INTRODUCTION

In hyperspectral image analysis an exploratory approach to analyse the image data is to conduct subspace projections. As linear projections often fail to capture the underlying structure of the data, we present kernel based subspace projections of PCA and Maximum Autocorrelation Factors (MAF). The MAF projection exploits the fact that interesting phenomena in images typically exhibit spatial autocorrelation.

The analysis is based on near-infrared hyperspectral images of maize grains demonstrating the superiority of the kernel-based MAF method.

HYPERSPECTRAL GRAIN DATA

A collection of 8 maize grains, front and backside, are used to generate a single hyperspectral image of 153 bands.

- Image size: 370 x 149 x 153
- Spectral range: 950 - 1700nm.

The hyperspectral image tensor is unfolded and represented as a n x p matrix X, where each row represents an observed pixel, i.e. X is a 55130 x 153 matrix.

Pre-Processing

- Light source and dark current compensation.
- Remove 900-950nm (poor SNR).
- No spectral scatter correction.

Subsampling

Appr. nT = 3000 random samples are used for extracting the projection vectors applied to all data pixels.

SUBSPACE PROJECTIONS

Linear Principal Component Analysis (PCA)

Eigenvalue problem formulation maximizing the variance

\[ \Sigma u_i = \lambda_i u_i, \quad \text{where} \quad \Sigma = \frac{1}{n} X^T X \]

The orthonormal projection eigenvectors are expressed as \( U = [u_1, u_2, \ldots, u_r] \) where \( r = \min(n, p) \). The subspace projection becomes \( \Sigma = U^T x \).

The dual formulation is given by

\[ \frac{1}{n} \Sigma u_i^T v_j = \lambda_i v_j \quad \Rightarrow \quad u_i \propto X^T v_i \quad \wedge \quad v_i \propto X u_i \]

Principal Components, PC1-PC3.

Linear Maximum Autocorrelation Factor (MAF)

Maximise autocorrelation \( \rho \) of linear combinations \( a^T x(r) \) of zero-mean spatial variables \( x(r) \) at location \( r \). The difference \( x(r) = x(r) - x(r+\Delta) \) has a covariance matrix \( \Sigma_u \), where \( \Delta \) is a displacement vector.

The autocorrelation \( \rho \) can be found as

\[ \rho = 1 \frac{a^T \Sigma_u a}{a^T a} \]

Properties

- Assumes 2nd order stationarity.
- Invariant to linear matrix transformation \( T x \), i.e. spectral scatter correction is not necessary.

Kernel PCA

Applying the kernel trick consist of mapping \( x \) into a higher dimensional feature space via the non-linear function \( \phi(x) \), i.e. \( x \rightarrow \phi(x) \).

The eigenvalue problem becomes

\[ \Sigma = \phi^T \Sigma \phi \]

By exploiting the dual formulation the subspace projection can be found as

\[ \Phi U = K V \Lambda \]

Projection is memory-based due to \( K = \{k(x_i, x_j)\} \), i.e. dependence on training dataset.

Gaussian kernel is given by \( k(x_i, x_j) = \exp(-\frac{1}{\sigma^2} \| x_i - x_j \|^2) \).

Kernel MAF

As for kernel PCA a similar framework can be derived for the kernel MAF method using the same Gaussian kernel.

CONCLUSION & REFERENCES

Conclusion

We have demonstrated how the kernel-based projections outperform the linear variants by their ability to suppress background noise, illumination and shadow effects.

The kernel MAF transform further provides a superior projection in terms of labelling different maize kernels parts with same colour.

References
