BLG453E
Homework-5

06.12.2019

Res. Asst. Yusuf Huseyin Sahin
sahinyu@itu.edu.tr

The doll was so utterly devoid of imagination that what we imagined for it was inexhaustible. For hours, for weeks on end, we must have been content to lay the first fine silk of our hearts in folds around this immobile mannequin, but I have to believe there were certain abysmally long afternoons when our twofold inspirations petered out and we suddenly sat in front of it, expecting some response. (...) It remained silent then, not because it felt superior, but silent because this was its established form of evasion and because it was made of useless and absolutely unresponsive material. It was silent, and the idea did not even occur to it that this silence must confer considerable importance on it in a world where destiny and indeed God himself have become famous mainly by not speaking to us. (...) That we did not then make you into an idol, you brat, and perish from fear of you, was because —I must tell you— it was not you we had in mind. We were thinking of something quite different, invisible, something we held at arm's length from you and from ourselves, furtively, with vague anticipation, something for which both of us were in a way only pretexts. We were thinking of a soul, the soul of the doll.

Some Reflections on Dolls, Rainer Maria Rilke

Ja, ich reiß' der Puppe den Kopf ab
Und dann beiß' ich der Puppe den Hals ab
Es geht mir nicht gut, nein!

Till Lindemann, Puppe

- You should write all your code in Python language.
- Cheating is highly discouraged.
- Ninova only stores files under 20 MB. If you could not upload your results, you can share them with me via Dropbox, or send me private YouTube video links for each part's results.

1 - Part 1: A (rag)doll with a soul (70 pts.)

Both Rilke and Lindemann were angry with the dolls, and you will be angry with them too. However, instead of a toy doll, in this homework we will work on a ragdoll which is used for character animation. With the homework document, I uploaded two video files: *walker.avi* and *walker_hand.avi*. The first file is a 12 seconds long video containing a walking ragdoll animation and the other file has the segmentation of the ragdoll's right hand. Some example images are given in Figure 1.

Figure 1: A frame from the ragdoll video and its right hand segmentation

To read the videos frame by frame you can benefit from the *moviepy* library. An example for reading frames is given below.

```
import moviepy.video.io.VideoFileClip as mpy
import cv2

walker = mpy.VideoFileClip("walker.avi")
walker_hand = mpy.VideoFileClip("walker_hand.avi")

frame_count = walker_hand.reader.nframes
video_fps = walker_hand.fps

for i in range(frame_count):
    walker_frame = walker.get_frame(i*1.0/video_fps) #To get frames by ID
    walker_hand_frame = walker_hand.get_frame(i*1.0/video_fps)

    walker_hand_frame = (walker_hand_frame>127) #The right hand
```

In this part of the homework you will implement Lucas-Kanade algorithm to track the hand. For every two consecutive frames;

- Use the frame from walker_hand as a mask to obtain information about the hand's position. You can use a rectangular area to work on.
- Implement Lucas-Kanade algorithm locally to find the OF vector of the hand.
- Using cv2.arrowedLine function, place an arrow according to the OF vector.

You can skip 50% of frames (e.g. every odd frame) to obtain more accurate and quicker results. Obtain a video with OF arrows.

2 - Part 2: "I'm Nobody! Who are you? / Are you -Nobody- too?" (30 pts.)

VGG Face Dataset[1] consists of images of many famous people collected from the internet. I processed and uploaded a subset of VGG-Face which contains 32x32 grayscale face images. Take all the images as 1024 sized vectors. Using **sklearn** library, employ PCA to project these images to a lower dimensional space (e.g. $\mathbb{R}^{100}$). If you have computation problems you can use a subset of the given set.

Then, project one of your 32x32 face image to the same space. Your eyes in the image should be approximately on 10th row. Use 10-nearest neighbors classification inside the projected space to find the celebrities who are most similar to you.

[1] Parkhi, O. M., Vedaldi, A., Zisserman, A. (2015, September). Deep face recognition. In bmvc (Vol. 1, No. 3, p. 6).