

ECE 417 MP 4: Video Synthesis

1. Introduction

This lab focused on creating facial animations that matches up with a given audio signal with the help of the Artificial Neural Network. The Artificial Neural Network first attempts to map the facial features with the given audio track and returns a new set of mesh features for each frame (with the help of the interpVert function). For each frame, we then use barycentric coordinate and bilinear interpolation to map each pixel to a new output pixel. Doing this for all frames will then result in a smooth animation that matches the original static image into a video with matched up audio.

2. Methods

There were two specific functions that were needed for the image warping function: barycentric and bilinear interpolation. The barycentric coordinates helped us map the correct triangle mesh to each of the input coordinates. This was done by first knowing that each pixel is mapped by: $x = X \lambda$ where X is the new located triangles and λ is the mapping between the old coordinates and the new coordinates. We can find λ by doing $\lambda = X^{-1} x$. Afterwards, we can obtain our new set of output coordinates by doing $u = U \lambda$ where U was our original located triangles and u would be our new located pixels. In the code provided, the barycentric coordinate function returns the corresponding triangle for each of the input coordinates. This is determined by checking whether the λ values are between 0 and 1 (this condition implies that the pixel is inside the triangle).

Once we obtained our new coordinates, we then use bilinear interpolation to obtain the corresponding pixel value for that coordinate. Bilinear interpolation looks at the pixels that are adjacent to the current pixel location and adds them up together to obtain a weighed pixel value based on its neighbored pixels.

3. Results

The result ended up being a clean video of the lips moving whenever the audio outputs a voice. The inside of the mouth is still black however as we didn't provide any inside features of the mouth.

4. Extra Credit

Teeth/Tongue

In the original case, for cases when there were no triangles found within the corresponding inputted pixel, the pixel value would then be set as black. This resulted in a black area mouth with no other physical features. Instead of setting the pixel as black, the pixel value is now set to a pre-loaded teeth image at that specific location.



Figure 1: Mouth with Teeth Inside

New Facial Animation

To create a new face, we need to first create a new mesh model for the new face. The new face that was taken was Obama. In order to adjust the mesh, the previous mesh was used with this new face except with some slight changes in the vertices position in order to match with the lip structure accurately.



Figure 2: New Mesh Model with New Face