Application of Data-Driven Decomposition to Landsat-TM Images for Crop Area Estimation

Maurice S. klein Gebbinck, Theo E. Schouten

University of Nijmegen, Computing Science Institute
P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
E-mail: mauricek@cs.kun.nl

December 4, 1996

Abstract

An accurate crop area estimation method based on satellite remote sensing imagery is needed to manage the agricultural subsidy system of the European Union. The area estimator can use either classification, which allocates a pixel to a single class, or decomposition, which divides a pixel between several classes, to determine the ground cover type(s) a pixel is composed of. While in early days classification was much used, recently the decomposition approach has gained more interest, however, only on a per pixel basis. In a previous study, we developed the data-driven decomposition method, which used spatial information to guide the decomposition process; on artificial Landsat-TM images this method proved to be far more accurate than techniques based on classification or pixel-based decomposition. To investigate whether data-driven decomposition also results in an improved area estimation when using real satellite images, the area of 17 agricultural lots was determined from a large scale topographical map. After co-registration with a corresponding Landsat-TM image, application of data-driven decomposition gave an estimation that was equally or more accurate than the estimates of a similar method based on classification in 14 of the 17 cases. Furthermore, data-driven classification also showed to be better suited for handling the small boundary structures that separated the agricultural fields. These results suggest that the accuracy of data-driven decomposition is higher than that of an area estimator based on classification when dealing with agricultural fields.

Keywords: remote sensing, Landsat-TM imagery, (crop) area estimation, subpixel resolution, mixed pixels, linear mixture model, classification, decomposition.

Classification: AMS-1991: 68U10, 68U30; CR-1991: I.4.7, I.4.6, I.4.8, I.4.9, J.2.

1 Introduction

For the European Union it is of great importance to be able to make an accurate estimation of the area of each agricultural field. Based on these statistics, the total yield per crop can be predicted, which gives a forecast of the amount of money that has to be spent on agricultural subsidies. Furthermore, by comparing the estimated area with the area supplied on a farmer's subsidy application, possibly fraudulent claims can be identified. With six million farmers occupying 150 million parcels, an automated system using satellite remote sensing images seems to be the only option to perform this task. Because of its data characteristics, the imagery of the Thematic Mapper (TM) scanner aboard the Landsat satellite is well suited for this kind of ground cover mapping applications.

The classical approach to crop area estimation is to allocate each pixel to one of a number of known classes and count the pixels per class afterwards. To reduce the number of misclassifications, it has been proposed to segment the image first so that entire regions can be classified (Janssen [4], Schoenmakers [8]). Although this strategy indeed improves the classification of the central pixels of an agricultural field, the boundary pixels, which usually are a mixture of several ground cover types, are still classified incorrectly. By adopting the linear mixture model of Horwitz *et al.* [3], however, these mixed pixels can be decomposed and partially allocated to several classes. The decomposition method has been applied in several studies (e.g. Cross *et al.* [1], Quarmby *et al.* [6], Settle and Drake [9], Foody and Cox [2]), but always on the basis of individual pixels.

To combine the advantages of a region-based approach and the decomposition technique, we developed the data-driven decomposition method [5], which proved to be far superior to the other methods on simulated satellite images. The purpose of the present study was to determine whether data-driven decomposition would also result in a more accurate area estimation than methods based on classification when dealing with real satellite images. Therefore, an experiment was carried out in which the area estimates of the various methods were compared quantitatively with the true area derived from a topographical map. Apart from that, the methods were assessed for their ability to detect and identify the small boundary structures that separate the agricultural fields because the more correctly these structures can be processed, the more accurately the area of a field can be estimated.

2 Data-driven decomposition

The data-driven decomposition method described in [5] models an agricultural field or any other object that is relatively large given the resolution of the scanner as an inner cluster of pure pixels surrounded by a layer of mixed pixels at the object boundaries. These pure pixel regions can sometimes be found using ancillary information, but if no such information is available, the image has to be segmented. A segmentation method which combines region growing and edge detection was developed by Schoenmakers [8] to segment large optical satellite images. Janssen [4] proposed an integrated segmentation and classification method, but, apart from the satellite image, his method needs data from a geographical information system as well. For data-driven decomposition any segmentation method can be chosen as long as it is able to determine the pure pixel regions reliably.

For optical satellite images such as Landsat-TM images, the pure pixels of an agricultural field are assumed to have a multivariate normal distribution. According to the linear mixture model introduced by Horwitz *et al.* [3], each mixed pixel $\mathbf{x}$ can now be described as

$$\mathbf{x} = \mathbf{M}\mathbf{f} + \mathbf{e}, \tag{1}$$

where vector $\mathbf{f}$ denotes the fractions of the contributing fields, and matrix $\mathbf{M}$ lists their mean reflectance. The variation in the reflectance of the fields is covered by the error vector $\mathbf{e}$, which has a multivariate normal distribution with a zero mean and a variance-covariance matrix that depends on the distributions and fractions of the objects. If this dependence on the fractions, which are to be estimated, is abandoned and the average of all variance-covariance matrices is taken, the maximum-likelihood solution of $\mathbf{e} = \mathbf{x} - \mathbf{M}\mathbf{f}$ gives an estimation of $\mathbf{f}$ that can be implemented as a series of fast matrix multiplications.

There are two aspects of our decomposition method that justify the predicate data-driven:

1. Spatial information is used to determine the most probable components of a mixed pixel. By searching the neighbourhood of a mixed pixel for pure pixel regions, the fields that probably contribute to the reflectance of the mixed pixel can be identified. The advantages of this approach are twofold: the accuracy of the estimation is higher because no fractions can be assigned to regions that make no contribution, and the image can be analysed using an unlimited number of classes because every mixed pixel can be decomposed using a different set of classes.[1] The problem that the neighbourhood of a mixed pixel usually contains

[1] If the number of components exceeds the number of spectral bands by more than one, the linear mixture model can no longer be solved uniquely.

other mixed pixels, which hampers the identification of the pure pixel regions, was solved by making the algorithm iterative; in the first iteration the mixed pixels whose fields can be found directly are decomposed, while in the next iterations the composition of the pixels processed before is used to decompose the more difficult mixed pixels.

2. The distributions of the agricultural fields are determined from the image itself. Instead of extracting the distribution of a field from a class database, the mean vector and variance-covariance matrix are calculated from its pure pixel region. This way, differences between the distributions of fields of the same class, which are due to local variations in ground slope, elevation, soil type, etc, can be accounted for. Fields whose pure pixel region is too small to calculate its distribution reliably are classified and assigned the corresponding distribution from the class database after all.

Agricultural fields are often separated by a strip of grass, a dirt road, a ditch, etc. These small boundary structures are usually made up of mixed pixels without a pure pixel region, so they cannot be identified nor can their distributions be calculated. Therefore, a set of edge classes is established whose members are always considered as possible components of the mixed pixel and whose distributions are taken from the class database. To facilitate a quantitative evaluation of the results, for each edge class a so-called global region was constructed.

Some pixels which are potentially mixed are in fact isolated pixels that do not belong to any nearby region. Such a pixel, which can be tracked because the likelihood of its decomposition between neighbouring regions is very low, is decomposed into a class not found in the neighbourhood but taken from the class database and one neighbouring region. To determine which combination of classes the isolated pixel is truly composed of, all combinations are tried and the one with the highest probability is selected. The same mechanism is used to limit the number of components a mixed pixel is decomposed into: the combination of one or two neighbouring regions and zero or one edge class(es) that has the highest probability is chosen. When dealing with Landsat-TM images, more components could be tried simultaneously, but since in most cases not all of them would belong to the true composition of the pixel, this would only decrease the accuracy of the estimation.

3 Set-up of the experiment

The study site that was selected is located near the village of Zeewolde in the south of the Flevoland polder. The land, which has been reclaimed from the IJsselmeer, is flat with an altitude of 3.30 m below sea level and is mainly used for the cultivation of arable crops. A network of roads, canals, and ditches divides the area into lots, which are basic units of approximately 500x1600 m^2 that can be rented or bought. Many farmers subdivide their lots into several agricultural fields, which are characterised by the cultivation of a single crop. The locations of the field boundaries are determined by the farmers themselves and are subject to change every year. In his study, Janssen [4] calculated that the average field size of the nearby Biddinghuizen test site was 6.9 ha. This relatively large size makes the area most suitable for data-driven decomposition since it permits an accurate calculation of the statistical distribution of the average field.

For a quantitative analysis of the different area estimation methods, supervised topographical data which has an accuracy that goes beyond the resolution of the scanner was needed. For this purpose, the boundaries of 17 lots were digitised from a 1:10,000 topographical map—see Figure 1. Assuming a digitising error of ± 0.5 mm, the boundaries had a positional accuracy of ± 5 m, which was several times higher than the 30 m resolution of Landsat-TM images. Co-registration of the topographical data with a Landsat-TM image of the same region was accomplished by an affine transformation, which translates, rotates, and scales the vector data. The alternative was to geocode the satellite image in the map's coordinate system, but this method would deteriorate the image quality because it requires the image to be resampled. The parameters of the affine transformation were calculated from four ground control points (GCP's) that were positioned in one of the corners of the digitised area each. The root mean square (RMS) error of this transformation was only 0.283 pixel, which was relatively low compared to the values (0.4 and 0.9) that were achieved by Janssen. Calculation of the scale factor gave a value of 29.99, which closely corresponded to the 30 m resolution of the Landsat-TM image. Based on these figures as well as on careful visual inspection of the lot boundaries superimposed on the satellite image—see Figure 2—, it was assumed that co-registration with an error in accuracy of less than one pixel was achieved.

To evaluate the accuracy of the area estimation methods, several options were considered:

I. Compare the estimated and true proportions per pixel. This approach was preferred because it provides the most detailed information about the performance of the area estimator. However, both the resolution of the topographical data and the accuracy of the co-registration were too low to determine the true composition of each pixel. A better co-registration may be achieved by using many more GCP's or by applying a different registration algorithm; more accurate topographical data may be acquired from the Land Registry in Apeldoorn (the Netherlands), where the "Grootschalige Basiskaart van Nederland", which has a resolution of ± 0.5 m, is maintained. However, this data is probably only commercially available, may not be digitised for the Zeewolde area, and may still be not detailed enough. Furthermore, it remains questionable whether the results will be meaningful because the fact that the vegetation often gets more sparse towards the field boundaries, where the mixed pixels are situated, is not reflected by the topographical data.

II. Compare the estimated and true area per field. An accurate estimation of the area per field is one of the goals of the EU because it is needed when checking a farmer's subsidy claim. Unfortunately, no topographical map includes field boundaries because they change too frequently. Janssen [4] derived the crop types and field boundaries of the Biddinghuizen test site from maps that were provided by the farmers themselves, but he found that the field boundaries could have been added just as well by means of on-screen digitising. Since the resulting positional accuracy is rather poor, a quantitative analysis on this basis would not be very accurate either.

III. Compare the estimated and true area per crop. Since an estimation of the area per crop is needed to make a prediction of the agricultural production, which is used to determine the subsidy for each crop, this is another goal of the EU. Because the topographical map was our only source of supervised information, the true area per crop had to be derived by summing the areas of all fields on which the crop was being grown. But since the field boundaries could not be located very accurately—see Alternative II—, this approach would not give very reliable results.

IV. Calculate the reliability of classification and decomposition per pixel. In this last approach it is only checked whether the area estimation method can process the mixed pixels reliably by comparing the distribution of the reliability measure with the distribution obtained from experiments with artificial satellite images carried out in a previous study [5]. Although this approach can always be used because no topographical information is needed, it also has some severe drawbacks. One of them is that classification and decomposition use different reliability measures, so their results cannot be compared. Another drawback is that the conclusions are at best indirectly related to the area estimation capabilities of the methods applied.

Although none of the four alternatives seemed to be particularly suitable, a slightly altered Approach II was adopted as the basis for our experiment. Instead of comparing the estimated and true area per field, a comparison was made between the estimated and true area per lot. For this purpose the accuracy of the topographical data was high enough: the true area of a lot could be calculated with an error of only $\pm 2.6\%$.[2] Because the area estimation methods calculated the area

[2] Given an average lot size of 500x1600 m^2 and a digitising error of ± 5 m, a rectangular lot is underestimated by at most $1 - (490\text{x}1590)/(500\text{x}1600) = 2.6\%$ and overestimated by at most $1 - (510\text{x}1610)/(500\text{x}1600) = 2.6\%$.

of fields rather than lots, some extra steps had to be taken to make a comparison possible. As a first step, a vector-to-raster conversion of the topographical data was carried out in order to find out which pixels belonged to what lot; the resulting image was used to make a mapping between the agricultural fields provided by the area estimators and the lots of the topographical data. Unfortunately, it was not possible to calculate the area of a lot merely by summing the areas of its fields because part of the area of the field boundaries intersecting the lot may be assigned to the edge classes. For example, a strip of grass separating two fields within a lot can be the cause of an area leakage to the global region of the edge class grassland, thus leading to an underestimation in case the strip of grass is uncharted. Our solution to this problem was to count the number of the pixels that were located completely inside the lot under observation and to add the pixel fractions that were assigned to one of the fields constituting the lot for each pixel that according to the vector-to-raster conversion was located on the lot's boundary. After a simple multiplication—a Landsat-TM pixel measures exactly 30x30 m^2—the area estimate could be compared directly to the supervised lot area that was derived from the topographical data.

Given the set-up of the experiment, it was decided to apply solely data-driven classification and data-driven decomposition because their model-driven counterparts were more suited for area estimation per class than for area estimation per lot. Moreover, the data-driven methods also resulted in a better performance when dealing with simulated images [5]. However, these area estimators required the image to be segmented into regions consisting of pure pixels before they could be applied. Rather than using a computer derived segmentation which possibly contained all sorts of errors and inaccuracies, the set of digitised boundaries was extended with the field boundaries that were digitised on-screen. After a vector-to-raster conversion of the entire set using a sufficient margin to account for the co-registration error and positional inaccuracy of the field boundaries, a supervised segmentation of high quality was available. It was calculated that 19.4% (15.6%) of the digitised area was occupied by pixels on field (lot) boundaries, which indicates that, even when surveying an area with relatively large fields like the Flevoland polder, the accuracy of the area estimation may be improved considerably by applying techniques specialised in mixed pixels.

The class database needed by the area estimation methods was dependent on the image that is to be processed not only because it must contain all ground cover types present in the image but also because the class distributions were biased by the sunlight intensity, atmospheric effects, sensor calibration, etc. The Landsat-TM image of the Zeewolde study site that was used was acquired on 24th May 1990; the image was cloudfree and displayed fewer scanner anomalies than more recent images, but it was taken rather early in the growing season. Because of this early acquisition date, many fields lay fallow while the remaining fields were covered with cole-seed, alfalfa, winter wheat, and grass. In addition to these ground cover types, the distributions of water, forest, and asphalt road were entered into the database because they were considered to be the edge classes together with grass and bare soil. For several classes, multiple distributions were entered in order to cover their spectral variety: shallow, deep, and canal water; coniferous and three types of deciduous forest ranging from green to brown; and three types of grass ranging from green to yellow. The distribution of each class was determined by a remote sensing expert by selecting a sufficiently large area completely covered by that class and calculating the mean vector and variance-covariance matrix of the corresponding pixels. Only for the classes canal water and asphalt road no such regions could be found; the distribution of these classes was determined by collecting a number of non-contiguous pure pixels of the same class and calculating the set's statistics.

To simplify a quantitative comparison of the results of the area estimators, it was desirable to have a single measure based on the area estimates of the individual lots. Furthermore, the measure was also needed to select the optimal configuration from a series of trial runs of the experiment in which the threshold setting and the set of edge classes were varied. A measure that met these requirements was the areal error measure introduced by Klein Gebbinck and Schouten [5]:

$$e_A = \frac{1}{2}\sum_i \left|\hat{A}_i - A_i\right|. \tag{2}$$

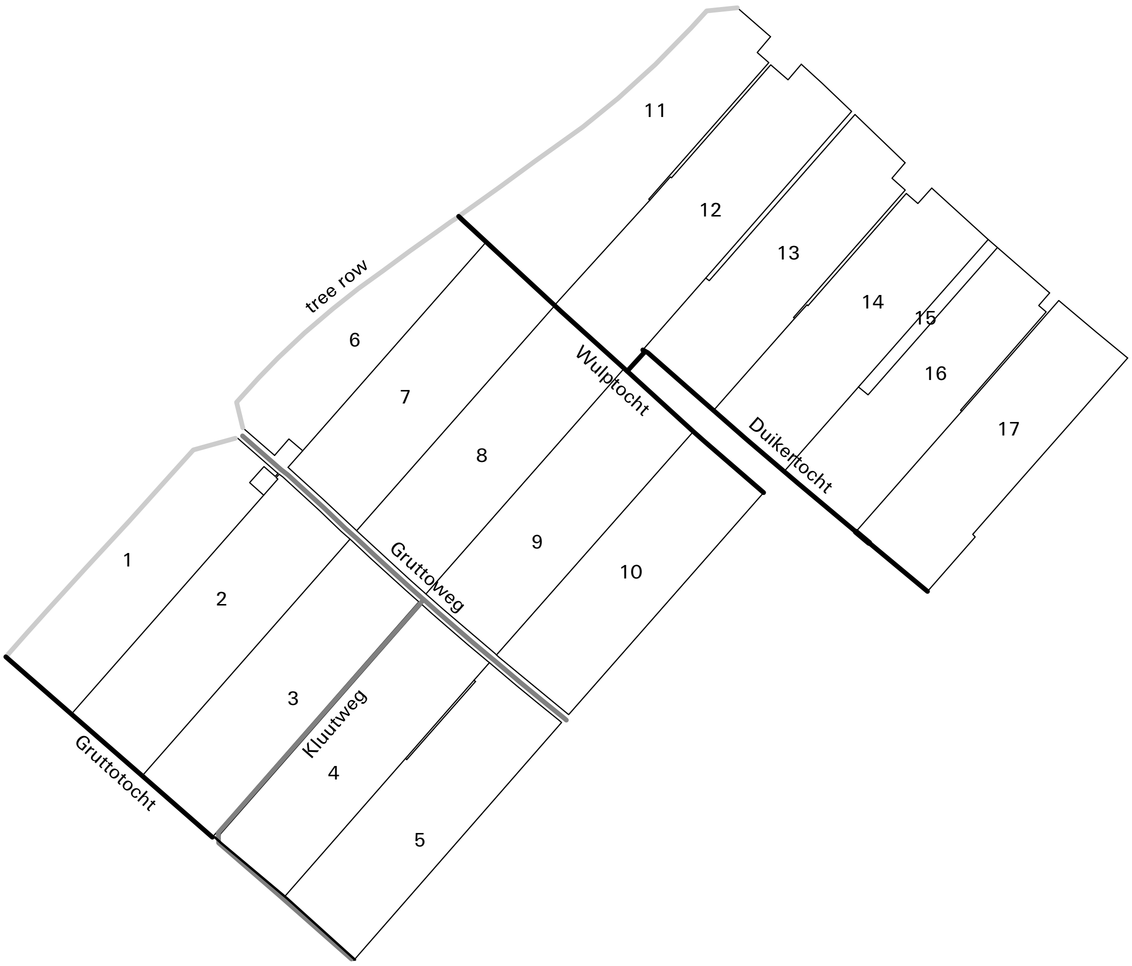

Figure 1: Digitised boundaries of the 17 lots used in the experiment. The lots are separated from each other by ditches or major boundary structures such as tree rows (light grey), asphalt roads (dark grey), and canals (black).

However, since in the experiment described in the present report the area per lot instead of the area per crop was estimated, A_i ($\hat{A}_i$) was chosen to represent the true (estimated) area of lot i instead of the true (estimated) area covered by class i. The reason for introducing the scalar $\frac{1}{2}$—underestimation of one crop automatically results in overestimation of another crop—was no longer valid: because the supervised information about the image was incomplete, the underestimation of lot i did no longer necessarily result in the overestimation of lot j but could also lead to the overestimation of some unknown boundary structure belonging to the background. However, since the scalar had no effect on the order of best area estimators, we saw no reason to remove it.

4 Results

The presentation of the results is split up into three parts. Section 4.1 lists the area estimates of the individual lots and compares them with their supervised area derived from the topographical data. In Section 4.2 it is analysed how well the area estimators could be used to process the boundary structures that separated the lots. Section 4.3, finally, describes the influence that spectral confusion had on the outcome of this experiment with regards to the area estimation accuracy and the ability to detect and identify the boundary structures.

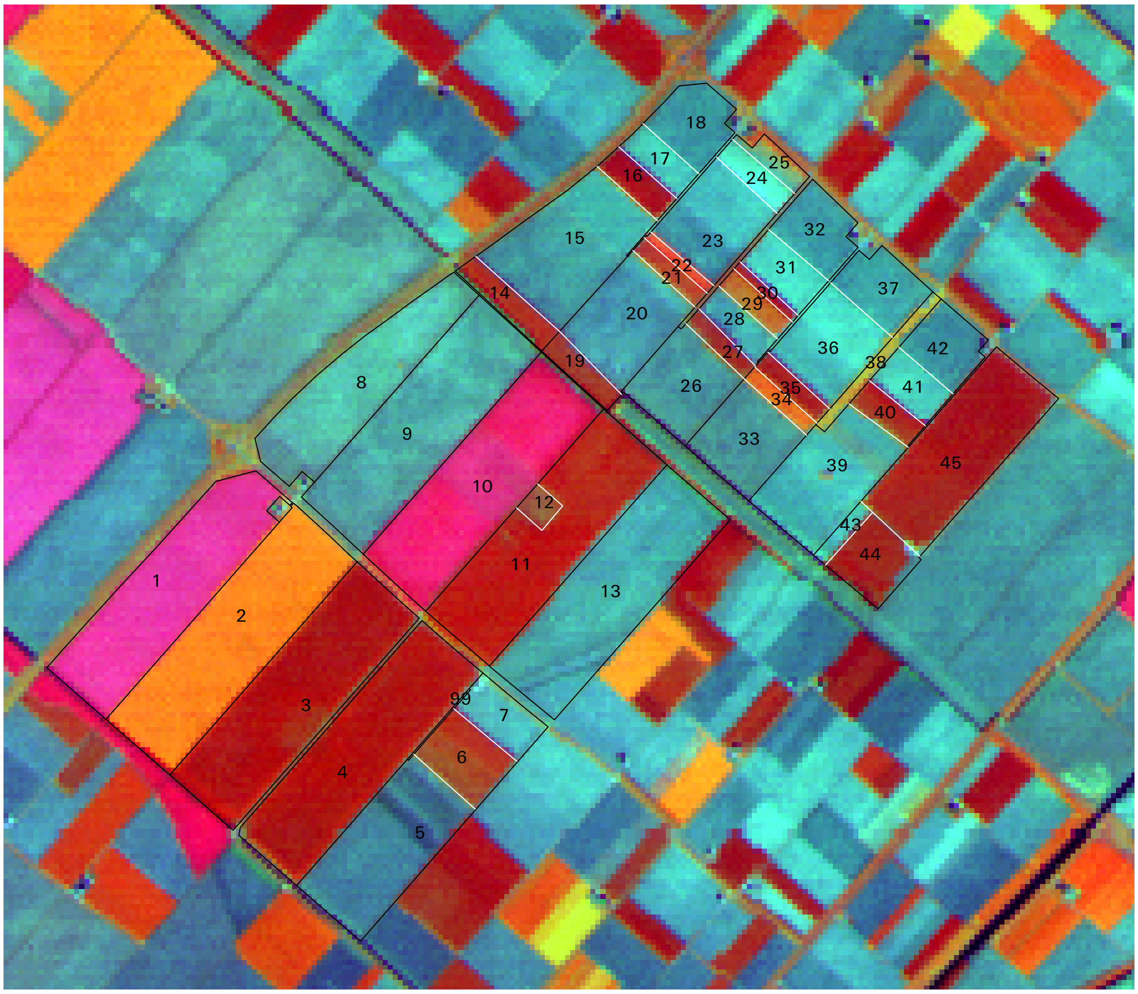

Figure 2: Field geometry superimposed on a false colour composite (RGB: 453) of the Landsat-TM image. The co-registered lot boundaries are given in black, while the added field boundaries are drawn in white.

lot	main class	area	decomposition		classification incl.		classification excl.	
			area	deviation	area	deviation	area	deviation
1	cole-seed	825.621	840.688	15.067	830	4.379	830	4.379
2	alfalfa	945.042	953.606	8.564	904	-41.042	904	-41.042
3	winter wheat	928.779	933.389	4.610	954	25.221	954	25.221
4	winter wheat	888.710	886.814	-1.896	887	-1.710	887	-1.710
5	bare soil	932.792	932.651	-0.141	904	-28.792	930	-2.792
6	bare soil	617.819	624.199	6.380	590	-27.819	631	13.181
7	bare soil	886.106	866.815	-19.291	810	-76.106	845	-41.106
8	cole-seed	890.743	862.565	-28.178	895	4.257	895	4.257
9	winter wheat	886.415	898.676	12.261	894	7.585	894	7.585
10	bare soil	879.785	883.969	4.184	864	-15.785	914	34.215
11	bare soil	1055.901	1055.531	-0.370	1013	-42.901	1037	-18.901
12	bare soil	1004.607	1009.773	5.166	969	-35.607	999	-5.607
13	bare soil	880.806	879.161	-1.645	840	-40.806	871	-9.806
14	bare soil	901.159	887.369	-13.790	847	-54.159	877	-24.159
15	grassland 3	76.922	75.065	-1.857	84	7.078	94	17.078
16	bare soil	826.095	835.939	9.844	795	-31.095	814	-12.095
17	winter wheat	883.555	883.794	0.239	882	-1.555	897	13.445
e_A				66.742		222.948		138.289

Table 1: Comparison of the lot area derived from the topographical data with their estimates by the decomposition and classification methods; all areas are given in pixels. The ground cover type that covered the largest part of the lot is given as its main class. Data-driven classification was applied with the class bare soil included in as well as excluded from the set of edge classes.

4.1 Estimation of the lot area

In Table 1, the areas of the 17 lots derived from the digitised vector data are given, together with their estimates by data-driven decomposition and data-driven classification. The set of edge classes that was used for these experiments consisted of all ground cover types described in Section 3 except for the classes cole-seed, alfalfa, and winter wheat; leaving out one or more classes resulted in a higher areal error. The optimal threshold setting for this set was 54.0 for decomposition whereas classification gave the best results with the threshold set at 150.0.[3] The areal error of data-driven decomposition was 66.742, which was less than one third of the 222.948 of data-driven classification. Looking at the results on the individual lots, it can be observed that in 13 cases decomposition led to a more accurate area estimation while for one lot the estimates of decomposition and classification were approximately equal (± 1 pixel). In only three cases did decomposition result in a less accurate estimation, but for each of these cases a suitable explanation was found.

The area of the first such lot—lot 1—was overestimated by classification but even more by decomposition. The reason for the increased overestimation was that the pixels at the south-west boundary were a mix of canal water, the ground cover type of lot 1, and the class of the undigitised field at the south-west. Since both lot 1 and the undigitised field were covered with cole-seed, decomposition erroneously assigned part of the area of the undigitised field to lot 1. Classification, on the other hand, assigned only a few of the boundary pixels to lot 1 since, due to spectral confusion, most pixels resembled winter wheat much more than cole-seed. As a result, overestimation was reduced considerably.

The area of lot 8 was slightly overestimated by classification but severely underestimated by decomposition. The reason for the underestimation was that the topographical data did not

[3] A threshold of 54.0 (150.0) corresponds to an average deviation of 3σ (5σ) per spectral band in case the bands are uncorrelated; see [5].

fully agree with the actual ground truth. As can be seen in Figure 2, lot 8 was not covered by cole-seed only but contained an uncharted small strip of bare soil near the north-east boundary. Consequently, every accurate estimation of the area covered by cole-seed was bound to be an underestimation of the area of the entire lot as derived from the topographical data.

The last lot for which the areal error was higher for decomposition than for classification was lot 9. Both methods overestimated the lot's area, but the estimation by the decomposition method deviated about five extra pixels. Paradoxically, the higher overestimation was caused by the higher accuracy of data-driven decomposition. North-east of the boundary of the lot a small strip of pixels enclosing a canal, an earth bank covered with grass, and some marsh land was present—see Figure 2. Due to spectral confusion, this particular mix of ground cover types closely resembled winter wheat[4], which was the class of lot 9. As a result, the area of this lot would always be overestimated. However, the pixels at the lot boundary itself, which consisted of winter wheat and canal water, were completely classified as asphalt, thus reducing the overestimation. Because data-driven decomposition processed these pixels more accurately and assigned part of their area to lot 9, a higher error was achieved.

In Table 1 it can be seen that the area of all lots which were largely covered by bare soil was severely underestimated by data-driven classification. The reason for this was that bare soil was assumed to be an edge class, and consequently many pure soil pixels were not assigned to the lot to which they belonged but to this edge class. Decomposition, on the other hand, could deal with the subtle differences between these two statistical distributions much better and did not seem to suffer from this dilemma at all. To verify this hypothesis, the same experiment was carried out after removing bare soil from the set of edge classes. The results of the classification approach, which are presented in the last two columns of Table 1, showed that the area estimates of the lots covered by bare soil were improved considerably. The areal error e_A fell from 222.948 to 138.289, and was only twice as high as the areal error of data-driven decomposition (66.742). However, qualitatively the results hardly changed: decomposition was still more accurate than classification in 12 cases, approximately equally accurate (± 1 pixel) in two cases, and less accurate—for the same reasons given in the previous paragraph—in three cases. The area of some lots which were underestimated at first now became somewhat overestimated, which might indicate that bare soil was an edge class after all. This supposition was supported by the fact that removal of bare soil from the set of edge classes also had a considerable negative influence on the accuracy of the area estimation by data-driven decomposition.

4.2 Processing of the boundary structures

Dual to finding where an agricultural field ends, which is necessary for area estimation, is finding where a boundary structure begins. The latter task, however, is much more difficult because the width of many structures usually is very small (< 1 pixel) and their composition can be rather complex, e.g. a small asphalt road which is flanked by ditches with grassland verges in between. Because of their size and complexity, these structures were not mapped in great detail on our topographical map. As a result, it was impossible to acquire supervised topographical data that was accurate enough for a quantitative analysis without doing a lot of labor-intensive field-work. Even if this data was acquired, evaluation of the results would remain problematic because of the limited accuracy of the co-registration. However, it was possible to address some qualitative issues regarding the processing of these structures, like, 'Was the boundary structure detected at all instead of being divided between the neighbouring fields' or, 'Could the main component (or even all components) of the structure be identified'. To answer these questions, the performance of the area estimation methods on three types of boundary structures was examined: a tree row, asphalt roads (Kluutweg, Gruttoweg), and canals (Gruttotocht, Wulptocht, Duikertocht)—see Figure 1.

Data-driven classification was able to find most boundary structures, but in some cases a considerable number of gaps was present. The small ditches between the lots, however, often

[4]The Mahalanobis distance between winter wheat and the mix of canal water (31%) and grassland 2 (69%) was only 0.067; see Section 4.3 for more details.

remained undetected, as was the part of the Gruttotocht along lots 1–3. Furthermore, classification found several structures at field boundaries inside a lot. Although no permanent structure such as a asphalt road or a ditch was present, it may be that a temporary structure such as a dirt road used for farming purposes was detected. Due to spectral confusion—see Section 4.3—, nearly all boundary pixels were assigned to the class asphalt road. The only exceptions were the pixels belonging to the tree row along lot 1 and most of the pixels of the Gruttoweg, which were assigned to the class grassland. Although it was questionable whether grassland truly was their main component, the 1:25,000 topographical map[5] showed that grassland verges indeed were part of these structures. Of the other ground cover types that were assumed to belong to the set of edge classes, only deciduous forest was assigned a small number of pixels, none of which were located at the tree row. Not a single pixel was assigned to the class water although water was a component of the numerous ditches and canals.

Detection of the boundary structures by data-driven decomposition turned out to be no big problem. Since all structures caused the pixels at the field boundaries to deviate from the mean values of the neighbouring fields, a fraction of their area was quickly assigned to one of the edge classes. Therefore, it was more interesting to verify whether the edge class that was chosen was indeed one of the components of the boundary structure. Analogue to classification, decomposition assigned fractions of a large number of pixels to the class asphalt road. Many of these pixels were not located on one of the asphalt roads but on one of the other boundary structures or field boundaries inside a lot. The cause of this behaviour was again spectral confusion, which will be discussed in Section 4.3. A closer look at the results of decomposition, however, showed that the selected boundary structures could be identified much better than before.

The pixels of the tree row along lots 6 and 11 were decomposed into deciduous forest and grassland, which were precisely their components according to the topographical map. The tree row along lot 1, however, was decomposed into grassland and water. The reason that these parts of the same tree row were decomposed differently was that the main class of lots 6 and 11 was bare soil whereas the class of lot 1 was cole-seed. Apparently, the latter mix of ground cover types resulted in a partially incorrect decomposition because of spectral confusion.

In Figure 1 it can be seen that the Kluutweg consisted of two parts. Decomposition of the pixels along lots 4–5 revealed the presence of the asphalt road, but the second part, which lay between lots 3 and 4, was thought to be composed of bare soil. The same error could be observed for several of the short road fragments between lots 11–17. The reason for this error as well as for most other cases of spectral confusion was the spectral similarity of the classes asphalt road and bare soil. The pixels belonging to the Gruttoweg structure could be processed without these problems. For nearly all these pixels, both the asphalt road as well as the grassland verge were identified.

Since not a single pixel of water was found by classification, it was expected that the decomposition method would have difficulties extracting the component water as well. This reservation proved to be justified partially since only a few of the ditches that separated the lots were found. When decomposing the pixels belonging to the canals, however, considerable fractions of water were identified. Nearly all pixels of the Duikertocht for instance were found to contain a large fraction of water. Of the Gruttotocht the part along lots 1–2 was identified as water; the fragment along lot 3 was thought to contain a much smaller proportion of water, while the part parallel to the Kluutweg was largely assigned to the class asphalt road. The Wulptocht, however, proved to be the most difficult boundary structure to decompose. Pixels consisting of a fair amount of water were found only along lot 9. Because of spectral confusion, deciduous forest was the main component found for the parts along lots 6–7 and 10, whereas the pixels along lot 8 were thought to contain large fractions of winter wheat, which was the class of the neighbouring field 19.

[5] Topographical map of The Netherlands, sheet 26D, 1991.

4.3 Spectral confusion

In the present report, the term spectral confusion is used to describe the resemblance of a class (mixture) to a different class (mixture). If a pixel is modelled to represent of a single class, spectral confusion is caused by overlapping class distributions that are due to the natural variation within a class. If a mixed pixel model is assumed, spectral confusion is possible if a pixel can be decomposed into a number of different combinations of classes. The possibility of spectral confusion depends entirely on the location of the class means in the feature space and their within-class variation. The occurrence of spectral confusion, however, depends on the scene that is processed: some combinations of classes never occur and therefore cannot give rise to spectral confusion.

Classification, which assumes a pure pixel model, displayed only one clear case of spectral confusion. According to a remote sensing expert, the class of field 22 was grassland, but the classification algorithm assigned it the class alfalfa. However, alfalfa was chosen over grassland 2 with only a minor preference: the Mahalanobis distance of the field's mean to alfalfa was 72.317, compared to 76.108 for grassland 2. The reason for this misclassification was that the number of pixels belonging to field 22 was only six, which apparently was too low for an accurate classification, given the within-class variation of grassland 2.

The results of data-driven classification showed that there was another case of spectral confusion though that explanation could be argued. Subject of this dispute were the mixed pixels that occurred at the field boundaries. As was described in the previous section, nearly all boundary structures were classified as asphalt road although many of them in fact were canals, ditches, etc. Because classification assumed all pixels to be pure, each mixed pixel could be regarded as a pure pixel of the class of its main component and the deviation from the class mean as caused by within-class variation. According to this view, the classification of a canal pixel as asphalt road was a case of spectral confusion. However, it could also be argued that a pixel that consisted of more than one ground cover type simply did not fit into the pure pixel model, as a result of which spectral confusion did not apply. Anyway, data-driven decomposition displayed similar problems with the identification of the boundary structures; therefore, a more in-depth analysis was necessary.

In Sections 4.1 and 4.2, several processing errors by data-driven decomposition were mentioned which were imputed to spectral confusion. The most frequently occurring error by far was the discovery of a large asphalt road component in pixels that belonged to a different type of boundary structure or were located on field boundaries within a lot. Especially ditches that were located next to lots of bare soil as well as field boundaries of bare soil and winter wheat were largely assigned to the class asphalt road. In Table 2 it can be seen that the spectral signatures of the classes bare soil and asphalt road were remarkably similar: for all bands the mean value of asphalt road was a bit lower except for band 4, which had a slightly higher value. Since a similar pattern could be observed for all other classes, it was assumed that a pixel consisting of bare soil and a small fraction of another class resembled asphalt road to a large degree. To verify this assumption, the class mean of asphalt road was decomposed into all combinations of two and three other classes. Strictly spoken, this experiment should have been carried out the other way around—decompose a previously generated set of samples obtained by combining a number of classes while varying their proportions—but with 15 classes this task would be enormous. Since it was impossible to quantitatively verify the validity of the assumed mechanism anyway because no supervised information about the composition of the incorrectly decomposed pixels was available, an experiment that would supply a plausible explanation for the suspected cases of spectral confusion was all that was looked for.

Decomposition of the mean vector of the class asphalt road in two other components resulted in a Mahalanobis distance of less than 10.0 if and only if one of the components was bare soil. The threshold of 10.0 was chosen after observing that decomposition of other class means in two components hardly ever resulted in a Mahalanobis distance of less than 10.0. A few of the combinations that resembled asphalt road closely were 71.5% bare soil and 28.5% canal water (distance: 2.634), and 80.9% bare soil and 19.1% winter wheat (distance: 4.602). Decomposition of asphalt road in three other components also gave a Mahalanobis distance of less than 10.0

class	band 1	band 2	band 3	band 4	band 5	band 7
shallow water	83.5	36.4	31.3	18.2	7.4	4.1
deep water	79.9	33.8	28.3	17.1	7.0	3.9
canal water	81.6	31.6	28.1	33.5	23.3	9.3
coniferous forest	75.4	28.5	25.2	50.3	28.3	9.9
deciduous forest 1	78.5	32.0	28.9	52.3	60.3	26.5
deciduous forest 2	77.3	31.2	26.0	74.3	41.3	12.4
deciduous forest 3	77.8	32.2	27.3	77.8	52.4	17.5
grassland 1	85.4	36.3	32.1	103.8	92.1	34.0
grassland 2	80.6	32.1	27.1	115.4	59.8	18.0
grassland 3	82.7	35.0	31.2	91.2	74.8	27.8
asphalt road	92.2	40.0	42.4	56.2	60.8	32.2
bare soil	96.8	42.2	46.8	55.8	75.3	46.3
cole-seed	86.0	45.7	46.3	111.5	54.3	19.4
alfalfa	82.5	34.9	29.1	119.9	68.9	20.0
winter wheat	80.7	31.6	27.1	90.3	47.6	14.5

Table 2: Class mean vector for all ground cover types considered. The spectral signatures of asphalt road and bare soil are very similar.

only if one of the components was bare soil. Out of the 78 possible 3-tuples containing bare soil, 71 combinations resulted in a distance below the threshold. The three combinations to which asphalt road had the lowest distance were: 58.0% bare soil, 28.2% canal water, 13.8% cole-seed (distance: 0.332); 65.2% bare soil, 21.7% shallow water, 13.1% alfalfa (distance: 0.707); and 67.2% bare soil, 15.4% shallow water, 17.4% winter wheat (distance: 0.946). These results made it likely that both ditches adjoining lots of bare soil as well as boundary pixels between fields of winter wheat and bare soil—the two main cases of spectral confusion observed previously—were largely assigned to the class asphalt road.

5 Conclusions and future work

The purpose of the study was to determine whether the area of an agricultural field could be estimated more accurately by a method based on decomposition or by a method based on classification. Therefore, the area estimation accuracy of data-driven decomposition and data-driven classification was assessed using a real Landsat-TM image of the Flevoland polder. To make a verification of the results possible, a small scale topographical map was digitised and co-registered to the satellite image with subpixel accuracy. From this supervised data, the true area of each lot could be calculated for use in a quantitative comparison of the accuracy of the area estimators. Furthermore, some qualitative information about the composition of the boundary structures was derived which was used to determine how good decomposition and classification were at detecting boundary structures and identifying their components.

For 14 of the 17 lots, data-driven decomposition provided an area estimate that was approximately as accurate to far more accurate than the estimate of data-driven classification. In the remaining three cases plausible explanations for the seemingly lower accuracy of the decomposition approach were found. A problem for the classification method was the class bare soil: if bare soil was considered to be an edge class, the area of bare soil lots was underestimated, but if bare soil was not included in the set of edge classes, the area of several lots was overestimated. Since decomposition performed well with bare soil being an edge class and gave more accurate area estimates for most lots, data-driven decomposition was considered to be a better area estimator for agricultural fields than data-driven classification. Whether the same conclusion can be drawn when dealing with much smaller agricultural fields is uncertain because the minimum area needed for a reliable estimation of a field's statistical distribution was not determined.

The three types of major boundary structures, which were tree rows, asphalt roads, and canals, were all largely detected by both decomposition and classification; the small ditches between the lots, however, often remained undetected. The identification of the main component(s) of a structure was more problematic, especially for data-driven classification. Although the results of data-driven decomposition were much better, the performance was still rather poor. The reason for these problems was that the resemblance of bare soil and asphalt road caused spectral confusion; as a result, boundary structures composed of water were hard to identify and structures containing bare soil were often assigned to the class asphalt road. The impact of spectral confusion was much harder on classification than on decomposition because the boundary structures were composed of mixed pixels mostly and they did not fit into the pure pixel model used by classification. Although neither decomposition nor classification was found to be suitable for dealing with boundary structures, their results may still be of use as an extra source of information for specialised boundary extraction applications like road network extraction.

The most obvious direction for future work is to repeat the experiment under different conditions. To determine the applicability of data-driven decomposition for other geographical regions, the influence of a smaller average lot size should be investigated. Another factor that is of importance is the ruggedness of the terrain because it may invalidate the linear mixture model. Furthermore, the image used should be acquired much later in the growing season in order to have a greater variety in ground cover types. As a consequence, far less fields will lie fallow, thus eliminating the problem of edge class membership of bare soil which made the operation of data-driven classification difficult.

Future work can also be directed towards improving the detection and identification of boundary structures because the better these structures can be processed, the more accurate the area estimation of the neighbouring fields will be. The results may be improved by combining the decomposition method with specialised road network extraction techniques like those used by Ruskoné *et al.* [7]. Since most boundary structures in Flevoland are linear and usually have a high contrast with their surroundings, application of these techniques may be successful even though satellite images instead of high resolution aerial images are used. Another improvement can be the introduction of mixed classes to reduce spectral confusion. If, for example, a mixture of 75% bare soil and 25% water is frequently confused with asphalt road, constructing such a class may well result in an improved identification. However, care must be taken that the additional classes do not introduce new cases of spectral confusion and that the number of mixed classes is kept small to ensure that the time needed for processing remains acceptable.

6 Acknowledgments

I would like to thank Gerard Nieuwenhuis, head of the Remote Sensing Department of the DLO Winand Staring Centre for Integrated Land, Soil and Water Research, for the opportunity to carry out these experiments on the department's data and Edwin Noordman and Henk Kramer of the same department for their help with the acquisition of the data set.

References

[1] A.M. Cross, J.J. Settle, N.A. Drake, and R.T.M. Paivinen. Subpixel measurement of tropical forest cover using AVHRR data. *International Journal of Remote Sensing*, 12(5):1119–1129, 1991.

[2] G.M. Foody and D.P. Cox. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing*, 15(3):619–631, 1994.

[3] H.M. Horwitz, R.F. Nalepka, P.D. Hyde, and J.P. Morgenstern. Estimating the proportions of objects within a single resolution element of a multispectral scanner. In *Proceedings of the 7th*

International Symposium on Remote Sensing of Environment, pages 1307–1320, 1971. Ann Arbor, Michigan.

[4] L.L.F. Janssen. *Methodology for updating terrain object data from remote sensing data.* PhD thesis, Wageningen Agricultural University, The Netherlands, 1994.

[5] M.S. klein Gebbinck and T.E. Schouten. Accurate area estimation by data-driven decomposition of mixed pixels. Technical Report CSI-R9622, Computing Science Institute, University of Nijmegen, 1996.

[6] N.A. Quarmby, J.R.G. Townshend, J.J. Settle, K.H. White, M. Milnes, T.L. Hindle, and N. Silleos. Linear mixture modelling applied to AVHRR data for crop area estimation. *International Journal of Remote Sensing*, 13(3):415–425, 1992.

[7] R. Ruskoné, S. Airault, and O. Jamet. Road network extraction by local context interpretation. In J. Desachy, editor, *Image and Signal Processing for Remote Sensing II*, pages 126–136. SPIE 2579, 1995. Paris.

[8] R.P.H.M. Schoenmakers. *Integrated methodology for segmentation of large optical satellite images in land applications of remote sensing.* PhD thesis, Department of Informatics, University of Nijmegen, The Netherlands, 1995.

[9] J.J. Settle and N.A. Drake. Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, 14(6):1159–1177, 1993.