T-61.5010 Information Visualization Examination

March 8th, 2010

To pass the course you must also pass the term project (assignment). Results of this examination are valid for one year after the examination date. To get full points you must complete all of the problems 1–5. All of the problems have an equal weight.

You can answer in Finnish, Swedish or English (preferred). Please write clearly and leave a wide left or right margin. No extra material (calculator, lecture notes etc.) is allowed. Please write your answers preferrably using a ballpoint pen, not a pencil.

Instructions for the essays: Write in full sentences and structure your answer in paragraphs. The essay should be written in a manner understandable to your fellow student (who would have the necessary prerequisite information to take this course, but has not taken it) who has asked you to tell him/her about the topic of the essay.

The results will be posted to the course Noppa home page on April 8th 2010, at latest. No other announcement will be made.

There are 4 (four) pages in this examination. You can keep this paper.

1 Multiple choices questions

The following questions have each different proposed answers. Only one of them is correct. **You have to give your answer along with your confidence ("High" or "Low") for each answer.** Grading for each of these multiple choices questions is then:

- +2 if answer is right and confidence is high
- +1 if answer is right and confidence is low
- 0 if answer is missing
- -1 if answer is wrong and confidence is low
- -2 if answer is wrong and confidence is high

Write on your answer sheet the correct answer (a, b, c, d,...), along with the confidence you have (High, Low) for that question; e.g "a, Low" is a proper way of answering a question. **Missing confidence for a question will be treated as "Low".**

Question 1

PCA (Principal Component Analysis) is a projection method with

a. bad recall and bad precision

b. good recall and good precision

c. good recall and bad precision

d. bad recall and good precision

Question 2

Which affirmation is correct ?

a. MDS (Multidimensional Scaling) preserves small distances

b. CCA (Curvilinear Component Analysis) does not preserve small distances

c. For MDS, the stress is always increasing with the projection dimension

d. MDS requires the coordinates of the original data to project (distances are not enough)

e. None of the above answers is correct

Question 3

The first principal component for PCA (Principal Component Analysis) in this figure (Figure 1) is given by:

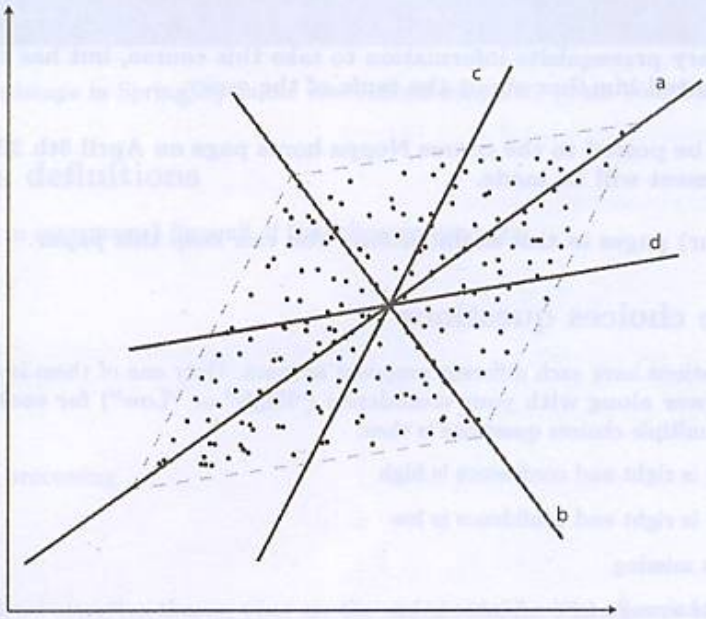

Figure 1: Which of the four lines is the first Principal Component ?

a. line a

b. line b

c. line c

d. line d

Question 4

We want to perform Laplacian Eigenmap (LE). We have the graph laplacian L:

$L = \begin{pmatrix} -2 & 1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 0 & 1 & -2 \end{pmatrix}$, and we compute the eigenvectors in the form of the matrix V and the eigenvalues in the form of the matrix D, for example in Octave/Matlab, as [V, D] = eig(L). We obtain

$D = \begin{pmatrix} -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, and $V = \begin{pmatrix} -0.4 & -0.4 & 0.4 & -0.4 & 0.4 & 0.4 \\ 0.4 & -0.2 & -0.5 & 0.2 & 0.5 & 0.4 \\ -0.4 & 0.6 & 0.1 & 0.6 & 0.1 & 0.4 \\ 0.4 & -0.4 & 0.4 & 0.4 & -0.4 & 0.4 \\ -0.4 & -0.2 & -0.5 & -0.2 & -0.5 & 0.4 \\ 0.4 & 0.6 & 0.1 & -0.6 & -0.1 & 0.4 \end{pmatrix}$.

Which of the following affirmations is correct:

a. Projecting to one dimension using LE is enough to preserve the proximity coded in L

b. Projecting to two dimensions using LE is enough to preserve the proximity coded in L

c. Projecting to three dimensions using LE is necessary to preserve the proximity coded in L

d. Laplacian Eigenmap does not work in this case

e. None of the above answers is correct

Question 5

The definition of the cost function σ_r for CCA (Curvilinear Component Analysis) is (where d denotes the Euclidean distance between points in the original space (x) and in the projection (y), and $F(d, \lambda_y)$ equals unity if $d < \lambda_y$ and 0 otherwise):

a. $\sigma_r = \sum_{i<j} (d(x_i, x_j) - d(y_i, y_j))^2 F(d(y_i, y_j), \lambda_y)$

b. $\sigma_r = \sum_{i<j} (d(x_i, x_j) - d(y_i, y_j))^2 F(d(x_i, x_j), \lambda_y)$

c. $\sigma_r = \sum_{i<j} (d(x_i, x_j) - d(y_i, y_j))^2$

d. $\sigma_r = \sum_{i<j} (d(x_i, x_j) - d(y_i, y_j)) F(d(y_i, y_j), \lambda_y)$

e. $\sigma_r = \sum_{i<j} (d(x_i, x_j) - d(y_i, y_j)) F(d(x_i, x_j), \lambda_y)$

2 Depth Cue theory

Based on the piece of art (Figure 2) below, discuss and explain the Depth Cue theory. List and explain at least 4 of them (there might more/less than 4 depth cues in this piece of art).

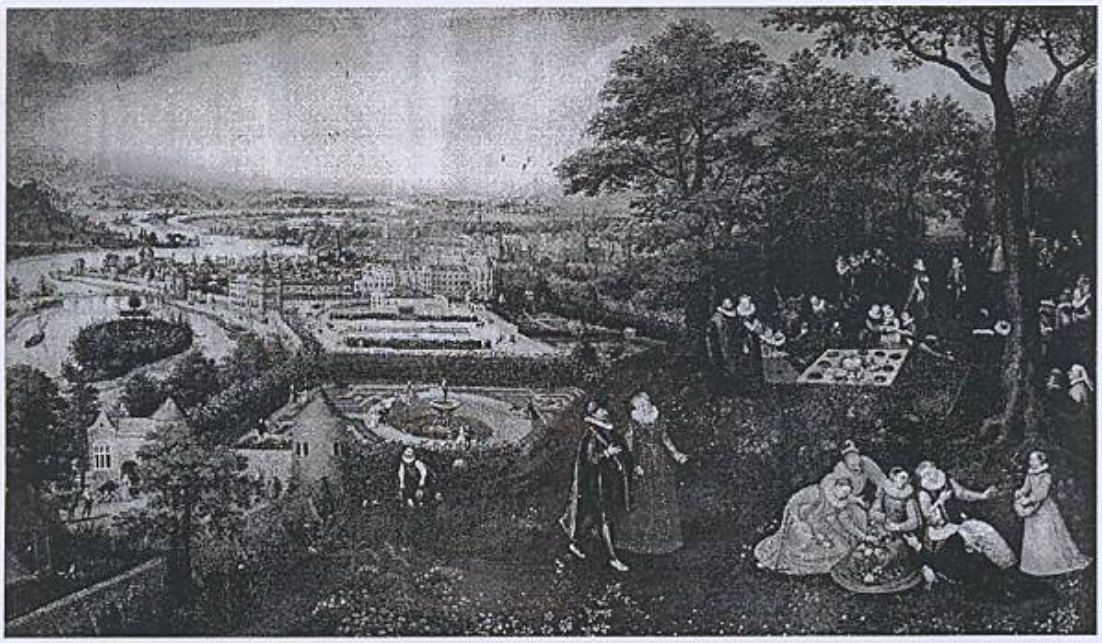

Figure 2: Landscape in Spring, by Lucas van Valkenborch, 1587 (from Web Gallery of Art).

3 Concepts definitions

Define and explain in maximum 3 lines all of the following concepts:

- Lie-factor
- Data-ink
- Gestalt laws
- Geons
- Pre-attentive processing

4 Essay 1

According to the visual attention theory, what are the ways or choices that allow information to pop-up (active from the low-level point of view).

5 Essay 2

Choose one of the two following subjects for the second essay:

- History and theory of data graphics according to Tufte;
- Describe and compare the two theories on object recognition.