

# DATA 605: Assignment 04

EigenShoes

Philip Tanofsky

15 February 2021

## EigenShoes

Starting point: <https://rpubs.com/R-Minator/eigenshoes>

```
# Libraries
library(jpeg)
library(EBImage)
library(OpenImageR)
```

```
##
## Attaching package: 'OpenImageR'

## The following objects are masked from 'package:EBImage':
##
##   readImage, writeImage
```

## Use of Graphics

Add graphics to the data set.

```
# Prepare for Image Processing
num <- 17
files <- list.files("/Users/philiptanofsky/Documents/School/CUNY/MSDS/Courses/DATA605/Week04/jpg/", pat
#files
```

## View Shoes Function

```
# Set Adjustment Parameters
height <- 1200
width <- 2500
scale <- 20

plot_jpeg <- function(path, add=FALSE) {
  # Read the file
```

```

jpg <- readJPEG(path, native=T)
# Get the resolution, [x, y]
res <- dim(jpg)[2:1]
# Initialize any empty plot are if add == False
if (!add) {
  plot(1, 1, xlim=c(1, res[1]), ylim=c(1, res[2]), asp=1, type='n',
       xaxs='i', yaxs='i', xaxt='n', yaxt='n', xlab='', ylab='', bty='n')
  rasterImage(jpg, 1, 1, res[1], res[2])
}
}

```

## Load the Data into an Array

```

im <- array(rep(0, length(files) * height/scale * width/scale * 3), dim=c(length(files), height/scale, width/scale, 3))

for (i in 1:num) {
  temp <- resize(readJPEG(paste0("/Users/philiptanofsky/Documents/School/CUNY/MSDS/Courses/DATA605/Week04/jpg", files[i])), height/scale, width/scale)
  im[i,,] <- array(temp, dim=c(1, height/scale, width/scale, 3))
}

```

## Vectorize

```

flat <- matrix(0, num, prod(dim(im)))

for (i in 1:num) {
  newim <- readJPEG(paste0("/Users/philiptanofsky/Documents/School/CUNY/MSDS/Courses/DATA605/Week04/jpg", files[i]))
  r <- as.vector(im[i,,1])
  g <- as.vector(im[i,,2])
  b <- as.vector(im[i,,3])
  flat[i,] <- t(c(r, g, b))
}
shoes <- as.data.frame(t(flat))

```

## Actual Plots

NOT SHOWING UP, PLEASE CHECK THIS CODE BELOW

```

# Old Shoes
par(mfrow=c(3,3))
par(mai=c(.3, .3, .3, .3))
for (i in 1:num) { # plot the first images only
  plot_jpeg(writeJPEG(im[i,,]))
}

```





## Get Eigen components from correlation structure

```
scaled <- scale(shoes, center=TRUE, scale=TRUE)
mean.shoe <- attr(scaled, "scaled:center") # saving for classification
std.shoe <- attr(scaled, "scaled:scale") # saving for classification ... later
```

## Calculate Covariance (Correlation)

```
sigma_ <- cor(scaled)

#sigma_
```

## Get the eigencomponents

```
myeigen <- eigen(sigma_)
cumsum(myeigen$values) / sum(myeigen$values)

## [1] 0.6928202 0.7940449 0.8451073 0.8723847 0.8913841 0.9076338 0.9216282
## [8] 0.9336889 0.9433872 0.9524455 0.9609037 0.9688907 0.9765235 0.9832209
## [15] 0.9894033 0.9953587 1.0000000
```

## Eigen shoes

```
scaling <- diag(myeigen$values[1:5]^(-1/2)) / (sqrt(nrow(scaled)-1))
eigenshoes <- scaled %*% myeigen$vectors[,1:5] %*% scaling
imageShow(array(eigenshoes[,1], c(60, 125, 3)))
```



## Generate Principal Components

Transform the images

```
# Generate variables
height <- 1200
width <- 2500
scale <- 20
newdata <- im
dim(newdata) <- c(length(files), height*width*3/scale^2)
mypca <- princomp(t(as.matrix(newdata)), scores=TRUE, cor=TRUE)
```

## Eigen shoes

Generate Eigen shoes

```

mypca2 <- t(mypca$scores)
dim(mypca2) <- c(length(files), height/scale, width/scale, 3)
par(mfrow=c(5,5))
par(mai=c(.001, .001, .001, .001))
# Plot the first 20 eigenshoes only
for (i in 1:num) {
  plot_jpeg(writeJPEG(mypca2[i,,], bg="white")) # Complete without reduction
}

```



## Variance Capture

```

a <- round(mypca$sdev[1:num]^2 / sum(mypca$sdev^2), 3)
cumsum(a)

```

```

## Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9 Comp.10
## 0.693 0.794 0.845 0.872 0.891 0.907 0.921 0.933 0.943 0.952
## Comp.11 Comp.12 Comp.13 Comp.14 Comp.15 Comp.16 Comp.17
## 0.960 0.968 0.976 0.983 0.989 0.995 1.000

```