GEOGRAFICKÝ ČASOPIS

51 1999 2

*Dick van der Zee**

ASSESSING THE ACCURACY OF AIRPHOTO INTERPRETATION: THE EXAMPLE OF FARMS, HOUSES AND RURAL ROADS

D. van der Zee: Assessing the accuracy of airphoto interpretation: the example of farms, houses and rural roads. Geografický časopis, *51*, 1999, 2, 7 figs., 4 tabs., 7 refs.

The assessment of the accuracy of airphoto interpretation is not a simple matter. This will be illustrated by a number of case studies dealing with the interpretation of farms, houses and rural roads. Right or wrong can be a relative question. In addition it can be viewed from two different points: number of interpreted objects found to be correct in reality or number of objects in reality that have been correctly interpreted. A number of different aspects is related to accuracy, of which reliability of the identification appears to be the most essential one. The influence of the skill and experience of the interpreter and the scale of the airphotos used on the accuracy of the interpretation can be great. If not point objects but line or area objects are concerned, the expression of accuracy in number of objects correctly interpreted may not be satisfactory. The percentage of correctly interpreted area or length may be more relevant. It is also important to analyse and quantify the type of misinterpretations.

Key words: airphoto interpretation, accuracy of interpretation, reliability of identification

* International Institute for Aerial Survey and Earth Science (ITC), Enschede, the Netherlands

INTRODUCTION

When trying to apply the interpretation of aerial photographs on recreational facilities, the question was raised, whether the different types of facilities can be interpreted with a sufficient level of consistency (Van der Zee 1992).

To really establish how accurate an interpretation has been, it will be necessary to compare it with other sources of information on the same category. Because the accuracy of secondary data sets is sometimes also questionable, field observations are the only way to really test the accuracy of the interpretation.

The number of field observations to do is a function of budget as well as of the level of accuracy wanted and the type of sampling. When using a random sampling technique statistical formulae can be applied to determine the proper size of the sample. This can also be applied to stratified random sampling. However, with airphoto interpretation often selective sampling is done on basis of the interpretation in order to reduce the number of field observations still further. But then the statistical formulae can not be applied any more.

The assessment of the accuracy of an airphoto interpretation is not a simple statistical matter. It is not just a question of right or wrong that can be easily expressed in a percentage. A percentage of what? A percentage of the total items interpreted that appeared to be correct? Or the percentage of the units in the field that appeared to be correctly interpreted. These are two sides of the medal, but they are different sides. In addition, it is also relevant to know what factors cause the rate of misinterpretations. An example may illustrate this issue. Therefore one special case study was done in the rural area around Enschede, the Netherlands, on recent airphotos and including a complete field verification in order to establish quantified accuracy rates. Because the number of recreational facilities found was not large enough for statistical purposes, and moreover very diverse, the accuracy test was focused on an item that was numerous: farms.

ACCURACY OF INTERPRETATION OF FARMS

In the interpretation of the area around Enschede, the Netherlands (see Fig. 1), it was tried to distinguish farms from houses and buildings with other functions on aerial photographs of 1:7500 taken in September 1988. Main criteria for identifying a farm as such are the presence of fodder silage pits or towers, manure pits, and numerous additional larger and smaller buildings such as barns and sheds. Also clear traces of access to adjoining fields are a good indication for a still active agricultural function.

Many cases are clear, but other cases challenge both the skills in interpretation, as wells as in definition of the category. They may not only be difficult to identify on the airphoto, but also hard to classify even after inspection in the field. What about the small farms with only a marginal agricultural function, and about the farmer that just quit farming but stays living on the farmstead and still keeps some animals for a hobby?

Is his place more a farm than that owned by someone with a job in town who also keeps sheep or horses as a hobby? There is a kind of twilight zone between the purely agricultural and the purely residential functions and it is in this zone that most

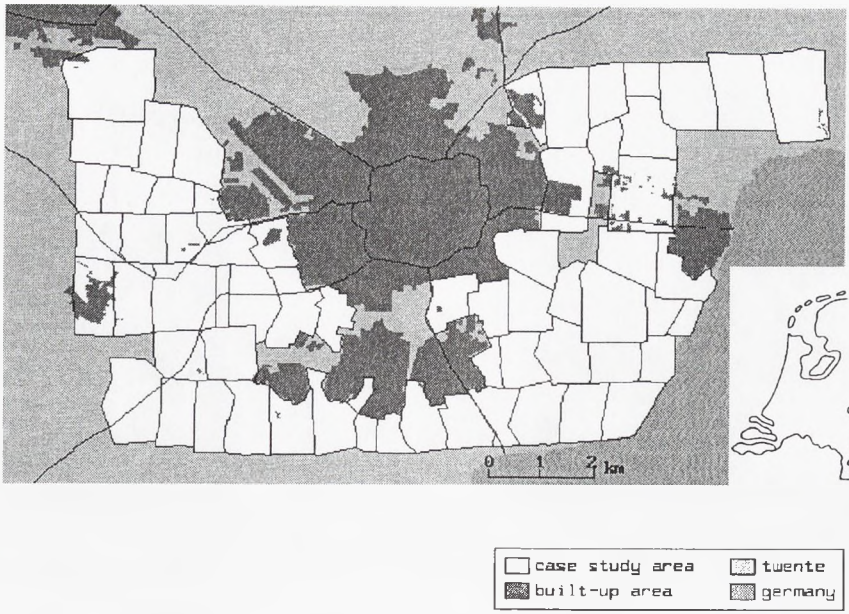


Fig. 1. The location of the study area around Enschede, the Netherlands.

of the misinterpretations and misclassifications will occur. See also Fig. 2.

The great variety in building styles adds to the confusion. Farms occur in a range from the traditional Twente style to the ultra-modern style of farm building. On the other hand old farmhouses built in the typical style of Twente are renovated and reconstructed as residence, new luxury residences are constructed in the style of such old farmhouses. They may be recognised as being non-agricultural because of the nicely designed gardens and the absence of any silage or manure pits. But some modern farms also have nice modern bungalow style residential units associated with them, surrounded by designed gardens. See also Fig. 3.

In total 293 farms were interpreted in the study area, and only 220 were observed in the field. Does this imply an accuracy of 75%? It does if all the 220 had been interpreted as farm, but that is not the case. Of the 293 farms interpreted only 209 also in reality were farms, 71%. But viewed from the other side, of the 220 farms found in the field, 209 had been interpreted as such, a score of 95%, not bad at all. See also Fig. 4.

RIGHT OR WRONG, A RELATIVE QUESTION?

Some of the cases of misinterpretation concerned former farm buildings, which now served as a base for an agricultural contractor. He has the barns and yard still full of large agricultural equipment. In some cases even from field observations it was not definitely clear whether a farm was still in function as such or not. It was not possible to carry out interviews at every place to find out about the main occupation of the residents.

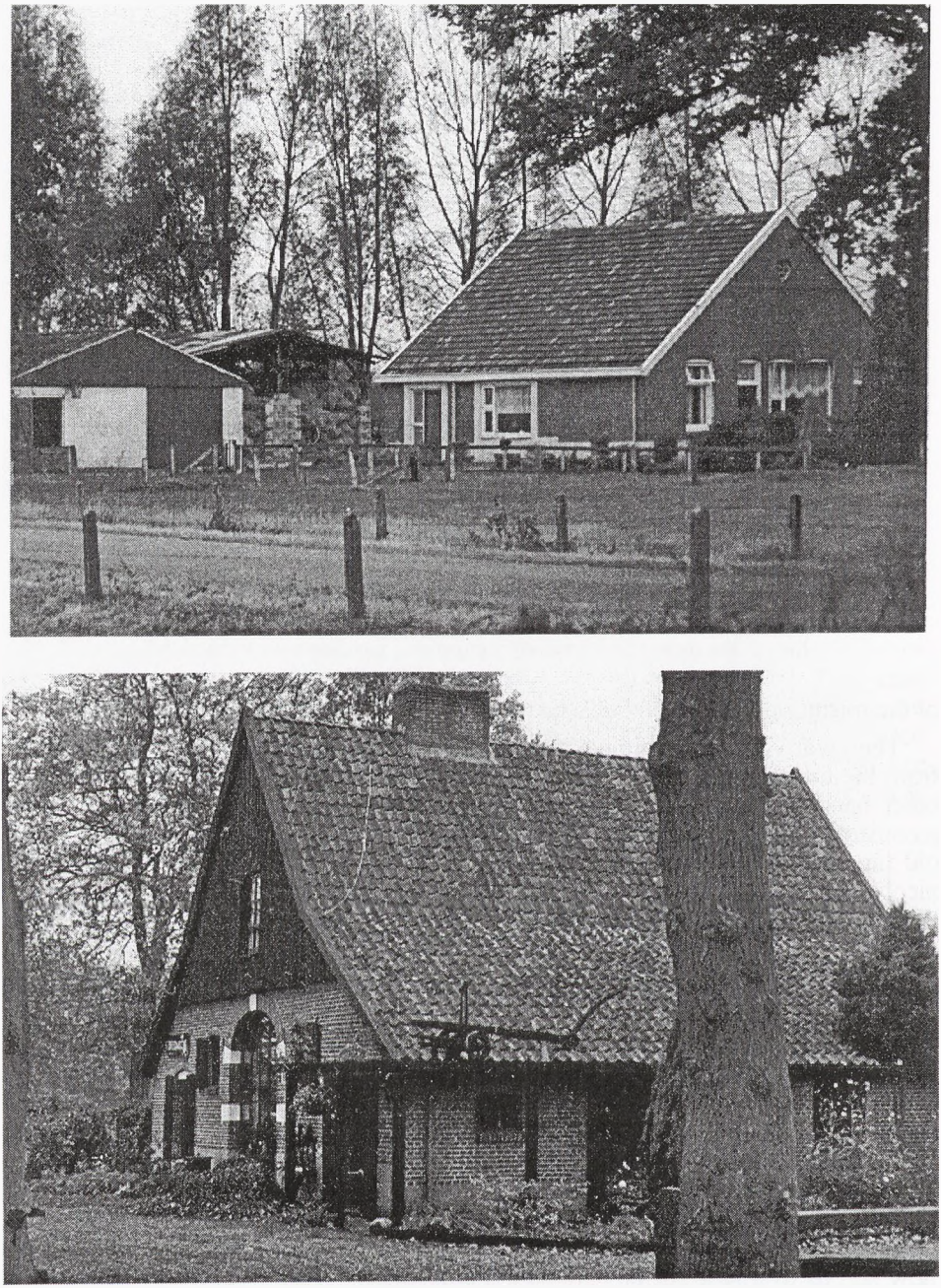
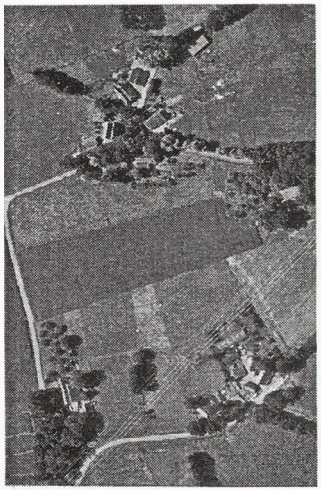
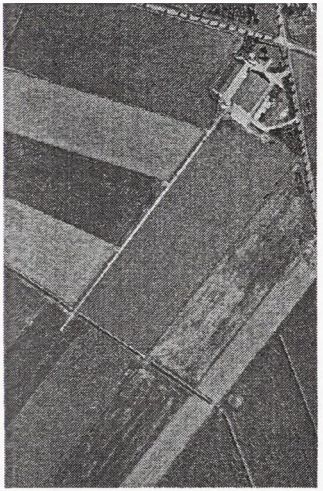
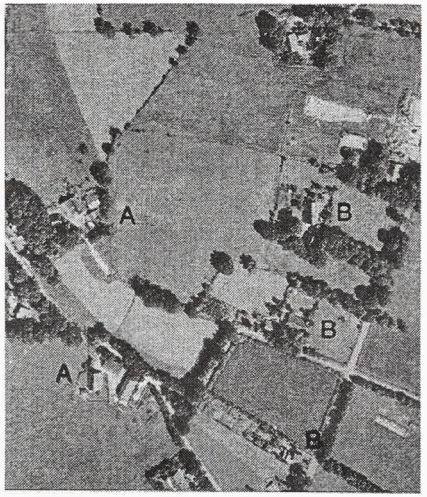


Fig. 2. Farm or non-farm also difficult to asses by field observations.
Up: residential with large sheds, one with hay and straw: farm?
Down: residential in traditional farm style.







0 100 m

Fig. 3. Examples of interpretation of farms in the Enschede area on sections of airphotos 1:7500. Left: Twente style farms, adapted to modern farming practises; middle: modern style farm; right: A= farms, B = residential (former farms or built in farm style).

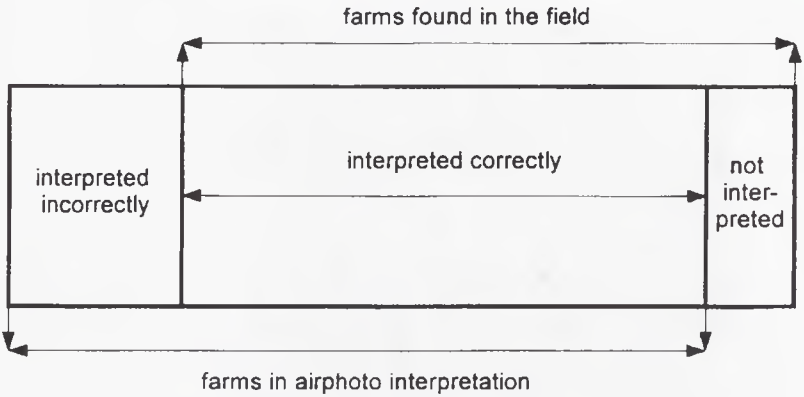


Fig. 4. Accuracy of interpretation of farms expressed in percentages.

It was also considered to be a misinterpretation when the interpreted farm appeared to be only a barn or a stable instead of a full farm. If these cases, which are not real farms but very closely associated with agriculture, are considered to be correct interpretations, the score rises from 71 to 78%.

If also the cases in which a residential function is associated with some (hobby) farm activities such as keeping horses or sheep, or with the presence of a big barn or workshop are counted as correct the score becomes 83%. The remainder consists of 45 cases in which the residential function is clearly dominant, although the building is in the Twente farm style, new or old, and of some cases in which a former farm has got the function of youth centre or children's farm or something similar. A misinterpretation of function rather than form, but still, 17% incorrect. This type of misinterpretation is hard to improve with sharpened criteria after re-interpretation. It is understandable but is it also acceptable?

Looking again at it from the other side, of the 220 farms observed in the field, some may not have been interpreted straightforwardly as farms. But, in a number of cases a doubt between agricultural and residential function was indicated, implying that they should be included in the sample for field verification. If these cases can also be considered as correct interpretations the score becomes 97%.

It depends on the aim of the survey what rate of misinterpretation is still acceptable. If the interpretation is meant to provide a database from which to take a sample for an enquiry survey on farm activities, it is not a real problem that also some non-agricultural units are included. They can be separated in a later stage. The number of sites not included in the interpretation could be more of a problem, but in this case the error remains within acceptable margins, especially if also question-marked interpretations are incorporated. If, however, the interpretation is meant to map the location of the pure agricultural units, the number of mistakes is far beyond the acceptable rate of error. Still, even if the aim is known, it is difficult to establish criteria for acceptable accuracy.

The expression from the other side, that is, the percentage of objects observed in the field that were interpreted correctly, may be interesting, but most often only the

first approach is used. It is seldom possible to carry out a complete field check, only a sample of the interpreted objects is checked, or a small pilot area is checked completely, and the accuracy found før that is extrapolated to the whole interpretation.

IDENTIFYING FARMS ON HIGH-RESOLUTION SATELLITE IMAGES

When in airphoto interpretation it is already difficult to achieve acceptable accuracy for objects such as farms, for satellite image interpretation this can be forgotten all together. This also will hold true for the promised high-resolution satellites, of which none is operational yet.

For the accuracy test only an image of the Russian KVR-1000 was available for comparison, re-sampled in this case to a 5 x 5 meter ground resolution.

When comparing the sections of airphotos with corresponding sections of this image (see Fig. 5), it is clear that, although individual buildings in many cases can be detected without many problems, identification of the function of these buildings is much more difficult on the satellite image. Therefore it may go without saying that satellite images with a coarser resolution, such as SPOT and Landsat Thematic Mapper, cannot be used for identifying buildings at all.

DIFFERENT ASPECTS OF ACCURACY

There is a number of different aspects related to accuracy, that can be expressed in terms of reliability and precision (Tempfli and Kure, 1980). These different aspects have been summarised in figure 6. In the case of the farm study mainly the accuracy of the *interpretation* of airphotos (or images) is concerned, and more specifically the accuracy of the *identification*. Thus it is the reliability aspect that is most relevant.

In the farm study *delineation* of the objects was not of interest, neither were *measurements*.

With respect to location, the *absolute precision* of the farms, that is, the exactness of locating the position of a given point on the photo image to a surveyed ground position, usually given in a standard X, Y, Z co-ordinate system, was not important.

It was the *relative precision* that was of interest. This relative precision refers to the location of individual points relative to one another in terms of distances, angles and height differences.

RELIABILITY OF IDENTIFICATION

Reliability is the degree to which results are consistent upon repetition of an experiment or test (Grinde and Kopf 1986). It can be considered for repeated interpretations applied to similar situations or by different persons (Chenoweth and Gobster 1986).

Thus, reliability may express the chance that the same interpreter will identify objects of the same type correctly and consistently in successive series of interpretations, or, that two or more interpreters will not significantly deviate from

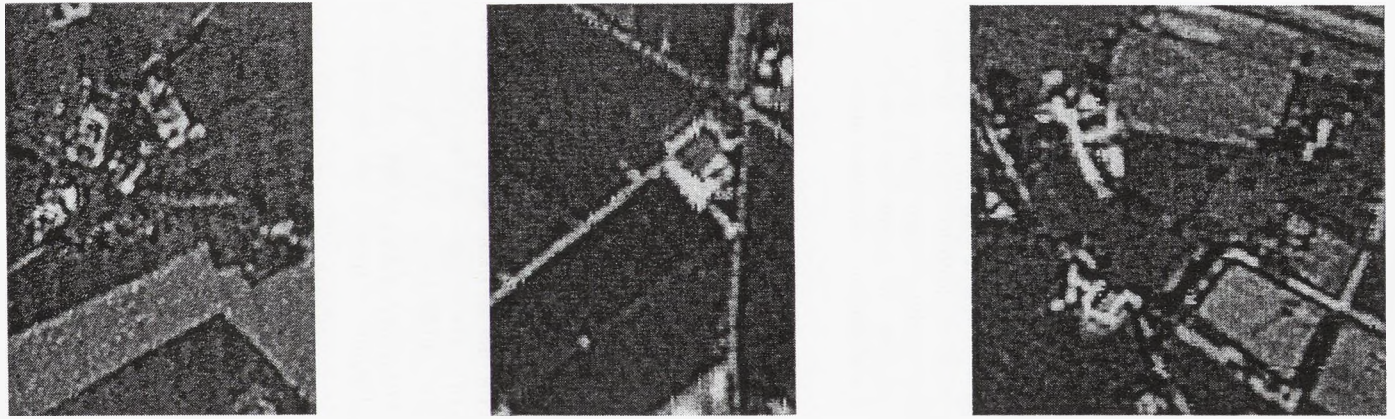


Fig. 5. Examples of interpretations of farms in the Enschede area on sections of a KVR-1000 satellite image, that correspond to the sections of airphotos of Fig. 3.

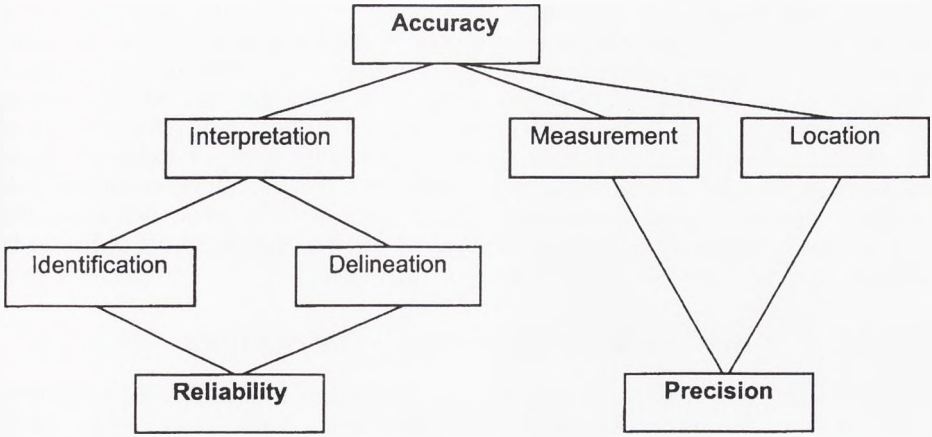


Fig. 6. The different aspects of accuracy.

each other in their identification of identical objects. Of importance are both the specialist reference level and the local reference level of the interpreters. It may be self evident that the results of a trained photo or image interpreter, that is familiar with the topic as well as with the area concerned, will be much better than that of a trainee, that is new to both the topic and the study area.

If two or more interpreters carry out an interpretation, it is therefore necessary to establish the consistency among them, that is, to calibrate the interpretation accuracy of the individual team members. This type of reliability is usually called *inter-observer reliability* or *between-observer reliability*. Agreement can be calculated by determining the number of times the interpreters agreed divided by the total number of interpreted objects by all interpreters (Vining and Stevens 1986). Thus, reliability is *replicability*, or the degree to which different users come up with the same results (Yeomans 1986). This can be rather easily empirically tested. To test the consistency of interpretation of one single interpreter is more difficult. In repeating the same interpretation, he will be influenced by his interpretation of the previous time. In applying the interpretation to another set of airphotos it never will be certain whether differences in interpretation are not caused by existing differences in the image.

In the farm study, an experienced interpreter that knew the area well did the airphoto interpretation. But the influence of the individual performance of different interpreters, as well as that of the scale and quality of the airphotos, can be very well illustrated by the example of house counts in Southern Italy and South-western Spain.

THE ACCURACY OF HOUSE COUNTS BY AIRPHOTO INTERPRETATION

The house counts were carried out by airphoto interpretation and checked by field observations in the context of fieldwork projects of the Rural Survey course of the

ITC. All settlements were concentrated settlements, none of the participants was familiar with the area, but all had been trained in the topic of study. For all settlements first segregation into residential and non-residential functions was made. A settlement plan was sketched in which *blocks* were delineated that had to serve as counting units as well as field observation units. For each block the number of houses was counted on the airphoto. Then for each settlement a number of blocks was taken at random for a field count. The difference between photo count and field count is expressed as a percentage of the photo count, in order to be able to use this as a correction factor. A positive percentage means that the photo count has to be reduced, a negative percentage that it has to be increased.

DIFFERENCES BETWEEN INTERPRETERS

In the case study in Southern Italy for four interpreters four different factors were found: -4,5%, 5.8%, 7.0% and 34.7%. When an accuracy range from +10% to -10% is accepted, only one interpreter needs to improve considerably. But when a range from +5% to -5% is required, only one stays within that range. However, the results do not necessarily reflect a comparison in skill and ability of the interpreters, because each interpreter looked at different settlements. In some settlements it was much easier to distinguish individual houses than in others, because of the difference in structure of the settlement and because of different quality of the airphotos. Some settlements were on first quality 1:17000 scale, others on second rate copies on 1:± 24000 scale.

In the first case study in South-western Spain three interpreters were all using 1:20000 airphotos of the same quality. Still, they obtained different overall results: 21.6%, 7.7% and -12.9%, not too good. Within the overall result for each interpreter there was a range that differed per individual settlement. Thus, also in this case, the differences might not necessarily reflect differences in skill and ability of the interpreters, but could possibly be explained by the fact that each interpreter looked at different settlements, with different character and structure. Therefore, to really compare the individual performance of interpreters their interpretations of the same settlement should be compared. This was done in the second case study in Spain. Four interpreters made interpretations of the same three settlements, and their results were compared. When comparing the overall personal results in the form of dwelling units counted per pair of interpreters, the smallest difference was 1.3%, the largest 13.0%. But, these differences varied per settlement. The results of this comparison could be used to bring the results of the individual interpreters even closer together.

THE EFFECT OF A FEEDBACK WITH FIELD OBSERVATIONS

In the Spain case studies the results of the field check were fed-back into a re-interpretation and a re-count in order to establish what rate of error was due to lack of experience and what simply is inherent to the fact that not everything can be deduced from an airphoto. After re-interpretation, in the first Spain case study the inaccuracy factor became 4.9%, 0.9% and -5.3%. Thus, after using a set of field observations as training set, the interpretation accuracy comes much more within acceptable ranges than before. Still, the difference between the individual interpreters remains more than 10%. Also in the second Spain case study the problem of lack of

local reference level was partly solved by the field checks in the first settlements, after which prior to further field observations, first the other settlements were re-interpreted. The difference between first and second interpretation varied from one person to the other, ranging from 1.6% to 24.7% difference, in absolute numbers 13 and 227 respectively. Those that were already reasonably accurate the first time could hardly improve any more, but those that had been less accurate showed great improvements. Their results became more consistent and the error rates smaller. The differences between interpreters in the re-interpretation became much smaller than in the first interpretation, although according to the statistical tests applied no significant difference between the first and second interpretation could be found.

THE INFLUENCE OF PHOTO SCALE ON THE INTERPRETATION RESULT

Of some settlements also airphotos on 1:10000 were available and interpreted in the first Spain case study. When comparing the results of both first interpretation and re-interpretation of the 1:20000 and the 1:10000, in two cases only slight differences were found: 1.1% and 0.9% for the first interpretation, -0.2% and 1.8% for the re-interpretation. This led to the conclusion that a larger scale not necessarily leads to a higher accuracy. It was observed, however, that the interpretation was much easier and less strenuous on the larger scale. Only the third interpreter showed large differences: -19.9% for the first interpretation, -5.1% for the re-interpretation, but it was concluded that with additional training sets also this interpreter could be brought more in line with the results of his colleagues. Also in the second Spain case study the two scales were again compared. Already after the interpretation of the first three settlements it became clear that there was no significant difference in the interpretation results, but that interpretation was much easier on the 1:10000. Therefore it was decided to only use 1:10000 for the rest of the study. Of those settlements of which no original 1:10000 airphotos were available, sections of the 1:20000 photos could be enlarged to that scale without loss of quality. Thus, only the 1:10000 interpretation eventually was checked in the field this time. The difference between the first interpretation and the field observation still showed large variation between individual interpreters: 28.3%, 11.0%, 7.6%, 1.9%, but the re-interpretation brought the rate of error within the range of +5% to -5% that was thought to be acceptable at that occasion: 4.8%, -0.5%, 0.5%, 0.3%.

The Spanish and Italian case studies made it clear, that the individual performance of the interpreter is of considerable influence on the results of something relatively simple as a count of houses. Therefore, in the interpretation of airphotos for something more complicated, such as recreational phenomena, the identification will be even more strongly influenced by specialist reference level and local reference level. Such reference level can only be gradually built up. Where different persons have to work together on an interpretation, they have to use a common training set in order to bring their accuracy on more or less the same level, or at least to be able to calibrate their personal deviations. And when it is possible to start with a small representative pilot area and then check that in the field first, the results of the interpretation of the rest of the area in general will be much more reliable.

ASSESSING THE ACCURACY OF AN INTERPRETATION OF RURAL ROADS

When not point objects such as farms or houses are concerned, but objects which cover a certain area or length, assessing the number of correctly interpreted objects as a percentage of the total number of objects may not be satisfactory. Depending, of course, on the aim of the survey, in such cases the percentage of area or length that has been correctly interpreted may be more relevant.

Then also the reliability of delineation may become relevant, because the precision of the measurements is largely dependent on that.

The interpretation of the road network in the Enschede case study area, on the 1:7500 set of airphotos (see Fig. 1) will serve as an example of accuracy assessment for linear objects.

THE INTERPRETATION ACCURACY OF ROADS IN THE ENSCHEDE CASE STUDY

Categories of roads

The categories distinguished and the interpretation criteria used are presented in Tab. 1.

When checking the interpretation in the field, the categories highway, railroad and main road and path were maintained. Of the other roads the type of pavement was recorded: asphalt, bricks, concrete, rubble or sand. Gravel roads were included under the rubble roads. Sand roads included also roads, which contained a lot of grass. The difference between a pure sand road and a real rubble-paved road is a gradual one, various stages between the two extremes were found.

Tab. 1. Categories of roads and interpretation criteria

Categories	Photo interpretation criteria
Highway	Separate traffic lanes, split-level crossings, intricate junctions.
Railroad	Straight, gentle curve, no junctions with other roads, light linear feature with dark linear centre.
Main road	Wide roads with intricate crossings and bifurcations, often white demarcation lines and symbols visible (triangles, arrows).
Other paved road	Varying widths and grey-tones, but in general thought to be darker than tracks.
Track or non-paved road	Lighter grey-tones than paved roads, edges less regular, sometimes two narrow light lines suggesting a grass-covered centre.
Path	Also lighter grey-tone but much narrower than tracks.
Tree-covered road	The linear arrangement of the trees suggests the presence of the road, but the type of road or path can not be assessed, this category therefore asks for more elaborate field checking.

The classification could have included also the width of the road and the condition of the pavement, but it was decided to keep it simple.

For comparison the results of both airphoto interpretation and field check were digitised into a GIS (ILWIS in this case). The "interpretation" road network was edited to arrive at the "field" road network, in order to avoid shifts in location by digitising the same line twice. In many cases only the classification had to be adapted, not the delineation. For interpretation examples see Fig. 7.

Then from the two vector maps the total length of the different categories was derived as presented in Tab. 2. Only totals can be compared in this way.

Highway and railroad have been interpreted correctly. In the main roads a small difference is found, due to reclassification rather than misinterpretation. It is in the other road categories that the major differences occur. The total length of roads interpreted and recorded in the field is different, a 9% under-representation. Some roads have disappeared because of town expansion and new roads have been constructed since the date of the aerial photography. But the main cause of the difference is that roads have not been interpreted at all. For a large part this can be attributed to the category paths. If these are not included, the total length of roads interpreted is 280 kilometre, and the total length of roads found in the field 281 kilometre. This is not a bad result at all, and the difference might be attributed to just a rounding error.

Tab. 2. Categories of roads interpreted and field-checked (in kilometres)

air photo interpretation			field check		
category	total length		category	total length	
	km	%		km	%
highway	5.1	1.7	highway	5.1	1.6
railroad	4.6	1.6	railroad	4.6	1.4
main road	16.5	5.6	main road	14.6	4.5
paved road	174.5	59.6	asphalt road	160.9	49.9
tree covered road	38.5	13.2	brick road	8.6	2.7
track	40.4	13.8	rubble road	25.6	7.9
path	13.5	4.6	concrete road	0.4	0.1
			sand road	61.2	19.0
			path	41.7	12.9
TOTAL	293.3	100	TOTAL	322.7	100

ANALYSIS OF MISINTERPRETATIONS

To analyse and quantify the misinterpretations, the two road networks have to be superimposed and compared. This was done by converting the vector maps into grid

maps with cells representing 25 x 25 meter in the terrain. Then a cross-table was made that expresses the coincidence of interpreted categories with field categories in numbers of grid-cells. Because in the process of transforming the maps into grid-cell maps a certain amount of generalisation takes place, the combinations can not be expressed in units of length any more. Instead the results are presented in percentages of grid-cells in table 3 for two approaches: how the field roads had been interpreted and what the interpreted roads were in reality.

For example, with respect to the main roads identified in the field it can be stated that 99% were also interpreted as main road. A very high accuracy therefore. That it is not 100% can be attributed to the fact that at road crossings the computer sometimes attributed the grid-cell to the other road category. When considering the roads that have been interpreted as main road, however, it appears that only 87 % were really main road and 13 % not. Even though in this case it was more a question of reconsideration of the classification after the field visit than a real misinterpretation, the accuracy factor becomes quite different. And so also the other categories can be viewed from both points of view.

Tab. 3. Comparison of interpretation and field observation of the road network (in percentages of grid-cells)

Interpretation categories	Road categories found in the field								
How the "field" roads had been interpreted	Main road	Asphalt road	Brick road	Sand road	Path	Rubble road	Concrete road	No road	
Main road	99.4	1.2	-	-	-	0.3	-	-	
Paved road	0.4	81.8	62.2	42.9	-	33.1	-	26.1	
Tree covered	-	11.3	8.8	19.0	10.6	19.1	7.1	5.0	
Track	0.2	1.9	6.1	32.5	15.9	29.2	85.7	68.9	
Path	-	0.1	2.0	-	31.0	2.6	-	-	
No road	-	3.7	20.7	5.5	42.5	15.7	7.1	-	
TOTAL	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
What the interpreted roads were in reality	Main road	Asphalt road	Brick road	Sand road	Path	Rubble road	Concrete road	No road	TOTAL
Main road	87.3	12.2	-	-	-	0.5	-	-	100
Paved road	0.0	75.0	3.0	15.6	-	5.4	-	1.0	100
Tree covered	-	43.9	1.8	29.3	10.9	13.2	0.1	0.8	100
Track	0.1	7.0	1.2	46.9	15.3	18.8	0.8	10.0	100
Path	-	1.0	1.2	-	92.5	5.3	0.0	-	100
No road	-	17.5	5.2	10.4	53.6	13.2	0.1	-	100

Of the asphalt roads identified in the field 82% had been interpreted as paved road, thus correctly, and 11% was obscured by trees on the photographs. About 4%

could not be interpreted because they were constructed after the date of the aerial photography. Of the brick-paved roads 62% was interpreted as paved road, nearly 9% was obscured by trees, and about 33% had been interpreted as track. The rather high percentage that was not interpreted at all is not a real misinterpretation though. The driveways on the large parking lots of a recreation area have been incorporated into the inventory after the field survey, but were not included in the airphoto interpretation. Apart from that, this category still has a much lower accuracy score than the asphalt roads. But, in absolute sense this category is only small. Therefore, why bother?

Of the roads interpreted as paved roads, 75% appeared to be asphalt, 3% brick, 5% rubble, and 16% sand roads. Only 1% was not found in the field any more, because they had been removed in the course of town expansion activities. When taking asphalt roads and brick roads together as paved roads, then 81% had been interpreted as such correctly. The major source of error has been the tree cover. But still, of all the roads interpreted as paved, only 78% were so in reality. Especially with respect to the smaller roads it has been difficult to interpret whether they are really paved or not. Also the sand roads or rubble roads in this area are quite straight, neat and wide. See also Fig. 7.

Of the sand roads 43% had been interpreted as paved roads, almost 20% were obscured by trees. Of the rubble roads 33% had been interpreted as paved road, 30% as track and 19% were obscured by trees. Taking both categories together as tracks, then 40% was interpreted as paved road and only 32% as track. Not a high accuracy. Also from the other point of view the accuracy is not really high. Of what was interpreted as tracks, 47% appeared to be sand roads and 19% rubble roads, together 66%. In addition 8% appeared to be paved road and 15% was classified in the field as path.

Of the paths 31% was interpreted as such, 11% was obscured by trees, and 16% was interpreted as track, rather a classification error than an interpretation error. But, 43% of the paths was not interpreted at all, either because they were not visible in the airphotos, or because during the interpretation they were not included as such, being considered an integral part of a park. After field work it was then decided to include such paths too. If the non-interpreted paths are left out of consideration, then 54% of the paths have been identified as such. Still not a very high score. But, when looking at it from the other side, what had been interpreted as path in 93% of the cases was found to be path in reality too.

Resuming, it can be stated that the total length of the road network, excluding paths, has been accurately interpreted, but that the subdivision into different categories cannot be made with sufficient reliability and consistency. Paths, at least in this part of the country, cannot be interpreted with sufficient accuracy, mainly because they are hidden under the trees and too narrow and/or winding to be revealed by alignment of trees.

COMPARING THE AIRPHOTO INTERPRETATION WITH THE TOPOGRAPHIC MAP

Why should the road network be interpreted from airphotos in the first place? Is not sufficient information already available on topographical maps? The maps

available for the Enschede case study that was carried out in 1989 were published in 1977, based on a revision in 1974. The road classification for the sake of comparison has been regrouped into main roads, secondary roads, other paved roads and non-paved roads. Paths are also indicated on the topographical map, but have not been included in the comparison.

Also these road networks have been incorporated into the GIS. Actually, the roads of the topographic map had been the first to be digitised; then editing this first set resulted in the other sets. The roads on the topographical map were then compared with both the interpretation results as the field observations. Of all roads interpreted 19% did not appear in the map, and of the roads found in the field 24% was not in the map, of which 36% main roads.

Also with respect to the classification differences can be observed. Of the main roads and also the secondary roads the interpretation corresponds reasonably to the classification on the map: 97% and 82% respectively (or 93% if tree covered roads can be included as correct interpretation). At the lower orders of roads the discrepancies are larger, but it cannot be established whether this is because of misinterpretation or because of road improvement since the time of mapping. It is of course more interesting to compare the map with the field observations, see Tab. 4.

The main roads in the map did correspond to those in the field. The small differences indicated in the table have to be attributed to the grid-cell allocation problem that was already mentioned. Of the secondary roads 93% corresponds to paved. Of the other paved roads of the map 59% was asphalt or brick, 25% rubble and 16% sand road or path. Of the non-paved roads in the map 19% in reality was (now) paved, 14% rubble and 67% sand road or path.

Of the main roads in the field 62% was indeed indicated as main road, 26% as secondary road and 9% did not appear. Of the paved roads (asphalt or brick) 65% was indicated as secondary road, 9% as other paved road, 10% as non-paved road and 16% did not appear. Of the rubble and sand roads 62% was indicated as non-paved, 11% as other paved and only 4% as secondary road. But 23% was not indicated on the map at all.

Therefore, although the map may not be up-to-date with respect to the actual condition of the roads, it is not really less accurate in this respect than the results of the airphoto interpretation. But with respect to the road network, irrespective of subdivision into categories, the map is far from complete. This difference will become much larger if the available topographical map is of a very old date and/or of a small scale, a situation not uncommon in many countries Thus one conclusion is that airphoto interpretation can be very useful for a rapid up dating of the map for the road network.

Another conclusion from the Enschede case study is that not too much accuracy can be expected with respect to a classification of roads into different categories in this case. However, the accuracy with which a classification of roads can be made by airphoto interpretation strongly depends on the types of road available, the criteria that can be applied, as well as the character of the area, and therefore will differ from one area to the other.

Fig. 7. Some examples of the different road categories on airphotos and KVR-1000 satelite image. On the airphotos (left and right) "A" are asphalted roads, "S" are sand roads. The roads in the satellite image (centre) are also partly obscured by trees and their shadows. The distinction between sand roads and asphalted roads was possible in this case, because a difference in grey tone can be observed. But whether this can be done consistently remains a question.

Tab. 4. The road network on the topographical map compared with field observations (in percentages of grid-cells)

Roads on the topographical map	Roads found in the field			
What the roads in the field are on the map	Main road	Paved road	Non-paved road	Path
Main road	62.4	0.1	0.2	-
Secondary road	25.8	64.7	3.9	1.2
Other paved road	0.8	9.3	11.0	2.3
Non-paved road	1.9	9.6	61.7	34.2
No road	9.1	16.1	23.3	62.3
TOTAL	100.0	100.0	100.0	100.0

What the roads on the map are in the field	Main road	Paved road	Non-paved road	Path	TOTAL
Main road	95.9	2.2	2.0	-	100
Secondary road	3.1	93.4	3.1	0.4	100
Other paved road	0.4	58.7	37.3	3.6	100
Non-paved road	0.3	18.7	64.8	16.2	100
No road	1.7	36.0	28.2	34.0	100

In many countries a much simpler classification, for which airphoto interpretation criteria can be established easily, will suffice. The resulting interpretation accuracy can then be expected to be higher. Unfortunately, the classification that one likes to apply is not always coinciding with the categories that can be easily and consistently distinguished from each other on airphotos, leave alone satellite images.

In the Enschede case study the high incidence of roads obscured by trees contributed to the inaccuracy. This may be a factor of importance in many other countries too. In an open polder or prairie landscape this factor will be of much less importance.

Moreover, it may go without saying that airphotos can only reveal the presence of roads. About the juridical, social and economic aspects of accessibility airphotos cannot give any information.

THE ACCURACY ASPECT IN THE ROADS CASE STUDY

In the case of the roads study not only the reliability of the identification is important, but also that of the delineation. Small inaccuracies in delineation should be eliminated in the analysis. The procedure to digitise one road network as a base and derive the others from that by editing was not only meant to avoid a lot of laborious digitising, but also to avoid the creation of slight shifts in location that actually do not exist. In the case study area with the road network consisting of rather smooth, straight or gently curving lines, such problems are not very large. When lines are more sinuous, special attention has to be paid to avoid mismatches. In such cases the scale becomes very important.

In the comparison of the different sets of the road network their lengths were used as parameter. This implies measurement and thus the aspect of precision.

With the present day GIS, lengths and areas can be very accurately measured. By default they are expressed in meters with up to five figures behind the point. Of course this is ridiculous, because of the lack of reliability of the delineation. For this reason rounding off in full meters, or even kilometres, or in hectares makes more sense.

With respect to location, also here the relative location was most important. Although by taking the road network from the topographical map as a base for digitising, the absolute location was rather precise.

The comparison of the road networks also illustrates that a numerical right/wrong rate is far from satisfactory. Also the possible error factors have to be distinguished and analysed (Dodt and Van der Zee 1974).

It should be established whether errors are due to:

- lack of information and/or inaccurate information regarding actual appearance of objects;

- ignorance of processes involved in taking and processing aerial photography;

- low-level photo-physiognomic differentiation of objects on available aerial photography;

- carelessness or tiredness on the part of the particular photo- interpreter.

From this analysis it then can be deduced whether a low degree of reliability can be subject to improvements by further training, or whether it is inherent to the type of object concerned.

CONCLUSIONS

A statement on the accuracy of an airphoto interpretation can not be simply made in terms of right/wrong statistics. Right or wrong often is a relative question and moreover can be viewed from two different points: number of interpreted objects found to be correct in reality or number of objects in reality that have been correctly interpreted.

In addition there is a number of aspects to accuracy, that often are interrelated: the reliability of the identification, the reliability of the delineation, the precision of measurements and the precision of location. Not all aspects are always applicable or relevant. This depends strongly on the categories on which the interpretation is focused as well as on the purpose of the interpretation and the scope the study.

Still, the reliability of the identification appears to play a key role.

In addition to assessing the rate of right or wrong, also the causes of misinterpretation should be analysed and if possible fed back into the interpretation.

Proper training of the interpreters and feedback of field observations into the interpretation can greatly improve the accuracy. Sometimes a reclassification may be necessary, when the classification wanted does not match which what is possible to be distinguished on aerial photographs. And in some circumstances it has to be accepted that airphoto interpretation can not achieve the required accuracy and other means of information collection have to be applied.

REFERENCES

CHENOWETH, R.E. & GOBSTER, P.H. (1986) Wildland description and analysis. In Smardon, R.C., Palmer, J.F., Felleman, J.P., eds. *Foundations for visual project analysis*. Chichester (Wiley), pp 81-101.

DODT, J., ZEE, D. van der (1974). Identification of rural land-use types. (A method of teaching in, and data acquisition by, photo-interpretation). *ITC Journal*, 5, 599-616.

GRINDE, K., KOPF, A. (1986). Illustrated glossary. In Smardon, R.C., Palmer, J.F., Felleman, J.P., eds. Foundations for visual project analysis. (Wiley), pp. 308-333.

TEMPFLI, K., KURE, J. (1980). Large scale surveys: aerial versus ground survey methods. *ITC Journal,* 4, 696-717.

VINING, J., STEVENS, J.J. (1986). The assessment of landscape quality: major methodological considerations. In Smardon, R.C., Palmer, J.F., Felleman, J.P., eds. *Foundations for visual project analysis.* Chichester (Wiley), pp. 167-186.

YEOMANS, W.C. (1986). Visual impact assessment: changes in natural and rural environment. In Smardon, R.C., Palmer, J.F., Felleman, J.P., eds. *Foundations for visual project analysis*. Chichester (Wiley), pp. 201-222.

ZEE, D.van der (1992). *Recreation studied from above. Airphoto interpretation as input into land evaluation for recreation.* Dissertation Agricultural University Wageningen.

Dick van der Zee

HODNOTENIE PRESNOSTI INTERPRETÁCIE LETECKÝCH SNÍMOK: PRÍKLAD FARIEM, DOMOV A VIDIECKYCH CIEST

Hodnotenie presnosti interpretácie leteckých snímok je zložité. Štúdia dokumentuje túto skutočnosť prostredníctvom niekoľkých ukážok interpretácie fariem, domov a vidieckych ciest.

Čo je správne a nesprávne môže byť relatívna otázka. Okrem toho treba uvedený problém vidieť prostredníctvom rôznych pohľadov: z hľadiska počtu interpretovaných objektov z existujúcich, alebo z hľadiska počtu správne interpretovaných objektov.

Na presnosť interpretácie vplýva veľa rôznych aspektov. Spoľahlivosť identifikácie je jeden z najdôležitejších.

Zručnosť a skúsenosť interpretátora, ako aj mierka použitých leteckých snímok, môžu mať veľký vplyv na presnosť interpretácie.

Ak je interpretácia orientovaná na líniové a plošné, nie bodové objekty, vyjadrenie presnosti množstva správne interpretovaných objektov môže byť neuspokojivé. Percento správne interpretovanej plochy alebo dĺžky môže byť dôležitejšie.

Pre zlepšenie presnosti interpretácie je veľmi dôležité analyzovať a kvantifikovať druh nesprávnych výsledkov interpretácie.

Preložil J. Feranec