CSE 573 - Computer Vision and Image Processing (Write - up)

- **Name:** Sujit Singh
- **UBIT Name:** sujitjit
- **UBIT:** # 50247206

Section 1.0

There are four types of filters created through createFilterBank function. These Filter Banks along with their properties are explained below:

1. **Gaussian Fitler:** Gaussian Filter is used to reduce the noise and details from the image. The scales of the Gaussian filter depends on the value of sigma. In Gaussian Scale the increase in scales smoothens the image and makes it more blur.
2. **Laplacian of Gaussian:** Laplacian of Gaussian detects the sudden change in intensity of the image, generally edges. The increase in the scale of the logScale diminishes the edges of the output image
3. **dxScale:** This scale enhances the edges along the x direction. With the increase in the scale, it increases the width of the edge making it more blur.
4. **dyScale:** This scale enhances the edges along the y direction. With the increase in the scale, it increases the width of the edge making it more blur.

Section 1.1

The below Figures, Fig 1.1 shows the actual image, and Fig 1.2 depicts the representation when filter-bank (Set of 20 filter bank) is applied to the image. The depiction is done using montage function in matlab.

Fig 1.1 Actual Image

Fig 1.2 Image after applying filter

Section 1.3

The representations from Fig 1.3 – Fig 1.8 shows computation of visual words in the dictionary

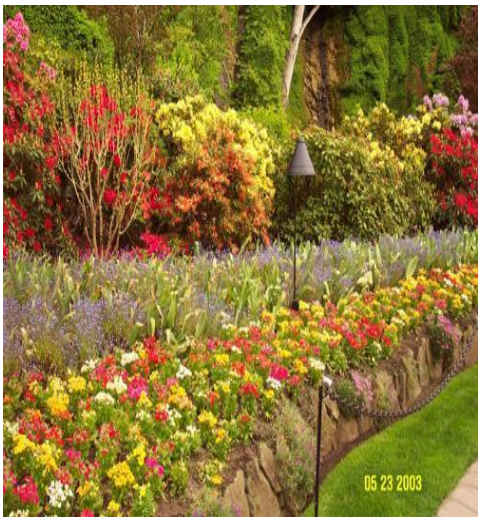

Fig 1.3 Actual Image

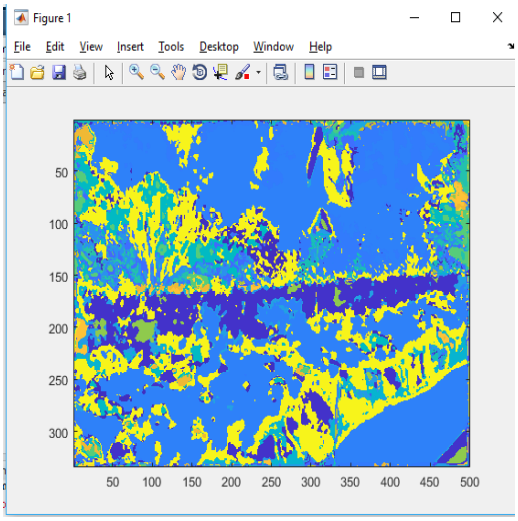

Fig 1.4 Visual word map

(**Fig-source :** garden/ sun_bavqqcxmiixxeirr.jpg)

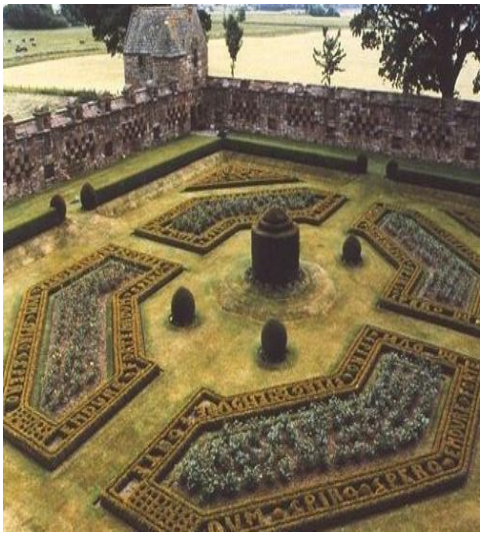

Fig 1.5 Actual Image

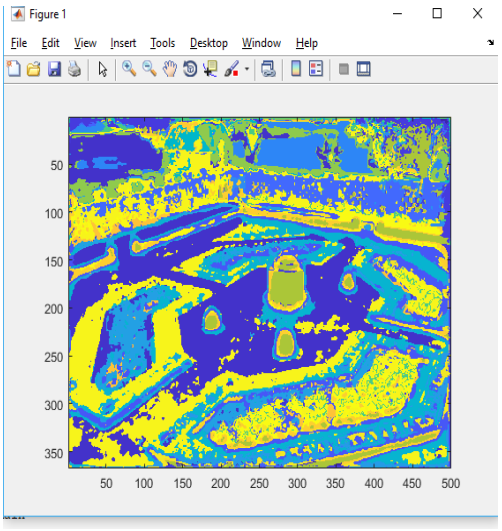

Fig 1.6 Visual word map

(**Fig-source :** garden/ sun_aanfdszdggeqrclh.jpg)

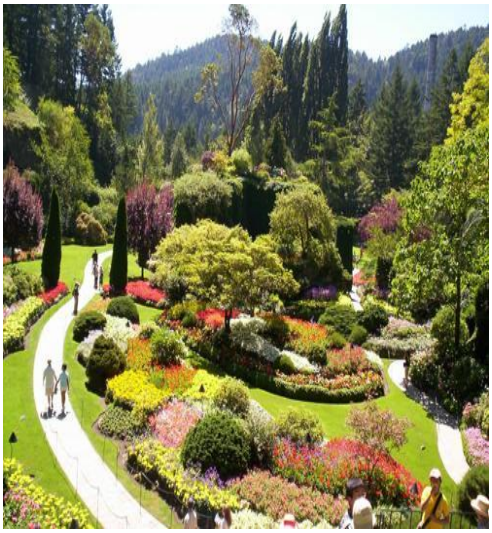

Fig 1.7 Actual Image

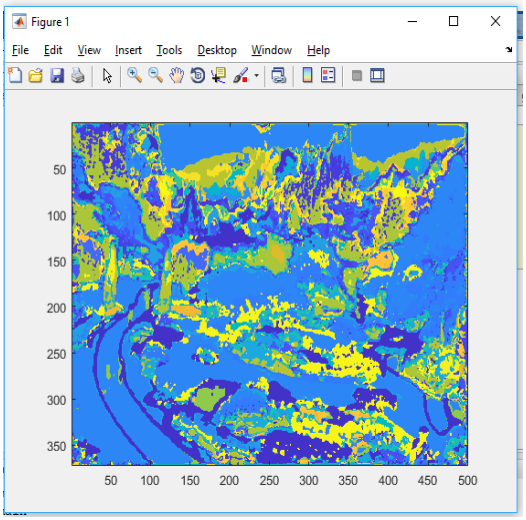

Fig 1.8 Visual word map

(**Fig-source :** garden/ sun_bgwhceaudanflans.jpg)

From the above figures, the complex RGB color pallets are mapped into a color group. This color group represents a particular domain in the visual dictionary such that it maps a respective texture in the RGB image to the corresponding code in the visual dictionary. Hence, a characteristics belonging to a particular category are clustered together. For eg: all the grasses in the garden will be clustered into a similar color pallet. Hence, it becomes easy for the system to understand a particular feature / characteristic of the segment of an image belonging to same category.

A further justification of visualization lies to the fact that an RGB image will have 256 x 256 x 256 possible combination of colors. It becomes a very tedious task for comparing histogram of thousands of image with such a high possibility color pallet. It increases the computation time and lowers precision. Visualization helps to reduce this enormous possibility by clustering the color pallet into k clusters using k-means. The effect of visualization is seen at various steps of our program making it compatible to train, test and compare 1600 images of such a high resolution. Hence, visualization is an important step in building a recognition system.

Section 2.5

The confusion matrix depicting the predicted and the guessed classes from the test set.

7	3	1	0	3	2	1	1
4	6	0	1	4	1	1	2
0	0	16	0	0	4	0	0
4	3	1	15	5	0	0	0
1	1	1	0	5	1	1	0
2	4	1	2	1	7	4	3
1	3	0	0	1	3	12	6
1	0	0	2	1	2	1	8

Accuracy of the recognition system is 47.5%

Section 2.6

From the confusion matrix obtained in the result, we can see that lowest two classes that are most incorrectly classified are second (value of the diagonal element corresponding to second column is 6) and fifth (value of the diagonal element corresponding to fifth column is 5) i.e. computer_room and library respectively. In both the classes, we find the environment of the closed room where we have furniture, tables, desks, racks etc. This corresponds to the same color pallet making it difficult for the recognition system to classify appropriately. On the other hand, the classes like garden and ice_skating have unique characteristics / features in this set of classes. Hence, they have higher possibility of getting a correct guess.

Section 2.7

We can bring about improvement in this recognition system by working on certain parameters such as histogram intersection, extracting features etc. There are various methods of histogram intersection[1] from among which we have used conventional histogram method. From the experimental results shown[1] the Color Ratio Gradient (CRG) Method gives the fastest result and the Gaussian Weighted Histogram Intersection (GWHI) Method gives the best precision. Color Edge Cooccurrence Histogram (CECH) Method gives the combination of both fast results and precise outcomes. We can implement the above histogram intersection techniques to improve the accuracy of our system in future.

References:

[1] W. Jia, H. Zhang, X. He and Q. Wu, "A Comparison on Histogram Based Image Matching Methods," *2006 IEEE International Conference on Video and Signal Based Surveillance*, Sydney, Australia, 2006, pp. 97-97. doi: 10.1109/AVSS.2006.5