

DIE VERWENDUNG VON ZUSATZINFORMATION ZUR VERBESSERUNG DER DIGITALEN MULTISPEKTRALEN KLASSIFIZIERUNG

Elisabeth Dennert-Möller
Institut für Photogrammetrie und Ingenieurvermessungen
Universität Hannover
Bundesrepublik Deutschland
Kommission III, W.G. III/5

ZUSAMMENFASSUNG

Die Qualität von Klassifizierungsergebnissen hängt eng zusammen mit der Information, die dem Klassifizierungssystem zur Verfügung steht. Die Einbeziehung von Zusatzin-formation - d.h. jeder Art von analog oder digital vorliegender Information, die Aussagen über das konkrete Klassifizierungsproblem liefert - verbessert die Trennbarkeit der Klassen. Es werden hier zwei verschiedene Methoden, Zusatzinformation in den Klassifizierungsprozeß einzubeziehen, vorgestellt: die Erweiterung des Merkmals-raums und der Einsatz von verallgemeinerten a priori-Wahrscheinlichkeiten. Ihre Wir-kungsweise wird anhand von Fernerkundungsaufnahmen aus Wattgebieten demonstriert.

ABSTRACT

The quality of classification results depends on the information available to the classi-fication system. Introduction of ancillary information - that means any kind of ana-logue or digital information, which directly or indirectly states anything about the classification problem - improves the separability of the classes. In this paper two different methods of introducing ancillary information into the classi-fication procedure are presented: the use of a priori probabilities and the addition of an artificial channel. Examples are given for both methods.

1. Einleitung

Die Einbeziehung von Zusatzinformation beruht auf dem Gedanken, dem Klassifizie-rungssystem möglichst viel verschiedenartige Information zuzuführen. Als **Zusatzin-formation** wird hier jede Art von Information bezeichnet, die über die eigentliche aus-zuwertende Fernerkundungsaufnahme hinaus **direkt** oder **indirekt** Aussagen über das konkrete Klassifizierungsproblem liefert. Dazu gehören Kenntnisse über die speziellen Aufnahmebedingungen, z.B. in bezug auf das Aufnahmegerät, die klimatischen Bedin-gungen zum Zeitpunkt der Aufnahme usw., und direkt oder indirekt problembezogene Information, z.B. über die Bodenverhältnisse, wie sie vor allem in Form von themati-schen Karten vorliegen kann. Auf welche Weise Zusatzinformation in das Klassifizie-rungssystem eingebracht werden kann, hängt vom speziellen Problem und der Art der Information ab. Dabei kann man die Zusatzinformation für eine Vorverarbeitung ver-wenden, für eine nachträgliche Verbesserung des Klassifizierungsergebnisses einsetzen /Ehlers et al. 1984, Itten, Fasler 1979/, oder in den eigentlichen Entscheidungsprozeß der Klassifizierung einbeziehen.

Im folgenden werden zwei Ansätze vorgestellt, die die Einbeziehung von Zusatzinfor-mation in die Maximum-Likelihood (ML) Klassifizierung ermöglichen. Am Beispiel von Auswertungen von Multispektralabtastreraufnahmen aus den Wattgebieten der deutschen Nordseeküste wird ihre Wirkungsweise demonstriert.

2. Erweiterung des Merkmalsraums um einen künstlichen Kanal

Liegt die Zusatzinformation für jeden Bildpunkt (b_x, b_y) des auszuwertenden Bildes vor, so liegt es nahe, sie als zusätzlichen (künstlichen) Spektralbereich in die Klassifizierung miteinzubeziehen. Das gilt vor allem, wenn es sich um ein zusätzlich vorliegendes Multispektralbild handelt, das sich durch eine geometrische Transformation in das gegebene Bildkoordinatensystem überführen läßt. Im Falle von Zusatzinformation in Form einer thematischen Karte muß diese zunächst digitalisiert und geometrisch angepaßt werden. Entsprechendes gilt für als Punktreihe nur an bestimmten Bildpunkten vorliegende Zusatzinformation. Diese muß durch eine für alle Bildpunkte definierte Funktion approximiert werden. Die Erweiterung des Merkmalsraums um diesen zusätzlichen Spektralbereich entspricht der Einführung einer $(n+1)$ ten Zufallsvariablen x_{n+1} , die im allgemeinen unabhängig von den spektralen Merkmalen x_1, \dots, x_n angenommen werden kann.

Mißt man die Trennbarkeit von je zwei zu unterscheidenden Klassen ω_i und ω_j mit dem Jeffrey-Matusita-Maß /Swain, King 1973/

$$J_{ij} = 2(1 - e^{-\alpha_{ij}})$$

mit

$$\alpha_{ij} = \frac{1}{8}(\underline{m}_i - \underline{m}_j)^T \left(\frac{\underline{C}_i + \underline{C}_j}{2} \right)^{-1} (\underline{m}_i - \underline{m}_j) + \frac{1}{2} \log \frac{\det\left(\frac{\underline{C}_i + \underline{C}_j}{2}\right)}{\sqrt{\det(\underline{C}_i) \cdot \det(\underline{C}_j)}}$$

so ergibt sich für den Fall der Erweiterung um einen zusätzlichen Kanal

$$\alpha_{ij}^* = \frac{1}{8}(\underline{m}_i^* - \underline{m}_j^*)^T \left(\frac{\underline{C}_i^* + \underline{C}_j^*}{2} \right)^{-1} (\underline{m}_i^* - \underline{m}_j^*) + \frac{1}{2} \log \frac{\det\left(\frac{\underline{C}_i^* + \underline{C}_j^*}{2}\right)}{\sqrt{\det(\underline{C}_i^*) \det(\underline{C}_j^*)}}$$

Hierbei bezeichnen \underline{m}_i , \underline{m}_j die Mittelvektoren und \underline{C}_i und \underline{C}_j die Kovarianzmatrizen der Klassen ω_i und ω_j und

$$\underline{m}_k^* = \begin{pmatrix} \underline{m}_k \\ \hline m_{k,n+1} \end{pmatrix} \quad \text{und} \quad \underline{C}_k^* = \begin{pmatrix} \underline{C}_k & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \dots 0 & \sigma_{k,n+1,n+1} \end{pmatrix}$$

die erweiterten Mittelvektoren und Kovarianzmatrizen für $k = i, j$. Es ist

$$\alpha_{ij}^* = \alpha_{ij} + \frac{[m_{i,n+1} - m_{j,n+1}]^2}{8[\sigma_{i,n+1}^2 + \sigma_{j,n+1}^2]} + \frac{1}{2} \log \frac{\sigma_{i,n+1}^2 + \sigma_{j,n+1}^2}{2[\sigma_{i,n+1} + \sigma_{j,n+1}]}$$

also

$$\alpha_{ij}^* \geq \alpha_{ij}$$

wobei das Gleichheitszeichen nur für

$$m_{i,n+1} = m_{j,n+1} \wedge \sigma_{i,n+1} = \sigma_{j,n+1}$$

gilt. Die Trennbarkeit von zwei Klassen, die nicht im künstlichen Kanal durch identische Mittelwerte und Varianzen gekennzeichnet sind, wird also verbessert.

Enthält der Zusatzkanal z.B. Höheninformation in Form eines digitalen Geländemodells, so gliedert sich die spektral zunächst einheitliche Klasse ω_i auf in verschiedene von ihrer Geländehöhe abhängige Unterklassen $\omega_{i1}, \dots, \omega_{ij}$, von denen gegebenenfalls jede durch ein Trainingsgebiet repräsentiert werden muß. Dagegen lassen sich zwei verschiedene Klassen, die z.B. aufgrund verschiedener Höhen spektral gleich erscheinen, auf diese Weise voneinander trennen. Ob die Erweiterung des Merkmalsraums durch einen künstlichen Kanal sinnvoll ist, läßt sich nicht allgemein beurteilen, sondern nur für das konkret vorliegende Klassifizierungsproblem mit der zur Verfügung stehenden Zusatzinformation.

3. Einführung von a priori-Wahrscheinlichkeiten

Eine andere Möglichkeit, Zusatzinformation in das Klassifizierungsverfahren einzubringen, ist der Einsatz der a priori Wahrscheinlichkeiten $p(\omega_i)$, die Bestandteil der ML-Entscheidungsfunktionen sind /Dennert-Möller 1983, Strahler 1980/ für diesen Zweck.

Kennt man in etwa die flächenmäßige Verteilung der einzelnen Klassen im Bild, z.B. aus einer thematischen Karte oder aus einer bereits vorliegenden Klassifizierung einer anderen Aufnahme, so läßt sich dieses Kenntnis dazu verwenden, die einzelnen Klassen verschieden zu gewichten: Setzt man in die Entscheidungsfunktionen verschiedene Werte für $p(\omega_i)$ ein, so ergibt sich bei festgehaltenem Wahrscheinlichkeitsniveau eine Vergrößerung derjenigen Klassenellipsoide, für die $p(\omega_i)$ groß ist, eine Verkleinerung für diejenigen, für die $p(\omega_i)$ klein ist.

Die ML-Entscheidungsregel läßt sich verallgemeinern, indem man annimmt, daß $p(\omega_i)$ von einer weiteren Variablen κ abhängt. Die a priori Wahrscheinlichkeit für das Auftreten einer Klasse ω_i wird nicht mehr als konstant angenommen, sondern variiert in Abhängigkeit von einer (bekannten) Gegebenheit κ ; anstelle von $p(\omega_i)$ soll $p(\omega_i, \kappa)$ a priori bekannt sein.

Die Variable κ zeigt den Status einer äußeren Bedingung an, die für die Wahrscheinlichkeit des Auftretens einer bestimmten Klasse eine Rolle spielt. κ kann z.B. eine bestimmte Bodenbeschaffenheit anzeigen, in deren Abhängigkeit bestimmte Pflanzenarten mehr oder weniger wahrscheinlich auftreten. Im Gegensatz zur Einführung einer Variablen als künstlichen Kanal wird hierbei vorausgesetzt, daß die klassenspezifischen Verteilungsfunktionen $p(\underline{x} | \omega_i)$ hiervon nicht beeinflußt werden:

$$p(\underline{x} | \omega_i) = p(\underline{x} | \omega_i, \kappa) .$$

Man kann zeigen /Dennert-Möller 1983/, daß zur Erweiterung der ML-Entscheidungsregel für diesen Fall lediglich die a priori Wahrscheinlichkeiten $p(\omega_i)$ durch die bedingten a priori Wahrscheinlichkeiten $p(\omega_i | \kappa)$ zu ersetzen sind.

In (4.2) wird der vereinfachte Zusammenhang zwischen der Strömungsgeschwindigkeit des auflaufenden Wassers und der Ablagerung von verschiedenen Sedimenttypen in Wattgebieten über die a priori-Wahrscheinlichkeiten in die Klassifizierung eingebracht. Die Variable κ steht in diesem Fall für die Bildkoordinaten und $p(\omega_i, \kappa)$ beschreibt den Zusammenhang zwischen der geometrischen Lage jedes Bildpunktes und der Wahrscheinlichkeit für das Auftreten einer der Sedimentklassen durch lineare Funktionen. Genauso kann man z.B. im Falle des Vorliegens eines digitalen Geländemodells für jeden Bildpunkt den Zusammenhang zwischen der Höhe und der Wahrscheinlichkeit eines gewissen Pflanzenbewuchses in die Klassifizierung einbeziehen.

4. Beispiele

4.1 Erhöhung der Klassentrennbarkeit durch Einbeziehung von Niedrigwasserzeiten

Am Beispiel der Auswertung einer LANDSAT-Aufnahme (Abb.1) der nordfriesischen Wattgebiete soll hier die Erweiterung des Merkmalsraums um einen künstlichen Kanal veranschaulicht werden.



Abb.1: Kanal 6 der LANDSAT MSS-Aufnahme vom 11.8.1975

Der Klassifizierung liegt eine geologische Fragestellung zugrunde. Es sollen fünf Watt-sedimente im Bild erkannt werden, die sich in ihrem prozentualen Gehalt an Schluff und Ton mit den in Tab.1 angegebenen Grenzen unterscheiden.

Klasse	Sedimenttyp	Gehalt an Schluff und Ton
1	Wattsand	< 5 %
2	schwach schlickiger Wattsand	5 - 10 %
3	sandiger Schlicksand	10 - 25 %
4	schlickiger Schlicksand	25 - 50 %
5	Schlick	> 50 %

Tab.1: Feingliederung der Wattsedimente gemäß Vorschlag des Kuratoriums für Forschung im Küsteningenieurwesen (KFKI)

Bei Verwendung der Kanäle 4, 5 und 7 der LANDSAT-Aufnahme ergeben sich die in Tab.2 angegebenen Trennbarkeitswerte als mittlere Klassifizierungsgüte in den Trainingsgebieten /Dennert-Möller 1983/.

S	1	2	3	4	5
1	79.7				
2	21.9	48.0			
3	0.2	10.6	77.5		
4	6.0	8.3	6.7	41.0	
5	0.6	3.8	14.9	17.0	69.0

Tab.2: Trennbarkeit der Klassen bei Verwendung der Kanäle 4, 5 und 7

Die Wattgebiete sind täglich zweimal vollständig von Wasser bedeckt. Da der Zeitpunkt des Niedrigwassers an verschiedenen Orten zu verschiedenen Zeitpunkten eintritt, ist in einer LANDSAT-Szene jeder Sedimenttyp mit verschiedenen Feuchtigkeitsstufen abgebildet. Das gilt auch für Ausschnitte der hier vorliegenden Größe. Da sich die Feuchtigkeit einer Oberfläche auf ihr Reflexionsverhalten auswirkt, liegt es nahe, diese als zusätzliches Merkmal in eine Klassifizierung einzubeziehen. Hierfür wurde ein "Niedrigwasserzeitenkanal" erzeugt, indem auf der Basis der Zeitpunkte des Eintritts des Niedrigwassers an 22 Punkten am Tag der Aufnahme innerhalb des LANDSAT-Bildes ein Ausgleichspolynom 2.Grades berechnet wurde, das für jeden Bildpunkt einen Niedrigwasserzeitpunkt liefert.

Durch eine geeignete lineare Transformation der Zeiten auf Grauwerte entsteht ein künstlicher Kanal, dessen Äquidensitenbild, überlagert mit einer Verarbeitung des Kanals 6 der LANDSAT-Aufnahme, in Abb.2 dargestellt ist.

Die Klassifizierung der Trainingsgebiete ergibt bei Verwendung des um diesen künstlichen Kanal erweiterten Merkmalraums die in Tab.3 angegebenen Trennbarkeiten.

S	1	2	3	4	5
1	91.6				
2	6.26	92.1			
3	0.3	1.1	98.5		
4	1.8	.3	0.0	95.8	
5	0.5	0.0	0.0	0.0	99.8

Tab.3: Trennbarkeit unter Verwendung von Zusatzinformation



Abb.2: Äquidensiten der transformierten Niedrigwasserzeiten, einer Verarbeitung des Kanals 6 überlagert

4.2 Einbeziehung des Zusammenhangs zwischen Strömungsgeschwindigkeit und Sedimentablagerung

Dem Beispiel liegt eine Flugzeugabstasteraufnahme des Wattgebiets Grildumersiel bei Wilhelmshaven vom 11.8.1976 zugrunde. Der eingesetzte BENDIX-M²S ist ein 11-Kanal-Zeilen-Abtaster, dessen Kanäle den Bereich von Ultraviolett bis zum thermalen Infrarot umfassen. Abb.3 zeigt die Panorama-entzerrten Bilder der Kanäle 7 bis 11, aufgenommen bei einer Flughöhe von 600 m.

Die Definition der Klassen wird nach meeresbiologischer Fragestellung /Dennert-Möller, Dörjes 1978/ vorgenommen:

Klasse	Watt-Typ
1	Schlickwatt
2	Sandwatt
3	trockener Sand
4	Zostera (Seegraswiesen)
5	Diatomeen (Kieselalgen)
6	Wasser

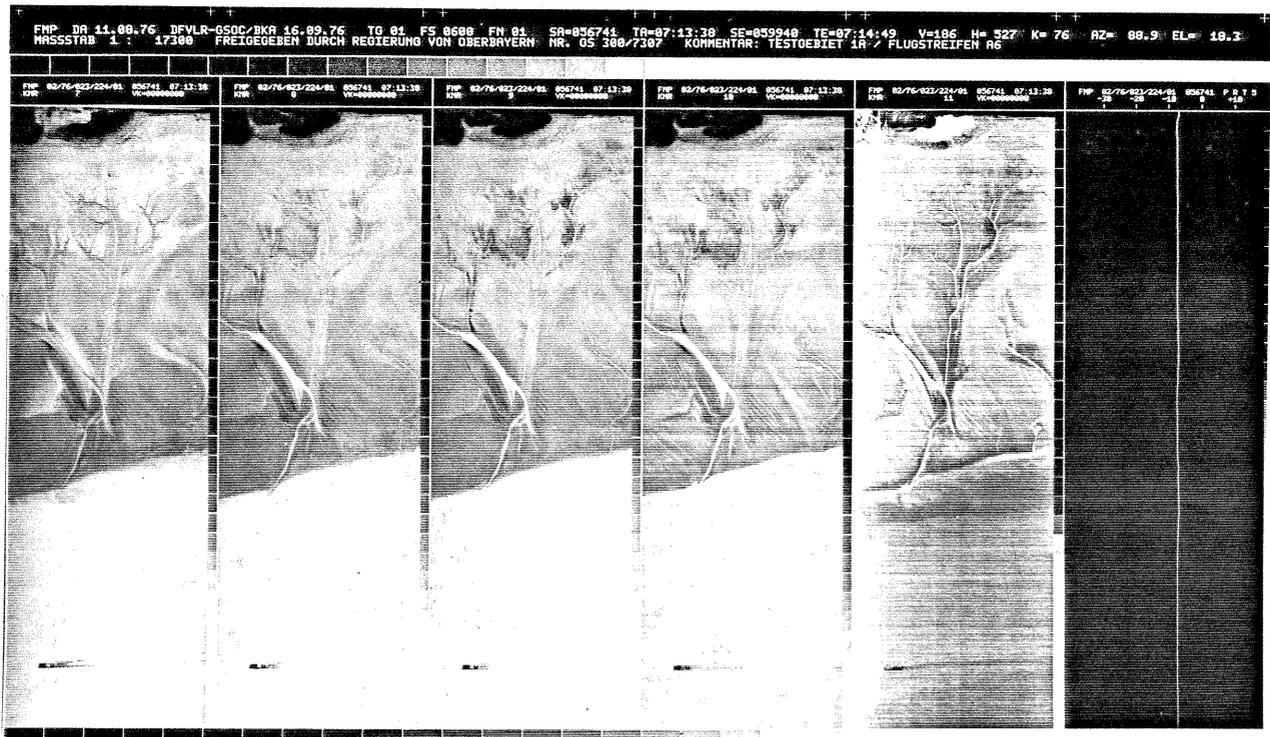


Abb.3: Kanäle 7 bis 11 der BENDIX M²S-Aufnahme vom 11.8.1976

Die Klassifizierung unter Verwendung der Kanäle 7 und 9 (Rot und Infrarot) führt auf die in Tab.4 aufgeführten Werte der Klassifizierungsgüte in den Trainings- und Kontrollgebieten.

Klasse		1	2	3	4	5	6	unkl.
1	T	99.1	.2		0.7			
	K	86.1	12.1	0.2	1.2			0.4
2	T		99.8	.1				
	K	1.3	90.5	6.4	0.1			2.2
3	T			100.				
	K			100.				
4	T	.8	.4		83.	15.6		.2
	K	1.2	0.3		81.3	15.4		2.6
5	T		.2		9.3	90.2		.2
	K	6.3	32.6	22.1	14.4	20.1		4.4
6	T						100.	
	K						100.	

Tab.4: Klassifizierung der Trainings- (T) und Kontrollgebiete (K)

Die Flugrichtung verläuft in etwa senkrecht zur Wasser/Watt-Kante. In Vereinfachung des Zusammenhangs zwischen der Strömungsgeschwindigkeit und den sich ablagernden Sedimenten /Reineck 1970/ kann man annehmen, daß die drei Streifen Sandwatt, Mischwatt, Schlickwatt senkrecht zur Flugrichtung liegen. Somit kann ein Zusammenhang zwischen der Zeilennummer und der in dieser Zeile möglichen Sedimente hergestellt werden. Als einfaches Modell wird angenommen, daß die Wahrscheinlichkeit des Auftretens der Klassen Schlickwatt und damit auch Zostera und Diatomeen von der Wasserkante zum Land hin linear ansteigt. Analog steigt die Wahrscheinlichkeit für das Auftreten der Klassen Sand und trockener Sand linear in umgekehrter Richtung und fällt wieder im Bereich, in der die Watt/Wasserkante verläuft. Für die Klasse Wasser kann angenommen werden, daß die Wahrscheinlichkeit ihres Auftretens im Wattbereich nahe bei Null liegt, im Bereich der Watt/Wasser-Kante steil ansteigt und anschließend in der großen Rinne gleich 1 ist, da hier keine andere Klasse auftreten kann. Der Verlauf der hier beschriebenen stückweise linearen stetigen Wahrscheinlichkeitsfunktionen $p(\omega_i, b_x)$ ist in Abb.4 dargestellt.

Diese Wahrscheinlichkeitsfunktionen können, wie in 3. beschrieben, als a priori Wahrscheinlichkeiten in der Klassifizierung miteinbezogen werden.

Die Ergebnisse der so durchgeführten Klassifizierung in den Trainings- und Kontrollgebieten sind in Tab.5 zusammengestellt.

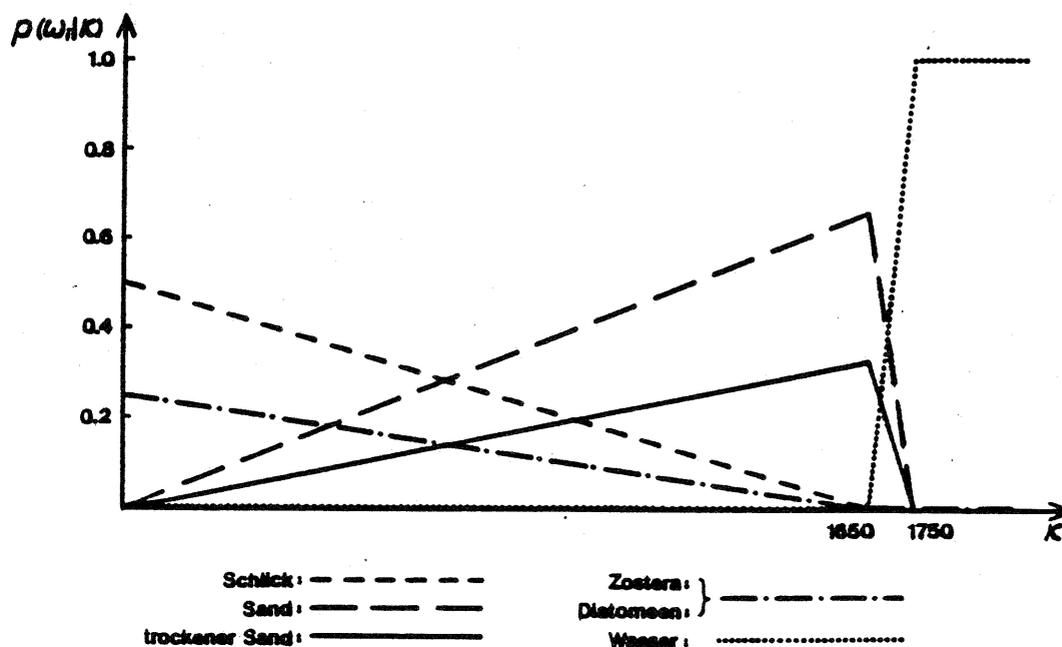


Abb.4: A priori-Wahrscheinlichkeiten der Klassen

Gegenüber den in Tab.4 dargestellten Ergebnissen ergibt sich vor allem eine Verbesserung der Trennung Schlick-Sand und eine Erhöhung der Anzahl unklassifizierter Bildpunkte. Die Klasse Diatomeen konnte auch hier nicht unterschieden werden. Im Hinblick auf die Klassifizierungsgüte sind die Ergebnisse vorsichtig zu interpretieren. Durch geeignete Wahl der a priori Wahrscheinlichkeitsfunktionen läßt sich fast immer erreichen, daß der Anteil der in den Kontrollgebieten richtig erkannten Bildelemente eine vorgegebene Schranke nicht unterschreitet. Die Zuverlässigkeit der angegebenen Werte läßt sich in etwa anhand der Übereinstimmung des gewählten Modells mit der Realität - soweit nachprüfbar - ermesen.

Klasse	Gebiet	Schlick	Sand	trock.Sand	Zostera	Diatomeen	Wasser	unkl.
Schlick	T	98.9	.4					.8
	K	88.2	.7	.3				10.8
Sand	T		98.9					.1
	K		92.6	1.0				6.4
trock.Sand	T			99.8				.2
	K			98.5				1.5
Zostera	T	.8			81.3	15.4		2.6
	K	1.0			81.9	8.9		2.4
Diatomeen	T				9.1	85.8		5.3
	K	.3	22.8	15.8	2.0	14.5		45.1
Wasser	T						100.0	
	K						100.0	

Tab.5: Klassifizierung der Trainings- (T) und Kontrollgebiete (K) unter Verwendung von a priori Wahrscheinlichkeiten

5. Zusammenfassung

Die Einbeziehung von (nicht-spektraler) Zusatzinformation in den Klassifizierungsprozeß unterstützt die Auswahl geeigneter Trainingsgebiete, ermöglicht die nachträgliche Korrektur von Fehlklassifizierungen oder erhöht die Trennbarkeit der Klassen. Letzteres läßt sich erreichen, indem die Zusatzinformation in Form eines künstlichen Spektralbereichs zur Erweiterung des Merkmalsraums benutzt wird. Die Erhöhung der Trennbarkeit der Klassen läßt sich hier unter Verwendung des Jeffrey-Matusita-Maßes allgemein beweisen. Als weitere Möglichkeit zur Einbeziehung von Zusatzinformation können die in der Regel als konstant angenommenen a priori-Wahrscheinlichkeiten als abhängig von einer zusätzlichen Variablen betrachtet werden. Durch die entsprechende Verallgemeinerung der Maximum-Likelihood-Entscheidungsregel wird die Voraussetzung geschaffen, funktionale Modelle äußerer Bedingungen in den Klassifizierungsprozeß miteinzubeziehen. Die Bereitstellung und Verknüpfung von Fernerkundungsdaten und geowissenschaftlicher Information in gemeinsamen Informationssystemen lassen hier weitere Leistungssteigerungen der digitalen Klassifizierung erwarten.

Literatur

- DAVIS, W.A. und PEET, F.G. 1977: A Method of Smoothing Digital Thematic Maps. Remote Sensing of Environment 6 No.1
- DENNERT-MÖLLER, E. 1983: Untersuchungen zur digitalen multispektralen Klassifizierung von Fernerkundungsaufnahmen mit Beispielen aus den Wattgebieten der deutschen Nordseeküste. Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen Nr.127. Dissertation.
- DENNERT-MÖLLER, E. und DÖRJES, J. 1978: Multispektrale Klassifizierung von Wattgebieten. W78-04 BMFT
- EHLERS, M., DENNERT-MÖLLER, E., KOLOUCH, D. und LOHMANN, P. 1984: Nonrecursive Filter Techniques in Digital Processing of Remote Sensing Imagery. Pres.Paper XVth Int.Congress of ISPRS, Rio de Janeiro, Brasilien, 17.-29.Juni

- HUTCHINSON, C.F. 1982: Techniques for Combining LANDSAT and Ancillary Data for Digital Classification Improvement. Photogrammetric Engineering and Remote Sensing 48, No.1
- ITTEN, K.I. und FASLER, F. 1979: Thematic Adaptive Spatial Filtering of LANDSAT Landuse Classification Results. Proc. of the 13th Int.Symp. ERIM, Vol.2
- REINECK, H.E. 1970: Das Watt. Frankfurt/M.
- STRAHLER, A.H. 1980: The Use of Prior Probabilitites in Maximum-Likelihood Classification. Remote Sensing of Environment 10.
- SWAIN, P.H. und KING, R.C. 1973: Two Effective Feature Selection Criteria for Multi-spectral Remote Sensing. Proc 1rst Int.Joint Conf. on Pattern Recognition. IEEE 73 CHO 82 1-9C