APPLIED SIGNAL PROCESSING 2011

EXERCISE 3

a) Image compression using PCA

Consider an $N_r \times N_c$ array $\mathbf{X}(i,j)$ representing an image. Determine the deviation from the mean $\mathbf{W}(i,j) = \mathbf{X}(i,j) - m$ where m is the mean value

$$m = \frac{1}{N_r N_c} \sum_{i,j} \mathbf{X}(i,j)$$

In order to compress the array by PCA, determine the singular value decomposition

$$\mathbf{W} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \sum_i^p \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

where $p = \min(N_r, N_c)$.

Plot the singular values σ_i. Introduce next a compressed approximation $\mathbf{W}_r$ of $\mathbf{W}$ using only the r first dominant principal components,

$$\mathbf{W}_r = \sum_i^r \sigma_i^r \mathbf{u}_i \mathbf{v}_i^T = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^T$$

where $\mathbf{U}_r = [\mathbf{u}_1 \cdots \mathbf{u}_r]$, $\mathbf{\Sigma}_r = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$ and $\mathbf{V}_r = [\mathbf{v}_1 \cdots \mathbf{v}_r]$. Next present the compressed image $\mathbf{X}_r(i,j) = \mathbf{W}_r(i,j) + m$ and compare with the original.

Calculate the amount of data needed to represent the compressed image. Compute the average root-mean-square approximation error

$$e(\mathbf{W} - \mathbf{W}_r) = \Big(\sum_{i,j} \frac{1}{N_r N_c} \big(\mathbf{W}(i,j) - \mathbf{W}_r(i,j)\big)^2 \Big)^{1/2}$$

and determine the relative root-mean-square approximation error
$e(\mathbf{W} - \mathbf{W}_r)/e(\mathbf{W})$.

b) Blind source signal separation using ICA

Take $M (\geq 2)$ signals $y_i(n)$ of equal lengths and mix them randomly to produce M mixed signals $w_i(n)$. Then use the FastICA algorithm to unmix the signals $w_i(n)$ to find independent signal components $s_i(n)$. Compare the calculated independent components with the original signals $y_i(n)$.

The calculations can be performed using the FastICA program package, which can be downloaded from the web.

Note: Observe that FastICA determines the independent components $s_i(n)$ in arbitrary order, the components are normalized to have unit variance (whereas there is no such assumption on the original signals $y_i(n)$), and they are arbitrary with respect to sign ($-s_i(n)$ and $s_i(n)$ are equally valid independent components). For comparison, $y_i(n)$ and $s_i(n)$ should therefore be scaled to have equal variances, and possible sign differences should observed.