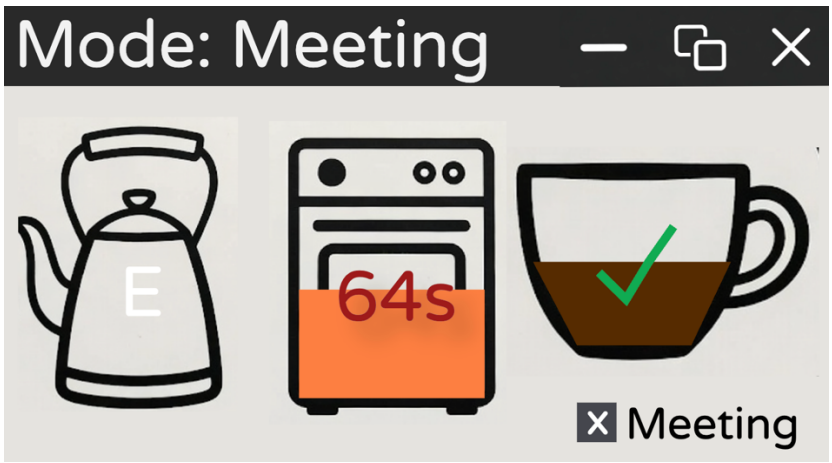


Figure 2. Icon representing design (kettle, oven, cup icons by grok, 2025)

The design 2 of the micro float panel aims to show the information of the devices controlled by the application with icons.  represents the condition of the devices connected with the system. Meeting mode selected would also be shown on the top of panel as Mode: Meeting.

For example, in figure 2  shows the water kettle indicated as empty, with the character “E” and kettle without water.  indicates the water kettle is now full.

And the oven condition represents with the image of  , the oven filled with orange showing it is in heating process, with the time 64s left till finish.

The right most cup image shows the coffee maker’s condition.  represents the espresso was prepared (with the green tick represent ready, half-filled cup represent espresso size).  represents the lungo size coffee prepared. And  represents no coffee available.

Meeting is the mode switching button. With the toggle choice, users can switch between meeting modes and non-meeting mode.

Choosing of Design:

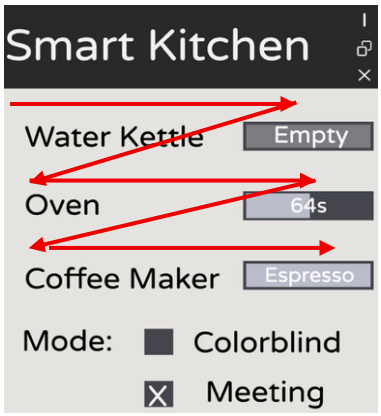
- **Micro floating panel Size**

In terms of size, the micro floating panel should be as small as possible with showing the simplest function of the interface, because smaller interface can bring less distraction to the users from the concentration of works. Design 1 takes about approximately 10% of the screen, while the design 2 should take about 8% only. Therefore, design 2 should be considered in terms of interface size.

- **Information density and Reading habit of human**

In terms of information density and reading habit of human, design 1 applied listing design, which fulfill the human habit of reading in order of the Z-pattern (Kholodov, 2023). By Gutenberg diagram, the left upper most corner should be most aware of, users can be acknowledged that

they are using the application with looking the Smart Kitchen at the left upmost corner, then view the condition of each device with the Z-path. For instance, “water kettle” from the left of first line then the “empty” shown at the right, showing the water kettle is now empty. The process of reading can be shown as below:



Although design 1 provides decent information density with the z-pattern reading listing design, design 2 provide even more information. The icons in design 2 provides information without process of “reading”. Users can simply read the icon to receive information of the corresponding device. For example,  illustrates the “oven”, “heating”, “64s left” by simply one reading process. But the design 1 required at least 2 reading processes, “oven” from the left of the row and “64s” from the right of the row. The reduce of reading steps can reduce the distraction of the WFH users.

Moreover, although design 1 provide the colorblind mode toggle button, the colorblind mode does not require frequent switch like meeting mode once it set. Therefore, the colorblind mode button can be said as secondary feature, which is not necessary for the objective of micro float interface: provides application’s key features only.

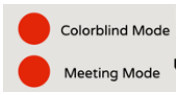
To conclude, icon-based design (design 2) chosen for the micro floating panel.

2.2 Main Dashboard Interface

The main dashboard is the main command interface that users mainly use for checking details or fine-tune a task. It stays hidden behind the micro float panel most of the time, so the screen stays clean for the other applications or browsers. The main components in the dashboard interface are the **Control Panel** which shows the information of the features of the application and **navigation menu** which allows the connect between different interface of our application.

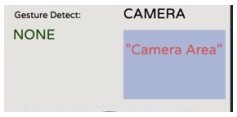
The control panel composed with the following features:

1. Mode Selection
2. Camera and feedback
3. Device condition

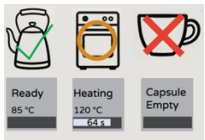
Mode Selection area 

: two circle buttons for choosing mode. If mode selected, the circle would be ticked. Users can see whether the current mode is colorblind mode or meeting mode or either. The red button allows users to switch mode. For example, 

indicated both colorblind mode and meeting mode were on.

Camera and feedback 

: The grey box gives insta-feedback of computer camera, allow users check whether the camera works and they are doing the correct gesture. Gesture Detect area would indicate what gesture now currently detected. Users can activate the desire device with voice command correctly after verifying their gesture confirmed.

Device Condition 

: similar with the micro floating panel, it shows the current condition of the devices. Users also able to change default temperature and the timer in the tag under the oven icon. The status of device also shown corresponded with icons. ✓Green tick represents the device ready for use. ◎Orange circle represents the device is processing. ✖Red cross represents the device not able to be used. In the example, the water kettle shows “Ready” “85°C" with empty progress bar, the oven indicates "Heating 120°C 64s" with an incomplete progress bar, and the coffee maker alerts "Capsule Empty" with both word description and red cross to prevent unnecessary activation.

Design 1:

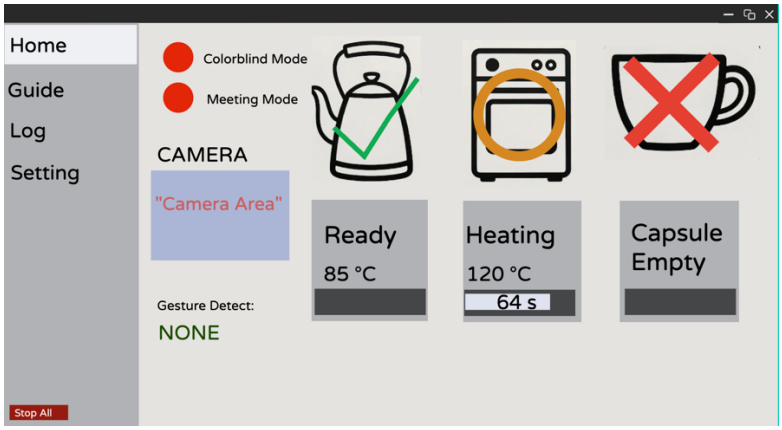


Figure 3. Left-side menu bar dashboard (kettle, oven, cup icons by grok, 2025)

The design 1 of the dashboard use vertical left-side menu bar as navigation bar. This layout is a familiar interface which like common web applications, users can easily switch between pages (e.g., Home, Guide, Log, Setting) without learning of using this interface. The menu bar optimizes minimal horizontal space, leaving a major part of the screen for display of 3 devices. The page currently shown indicated by the white background under the word. In figure 3, 

differed from the other options, indicating user currently in the home page. The vertical design also reduces mis-clicks of emergency “stops all” button. By placing important options in a linear format, WFH users working efficiency boosted during multitasking.

Design 2:

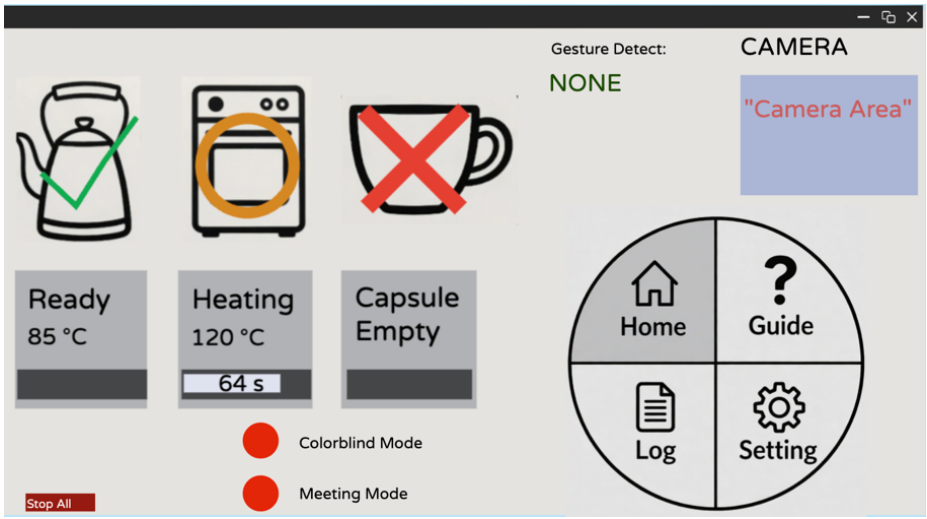


Figure 4. Circular menu dashboard (kettle, oven, cup, menu icons by grok, 2025)

The design 2 of the dashboard uses a circular menu 

at the bottom-right for navigation. The left upper corner of the circle shaded area 

indicated that the user currently in home

page. This circular menu aims to minimize need of users' cursor movement by condensing options (e.g., Home, Guide, Log, Setting) into a radial layout. However, users usually unfamiliar with circular menus. Like design 1, device statuses are displayed, but on the left side of the interface. This design prioritizes Fitts's law, by decreasing the movement required by the circular menu, which increase 15 - 20% productivity of users (Callahan, 1988). Chance of mis-clicking also minimized under this design, due to the angle can well separate different options in circular menu.

Choosing of Design:

- Utilization of Interface space

In terms of space utilization, design 1 optimizes minimal horizontal space by having a vertical bar at left most, leaving a major part of the screen for information display. The information can be well arranged in the interface evenly. Although design 2 has similar major part design with design 1, the circular menu design has a critical weakness point in terms of space utilization. The circular menu wasted the corner area around the menu. For example, if the radius of circular menu is r, the area used π x r^2, but the actual area consumed is (2 x r) ^2, the different in area wasted. Although we can design the menu can be hided and only called out when used, consider to the function and the frequency of calling out the menu, Users usually can control all functions by hand and voice command, hence the frequency of calling out the menu is low. It is pointless to have an extra step to call out the menu. Design 1 should be considered in terms of space utilization.

- Scalability for further development

In terms of possible scalability, design 1 applied vertical bar design, can be modified easily, extra row can be simply added to the bar if further pages implement in the future. However, in design 2 circular menu, the menu items are limited due to each page share a part of circle. For example, 4 pages share 25% each of the circle, if the number of pages increased to 10, each page shared only 10% of total area, which is only 36 degrees. It is not possible to show the page name with limited area. Design 1 should be considered in terms of scalability.

- Fitts's Law application and minimize cognitive load

In terms of usability, vertical menu bar in design 1 is a familiar interface which like commonly used in many web applications, users are already well-trained, they can easily control the menu bar without learning of using this interface. The left most navigation bar also applies Fitts's law, users move their curser to the left side count as infinite cursor move. By Fitts's Law,

$$T = a + b \times \log_2(D \div W + 1)$$

By fixing the menu at the most edge at left side of screen, D can be minimized for any cursor position. Therefore, the T (time consumed) also be reduced. Comparing with design 2, prioritizes Fitts's law by decreasing the movement required by the circular menu, which increase 15 - 20% productivity of users (Callahan, 1988). The circular menu not often used in other applications. Our target users (WFH employees and housewives) may not be familiar with these kinds of controls. Design 1 considered as a better choice in terms on the Fitts's law.

To conclude, left side menu bar design (design 1) chosen for the dashboard design, due to its benefits brought by better utilization of interface, better scalability for application development and the fulfillment of Fitts's law.

2.3 Guide Page

The guide page aims to provide users with an interactive tutorial on hand gestures control follow with bilingual voice commands instructions, ensuring they can quickly be familiar with the system's multimodal features.

In this page has a live camera area 

for real-time feedback practice, images for simple visual reference for users, and detection feedback for gestures and voice. This design accessible with voice by offering bilingual (English, Cantonese/Mandarin) commands. Safety features like emergency stop also emphasized in this page. Users informed the emergency stop works with either voice command or hand gesture in any mode.

Design 1:

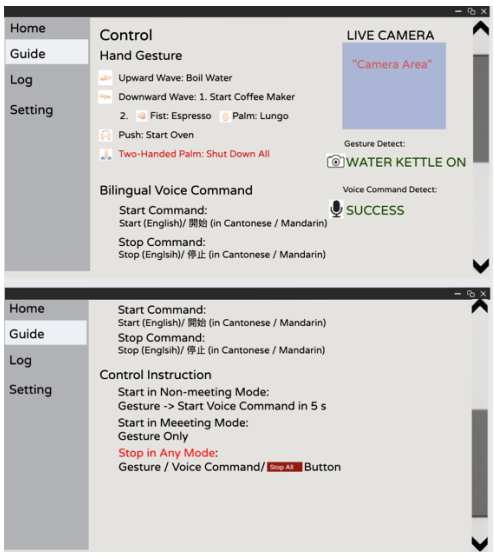


Figure 5. Simplified guide page with scroll bar (icon by Unicode Consortium, 2022)

The design 1 of the guide page expands on control instructions with scroll bar, providing further text explanations for simple control instruction. While the live camera and demo areas on the right top of the page, next to the hand gesture instruction. The area includes the real-time detection feedback from computer camera 

for gestures and voice but prioritizes textual clarity for users 

. The guide on the left shows simple gesture image with textual guidance (e.g.  upward wave image represents boil water command). At the bottom of design 1, it also lists emergency stop as "Gesture / Voice Command / Stop All Button" which is red in color for emphasis. Design 1 is more static, potentially require less interactivity for users quick learning on the control of application, but offers comprehensive bilingual support and safety reminders to prevent mis-activation during WFH meetings.

Design 2:

Figure 6. Interactive guide page (hand images by grok, 2025)

The design 2 of the guide page includes a live camera area 

on the left for immediate gesture feedback similar with design 1, a demo image on the right showing the correct hand pose from the perspective of user (first person angle). For example, in figure 6, upward wave hand shown as the demo image and real-time feedback below (e.g., “Gesture Detect: NONE” 

in red turning to “WATER KETTLE ON” 

in green). In Feature 1: Water Kettle (figure 6), users see the upward wave demo and must replicate it, following with voice command detection “Start”, activating only after the gesture. The page includes bilingual commands (e.g., “Start (English)/開始 (in Cantonese / Mandarin)”) 

and control

instructions. On the top right corner, a progress indication 

showing the current process in indigo in color and total page number of “4”. Design 2 promotes hands-on learning, reducing errors in actual use.

Choosing of Design:

- Learning Efficiency and Universal Usability

Design 1 more tends to be a guidebook for users, users can read the guide anytime as they wish, especially when they forget about the hand gesture control. Design 2 enforces an interactive, step-by-step practice for performing each gesture with the live camera feedback. With the interactive tutorial in design 2, novice or first-time users can learn different hand command gesture hand-by-hand, they can be promoted into knowledgeable intermittent users after the tutorial. In contrast, Design 1 relies on passive reading, users are more likely to forget the gesture control process, it is harder to promote the novice users into knowledgeable intermittent users. Therefore, design 2 promotes higher learning efficiency for users with the interactive design.

- Informative Feedback and Instruction clarity

Besides textual instruction for users, images were used in both designs to provide clear guidance on gesture and voice command control. In terms of feedback given, both designs can receive feedback from application via the area of gesture/ voice detection 

, in non-colorblind mode, positive feedback (e.g. “Success”, “Water kettle on” would be shown with green in color, while negative feedback like “None” shown in red. The detection area output decent feedback for users.

However, design 1 prioritized the quick guidance for users. Therefore, less complicated icon used to represent the motions, like  represents upward wave and represents downward wave. These icons may affect to guide the smart users but confuse the less smart users. By comparison, design 2 shows the demo motions in the 

demo area, with extra interactive prompt like “Please Upward Wave” to guide users’ actions. Design 2 obviously gives clearer guidance, with more interaction with users.

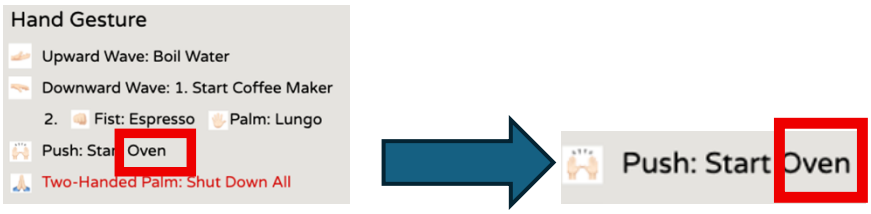
Moreover, at the top right of design 2, process indication shown with 

. Users can be acknowledged with how much they have learnt to control the application. Unlike with

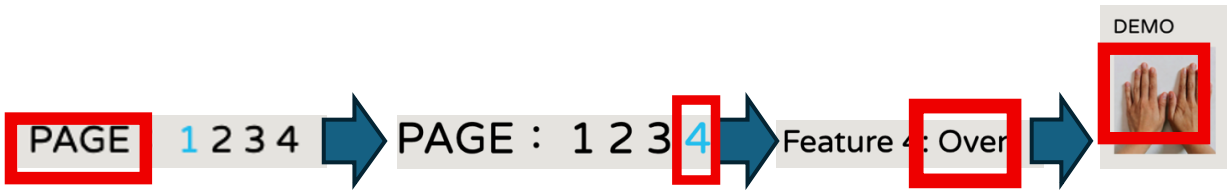
design 1, users simply screen the instructions, less interactive features to tell the learning progress.

- Revision and Memory Load

In design 1, the 1- page guidance offers faster revision for users. Users can simply screen for the answer they desire. For example, if user wish to search the gesture command for starting oven, he/ she only screen for the keyword “oven” in the interface. Then he/ she would find the gesture for start oven. The focusing process of user shown as below:



Comparing with the above, design 2 provide more steps for searching specific gesture command for users. Users must search the gesture page by page. Until he/ she finds the page desired, which takes more steps and more complicated process required. Focus of user shown as below:



Users must search for the page tag in design 2, next with the page number, then confirm the device is correct, last to look for the demo motion. It may take even more steps to achieve the simple search for specific gesture if the user does not recognize the exact page they desire. Obviously, it is less convenient for the users to search for specific motion control in design 2. In terms of short-term memory load for revising the control processes, design 1 is leading ahead.

Due to the importance of promoting novice users into knowledgeable intermittent users, our group would consider using design 2 (interactive guide) as the primary guide page for tutorial. However, we also believe quick access guidance from design 1 is also necessary due to its convenient control revision, so that design 1 would also be kept as secondary guide for user.

3. Final Designs

The final interface design is integrated the strengths of each component design comparison to form a **cohesive, efficient, and user-friendly** for different users' multimodal system. To achieve the above objectives, the design **prioritizes minimal distraction, allows rapid task execution, utilizes dashboard spaces, provides informative feedback and universal usability** for WFH users and housewives under high cognitive load.

Main Dashboard:

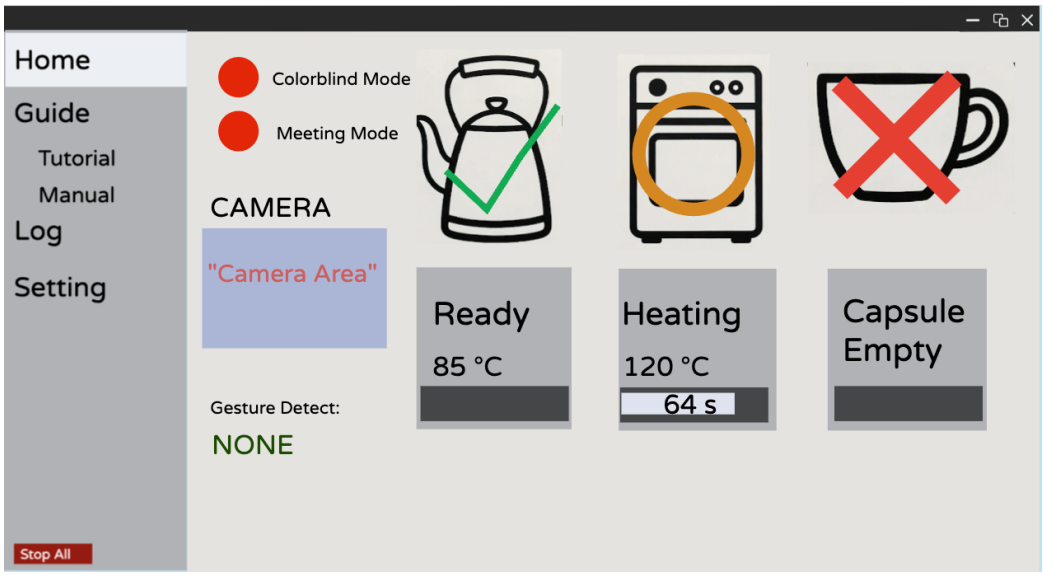


Figure 7. Home page final design (kettle, oven, cup icons by grok, 2025)

Left-side vertical menu bar chosen as the navigation structure This decision applied Fitts's Law by fixing the menu at the left most of the screen edge (D tends to 0), decrease cursor targeting time of users. The vertical layout supports scalability for future development. The central control panel shows condition of devices in a horizontal Z-pattern flow, which fits with human reading habits. The emergency stop button at the bottom left allows **rapid access** without mis clicking (far from other features). Camera feedback area is embedded in the top-right corner on main dashboard, allowing users to verify gesture input. Two mode switching button on the top left corner provide quick access for switching modes. Top right-most icons enable users switching interface size, refers to minimize, micro floating <-> full screen dashboard, and exit.

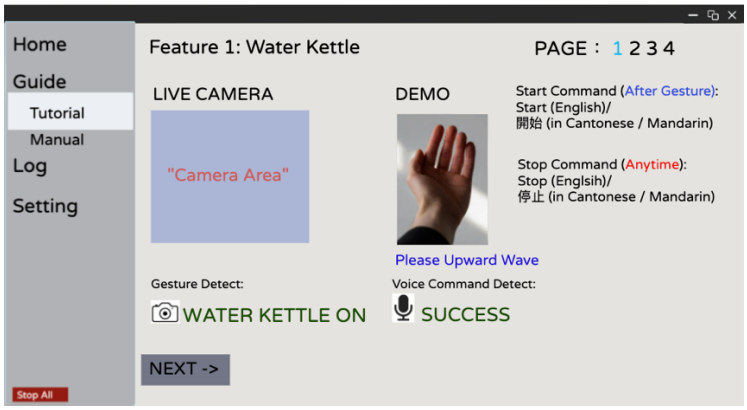


Figure 8. Tutorial final design (hand images by grok, 2025)

Users can access primary guide page via the navigation menu “tutorial” button. The primary guide page adopts interactive tutorial to promote novice user into knowledge intermittent user. Users are guided through four main steps (one per device), requiring active gesture replication under live camera. The layout follows a left camera, with a right demo image.

Progress at the top right indicates completion (e.g. “2/4”), reducing memory load and providing clear goal for users. Upon successful gesture follows with voice command (“Start” / “開始”), the system textual confirmation (e.g. “WATER KETTLE ON”). This hands-on practice enhances motor memory retention, compared to passive reading with the guidebook design.

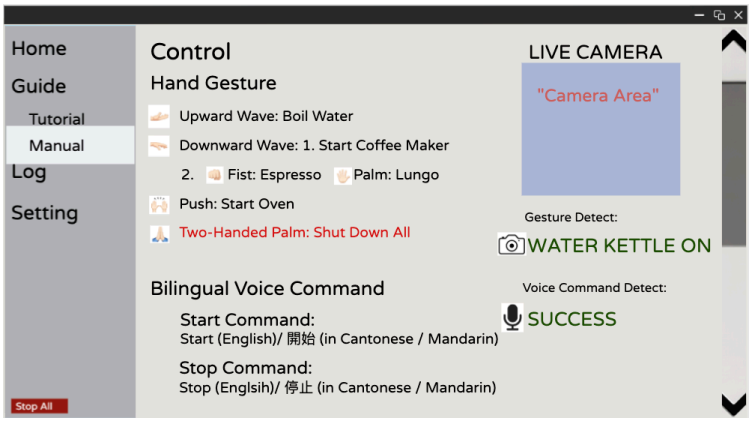


Figure 9. Manual final design (icon by Unicode Consortium, 2022)

Users can access secondary guide page via the navigation menu “Manual” button. Despite the interactive tutorial, **Scrollable Manual** is retained as a **secondary reference tool**. This single page, vertically scrollable guide allows **expert users** needing **rapid lookup** revising commands control.

The **Z-pattern scanning** also applied in the page, allows users to locate a command in short moment. This dual-mode learning system (interactive + guidebook) balances **learning efficiency** and **revisability**.

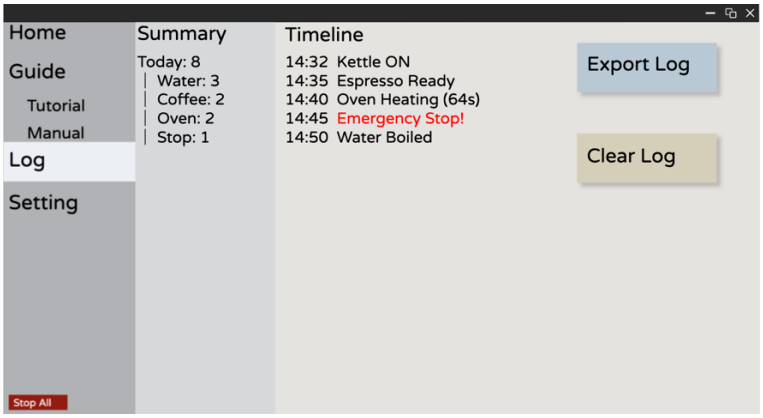


Figure 10. Log page final design (icon by Unicode Consortium, 2022)

Users can access log page via the navigation menu “Log” button. The operation log adopts a chronological layout with summary (actions taken) and timestamps. Buttons on the right allow **outputting log** as CSV file and **clearing log** records respectively.

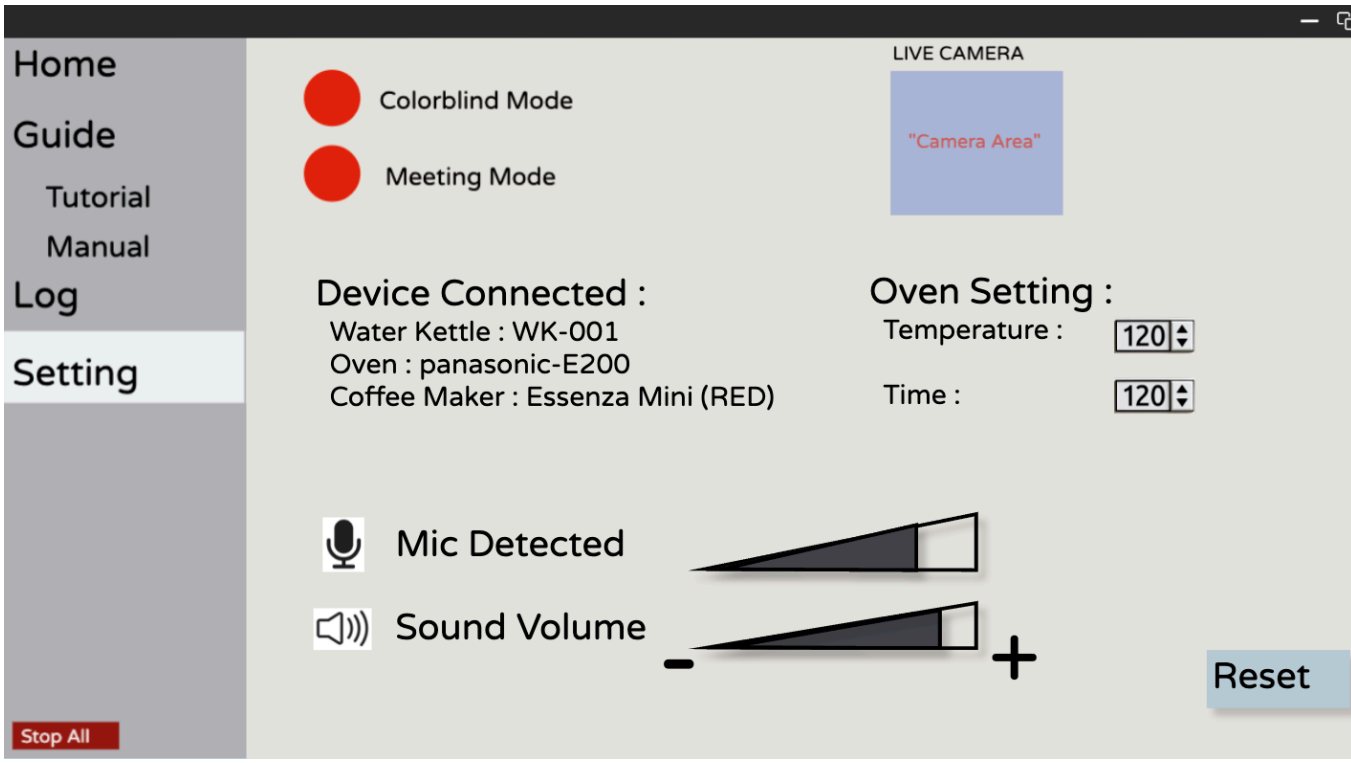


Figure 11. Setting page final design (icon by Unicode Consortium, 2022)

Users can access setting page via the navigation menu “Setting” button. The page supports **mode switching**, reviewing the mode of **connected devices**, and **functional check** of input/ output of application (voice, camera input and voice output). Reset button provides quick reset to default setting. The horizontal scroll aligns with volume control, which familiar to WFH users. Oven temperature and default timer also can be set.

Micro Floating Panel:

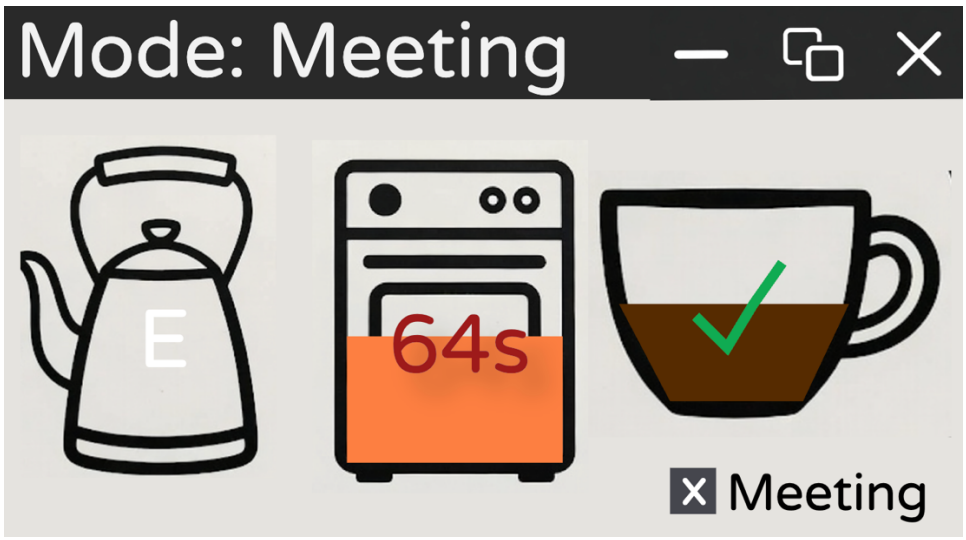


Figure 12. Micro Floating Panel (kettle, oven, cup icons by grok, 2025)

The **final micro float panel is icon-based**, occupying only **~8% screen**. Main information indicated via **single-glance icons**.

- Kettle: E (empty) | F (Full/ Still have water) | B (Water boiling)
- Oven: 64s (countdown) | progress (icon progress bar)
- Coffee: green tick (ready)| half-filled cup (Espresso Ready) | fill-filled cup Lungo Ready | Cross (No Capsule)

For example, in figure 12, showing the kettle is empty, oven heating with 64s left, espresso ready. The **meeting mode toggle** is prominently placed at the bottom, supporting **one-click activation** to prevent voice command mis-activating devices during calls. The panel can be hidden or expanded into full screen with control button on the right upper corner.

Colorblind Mode:

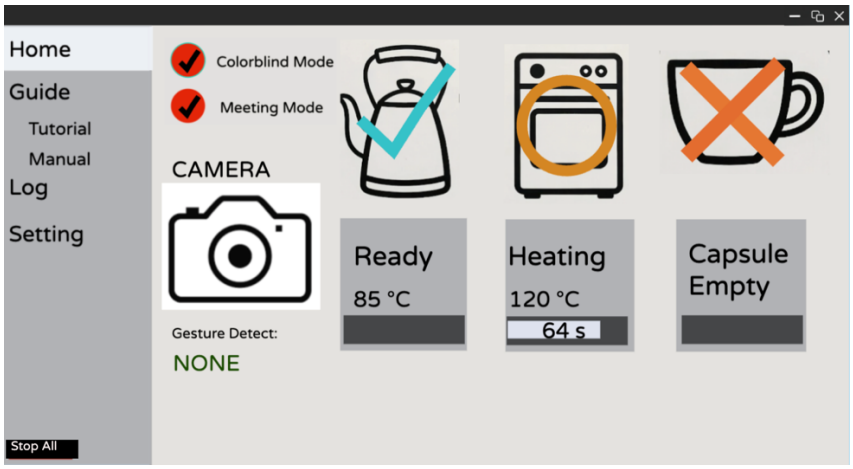


Figure 13. Colorblind mode homepage (kettle, oven, cup icons by grok, 2025)

Colorblind Mode toggled via the toggle button at home page or setting page. It integrated as a **global setting** with **immediate visual transformation for the colorblind users**. Red-greens contrast are replaced with shape + color redundancy:

- Blue ticks **(ready)**
- Orange circle **(processing)**
- Deep orange cross **(error)**

Also, the emergency stop button would change to black background color to enhance the contrast white words. The camera area with show a camera icon instead of red words when the camera is off. This ensures supports for users with color vision deficiency without workflow disruption.

Overall, our system flow shown as below flowchart:

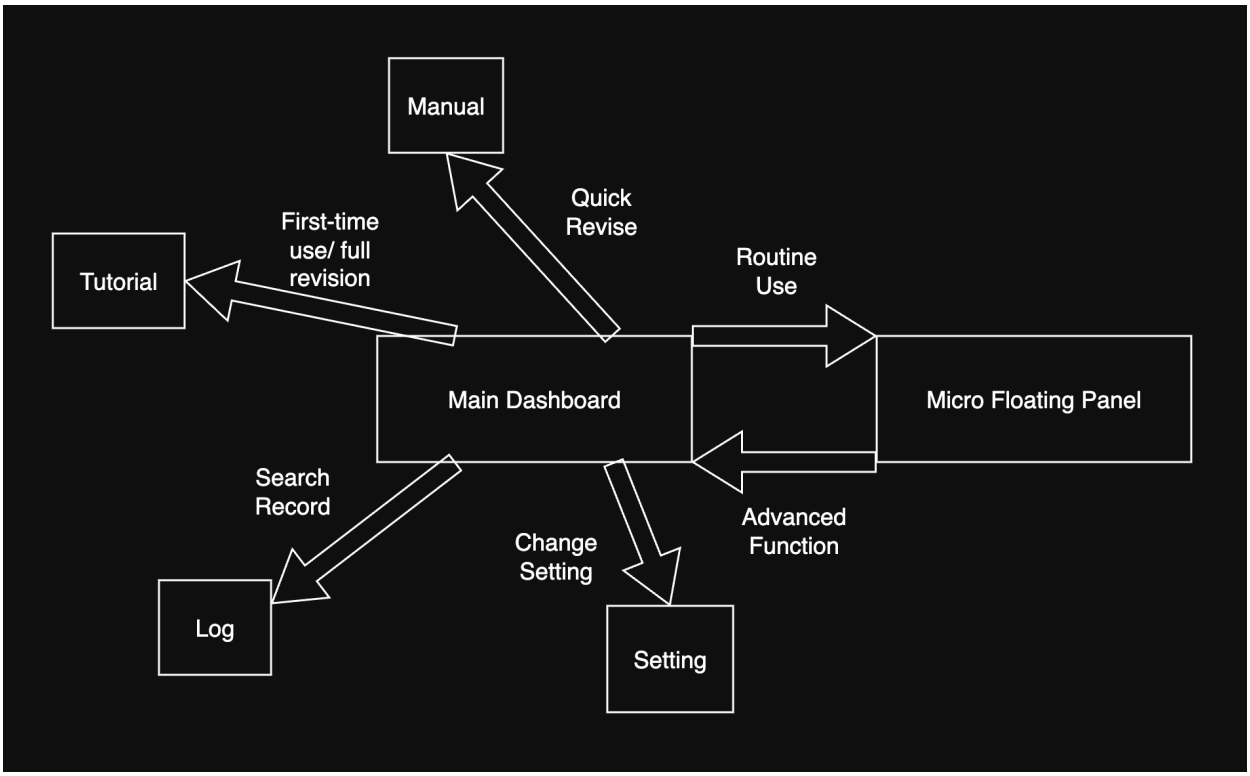


Figure 14. System Flowchart

By the core dashboard as main control panel, users most frequently switch between micro floating panel with it. When routine use, the system serves in micro floating panel. Advanced function can be accessed with if switch back to main dashboard.

From the navigation menu, Manual, Tutorial, Log, Setting pages could be accessed. For first-time users or full revision, they would access to our tutorial page for the interactive tutorial. For users want to quick revise/ specific command search, they can access the scroll-bar manual. More setting and information on system can be accessed via setting page. And the log page provide record of the system and command inputs.

4. Conclusion

Our group integrated the thought of needs of the WFH workers and housewives on multi-tasking of housework and job tasks. Designs of different interface fulfilling the needs for users. We have created few prototype designs, with different pros and cons based on the nature of interfaces. The are some main concerns on the design:

- Difficulties on users learning
- Time for performance
- Emergency shut down
- Convenience for all users

Final interface design is integrated the strengths after examining on different designs. To create a **cohesive, efficient, and user-friendly** for different users' multimodal system, our designs mainly applied Fitts's law, Gutenberg diagram (Z-pattern) and Human Computer Interaction (HCI).

Also, we aimed to remote control the kitchen tasks, gesture + voice driven control with minimal visual distraction should be performed by the system. In terms of **HCI**, our design provides **interactive training** for users, allowing them to familiar with the controls. The micro floating panel allows users have less visual distraction with minimal consumption on screen. For **colorblind users**, we also design alternative mode for accessibility.

Our next step would be using ml5.js to create the implement the design and keep improving it if further problems found. Hopefully the application can good support on target users.

5. Current member responsibility

Student Name:	Student ID:	Contribution	Work done
Wong Siu Lun	58383709	0.9	Interface design, Research studies, design synthesis, Initial Project report, Design report, Design comparison
Tsoi Ho Chun	59380140	0.1	Design comparison

6. Reference

Callahan, J., Hopkins, D., Weiser, M., & Shneiderman, B. (1988). An empirical comparison of pie vs. linear menus. In Proceedings of the SIGCHI conference on human factors in computing systems (CHI ’88) (pp. 95–100). Association for Computing Machinery. https://dl.acm.org/doi/pdf/10.1145/57167.57182

Kholodov, N. (2023, January 18). Z-shaped pattern for reading web content. UX Planet. https://uxplanet.org/z-shaped-pattern-for-reading-web-content-ce1135f92f1c

Unicode Consortium. (2022). Emoji 15.0. Retrieved October 3, 2025, from https://unicode.org/emoji/charts/full-emoji-list.html