An image analyzer is used to synthesize the colors of each pixel constituting the image in the Red, Green, Blue (RGB) color space. The color of each pixel corresponding to a specific mixture of these 3 colors in various proportions, the role of this analyzer is to measure the levels of the R, G, B components. These levels, respectively denoted NR (Red Level), NG (Green Level) and NB (Blue Level), are coded between a minimum (0%) and a maximum (100%). For example, for a yellow colored pixel, the analyzer should indicate: NR=NG=1 and NB=0. For a cyan pixel: NR=0, NG=NB=1.

In order to calibrate the device, a “reference image” is analyzed 100 consecutive times. The definition of this "reference image" is 5 pixels in width and 5 pixels in height. Every measurement, the analyzer will generate 3 matrices NR, NG, NB of dimension 5x5. In each of these matrices show the level of R, G or B of the corresponding pixel.

For example, a measure could yield the following matrices:

$$NR=\begin{pmatrix} 1&1&1&1&1\\ 1&1&1&1&1\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0 \end{pmatrix} \quad NG=\begin{pmatrix} 1&1&1&1&1\\ 1&1&1&1&1\\ 1&1&1&1&1\\ 1&1&1&1&1\\ 1&1&1&1&1 \end{pmatrix} \quad NB=\begin{pmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 1&1&1&1&1\\ 1&1&1&1&1\\ 1&1&1&1&1 \end{pmatrix}$$

This would then correspond to a two-color image whose top 10 pixels (first 2 lines) would be yellow and the bottom 15 pixels (last 3 rows) would be cyan.

The results of these 100 analyzes are saved in the analyseur.sod file.

To extract under Scilab the matrices NRi, NGi and NBi from the ith analysis for i ∈ {1, ..100},

Just type the following line of code:

```
load('parser.sod');

NRi = NR{i}.entries;

NGi = NG{i}.entries;

NBi = NB{i}.entries;
```

Data analysis

Issue 1

From the measurements provided by the image analyzer, construct a data matrix where the individuals would correspond to the various analyzes carried out (i.e. 100 individuals). We will name M the matrix thus obtained.

Issue 2

Construct the average individual of individuals 21 and 22.

Issue 3

We want to disregard the 10 analyzes that have the most bad measurements and therefore, the 10 most different from the other 90. Suggest a method to do and identify the individuals to be removed.

Issue 4

Build the new data matrix that does not take into account these 10 “bad” ones analyzes. This matrix will be called M90.

Issue 5

Build the average individual of M90 and deduce the matrices NRaverage, NGaverage and NBaverage corresponding.

Question 6

Build the reduced centered data matrix of the M90 matrix. We denote this matrix by X.

Question 7

Deduce how many axes are then necessary to restore at least 80% of the inertia of the cloud of the 90 individuals, specifying the quality of the representation thus obtained (in %) .