ECE651 - Detection Theory
Computer Assignment 2, Due December 7th, 2017
Dr. M.R. Azimi

The purpose of this computer assignment is to apply the *matched subspace detector* to the same mammogram imagery data set used in Computer Assignment 1 and to compare the detection method and results with those developed in Computer Assignment 1. Again, the objective of the assignment is to develop a detection system capable of discriminating calcified masses from normal tissue in mammogram imagery.

Matched Subspace Detection

1. Process the images by considering each column of length N within an Region of Interest (ROI) in the image [1] to be a realization of a Normally distributed random vector. This processing scheme is displayed pictorially in Figure 1.

2. From a subset of images extract $M \gg N$ columns from some selected and representative set of ROIs containing calcified masses (training data), $\{\mathbf{s}_i\}_{i=1}^{M}$, and form the data matrix $S = [\mathbf{s}_1 \ \cdots \ \mathbf{s}_M] \in \mathbb{R}^{N\times M}$.

3. Find the orthonormal matrices $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_N] \in \mathbb{R}^{N\times N}$ and $V = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_N] \in \mathbb{R}^{M\times N}$ as well as the matrix $\Sigma = \text{diag}\{\lambda_1, \ldots, \lambda_N\}$ such that $S = U\Lambda V^T$, i.e. obtain the "thin" singular value decomposition of the data matrix S. Note that these matrices can be computed using Matlab command *svd(S,'econ')*.

4. Assuming that $\lambda_1 > \cdots > \lambda_N$, extract the $p < N$ principal basis vectors from matrix U associated with the p largest singular values and form the *signal subspace* matrix $H = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_p] \in \mathbb{R}^{N\times p}$. If unsure of a suitable value for p, one possibility is to measure the percentage of the energy captured in the first p basis vectors by measuring the quantity

$$\eta = \frac{\sum_{i=1}^{p} \lambda_i^2}{\sum_{i=1}^{N} \lambda_i^2} \tag{1}$$

   and choose the smallest p that achieves a value of η sufficiently close to one.

5. For each column within the image, $\mathbf{x}$, consider the following hypothesis test

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} = \mathbf{n} \\ \mathcal{H}_1 &: \mathbf{x} = H\boldsymbol{\theta} + \mathbf{n} \end{aligned}$$

[1] The size of the ROI should be determined experimentally. You can use the same size used before for benchmarking purposes.

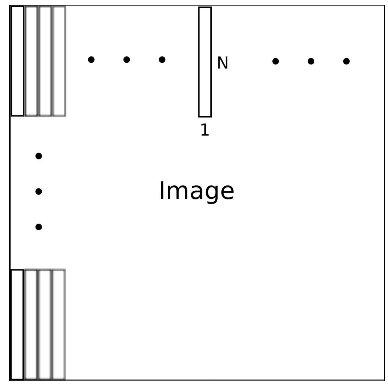

Figure 1: Processing using Columns within the Image.

where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I)$ and $\boldsymbol{\theta} \in \mathbb{R}^p$ is a deterministic but unknown vector of parameters. If unsure of a suitable value for σ^2, extract columns of normal tissue and estimate this parameter. Note that if given K realizations of normal tissue, $\{\mathbf{n}_i\}_{i=1}^K$, then a suitable estimate of this parameter is given as

$$\hat{\sigma}^2 = \frac{1}{NK} \sum_{i=1}^{K} ||\mathbf{n}_i||^2 \tag{2}$$

6. Empirically generate the ROC curve for the matched subspace detector by plotting P_{FA} vs. P_D for a varying threshold and compare this to the theoretical ROC curve. Note that the Matlab commands *chi2cdf* and *ncx2cdf* give the cumulative distribution functions for the central and non-central chi-squared distributions, respectively. For a suitably chosen threshold, give the P_D and P_{FA} for the detector.

7. Compare the matched subspace detector built in this computer assignment to those detectors developed in Computer Assignment 1 in terms of both performance and implementation.