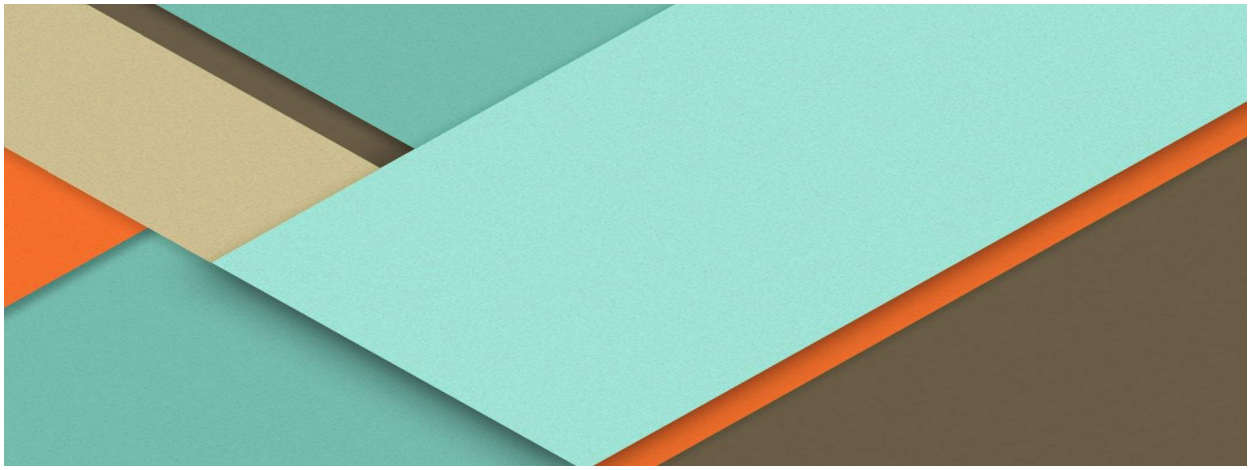

Road Extraction Using K-Means Clustering and Morphological Operations

Project ID: 1 - Team Firefly

Mehul Arora (2021701033)
Neha S (2021702012)
Radha Krishna B (2020801013)
Suvadeep Maiti (2021702021)
Umamahesh J (2019900058)

Github Link :https://github.com/Digital-Image-Processing-IIITH/dip-project-firefly

Problem Statement :

Road extraction plays a very important role in vehicle navigation systems, urban planning, disaster management and traffic management systems. In this project, we will implement road extraction based on the paper " Road Extraction Using K-Means Clustering and Morphological Operations " by Rohit Maurya, Dr. Shalini Singh, Dr. P.R Gupta, Manish Kumar Sharma.

The road extraction involves the two main steps: the detection of roads that might have the other non road parts like buildings and parking lots followed by morphological operations to remove the non road parts based on their features. We use the K-Means clustering to detect the road area and may be some non road area. Morphological operations are used to remove the non road area based on the assumptions that road regions are an elongated area that has the largest connected component.

Test Image | Expected Output Image | Reference Image

Goals and Approach :

Goal of this project is road extraction from a digital image.

The basic approach is as follows and has four major sections:

- K-Means clustering based segmentation is used to find road clusters.
- Cluster which has the longest connected component is identified as a road cluster, since roads usually appear as elongated regions.
- Morphological operations are used to filter the areas which have similar features as roads, like buildings, parking lots. The concept that roads appear as elongated areas and others are usually open areas, is used for filtering.
- Result evaluation can be done by comparison with a manually plotted high quality reference.

K-Means clustering:

It is an unsupervised clustering algorithm that classifies the input set of data into a specified number of clusters, based on their distance from each other. A Euclidean distance metric needs to be provided to specify how close two objects are to each other. The points are clustered around centroid $\mu_i \ \forall \ i = 1,...k$ which are obtained by minimising

$$V = \sum_{i=1}^{k} \sum_{x_j \epsilon S_i} (x_j - \mu_i)^2$$

where there are k-clusters S_i and μ_i is the centroid intensity and xj is the intensity of a point in the cluster S_i.

Road Cluster Identification:

Road clusters need to be identified from the obtained clusters. Considering that roads appear as long elongated segments, the cluster which has the longest connected components is chosen as the road cluster.

Road Cluster filtering:

Following operations are performed:

1. Dilation - to fill any gaps between road pixels
2. Non-road area removal - This involves steps:
    - Find connected components
    - Compute area of each component
    - Remove non-road parts, i.e. small objects.

Result Evaluation:

This is done using performance metrics like completeness, correctness and quality measure evaluation.

Expected Deliverables:

- Collection of dataset and preprocessing to create reference images.
- Performing image segmentation using K-means clustering.
- Road cluster identification for the clustered image.
- Road cluster filtering (Dilation and Non-road area removal).
- Performance metrics (Completeness, Correctness and Quality) for Result Evaluation.
- If feasible in the time constraints, we can explore methods to locate road networks in complex settings like urban areas.

Dataset:

We plan to use our own customized dataset for finding the road networks and then testing with the reference image. The dataset will be fabricated from Google maps. The test images will be taken directly by turning off the labels in Google maps. The reference images will be preprocessed by taking the transit images and converting them grayscale after preprocessing.

The dataset will consist of 8 test images and their corresponding reference images. Out of 8, 5 images will be taken from rural areas with relatively simpler road networks. The other 3 images will be of complex nature, with road networks in urban areas, tangled with buildings and water bodies.

Milestones & Timeline

Timeline	Milestones
5th Nov	Project Allocation
Week-0	Paper relevant work reading, discussions and initial project layout
9th Nov	Project proposal submission
Week-1	Data collection, pre-processing, Implementation of K-means Clustering
18th Nov	Mid evaluation
Week-2	road cluster identification Road cluster filtering, Creating reference images
Week-3	Result evaluation, PPT Preparation
1st Dec	Code, PPT submission
2nd Dec - 4th Dec	PPT presentation