DATA605 - Assignment 4

Nick Oliver

Homework 4

Problem

With the attached data file, build and visualize eigenimagery that accounts for 80% of the variability. Provide full R code and discussion.

Solution

I borrowed heavily from Diego Herrero's Face Recognition using Eigenfaces[1] and also used code from the Wikipedia page on Eigenfaces[2]

Load Libraries

Only using OpenImageR library to load images

```
library(OpenImageR)
```

Load images and extract data into a matrix

```
basePath <- '~/jpg/'
files <- list.files('~/jpg')
data <- matrix(0, length(files), prod(c(1200,2500,3)))

for (i in 1:length(files)) {
  img <- readImage(paste0(basePath,files[i]))
  r  <- as.vector(img[,,1])
  g  <- as.vector(img[,,2])
  b  <- as.vector(img[,,3])

  data[i,] <- t(c(r, g, b))
}
```

Convert vectors into a dataframe

```
shoes <- data.frame(x = t(data))
```

scale and center

```
scaled <- scale(shoes, center = TRUE, scale = TRUE)
```

[1] https://rpubs.com/dherrero12/543854/
[2] https://en.wikipedia.org/wiki/Eigenface

Calculate covariance

```
sig <- cov(scaled)
```

Compute eigenvectors and values

```
eig <- eigen(sig)
eigenvalues  <- eig$values
eigenvectors <- eig$vectors
```

Evaluate the number of principal components needed to represent 80% Total variance.

```
eigsum <- sum(eigenvalues);
csum <- 0
k80 <- 0
tv <- 0
for(i in 1:17){
   csum <- csum + eigenvalues[i]
    tv <- csum / 17
    print(tv)
    if(tv > 0.80){
      k80 <- i
      break
    }
}
```

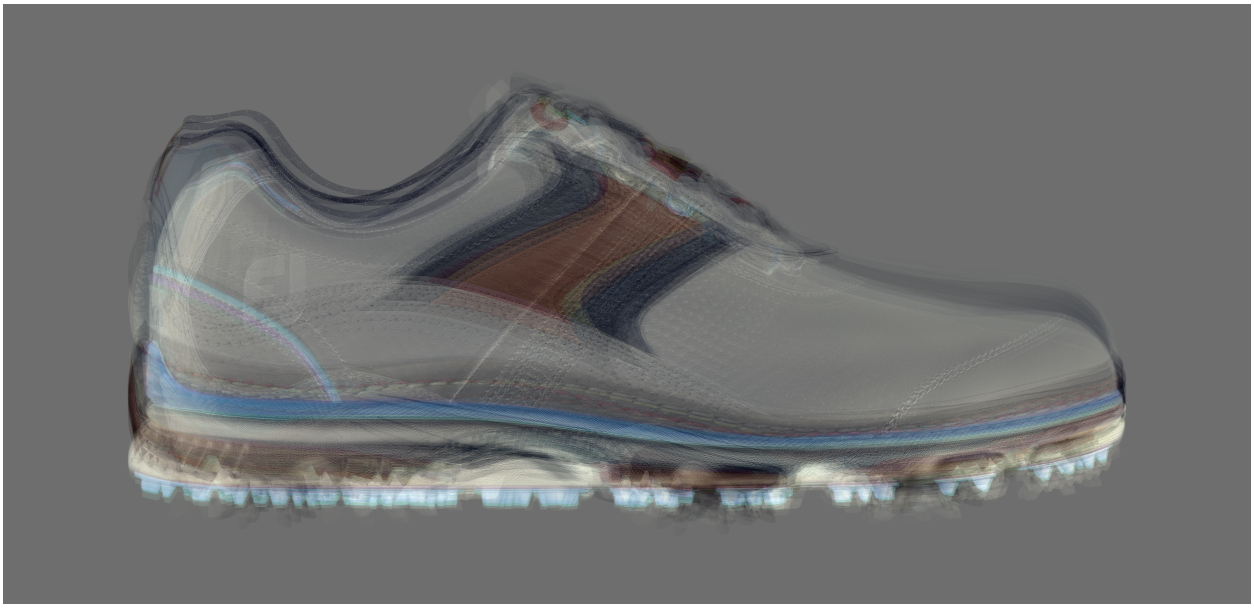
```
## [1] 0.6833138
## [1] 0.782474
## [1] 0.8353528
```

```
k80
```

```
## [1] 3
```

```
shoesScaled <- diag(eigenvalues[1:k80]^(-1/2)) / (sqrt(nrow(scaled)-1))
eigenShoes <- scaled %*% eigenvectors[,1:k80] %*% shoesScaled
```

```
  eigenimage <- array(eigenShoes[,3], c(1200,2500,3))
  imageShow(eigenimage)
```



References

1. https://rpubs.com/dherrero12/543854/
2. https://en.wikipedia.org/wiki/Eigenface