**Chapter 2**

**Fundamental Concepts of Background Subtraction, K-means Clustering and Image Contouring**

For any application requiring Image processing of video, background subtraction and modeling, clustering, checking the similarity of images, de-noising and contouring are all very basic pre-requisites. This chapter gives an overview of all these concepts that will help elucidate the design and implementation details covered in the chapters that follow.

**2.1 Background Modeling**

When video is captured in real life scenes and objects in the scenes are analyzed, there are two types of objects. Background objects and foreground objects. For example, consider a typical traffic scene. Vehicles, pedestrians and maybe even some birds are all classified as objects that move. The objects such as buildings, trees are all stationary. Therefore they are considered background objects. While this may seem quite simplistic at first, it is not so straight forward. Now, imagine that there are people in the background that are talking using hand gestures or that there are tree branches, the position of which has been displaced by a gush of wind. Even though these objects are strictly speaking moving objects, they are considered non-static background objects. Therefore, foreground objects are objects of interest with respect to a specific application, those that move a significant amount against a modeled background. Background Modeling refers to the process of constructing this background. There are several ways to accomplish this. We have used the running average method, where every frame is averaged at a rate determined by a regulator value. If that value is equal to a certain threshold, that threshold being decided by the domain of application, the non-static background objects are also included in the background so modeled [15][16]. There are other methods of background modeling such as the Gaussian Mixture Model method [23]. There are also lesser known but reasonably efficient methods such as the codebook construction method [24]. However, as stated above our aim is to achieve efficiency while maintaining simplicity. Hence, we have a adopted a statistical approach that is computationally economical and efficient.

**2.2 Background Subtraction**

Background Subtraction refers to the method of literally just subtracting all the frames constituting a video against a modeled background of the scene. Thus, all the moving objects or foreground objects are separated out and the result appears as a black background with white objects moving against it. There are various techniques that can be used to carry out background subtraction [17]. One very popular method is known as the Gaussian Mixture model which uses a mixture of K Gaussian distributions to model the foreground [1]. The result looks as shown in figure 2.1.

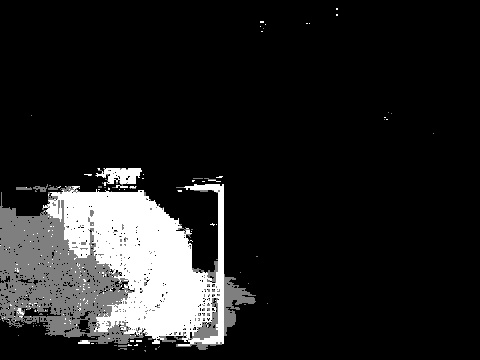


Figure 2.1: Background subtracted image using GMM

However, as is evident from the figure the Gaussian Mixture Model approach suffers from lack of detail and background noise. Our approach resolves both these shortcomings as will be explained in later chapters.

**2.3 K-Means Clustering**

K –Means Clustering refers to a method of clustering a dataset based on the Euclidean distance when the dataset is mapped on to a co-ordinate system. For example, consider that you want to group a sample of 10 people into 2 groups based on their average height. First all the 10 people are plotted on the XY plane against their height. Next, two random points with maximum separation are chosen as the centroids of the two clusters (since we need all the points to be grouped into two clusters). Of course, there are improved versions of the algorithm where the number of clusters is decided adaptively specific to the application [25][26]. But we don’t need this feature in our method and hence are going to decide the number of clusters as two. Another point among the eight remaining points is chosen. It is grouped along with the cluster that it is closest to. Subsequently, for that cluster, the new average is calculated by averaging the newly added point with the old one. In this manner, all the points are grouped into two clusters thus accomplishing our demarcation based on height. Figure 2.2 illustrates points before clustering and figure 2.3 illustrates points after clustering. X represents the person number and Y the height in meters.

Figure 2.2: Height of 10 boys in metres

Figure 2.3: 10 Boys in two clusters according to height

**2.4 Image Contouring**

In image processing often times we need to identify objects in the most computationally efficient way because image processing in general is highly computer intensive; therefore, it is beneficial to accomplish certain functions in less computer intensive ways whenever possible. For example, consider that in an image of still coins of different sizes, the size being representative of their values, you are asked to find the coin with the median value. In statistics, median means the middle value in the sorted set of values. If we were to perform this action using standard image processing techniques, we could threshold the image, find image gradients and compare the shapes to the shapes of numbers identified using a digit classifier. Another way of doing it would be to compare the coin images with a database of stored coin images using a structural similarity index. A third way would be to compare color histograms of areas in the image to those already obtained before we start processing. However, while all these tasks present a definite way to obtaining the desired result, they all involve complex image processing. Now, what if we could accomplish the same with a much simpler technique? Introducing Image contouring; it is the process of identifying contours or outlines of objects in an image. So in our coin problem if we could identify a set of contours, sort them by area after approximating them into standard polygonal shapes and then simply pick out the middle term in the set, we’re done! Likewise, Image contouring can be used to simplify a lot of problems in real life as well. In our method for example, we have used contouring to find the vehicle of interest in the frame. In a frame sized window, the object corresponding to the largest contour is most likely to be the vehicle! This is a reasonably safe assumption. Moreover, this method is highly computationally simple and economical. Figure 2.3 shows a contoured image of a car and few pedestrians.

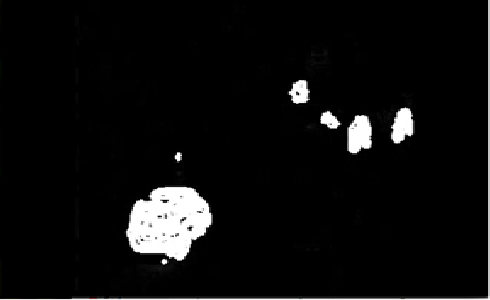


Figure 2.3: Contoured image of car and pedestrians

**2.5 Summary**

In this chapter, the basic concepts required for understanding the details of implementation of the project and the mathematics behind it has been briefly explained. In order to understand how these techniques are relatively advantageous, a background in basic image processing concepts used in this project is necessary. Also, K- means Clustering has been explained with respect to a random data set involving the heights of persons. However, how this applies to color quantization is explained in the following chapters.