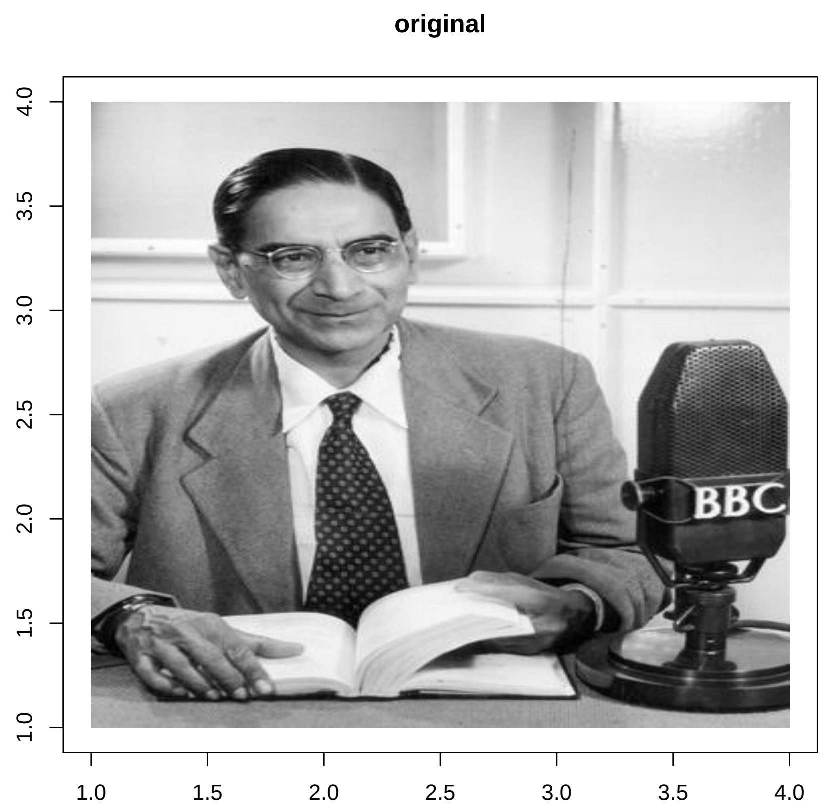
The first stage of our analysis requires us to read the image and store it as a vector/matrix,with the help of R we build ourselves a function which is capable of reconstructing an image from the vector/matrix.To check the performance of the function created by us , we render the image of P.C.Mahalonobis.

```
## Warning:  package 'jpeg' was built under R version 4.0.3
```

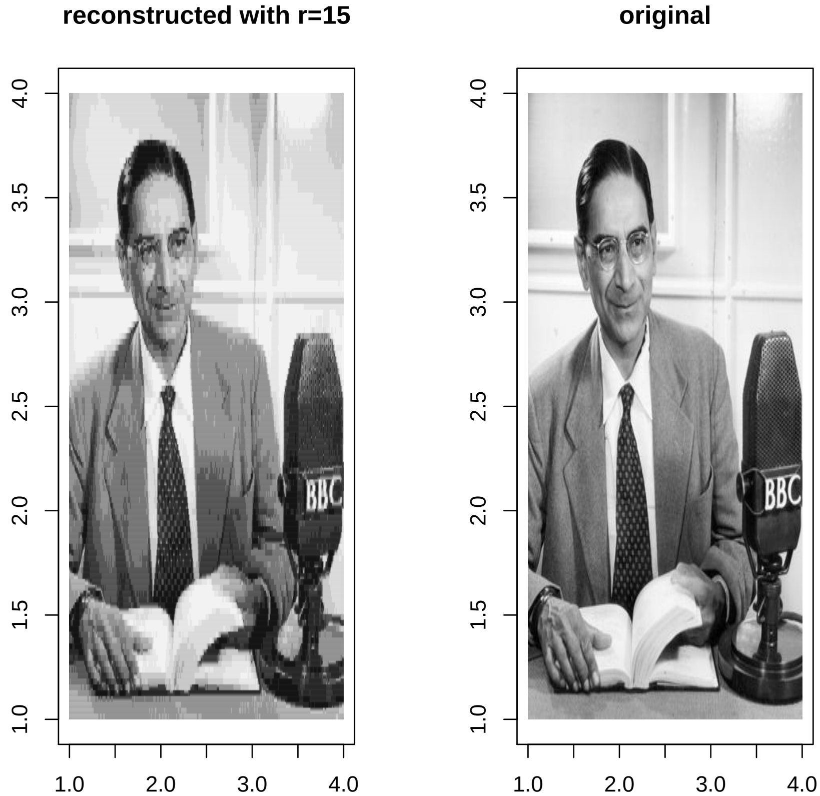


Keeping the technique of the image vector quantization as our quidelines we proceed to see how the choice of number of clusters and also the choice of number of adjacent pixcel values to be taken as Vi's impacts our analysis.

Before proceeding we first implement a particular version of the VQ method keeping q=4,r=15.

We also plot the original & as well as the reconstructed image side by side for comparisons and analysis.

```
## Warning:  did not converge in 10 iterations
```



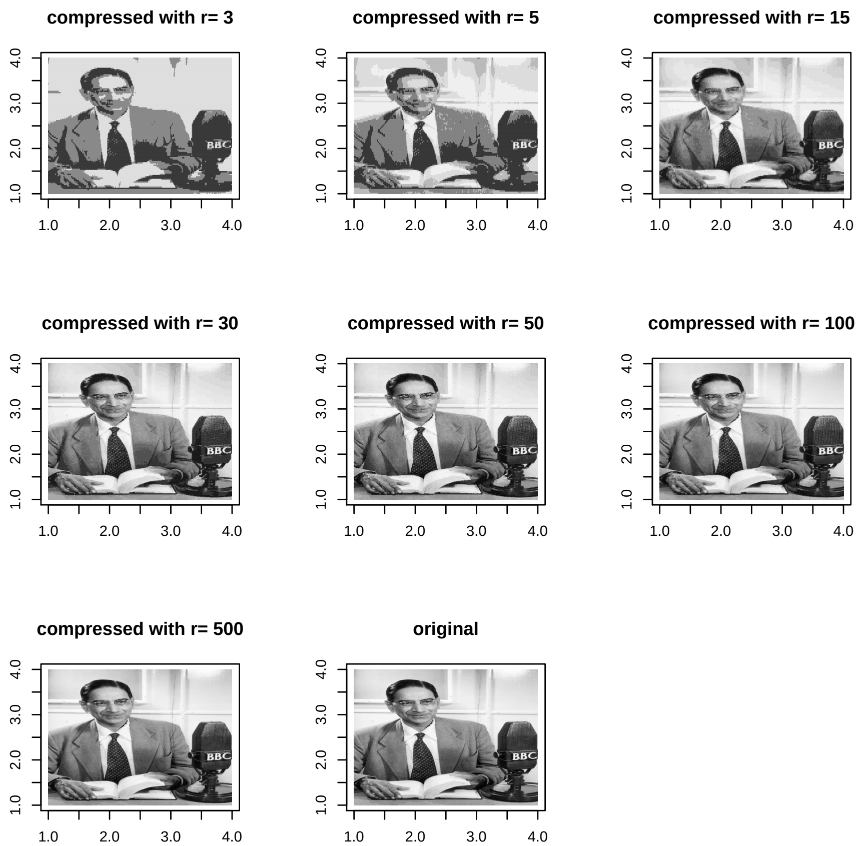
We can see taking r=15 ,i.e, number of clusters to be 15 the reconstructed image is quite good and most of the relevant details are captured and minute detailing are missing.Thus there exists a lot of scope regarding the reduction of dimensions aspect.Proceeding further we also take in account the euclidean distance between the original image and estimated reconstructed image as a measure of similarity/dissimilarity.

Now to run the VQ method for different r , keeping the q=4 and plot the images in a page to help us draw conclusions and compare.Also find the Euclidean Distance each time.

Reconstructed Images with different r:

```
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
```

```
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
```



We can see that with the increasing number of clusters the detailing in the picture approximately matches the original picture the images with $r>50$ almost have the same level of detailing as that of the original picture.While the image reconstructed with r =3,5 hardly have any detailing & noisy and only can potray that its a picture of P.C.Mahalonobis to only who can identify and more than an image it looks smudged painting.

If we look at the Euclidean distance for r = 3,5,15,30,50,100,500 respectively

```
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
```

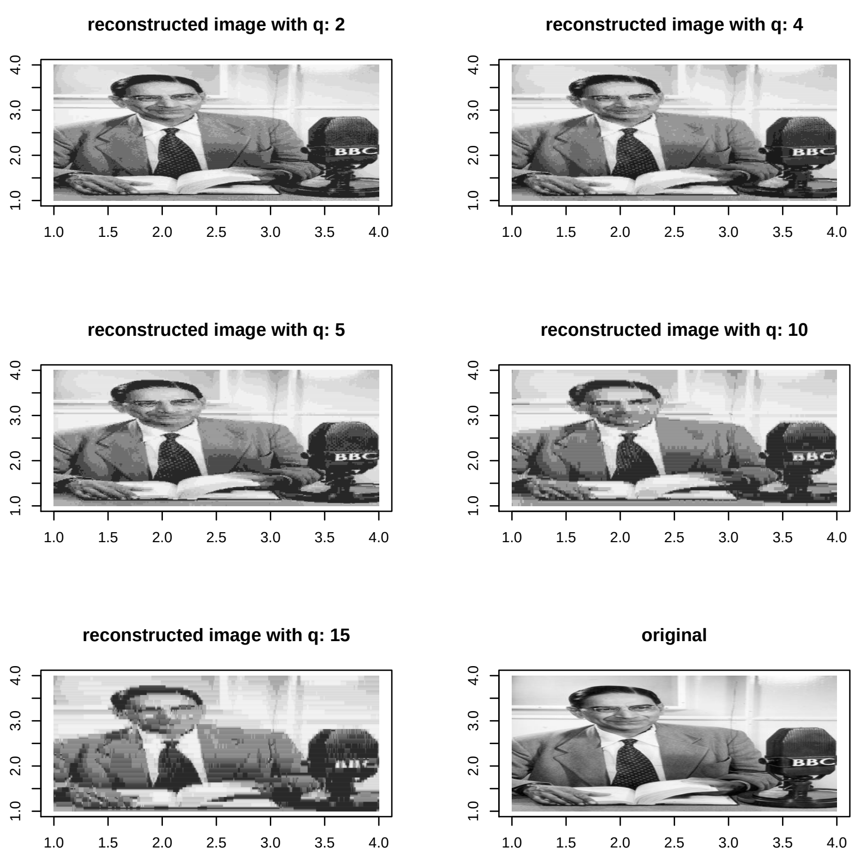
```
## Warning:  did not converge in 10 iterations

## [1] 39.80147 32.53354 25.76398 22.97205 21.26647 20.05834 18.44380
```

We can see that with increasing r the Euclidean Distance decreases ,i.e, the reconstructed image starts to become a well representation of the original image.

Now to run the VQ method for different q, keeping the r=15 and plot the images in a page to help us draw conclusions and compare.Also find the Euclidean Distance each time.

Reconstructed Images with different q:

```
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
```



We can see that with the increasing number of q ,i.e, size of Vi the detailing in the picture drastically detoriates with respect to the original picture & becomes tremendously noisy, the images with higher q have little to no detailing as that of the original picture.While the image reconstructed with q=2,4 still preserves detailing & lower q instead produces an image almost same as original picture.

If we look at the Euclidean distance for q = 2,4,5,10,15 respectively

```
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations
## Warning:  did not converge in 10 iterations

## [1] 13.20416 27.24943 22.15324 26.90501 49.10364
```

We can see that with increasing r the Euclidean Distance increases ,i.e, the reconstructed image starts to deviate from the original image with a lot of noise.But for q=4 it is larger than than q=5, but it is for totally a different reason, as for q=4 after the reconstruction ,image turns out to be of 428*346 instead of 430*346 as 430 is not divisible by 4.So this creates larger euclidean distance.

Conclusions:

The above problem actually proves the intuitive idea about the analysis. If we consider the image data is clustered into different classes and we increase the number of classes we are actually considering details of the image,we are denoting some points as a cluster if it differs marginally , where else for fewer cluster we try to cluster marginally different classes into one large class,this effect can also thought to be as an smoothing out effect.This is quite evident from the above problem.Also as we consider more clusters and thus acknowledge the small dissimilarities we approximately reach closer to the original image and thus we see a decrease in euclidean distance with increasing clusters.

Similarly the basic VQ method depends on the believe that the adjacent pixel values does not differ much or are quite uniform in nature.But as we increase the size of q we are believing that a larger number of adjacent pixel values does not differ much which is not that much of a valid argument and thus unlike r with increasing q we see that the details are being lost.Smaller q on the other hand gives the closest approximation of the original image as the concept of small q means that we believe only the closest pixel's have similar values and this assumption is valid which we have intuitively proved earlier.Thus smaller q gives better reconstructed image and hence smaller euclidean distance,& we see a increase in euclidean distance with increasing q.

So we must choose a pair of q & r such that our basic idea of ease of usability of the data and lesser storage capacity is met keeping intact the vital characteristics of the data(here it is the detail of the image).Thus the storage,ease is a trade of with the details of the image.