TECHNICAL UNIVERSITY OF DENMARK

**30540 - Mapping from Aerial and Satellite Images**
**F19**

---

# Report 6
# Change Detection and the MAD algorithm

---

*Author:*
Jonathan GUNDORPH
s153114

*Course responsible:*
Allan AASBJERG NIELSEN
Daniel Haugård

March 20, 2019

# 1 Introduction

In Machine Learning and data mining, there exist many different algorithms and approaches for analyzing a dataset. The approach that one might wish to use depends heavily on the problem. If the user has an existing ground-truth, i.e. a "target" of how the outcome of the classification needs to look like, he would use *supervised learning*. If not, he would rather look at the statistical similarities of the data set, an approach known as *unsupervised learning*. The concept of this report will cover an algorithm under the unsupervised learning category, namely the *Multivariate Alteration Detection (MAD) Method for Change Detection in Multi- and Hyperspectral Data*.

The idea behind the MAD method is to place increasingly emphasis on "difficult observations", i.e. observations whose change over time is uncertain or difficult to see. It builds upon the already established technique of Canonical Correlation Analysis (CCA), where the multivariate data collected at the same geographical region, measured over multiple time-scales, is used to derive the canonical variables, who are then subtracted from each other. The orthogonal information then contains information on the differences over time, and maximizes this difference in all the variables (spectral bands, in this case). Change detection in this fashion is very robust and invariant to separate linear (affine) transformations in the originally measured variables over time.

## 1.1 Multivariate Change Detection

In multivariate change detection in spatial remote sensing, it is customary to calculate the difference between two images (or in this case, two spectral bands). The reason for this is to highlight the areas which exhibit change, as those who does not will have a value of zero or close to zero. Given two multivariate images with k spectral bands, taken over the same region, a simple spectral change detection is the vector of band-wise differences, known as the *change vector*:

$$\mathbf{X} - \mathbf{Y} = [X_1 - Y_1, ..., X_k - Y_k]^T \tag{1}$$

It is important to normalize the data to a common zero and scale or calibrate over time beforehand. Then, to place emphasis on the change, one will wish to maximize the measure of change in the simple multispectral difference image. There is several efficient ways to do this, such as maximizing the deviations from no change under a constraint, equivalent to a principal component analysis. However in our case, we wish to look at the linear combinations

$$\mathbf{a}^T\mathbf{X} = a_1 X_1 + ... + a_p X_p \tag{2}$$

$$\mathbf{b}^T\mathbf{Y} = b_1 Y_1 + ... + b_q Y_q \tag{3}$$

and the difference between them $\mathbf{a}^T\mathbf{X}$ - $\mathbf{b}^T\mathbf{Y}$. We then maximize the variance of this difference and request unit variance of the two terms. We also request that $\mathbf{a}^T\mathbf{X}$ and $\mathbf{b}^T\mathbf{Y}$ are positively correlated. So we wish to determine the linear combinations with extreme correlations, and we will do so with the aid of the Canonical Correlation Analysis (CCA).

# 2 Method and results

## 2.1 Canonical Correlation Analysis (CCA)

CCA determines the relationship between two groups of multiple variables. It finds the two sets of linear combinations of these original variables, one for each group. The first two linear combinations are known as the first canonical variates, and they contain the largest correlation. The second two combinations has the largest correlation based on the condition that they are orthogonal to the first set of variables. Higher order CC's are defined similarly.

### 2.1.1 Calculating the Canonical variables

We denote the variance-covariance matrix (dispersion matrix) of the one set of variables ($\mathbf{X}$) $\sigma_{11}$, the dispersion of the other set of variables ($\mathbf{Y}$) $\sigma_{22}$, the covariance between them $\sigma_{12}$ and the canonical correlation $\rho = Corr(\mathbf{a}^T\mathbf{X}, \mathbf{b}^T\mathrm{Y})$. We then wish to solve the eigenvalue problem for either variable $\mathbf{a}$ or $\mathbf{b}$:

$$\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\mathbf{a} = \rho^2\Sigma_{11}\mathbf{a} \tag{4}$$

$$\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{b} = \rho^2\Sigma_{22}\mathbf{b} \tag{5}$$

$\mathbf{a}$ and $\mathbf{b}$ are then the eigenvectors for each set, $\mathbf{X}$ and $\mathbf{Y}$, corresponding to their respective eigenvalues $\rho_i^2$. Multiplying these eigenvectors onto the original data returns the canonical variables, $\mathbf{a}^T\mathbf{X}$ and $\mathbf{b}^T\mathbf{Y}$.

## 2.2 MAD transformation

Based on the CCA approach from last chapter, the MAD transformation is defined as the set of differences of the canonical variables, as in

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \mathbf{a}_p^T \text{ X - } \mathbf{b}_p^T \text{ Y} \\ \vdots \\ \vdots \\ \mathbf{a}_p^T \mathbf{X} - \mathbf{b}_p^T \mathbf{Y} \end{bmatrix} (6)$$

Visualizing these differences will then reveal the changes in the image. One can also check if the canonical variables is calculated correctly, by examining the structure of the variance-covariance matrix of this difference matrix. This structure should have 1 in the diagonal and the variance of the difference in the subdiagonals for each set. After calculating the matrix, one can see that this is the case:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.9529 | 0.0000 | 0.0000 |
| 0.0000 | 1.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.6408 | -0.0000 |
| 0.0000 | -0.0000 | 1.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.4334 |
| 0.0000 | -0.0000 | 0.0000 | 1.0000 | 0.0000 | -0.0000 | 0.0000 | -0.0000 | -0.0000 |
| 0.0000 | -0.0000 | -0.0000 | 0.0000 | 1.0000 | -0.0000 | 0.0000 | -0.0000 | -0.0000 |
| -0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | 1.0000 | -0.0000 | 0.0000 | 0.0000 |
| 0.9529 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0000 | 1.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.6408 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | 1.0000 | 0.0000 |
| 0.0000 | -0.0000 | 0.4334 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 | 1.0000 |
| 0.0000 | 0.0000 | 0.0000 | 0.3057 | 0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |

# 3 Results

After performing the MAD transformation on the spatial data, the following picture grid can be produced, highlighting the differences in the images over time:

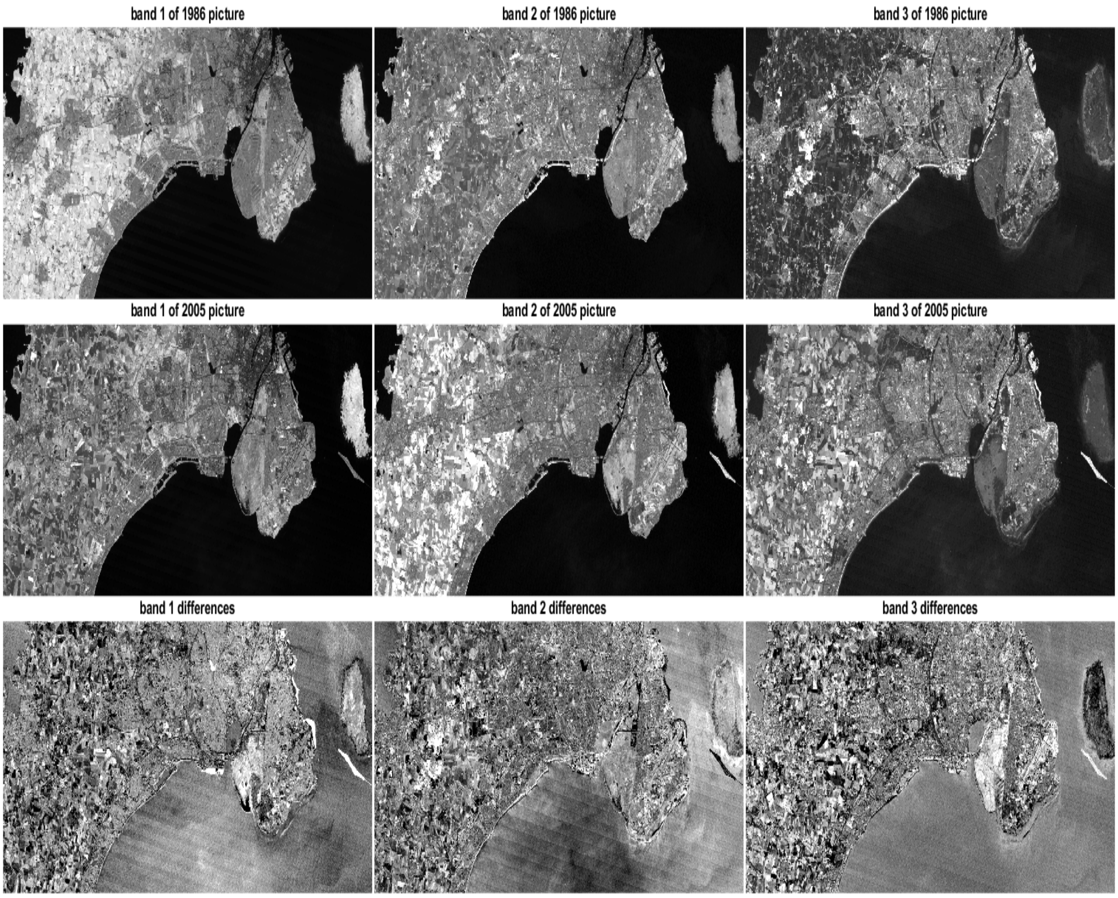


**Figure 1:** Band 1, 2 and 3 and their differences

The picture just shows the first three bands as an example of how it would look. For example, notice the highlighted areas in band 1 differences, where an artifical island has sprung up on Amager. Also, the Peberholm artifical island, built for the Øresunds bridge, has been constructed and is clearly visible in all three bands. Also, there seems to have been some land expansion down at Brøndby Strand and Avedøre Holme, which also jumps out quite visibly in the picture.

# 4 Conclusion

All in all, change detection is a very useful tool and can be used for many things when examining spatial data in Remote Sensing. Its applicabilities has been demonstrated here for the Copenhagen area, where city expansion, for example, is very visible in the MAD transformation visualization.