REPORT

1. Part 1

For each category, compute the mean image and the first 20 principal components. Plot the error resulting from representing the images of each category using the first 20 principal components against the category.

I have computed the following mean images for each of the 10 categories. On some of the images you can even barely discern the object in question, especially the automobile, frog, and horse. The images for the airplane and ship have a lot of what can be understood as the sky or water, respectively.

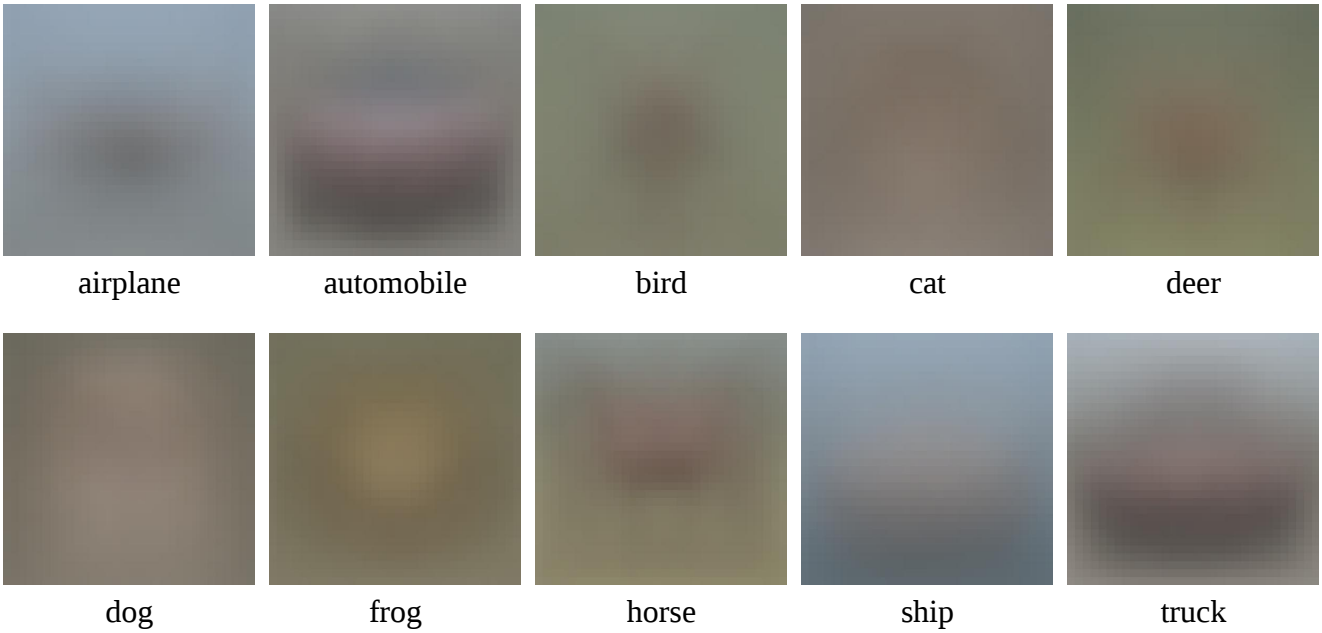


Fig. 1. Mean Images

I have also computed the absolute and relative errors of representing the images of each category using the first 20 principal components for that category. You can see that the shape of the bar charts is similar in both cases:

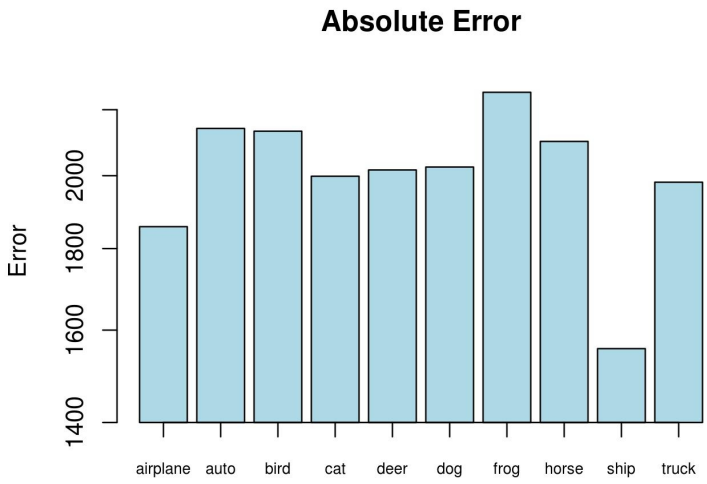


Fig.2a Abs. error of representing each category with first 20 PCs

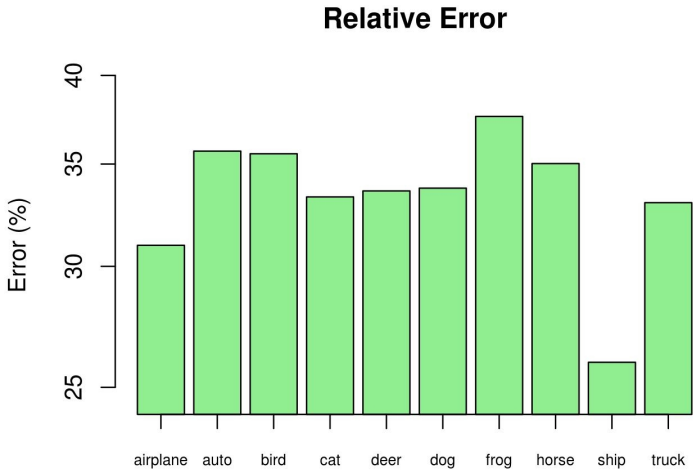
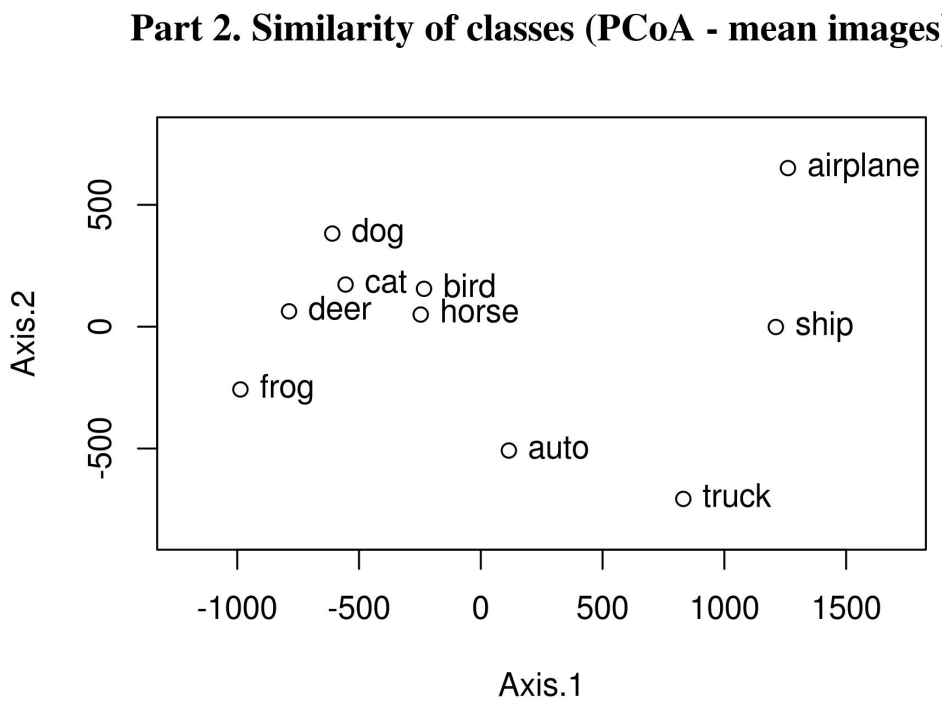


Fig.2b Rel. error of representing each category with first 20 PCs

Part 2

Compute the distances between mean images for each pair of classes. Use principal coordinate analysis to make a 2D map of the means of each categories. For this exercise, compute distances by thinking of the images as vectors.

The results are shown below. You can clearly see that all the animate subjects are grouped to the left of the zero on axis 1 with relatively small distances among them, and all the inanimate subjects are to the right with larger distances (bigger difference in images):



Part 3

Here is another measure of the similarity of two classes. For class A and class B, define E(A | B) to be the average error obtained by representing all the images of class A using the mean of class A and the first 20 principal components of class B. Now define the similarity between classes to be (1/2)(E(A | B) + E(B | A)). If A and B are very similar, then this error should be small, because A's principal components should be good at representing B. But if they are very different, then A's principal components should represent B poorly. In turn, the similarity measure should be big. Use principal coordinate analysis to make a 2D map of the classes. Compare this map to the map in the previous exercise? are they different? Why?

The results are shown on the figure below. The categories are located in a different way because a different similarity measure was used. However, you can still see that all the animate subjects are located very close to each other to the right from the zero on axis 1 while the inanimate subjects are to the left from that zero, although the distances between them are a lot bigger which tells us about how different their pictures are. The general trends of grouping and the magnitude of distances within the two groups are similar between parts 2 and 3 which can be construed as a sign of consistency of the results