Robot Positioning Using an Omnidirectional Camera

Report for Lab Exercise 5 for the Computer Vision Course

A.P. van Ree (10547320)
T.M. Meijers (10647023)
University of Amsterdam
The Netherlands

December 12, 2014

1 Introduction

Written exercises: This report contains the theory questions about PCA which is discussed in Lab exercise 5: *Robot Positioning Using an Omnidirectional Camera*.

Programming exercises: The code is commented well enough so that every step is clear, and the structure of the code is well organized. The main file (*main.m*) is organized into sections which correspond to the structure in the assignment. The helper functions we have written (*parse_data.m* and *pca.m*) are commented well enough to explain their working as well. The code has two main sections, 2 and 3 (from the assignment). Both have an initial subsection (2.1 and 3.1) which should be run to load all the data structures before running other subsections of the corresponding main section.

2 Principal Component Analysis

2.1

The conditions under which the sum of squared differences between pixels of two images I_1 and I_2 reduces to a correlation-based similarity measure, are the following:

1. Only one object in the image, more than one would be too complicated.
2. The camera that captures the objects is fixed and under weak perspective. This means that the camera is always in the same position, aiming the same way etc.
3. Normalized imgages of object – no more whitespace than needed around the object. This makes sure that the only thing in the image is actually the object of interest.
4. $\sum_{i=1}^{N}\sum_{j=1}^{N} I(i,j)^2 = 1$, meaning that the energy of the pixel values in all images is normalized to one. Because of this the images can be compared accurately, without having to care about the difference in pixel intensity. They all the the same rage in each image.
5. The object is displayed completely in all images. Otherwise the images could be taken in the same place but would display different parts of the whole object, giving the wrong results.

2.2

The images all have one object, the inside of a hall in a building on the Kruislaan in Amsterdam. The camera is always in the same (fixed) position, in the same angle. Also, non of the images have whitespace around them, although some of them have black borders around them because they have been rotated. These black borders lead to weaker results when using PCA. The maximum value of each pixel is normalized to one, and last the hall that is captured in the images is displayed completely in all of these images.

2.3

Yes, by representing the images as vectors and calculating the correlation one takes the sum of the product of the corresponding elements in the vectors and this is exactly what a dot product is.

2.4

Because reshape allows one to reshape a matrix into a vector (by concatenating all rows of the matrix).

Second set of theory questions

2.5

Shlens uses a datamatrix that is transposed in comparison to Trucco, which is why he uses the transposed relative method of what Trucco uses for X.

2.6

This is because r is a constant factor it can be left out of the summation (as Leo explained during his lecture). Because we want to project every vector onto the normalized vector r, we first make a covariance matrix and then 'project' this onto r again.

2.7

Of course, because as Shlen states: when a matrix orthogonally diagonalizable, it is always symmetric.

2.8

Yes, with the matrix S we rotate the frame to a standard coordinate frame as explained. The diagonal matrix then gives us the eigenvalues which in turn tell us how much each direction (eigenvector) should be 'stretched'. We then rotate the whole frame back.