DSAA PROJECT

Team Members

Sarthak Singh, 20171118
Akshay Kharbanda, 20171037
Freya Mehta, 20171184

Problem Statement

These days, the demand for online lectures is increasing. For better visual experience,

along with the video of the lecture, soft copy of the slides is also being embedded into the

video. But most of the universities are manually matching slides from the video to the soft copy which is a laborious task. So the problem statement is to automate this slide matching process.

So to be more precise, you are given a set of noisy slide images extracted from the video and a set of slides from the original ppt. You need to build a mapping for each of the sampled noisy slides with the corresponding original slide.

For example in the dataset given, consider the slides in any of the folders. You will see

4-5 frames sampled from the lecture for which the corresponding ground truth slide is ppt.jpg. You may evaluate the performance of your algorithm on the given data. We would be testing on a more robust dataset.

To run the file

- **Execute the command : python3 <rollno>.py <path/to/slides/directory> <path/to/frames/directory>**

- **NOTE : (In the path, remember to specify the "/" at the end of the path to ensure the script runs!)**

Output

- **Output of the script is directed to <rollno>.txt**

Report

Cross Correlation is a valid technique to relate provide similarity and best match between two images.

In the original dataset, we made a bash script to separate the slide images into the slides directory, and the other frames in the frames directory.

Initially we made a function “load_image”, which opens the image and loads it into a numpy array.

Then, we made a function “corr” to take the cross correlation between the frame and all the slides, using the formula for cross-correlation of 2 matrices. It returns the max normalised cross-correlation value.

Then we run a loop for each frame, and check it’s cross-correlation value with all the slides in the dataset by calling the foo function for each frame and slide combination.

The maximum cross-correlated value is then the cross-correlated value with the matched slide.

We then write this into the <roll_num>.txt file.