Fisher linear discriminant analysis

- Eigenimage method maximizes "scatter" within the linear subspace over the entire image set – regardless of classification task

$$W_{opt} = \arg\max_{W} \left(\det\left(W R W^{H} \right) \right)$$

- Fisher linear discriminant analysis (1936): maximize between-class scatter, while minimizing within-class scatter

$$W_{opt} = \arg\max_{W} \left(\frac{\det\left(W R_B W^{H} \right)}{\det\left(W R_W W^{H} \right)} \right)$$

$$R_B = \sum_{i=1}^{c} N_i \left(\vec{\mu}_i - \vec{\mu} \right) \left(\vec{\mu}_i - \vec{\mu} \right)^{H}$$

Samples in class i

Mean in class i

$$R_W = \sum_{i=1}^{c} \sum_{\vec{\Gamma}_l \in Class(i)} \left(\vec{\Gamma}_l - \vec{\mu}_i \right) \left(\vec{\Gamma}_l - \vec{\mu}_i \right)^{H}$$

Fisher linear discriminant analysis (cont.)

- Solution: Generalized eigenvectors $\overrightarrow{w_i}$ corresponding to the J largest eigenvalues $\{\lambda_i \mid i = 1,2,...,J\}$, i.e.

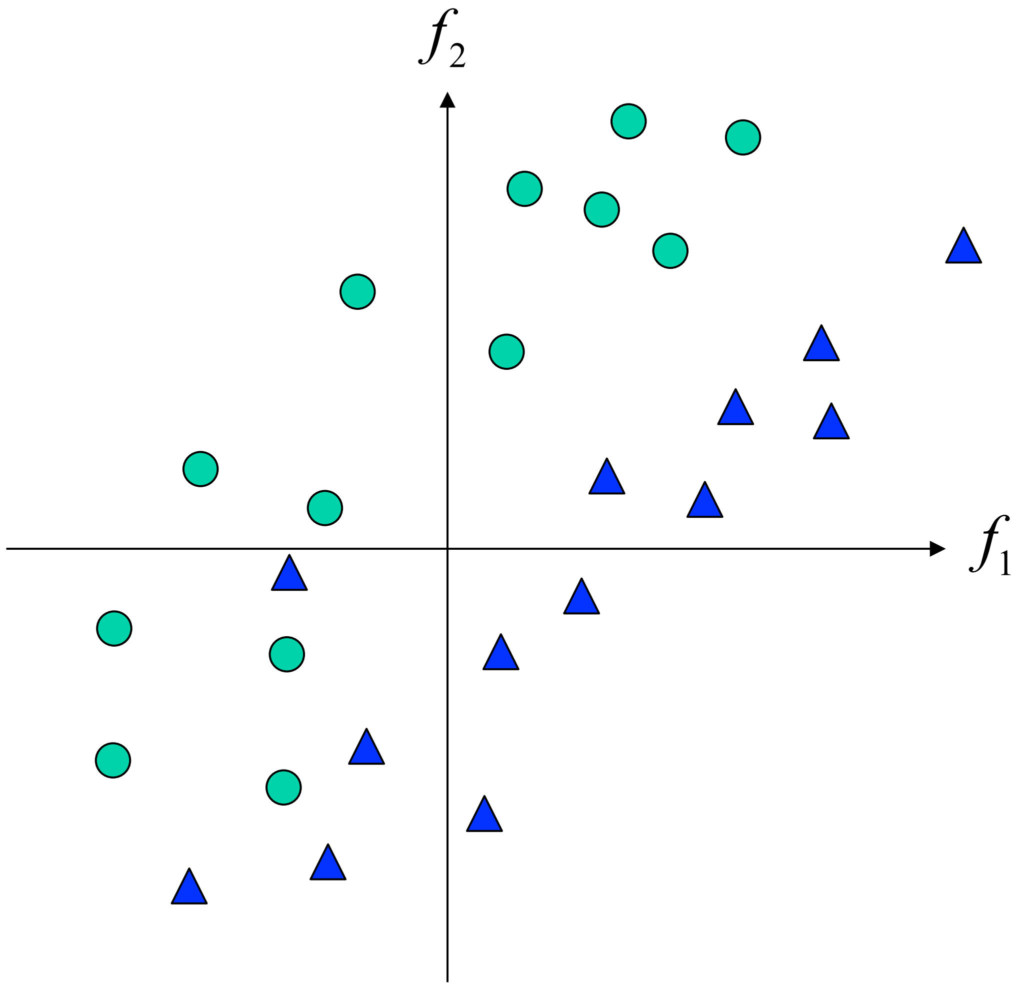
$$R_B \overrightarrow{w_i} = \lambda_i R_W \overrightarrow{w_i} \quad , \quad i = 1,2,...,J$$

- Problem: within-class scatter matrix R_w at most of rank L-c, hence usually singular.
- Apply KLT first to reduce dimensionality of feature space to L-c (or less), proceed with Fisher LDA in lower-dimensional space

Eigenimages vs. Fisherimages

2-d example:

Goal: project samples on a 1-d subspace, then perform classification.

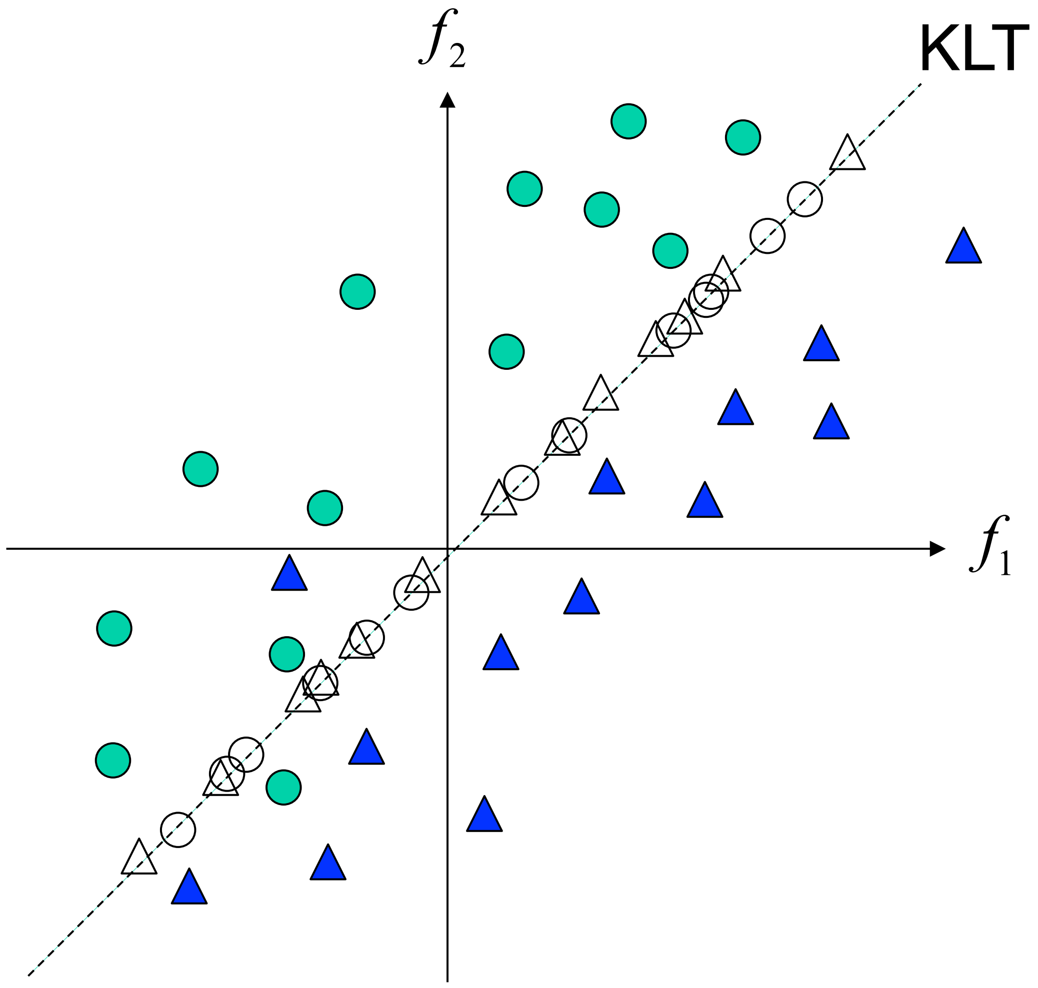


Eigenimages vs. Fisherimages

<u>2-d example:</u>

Goal: project samples on a 1-d subspace, then perform classification.

The KLT preserves maximum energy, but the 2 classes are no longer distinguishable.



Eigenimages vs. Fisherimages

<u>2-d example:</u>

Goal: project samples on a 1-d subspace, then perform classification.

The KLT preserves maximum energy, but the 2 classes are no longer distinguishable.

Fisher LDA separates the classes by choosing a better 1-d subspace.