

## DSAA Project Report

# Slide Matching

**Team: Get Rect**

Syed Sumaid Ali - 20171092

Rohan Chacko - 20171061

AadilMehdi Sanchawala - 20171043

## Requirements

1. Language
  - a. python3
2. Packages
  - a. opencv-python
  - b. numpy

## Commands to Run

1. `python3 move.py`

This will copy frames and slides from each sub-directory of 'Dataset' to 'frames' directory and 'slides' directory respectively. 'frames' and 'slides' are sub-directories of 'full-test' directory. Each filename is prefixed with the directory name it was copied from.

2. `python3 20171043_20171061_20171092.py <path_to_slides> <path_to_frames>`

Final output will be printed on terminal and also written to output file:

20171043\_20171061\_20171092.txt

## Algorithm

The algorithm used to match the slides with the frames is Normalised Correlation Coefficient.

The correlation coefficient of two random variables is a measure of their strength and direction linear dependence.

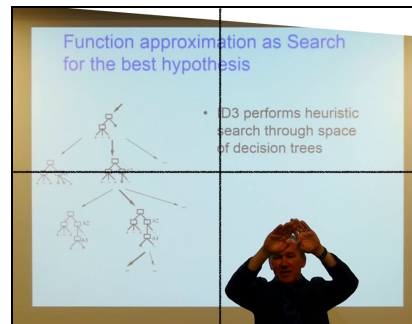
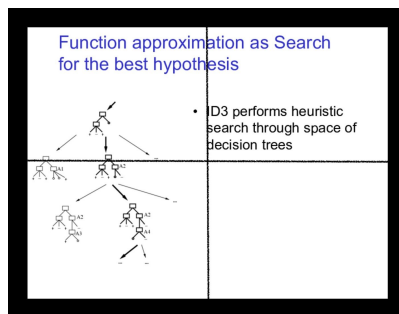
$$R(x, y) = \frac{\sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x', y'} T'(x', y')^2 \cdot \sum_{x', y'} I'(x + x', y + y')^2}}$$

where,

$$\begin{aligned} T'(x', y') &= T(x', y') - 1/(w \cdot h) \cdot \sum_{x'', y''} T(x'', y'') \\ I'(x + x', y + y') &= I(x + x', y + y') - 1/(w \cdot h) \cdot \sum_{x'', y''} I(x + x'', y + y'') \end{aligned}$$

This algorithm is similar to cross-correlation but the random variables are normalised with their covariances.

Using this as a base for matching images, we segment each of the query image(frames) and the template image(slides) into 4.



After segmentation, we apply Template Matching with Normalised Correlation Coefficient method, used to match the query image to the template image.

We do this process for all the templates per query image and choose the template returning the highest values of the Cross Coefficient matrix and return that template as our match for the query.

## Previous Versions and Observations

We tried using the following different methods to achieve the result of matching frames with slides

- **Mean Square Error**
  - We take the mean square error of the template image and the query image and choose the template having the least mean square error as the match to the query frame
  - Observations :
    - Many test cases with professor failed
    - Test Cases with slightly different orientation failed
    - Slight blurring of frames also causes wrong identification
    - Accuracy ~ 40%
- **Normalised Cross Correlation without Segmentation**
  - We used template matching of OpenCV with Normalised Cross Correlation as the method of comparison and supplied the entire template and query images.
  - Observations :
    - An improvement over naive method - Able to recognize slide with slight differences
    - Test cases with professor still fail
    - Accuracy ~ 85%
- **Normalised Correlation Coefficient without Segmentation**
  - We used template matching of OpenCV with Normalised Correlation Coefficient as the method of comparison and supplied the entire template and query images.
  - Observations :
    - Accuracy slightly improved over previous method
    - Accuracy ~ 89%
- **Normalised Correlation Coefficient with Segmentation**
  - We used template matching of OpenCV with Normalised Correlation Coefficient as the method of comparison. We segmented the template and query images

into 4 blocks and compared the corresponding blocks using the abovementioned algorithm.

- Observations :
  - Huge improvement over all previous versions
  - Segmentation helps with edge cases like professor also
  - Accuracy ~ 97.36%

## Final Observations

The final frame - slide matches/mismatches can be categorized into the following:

- a) Certain frames that have a dense amount of red colored text cannot be matched with its corresponding slides (which do not have the red colored text) because if we do consider removing the red colored text, then the frames-slides which do have a moderate amount of red colored text would not be correctly matched. The same holds for slides which have red colored text and frames which do not.
- b) Certain frames that have partial amount of content compared to the corresponding slide matches with a different slide that have the exact amount of textual content as that particular frame. This indicates that our program matches the frame with the best corresponding slide but due to issues with the dataset, it considers as an error.
- c) Certain slides that are the same are present in two different directories causing the program consider it as an error.
- d) There are certain frames that are incorrectly matched to slides due to effects of homography, highly varying intensity across the frame, etc.

## Summary

Total failed test cases : 54/835

Test cases failure because of errors in data set : 32/835

Accuracy ( After adjusting for data set failures ) : 97.36%

We have analyzed each failed test case and provided a detailed report in 'Observation Report.pdf'.