Centre for Mathematical Sciences, Mathematics, LTH, August 2014

Image Analysis, Assignment 1

1 Instructions

In this assignment you will study camera basic histogram equalization, linear algebra for images and start your work on your own Optical Character Recognition system. The data for the assignments is on the course homepage:
`http://www.ctr.maths.lu.se/course/newimagean/2014/`

The assignment is due on thursday of study week 3. Make sure you answer all questions and provide complete solutions to the exercises. You may hand in hand-written solutions and/or printouts in the slot marked "bildanalys" on the 3rd floor of the mathematics building and/or as a pdf by mail to fman20@maths.lth.se. Write your name and the assignment number in the subject line. For each exercise, we have tried to make it clear what should be included in the report. In addition, all the code should be submitted as m-files by mail. Make sure that your matlab scripts are well commented and can be executed directly (that is, without loading any data, setting parameters etc. Such things should be done in the script).

You will have time to work with the assignments during the computer laboratory sessions and the exercise sessions. These sessions are intended to provide an opportunity for asking questions to the lecturer on things you have problems with or just to work with the assignment. During the laboratory sessions you should work on the exercises marked "Computer Exercise". The rest of the exercises are intended to provide hints and prepare you for the computer exercises. You are expected to have solved these before you go to the lab sessions. The report should be written individually, however you are encouraged to work together (in the lab session you might have to work in pairs). Keep in mind that everyone is responsible for their own report and should be able to explain all the solutions.

2 Image sampling

Consider a continuous image. The intensity in a point (x, y) in the image is given by the function

$$f(x, y) = 4x(1 - x), \quad 0 \leq x \leq 1, 0 \leq y \leq 1.$$

What does the image look like?

Sample the image evenly to a discrete image with 5×5 pixels. Let the lower left pixel be a sample from the point $(0, 0)$ in the continuous image and the upper right pixel a sample from $(1, 1)$. Quantify the discrete image with 16 different gray levels from 0 to 15. What is the result?

In the report, write out the resulting 5×5 image matrix.

3 Histogram equalization

An image (in a continuous representation) has gray level histogram

$$p_r = \frac{3}{2}\sqrt{r}, \quad r \in [0, 1] \ .$$

What gray level transform $s = T(r)$ should be used so that the resulting histogram p_s is uniform, i.e.

$$p_s = 1, \quad s \in [0, 1] \ ?$$

In the report, specify the transformation $s = T(r)$.

Neighbourhood of pixels

Consider the following image

$$\begin{pmatrix} 3 & 3 & 2 & 2 & 2 & 3 & 3 & 3 & 0 & 2 & 2 & 2 \\ 0 & 3 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 3 & 2 & 3 \\ 1 & 2 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 2 & 3 \\ 3 & 0 & 3 & 1 & 0 & 3 & 3 & 0 & 2 & 3 & 3 & 2 \\ 3 & 2 & 1 & 0 & 1 & 3 & 1 & 3 & 2 & 0 & 3 & 2 \\ 0 & 3 & 3 & 1 & 1 & 1 & 1 & 0 & 2 & 3 & 3 & 3 \\ 2 & 0 & 2 & 1 & 1 & 0 & 1 & 0 & 2 & 2 & 0 & 2 \\ 2 & 1 & 3 & 0 & 0 & 3 & 2 & 1 & 2 & 0 & 3 & 2 \\ 2 & 0 & 3 & 1 & 0 & 3 & 2 & 3 & 2 & 2 & 0 & 3 \\ 2 & 2 & 2 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 2 & 1 \\ 0 & 2 & 3 & 0 & 0 & 1 & 0 & 0 & 0 & 3 & 3 & 2 \\ 0 & 2 & 3 & 0 & 3 & 3 & 2 & 1 & 2 & 1 & 0 & 2 \end{pmatrix} .$$

a) Mark each element that has intensity 0 or 1 with a circle.

For a pixel with coordinates (m, n) the 4-neighbours are the four pixels with coordinates $(m \pm 1, n)$ and $(m, n \pm 1)$.

b) Fill the circles of those pixels that has intensity 0 or 1 and has at least two 4-neighbours that also have intensities 0 or 1.

4 Scalar products and norm on images

How is the scalar product defined for images? Given three images

$$u = \begin{bmatrix} 1 & -3 \\ 4 & -1 \end{bmatrix},$$

$$v = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix},$$

$$w = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} .$$

Calculate $||u||$, $||v||$, $||w||$, $u \cdot v$, $u \cdot w$, $v \cdot w$. Are the vectors $\{v, w\}$ orthonormal? What is the orthogonal projection of u on the supspace spanned by $\{v, w\}$?

5 Dimensionality

The set of images with 2×2 pixels is a vector space. How many dimensions does this vector space have? How many dimensions does the space of 2000×3000 images have?
In the report, answers are enough.

6 Image compression

A small camera delivers low resolution images with 3×4 pixels. Before transmitting the image to a computer, one would like to compress the images consisting of 12 intensities to 4 numbers. After studying numerous images and using principal component analysis one has determined that the following four images represent typical images well,

$$\phi_1 = \frac{1}{3}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \phi_2 = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \phi_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \phi_4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}.$$

Show that these four images are orthonormal in the scalar product

$$(f, g) = \sum_i \sum_j f(i, j)g(i, j) \ .$$

How should we determine the four numbers (coordinates) x_1, x_2, x_3, x_4 such that the approximate image

$$f_a = x_1\phi_1 + x_2\phi_2 + x_3\phi_3 + x_4\phi_4$$

is as close to f as possible, i.e. such that $|f - f_a|^2 = (f - f_a, f - f_a)$ is as small as possible?
Calculate x_1, x_2, x_3, x_4 for the image

$$f = \begin{pmatrix} -2 & 6 & 3 \\ 13 & 7 & 5 \\ 7 & 1 & 8 \\ -3 & 3 & 4 \end{pmatrix} \ .$$

7 Image bases

On the web page for the course there is a file, zip-file, `assignment1bases.mat` . with two variables `bases` and `stacks`.
The variable `stacks` is a cell array. It contains two stacks of images

- 500 test images of faces of size 19×19. These are stored in a variable `stack{1}`, which is a tensor of size $19 \times 19 \times 500$.
- 500 other test images of size 19×19. These are stored in a variable `stacks{2}`, which is a tensor of size $19 \times 19 \times 500$.

In the previous exercise we saw how to project an image onto a low-dimensional subspace defined by a set of basis images. Is there a difference between different bases? What is the best basis? How can one calculate a good basis?

The variable `bases` is also a cell array. It contains three sets of bases for 3 different sub-spaces of dimension 4. The first basis is stored in a variable `bases{1}`, which is a tensor of size $19 \times 19 \times 4$. Thus the four basis images are `bases{1}(:,:,1)`, `bases{1}(:,:,2)`, `bases{1}(:,:,3)`,`bases{1}(:,:,4)`.

Write a matlab function that projects an image u onto a basis (e_1, e_2, e_3, e_4) and returns the projection u_p and error norm r, i.e. the norm of the difference $r = |u - u_p|$.

Then write a script that tests all of the 500 test images in a test set and returns the mean of the error norms. Calculate this mean for each of the 2 test sets on each of the three bases.

In the report describe in your own words what the images look like in the two test sets. Also in the report plot the four images in each of the three bases and describe what they look like. Then print the mean of the error norms for the six combinations (2 test sets against the three bases). Which basis works best for test set 1? Why? Which basis works best for test set 2? Why?

Discussion? How should one calculate the best basis for a stack of images?

8 Segmentation part of OCR

On the web page for the course in image analysis there is a zip-file, `assignment1ocr.zip` . By downloading and unpacking the file one obtains a folder with (i) examples of images of text and (ii) matlab-scripts.

Study the script `inl1_stomme.m`, that reads in one of the images. The image contain text (dark against a light background) Write a matlab program `im2segment` that takes such an image matrix I as input and returns a segmentation, i.e. a set of images $S = (S_1, \ldots, S_n)$ one for each letter in the image. Each such image matrix S_i should be a matrix with ones at the pixels for that letter and zeroes for all other pixels. A suggestion is to use a so called 'cell array' in matlab as the data structure for the output.

```
S = cell(1,n);
S{1} = bild1;
...
S{n} = bildn;
```

A suggested name for the script is

```
function S = im2segment(I);
```

In the report, print your code and the results of using your algorithm, i.e. supply examples of input data (e.g. as images) and result after applying your segmentation algorithm (e.g. as a number of images, one for each segment).