# Optional Projects
# CS 558 – Computer Vision

Akshay Atam
aatam@stevens.edu

## Problem Statement

These projects are a part of the Optional End of Course Projects which includes projects from four different topics: Image Data Association, Object Segmentation, Advanced Image Filtering, and Non-Iterative Super-pixel Segmentation. The dataset used for these projects are namely fountain-P11, Herz-Jesu-P8, entry-P10, and Castle-P19 image sets from the Stretcha MVS evaluation website. Along with these image sets, images used in Assignment 3 and Assignment 4 have also been taken for Non-Iterative Super-pixel Segmentation.

Image Data Association refers to the task of applying unsupervised clustering to a superset of images created from the above mentioned Stretcha MVS evaluation datasets. The goal of this task is to correctly cluster the image sets into disjoint subsets.

The goal of Object Segmentation is to develop a framework to perform segmentation of a specific foreground object. Castle-P19 image set is used to segment a foreground object (in this case a tractor) from the background.

Advanced Image Filtering is the task of applying a specific filter – a multi-dimensional Gaussian filter using the Permutohedral Lattice in this case – on images used for Assignment 4.

Finally, the last assignment aims to implement Non-iterative Super-pixel Segmentation (SNIC) on images from Assignment 3 and 4 and compare it with Simple Linear Iterative Clustering (SLIC) outputs from Assignment 4 and the original SLIC version.

A total of three projects, apart from Advanced Image Filtering, have been completed and presented in this report. In association with the report, the resulting outputs and the code files have been attached for review.

## Background

All that follows on these tasks I gained from Stevens Institute of Technology's CS 558 Computer Vision course, guided by Professor Enrique Dunn.

The project falls under the domain of Computer Vision. This domain of Machine Learning deals with machines that use images, videos, and other visual aids to capture relevant information and make recommendations based on that information. It works in a similar fashion as the human vision by recognizing patterns in images and getting an inference.

The rest of the report is as follows. Project 1 talks about Image Data Association. Project 2 focusses on Object Segmentation. Project 3 comprises of Non-Iterative Super-pixel Segmentation. Next, there is an output section with the output of individual algorithms. Finally, the report includes conclusions and future scope of improvement. The individual jupyter notebooks and output folders is also provided with the report.

# Project 1: Image Data Association

This task involves applying Mean Shift algorithm on two supersets of data. The first superset consists of images from fountain-P11, Herz-Jesu-P8 and entry-P10 image sets. The other superset consists of images from entry-P10 and castle-P19 image sets. Both the supersets contain 29 images.

The goal of the images is to create clusters of similar images into a single aggregated image. The global descriptor used in this task was 4-bin color histogram. The parameters used to obtain the output were threshold distance and kernel bandwidth.

Other higher bin size for the color histogram were also used however, none of them gave promising results. It was observed that increasing the bin size made more clusters with some clusters having only one image. Further, the distance threshold and kernel bandwidth were also tuned to obtain the best clusters. It was observed that the number of clusters decreased as we increased the distance and increased when the threshold distance was kept small.

Upon multiple iterations of testing different distance and kernel bandwidth measures, the algorithm was able to make five clusters when threshold distance was set to 49500 with kernel bandwidth of 27500. The obtained clusters were stitched together in one single image using a scaling factor that was able to attach the images side-by-side. The output clusters are stored in problem1 folder under output directory.

# Project 2: Object Segmentation

This task required object segmentation on the castle-P19 image dataset. The goal was to separate the foreground object (tractor) from the background. To achieve this, the algorithm needed to know which pixels to consider to correctly separate the foreground. To achieve this, a software called GIMP was used to fill the background with a specific color (the color blue was chosen however, any color could be chosen).

The approach of this task was similar to Pixel Classification task that was given on Assignment 4. So, the approach that was taken to perform this task was identical to that of Assignment 4, we needed to detect pixels that corresponded to background and the image of interest.

However, the results obtained were far from perfect. Using any color on the output image resulted in visually unappealing results so the background was made black so that we could see the output clearly and identify gaps in the implementation. To avoid the report getting filled with images, a subset of 3 images is shown in this report. The rest of the images, including the subset images, can be looked in the output directory under problem2 folder.

# Project 3: Simple Non-Iterative Clustering

This task required to implement the SNIC algorithm and comparing the results from images used in Assignment 3. The basic idea of SNIC was obtained by going through with the paper and understanding the algorithm provided.

A priority queue was used to store the K elements that was the seed for the algorithm. The use of queues proved to be helpful as this algorithm is fast and runs on a single iteration. The paper also introduced a new distance computation measure using the spatial coordinates and CIELAB color space. The queue was then populated with pixels that were 4 or 8-connected to a current superpixel. Since there was a choice to choose the connected pixels, I chose to get pixels that are 4-distance away.

Some additional helper functions were used to compute the grid, linear interpolation and square norm for getting the distance measure and calculation of the distance using the normalization factors $s$ and $m$ as proposed in the paper.

A label map was first created and was initialized with -1. The algorithm states to initialize with 0 but upon running the algorithm it gave an error on division. The priority queue was initialized with the position, color and index of the centroids and the loop was run while the queue was not empty. The element with the smallest distance to a centroid was chosen and was done for all centroids, keeping track of the labels and the distance map. Finally, the resulting image was created and boundaries were created on adjacent superpixel segments.

Additionally, the SNIC algorithm was compared with its iterative counterpart SLIC that was in Assignment 4 and the original GUI version of SLIC. It was observed that the Assignment 4 approach of SLIC used a block size of 50 and max_iter = 3, the resulting superpixels were square in shape. A single image has been shown in this report. The rest are available in output directory under problem3 folder.

# Output

Following are the results from each of the projects:

## Project 1: Image Data Association

Cluster 1: 8 images

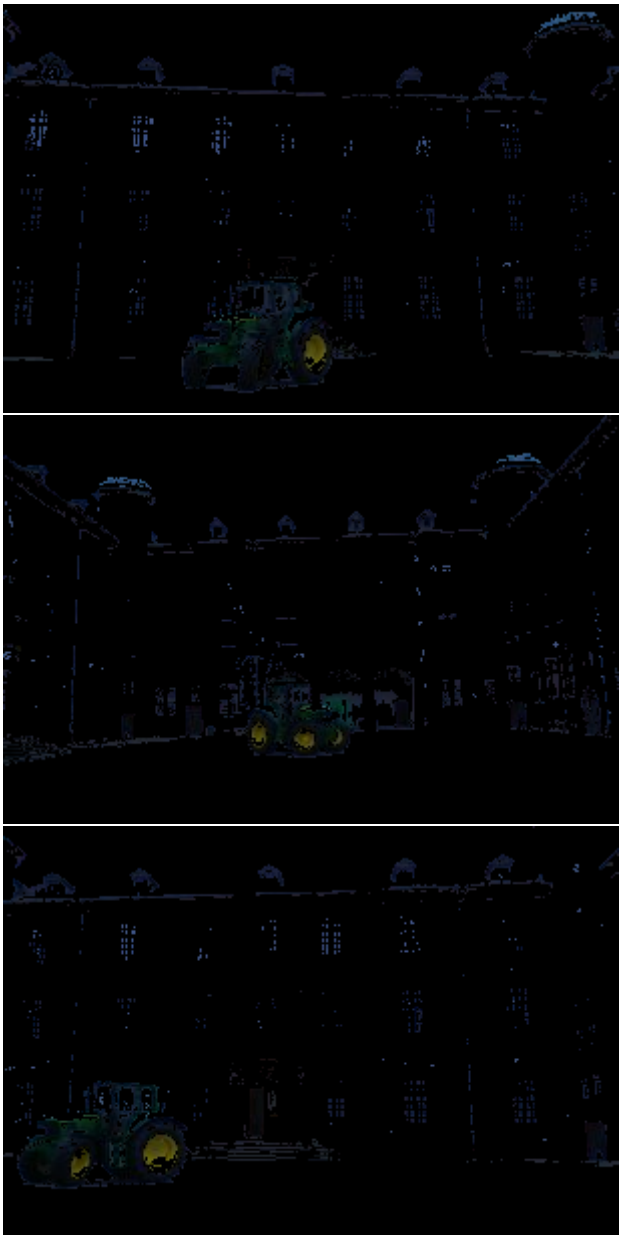
Cluster 2: 14 images

Cluster 3: 11 images

Cluster 4: 15 images

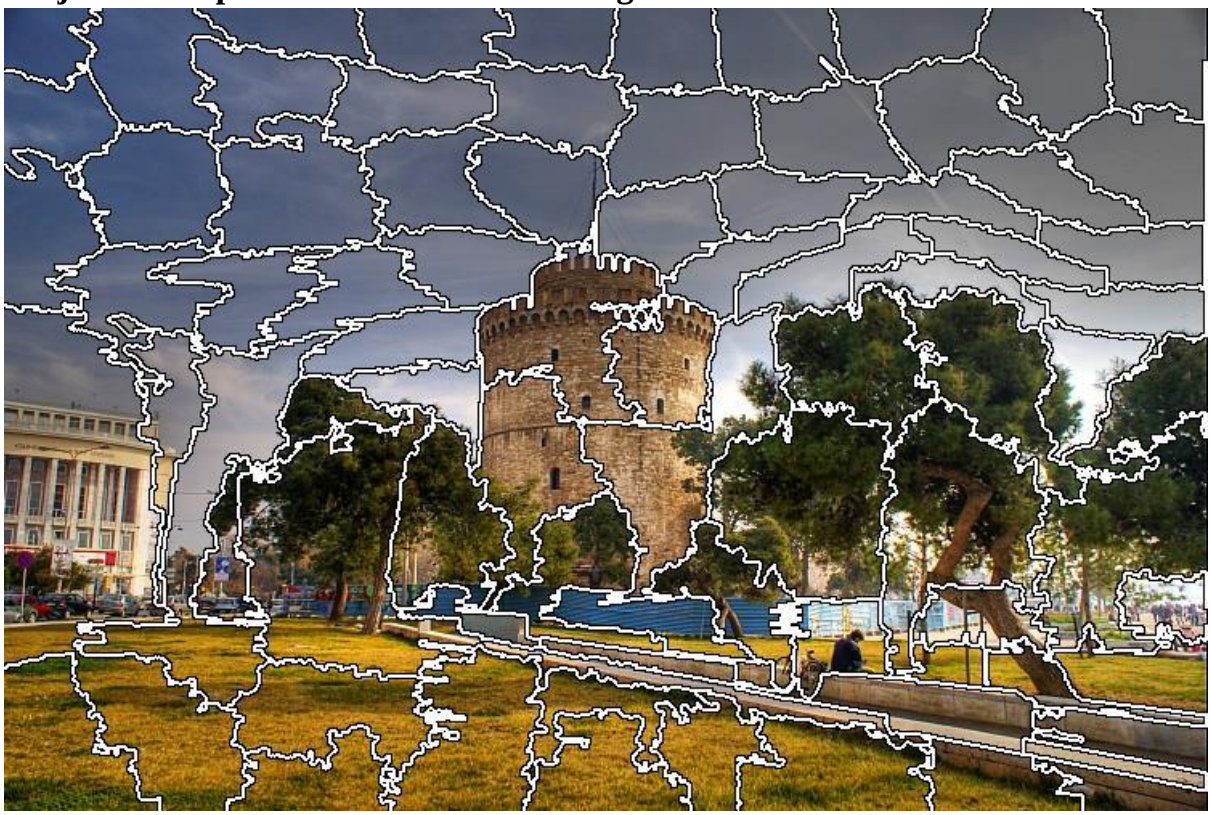
Cluster 5: 10 images

## Project 2: Object Segmentation

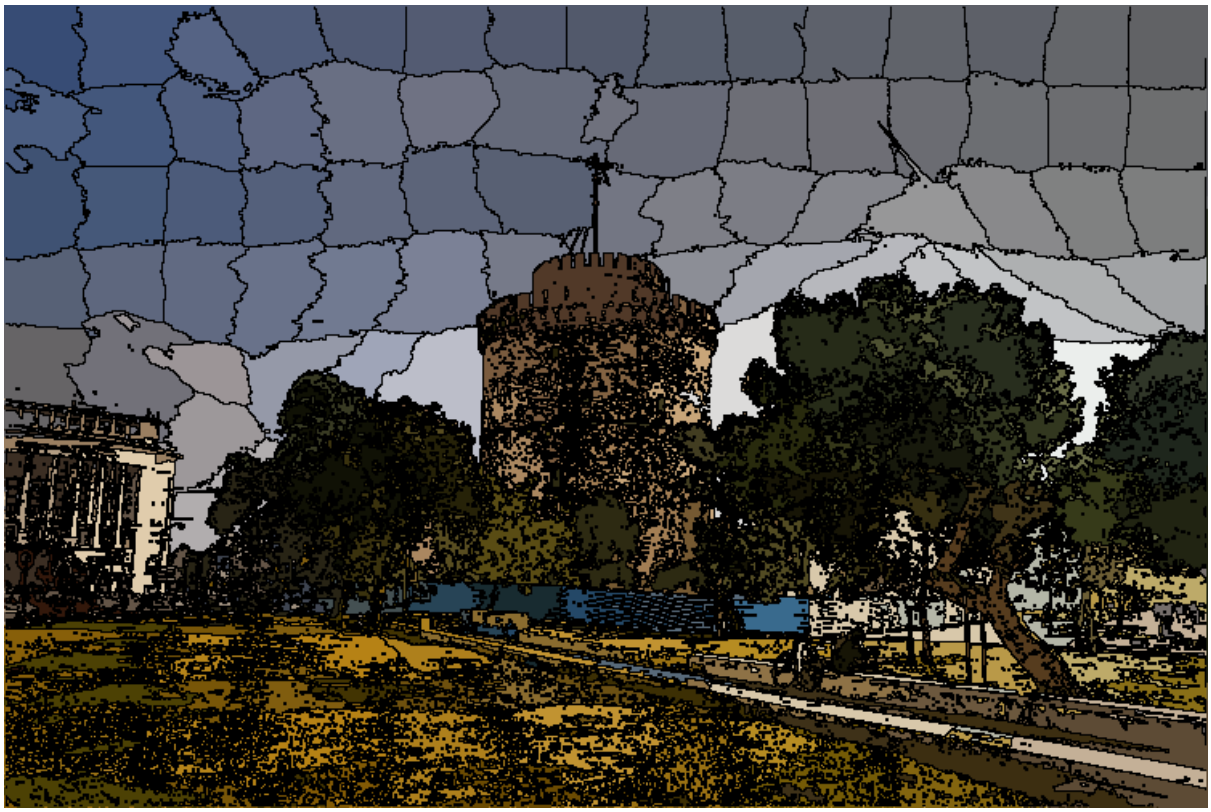


Obtained images

## Project 3: Simple Non-Iterative Clustering



SLIC GUI version



SLIC Assignment 4 implementation

SNIC Implementation

# Conclusions and Future Scope

This report is a combination of three projects performed as part of the Optional End of Course Projects in CS 558 – Computer Vision. Each project showed a different domain in the realm of Computer vision and a lot of understanding was gained in the process. However, the results are far from perfect and there is scope for improvement.

In project 1, there should've been four clusters each pertaining to the respective image set but five were obtained. This requires in-depth review on the approach of using color histogram as the image descriptor. Any other feature descriptor or a combination of multiple descriptors and a robust classification algorithm would yield better results.

The scope of improvement is maximum for the second task. The output image was far from perfect and the foreground image was not segmented properly. A different approach for clustering could be performed and evaluated to see if it gives a better result, preferably the use of Mean-Shift algorithm or Hierarchical Agglomerative Clustering.

Finally, performing Simple Non-Iterative Clustering was a challenging, yet rewarding, experience. The use of priority queue and a non-iterative approach proved to be fasted that using SLIC. The scope of improvement in this task would be to implement the SNIC-based polygonal partitioning and compare with the results of SLIC and other algorithms.