HYBASE: HYperspectral BAnd SElection

Piet B.W. Schwering, Herman H.P.Th. Bekman and Harm H. van Seijen

TNO Defence, Security and Safety, P.O. Box 96864, NL-2509 JG The Hague, Netherlands, piet.schwering@tno.nl

ABSTRACT

Band selection is essential in the design of multispectral sensor systems. This paper describes the TNO hyperspectral band selection tool HYBASE. It calculates the optimum band positions given the number of bands and the width of the spectral bands. HYBASE is used to assess the minimum number of spectral bands that is required to get the best target background contrast. The band selection algorithm is described along with a description of the graphical user interface. HYBASE is tested on a representative dataset. The test results shed new light on the optimum band selection. HYBASE is developed for the Royal Netherlands Army to investigate the benefit brought by hyper- or multispectral sensors in comparison to present day broad band sensors. HYBASE is tested in European field trials.

Keywords: Hyperspectral Imaging, Multispectral, Band selection, Infrared, Detection, Recognition, Identification.

1. INTRODUCTION

The ever growing complexity of the modern battlefield makes it more difficult to generate a recognized environmental picture. Constraints like less availability of qualified personnel and changing battleground conditions contribute to this growth. This resulted in a need for sensors which are able to discriminate threats better than sensors that are built today and sensors that can be adapted to a new operational scenario. There are several emerging technologies for discriminative imaging; one of them is multi- and hyperspectral sensors.

Present day hyperspectral sensors may contain a high number of spectral bands ranging typically from 100 to 200. The penalty for using such a high number of spectral bands is that the signal to noise ratio decreases. Also operational systems will be complex and expensive because of the wavelength discrimination element inside the hyperspectral sensor, this is especially true for the infrared wavelength range. If the complete hyperspectral image cube has to be processed for the detection of targets making use of both spectral and spatial target characteristics the huge amount of data of a hyperspectral image cube is troublesome. This complicates a near real time image processing solution. Band selection is therefore seen as an important step in realizing operational hyper/multi spectral imaging solutions

Most research involving band selection has been focusing only on the location of the bands. However, for a multispectal configuration very narrow bands are not practical, because this would require long integration times to get a good signal-to-noise ratio. Our research therefore not only looks at the location of the bands but also at the width of the bands. In previous research a first attempt has been made by developing an algorithm that first determines the best locations for the bands and than the best width of the bands. This however does not allow for a comparison between broad and narrow bands. Therefore a new algorithm has been written that is capable of finding a specified number of bands of a specified width. Two version of this algorithm have been developed, a fast one that can quickly find a solution but does not guarantee to find the best bands and an optimal algorithm, which searches all possible combinations but as a consequence takes a lot longer and can only be used if the number of required bands is small.

2. BAND SELECTION ALGORITHMS

In our previous research (Withagen et al., 2001) a first attempt has been made by developing an algorithm that first determines the best locations for the bands and than the best width of the bands. This however does not allow for a comparison between broad and narrow bands. Therefore a new algorithm has been written that is capable of finding a

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV,
edited by Sylvia S. Shen, Paul E. Lewis, Proc. of SPIE Vol. 7334, 733422 · © 2009 SPIE
CCC code: 0277-786X/09/$18 · doi: 10.1117/12.818468

specified number of bands of a specified width. Two versions of this algorithm have been developed, a fast one that can quickly find a solution but does not guarantee to find the best bands and an optimal algorithm, which searches all possible combinations but as a consequence takes a lot longer and can only be used if the number of required bands is small.

To evaluate a band combination the band selection algorithm uses a distance measure that quantifies the separation between classes. We used two different distance measures (Landgrebe, 2003). The Mahalanobis distance is defined as:

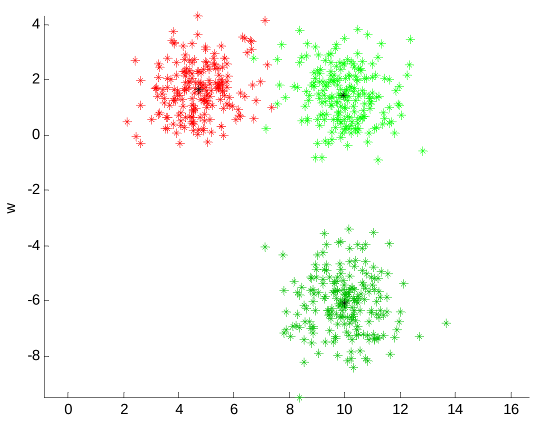
$$D = \sqrt{[\mu_1 - \mu_2]\Sigma^{-1}[\mu_1 - \mu_2]^T}$$

Where μ_1 and μ_2 are the class averages of class 1 (target) and 2 (background) and Σ is the covariance matrix of the classes.

When using the Mahalanobis distance measure one has to keep in mind that the following assumptions are made:

- The distributions of the classes are multivariate Gaussian distributions.
- The covariance matrix of these distributions is the same for all classes.
- The total number of pixels is large enough to accurately describe the covariance matrix (a rule of thumb is that the number of pixels should be at least 10 times the number of dimensions).

This distance measure is implemented by first transforming the feature-space and then calculating the Euclidian distance between the centres of the classes in this transformed feature space. The transformation makes use of the average covariance matrix of the different classes involved. The data is transformed to a different feature-space by multiplying it with the eigenvectors of this covariance matrix. The effect of this transformation is that the data is de-correlated. (for color images, please see electronic version of manuscript)



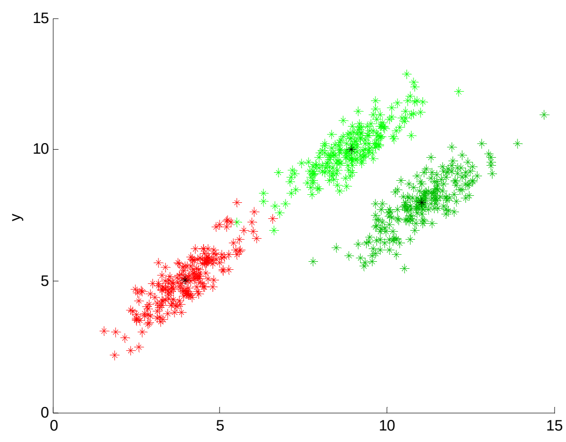


Figure 1: Example of a 2D feature-space (left) and its transformation (right).

The advantage of this can be seen in Figure 1. In the original feature-space (left figure) the distance between BG1 (light green) and BG2 (dark green) is larger than the distance between BG1 (light green) and T1 (red), and hence the separability between BG1 and BG2 is better. In the transformed feature-space (right figure) the classes which are most easy to separate also have the highest (Euclidian) distance. In the transformed feature-space the (Euclidian) distance is calculated between the centres of the different classes. The resulting set of distances is stored in a distance matrix. From this matrix a final single distance value can be derived in several ways depending on the experimental requirements (for example the minimum value of this matrix can be taken). Because we want to distinguish between background and target classes we choose the smallest distance between a target and a background class. This final distance value we will refer to as the quality (of a band combination). In the band selection algorithm this quality is maximized.

The Bhattacharyya distance is another distance measure that measures the distance between two multivariate Gaussian distributions. It is defined as:

$$B = \frac{1}{8}[\mu_1 - \mu_2]\left[\frac{\Sigma_1 + \Sigma_2}{2}\right]^{-1}[\mu_1 - \mu_2]^T + \frac{1}{2}\ln\frac{|1/2[\Sigma_1 + \Sigma_2]|}{\sqrt{|\Sigma_1||\Sigma_2|}}$$

One important difference with the Mahalanobis distance is that it does not take the average covariance matrix of the classes but keeps the class covariance matrices. The price that has to be paid is that now each class has to contain enough pixels to describe its covariance matrix accurately.

3. OPTIMIZATION OF BAND SELECTION

To find the best band combination two algorithms have been developed. One algorithm searches all possible band combinations. This algorithm can only be used if the number of required bands is small (<4) because calculation times increase exponentially with the number of bands. Therefore another algorithm has been written that searches in a more time efficient way, but as a consequence it is not guaranteed to find the best band combination.

The algorithms are implemented in Matlab® and make use of the toolbox PR Tools, a toolbox offered for free for academic research by the University of Delft in The Netherlands (reference 3). The main data-object of PRTools is called the dataset. In this dataset a large number of objects can be stored, each object consisting of a certain amount of features. We use this dataset-type to store our pixel-data. The dataset-object also makes it possible to label each object with an integer value, which can be used to divide the pixels in different classes.

We have analyzed the effects of two algorithms:

- Algorithm 1 is the fast algorithm. The way it selects its bands is by first selecting the band with the highest quality. Then it searches for a second band that in combination with the already found band gives the highest quality. Then it searches for a third band in the same way and so on till the number of required bands are found.
- Algorithm 2 (called the optimum algorithm) searches every possible combination of bands, which guarantees that it will find the band combination with the highest quality. Because calculation times increase exponentially with the number of bands, it can only be used if the number of bands required is small. The algorithm can also be used in a sub-optimal way, by defining a step-parameter (see below) higher than 1, in which case the algorithm gets faster. Besides a potentially better solution, algorithm 2 has another advantage with respect to algorithm 1: because it calculates the quality for each combination, an overview of all qualities can be made, giving extra insight in the problem.

To do the classification, the QDC (quadratic discriminant classifier) classifier provided by PRTools is used. This is a normal densities based quadratic classifier. For the algorithms several inputs are needed: number of background classes, bandwidth, shape, distance type, overlap, quality criteria. For algorithm 2 we also set a step and time estimation parameter. For *TimeEstimation*, if 1, the algorithm makes an estimate of the calculation time by calculating how many combinations it will have to evaluate and multiplying this with the quality evaluation-time, which it gets by making 5 evaluations and taking the average. A *Step parameter* (default is 1) is defined to use the algorithm in a faster, sub-optimal way. The idea behind this parameter is that when for example a bandwidth of 30 features is used, the band consisting of features 1 through 30 will almost be exactly the same as band 2-31. By setting a step of for example 3, the algorithm will only take into account bands 1-30, 4-33, 7-36 and so on, which can greatly decrease the calculation-time without sacrificing much of the optimality of the solution.

Table 1: Calculation times for algorithm 1, with max_overlap = 0 and band_width = 1. The number of pixels used is 1000.

number of bands	calculation time Bhattacharyya [s]	calculation time Mahalanobis [s]
2	8.5 16.0	
4	19.0 31.9	
6	30.9 47.7	
8	43.5 63.7	
10	57.3 80.3	

If the overlap parameter is set to 0 (no overlap allowed), the bandwidth has also some influence on the calculation time. The larger the bandwidth the faster the algorithm will be because after the first band has been picked all features that make up this band are excluded for the following bands so effectively the total number of features decreases. Table 1 shows the calculation times of the Matlab implementation of algorithm 1 for several numbers of bands using the Bhattacharyya or the Mahalanobis distance. The bandwidth used is 1. The number of pixels used is 1000. From this table can be concluded that the calculation of the Bhattacharyya distance takes on average about 30% less time then the calculation of the Mahalanobis distance.

Calculation times for algorithm 2 are a lot higher than those of algorithm 1. The relation between the calculation time, the number of bands and the total number of features is:

$$calc\_time \sim no\_pixels \cdot \frac{total\_features!}{no\_bands!(total\_features - no\_bands)!}$$

Depending on the value of the max_overlap parameter, the bandwidth also has a big influence on the calculation time. In table 2 some calculation times are given for three different bandwidths and 2 different numbers of bands. max_overlap is set at 0 and the distance measure is Bhattacharyya. The Mahalanobis distance measure shows the same pattern but the times are about 30% higher.

Table 2: Calculation times (hh:mm:ss) for algorithm 2, using the Bhattacharyya distance measure, with max_overlap = 0 and band_width = 1, 10 and 30. The number of pixels used is 1000. For step parameter = 1,2.

number of bands	**Step: 1**			**Step: 2**		
	1	**10**	**30**	**1**	**10**	**30**
2	00:10:03	00:07:20	00:04:58 00:	02:10 00:	01:50 00:	00:54
3	10:33:20	07:30:00	02:08:20 01:	13:20 00:	48:57 00:	16:10

4. GRAPHICAL USER INTERFACE (GUI)

The number of required spectral bands is assessed with HYBASE in a number of steps in a Matlab® environment. The first requirement is that the user has to input a hyperspectral image cube in which target and background are present. Subsequently regions in the image are selected and attributed to either background or target. Target boxes are coloured in red, background boxes in white. When all relevant target and background areas are selected the spectra of all pixels inside either the target or the background boxes are plotted at the bottom panel of the GUI (see Figure 2).

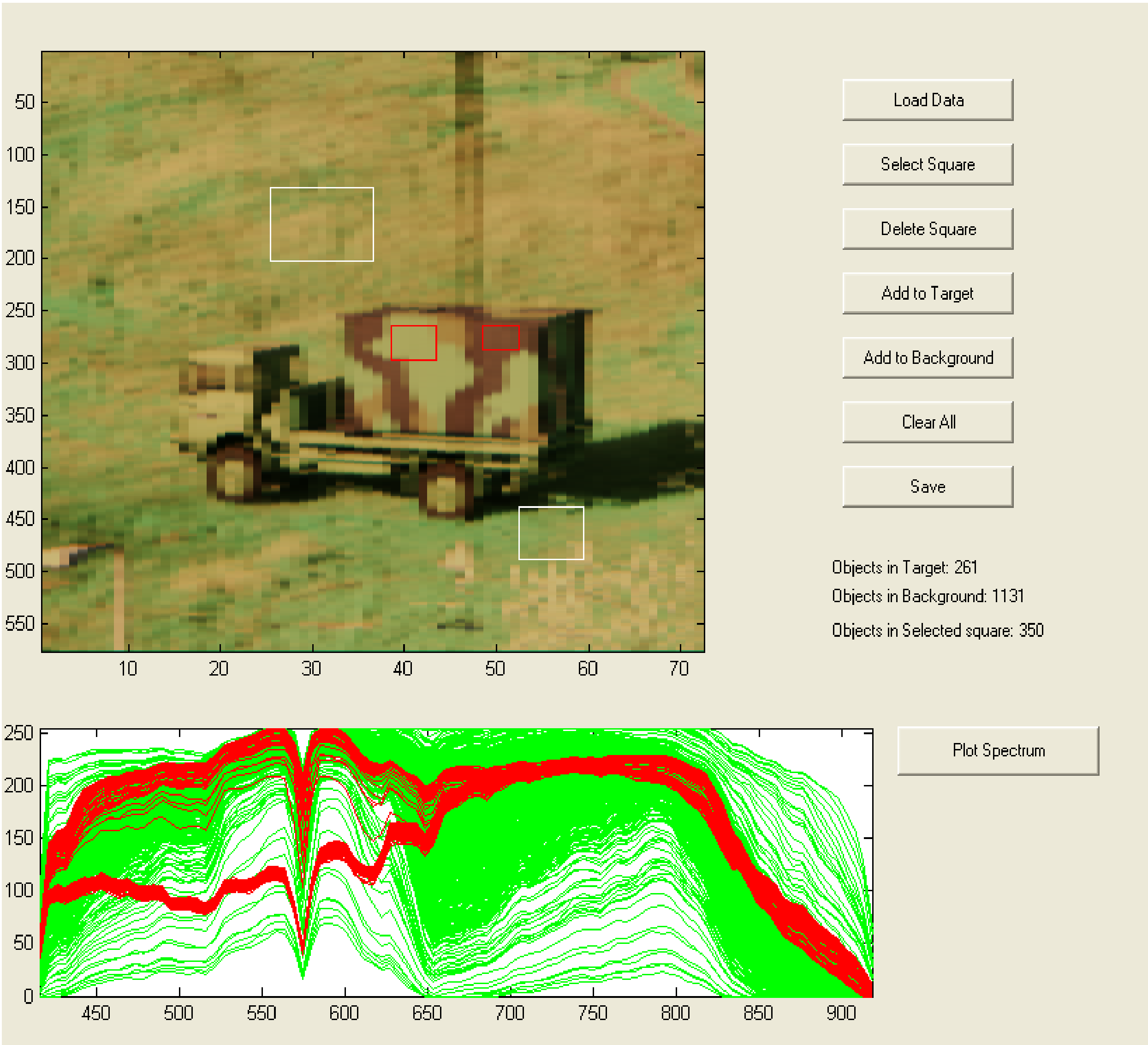


Figure 2: *Spectra of target areas and background areas are selected via Matlab® GUI*

(for color images, please see electronic version of manuscript)

Within the feature (or spectral band) selection tool the feature width and the maximum allowed overlap are input. Now the optimum position of these features is calculated. The result of the optimum band positions is plotted in the top panel of the GUI (see Figure 3), by vertical lines that are drawn over the spectra. Each band starts with a blue vertical line and ends with a black vertical line. The optimum spectral band positions are also outputted to the Matlab® command line.

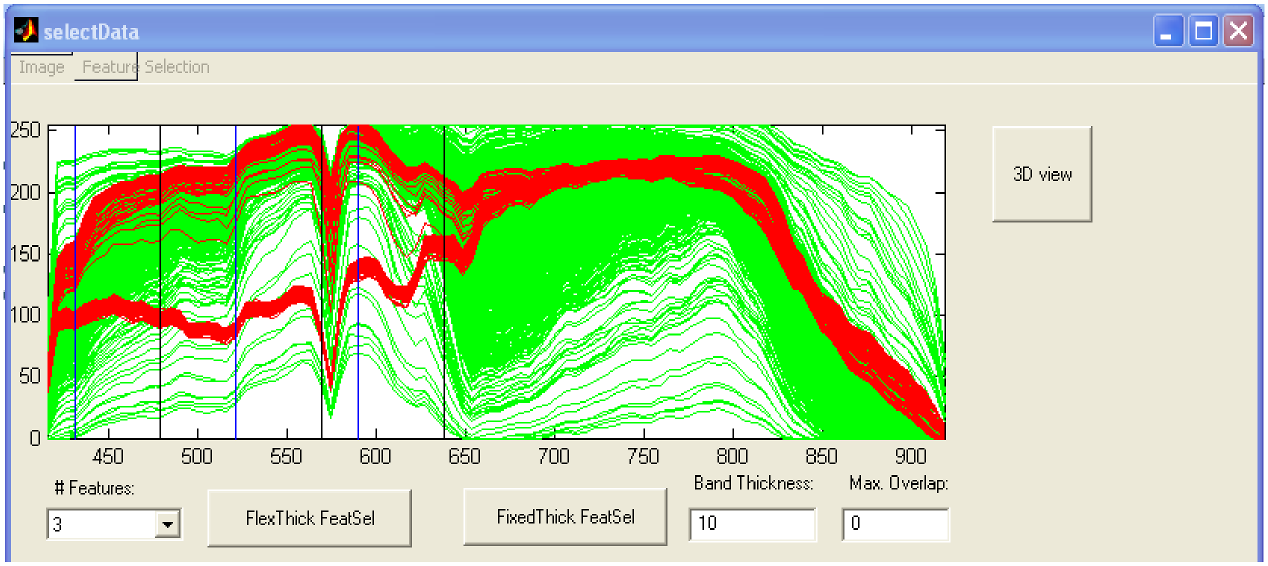


Figure 3: *Optimum position of spectral bands are indicated by vertical lines. Blue lines mark start of new band, black lines mark end of band.* *(for color image, please see electronic version of manuscript)*

If exactly three features are selected the position of the pixels in 3-dimensional feature space are plotted (see Figure 4). The user can rotate the cube for better visualization of the separation that has been achieved between target and background pixels using the selected number of features.

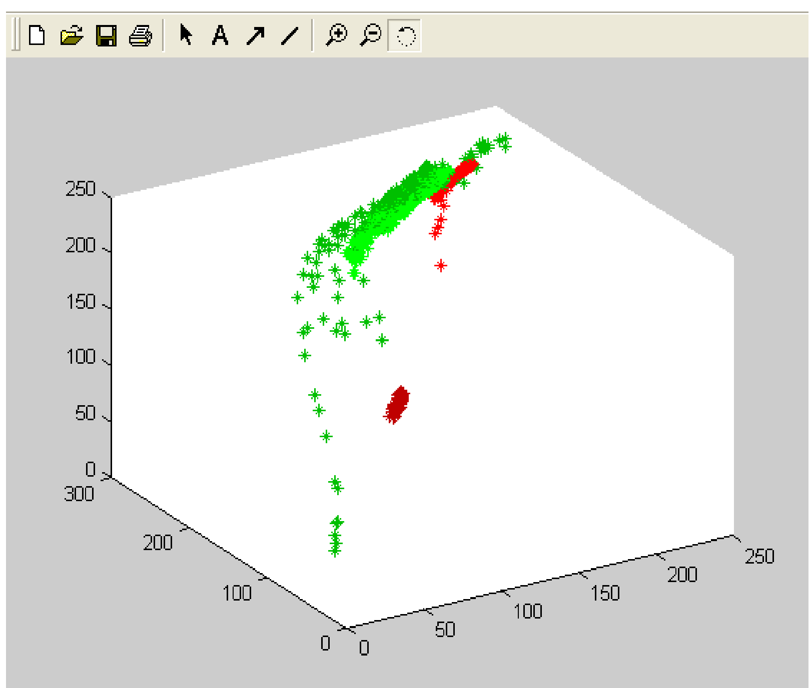


Figure 4: *Projection of pixels in 3-dimensional space spanned by three selected features (=spectral bands)*

5. BAND SELECTION RESULTS

To see how algorithm 1 (fast) performs compared to the algorithm 2 (slow, but optimal) a comparison has been made for a representative data set for the case that several targets are used with the following input settings:

bandwidth = 30
number of bands = 3
step = 2
overlap = 0

The comparison has been made for the Bhattacharyya as well as the Mahalanobis distance. Table 3 summarizes the result.

Table 3: Comparison between algorithm 1 and algorithm 2.

	algorithm 1 bhattacharyya	**algorithm 2 bhattacharyya**	**algorithm 1 mahalonobis**	**algorithm 2 mahalonobis**
Quality	0.4745	0.5450	2.3396	2.7131
QDC class. error *(misclassified pixels)*	75	76	81	78
band 1 *(feature numbers)*	80 - 109	81 – 110	6 - 35	7 - 36
band 2 *(feature numbers)*	114 - 143	117 – 146	110 - 139	95 -124
band 3 *(feature numbers)*	168 - 197	169 - 198	169 - 198	169 -198

Algorithm 1 performs quite well in finding the maximum quality. For the Bhattacharyya distance, algorithm 1 is 13% off, and for the Mahalanobis distance it is 14% off. The difference in classification error is even smaller.

The most surprising thing is that some of the found bands are really different, comparing band 1 for the Bhattacharyya distance and the Mahalanobis distance of algorithm 2 shows very different bands, while the classification error is similar. This raises the question if there are more band combinations that give similar results.

To investigate this, the quality of all band combinations (34220 in total) has been plotted for the Bhattacharyya distance and the Mahalanobis distance. These plots offer a revealing view on the significance of 'best bands'. There are in fact a lot of different band combinations that have a quality close to the maximum value, especially in the case of the Bhattacharyya distance. The periodic nature of the figures arises from the systematic way in which the band combinations were chosen. Because of that a certain band reoccurs every so often.

Having in mind that there is no direct translation of the distance measure into the classification result, it makes sense to not only look at the band combination with the highest distance, but also at the ones that come close to that. If the bands are plotted that are within 10% of the maximum value for the Bhattacharyya distance., the band around 11 µm has the highest contribution to the quality, since it is always present, see figure 5. When this band is chosen in combination with a band between 10 and 10.5 µm, the choice of the third band doesn't matter anymore. It can be anywhere between 8 and 9.7 µm. So the contribution of this third band is minimal and in this case just taking the first two bands would probably give a similar classification result.

In figure 5 (right side) also the pixel classification results are plotted as well. The red line in those graphs represents the classification error of the band combination with the highest quality. Its classification result is average compared to the classification when the other band combinations with quality within 10% of the maximum is being used.

Classification results have also been compared by using the Mahalanobis distance. This time there are only a total of 26 band combinations that are within 10% of the maximum and the bands are all around the same wavelengths. Surprisingly, these bands do not show up in the set of best bands found using the Bhattacharyya criterion. Still, the bands found with the Mahalanobis criterion give a comparable classification result. Apparently, the boundary of 10% within the maximum could be set lower to include even more band combinations.

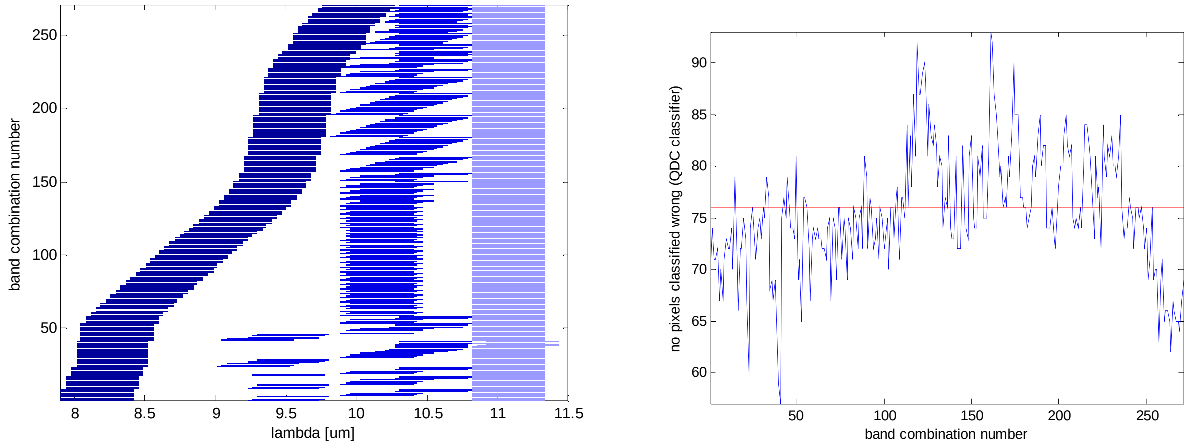


Figure 5: *Bands that have a quality (using the Bhattacharyya distance) within 10% of the maximum quality and associated miss-classifications*

6. APPLICATION OF BAND SELECTION

HYBASE is typically used in a system design study and these outputs can feed operational studies. Figure 6 shows the location of HYBASE in this design chain. Based on a hyperspectral data set in a relevant scenario one can make an analysis with HYBASE of the minimum number of required spectral bands, their widths and positions for the targets/backgrounds studied.

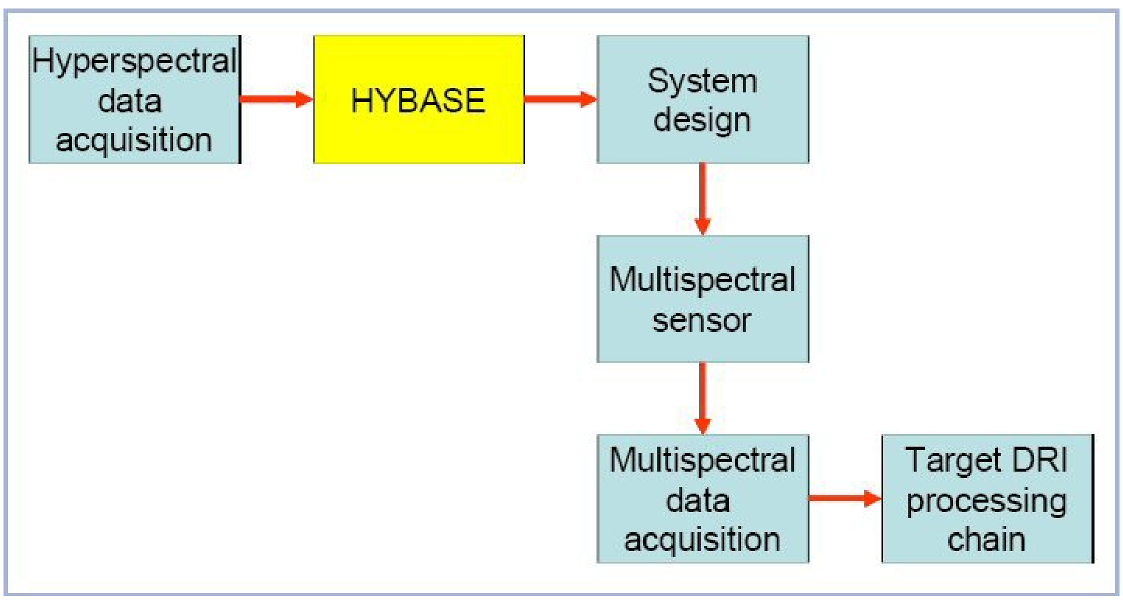


Figure 6: *Typical usage of HYBASE in system design*

When a multispectral sensor has been designed and realized the system will be used for data acquisitions. Multispectral target detection/classification tools will then be used to extract the relevant data. The target DRI (detection, recognition, identificaton) processing chain first pre- processes the acquired hypercubes (georeferencing, noise reduction, data normalization, temperature emissivity separation). Then targets are being detected using anomaly and signature based detection method in combination with change detection. Spatial information is used to reduce false alarm rates. Additional sensor data from e.g. high resolution imagers, radar and/or 3D laser radar is being used to classify and identify targets in a decision fusion process. Many of these algorithms run near real-time. Potential applications of sensor combinations are described in Schwering et al. (2007).

7. DISCUSSIONS AND CONCLUSIONS

An effective approach to optimum band selection in hyperspectral imaging has been demonstrated.

Target detection/recognition/identification is part of a more extensive software processing chain as depicted in Figure 6. First the acquired hypercubes are being pre-processed (georeferencing, noise reduction, data normalization, temperature emissivity separation). Then targets are being detected using anomaly and signature based detection method in combination with change detection. Spatial information is used to reduce false alarm rates. Additional sensor data from e.g. high resolution imagers, radar and/or 3D laser radar is being used to classify and identify targets in a decision fusion process. TNO has access to many of the tools in the processing chain which TNO can offer in addition to the HYBASE band selection tool.

Below are listed the conclusions from our research. What has to be kept in mind is that the conclusions are based upon one dataset with a frequency range from 7.7 µm to 12.1 µm, so mainly emissivity is measured:

- Algorithm 1 (the fast algorithm) performs good compared to algorithm 2 (the optimal algorithm). The band combinations found by algorithm 1 have a quality value within 15% of the quality found by algorithm 2, while the calculation time is a lot smaller.
- Using the Bhattacharyya distance as a measure for the separation of the different classes gives comparable results as the Mahalanobis distance.
- Although no thorough study has been done between the relation of the quality and the classification error, in some cases the difference in classification error can be very big for similar qualities (up to 100% difference)
- Often, there is a whole set of different band combinations that have a comparable quality and classification result. This set is revealed by plotting the band combinations having a quality within a certain percentage of the maximum quality.
- As a consequence of the above two points, the band combination with the highest quality does not necessarily have the lowest classification error.
- The location of the best bands depends strongly on the choice of target and backgrounds.
- For a good classification result clean spectra of the targets are required. Target masks for semi-hidden targets are useless, since they contain target as well as background pixels.
- If the number of bands increases the quality increases and the classification error decreases. Although other research shows that there is an optimal number of bands for the classification error, this did not show up in our results. This optimum is due to the fact, that when the number of bands increases, statistical values used to describe the feature-space like the covariance matrix can be predicted less accurate. That this optimum did not turn up in our results is probably due to the fact that we used the areas that were classified also to train the classifier.

The influence of the bandwidth on the quality is substantially less than the influence of the number of bands. This is probably because the spectra in the thermal infrared region (7.7 µm to 12.1 µm) involved did not have any sharp features. There is also no clear relation between the bandwidth and the quality, sometimes the quality increases.

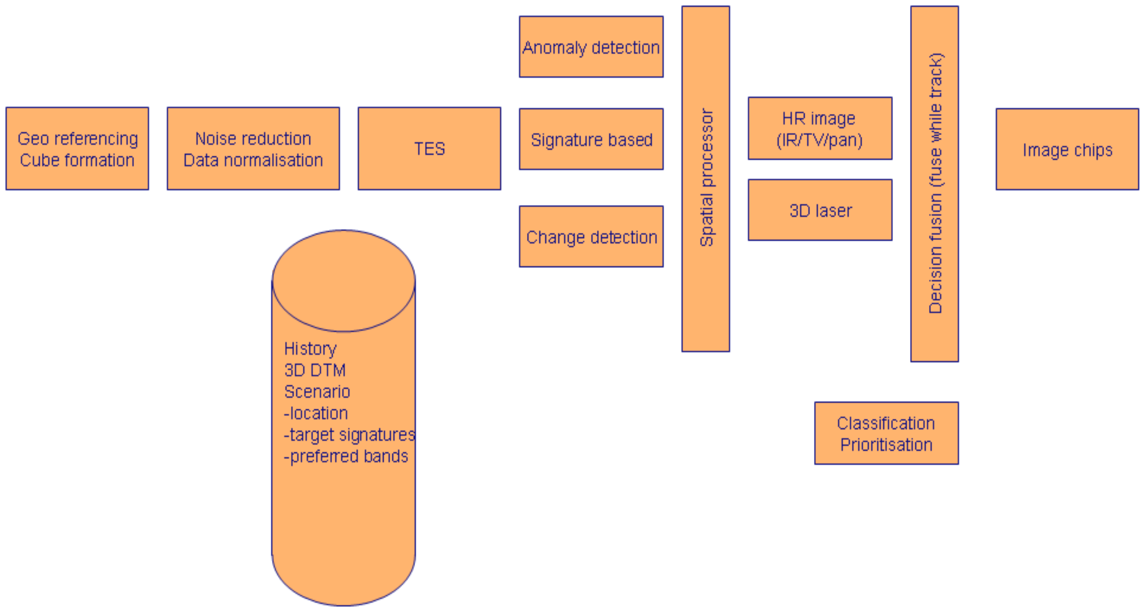


Figure 7: Multi spectral image processing chain

If the complete hyperspectral image cube has to be processed for the detection of targets making use of both spectral and spatial target characteristics the huge amount of data of a hyperspectral image cube is troublesome. This complicates a near real time image processing solution. Band selection is therefore an important step in realizing operational hyper/multi spectral imaging solutions. In Figure 7 we present the basic TNO processing chain for automatic target data processing of hyperspectral image information. This serves as the big picture in the research, consisting of real-time on-line steps, combined with supporting off-line data mining activities.

Most research involving band selection has been focusing only on the location of the bands. However, for a multispectal configuration very narrow bands are not practical, because this would require large integration times to get a good signal-to-noise ratio. Our research therefore not only looked at the location of the bands but also at the width of the bands.

8. ACKNOWLEDGEMENTS

We gratefully acknowledge the stimulating discussions with our colleague C.J. den Hollander and financial support from the NL MOD in programme V513.

9. REFERENCES

1. P.J. Withagen, E. den Breejen, et. al. " Band selection from a hyperspectral data-cube for a real-time multispectral 3CCD camera", Proceedings of SPIE Vol. 4381, pp 84-93, 2001
2. Landgrebe, D.A., "Signal Theory Methods In Multispectral Remote Sensing", Hoboken, NJ: Wiley, 2003
3. University of Delft, www.prtools.org
4. Schwering P.B.W., van den Broek S.P., van Iersel M., 2007, in 'Infrared Technology and Applications XXXIII', eds. B.F. Andresen, SPIE Vol. 6542-100, 654230, Orlando Florida (USA), April 9-13, 2007: "EO System Concepts in the Littoral"

5. Briottet X, Boucher, Y., Dimmeler, A., Malaplate, A., Cini, A., Diani, M., Bekman, H., Schwering, P., Skauli, T., Kas en, I., R enhorn, I., Klas en, L., Gil more, M., Ox ford, D.: 2006, in ' Targets an d Back grounds: XII, Characterization an d R epresentation', ed. W .R. W atkins, D . Cl ement, SP IE V ol. 6 239-paper163, O rlando Florida (USA), April 17-18 2006: "Military Applications of Hyperspectral Imagery"