Air Tabla

SM2702 Interdisciplinary Practices

List of disciplines used:
Art (Music), Computer Science: Artificial Intelligence (Computer Vision)

RAINA, Ritvik (55828826)
SHAHUL HAMEED, Shoaib (54871216)

Introduction

For this assignment, we were particular on highlighting the contrast between disciplines. Over the weeks of this course, we came to realize that the interdisciplinary borders have been breached in objects we use every day. So, we chose 2 accessible disciplines - music and computer science, which already have a long-running connection and have been complementing each other for at least the past 3 decades. Furthermore, we were keen to display a feature which will make our work different, which may allow others to learn something or even instill and interest. Our background – Indian culture was a concept which we could portray through our work, and we did just that.

The tabla is a traditional Indian folk instrument which falls under the percussion family. Imitating the drums, it is essentially a twin drum set which generally accompanies vocals and a band of instruments and is played using the fingers and palms. The fingers strike the centre of either drums to produce sounds in various scales, notes and tempo as well (Mishra, 2015).

Tabla being played (picture from Old Delhi Music)

The sound which is produced by the tabla is native and unique to understanding the essence which Indian music has. Listeners generally find the sound catchy, upbeat, soothing, exciting and even sometimes calming. It appeals to both folk and modern Indian music and is still widely being used for different performances, musicians, and even in the Indian film industry – Bollywood. When noticing the transition from folk to modern music, there has definitely been digitalization of this instrument. There are even various digital, easy-to-use versions of the tabla for sale.

The Air Tabla

While looking at the different avenues through which digitalization has taken place, it was evident that this was an interdisciplinary revolution. We identified the possibility of implementing it in a different manner where we use the same disciplines – music and computer science but using a more in-depth concept of computer vision to create the Air Tabla.

The Air Tabla is a Python program which runs to produce a program which uses the devices camera and the screen as well. The screen shows an image of a tabla set embedded onto the camera's display.

Air Tabla's screen

Computer vision by itself is considered interdisciplinary by today's standards. It is based on artificial intelligence to recognize natural images, interpret and respond to it. The premise of the program works in a manner to identify the movement along the images of the tabla (with a red object) which mimics the striking action and produces the sound of the tabla. The tabla has two drums which are used for different purposes depending on the context in which it is being played. There are various keys for each drum and our code can support one key for each drum. The keys (called *bols*) can be swapped for one another by simple alterations to the code. Our code contains all 12 bols which can be played separately and/or in unison.

Our demo can be viewed via this link:
https://drive.google.com/drive/folders/15UrFgfsl3o5V4w-0nplK6RJpc2qvVewn

The Disciplines

Indian Music – The Art

India is a country popular and well known for its rich heritage from different aspects. Music has been an ongoing and unprecedented part which has always been integral to the culture.
The term *raga* has wide-spread meaning and is excessively used while practicing Indian instruments, but it is simply a set of certain keys which are played in a scale and in different tunes. They produce a platform for the musician to explore his/her creative potential. The effect of playing a *raga* is experienced by the listener who should be able to distinguish the nuances which are being played.

It may seem similar to the scales used in classical western music, but *ragas* are more specific and use a framework which ascends and descends whilst the keys are played. (Vishwanathan, 1997) This allows many different *ragas* to arise from each scale.

Here are some of the scales of some *ragas* with reference to the western classical scale:

Ragas	Scale						
Asavari	C	D	E♭	F	G	A♭	B♭
Bilawal	C	D	E	F	G	A	B
Bhairav	C	D♭	E	F	G	A♭	B
Bhairavi	C	D♭	E♭	F	G	A♭	B♭
Kafi	C	D	E♭	F	G	A	B♭
Kalyan	C	D	E	F♯	G	A	B
Khamaj	C	D	E	F	G	A	B♭
Purvi	C	D♭	E	F♯	G	A♭	B
Todi	C	D♭	E♭	F♯	G	A♭	B♭

Some *ragas* and their scales (from Raag Hindustani)

There are over 80 different *ragas*, and it is always possible for experienced and qualified musicians to 'create' their own *raga*.

Listeners and connoisseurs of Indian music often exclaim that *ragas* tend to have an emotional effect. The basic premise of behavioral therapy is that cognitions influence emotion, which in turn influences an individual's behavior. Music is the language of sentiment, and no other art form elicits such strong feelings as music. (Sanivarapu, 2015)

Within ragas, through the various vocalists and instruments which are played, the tabla has a specific aesthetic or flavor which is termed as *rasa*. (Pathloth, 2020)

A popular Vedic depiction of the *rasas* – from The Natya Sastra

The 8 primary rasas are:

- *Sringara*: romance, love, attractiveness
- *Hasya*: laughter, mirth, comedy
- *Rudra*: Fury
- *Karunya*: compassion, mercy
- *Bibhatsam*: disgust, aversion
- *Bhayanaka*: horror, terror
- *Vira*: heroism
- *Adbhuta*: wonder, amazement

A ninth: *Shanta Rasa* – tranquility, was added as a seeming balance to all of the other 8 *rasas*.

Through the different *ragas* which can be played by the tabla, each of these emotions can be very prominently portrayed. Many performers and famous tabla players have spoken about how their years of training involve a deep study into the understanding towards *ragas* and *rasas*. It helps to form a connection with what they refer to as the ‘soul’ of the performing art. (Pathloth, 2020)

The Tabla

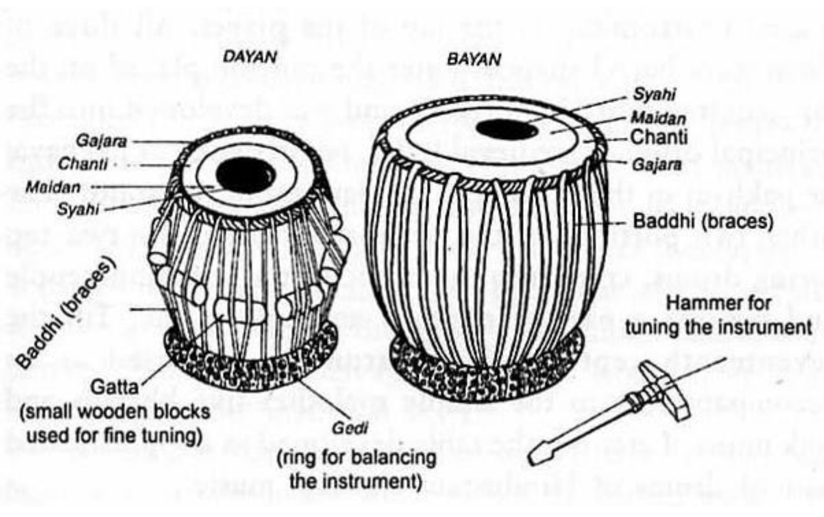
The pakhawaj, a two-headed barrel-shaped drum, was the first percussion instrument used in North Indian classical music. The pakhawaj used to accompany other instruments and also vocalists. According to modern research, the tabla was created in the first half of the 18th century by a drummer named Amir Khusru, who was tasked with creating a more subtle and melodic percussion instrument to complement the new style of music.

Drums and other instruments which are hit involve a lot of acoustic science which go into them since their needs to be cohesion in the music produced due to the fact there are multiple hits which produce the variety. The Tabla has been a very cohesive instrument which has allowed Indian musicians to expand their horizons on the different *ragas* they can play and *rasas* which they can influence across.

There are various different types of tabla, all which follow the same twin drum set however have different beats – which are used for different occasions. The two drums are called the *dayan* (smaller) and *bayan* (bigger).

As mentioned before the keys of the tabla are the different rhythmic patterns which are called *bols*. They are used in uniquely according to the situation and create the beat of the tabla. When played in conjunction or if mixed and matched they create wonderful upbeat rhythms.

The different bols for the *dayan* are: *dhe, dhun, duh, na, re, ta, tat, ti and tin.* And for the *bayan*: *ghe, kath and ke*. The are some bols which are played in harmony on both the drums which are: *dha and dhin*. Bols are also played in successive manners and in flams which are: *dhe, re, ke and ta* & *ge + na = ghran* and *ti + ke + dhe = trikhe.*



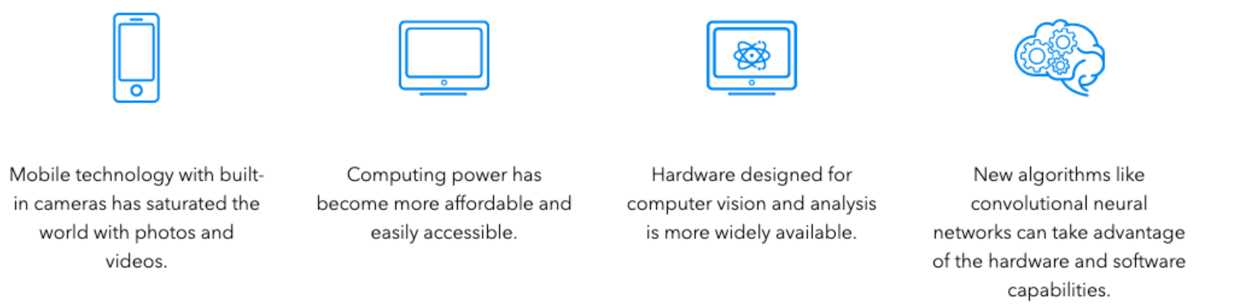
Parts of a Tabla (picture from India Instruments)

Computer Vision – The Science

Computer Vision is a sub-field of Artificial Intelligence, which enable computers to interpret and understand the visual world, in short, automate the human visual cortex. Using digital images from cameras and videos and deep learning models, machines can accurately identify and classify objects — and then react to what they “see” (SAS Institute, 2021).

Post the 1960s, Computer Vision was a major topic of research in universities pioneering in A.I. The goal was to copy the human visual system to endow robots with intelligent behavior (Szeliski, 2010). In 1966, the “Summer Vision” Project at M.I.T, achieved this by attaching a camera to a computer and having it give a description of what it “saw” (Papert, 1966). The 1970s, brought with them the first commercially usable computer vision algorithms. Which were used to interpret typed or handwritten text using OCR (Optical Character Recognition). Following the Internet Boom in the 1990s, large sets of images were made available online for analysis, ushering in the era of facial recognition programs (SAS Institute, 2021).

Today, a number of factors have converged to bring about a renaissance in computer vision:



(Source)

And these factors have culminated in making Computer Vision Practices accessible from the comfort of your home. OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products (OpenCV, 2021). And for our project we used this Library to build a Virtual Tabla Set.

Our Algorithm

Integrated Development Environment Used – Visual Studio Code
Operating System - Windows

Our program uses OpenCV to access the system (laptop's) camera and scans for the color red in two specific regions of the camera display. These regions show the pictures of the two instruments in a Tabla Set. On detecting the color red, in the specified regions, it produces distinct sounds for each region. The sounds can be changed in the code.

Below are the steps we followed while preparing our algorithm:

1. Before anything, we need to prepare the computer to be able to run Computer Vision Algorithms. As such we need to download the required files via command line (or Terminal) Codes.

    a. pip3 install opencv-python
    b. pip3 install pygame

    This downloads OpenCV and Pygame onto the system

2. We need to import some libraries from the downloaded modules to set up the right environment for the code to run.

```
# Importing the necessary libraries
import numpy as np
import time
import cv2
from pygame import mixer
```

NumPy – Converting python list to numpy array and performing required matrix operations.
Time – It provides many ways of representing time in code, such as objects, numbers, and strings.
cv2 – loads OpenCV into the IDE
Pygame - It includes computer graphics and sound libraries designed to be used with the Python programming language.

3. Next, we've prepared some functions that: Detect the color red and Play the specific sounds

```python
# This function is used to check if red color is present in the small region
def detect_in_region(frame,sound):

    hsv = cv2.cvtColor(frame, cv2.COLOR_BGR2HSV) # Converting to HSV
    mask = cv2.inRange(hsv, redLower, redUpper)  # Creating mask
    detected = np.sum(mask)                      # Calculating the number of red pixels

    # Call the function to play the Tabla bol
    play_bol(detected,sound)

    return mask
```

When detect_in_region is called, it uses cv2 to detect all the colors in the specified region of the camera (cv2.cvtColor). It then creates a mask of red color within a given HSV (Hue, Saturation, Brightness) range of the color red, given by redLower and redUpper. Finally, it calculates the number of the red pixels in the mask. And calls the next function to play the sounds.

```python
# This function plays the corresponding Tabla bol if a red color object is detected in the region
def play_bol(detected,sound):

    # Checks if the detected red color is greater that a preset value
    play = (detected) > tabla_thickness[0]*tabla_thickness[1]*0.8

    # If it is detected play the corresponding Tabla bol
    if play and sound==1:
        tabla_2.play()
        time.sleep(0.1)
    elif play and sound==2:
        tabla_1.play()
        time.sleep(0.1)
```

This function accepts the number of red pixel and checks if they are greater than a preset value that we've initialized (32,000 pixels). If greater than 32000, it sets play = true. 'sound', gets initialized with a value of 1 or 2 based on which one of the two regions was interfered with. Accordingly, it plays sound 1 or sound 2. 'time.sleep()' controls the speed of the sounds produced.

4. The main code

```python
# Obtain input from the webcam
camera = cv2.VideoCapture(0)
ret,frame = camera.read()
H,W = frame.shape[:2]
```

The code uses OpenCV to access the laptop camera, and checks to see if the camera frame is read correctly (camera.read() = True). It then finds the height and width of the frame using frame.shape[:2].

```
# Read the image of the two Tablas
tabla1 = cv2.resize(cv2.imread('./images/Tabla1.png'),(200,200),interpolation=cv2.INTER_CUBIC)
tabla2 = cv2.resize(cv2.imread('./images/Tabla2.png'),(200,200),interpolation=cv2.INTER_CUBIC)


# Set the region area for detecting red color
tabla1_center = [np.shape(frame)[1]*2//8,np.shape(frame)[0]*6//8]
tabla2_center = [np.shape(frame)[1]*6//8,np.shape(frame)[0]*6//8]

tabla_thickness = [200,200]
tabla1_top = [tabla1_center[0]-tabla_thickness[0]//2,tabla1_center[1]-tabla_thickness[1]//2]
tabla1_btm = [tabla1_center[0]+tabla_thickness[0]//2,tabla1_center[1]+tabla_thickness[1]//2]

tabla2_thickness = [200,200]
tabla2_top = [tabla2_center[0]-tabla2_thickness[0]//2,tabla2_center[1]-tabla2_thickness[1]//2]
tabla2_btm = [tabla2_center[0]+tabla2_thickness[0]//2,tabla2_center[1]+tabla2_thickness[1]//2]
```

The first two lines, resize the images that have been extracted with the Height and Width as 200x200, for both images. The rest of the code sets up the detection regions for each of the images np.shape(frame)[1] is the width and np.shape(frame)[0] is the height. 'tabla_thickness' is the resized image dimensions and are set to [200,200]. The code uses simple arithmetic to calculate the image centre, top and bottom.

5. Infinite Loop

```
while True:

    # Select the current frame
    ret, frame = camera.read()
    frame = cv2.flip(frame,1)

    if not(ret):
        break

    # Select region corresponding to Tabla 2
    tabla2_region = np.copy(frame[tabla2_top[1]:tabla2_btm[1],tabla2_top[0]:tabla2_btm[0]])
    mask = detect_in_region(tabla2_region,1)

    # Select region corresponding to Tabla 1
    tabla1_region = np.copy(frame[tabla1_top[1]:tabla1_btm[1],tabla1_top[0]:tabla1_btm[0]])
    mask = detect_in_region(tabla1_region,2)

    # Output project title
    cv2.putText(frame,'SM2702 Project 2 - Tabla',(10,30),2,1,(20,20,20),2)

    # If flag is selected, display the region under detection
    if verbose:
        frame[tabla2_top[1]:tabla2_btm[1],tabla2_top[0]:tabla2_btm[0]] = cv2.bitwise_and(frame[tabla2_top[1]:tabla2_btm[1],tabla2_top[0]:tabla2_btm[0]],fra
        frame[tabla1_top[1]:tabla1_btm[1],tabla1_top[0]:tabla1_btm[0]] = cv2.bitwise_and(frame[tabla1_top[1]:tabla1_btm[1],tabla1_top[0]:tabla1_btm[0]],fra

    # If flag is not selected, display the Tablas
    else:
        frame[tabla2_top[1]:tabla2_btm[1],tabla2_top[0]:tabla2_btm[0]] = cv2.addWeighted(tabla2, 1, frame[tabla2_top[1]:tabla2_btm[1],tabla2_top[0]:tabla2_
        frame[tabla1_top[1]:tabla1_btm[1],tabla1_top[0]:tabla1_btm[0]] = cv2.addWeighted(tabla1, 1, frame[tabla1_top[1]:tabla1_btm[1],tabla1_top[0]:tabla1_


    cv2.imshow('Output',frame)
    key = cv2.waitKey(1) & 0xFF

    # 'Q' to exit
    if key == ord("q"):
        break
```

The code now runs an infinite loop. This portion of the code is used to establish the detection regions for the camera in real-time and also waits until we give it the order to terminate by pressing the “Q” Key.

This document explains the main portions of the algorithm, the rest is explained via comments in the ‘.py’ file.

Interdisciplinary Aspect

Throughout the evolution of music, science; with concentrated reference to computer science and technology, has always been a progressive platform which has allowed the digitalization of music. This is the relationship we are trying to highlight with the Air Tabla.

According to Nicholas Reynolds, the interface in the form of AI has helped in the development of music education. Brining in this avenue of education and learning really adds a lot of value to the Air Tabla which open it up to becoming a vital tool for the provision of music lessons, especially given how the pandemic has been redefining the status quo. The delivery of audio content has proven to have been much more effective when it has both a sound and visual digital medium.

Furthermore, the interdisciplinary appreciation between AI and music has allowed various music new age producers to get much more creative and increase their song production rate. Computer vision allows them to reach out and explore various genres. From a study performed at Carnegie Mellon, it is interesting to notice the use of Computer Vision to represent music as images and how that has been vital in the growth of further expanding this relationship.

At the IEEE Conference on Computer Vision and Pattern Recognition, many questions were posed, there was a lot of conjecture on this concept becoming a reality. Earlier, a lot more focus was spent on the reverse process, capturing the EM ways of audio and presenting it in a 2D wave pattern. Even in this context, the interdisciplinary connection between music and computer vision came into life.

Purpose & Motivation

While starting out this project and looking for a spark to start our proceedings, an artwork presented during Mr Scott Hessels’ SM1701 course – A Very Nervous System by David Rokeby became more and more intriguing. The idea of producing music with movement from the body; there was something so unique and innovative about it which made us proceed with an idea on those lines. Rokeby used cameras an image-processors to translate the movement to sound.

As a more modern extension of his work, we realized the use of Computer Vision to allow us to produce a more flexible and customizable product.

David Rokeby demonstrating A Very Nervous System
(image from SM1701 course notes)

In an essay accompanying his work he states,

"The interaction between technologies, human thought, and an individual proves the unique character of the given work and emphasizes its outstanding significance for the further development of art."

The further development of the art, as he may have foreseen is the digitalization of music, allowing artists to explore various styles and genres with more ease. We felt whilst following the trajectory, which was started by an artist like Rokeby, our Air Tabla carries a form of validation. Our purpose to make this was also to raise awareness about Indian music and the various instruments used with the essence and story they each contain. Our medium is the disciplines of music itself, along with Computer Vision (AI). In Rokeby's essay – Transforming Mirrors, he references a famous quote by Marshall McLuhan, "The Medium is the Message." We can relate with McLuhan's statement by looking at the fact where we want to increase awareness about a form of music and allow it to be explored in various fields by the use of that form of music with the advancements in technology.

References

Computer vision: What it is and why it matters. (n.d.). Retrieved May 02, 2021, from https://www.sas.com/en_hk/insights/analytics/computer-vision.html#:~:text=Computer%20vision%20is%20a%20field,to%20what%20they%20%E2%80%9Csee.%E2%80%9D

Courtney, Todd A. Dombrowski, David. "Basic Technique of Tabla bols: Dhaa, Dhin, Ga, Ka, Naa, Na, Taa, Tak,TiRaKiTa, Tin, Tu, etc". chandrakantha.com.

Cunningham, E. (2013, August 02). Music in India: The art Of Konokol. Retrieved May 02, 2021, from https://theculturetrip.com/asia/india/articles/the-art-of-konokol-the-practice-of-south-indian-carnatic-music/

Datta, A., Solanki, S., Sengupta, R., Chakraborty, S., Mahto, K., & Patranabis, A. (2018). *Signal analysis of hindustani classical music*.

David Rokeby : Transforming Mirrors. (n.d.). Retrieved May 02, 2021, from http://www.davidrokeby.com/mirrorsintro.html

David Rokeby : Very nervous system. (1996). Retrieved May 02, 2021, from http://www.davidrokeby.com/vns.html

Indian Classical Music. (n.d.). Retrieved May 02, 2021, from https://raag-hindustani.com/

Miśra, V. (2015). *Art and science of playing tabla*. Delhi: Publications Division, Ministry of Information and Broadcasting, Government of India.

Old Delhi Music. (n.d.). Retrieved May 02, 2021, from https://olddelhimusic.com/

Open CV (2020, November 04). Retrieved May 02, 2021, from https://opencv.org/about/

Papert, S. (1966). *The Summer Vision Project*. M.I.T.

Pathloth, V., Dr. (2019). Taking center stage: The Poet-Saint and the IMPERSONATOR. *Impersonations: The Artifice of Brahmin Masculinity in South Indian Music,* 33-54. doi:10.1525/luminos.72.b

Pottekkat, N. (n.d.). Virtual Drums. Retrieved May 02, 2021, from https://github.com/navendu-pottekkat/virtual-drums

Pudaruth, S. K. (2016). A reflection on the aesthetics of Indian music, with special reference to Hindustani raga-sangita. *SAGE Open, 6*(4), 215824401667451. doi:10.1177/2158244016674512

Reynolds, N. (2010). Technology and computers in music and music education. *Key Competencies in the Knowledge Society,* 333-343. doi:10.1007/978-3-642-15378-5_32

Sanivarapu, S. (2015). India′s rich musical heritage has a lot to offer to modern psychiatry. *Indian Journal of Psychiatry, 57*(2), 210. doi:10.4103/0019-5545.158201

Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Springer.

Viswanathan, T. (1977). The analysis of Raga Alapana in South Indian music. *Asian Music, 9*(1), 13. doi:10.2307/833817

Yan Ke, Hoiem, D., & Sukthankar, R. (2005). Computer vision for music identification. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. doi:10.1109/cvpr.2005.105