TEAM

20171075 Teja Dhondu
20171077 Swetanjal Datta
20171079 Nishanth Sharma

Automation of slide matching

Approach

1. Going through the dataset, we observed that one of our main hurdles in achieving a high matching accuracy was the noisy images captured by the camera. The clicked images lacked good resolution and sometimes contained motion blur and therefore running a standard cross correlation function to match the frames with the given slides did not give us a good accuracy.
2. We tried out different noise removal techniques to enhance the clarity of the frames so that the essential features could be made more prominent for detection.
3. We found one of the standard digital image processing edge detection techniques called Laplace of Gaussian(LoG) enhancing the quality of frames substantially, making the contents on the slides prominent by highlighting the edges ***in all directions.***
4. To ensure a higher matching accuracy we apply the LoG filter to the original slides as well, and then do the matching using the inbuilt openCV matchTemplate() with parameter TM_CCORR_NORMED.
5. Thresholding was done on both the frames and the slides after applying the LoG filter to further enhance the prominence.

Implementation Details

- Laplacian of Gaussian(LOG) filter was implemented with kernel size of 101x101 and standard deviation(sigma) of 3. We used sigma of size 3 as it allowed us to detect the

edges of characters. An LOG filter is the double derivative of a Gaussian filter. It resembles a hat-like structure in 3D space. It is effective for edge detection as the Gaussian part of the filter smooths the data and the laplacian part detects edges. Unlike Sobel filter, it detects edges in all directions. We implemented the filter using fspecial in matlab and exported it to a .mat file

- After edge detection, we used thresholding. This was so that we could convert the images to binary data images. Each pixels would be either black or white. White indicates the pixel is a part of an edge. Doing this abstracts out a lot of information about the intensities of the pixels in the images and allows us to only compare the key features.
- Finally we computed cross correlation between each pair of images and matched those pairs with greatest correlation.

Observations & Results

1. On using the approach for mapping slides in dataset to frames we have obtained an accuracy of about 95.4491%. As slides usually have text color which is in contrast to the background color, therefore a high accuracy could be obtained by using the above approach. It took approximately 3 hours to run the entire dataset.

2. There can be cases where the various parts of slide is not seen yet or blocked by various objects. If most part of slide is visible then good amount similarity can be observed and hence images can be mapped accurately. Otherwise it might not be mapped properly. To improve the accuracy we can try implementing a BVH algorithm and map slides only when a threshold similarity is observed.

3. The current algorithm has a good amount of accuracy, but is a bit slow. For better efficiency in time, filter2D can be implemented in Fourier domain. filter2D implements convolution in time domain. Instead, we may use the Convolution Theorem to perform the computation in Fourier Domain and then transforming back to time domain. This would decrease the number of computations by a huge margin

4. If it is certain that size of frame and slide have same dimension size we can just use l2norm for similarity measure.