Image Processing - Exercise 1

Liel Amar, liel.amar, 211771993

Introduction

In this assignment, we were to create an algorithm which detects scene cuts in given videos. This is done through the usage of Histograms, and more specifically – detecting large, instantaneous changes, between the cumulative Histograms of two consecutive frames.

As for the differences between the two given categories of videos, we see that the first category (category 1) includes files that include 2 videos with 1 scene cut in between. In this category, there aren't any changes in terms of the color/gray-level. However, in the 2nd category, in addition to the single cut appearing, there's a change in the color/gray-level, which results in more drastic changes to the Histograms, but not as drastic as a scene cut.

Algorithm

The algorithm I used to detect scene cuts is the following:

1. Reading, importing and gray-scaling the wanted video
2. Looping over all frames, and for each, saving its cumulative histogram
3. Calculating the delta (difference) between the histograms of every 2 consecutive frames, using the L1 norm.
4. Finding the index of the maximum delta, which represents the point with the largest difference between two consecutive frames, marking the most likely scene cut.
5. Return the frame matching the same index, and the next frame, marking the scene cut.

Although the two video categories differ a bit, the algorithm provided above is able to detect the scene cut in both. This is due to the fact that in step 5, we find the maximum delta between 2 consecutive frames, and use it as the "threshold" to detect cuts. With this logic, when we look at category 1, we see that indeed most deltas are relatively small, until there is a scene cut, in which the two cumulative histograms of the last

frame of the first video, and the first frame of the second video vary the most - resulting in the largest delta which is later on selected on step 5. Similarly, in category 2, we see that while there's a change to the color/gray-level, which results in a higher delta than normal, it is still smaller compared to the delta between two different clips. Thus, selecting the maximum delta still results in us detecting the correct frame in which a scene cut happens.

Implementation Details

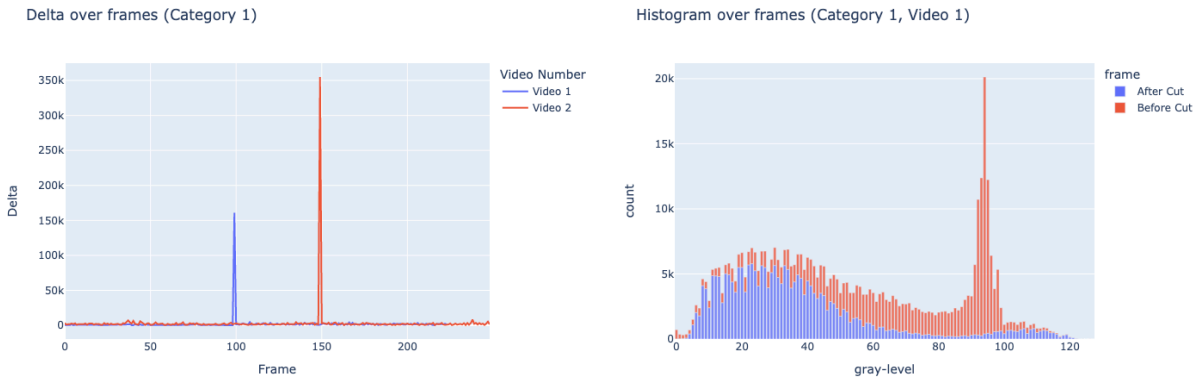
To implement the above algorithm, I used two existing libraries, which are: *numpy*, and *mediapy*. I chose these libraries specifically because either I have experience using them, or they were recommended to me by the course staff.

Let's describe how each step was implemented:

1. Using mediapy, and specifically the #read_video function, we load the video as a list of 3d matrices, each representing a frame as (x,y) coordinates. In addition, the output format argument was used to convert the imported video into a grayscale version
2. Then, I created an empty list of histograms, looped over every frame in the grayscale video, used numpy to calculate the cumulative histogram with the bins hyperparameter (set to 256), and appended the result into the list of histograms. This is done by first calling numpy.histogram to get the histogram of the current frame, and then numpy.cumsum to get the cumulative histogram.
3. Similarly, I created an empty list of deltas and then looped over all the consecutive pairs of histograms. For each pair of histograms (vectors), I calculated the L1 norm by subtracting the vectors, taking the absolute value with numpy.abs and then summing the differences with numpy.sum. Lastly, I appended each delta to the list of deltas.
4. To find the index of the maximum delta, I used numpy.argmax on the list of deltas, which returns the index of the largest number within that list. The list represents the difference between each pair of consecutive frames' histograms, and thus this is the wanted threshold to detect the scene cut.
5. Lastly, we return the index obtained in the previous step, which represents the frame in which the scene cut starts, and the following frame as well due to the assignment's requirements.

Category 1 Results

The results for the scene cut detection of category one are:
video1_category1.mp4 cuts at frames 99-100
video2_category1.mp4 cuts at frames 149-150

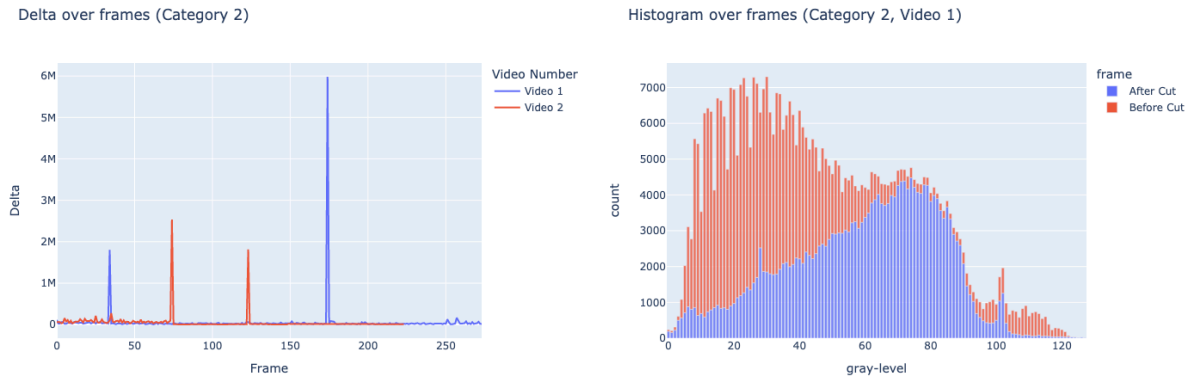


As can be seen, in Category 1, which only includes one scene cut and no more, the delta between every 2 consecutive frames is relatively low, until we get to frames 99-100 in video 1 and frames 149-150 in video 2. These frames have a huge spike (150,000 and 350,000), which represent a large difference in the histograms, suggesting there's a scene cut. When we look at the Histograms themselves and examine the gray-level before and after the cut, we see that before the cut we have whiter pixels (higher gray level), unlike after the cut, in which we have many more dark pixels (lower gray level). From this we can indeed infer there's a big difference between the two frames. A similar thing can be found in Video 2 Category 1.

Category 2 Results

The results for the scene cut detection of category two are:
video3_category2.mp4 cuts at frames 174-175
video4_category2.mp4 cuts at frames 74-75

While watching the videos in category 2, I noticed that there's a change in the color/gray-level during one clip out of the two in each video, unlike category 1.



Category 2 is a bit different than Category 1, as in this category we have 2 spikes for each video. One spike represents the scene cut as before - this being the bigger spike. In addition to that, we have another, smaller spike, which represents the change in color/gray-level in a video. The change in color/gray-level does not affect the delta between 2 histograms that much as the same video remains, and only the depth of colors change - resulting in a smaller spike. This suggests that choosing the largest delta indeed results in the correct scene cut position.
When we look at the Histograms themselves and examine the gray-level before and after the cut, we see that before the cut we have darker pixels (lower gray level), unlike after the cut, in which we have many more light pixels (higher gray level). From this we can indeed infer there's a big difference between the two frames. A similar thing can be found in Video 2 Category 2.

Conclusion

To sum up all our findings and insights, we can conclude that changes between frames in the same clip usually result in low deltas when looking at the cumulative histograms.
However, changing the color/gray-level in the middle of a clip can result in "false" spikes that might affect scene cut detection.
Due to our limitation of two clips per video, and with the find that showed that the scene cut spike (delta) is larger than any color/gray-level change, we know that the scene cut is necessarily happening where the largest delta between two frames occur.