Band selection from a hyperspectral data-cube for a real-time multispectral 3CCD camera

Paul J. Withagen, Eric den Breejen, Eric M. Franken, Arie N. de Jong, Hans Winkel

Electro-Optical Systems, TNO Physics and Electronics Laboratory, PO Box 96864, NL-2509 JG The Hague, The Netherlands

ABSTRACT

Given a specific task, like detection of hidden objects (i.e. vehicles and landmines) in a natural background, hyperspectral data gives a significant advantage over RGB-color or gray-value images. It introduces however, a trade-off between cost, speed, signal-to-noise ratio, spectral resolution, and spatial resolution. Our research concentrates on making an optimal choice for spectral bands in an imaging system with a high frame rate and spatial resolution.

This can be done using a real-time multispectral 3CCD camera, which records a scene with three detectors, each accurately set to a wavelength by selected optical filters. This leads to the subject of this paper: how to select three optimal bands from hyperspectral data to perform a certain task. The choice of these bands includes two aspects, the center wavelength, and the spectral width. A band-selection and band-broadening procedure has been developed, based on statistical pattern recognition techniques.

We will demonstrate our proposed band selection algorithm, and present its classification results compared to red-green-blue and red-green-near-infrared data for a military vehicle in a natural background and for surface laid landmines in vegetation.

Keywords: Hyperspectral, multispectral, band selection, camouflage, real-time, feature selection, band broadening

1. INTRODUCTION

To test the concept of band selection from hyperspectral data for a multispectral camera in order to increase the detection performance, we studied two tasks:

1. The detection of a military vehicle in a natural surrounding
2. The detection of surface laid landmines in vegetation

For these tasks hyperspectral data is recorded and band-selection is performed. On the full scene images for the three selected bands, automatic classification is performed.

The band-selection is performed for our 3CCD camera [2]. A camera with a beam splitter and three Charge Coupled Device (CCD) arrays, each operating like a normal gray-value camera, but sensible to light from a predefined wavelength band, using optical cut-off filters. For this camera three bands are selected for each task. The performance of the classification algorithm is compared to the performance of the same algorithm on normal Red, Green, and Blue (RGB) data and on data from our DuncanTech Red, Green, and Near InfraRed (RGNIR) camera as we got it from the factory.

In section two the used sensor systems will be described. Then a description of the data is given in section three and the band selection algorithm is described in section four. In section five the classification algorithms are described and in section six the results of band selection and classification are given. Finally conclusions are drawn in section seven.

2. SENSOR SYSTEMS

To record hyperspectral data we have an ImSpector system [4], which is a combination of a grating and prisms. It efficiently produces a spectrum of an entire line and projects this on a CCD camera, see figure 1. The system is attached to a Sagebrush pan-tilt unit, so a hyperspectral image cube can be build by

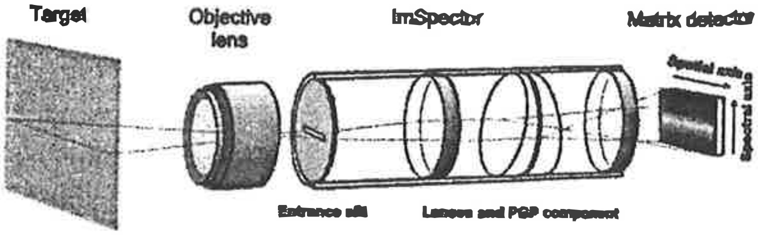


Figure 1 Basic operation of an ImSpector hyperspectral imager. Courtesy Specim.

Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII, Sylvia S. Shen, Michael R. Descour, Editors, Proceedings of SPIE Vol. 4381 (2001)

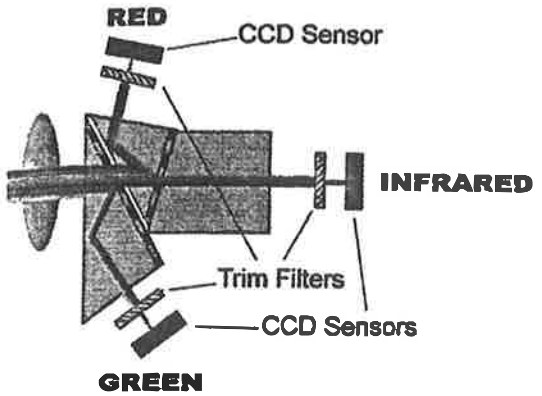


Figure 2 Basic operation of the DuncanTech 3CCD camera. Courtesy DuncanTech.

scanning in the direction perpendicular to the spatial plane of the ImSpector.

The ImSpector (type V9) creates a spectrum from 430 to 900 nm with a spectral resolution of 6 nm. This covers the entire visual- and near-infrared range, especially suitable for use in areas containing vegetation. The spectrum is projected on a BPW34 ½ inch black-and-white CCD camera, and the output signal is digitized using an eight-bit frame-grabber.

The spectral ax has been calibrated using three small-band optical filters distributed over the wavelength-range.

The hyperspectral image cubes have *576xNx90* points, with 576 the number of spatial points in one scan-line, *N* the number of recorded scan positions, and 90 the number of spectral bands.

The second system is a DuncanTech 3CCD CIR camera [2]. This camera contains a beam-splitter consisting of prisms with an optical coating. The input beam is split into three parts with wavelengths 500-580 nm (green), 620-700 nm (red), and 740-900 nm (near infrared). For each of these bands, a separate CCD element records the signal at 7.5 frames of 1392x1040 pixels per second with 8 or 10 bits per pixel digital output. For each of the beams, an optical trim-filter can be used to accurately select a sub-band, see figure 2.

The third system is a normal RGB color camera with 768x576 pixels and analog PAL output.

The data from the ImSpector system can be used to simulate cameras with various spectral responses, for example an RGB camera and any configuration of the 3CCD camera. After band-selection is done, the 3CCD camera can be adapted to the calculated specifications. Depending on the desired wavelengths, only the trim-filters need to be replaced, or in addition the beam splitter needs to be adjusted with different optical coatings.

3. DATA DESCRIPTION

The data used for this article was recorded from the TNO-FEL laboratory in The Hague, The Netherlands, at the end of October / beginning of November 2000 with a partly clouded sky. Below, a short description of the two datasets will be given.

3.1. "Truck"-dataset

The scene consists of a truck painted in camouflage colors. The truck was surrounded by different kinds of grass. See figure 4 for recorded RGB, RGNIR, and ImSpector images of the truck. Because of changing light-conditions during the 10 minutes it took to record all ImSpector images, a darker bar can be seen in the two ImSpector images.

The ImSpector was attached to a pan-tilt unit, which was at its turn fixed to a platform on which the DuncanTech 3CCD CIR camera and a color camera were attached. All three cameras were aligned and had a field of view of about 6x4.8 degrees (21x16 degrees for the 3CCD camera, so only a small region of the entire image could be used, the camera was not focussed optimally for this region). The ImSpector had its scan-line vertical, so it had to be moved in the horizontal direction by the pan-tilt unit to create a 2D image. The width of the ImSpector scan-line is 0.1 degrees, 72 steps are used to create the hyperspectral image cube. The positioning of the pan-tilt unit and digitizing of the signal from all three cameras was performed automatically. All cameras produced 8 bit output images that were recorded by a PC with a frame grabber. In this way a hyperspectral image cube and two color images could be recorded directly after each other, with a minimum of scene variation. All images were warped to the ImSpector image.

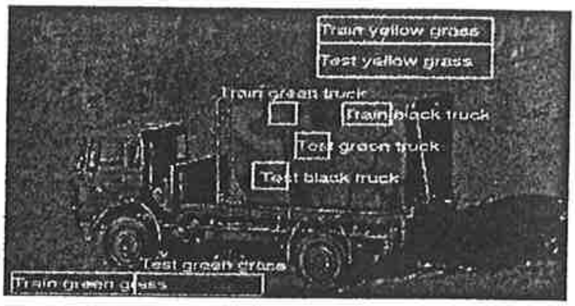


Figure 3 Train- and test-sets for the "Truck"-data.

By hand four classes were defined (see figure 3):

- Black camouflage paint (205 pixels train, 63 pixels test)
- Green camouflage paint (246 pixels train, 100 pixels test)

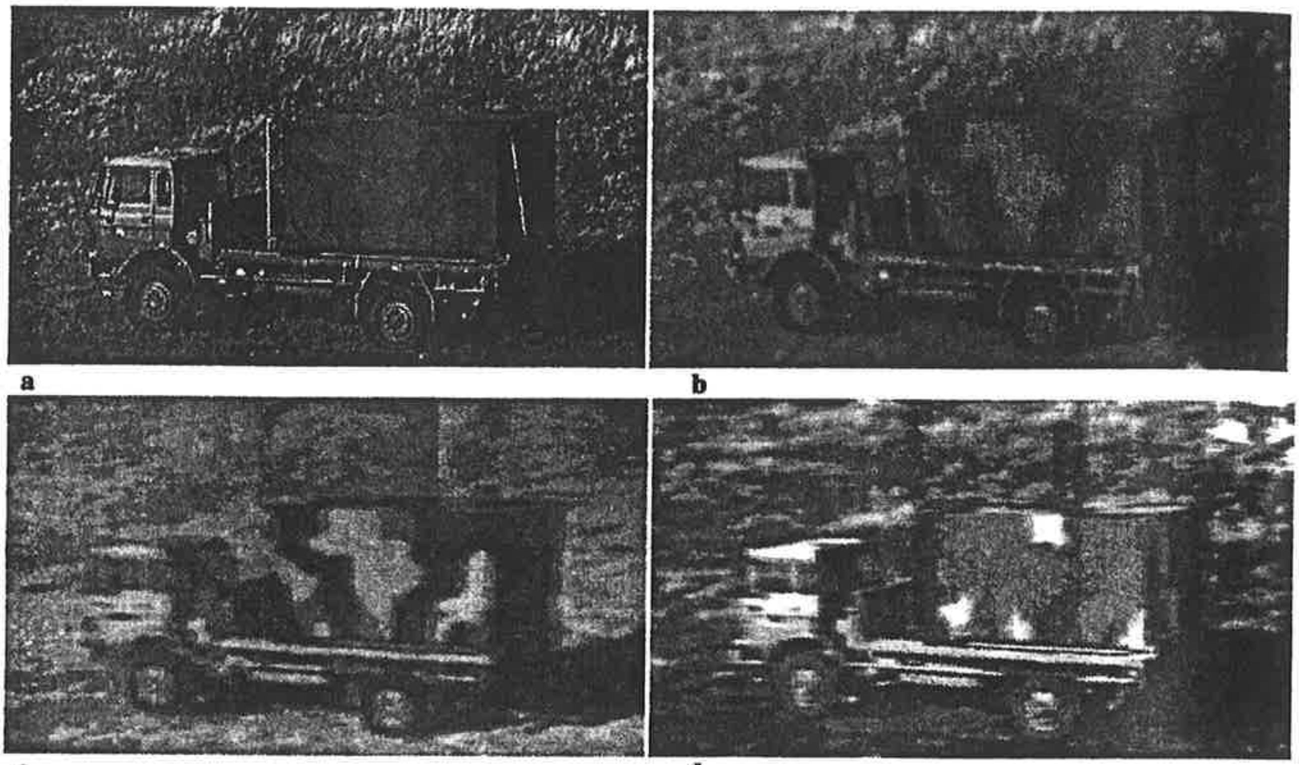


Figure 4 Four images of the "Truck"-dataset, a) RGB image made by a normal color camera, b) part of the RGNIR image made by the DuncanTech 3CCD camera (the camera was slightly out of focus), c) and d) show two typical ImSpector images for different bands.

- Green grass (231 pixels train, 220 pixels test)
- Yellow grass (697 pixels train, 340 pixels test)

For each of these classes a rectangle of pixels was defined for training data, and one rectangle of pixels for testing data. The classification results in table 1 are only for these (small) test-areas.

3.2. "Mines"-dataset

The scene consists of mine-like objects laid in a regular pattern of 5x5 mines with a spacing of 45 cm on a field of grass. All objects were different in model, shape, and color.

The ImSpector was attached to a movable platform (see [5]) at a height of 1.9 meters above the grass. The elevation of the ImSpector was 25 degrees and the field of view of the (horizontal) scan-line 22 degrees. So the scan-line was at a distance of 4 meters in front of the platform and had a width of 1.7 meters. The platform was moved in steps of one cm, and at every scan-line, 5 seconds of video was recorded. This video was then digitized using a frame-grabber, and averaged to produce one noise-reduced image per position. A total of 215 positions have been recorded. See figure 6 for some of the images. With the DuncanTech 3CCD camera, an image has been recorded from the same camera location, but because it had a much wider field of view, only a part of the image was used. From a different camera location, an image was taken with a digital color camera. The ImSpector data was recorded between 12:00 and 13:00, and the RGB and RGNIR data between 14:00 and 14:30. All images were warped to the ImSpector image.

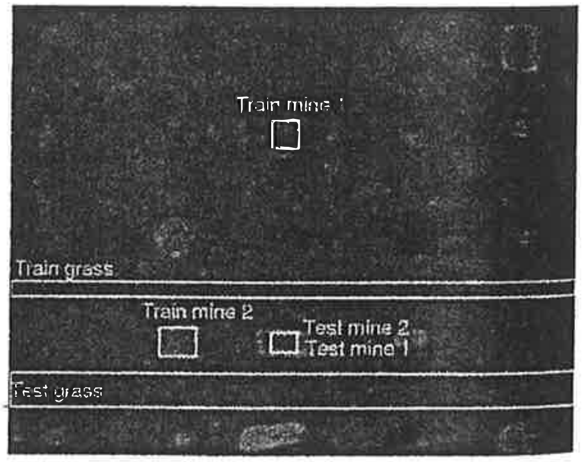


Figure 5 Train- and test-sets for the "Mines"-data.

By hand three classes were defined (see figure 5):

- Mine 1 (370 pixels train, 286 pixels test)
- Mine 2 (570 pixels train, 286 pixels test)
- Background (4557 pixels train, 9765 pixels test)

For each of these classes a rectangle of pixels was defined for training data, and one rectangle of pixels for testing data (there was only one test-area for both mines classes). The classification results in table 1 are only for these (small) test-areas.

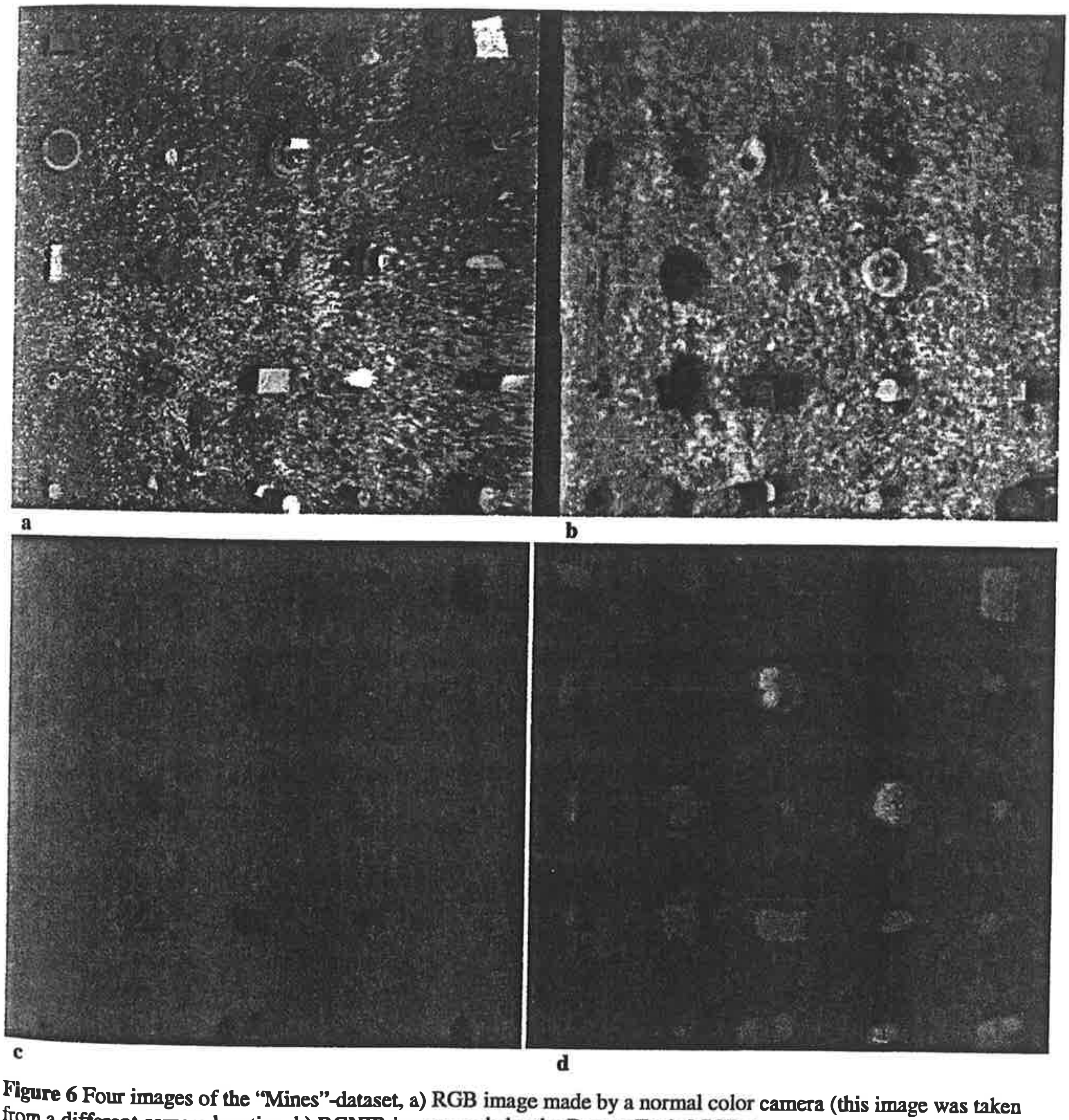


Figure 6 Four images of the "Mines"-dataset, a) RGB image made by a normal color camera (this image was taken from a different camera location, b) RGNIR image made by the DuncanTech 3CCD camera, c) and d) show two typical ImSpector images for different bands. All images were warped to the ImSpector image.

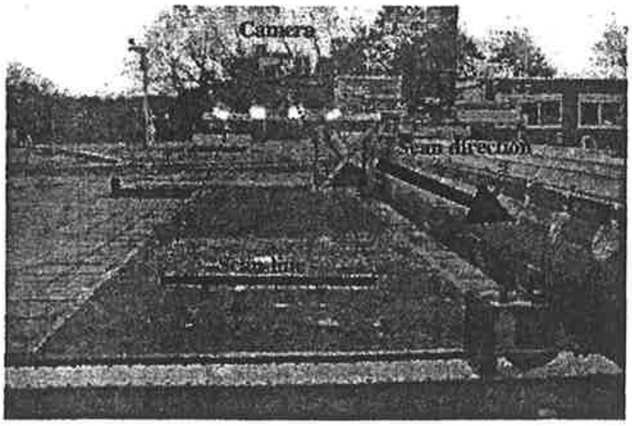


Figure 7 An image of the moving platform with a camera mounted on it for the collection of the "Mines"-dataset.

4. BAND SELECTION

In order to make a real-time multispectral imaging system optimal for the selected applications, we need to know how to choose three optimal wavelength bands for the multispectral camera

The problem at hand is not only to select the spectral wavelength at which the difference between the object and the background is optimal. Also the optimal width of the band has to be determined. This makes it a non-standard feature selection problem.

First we will describe the band selection algorithm we used to find the optimal center wavelength and bandwidth of all three bands, and after that the distance measure we used for the band selection.

To be able to compare the classification results of the three optimal bands to standard cameras, results will be given on full scene for:

- Three optimal bands in the 430-900 nm range
- Simulated RGB camera
- Simulated standard configuration of the 3CCD camera (so with standard trim filters, RGNIR)
- Real RGB camera
- Real RGNIR camera

for both the truck and the mines data sets.

4.1. Band selection

In order to obtain an optimal solution, we should make a feature-set containing all bands, with all bandwidths. When we limit the width of the selected bands to eleven times the spectral resolution, this would for our dataset containing 90 bands mean $11 \times 90 - 11 \times (11-1)/2 = 935$ bands. To select the optimal combination of bands, $935 \times 934 \times 933/6 = 1.4\ 10^8$ evaluations would be necessary.

To avoid this huge amount of computations, we developed an algorithm that combines forward feature selection [1] and band broadening. This algorithm looks like:

1. Calculate how each of the narrow bands performs in combination with the already selected bands (none in the first iteration) using a performance criterion. Select the best-performing band on the train regions.
2. Calculate the criterion value for the already selected bands and broader versions of the band selected in the previous step. Select the best band-width.
3. If the desired number of bands is not obtained, go to step 1, otherwise band selection is ready.

To obtain broader bands in step 2 all combinations of neighboring bands that contain the band selected in step 1 are possible. For computational reasons only bands within a constant width of the selected wavelength band are considered. A broad band is constructed by simply adding the intensity-values from a range of bands for each pixel, and dividing this by the number of bands added to each other.

For the same maximal bandwidth, this algorithm requires $((11-1)/2)^2$ evaluations per added band (we limit the search area to a region with the width of the maximal bandwidth, centered around the narrow band). This gives a total of $25^3 = 15.625$ evaluations.

The proposed forward feature selection procedure requires only a small number of computations in relation to other feature selection algorithms like backward and branch-and-bound feature selection. The drawback of this procedure is that it may produce a sub-optimal solution.

4.2. Distance measure

The use of the feature selection algorithm reduces the number of computations considerably, but there are still too many evaluations to train and evaluate a classifier for each evaluation. To reduce the amount of computations even further, instead of training and evaluating a classifier, the Mahalanobis distance is used to evaluate the combinations of bands. This is a measure for the distance in feature-space between a pixel and the center of all classes. The distance is then corrected for the mean and covariance of the data, so

$$D_M(x,A) = (\mu_A - x)^T G^{-1} (\mu_A - x)$$

Where A is a class, x is the pixel we want to calculate the distance for, μ_A is the mean of class A, and G is the covariance matrix of all data, see also [6]. Feature selection is performed on a training-set, which is manually selected from the image. The Mahalanobis distances of all points in one class to the mean of all other classes are calculated. The best combination of bands gives a maximum for the summation of all these distances for all classes.

5. CLASSIFICATION

After band-selection, three numbers will result for each pixel. These numbers are the amount of reflection in a particular wavelength band. These three numbers should be specific for the object we would like to detect. To test how specific these numbers are for the detection task we used the manually selected rectangular training- and test regions. All classifiers are trained on the training regions, and are then used to classify the entire image. The classification error was calculated for the relatively small test regions.

The pixels that have to be classified are from an image, so points that are close to each other are likely to have the same class. This information can be taken into account using a 2-stage classification procedure. This procedure starts with normal classification. Then all points with low classification accuracy are classified again, but now with an adjusted classifier. This classifier is adjusted so the a priori probability for the class which is most frequent under the eight neighbors is half, while the other classes all have equal probability, and together have a probability of half. In our case, those pixels that have a classification accuracy lower then the median of the classification accuracies of all pixels in the image are classified again in the second stage.

The classification is done using a linear discriminant function based on normal distributions of the classes, see [1,6]. A linear discriminant function D is a function of the form $D_A = a_{Ak}x_{Ak} + a_{Ak-1}x_{Ak-1} + \ldots + a_{A0}x_{A0}$, in which x_{Ak} is feature (or band) k of class A and a_{Ak} is a parameter adjusted by training (see [3]). This function is derived for each class. The object will be assigned to the class for which the discriminant function is the largest. The calculation of all discriminant functions for both classification stages is an offline procedure, and can be done at the same time as training the initial classifier. See figure 9 for a comparison between 1-stage and 2-stage classification results.

6. RESULTS

The results after feature selection (to obtain the optimal three bands and bandwidths), training the classifier, and classification, are shown in figures 8 and 10 for the "Mines"-dataset and the "Truck"-dataset respectively. The percentages of errors in the test regions are given in table 1. Only errors between mine versus background and truck versus background were taken into account, so for example a green-grass pixel classified as yellow-grass is not considered a wrong classification.

The test-regions were chosen by hand, and consisted of all the pixels in a rectangle for each class. Because these rectangles were quite small for some classes, the numbers given in table 1 only give an impression of the quality of the classifier. Because the training- and test regions were the same for the simulated RGB, simulated RGNIR, and three-optimal-bands data, these numbers in table 1 can be used to compare these three cases. The training- and test regions of the real RGB and real RGNIR data were slightly different because the warping was not perfect.

It may however be better to decide on the ability of the classifier to distinguish between different classes by judging from the complete images as presented in figures 8 and 10.

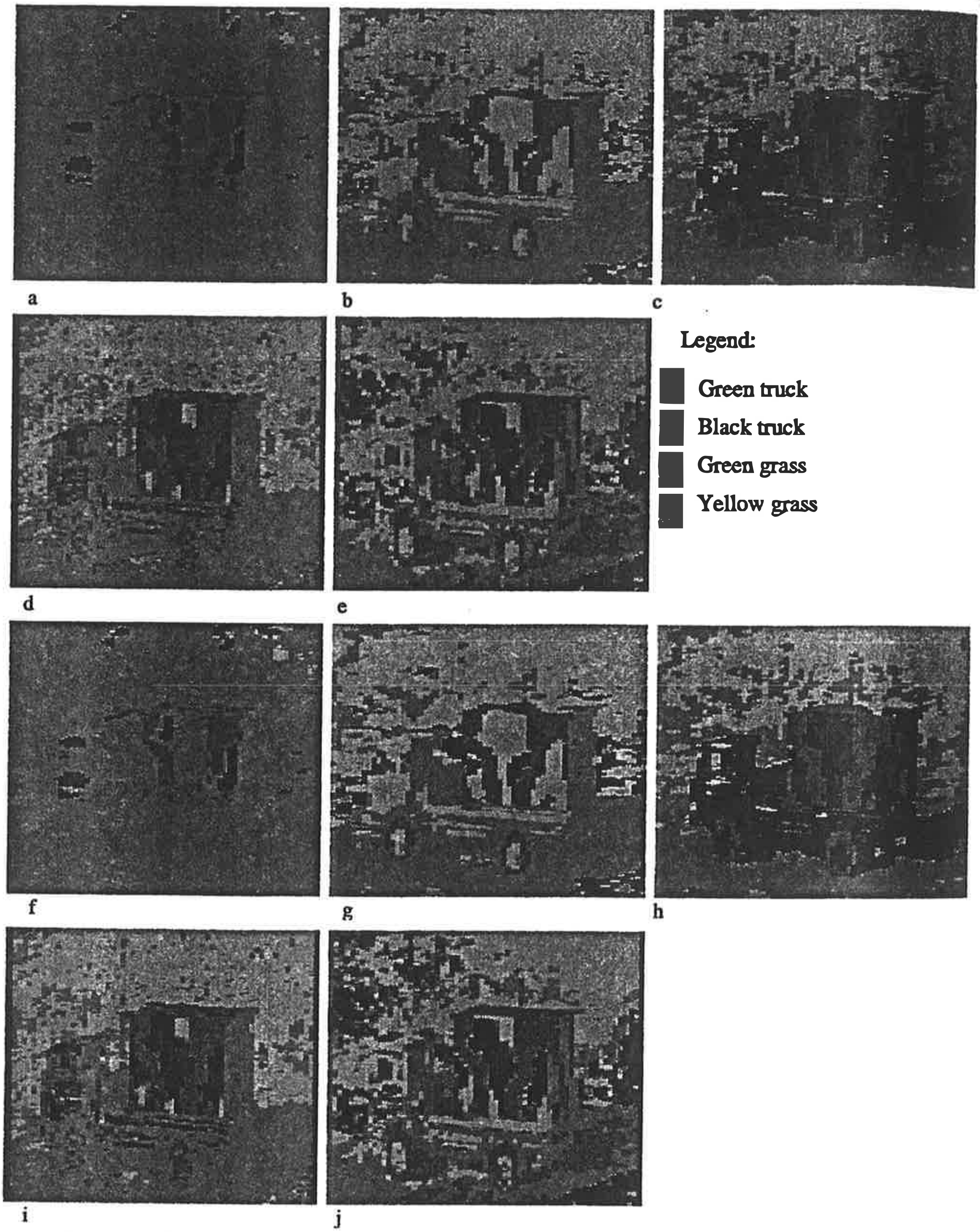
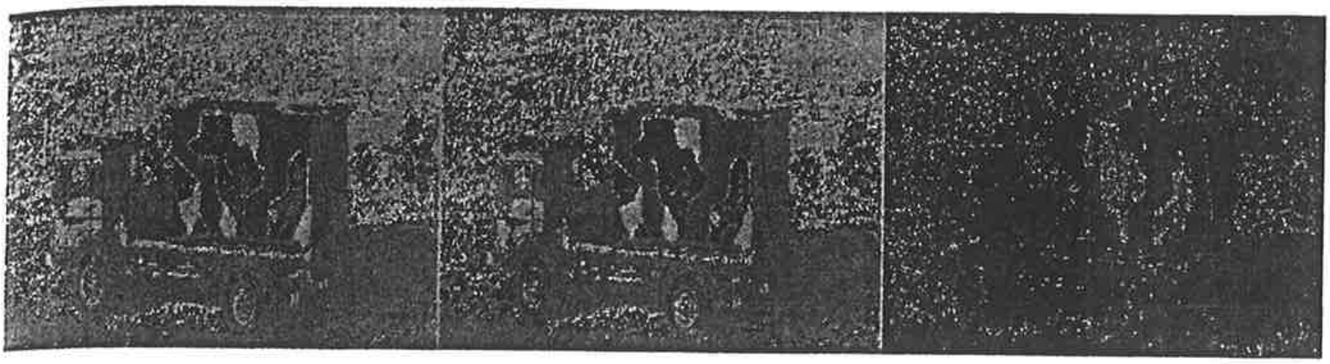


Figure 8 Classification results for the Truck-data. a) Shows the 1-stage classification result on simulated RGB data, b) on simulated RGNIR data, and c) on the three bands selected using band-selection, d) on real RGB data, and e) on real RGNIR data. f) Shows the 2-stage classification result on simulated RGB data, g) on simulated RGNIR data, and h) on the three bands selected using band-selection, i) on real RGB data, and j) on real RGNIR data.



a b c

Figure 9 Differences between 1-stage and 2-stage classification for real-RGB "Truck"-data. a) Shows the classification result after one classification stage, b) after 2 classification stages, and in image c) all white pixels changed class between the first and second classification-stage.

When comparing the results for real and simulated cameras, quite a large difference can be seen. This is probably caused by Imperfect simulation of the cameras and the difference in spatial resolution.

Judging from the simulated "Truck"-data, 3 optimal bands is considerably better then RGB and RGNIR. The "Mines"-data does not show this as clear, probably because all mine-like objects were different in size, shape, material, and color, so training on only two mines is far too less.

Some errors can also be seen on locations where there are shades in the images. This is because the effect has not been taken into account yet.

Instead of improving the results, the second stage of the 2-stage classification algorithm decreases the classification results in some cases, for example for the real RGB "Mines"-data.

Table 1 Classification errors for the test regions using three bands and 2-stage linear classification.

	"Mines"-dataset		"Truck"-dataset	
Data	**1-stage**	**2-stage**	**1-stage**	**2-stage**
Simulated RGB	3 %,	1 %	0 %	21 %
Simulated RGNIR	3 %	6 %	0 %	7 %
3 Optimal bands	0 %	0 %	0 %	0 %
Real RGB	6 %	5 %	0 %	0 %
Real RGNIR	0 %	0 %	0 %	0 %

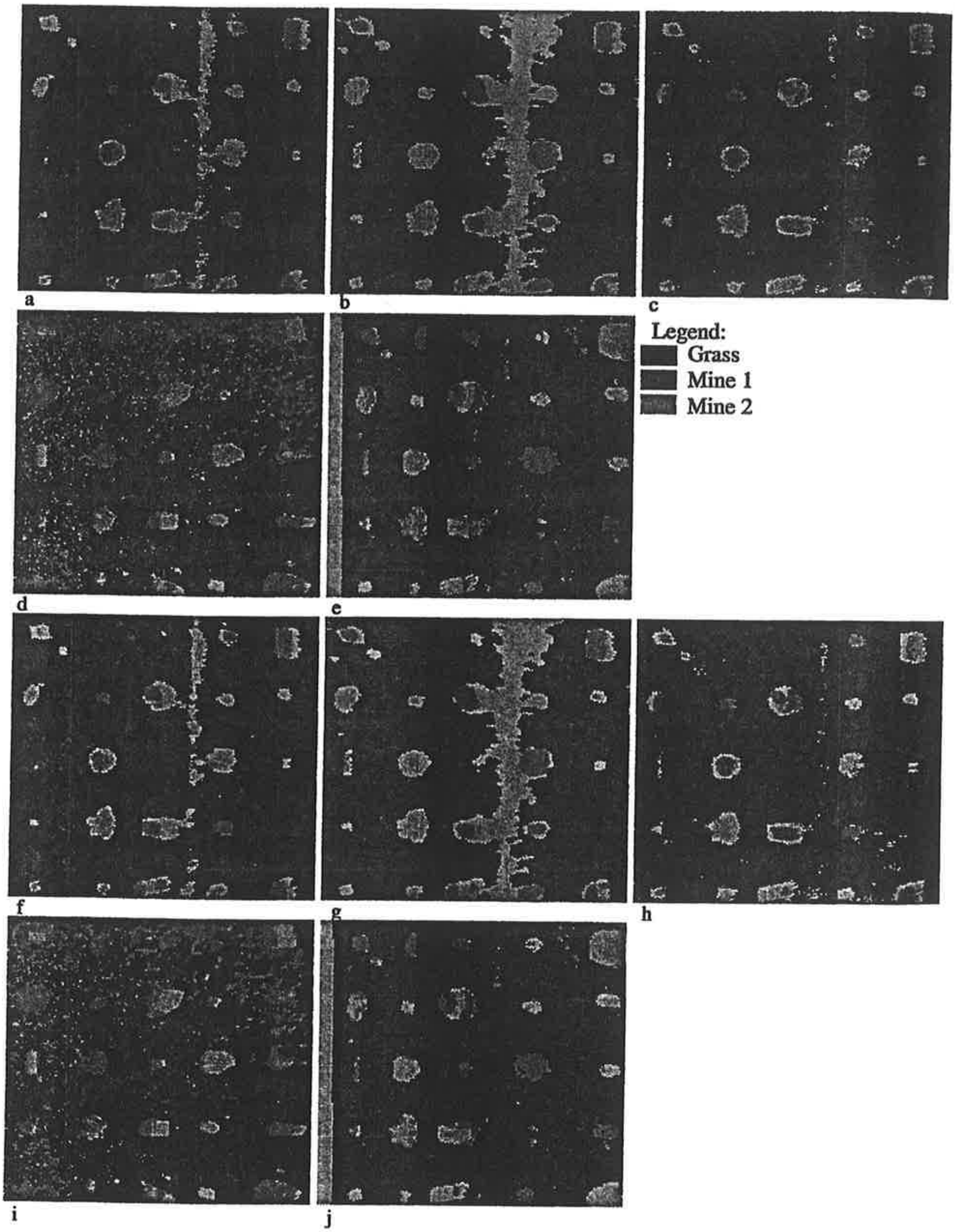


Figure 10 Classification results for the Mines-data. a) Shows the 1-stage classification result on simulated RGB data, b) on simulated RGNIR data, and c) on the three bands selected using band-selection, d) on real RGB data, and e) on real RGNIR data. f) Shows the 2-stage classification result on simulated RGB data, g) on simulated RGNIR data, and h) on the three bands selected using band-selection, i) on real RGB data, and j) on real RGNIR data.

7. CONCLUSIONS AND DISCUSSION

The goal of this research was to investigate whether the use of band-selection for a real-time multispectral camera would improve the detectability of military vehicles in a natural background and surface-laid landmines in vegetation. In order to do this; we compared automatic classification performance for real and simulated RGB (Red, Green, and Blue), real and simulated RGNIR (Red, Green, Near InfraRed) and three optimal bands in the 430 to 900 nm wavelength range, chosen automatically using band selection.

We developed a band-selection algorithm capable of finding three good bands (both the center wavelength and the bandwidth) from a hyperspectral datacube. Using these three optimal bands, a two-stage classifier was trained, and used to classify two datasets. The two-stage classifier was designed to incorporate spatial information in the classification process.

For the "Truck"-dataset, the use of 3-optimal-bands gave quite an improvement in detection results. The improvement was smaller for the "Mines"-data, probably due to the wide variety of objects. The use of the 2-stage classification did not always make an improvement to the results; sometimes it made the classification results worse.

It has to be noticed that automatic classification is only one of the applications where the selections of bands gives better results. Another, equally important application is detection and classification by human operators. The improvement of the use of selected bands is much harder to measure for this application, so it has not been investigated in this research yet. The false-color images the multispectral camera creates might be difficult to interpret by humans, this requires training for the operators and further study is necessary to select the optimal color-combination in which the bands are displayed

Further research is also necessary to evaluate the improvement of the multispectral camera over RGB and RGNIR for other (preferably larger) and more diverse datasets. For the "Mines"-data for example, a more representative training and evaluation data set for both object and background classes need to be recorded.

The occurrence of shadings introduces errors in the current classification. So some kind of normalization should be introduced. This normalization should not change the colors of the pixels, only the intensity, this should be investigated further. We used equal apriory probabilities for all classes in the first classification stage. This might not be the case in real applications. This has to be investigated, together with the cost of false classification of objects and the cost of missing objects.

It seems that the use of a 2-stage classifier might not be the best method to introduce spatial information in the classification process. Other algorithms that are capable of doing this should be investigate.

The abundant information in the hyperspectral data cubes needs to be studied in more detail to obtain optimal band selection procedures for both automatic classification and human interpretation.

8. REFERENCES

1 C.H. Chen, L.F. Pau, P.S.P. Wrang (Editors), "Handbook of pattern recognition & computer vision", World Scientific, Singapore, 1993

2 DuncanTech 3CCD camera: www.duncantech.com

3 K.S. Fu (Editor), "Digital pattern recognition", Spriner-Verlag, New York, 1976

4 ImSpector User manual 2.0, June 1999, www.specim.fi

5 W. de Jong, H.A. Lensen, Y.H.L. Jansen, "Sophisticated test facility to detect land mines", Proceedings of SPIE Vol. 3710, pp. 1409-1418, 1999

6 R. Schalkoff, "Pattern recognition: statistical, structural, and neural approaches", John Wiley & Sons, New York, 1992