EXAMINATION OF REMOTE SENSING
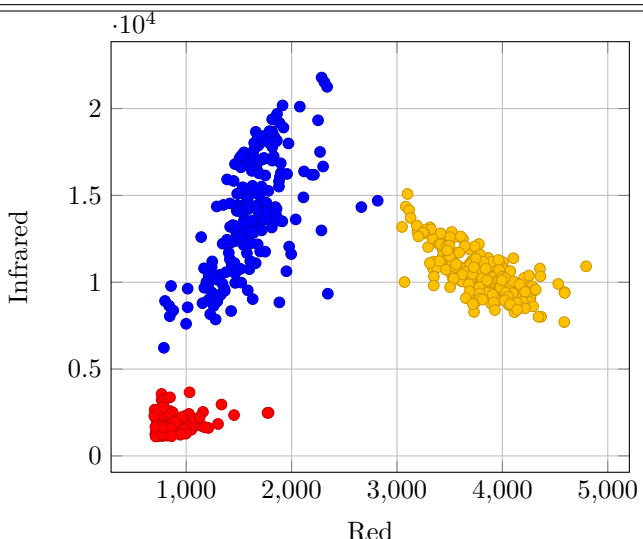
Last name:_____, First name:_____

2014-2015

The matlab files should be send at the end of the exam to `mathieu.fauvel@ensat.fr` and should be named {last-name_function*.m}

1 Multiple choice questions (/6 pts)

+1pt for right answer, -0.5pt for false answer and 0pt for no answer. One right response by question.

Questions	a	b	c	d	pts
1. In the figure above, the data are in a= reflectance, b= nanometer, c= without unit, d= micrometer					
2. In the figure above, the class corresponding to blue points is a= Bitumen, b= Shadow, c=Vegetation, d= Bare soil.					
3. In the figure above, the class corresponding to red points is a= Bitumen, b= Shadow, c=Vegetation, d= Bare soil.					
4. The figure above is a= a scatter plot, b= an histogram, c= mean spectra, d= a thematic map.					
5. A multispectral pixel PLEIADE corresponds to a square of length 2 meters. Knowing that a panchromatic pixel PLEIADE include 16 multispectral pixels, the spatial resolution of the panchromatic image is a= 1m, b= 50cm, c= 0.125 cm, d=4m.					
6. For a GMM classifier and pixel $\mathbf{x} \in \mathbb{R}^d$, if we suppose that the covariance matrix of each class is diagonal, then the number of the parameters to estimate the mean vector and the covariance matrix of a given class is a $= d$, b=d^2, c=2d, d= $d + d(d+1)/2$.					

2 Kernel k-NN (/6 pts)

k-nearest neighor classifier is a very simple technique that classifies a pixel $\mathbf{x}$ to a class c such as the distance to a referenced pixel of class c is minimal. The algorithm can be written in a few steps:

- Compute the distance $d(\mathbf{x}, \mathbf{x}_i)$ of each pixel $\mathbf{x}_i$, for each class, with $d(\mathbf{x}, \mathbf{x}_i) = \|\mathbf{x} - \mathbf{x}_i\|^2$.

- Select the class c for which the distance is minimal if $k = 1$, or using a majority vote among the k neighbors.

- Set the class of $\mathbf{x}$ to c.

The objective of the exercice is to extend this algorithm using kernel function and to compare the results in terms of classification accuracies. We are going to use the *kernel trick*, *i.e.*, the linear dot product will be changed to a kernel evaluation.

1. Write the distance function in terms of dot products ($\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$).

2. Kernelize the decision function: write the distance function using kernel evaluation rather than the dot product ($\langle \mathbf{x}, \mathbf{x}_i \rangle \to k(\mathbf{x}, \mathbf{x}_i)$). Then, for the specific case of Gaussian kernel ($k(\mathbf{x}, \mathbf{x}) = 1$), factorize the decision function.

3. Implement the new decision rule in Matlab using the Gaussian kernel.

4. Test it on the University of Pavia: learn the optimal hyperparameter on the train set and classify the test set.

5. Compare the results in terms overall accuracy with the conventional k-NN.

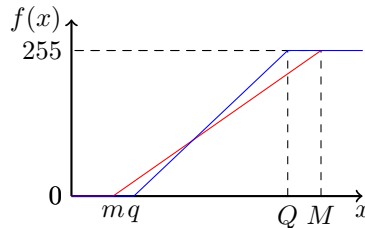## 3 Histogram stretching with saturation (/4 pts)

Figure 1: Histogram stretching.

In the lab-work 1, for the visualization of color composition, you have stretched the data between 0 and 255 using the following function:

$$f(x) = 255 \times \frac{(x - m)}{M - m}.$$

This approach is sensitive to noise and not optimal in term of contrast enhancement since the maximum and minimum of the spectral channel can be outliers of the true distribution of image.

In this work, you will implement the following stretching function (see figure 1):

$$f(x) = \begin{cases} 255 & \text{if } x > Q \\ 0 & \text{if } x < q \\ 255 \times \frac{(x-q)}{Q-q} & \text{otherwise} \end{cases} .$$

Q and q will be computed so that the 2th percentile and 98th percentile of the data range are not considered. The matlab function *prctile* will be useful.

Write a matlab function with input arguments the image and the three bands to visualize.

Percentile [Wikipedia]: *A percentile is a measure used in statistics indicating the value below which a given percentage of observations in a group of observations fall. For example, the 20th percentile is the value (or score) below which 20 percent of the observations may be found. The term percentile and the related term percentile rank are often used in the reporting of scores from norm-referenced tests. For example, if a score is in the 86th percentile, it is higher than 86% of the other scores.*
The 25th percentile is also known as the first quartile, the 50th percentile as the median or second quartile, and the 75th percentile as the third quartile. In general, percentiles and quartiles are specific types of quantiles.