

TSUMU - Tack Stack User-interface for Music Utilities

the Brick*

Department of Information Science, Nara Institute of Science and Technology

1 Introduction

TSUMU (Tack Stack User-interface for Music Utilities) is novel interface for media players using a camera and hand actions.

The music players are often used with other applications at the same time. However, conventional interfaces for music players use a mouse and a keyboard. For example, how do you do if you want to stop the music on operating a word processor? TSUMU makes it possible to control the player with more simple action than conventional interfaces.

2 How to Use TSUMU

There are two modes in TSUMU. One is “List Mode”, in which playlists are edited. Another is “Play Mode”, in which songs are controlled. Operating instructions of each mode are described below.

2.1 List Mode

In List Mode, playlists are edited. The playlists are depicted as blocks, and the blocks are “grabbed” and dragged by detecting the “grabbing action”. Figure 1 is a screenshot in List Mode. A cursor is drawn at a point of a gravity of the hand region.

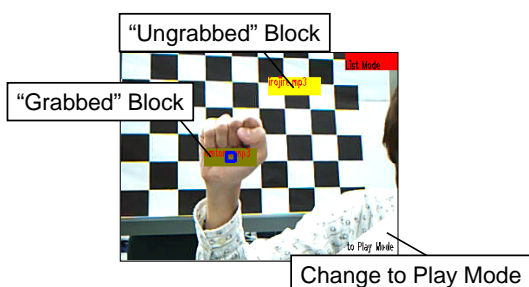


Figure 1: A screenshot in List Mode. Each block corresponds to a playlist. The block can be “grabbed” by a hand of the user, and the mode is changed by moving the block onto the button.

2.1.1 Loading Music Files

A current directory is scanned to find music files when the player is launched. Please note that the name of files should not include a space ‘ ’. The music files which is detected are displayed on the command prompt window. And the files are loaded as discrete playlists.

2.1.2 Available Hand Actions

Two hand actions, “Fist” and “Hand” are available in List Mode. Fist is used to grab blocks, and the blocks are released when the hand turns the action into Hand. Figure 2 shows pictures of the available hand actions.

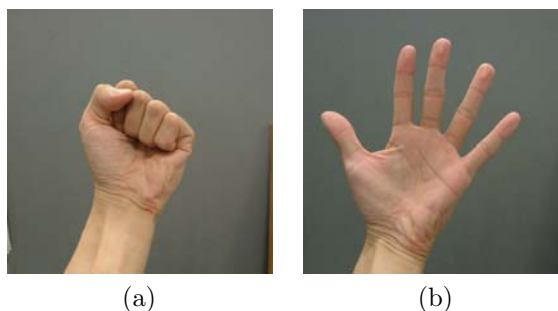


Figure 2: Available hand actions. (a) “Fist”, the block is “grabbed” when the hand shapes “Fist”. (b) “Hand”, the block is “ungrabbed” when the hand shapes it.

2.1.3 Editing Playlist

Editing playlist is implemented in stacking the block on other blocks. In Figure 3, There are two discrete playlists as blocks. These blocks are merged due to stacking a block on the other one. The block is dragged to “Change to Play Mode” button to play music according to the playlist on a music player.

2.2 Play Mode

A music player is launched with the “dragged” playlist and the player is operated in Play Mode. Four operations, “Play”, “Stop”, “Move to Next

*Yuki Uranishi, Department of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

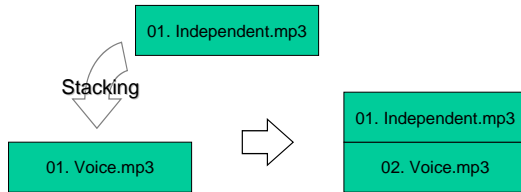


Figure 3: An example of merging two playlists into one. Two blocks are merged when a block is stacked on other blocks.

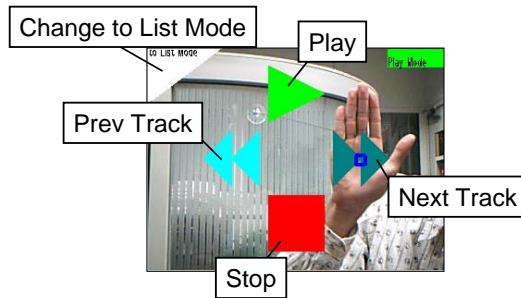


Figure 4: A screenshot in Play Mode. There are five buttons on the screen. Four buttons are used to operate the player and the other one is used to change the mode into List Mode.

Track” and “Move to Previous Track” are available. Figure 4 shows a screenshot in Play Mode. The player is controlled using the hand action.

3 Algorithm

Algorithms to implement above functions are described in this section.

3.1 Image Acquisition and Display

Open Source Computer Vision Library [1] is used for acquiring and displaying images. A camera is set on the ordinary arrangement.

3.2 Hand Detection

An algorithm to detect a hand region from images is described below. At first, Hand color regions are extracted using the color of each pixel. Pixels which have a color of a palm on Hue-Saturation plane are extracted, and the continuous pixels are labeled. Then, the largest region in the image is detected as a hand region.

3.3 Hand Motion Analysis

A ratio of the size of the actual hand region to the size of the pseudo convex hull of the hand re-

gion. Figure 5 shows the convex hull of “Fist” and “Hand”.

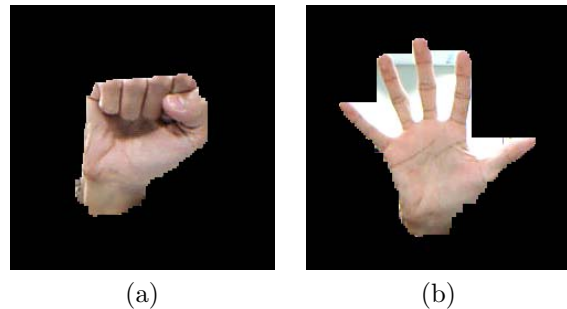


Figure 5: A convex hull of each hand action. (a) “Fist”, the block is “grabbed” when the hand shapes “Fist”. (b) “Hand”, the block is “ungrabbed” when the hand shapes it.

When the hand shapes “Fist”, a ratio of the size of actual hand region to the size of the pseudo convex hull of the region is approximate 1.0. On the other hand, when it shapes “Hand” then the ratio of the size of the actual region to the size of the pseudo convex hull is smaller. The hand motion is analyzed using the ratio.

3.4 Media Control

Music files are controlled using Media Control Interface (MCI)[2].

4 Summary

We proposed the novel interface using hand actions for editing playlist and controlling music players. This interface, TSUMU, makes it possible to control the player with more simple action than conventional interfaces.

Future work will aim at analyzing further hand action, such as waving a hand. In addition, we are also planning to implement volume control using a hand action.

Acknowledgements

The author of this interface wish to acknowledge the support supplied by Dr. Damien Douxchamps during the development of the interface.

References

- [1] Open Source Computer Vision Library (OpenCV) <http://www.intel.com/technology/computing/opencv/>
- [2] MSDN Home Page <http://msdn.microsoft.com/>