

Errors due to Integration of Image Data to Geographical Information Systems

Gerhard Schweinfurth
Institut für Photogrammetrie und Fernerkundung
Universität Karlsruhe
Englerstr. 7, D-7500 Karlsruhe 1
Federal Republic of Germany

Intercommission III/IV

Abstract

Traditionally Geographical Information Systems are assigned for registration and processing data, captured in form of existing maps and field observations (vector data). Nowadays there is an increasing use in processing imagery data (raster data) in these systems, too.

The combination of satellite data (LANDSAT Thematic Mapper) and digital aerial photographs with spatial data is demonstrated by several examples.

Errors occurring by this process result, besides obvious sources, from natural variations, from original measurements or arise through processing. These errors were analyzed and a strategy is pointed out to minimize them.

1 Introduction

Geographical information systems are powerful but expensive tools. The expense arises not only from the cost of the hardware, software, and trained personnel, but also from the not inconsiderable costs of data collection, data capture, and processing. According to the use of information systems, costs of data collection often exceed the money, which have to be raised up for providing and operating such a system. 'It is implicit that the collection and processing of environmental data leads to improvements in environmental management and control. This can only be so if data that are collected, entered, stored, and processed are sufficiently reliable and error-free for the purposes for which they are required' [Burrough 86].

Users of geographical information systems assume very often that all the data, which are stored in these systems, were totally error-free. Apparently there are neither input- nor measuring errors, measures of accuracy are unusually. 'Adding two maps together by means of a simple overlay, imply that both source maps can be treated as perfect, completely deterministic documents with uniform levels of data quality' [Burrough 86].

The high quality of maps, cartographers produce by means of corresponding tools of geographical information systems, causes the false impression that exactly drawn lines are always derived from error-free data. Nevertheless smooth contour lines can be computed by arbitrary clustered points without representing the natural feature of the countryside.

Till today only a few authors have pointed out these problems. 'It is remarkable that there have been so few studies on the whole problem of residual variation and how errors arise, or are created and propagated in geographical information processing, and what the effects of these errors might be on the results of the studies made' [Burrough 86].

2 Sources of possible errors in GIS

Burrough divides three main groups of factors governing the errors that may be associated with geographical information processing (table 1).

<p>Obvious sources of error</p> <p>Age of data Areal coverage – partial or complete Map scale Density of observations Relevance Format Accessibility Cost</p>
<p>Errors resulting from natural variations or from original measurements</p> <p>Positional accuracy Accuracy of content – qualitative und quantitative Sources of variation in data: – data entry or output faults – observer bias – natural variation</p>
<p>Errors arising through processing</p> <p>Numerical errors in the computer: – the limitations of computer representations of numbers Faults arising through topological analyses: – misuse of logic – problems associated with map overlay Classification and generalization problems: – methodology – class interval definition – interpolation</p>

Table 1: Sources of possible errors in GIS [Burrough 86]

'The word 'error' is used here in its widest sense to include not only 'mistakes' but also to include the statistical concept of error meaning 'variation'. The first group include topics that

are most obvious and easy to check on; the next group contains more subtle sources of error that can often only be detected while working intimately with the data. The third group is perhaps the most important because it includes the mistakes, errors, and misapprehensions that can arise as result of carrying out certain kinds of processing. These errors are the most difficult to spot because they require an intimate knowledge of not only the data, but also of the data structures and the algorithms used' [Burrough 86].

3 Errors occuring by combining imagery and spatial data

Errors resulting from natural variations and arising through processing often occur by integrating imagery data into geographical information systems. Some of these typical errors shall be demonstrated by means of three studies, which are pre-investigations to establish a geographical information system based on digital imagery and spatial data.

3.1 Testarea 'Karlsruhe'

The topographic dataset of the 'Karlsruhe testarea' such as settlement area, woodland, waters, main roads etc. are digitized from a topographic map at a scale of 1:100.000 (TK 100) and stored as vectors. By means of a diploma thesis LANDSAT TM-data were transformed into the same projection using identical points and overlaid with the rasterized data of the TK 100 [Roth 87]

Figure 1 shows the cross-section 'woodland-settlement'. All settlement and industrial areas within the woodland are detected, but at the borders of the wood there are many colour-margins.

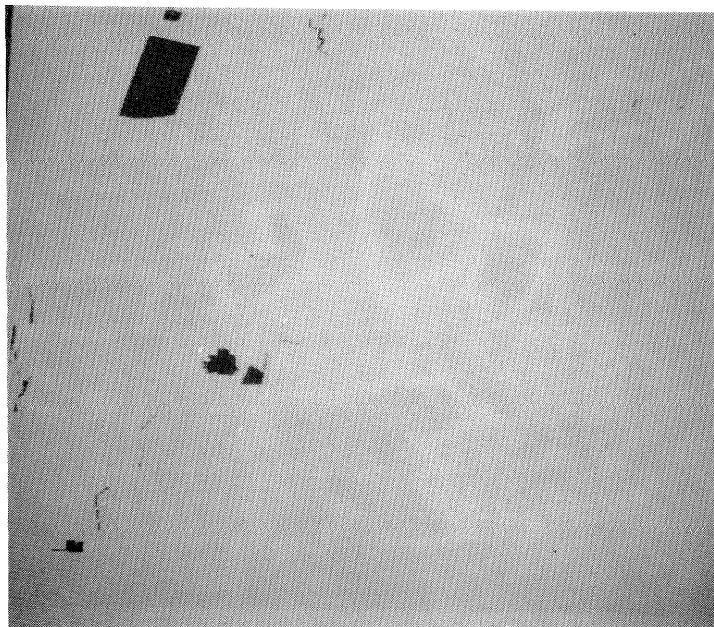


Figure 1: Cross-section 'woodland-settlement'

The sources of these errors cannot be separated exactly. Are they *digitizing-, registration-, transformation- or interpolation-errors*?

Figure 2 demonstrates another problem: *generalisation-errors* within the basic map.

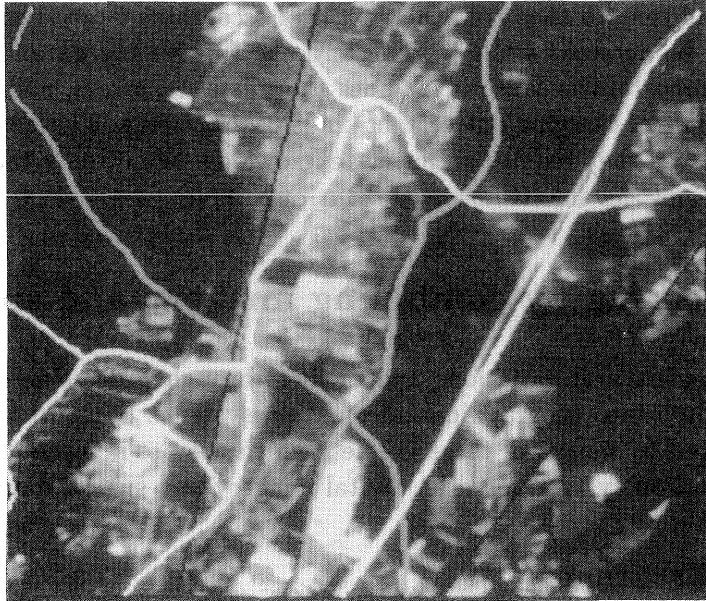


Figure 2: Different plotted lines of the highway

3.2 Testarea 'Filderraum'

A research project sponsored by the German Research Community (DFG) investigates possibilities to registrate changes of landuse by LANDSAT TM-data [Baumgart 87]. The testarea is the so-called 'Filderraum' southeast of Stuttgart city.



Figure 3: Classification errors due to inaccurate classes

Figure 3 shows an enlarged section of the classification result. In the first instance the classified settlements correspond very well to the digitized settlement areas. In detail there are a lot of inaccurate classifications, however.

These errors originate in inexact classes, e.g. the class 'settlement' includes gardens and parks, too.

3.3 Thematic Orthophotomap

The third example elucidates errors caused by overlaying large scaled aerial image-data with spatial data.

Figure 4 shows a superimposition of a land use plan (FNP) and a digital orthophoto [Schulz 87].

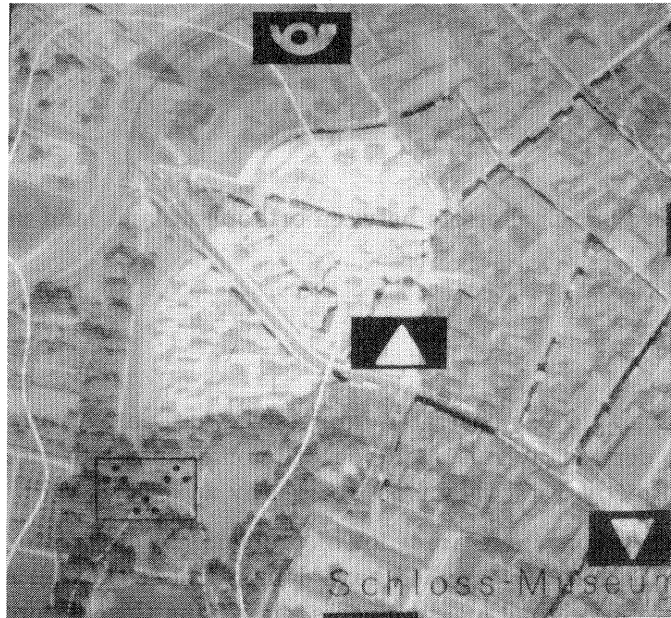


Figure 4: Superimposition land use plan - digital orthophoto

The FNP is based on the German base map (DGK 5) at the scale of 1:5.000, whereas the orthophoto is derived from an aerial photograph at 1:13.000 and the DTM is in a 50 m raster. Difficult to see by the black-white reproduction there are conflicts between orthophoto and FNP especially in streets. Whereas the positional errors in small streets are caused on height differences of the houses and on radial displacements, the origin of the eminent differences in the main roads only can explain by inaccuracies of the DTM and the computation of the orthophoto.

Figure 5 shows the positional displacements of skyscrapers very impressive, and at the same time errors due to digitizing the land use plan and to computing the orthophoto.



Figure 5: Positional errors between orthophoto and land use plan

4 Conclusions

Especially the last example demonstrates impressively that data overlay is only an aid to interpretate. Results derived by this method are not qualified to store in a geographical information system directly. First they must be revised by an expert and if necessary corrected.

Due to transformation errors there are a lot of problems by integrating imagery data into information systems. Besides tests applied have demonstrated that all errors resulting from natural variations or from original measurements and arising through processing are possible.

Caused by these sources of possible errors the accuracy of the content of geographical information systems is open to question. Of course these systems contain eminent information to use them, however, they have to be examined exactly.

For that reason contents and size of geographical information systems have to be defined very carefully before installing. There is no system to serve all purposes.

5 Outlook

How carefully and detailed a geographical information system has to be planned demonstrate the efforts to install an Official Topographic-Cartographic Information System (ATKIS) in the Federal Republic of Germany [Harbeck 87, Grünreich 88]. Figure 6 shows the structure of this planned information system.

Only to generate maps of different scales by means of 'digital map-models' (DKM) several 'digital landscape-models' (DLM) have to be established. An object-catalogue describes minute, what to take up into the corresponding model (figure 7).

Future will show though these rigid definitions are useful and can be preserved consequently.

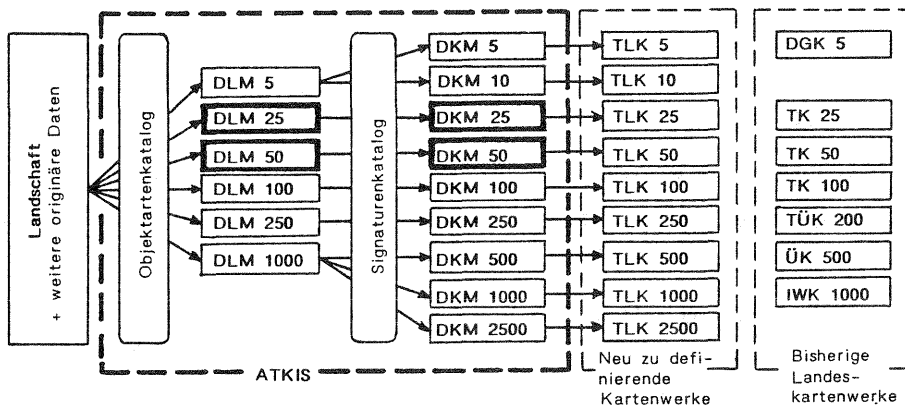


Figure 6: Structure of ATKIS and relations between DLM, DKM and maps [Harbeck 87]

Objektartenkatalog		Stand 08.86	Objektbereich Verkehr	Objektgruppe Straßenverkehr	
Schlüssel	Objektart Attribut, Name, Hierarchischer Hinweis Attributwert			DLM 5	DLM 2
	Straße				
	Anzahl der Fahrspuren			1	1
	tatsächliche Anzahl			1	1
	Klassenbildung (z.B. 3 und weniger, 4 und mehr)			0	0
	Qualität der Richtungsfahrbahnen			1	1
	physisch getrennt			1	1
	ungetrennt			1	1
	Gesamtfahrbahnbreite			1	1
	tatsächliche Breite in m			1	1
	Lage in Bezug zur Erdoberfläche			1	1
	im Geländeneiveau, auf Damm oder Einschnitt			1	1
	aufgeständert (Hochstraße)			1	1
	im Tunnel			1	1
	Verkehrsbedeutung, außerörtlich			1	1
	Fernverkehr			1	1
	Regionalverkehr			1	1
	zwischenörtlicher Verkehr			1	1
	Verkehrsbedeutung, innerörtlich			1	1
	Durchgangsverkehr			1	1
	Sammelverkehr			1	1
	Anliegerverkehr			1	1
	Gesetzliche Gliederung/Widmung			1	1
	Bundesautobahn			1	1
	Bundesautobahn, zugleich Europastraße			1	1
	Bundesstraße			1	1
	Bundesstraße, zugleich Europastraße			1	1
	Landesstraße			1	1
	Kreisstraße			1	1
	Sonstige Straße			1	0

0 = Keine Aufnahme 1 = Unbedingte Aufnahme 2 = Bedingte Aufnahme durch Verknüpfung mehrerer Merkmale
3 = Bedingte Aufnahme durch definierte Auswahl bei der Erfassung

Figure 7: Section of the object catalogue [Harbeck 87]

References

- [Baumgart 87] Baumgart, J.: Digitale Erfassung von Veränderungen der Landnutzung.- DFG-Provisional Report, IPF Universität Karlsruhe, 1987.
- [Burrough 86] Burrough, P.A.: Principles of Geographical Information Systems for Land Resources Assessment.- In: Monographs on Soil and Resources Survey, Oxford: Clarendon 1986, 194 p.
- [Grünreich 88] Grünreich, D.: Das Amtliche Topographisch-Kartographische Informationssystem (ATKIS) – Konzeption und Realisierungsstand.- In: Geo-Informationssysteme in der öffentl. Verwaltung, Tagungsband C, IPF Universität Karlsruhe, 1988, 95-105.
- [Harbeck 87] Harbeck, R.: Das AdV-Vorhaben „Amtliches Topographisch-Kartographisches Informationssystem (ATKIS)“ – Inhaltliche Konzeption.- Nachrichten aus dem Karten- und Vermessungswesen, Reihe I, Heft 99, IfAG Frankfurt a.M. 1987, 7-14.
- [Roth 87] Roth, A.: Entwicklung zur Überlagerung und Kombination von kartographischen und Fernerkundungsdaten als Vorstufe eines Landinformationssystems.- Unpublished diploma thesis, IPF Universität Karlsruhe, 1987.
- [Schulz 87] Schulz, E.: Digitale Erstellung einer Thematischen Orthophotokarte unter Verwendung eines Flächennutzungsplans.- Unpublished diploma thesis, IPF Universität Karlsruhe, 1987.