Depth-Map Generation

Project - 29

Team Members:

J.S.S Sravan Chandra - 201501025
Kamineni Vikas - 201501092
Avinash Vadlamudi - 201501164

Abstract:

To estimate depth information from a single view of rgb image. This method is based on the new image classification technique able to classify digital images as indoor, outdoor with geometric elements or outdoor without geometric elements. Using the information collected in the classification step a suitable depth map is estimated. The proposed technique is fully unsupervised and is able to generate depth map from a single view of the scene. This method requires less computational effort.

Introduction:

3D images have become more and more popular in everyday life. This paper implements a method based on image classification technique able to classify digital images as indoor, outdoor without geometric elements and outdoor with geometric elements. The input image is processed by the following steps:

1) Color based Segmentation
2) Rule-based region detection to find specific areas
3) Image classification
4) Approximated Depth map estimation

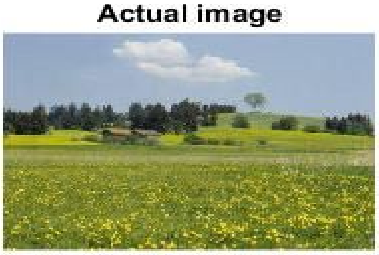
Method:

Color based segmentation:

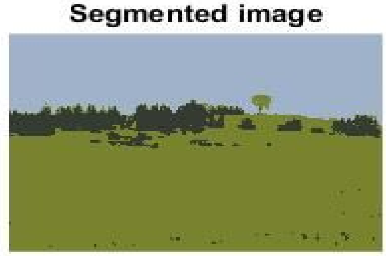
First, we need to do color segmentation on the image using mean-shift algorithm. In this algorithm , we initialise a random data point which is not clustered before as mean for a spatial bandwidth and then we find the mean of all the points around the mean point which fall in the spatial bandwidth region and then we shift the mean to this computed mean and iterate the method till it converges. We repeat this until each data point comes under any of the clusters.

Benefit of mean shift algorithm over k-means is that in k-means we need to give the k value which indicates how many clusters need to formed, but in mean shift algorithm, number of clusters depends on the given spatial bandwidth, so we don’t need to input the number of clusters.

Actual image



Segmented image



Regions Detection:

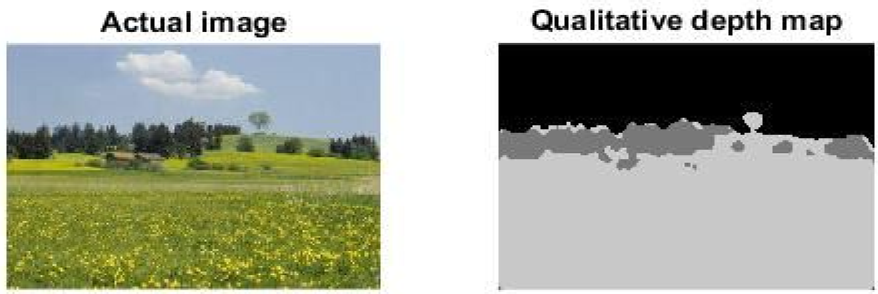
The semantic region detection can be based on color-based rules aimed to characterize specific regions such as: Sky, Far Mountain, Near Mountain, Land and Other.

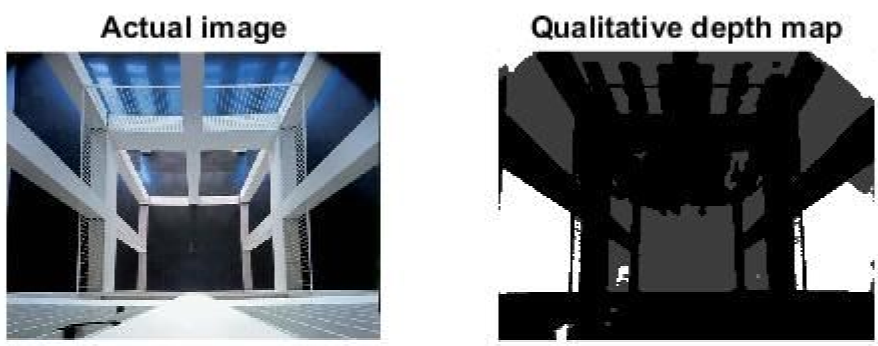
For this we need to perform the following steps:

1) Apply a 5X5 median filter to the segmented image
2) Compute HSV image from segmented image
3) The above specific regions are detected by following color-based rules. These rules are made by experimenting with specific set of images.(Actually, the paper has implemented these by training on a data set, which we don’t have access to)
4) After this , each region is assigned a specific gray value based on following order:

Gray(Other) > Gray(Land) > Gray(Near Mountain) > Gray(Far Mountain) > Gray(Sky)

5) Apply median filter again to remove the outliers in the above classified image.


Actual image
Qualitative depth map


Actual image
Qualitative depth map


Actual image
Qualitative depth map

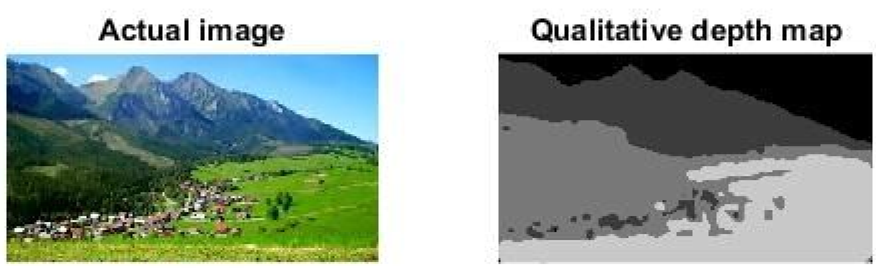

Actual image
Qualitative depth map

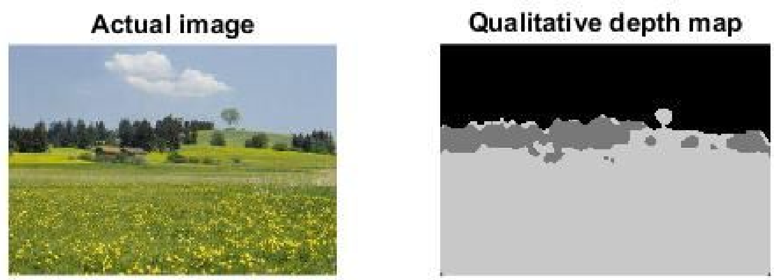
Image Classification:

Now we need to classify the image into *outdoor without geometric elements*, *outdoor with geometric elements* or *indoor*.

We do this by selecting some column regions in the final obtained qualitative depth map and find the number of region changes while we traverse the column from top to bottom. If the top region is sky and if the number of region changes is less than a particular threshold, we increment the count of R by 1 or else if the sky is top region we increment the count of R2 by 1.

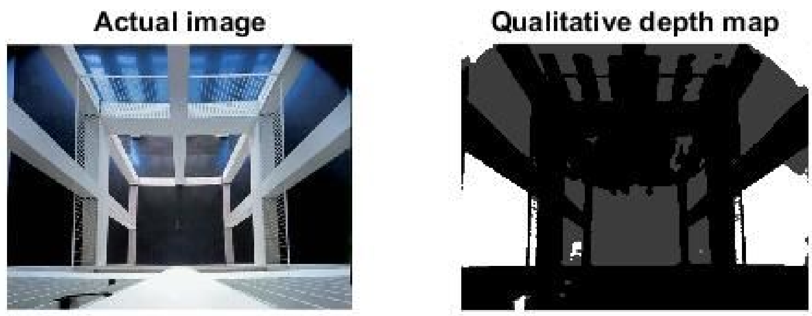
Now if the R value is greater than k1*N, where N is the number of analyzed column regions and k1 is a threshold in (0,1) , we classify the image as outdoor without geometric elements. Otherwise if the value of R2 is greater than k2*N where k2 is a threshold in (0,1) , we classify the image as outdoor with geometric elements, else it is classified as indoor.

The values of k1 and k2 are obtained by experimenting on a specific set of images.(These values are also not exact because our data set is limited)



This image is classified as **outdoor without geometric elements or landscape.**

Above image is classified as **outdoor with geometric elements**.



Above image is classified as **indoor**.

Above image is classified as **outdoor without geometric elements**.

Geometric Depth map generation:

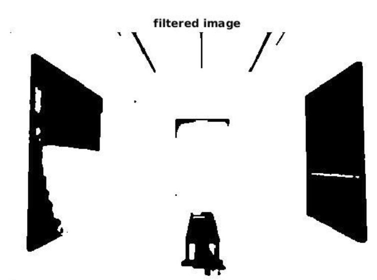
To generate the geometric depth map we followed a series of steps

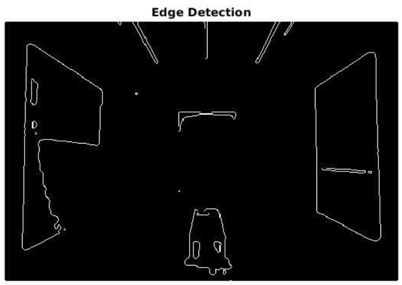
1) **Vanishing point detection:**

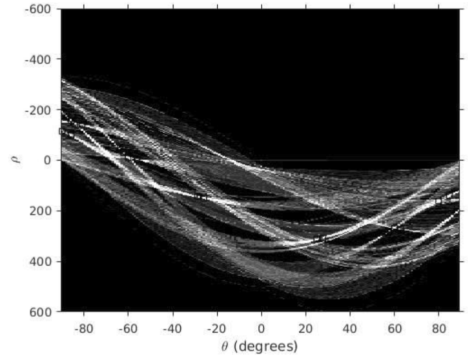
If the image is outdoor without geometric elements we consider vanishing points as the lowest point of intersection between land ∪ other and other regions (x_b,y_b) the coordinates of the Vanishing point is fixed to $(W/2,y_b)$.

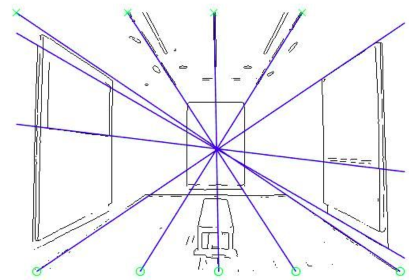
Otherwise

1. Edge detection using a 3x3 Sobel masks. The resulting images, I_{Sx} and I_{Sy} , are then normalized and converted into a binary image I E, eliminating redundant information.
2. Noise reduction of I_{Sx} and I_{Sy} using a standard low-pass filter 5x5.
3. Detection of line through hough lines by keeping voting threshold.
4. Compute of intersection between each pair of straight lines.
5. The Vanishing Point is chosen as the intersection point with the greatest number of intersections around it, while the vanishing lines detected are the main straight lines passing close to Vanishing Point.









2)**Gradient Plane generation:**

- Xvp<=0 AND (H-1/W-1) * Xvp<Yvp<-(H-1/W-1) * Xvp+H-1 Left Case
- Xvp>=W-1 AND -(H-1/W-1) * Xvp+H-1<Yvp<(H-1/W-1) * Xvp Right Case
- Yvp<=0 AND (W-1/H-1) * Yvp <= Xvp <=(W-1/H-1) * (H-1-Yvp) Up Case
- 0<Xvp<W-1 AND 0<Yvp<H-1 Inside Case

  Where Xvp and Yvp are X and Y coordinates of vanishing points.

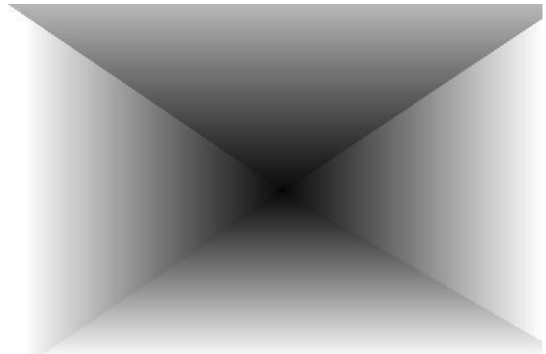
3) **Depth gradient Assignment:**

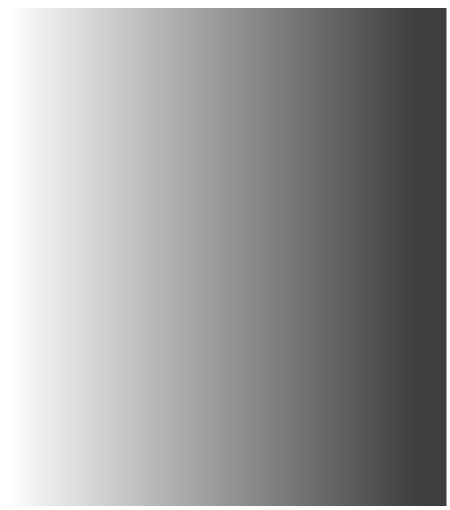
Two main assumptions are used:

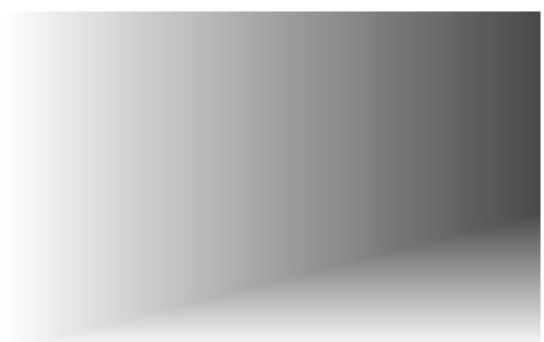
1. Higher depth level corresponds to lower grey values;

2. The vanishing point is the most distant point from the observer (this assumption is almost always true).

The depth levels value is approximated by a linear variation







Fusion of Geometric and Qualitative Depth Map:

Now , we generate the final depth map using the above two depth maps according to the following conditions.

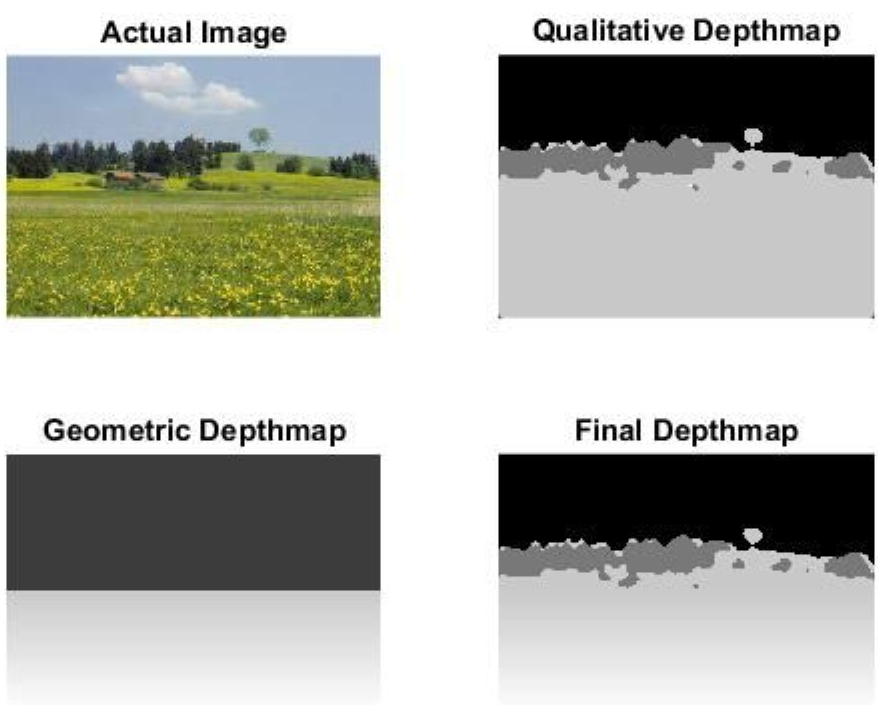
Let $M1(x,y)$ be the Geometric Depth Map
$M2(x,y)$ be the Qualitative Depth Map

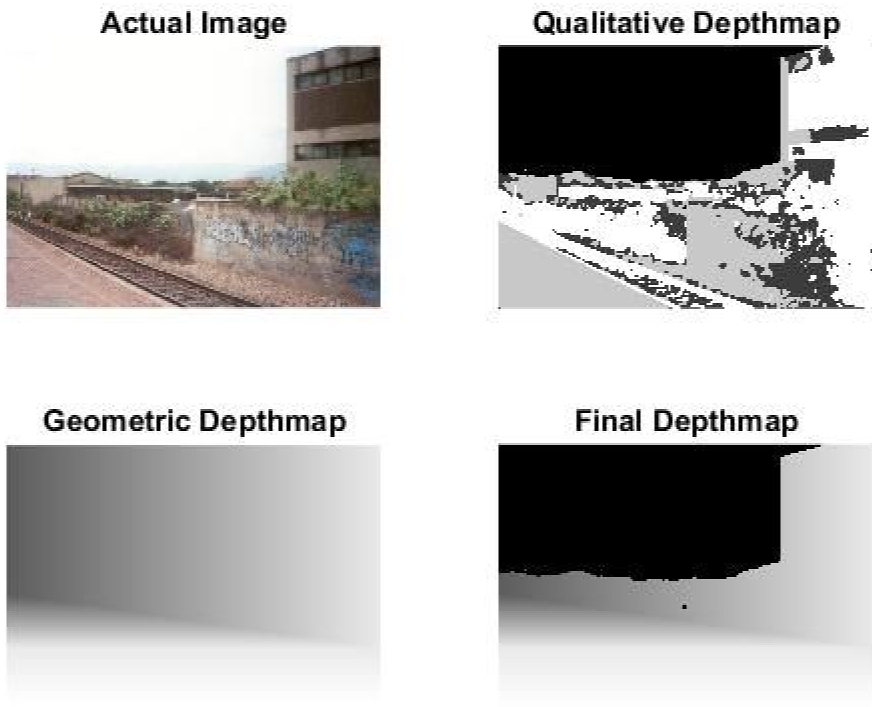
The Fusion depends on the image classification:

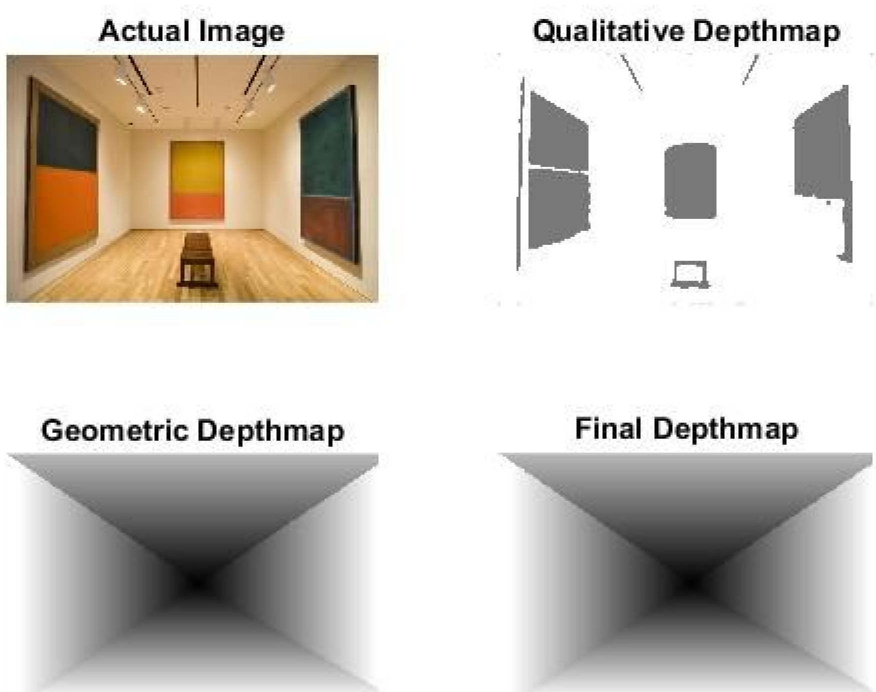
1. If the image belongs to the indoor category then $M(x,y)$ coincides with $M1(x,y)$:
    a. $M(x,y) = M1(x,y)\ \forall (x,y) : 0<=x<=W-1\ 0<=y<=H-1$.

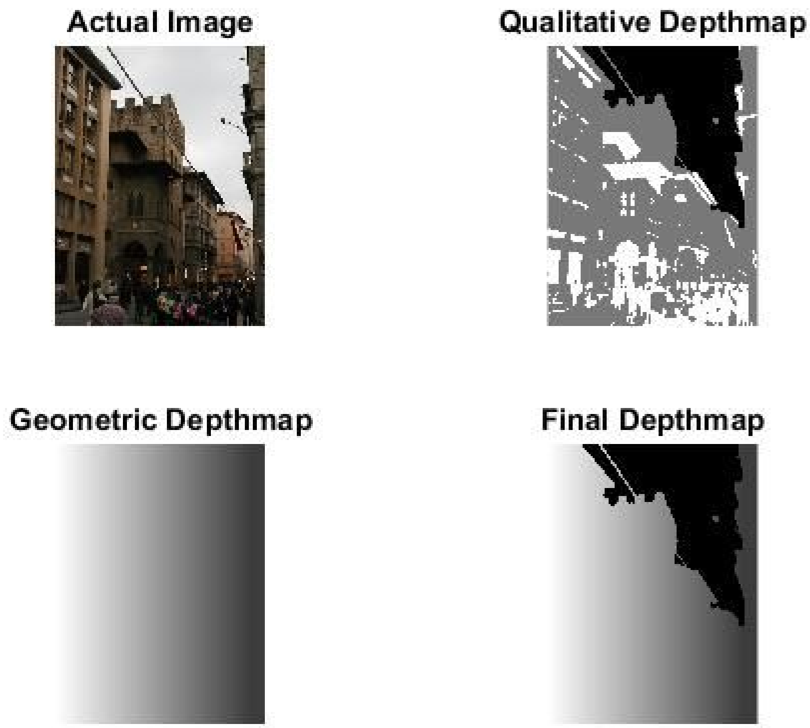
2. If the image is classified as Outdoor with absence of meaningful geometric components (landscape) then the image M(x,y) is obtained as follows:
    a. $M(x,y) = M1(x,y) \forall (x,y) \in Land$ and $\forall (x,y) \in Other$
    b. $M(x,y) = M2(x,y) \ \forall (x,y) \notin Land$ and $\forall (x,y) \notin Other$
3. If the image is classified as Outdoor with geometric characteristics then the image M(x,y) is obtained as follows:
    a. $M(x,y) = M2(x,y) \ \forall (x,y) \in Sky$.
    b. $M(x,y) = M1(x,y) \ \forall (x,y) \notin Sky$.

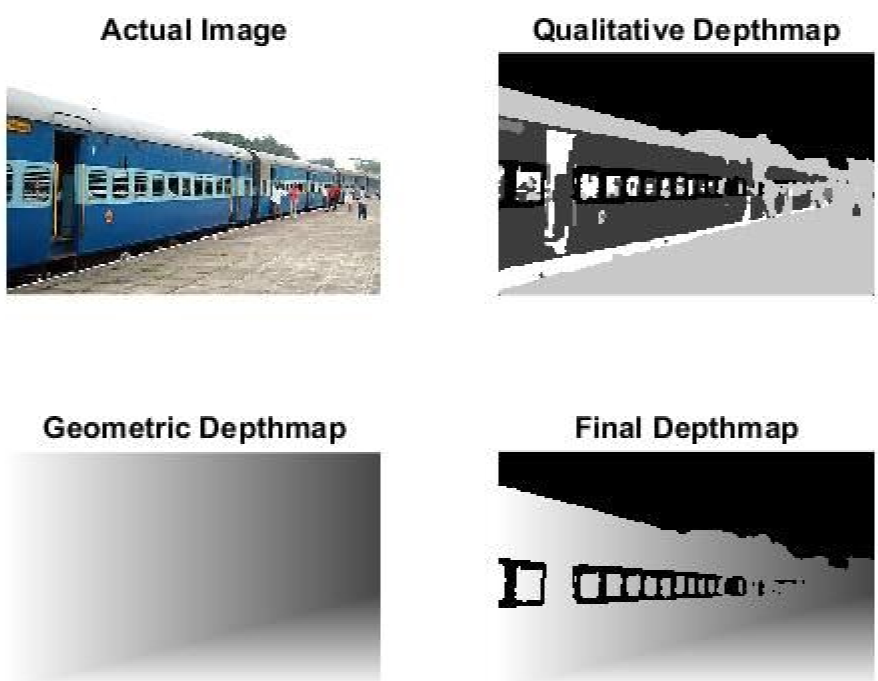
Results:




Actual Image
Qualitative Depthmap
Geometric Depthmap
Final Depthmap

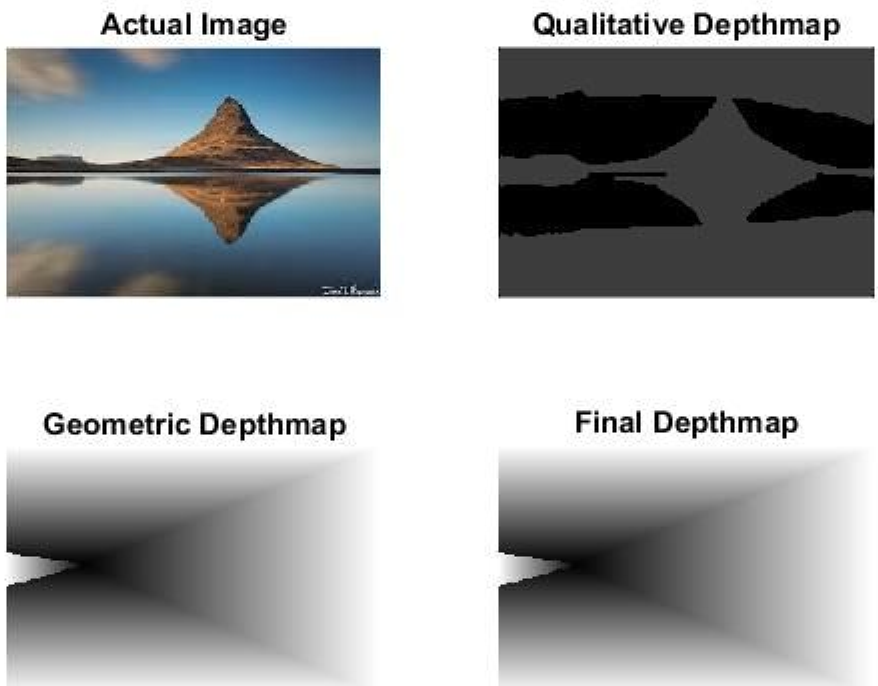

Actual Image
Qualitative Depthmap
Geometric Depthmap
Final Depthmap


Actual Image
Qualitative Depthmap
Geometric Depthmap
Final Depthmap

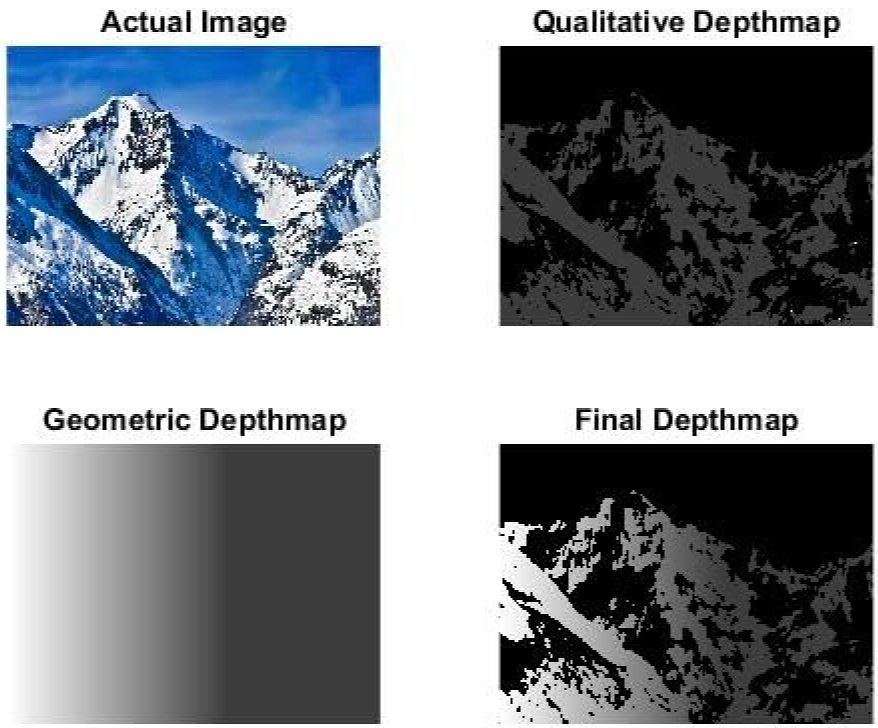

Actual Image
Qualitative Depthmap
Geometric Depthmap
Final Depthmap

Limitations:

1) We don’t have the training data set so we cannot get the appropriate color values which can classify all the regions correctly for all types of images and Geometric Depth Map heuristics.
2) Similarly we cannot get the appropriate k1 and k2 values which could classify the image correctly for any image.



For this image, we can’t segment the water clearly and also some part of sky region is not detected as sky region. Also, the Geometric depth map is not correct because we don’t have huge train data to estimate the proper heuristics.



In this image, regions aren't properly segmented.