CMPT365 Assignment1 Report

Yubo Hu yuboh@sfu.ca
Chauncey Liu cla284@sfu.ca

1. The brief introduction

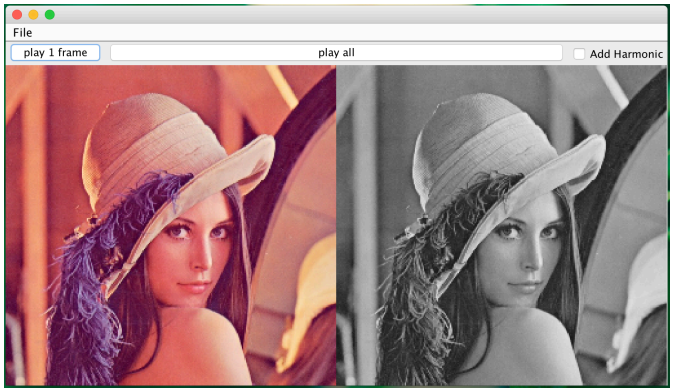
The main idea of the assignment is to find a proper way to convert information from video-based to audio-based. The audio generated can help human's brain reconstruct the image. Such a method can be used to enable blind people to "watch" the video.

Due to information processing speed is limited for ears, the amount of information cannot be too large to process. At the same time, the audio generated should have the proper loudness and frequency rather than sounds harsh. And the most important thing is how to find a best conversion using which human brain can reconstruct the original image from the audio generated as accurate as possible.

2. Our solution

Firstly, we subsample the image to 64*64 pixels and convert the color image into a luminance image with the NTSC formula:

$$\mathrm{Y}' = 0.299 * \mathrm{R}' + 0.587 * \mathrm{G}' + 0.114 * \mathrm{B}'$$



Secondly for each column of the luminance image, we assign the lowest frequency to the pixel on the first row and highest frequency to the pixel on the last low.

$$singal = \sin(2\pi * frequency * t)$$

$$frequency = 20 + 60 * rowNumber$$

And then we add up all signals we get from current column, using the intensity of pixel as the soundness. Here are formulas to perform the calculation:

$$chord\ for\ each\ column = \sum_{n=1}^{64} (\text{pixel intensity} * \text{signal})$$

Finally, we continuously play the chords derived from each column to make an audio.

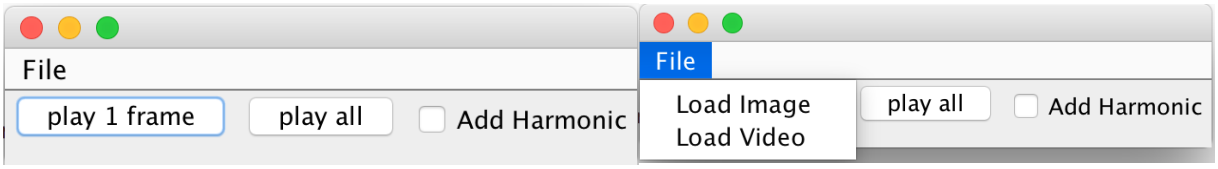
3. Program works on

	C++	Java	Other
Language		√	
	Any Res.	Input must be 64 *64	
Works on	√		
	Any video	On particular video	An image
Works on		√	√

4. User interface

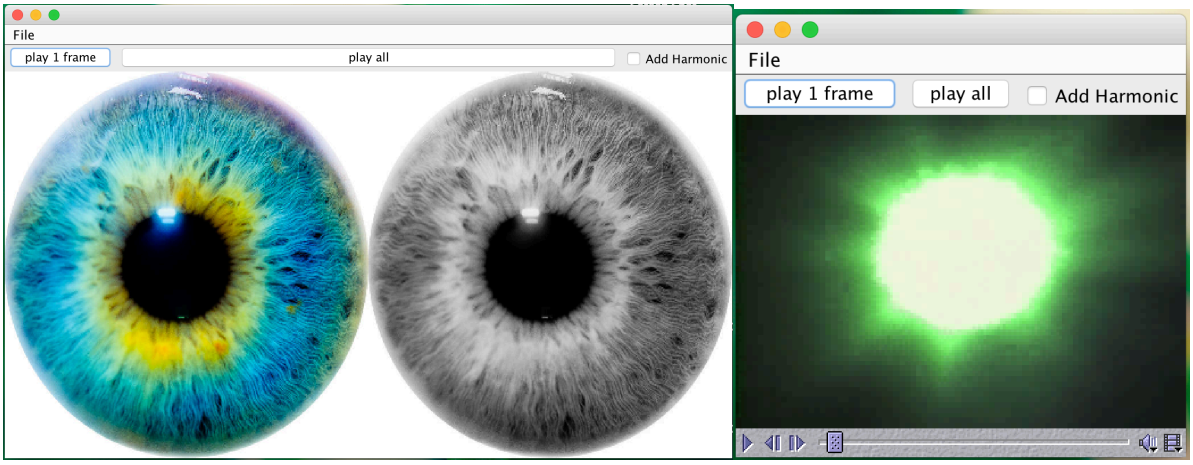
4.1 Initial interface

By clicking the “File” button in the menu bar, we can choose either to load an image or to load a video.



4.2 Image loading interface

After the image is loaded, by clicking “play 1 frame” the sound will be generated from the image and played.



4.3 Video loading interface

After the video is loaded, by clicking “play 1 frame” the sound will be generated from the current frame of the video and played. You can also drag the sliding control bar to choose which frame to display.

4.4 Extra feature

To make the audio generated more acceptable, we add a function called “Add Harmonic”. By checking the checkbox “Add Harmonic”, we will add harmonic to audio.

5. Summary

The method we apply in this assignment can generate the audio that extract some features from the original image. With the information in the audio we will be able to differentiate the dark image and bright image by the soundness of audio.

It is still hard to reconstruct the original image from the sound information. And the audio sounds irregular (like noise), which might make human feel uncomfortable. Another disadvantage is that the color information of the image is lost, which is also an important feature of the image.

To make this method helpful to blind people, we still need to improve the formula to get a harmonious audio. And we also need to extract the information from image like pixel color variety and convert it to a more obvious feature in audio.