GESTURE RECOGNITION SYSTEM

1. **INTRODUCTION**

The Gesture Recognition System for Media Player Control is an innovative project designed to enhance user experience by revolutionizing the way users interact with media player. This project involves a through and comprehensive research work on the scope and the feasibility in order to leverages advance gesture tracking and recognizing technique for enhancing the way the user interacts. The developed system is the result of deep understanding of requirements of users, ensuring that it seamlessly meets their needs of user while providing an intuitive and efficient interacting medium. Each UI component aims to address and facilitate the nuanced needs of the user to make the final system as much user centric as possible. The project has been evolved through continued refinement phases, each was committed to further enhancing the functionality and usability of our system. The document provides different low fidelity prototypes developed during each phase to show the evolution of the existing system. This document will also provide the valuable insights into various rationales and techniques behind each UI design component and the overall evolution of the system to the existing system by providing the motive behind each and every modification.

1. **LOW FIDELITY PROTOTYPES**

This section will provide the detail of various low fidelity prototypes that were developed in order to achieve goals of the project. The reason for using the low fidelity prototypes is that these prototypes are simple and rough representations of the final product, focusing on essential features and functionality without getting bogged down in details. Thus, Low-fidelity prototypes are instrumental in examining diverse design options, honing concepts, and ultimately achieving the best solution for user experience. Each developed prototype for this project has certain advantages and disadvantages associated with it, which are also described in this section. These advantages and disadvantages acted as rational behind modification of the given prototype to a better prototype. For this project 3 different type of prototypes were used :

* Sketch
* Storyboard
* Wireframe

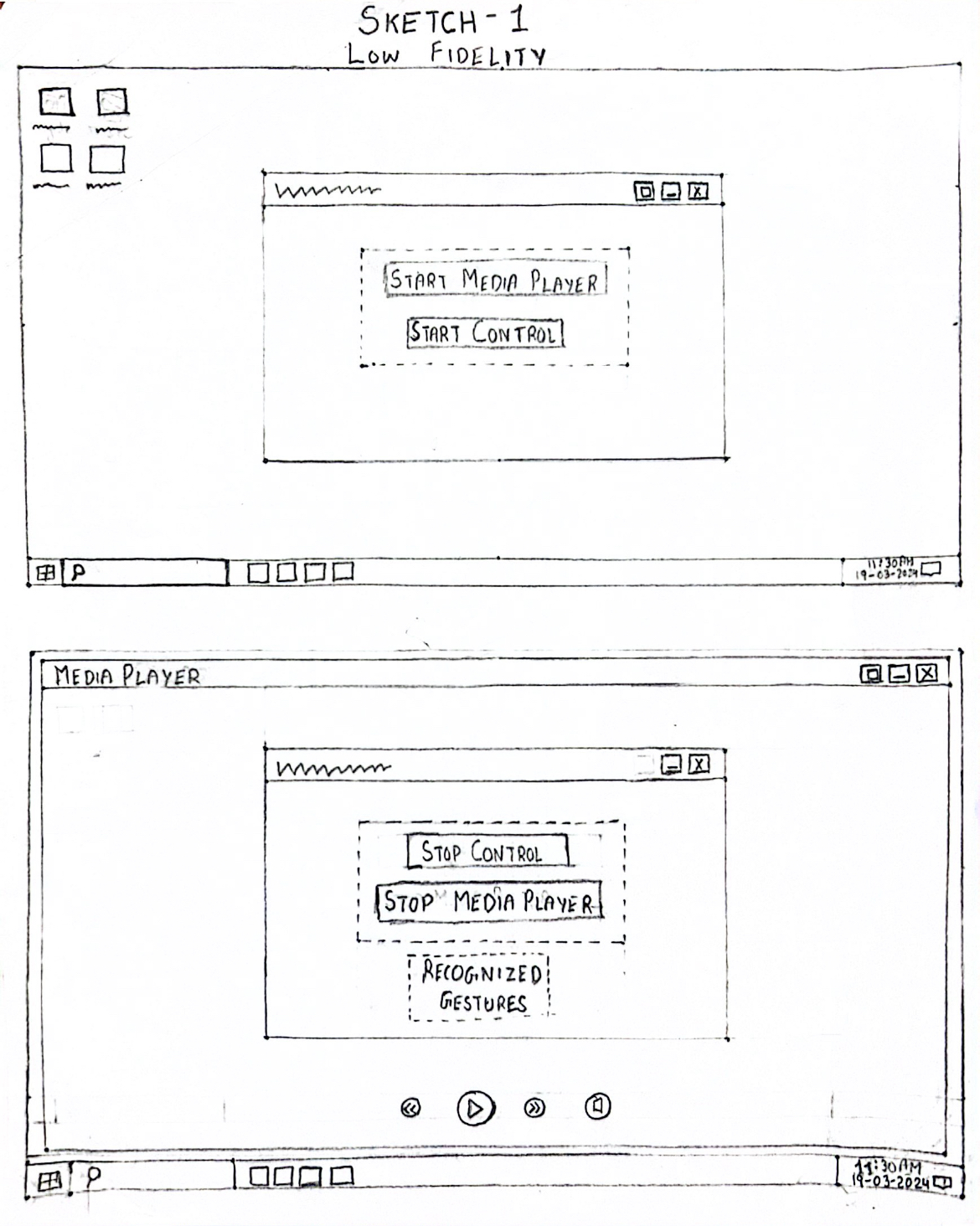
2.1. **Sketch**

Sketches are commonly generated during the initial phases of the design process to swiftly and casually explore concepts, ideas, and potential solutions for a UI design project. During the development phase of the User Interface for the hand gesture recognition project, multiple sketches were developed and put forward to address the user needs in the best possible way. Each new sketch isd esigned to address the shortcomings of the previous sketch prototype.

For this project specifically , 3 prototypes were designed and the 3rd one was chosen to be implemented as the most optimal User Interface to the given project. Each of these sketches are described as below:

**2.1.1 Sketch 1**

Sketch 1 illustrates the arrangement and elements of two windows within a user interface design.



* Window 1: This window is displayed to user when user opens the application.

Start Media Player Button: This button initiates the media player application, allowing the user to begin playing media content.

Start Control Button : Clicking this button activates the gesture control feature triggering the AI model to execute, enabling the user to interact with the media player using gestures.

* Window 2:

Stop Control Button: This button halts the gesture control functionality, effectively ending the user's ability to interact with the media player via gestures.

Stop Media Player Button: Clicking this button terminates the media player application, stopping the playback of media content.

Recognized Gesture display: This component displays information related to recognized gestures, providing feedback to the user about the system's interpretation of current gesture.

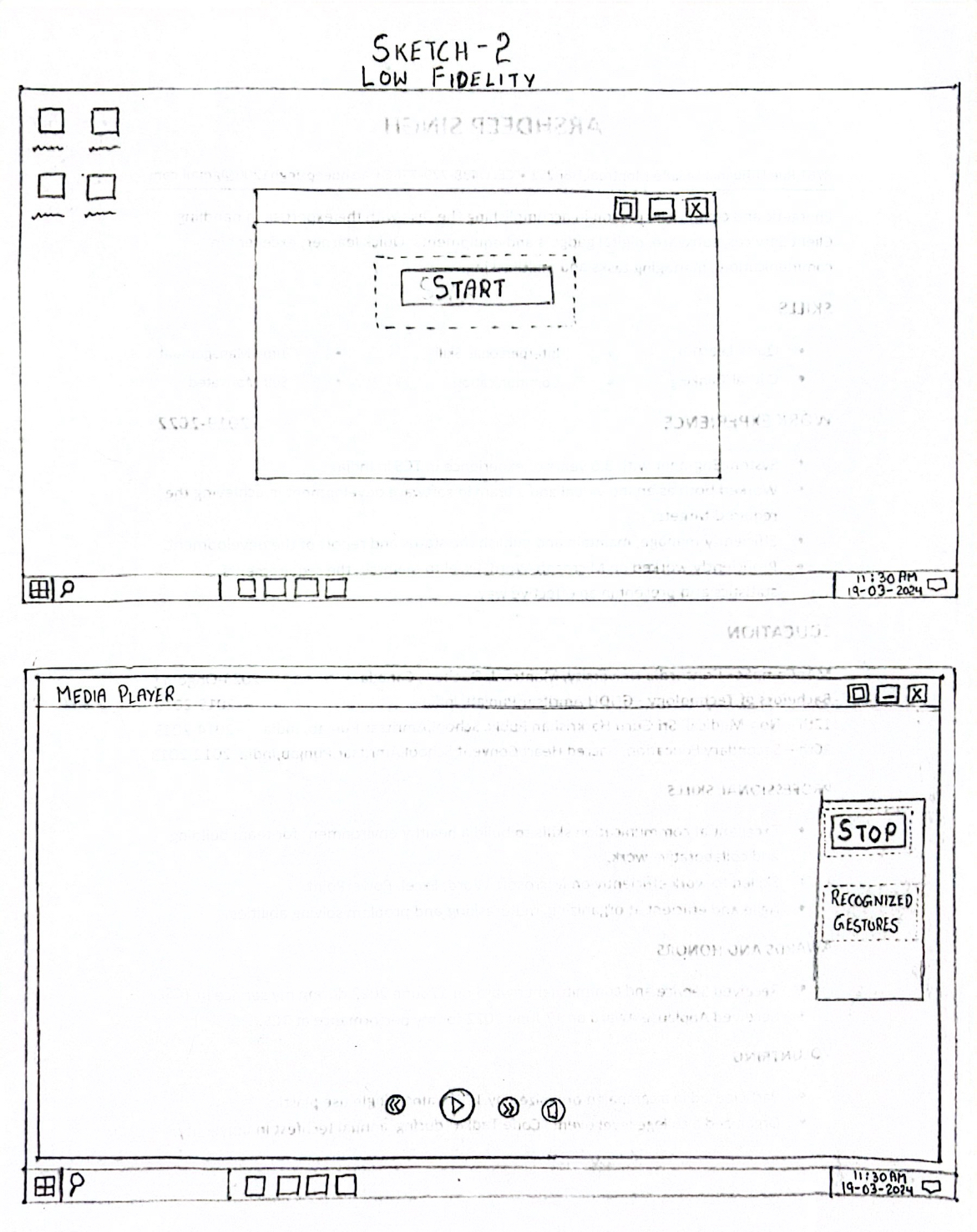
**Drawbacks**

* **Hindering User Visual Area due to Big Window Size:** The large size of the window required to accommodate all functions can degrade the user's visual experience. With a significant portion of the screen occupied by UI elements, such as buttons and Recognized Gesture display, the media player may be obscured. This limitation could hinder the user's ability to fully appreciate or engage with the media content, resulting in a less immersive experience.
* **Excessive Text and Unnecessary Functions:** The presence of excessive text and unnecessary functions, such as starting the media player and control separately, can lead to user confusion and frustration. The need to perform multiple steps, such as initiating the media player and control functions separately, adds unnecessary complexity to the user interaction process.

2.1.2 **Sketch 2**

Sketch 2 introduces a more streamlined approach to user interaction by consolidating multiple functionalities into a single button located within Window 1. "Start" button, serves as a dual purpose of initiating both the media player and gesture control features, simplifies the user experience significantly.

Window 2 is displayed as a vertical layout window on the right side to the media player to reduce the overshadowing of the content of the media player. Here, window has a "Stop" button, enabling user to cease both gesture control and media player operations simultaneously. Moreover, a Recognized Gesture display is included to provide immediate feedback on the system's interpretation of user gestures in real-time. This layout effectively utilizes screen space by organizing related functions in a concise and easily accessible manner, thereby improving usability and ensuring seamless navigation and control over the system's operations.



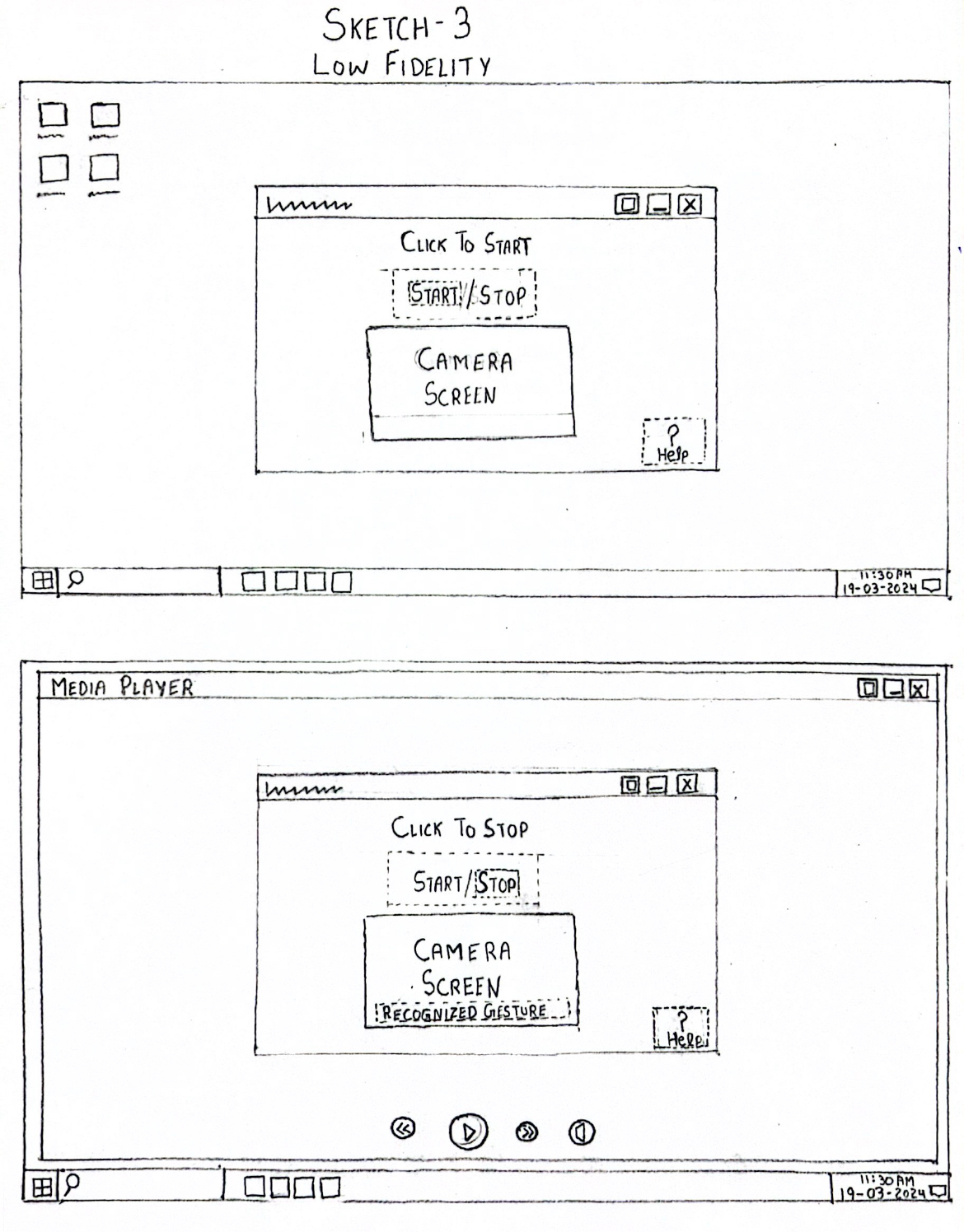
**Drawbacks**

* **Limited space for UI in the 2nd window:** The design of Sketch 2 presents a constraint in terms of available space within the 2nd window. Users may find it challenging to access or interact with certain elements due to the confined space, leading to a less optimal user experience.
* **No visual feedback to indicate how gestures are recognized:** There is still an absence of visual feedback to inform users about how their gestures are recognized by the system. Without this feedback mechanism, users may feel uncertain or disconnected from the interaction process, as they are unable to visually perceive whether their gestures are being accurately interpreted by the system. This lack of feedback could lead to user confusion, detracting from the effectiveness of the gesture control functionality.
* **Need to design two different windows with completely different formats:** Sketch 2 requires the design and implementation of two distinct windows with disparate formats, each serving different purposes. This introduces complexity in the design process, as designers must ensure consistency in user interface elements, visual styles, and interaction patterns across both windows. Managing these differences effectively while maintaining a cohesive user experience can pose challenges and may result in inconsistencies or usability issues.
* **Fixed window in the 2nd window consumes screen area:** The 2nd window in Sketch 2 is designed to remain fixed on the screen until gesture control is stopped, occupying valuable screen area throughout the interaction session. This fixed window may obstruct content or UI elements in media player, degrading overall users experience.

2.1.3 **Sketch 3**

Sketch 3 represents the finalized design chosen for the project, incorporating several key components to enhance user interaction and usability:

* **Single window design:** The user interface adopts a single window layout, promoting simplicity and ease of use. Initially, upon launching the application, a "Start" button is prominently displayed, allowing users to initiate gesture control functionality. As the gesture control process commences, the "Start" button dynamically transitions into a "Stop" button, enabling users to cease the application when needed. This streamlined approach minimizes clutter and optimizes space within the window, ensuring a clean and intuitive user experience.
* **Camera Screen:** The central component of the interface is the camera screen, which provides users with visual feedback on how their gestures are being recognized by the system. This real-time display allows users to observe their hand movements and gestures as captured by the camera, facilitating interaction and engagement with the system. Additionally, the camera screen also showcases the actions associated with recognized gestures, offering users immediate feedback on the outcomes of their interactions.
* **Help Button:** Positioned at the bottom right corner of the window, the "Help" button serves as a valuable resource for users seeking guidance on the system's functionality. Upon clicking the button, a new window overlays the existing interface, presenting users with comprehensive information about the gestures recognized by the system and the corresponding actions associated with each gesture. This intuitive design ensures that users can easily access assistance and support whenever needed, promoting a seamless and user-friendly experience.



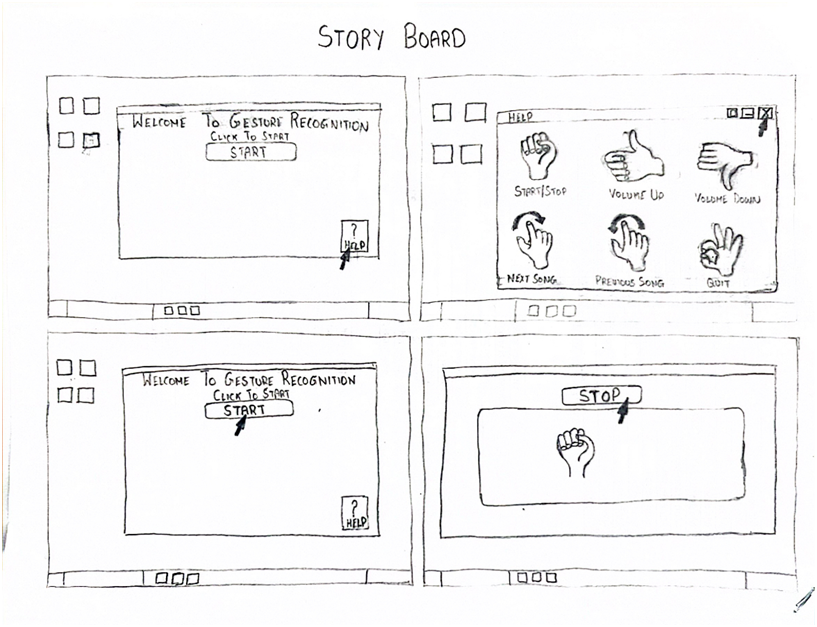
**Advantages:** Sketch 3 offers several advantages that enhance user experience and usability:

1. **Simplified Learning with Visual Feedback**: Through the camera screen, Sketch 3 delivers visual feedback to users, fostering quicker and smoother learning, particularly for newcomers. This real-time display grants users insight into how the system interprets their gestures, facilitating an intuitive grasp of the interaction process.
2. **Enhanced Understanding with Gesture Recognition Visualization**: By overlaying frames over fingers on the camera screen, Sketch 3 offers users a deeper understanding of gesture recognition. This feature showcases the intricate movements of the hand and the corresponding recognition frames, bolstering transparency and bolstering user confidence in the system's accuracy and reliability.
3. **Optimized Window Management**: Unlike previous iterations, Sketch 3 allows users to minimize the main window even during gesture control, preserving valuable screen space. This functionality ensures uninterrupted access to other applications or content, bolstering user flexibility and multitasking capabilities.
4. **Efficiency Through Streamlined Interface**: With a concise presentation of essential information and functions, Sketch 3 minimizes distractions. By prioritizing clarity and simplicity, the interface enables users to focus on tasks without unnecessary clutter, fostering efficiency and satisfaction in user interactions.

2.2 **Storyboard**

A storyboard in UI designing prototype serving as a graphical depiction of the user's progression within an application or system. It commonly comprises a sequence of illustrated frames or sketches, showcasing the series of interactions and events encountered by the user as they navigate through the interface.

Following story board prototype will describe the story board of how user will interact with the final developed system for gesture control media player:



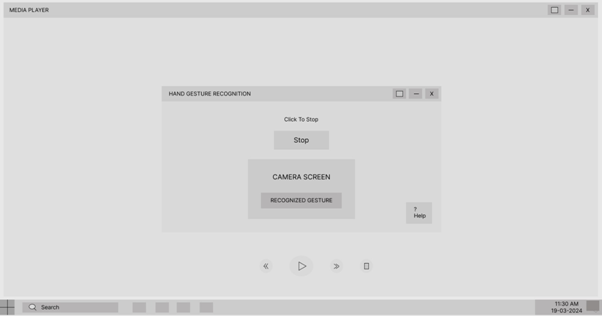
1. **User Opens Application and Accesses Help**: The user initiates the application and clicks on the help button to seek guidance on using gestures. This action triggers the opening of a new window, which contains information about various gestures and their associated actions. The window serves as a guide for the user to understand how to control the media player using gestures.
2. **Exploring Gesture Guide and Returning to Start Window:** After reviewing the information in the Help window, the user closes it and returns to the start window. This step indicates that the user has familiarized themselves with the available gestures and is now ready to proceed with controlling the media player.
3. **Starting Media Player and Gesture Control**: The user clicks on the Start button to initiate both the media player and gesture control functionalities. The camera screen opens within the same window, allowing the user to view the live video captured by the camera. Additionally, the Stop button becomes visible, indicating that the media player and gesture control are now active.
4. **Interacting with the Camera Screen:** Within the camera screen, the user can see themselves in real-time and observe how the camera keeps them in the frame even if they move. The system displays a frame over the user's hands to provide visual feedback, indicating that gestures are being recognized. Once the user has finished using the system, they can click on the Stop button to halt the media player and gesture control functionalities.

2.3 **Wireframe**

Wireframe are created during the early stages of the design process to outline the structure, layout, and functionality of a digital product or application. Low fidelity wireframes define layout and structure, facilitating early feedback and validation of concepts.

The following low-fidelity wireframe provided depicts the final user interface design based on Sketch 3.





It features the following key components:

**Single Window Design:** The wireframe adopts a single-window approach, simplifying the user interface by consolidating all functionalities within one interface. Initially, the window displays a "Start" button, which serves as the entry point for activating both the media player and gesture control.

**Camera Screen:** Within the window, there is a dedicated area referred to as the "Camera Screen." This screen provides users with live video feedback captured by the camera. Users can observe their hand movements in real-time, with the system ensuring that user remains within the frame, even during motion.

**Help Button:** Positioned at the bottom right corner of the window, the "Help" button allows users to access guidance on the various gestures recognized by the system and their associated actions. Clicking this button opens a new window overlaying the existing interface, providing users with comprehensive instructions and explanations.

1. **UI COMPONENTS**

This section will highlight all the UI components that were implemented in the finalized project. Section will also justify the rationale behind the choice of UI component and the way this is designed to keep the system user centric as much as possible.

* **Camera Motion Detector**

Camera Motion detection is implemented as a crucial UI component to improve the robustness of system and enhance user experience. Project uses an AI-based model for tracking the user's motion and centralizing the camera. Camera Motion Detector component ensures that the user remains properly framed within the camera's view, enhancing the overall user experience and usability of the system. This UI component seamlessly interacts with the gesture recognition system by dynamically adjusting the camera frame to the user's movements. This ensures precise capture and interpretation of the user's gestures and actions, thereby enabling effortless control over the media player.

**Rational**: Rational for implementing this UI component is to ensure the system is robust, even if the user is in consistent motion. This will allow user to control the music player even while performing other activities

* **Hand tracking**

Hand posture tracking is a technique in which the posture of the hand is represented in the form of a frame. Hand posture tracking is a method used to monitor and evaluate the real-time position and alignment of an individual's hand. It involves capturing the hand's movements and representing them visually as a frame or image. This frame provides a snapshot of the hand's configuration at a specific moment, including details such as finger positions, joint angles, and the overall shape of the hand. By continuously updating this frame as the hand moves, hand posture tracking assists the system in recognizing and interpreting various hand gestures and actions performed by the user.

**Rational:** The purpose of overlaying a frame over the user's hand is to provide a visual aid that enhances their comprehension of the system. This frame offers a clear representation of the hand's boundaries and positioning within the camera's view. By delineating the contours of the hand, including its fingers and palm, the frame allows users to visualize how their hand movements are tracked and interpreted by the system. Moreover, the frame functions as a feedback mechanism, indicating the accuracy of the system in detecting and recognizing the user's hand gestures.

* **User Support**

The "Help" button, located in the bottom right corner of the application, functions as a UI element intended to offer users assistance and direction regarding the hand gestures and associated actions. Upon clicking this button, a window or dialog box is triggered, presenting users with detailed information about the recognized hand gestures and their corresponding functionalities.

**Rational:** The rationale behind incorporating this UI component is to enhance user experience and facilitate ease of use. By providing users with access to a detailed guide on hand gestures and their functionalities, the system empowers users to interact more effectively with the application. This feature serves as a valuable resource for users who may be unfamiliar with the specific gestures supported by the system or who require clarification on how to perform certain actions.

* **Buttons**

Project implements a "Start" button as the primary interaction point for initiating camera usage, running the gesture recognition model, and hand framing model The integration of a "Start" button as the central point of interaction to activate the camera, launch the gesture recognition model, and frame the hand exemplifies a minimalist design approach aimed at streamlining user engagement. This strategy emphasizes simplicity by providing users with a straightforward and clear method to initiate camera usage and gesture recognition. With the presentation of a singular, dedicated button, the user interface maintains an uncluttered and intuitive layout, reducing cognitive load and potential confusion. Users are granted control over the activation process, allowing them to commence camera operation and gesture tracking according to their preferences, in line with principles of user-centered design. Furthermore, the "Start" button effectively directs user attention to the specific task at hand, promoting focused interaction devoid of unnecessary distractions. Its user-driven functionality empowers users to engage with the system on their own terms, fostering a sense of agency. Additionally, the "Start" button serves as a feedback mechanism, delivering immediate confirmation to users upon successful activation of the camera and gesture recognition features, thereby bolstering user confidence and overall usability.

Stop button is used to close the camera and running model simultaneously. The addition of a "Stop" button within the user interface serves to seamlessly halt camera operation and terminate the execution of the gesture recognition model simultaneously. This design decision emphasizes user convenience and accessibility by offering a straightforward method to disengage from the system's functions with just a single click. By integrating a dedicated button for stopping operations, users benefit from a clear and intuitive means to conclude their interaction with the system, thereby enhancing the overall user experience. This streamlined approach reduces user effort and cognitive load, enabling users to deactivate the camera and model quickly and easily without navigating complex menus or interfaces. Moreover, the presence of the "Stop" button instills a sense of control and confidence among users, empowering them to manage their engagement with the system at their preferred pace. Ultimately, the implementation of the "Stop" button as part of the user interface promotes usability and user satisfaction by facilitating smooth and efficient interaction with the system.

* **Gesture tracking using bubbles**

Gesture tracking using bubbles is a visual UI feature implemented to provide users with real-time feedback regarding the trajectory of their hand movements while performing gestures. By displaying bubbles or markers that follow the path of the user's hand across the screen, this UI component enhances user interaction by offering a clear and intuitive representation of gesture tracking.

Furthermore, gesture tracking using bubbles contributes to user engagement and satisfaction by providing a visually engaging and interactive element within the interface. Users can easily correlate their hand movements with the corresponding bubbles on the screen, fostering a sense of connection and control over the interaction process. This enhances the overall usability of the system by empowering users to effectively navigate and manipulate the interface through natural hand gestures.

**Rational:** The rational behind integrating gesture tracking using bubbles lies in its ability to simplify the user's job by visually indicating the movement of their hand during gesture execution. This visual feedback mechanism serves to guide users and help them understand how their hand movements are being tracked and interpreted by the system. By seeing the bubbles move in sync with their hand, users gain immediate insight into the system's recognition of their gestures, facilitating a smoother and more intuitive interaction experience.

