Real Time Virtual Piano Using Hand Gesture

Prof. Varsha K.Patl, Sagar Mahindrakar (sagarmahindrakar7395@gmail.com), Pratik Gaikwad

Department of Electronics and Telecommunication Engineering, All India Shri Shivaji Memorial Society's Institute of Information Technology, Pune, Maharashtra, India

ABSTRACT

Abstract— It is said that there is a need of portable music devices in order to save cost, provide a similar level of satisfaction to the user as playing on a physical instrument set does. Knowing this, why should everyone have to own a costly instrument set to start playing? For our final project, I have the intention of implementing a gesture-controlled virtual piano set on computer using Image Processing to provide a simple and fun way for people who wish to learn or simply enjoy playing the piano. More specifically, the user's hand motion would be tracked by a camera system. The piano set uses different coloured keys to ease the motion tracking. The motion tracking system will distinguish between hitting motions in different areas in front of the user and emulate sound of distinct notes based on hitting area. At a basic level, the implementation would consist of a limited number of piano keys and basic playback of sounds of the notes the user plays. Depending on available time, other features could be added. As an example, a learning mode in which the user loads preprocessed songs and is guided which keys to play to get the corresponding output. Another idea is to provide the user with an increased number of piano set notes for increased sound options. We expect that by successfully implementing the basic functionality, the system will be robust enough so that a wide range of additional features could be completed easily

Key Words: Keywords—image processing, live object detecting in matlab, live detecting object in OpenCV, colour detection in matalb.

I. INTRODUCTION

Piano is a large bulky instrument which is not very affordable to many due to its high cost. Moreover not all institutions can afford a large number of instruments due to space and cost constraints. The real-time virtual piano player is portable and cheap. The system captures the video of the hands playing the piano and creates suitable musical notes corresponding to the keys played by the user.

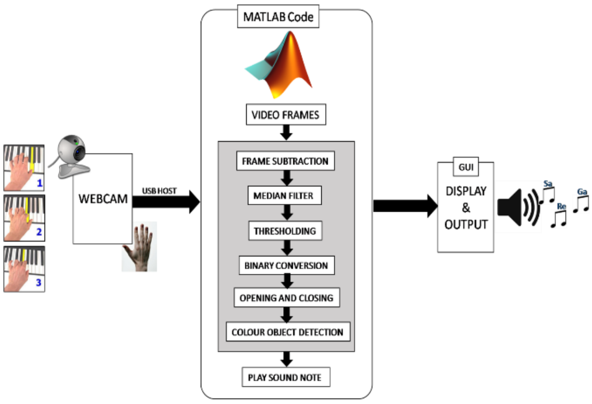

Fig-1 Block Diagram

II. BLOCK DIAGRAM DISCRIPTION

A web-camera is mounted vertically on a mount. It captures the area over which the piano keys are to be played. This video stream is converted to video frames. The input image frame then undergoes series of extraction and segmentation operations. The output result of extraction and segmentation is compared with the previous image frame using image subtraction. This enables us to track the blue colour object in the video. We can also configure the code to be able to track different colours like red or green. When the colour object is tracked in the particular segment, the corresponding sound note is played. These sound notes are generated using mathematical equations. We can play Sa, Re, Ga, Ma, Pa, Dha, Ni using this approach.

2.1 Webcam

Webcams are low in cost and their flexibility are very high. The webcam captures the video. We in our project are using a face2face C12.0 webcam.

Fig: 2.2.1.iBallC-12 720p

2.2 MATLAB Code

The MALTAB code is designed to acquire video frames using the Image Acquisition Toolbox. The Logitech C-270 720p web-camera is interfaced with the code. This camera is vertically mounted on an aluminium stand. When initiated by the GUI, the webcam starts to acquire video frames and perform operations so as to obtain the desired result.

We use a series of operations on the video frames as listed below:

- Frame Subtraction
- Median Filtering
- Thresholding
- Binary Conversion
- Opening and Closing
- Colour Object Detection

2.3 Steps in Detail

The above steps have been explained in detail below:

A) Frame Subtraction

Image subtraction or pixel subtraction is a method in wich the digital value of one pixel or the complete image is differentiated from another image. The changes between two sequential image is done by frame subtraction.

This method is use to recognise if there is moving object present. For detecting moving parts in videos from fixed cameras the background subtraction method are mostly used.

We use this technique to be able to track our blue colour object. This approach has been shown in figure 2.3.1

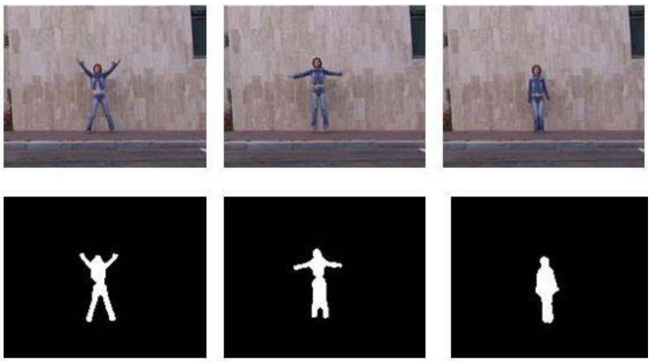

Fig: 2.3.1 Frame Subtraction

B) Median Filter

This filter is a digital nonlinear filter. It is a technique which to removes the noise. It is a pre-processing method and mainly used to support the results of later processing.

The principal of the median filter is that it run through the signal entry by entry, separating each entry with the median of adjacent entries. The behaviour of neighbours is called the "window", which moves, entry by entry, over the complete signal.

For 2D signals such as images, complex window patterns are possible (such as "box" or "cross" patterns). The consideration is that if the window presents are odd in number of entries, then the median is easy to define: After all the entries the middle value in the window are sorted numerically.

As seen from the figure below, we can note the manner in which noise is removed.

C) Grey Scale Conversion

Grayscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Grey scale is used for following reasons:

Signal to noise ratio: Colour information doesn't help us identify important features like edges, crests of the license plate.

Complexity of the code: To want to find the edges based on luminance and chrominance add to an additional work (debugging, additional pain in supporting the software, etc.).Using grey scale images helps to simplify these tasks.

Speed: With modern computers, and with parallel programming, it's possible to perform simple pixel-by-pixel processing of a megapixel image in milliseconds. But recognition segmentation and other tasks can take much longer than that. Hence using gray scale will help to reduce the computational power and hence increase the processing speed.

Fig: 2.3.2 Conversion to Grayscale

D) Binary Conversion

A binary image is a digital image that has only two possible values for each pixel. Typically, the two colours used for a binary image are black and white, though any two colours can be used. The colour used for the object in the image is the foreground colour while the rest of the image is the background colour.

Binary images are also called bi-level or two-level. This means that each pixel is stored as a single bit 0 or 1. The names black and white monochrome monochromatic are often used for this concept, but may also designate any images that have only one sample per pixel, such as grayscale images.

We are using 640x480 video frames. A 640×480 image requires 37.5kib of storage and is stored in bitmap format.

Because of this small size of image, it is easier to perform functions on these. From a grayscale image, binary conversion can be done using a suitable threshold to create binary images.

E) Opening and Closing

Erosion is mainly used for eliminating the irrelevant details of an image.

Dilation is mainly used to enclose holes in a single region, reduce the gaps between different regions and to fill the intrusions into boundaries of a region.

Opening is defined as erosion followed by a dilation using the same structuring element for both operations. The opening operator therefore requires two inputs: an image to be opened, and a structuring element. Grey level opening consists simply of grey level erosion followed by a grey level dilation.

Closing is opening performed in reverse. It is defined simply as dilation followed by erosion using the same structuring element for both operations. Grey level closing consists straightforwardly of a grey level dilation followed by grey level erosion.

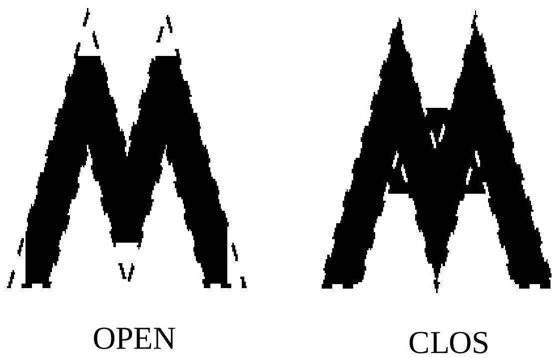

Fig: 2.3.3 Morphological Operation – Opening & Closing

F) Colour Object Tracking

Colour tracking is the ability to take an image, isolate a particular colour and extract information about the location of a region of that image that contains just that colour.

Fig-2.2.4 median filter

Colour object tracking involve an initial contour initialized from the previous frame to its new position in the current frame. It mainly involves colours which use the Red, Green, Blue colours in proportion.

We have used the blue colour object detection algorithm but it can also be modified for different colours such as red and green as shown.

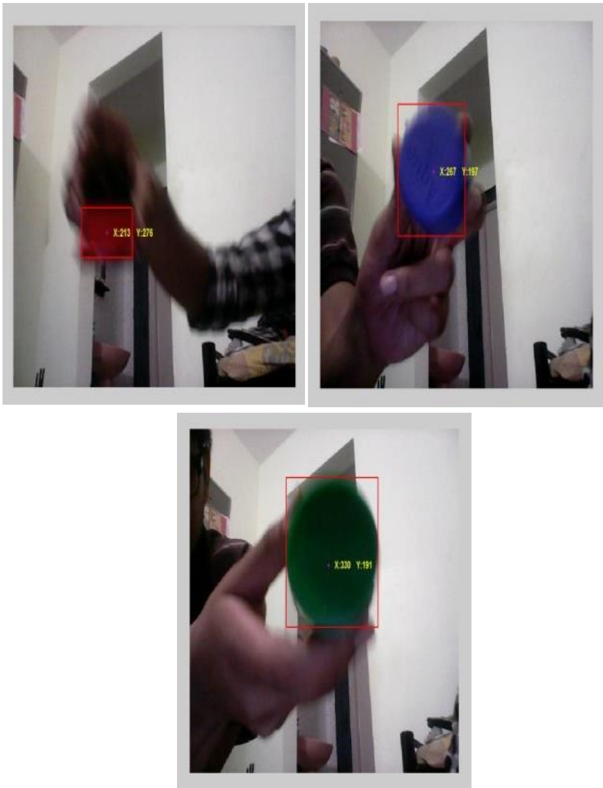

Fig: 2.2.4 Colour Object Tracking

To generate sounds, we are using frequency equations. We could have used recorded sound notes but this decreases the speed of the code. To enhance this, we have used frequency equations which are used with different frequencies. Each of Sa, Re, Ga, Ma, Pa, Dha, Ni has different frequencies.

The following equations enable us to produce the sound notes when using a sampling rate of 44100 Hz.

Table 2.1. Frequency Equations for Sound

SIN y = sin(2.*pi.*f.*t)
SQUARE y = square(2.*pi.*f.*t)
SAW TOOTH y = sawtooth(2.*pi.*f.*t)

Here,

t = time period which varies from 0 to r. ‘r’ is the key note length.

f = frequency of different sound notes.

The live video and the piano keyset can be viewed by the user on screen along with audio output.

In the ‘natural harmonic scale’, Re, Ga, Ma, Pa, Dha, and Ni are related to Sa as given below:

If Sa is assumed to have a frequency say 240 Hz then Re is 270 Hz, Ga is 300 Hz, ma is 320 Hz, Pa is 360 Hz, Dha is

400 Hz and Ni is 450 Hz. Sa of upper octet is then 480 Hz. The octet can be tabulated as follows:

Scales	
Re	Sa x 9/8
Ga	Sa x 5/4
Ma	Sa x 4/3
Pa	Sa x 3/2
Dha	Sa x 5/3
Ni	Sa x 15/8
Sa	Sa x 2/1

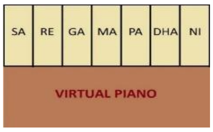

Fig: 2.2.8. Sheet Layout of Piano Keys

2.4 iBall face2face C-12 720p Webcam

We are using a iBall face2face C-12 720p Webcam to capture the video. This camera can be connected to the BBB using a USB cable. It provides video streaming of the resolution 1280x720. It works well with the MATLAB software as well as the BBB (as seen after testing). It provides a frame rate of 30 fps. It provides a still image resolution of 3 megapixels.

Fig-2.2.5 webcam

III. SOFTWARE DESIGN

3.1 MATLAB

MATLAB (MATrix LABoratory) is a multi – paradigm numerical computing environment and fourth-generation programming language. It allows real time video processing. MATLAB contains many inbuilt functions making it user friendly. It allows the user to perform complex processes with ease. The Image Acquisition Toolbox of MATLAB allows plenty of image processing functions to be performed.

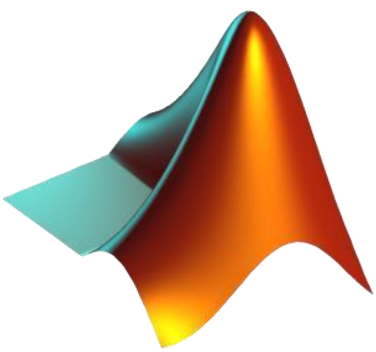

Fig-3.1.1 MATLAB

We are making use of MATLAB on a Windows platform. This enabled us to create a graphical user interface which provides options for the user and also can interact with the system.

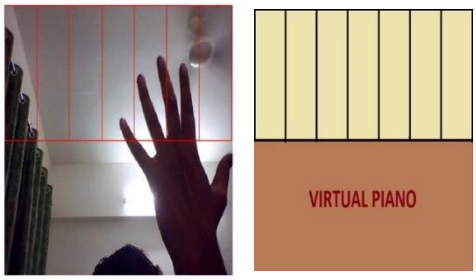

Fig-3.1.2 Calibration

We are using the above template coded in MATLAB for calibration. This template is being used to calibrate the keys of the piano with the code. The screen has been divided into 7 segments to represent each of Sa, Re, Ga, Ma, Pa, Dha, Ni respectively.

Similarly, we have designed a hard copy of the template. We try to match these two as close as possible to calibrate the piano correctly.

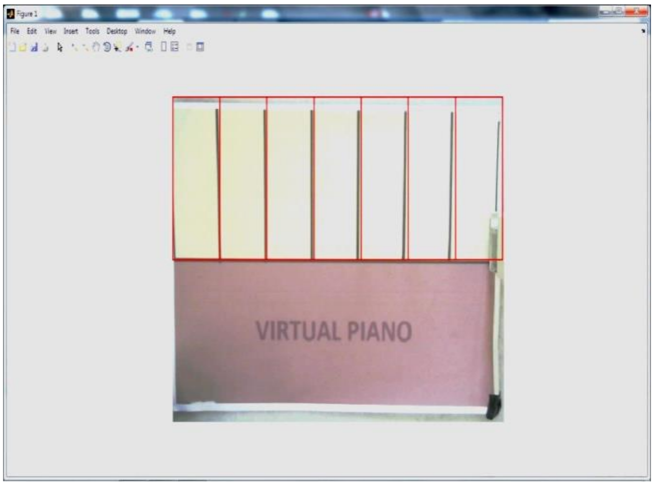

Fig-3.1.3 Calibration through GUI

The camera mount is made of aluminium. We had thought of using wood for the same but we used aluminium to make the setup light weight and sturdy.

As we can see from the figure, the webcam is vertically mounted on the aluminium setup.

The camera is set at a position such that it captures the template of the piano and can match it with the code.

Fig-3.1.4. Camera mount

We are using blue colour object tracking.

When the blue object is detected in a particular segment, the corresponding sound note is played. For example, in the figure, ‘dha’ will be played.

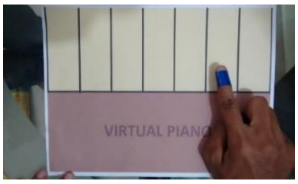

Fig-3.1.5 Piano on Paper

3.2Virtual Piano GUI

Fig-3.2. Virtual Piano GUI

IV. CONCLUSION

We have successfully designed a GUI & created a prototype of a virtual piano. We have successfully interfaced the camera to function as required with the MATLAB software.

We learnt, understood and implemented the core concepts of processing the image. Trying the implementation through the camera & DSP processor in the computer gave us a considerable amount of hands on knowledge of the processor. We used MATLAB’s Image Acquisition Toolbox and the Matlab GUIDE for successfully implementing the code. The GUI, as we desired, is user friendly and less complicated.

This report contains a detailed overview of what is required to bring about the success of the project. The implementation of this idea will enable many to learn or enjoy the pleasure of playing a piano virtually.

This project provides a portable solution in contrast to carrying the large bulky instrument. Through this project, we learnt the fundamentals of image processing and also the functionalities of the MATLAB.

ACKNOWLEDGMENT

It is my great pleasure in expressing sincere and deep gratitude towards my guide Mrs.V.K.PATIL,Assistant Professor Electronics & Telecommunication Engineering Department for her valuable guidance and constant support throughout this work and help to peruse additional studies in adaptive signal processing and MATLAB.

We take this opportunity to thank Head of the Department Mrs.M.P.Sardey and Project coordinator Mrs.V.K.Patil and all staff members ofdepartmentofElectronics &Telecommunication Engineering AISSMS IOIT, Pune, for cooperation provided by them in many ways.

The motivation factor for this work was the inspiration given to me by our honorable principal Dr. P.B.Mane.

Lastly I am thankful to those who have directly or indirectly supported for our work.

V. REFERENCES

[1] Rafael C. Gonzalez (University of Tennessee), Richard E. Woods (MedData Interactive) andSteven L. Eddins (The MathWorks, Inc.), in ‘Digital Image Processing Using MATLAB’ Second Edition, 2009 by Gatesmark, LLC

[2]Alasdair McAndrew, in ‘An Introduction to Digital Image Processing with Matlab, Notes for SCM 2511 Image Processing1’SchoolofComputerScienceand Mathematics, Victoria University of Technology.

[3] Shubham Srivastava, Pratibha Singh, “Real-Time Object Tracking Using Colour Feature”, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-3, Issue-8, pp. 56 – 59, January 2014.