

5 DATENINTERPRETATION UND DATENANALYSE

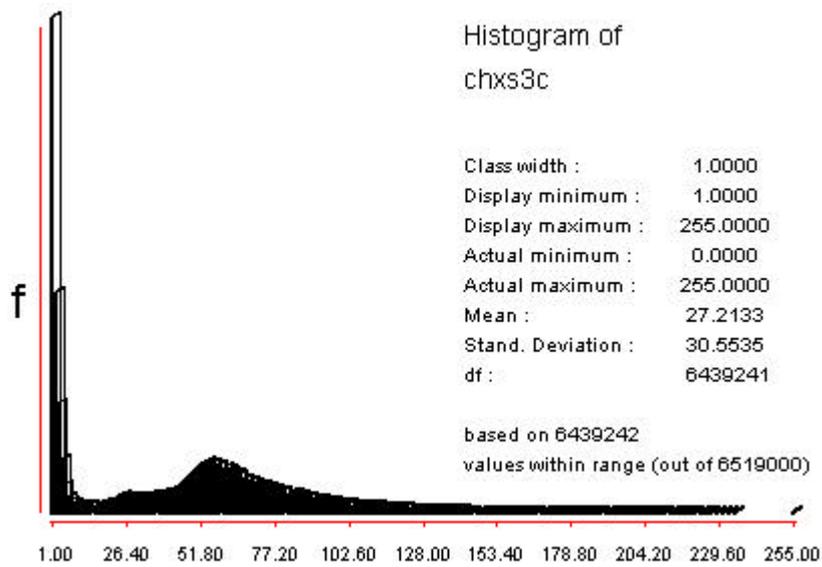
Um die in den Fernerkundungsdaten gespeicherten Informationen für bestimmte Fragestellungen nutzbar zu machen, können verschiedene Auswertungsverfahren angewandt werden. Zu den wichtigsten gehören im Rahmen dieser Arbeit die visuelle Bildinterpretation und die digitale Datenanalyse.

5.1 VERFAHREN ZUR VERBESSERUNG DER VISUELLEN DARSTELLUNG

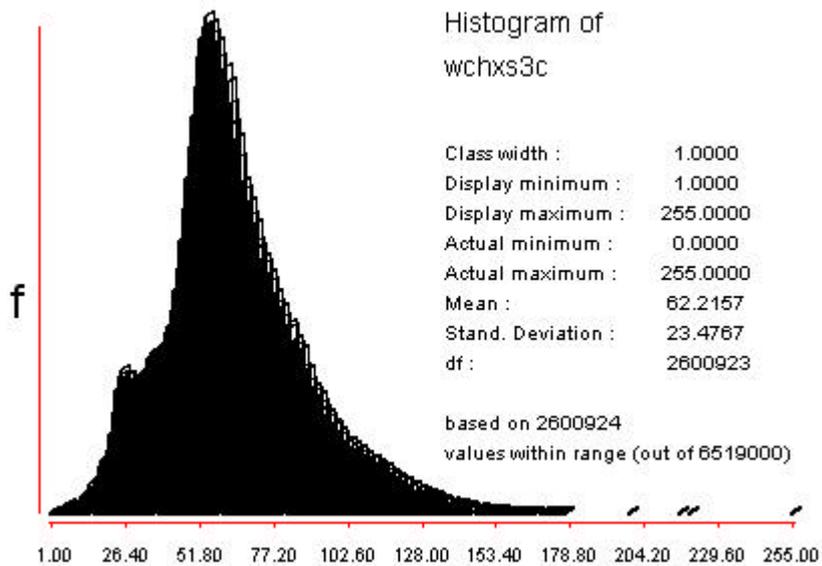
Die visuelle Bildinterpretation ist ein *typisch menschliches Auswerteverfahren* (ALBERTZ, S.139). Dabei wird die optische Wahrnehmungsmöglichkeit und die Erfahrung des Bildbetrachters genutzt, um Objekte wie Straßen, Berggipfel, einheitliche Geländeflächen u.ä. zu erkennen und aus ihren Beziehungen zueinander bestimmte Schlußfolgerungen zu ziehen. Um entweder möglichst viele oder aber gezielte Informationen aus den verschiedenen Spektralbereichen optisch zugänglich zu machen, wurden hier unterschiedliche Methoden zur Verbesserung der visuellen Darstellung (*image enhancement*) angewandt.

5.1.1 ENTFERNUNG REDUNDANTER OBJEKTFLÄCHEN

Die **SPOT**-Szene der Insel El Hierro stellt neben dem eigentlichen Inselkörper den umgebenden Atlantik sowie zwei Wolkenfelder im Norden und im Westen der Insel dar (siehe Abb. 2-1). Diese Objekte beanspruchen in dem 651.9 km² umfassenden Ausschnitt der Satellitenszene eine Fläche von 393 km². Die Wasserflächen sowie durch Wolken verursachte Schatten weisen im Image aus dem **nIR** überwiegend Grauwerte (**DNs**) von 0 bis 25 auf, bei Wolken und Dunst bewegen sich diese Werte überwiegend in den **DNs** von 130 bis 255. Sie sind nicht nur für die Vegetationsanalyse irrelevant, sondern sie stören durch die Menge der Informationen an beiden Enden der Grauwertskala eine kontrastreiche Darstellung der Inseloberfläche. Dies wird schon an den nachfolgenden Histogrammen des infraroten Kanals (Histogramm 5-1 und Histogramm 5-2) deutlich, stellt sich aber besonders in der Kombination (*color composit*) mehrerer Spektralkanäle dar (Abb. 5-1 und Abb. 5-2). Die genannten Grauwertinformation können auch von anderen Objekten auf der Insel reflektiert werden. Um die entsprechenden Pixelinformationen auf dem Inselkörper nicht zu verlieren, wurde durch wiederholte Reklassifikation ein *Boole'sches* Image erzeugt, in dem den Wasserflächen der Wert 0, allen anderen Flächen der Wert 1 zugewiesen wurde. Die Wolken und Wolkenschatten wurden als Polygone manuell digitalisiert und, ebenfalls mit dem Wert 0, dem *Boole'schen* Image hinzugefügt. Durch eine einfache Multiplikation kann mit dieser „Maske“ allen redundanten Flächen eine „no data“-Information zugewiesen werden, während alle anderen Pixel ihren Z-Wert behalten. Das *Boole'sche* Image stand für alle weiteren Operationen zur Verfügung.



Histogramm 5-1 Grauwertinformationen des **nIR**-Bandes vor der Entfernung von Wasserflächen und Wolken.



Histogramm 5-2 Verteilung der Spektralinformationen des **nIR**-Bandes nach Entfernung der Wasser- und Wolkenflächen.

Window from compo c: 1768 r: 736 to c: 2032 r: 984

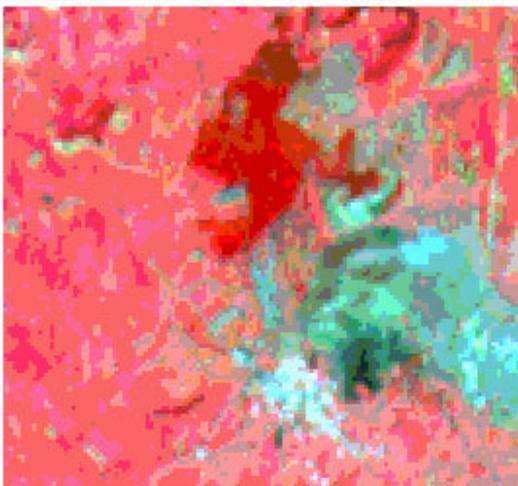


Abb. 5-1 Ausschnitt aus einer Kombination (*color composit*) der Original-Images Xs1, Xs2 und XS3 vor der Entfernung von Wasser- und Wolkenflächen.

Window from compm c: 1768 r: 736 to c: 2032 r: 984

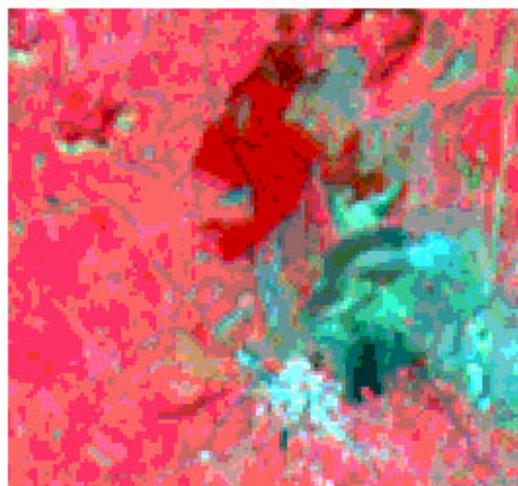


Abb. 5-2 Ausschnitt wie in Abb. 5-1 in der gleichen Kanalkombination nach Entfernung der Wasser- und Wolkenflächen. Es ist eine Verbesserung der Kontraste und der Informationsinhalte erkennbar.

5.1.2 KONTRAST-MANIPULATIONEN

In den Original-Images der einzelnen Spektralkanäle treten in der Regel nur die **DNs** eines begrenzten Bereiches gehäuft auf, während Teile der 256 möglichen Grauwerte gar nicht oder nur selten in Erscheinung treten. Durch die Verteilung der im Image vorhandenen **DNs** auf der gesamten Skala (*contrast stretching*) wird eine deutlich sichtbare Steigerung der Kontraste und damit eine Verbesserung der visuellen Interpretation unterschiedlicher Flächen erreicht. Weil zunächst wegen der relativ guten Geländekenntnis versucht werden sollte, einzelne Unterscheidungsflächen (Trainingsgebiete) für eine überwachte Klassifikation (siehe Kapitel 5.2.2) der Vegetation deutlicher darzustellen, wurden unterschiedliche Möglichkeiten der Kontraststreckung untersucht. Einige der vielen möglichen Verfahren (vgl. v.a. RICHARDS, S. 93 ff und LILLESAND/KIEFER, S. 543 ff) sollen hier am Beispiel des Image aus dem panchromatischen Kanal nach der Entfernung der Wasser- und Wolkenflächen beschrieben werden. In den folgenden Abbildungen werden die verschiedenen graphischen Darstellungen der Grauwert-Histogramme des Gesamtbildes nach erfolgter Streckung gezeigt. Von der Neuverteilung der Grauwerte und der Häufigkeitsdarstellung im Histogramm wurde in jedem Fall der Z-Wert (**DN**) = 0 ausgenommen, der die „no data“-Flächen des Bildhintergrundes repräsentiert. Zur Darstellung des jeweiligen Image dient ein Bildausschnitt der Umgebung des Dorfes San Andrés, wie schon in Abb. 5-1 und Abb. 5-2. Die Häuser des Dorfes stellen sich als hellgraue Pixel dar, bei der nördlich gelegenen dunkelgrauen Fläche handelt es sich um einen dichten Bestand aus angepflanzten *Pinus radiata*. Die umgebenden Wiesen (linke Bildhälfte) und die im Bild östlich und nordöstlich des Dorfes gelegenen fast vegetationslosen Vulkankegel sind in mittleren Grautönen abgebildet.

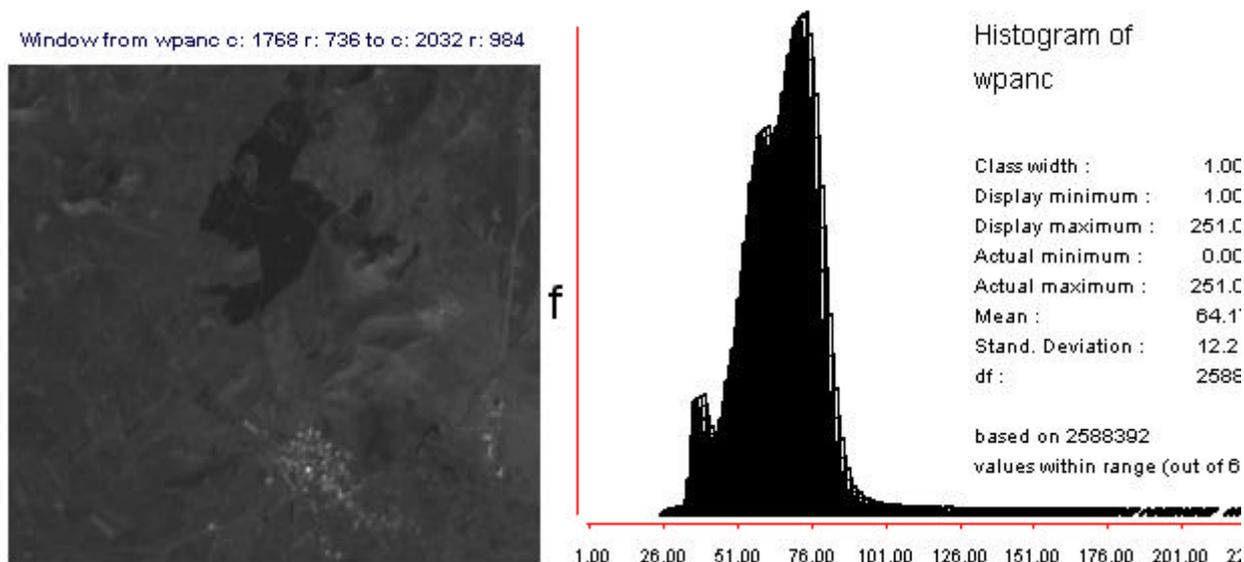


Abb. 5-3 Histogramm und Ausschnitt des panchromatischen Image nach Entfernung der Wasserflächen und der Wolken, auf deren Darstellung auch im Histogramm verzichtet wird. Die verbliebenen Grauwerte entsprechen den Originaldaten.

Die Verteilung aller im Original-Image auftretenden Grauwerte auf der Skala von 0-255 wird als lineare Streckung (*linear stretch*) bezeichnet. Weil bei dieser Art der Streckung, wie auch bei allen nachfolgenden, auf die Einbeziehung des Hintergrundwertes ($DN = 0$) verzichtet und die Originalwerte auf der Skala von 1-255 gestreckt wird, gilt für den *linear stretch* in diesem Fall der abgewandelte Algorithmus (nach LILLESAND/KIEFER, S. 547):

$$DN' = \left[\frac{DN - (MIN + 1)}{MAX - (MIN + 1)} \right] 255$$

wobei

DN' = dem Pixel im Ausgabe-Image zugewiesener Grauwert,

DN = Originalwert des Pixels im Eingabe-Image,

MIN = Minimale DN des Eingabe-Image, der im Ausgabe-Image der Wert 1 zugewiesen wird (im Beispiel die DN 24),

MAX = Maximale DN des Eingabe-Image, der im Ausgabe-Image der Wert 255 zugewiesen wird (in diesem Fall die DN 251).

Weil in der Grauwertskala der Wert 0 als schwarz und der Wert 255 als weiß dargestellt werden, erscheinen dunkle Pixel durch die lineare Streckung dunkler und helle Pixel heller. Die restlichen Grauwerte werden je nach ihrem Auftreten auf der Skala verteilt. Im Original-Image nicht vorhandene DN s werden auch im Histogramm des Ausgabe-Image als „Lücken“ dargestellt. Da sich der größte Anteil der Grauwerte aber noch immer im Bereich der DN 's von 10-66 konzentriert, ist nur eine unbedeutend verbesserte Darstellung möglich (Abb. 5-4).

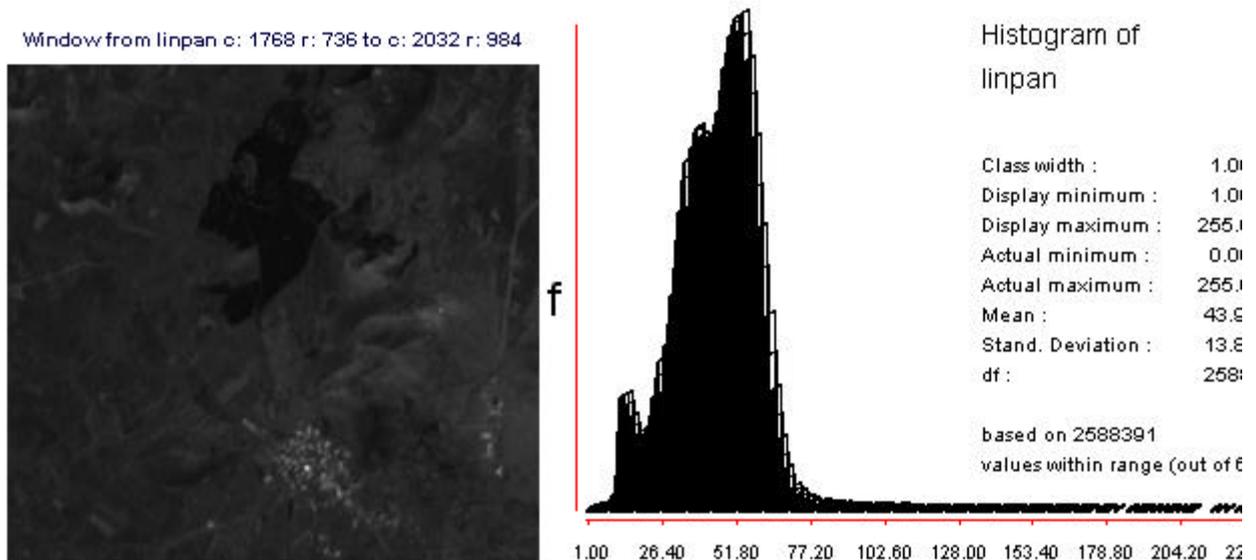


Abb. 5-4 Ausschnitt und Histogramm, unter Verzicht auf $DN=0$, nach linearer Streckung.

Eine deutliche Verbesserung der Darstellung wird durch eine lineare Streckung mit einem bestimmten Sättigungswert (*linear stretch with saturation*) an den beiden Enden der Grauskala erreicht (Abb. 5-5). Damit wird bestimmt, auf wieviel Prozent der Pixel im dunklen bzw. hellen Darstellungsbereich verzichtet werden kann. Diese Pixel werden dann den minimalen resp. maximalen DN 's des Ausgabe-Image zugeordnet, während die übrigen Pixel nach der o.a. Formel auf der Grauwertskala angeordnet werden.

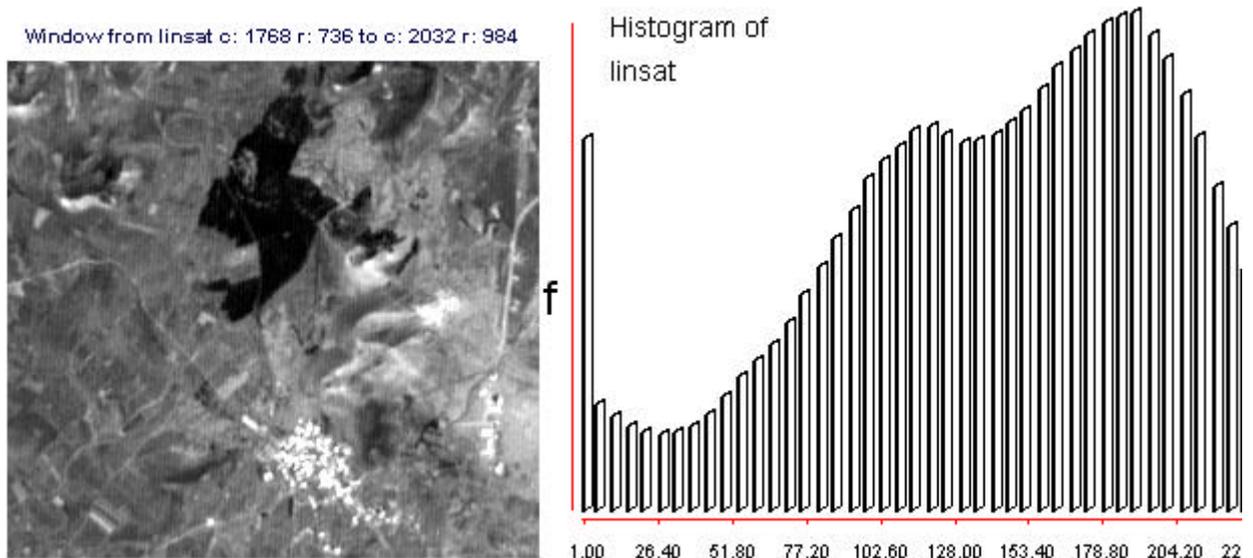


Abb. 5-5 Ausschnitt und Histogramm nach linearer Streckung mit 2.5 % Sättigung an jedem Ende der Grauskala.

Der Nachteil linearer Streckungen besteht darin, daß alle DN s des Eingabe-Image ohne Rücksicht auf die Häufigkeit ihres Auftretens gleichwertig auf der Grauwertskala verteilt werden (vgl. dazu LILLESAND/KIEFER, S. 549, RICHARDS, S. 97). Dies kann zu geringeren Differenzierungsmöglichkeiten der Flächen beim *linear stretch* oder, wegen der Konzentration größerer Datenmengen an den Enden der Grauskala beim *linear stretch with saturation*, zu

Verlusten der Unterscheidungsmöglichkeiten in hellen oder dunklen Bildbereichen führen. Durch eine der Häufigkeit der **DN** angepaßten Streckung (*histrogram equalization*) wird dieser Effekt vermieden. *In this approach, image values are assigned to the display levels on the basis of their frequency of occurrence.* (LILLESAND/KIEFER, S. 549). In diesem Fall wird den **DNs** von 24-50 nur der schmale Skalenraum von **DN's** 1-25 zur Verfügung gestellt, den **DN** von 90-251 der Bereich von **DN's** 250-255, während die größte Menge der **DNs** (50-100) über die verbleibende Skalenbreite gestreckt wird (siehe Abb. 5-6).

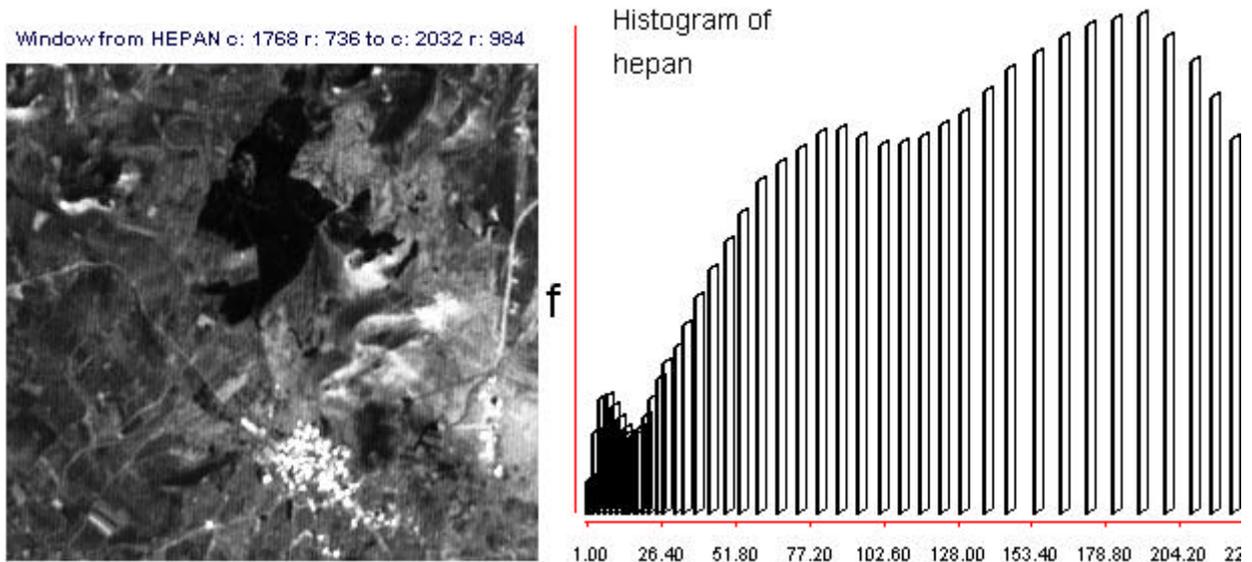


Abb. 5-6 Ausschnitt und Histogramm nach Streckung mit *histrogram equalization*.

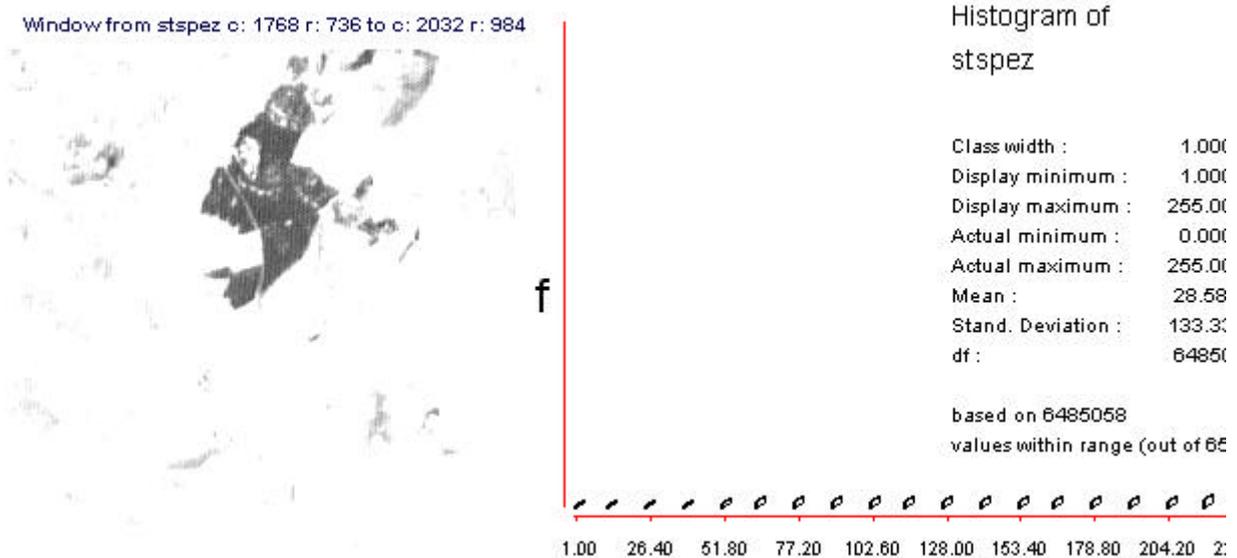


Abb. 5-7 Ausschnitt und Histogramm nach Streckung der **DNs** 30-50.

Zur besseren Abstufung der Grauwerte in bestimmten, homogen erscheinenden Flächen, können einzelne Bereiche der **DNs** gesondert gestreckt werden. Dies wurde versucht um eine genauere Differenzierung der Waldflächen zu erreichen, deren Z-Werte im panchro-

matischen Image auf die **DNs** von ca. 30-50 beschränkt sind. Nach einer Streckung dieser Grauwerte sind im Ausgabe-Image einige vorher nicht sichtbare Unterscheidungen in den Waldflächen dargestellt (Abb. 5-7).

In der Literatur wird wiederholt darauf hingewiesen *that stretched images are for display only. Because they have been manipulated, they are not reliable for analysis. Use only raw data for analysis unless you have a clear reason for using stretched data.* (EASTMAN 1997a, S. 85). In der Regel sollten diese Images nur zur visuellen Interpretation und zur besseren Unterscheidung beim digitalisieren von Trainingsgebieten dienen. Dennoch ist möglich, einzelne Images nach der Kontraststreckung mit Original-Images zu kombinieren oder alle für das *composit* verwendeten Bilder vor der Operation zu strecken (vgl. dazu auch Kapitel 5.1.4). Allerdings müssen die Streckungen sehr kontrolliert durchgeführt und visuell auch im *composit* überprüft werden, um Informationsverluste zu vermeiden. Welche Bereiche der Grauwertskala in welchem Spektralkanal vor der weiteren Analyse gestreckt werden sollten, richtet sich nach der Intention der Analyse. So können z.B. Waldflächen genauer unterschieden werden, wenn nur die Skalenbereiche aller Kanäle gestreckt werden, die entsprechende Spektralinformationen enthalten (Vgl. dazu Abb. 5-7).

Im Rahmen dieser Arbeit ließen sich auch nach allen Versuchen der Kontraststreckung¹ und trotz der Geländekenntnis nur wenige Gruppen der Bodenbedeckung (Wald, Wiese, Lava ohne Bewuchs, Siedlung) eindeutig identifizieren. Auch andere Methoden, welche der Verbesserung der Darstellbarkeit dienen (z.B. Einbeziehung gefilterter Datensätze², oder durch die Ermittlung eines Vegetationsindex erzeugter Images in ein *color composit*), führten nicht zum Erfolg. Die Klassifizierung wurde deshalb später halbautomatisch mit dem **IDRISI**-Modul ISOCLUST durchgeführt (Kapitel 5.2.3). Das für diesen Vorgang benötigte *color composit* zeigte allerdings in allen Gebieten die besten Differenzierungen nachdem die Original-Images **aller** zur Verfügung stehenden Spektralkanäle gestreckt und nach einer Farbraumtransformation (Kapitel 5.1.3.1) unter Einbeziehung des panchromatischen Bildes neu kombiniert wurden. Bei der Erzeugung des Falschfarben-*composit* wurde später eine lineare Streckung mit 1 % Sättigung gewählt. Abb. 5-23 zeigt eine Gesamtansicht dieses Image.

Die Streckung der einzelnen Bilder erfolgte nicht automatisch, sondern auf der Grundlage der tabellarischen Histogramme jedes Image. In jedem Bild wurden an beiden Enden der Grauwertskala 0.1 % der **DNs** von der Streckung ausgenommen und die restlichen Werte über die Skala von 1-255 linear gestreckt. Die so veränderten Datensätze waren die Grundlage für alle weiteren Operationen.

¹ Ein weiteres Verfahren der Kontrastmanipulation ist das sogen. *density* bzw. *level slicing*“ (vgl. Roberts, S. 106 ff und LILLESAND/KIEFER, S.543), auf das hier nicht näher eingegangen wird. Eine weitere interessante Möglichkeit der Kontrastverbesserung, *direct decorrelation stretch (DDS)* wird von JIAN GUO/MOORE erörtert.

Im einzelnen wurden gestreckt:

Kanal	Streckung von bis
xs1 (Grün)	29 - 160
xs2 (Rot)	19 - 131
xs3 (nIR)	3 - 164
panchromatisch	28 - 158

Das Beispiel aus dem panchromatischen Kanal ist in Abb. 5-8 dargestellt.

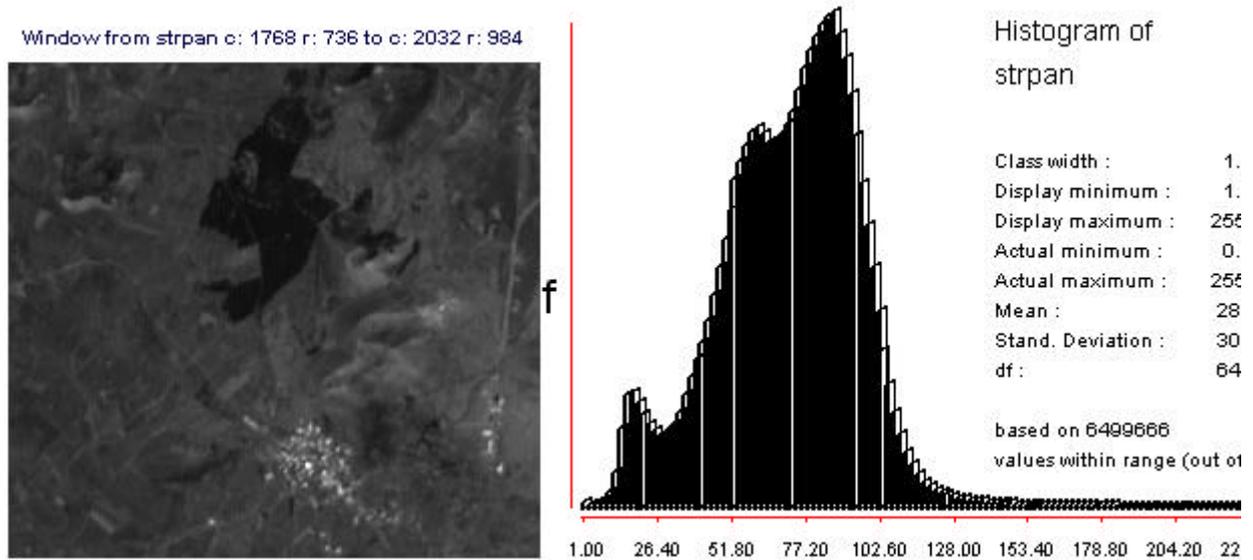


Abb. 5-8 Ausschnitt und Histogramm nach linearer Streckung der DN's 28-158.

Abb. 5-9 zeigt einen Ausschnitt aus einem *color composit* der gestreckten Multispektralkanäle. Im Vergleich zu Abb. 5-2 ist eine weitere leichte Verbesserung der Kontraste erkennbar. Die besten Differenzierungsmöglichkeiten zeigen sich aber, wie bereits erwähnt, erst nach Einführung der Spektralinformationen aus dem panchromatischen Bild.

Window from scomp c: 1768 r: 736 to c: 2032 r: 984

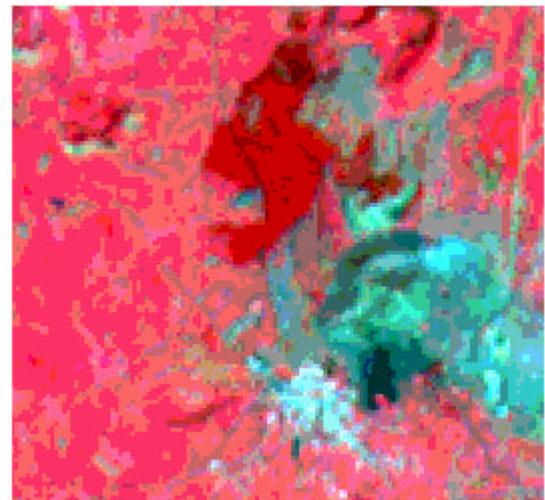


Abb. 5-9 Falschfarben-Composit aus der Kombination der gestreckten Kanäle xs1,xs2,xs3.

² Zu Filterverfahren z.B. RICHARDS S. 113 ff.

5.1.3 TRANSFORMATION DER SPEKTRALINFORMATIONEN

Für die optische Verbesserung von Fernerkundungsdaten stehen neben Kontrastmanipulationen der einzelnen Spektralkanäle eine Anzahl weiterer Verfahren zur Verfügung, mit denen die Spektralinformationen mehrerer Datensätze neu berechnet werden können. *The multispectral or vector character of most remote sensing image data renders it amenable to spectral transformations that generate new sets of image components or bands* (ROBERTS, S. 133). Anders als bei den Kontraststreckungen werden die Dateninhalte bei diesen Operationen nicht verändert, die Ausgabe-Images können daher auch zu Analysen herangezogen werden.

Eines dieser Verfahren ist die bereits erwähnte Hauptkomponentenanalyse (*principal component analysis*, **PCA**), deren Prinzip darin besteht, die *Grauwerte mehrerer Spektralkanäle so zu transformieren, daß eine geringe Zahl von Kanälen, nämlich die Hauptkomponenten, für die Bilddarstellung oder Klassifizierung ausreichen* (LÖFFLER, S. 71, vgl. dazu auch LILLESAND/KIEFER, S. 572 ff und RICHARDS, S. 133 ff). Dieses Verfahren wird in der Regel angewendet, wenn Fernerkundungsdaten aus mehreren ähnlichen Spektralbereichen vorliegen. Dies ist z.B. bei Landsat-**TM**-Daten gegeben oder wenn Datensätze verschiedener Aufnahmesysteme miteinander kombiniert werden sollen. Häufig ähneln sich dabei die Informationen der einzelnen Spektralkanäle, die Informationen aus dem **nIR** (Kanal 4) sind oft nahezu identisch mit den Aussagen der Kanäle des mittleren **InfraRot** (Kanäle 5 und 7). Bei der **PCA** werden die Daten aus bis zu sieben Kanälen nach ihren Korrelationen so geordnet, daß die ersten drei der erzeugten Images über 95 % aller Informationen enthalten. Weil hier nur **SPOT**-Daten aus drei Spektralbereichen verarbeitet wurden, kam dieses Verfahren nicht zur Anwendung.

5.1.3.1 FARBRAUMTRANSFORMATION

Typischerweise werden digitale Bilddatensätze in einer additiven Farbmischung, bestehend aus den Farben Rot, Grün und Blau, dargestellt (**RGB**-Modell). Für jede dieser Farben stehen auf einem Monitor mit einer 8-bit-Farbtiefe 256 Abstufungsmöglichkeiten (von 0-255) zur Verfügung, womit 256^3 (d.h. 16,7 Mio.) Farbschattierungen dargestellt werden können, von denen jede eine festgelegte Koordinatenposition in einem dreidimensionalen Körper einnimmt. Die in den verschiedenen Kanälen wiedergegebenen Reflexionen des Lichtes werden von den Bildverarbeitungs-Systemen deshalb ebenfalls in **DNs** von 0-255 ausgedrückt. Weil bei einem *color composit* die Informationen von drei Spektralkanälen nach einem bestimmten Algorithmus gemischt werden (vgl. Kapitel. 5.1.4), ist damit eine gezielte Darstellung der erzeugten Bilder möglich. Werden die Kanäle des sichtbaren Blau, Grün und Rot der entsprechenden Farbkomponente zugewiesen (für Landsat-**TM** Daten gilt: **Rot**=Kanal 3, **Grün**=Kanal 2, **Blau**=Kanal 1), wird das Ausgabe-Image in Normalfarben dargestellt, ähnlich einer Farbfotografie. Wird die Zuordnung **Rot**=**nIR**-Kanal, **Grün**=roter Kanal, **Blau**=grüner

Kanal vorgenommen, ist das Ergebnis ein Falschfarbenbild, das einer Infrarotaufnahme ähnelt. Andere Kanalkombinationen ergeben eher zufällige Farbmuster, sie können aber zur Hervorhebung bestimmter Informationen dienen (vgl. LÖFFLER, Farbtafel 1 neben S. 64).

Eine andere Möglichkeit der Farbdarstellung für digitale Bilddaten ist das mehrdimensionale hexagonale **HLS-** (oder **IHS-**) Modell, mit dem unendlich viele Farben dargestellt werden können. Bei diesem Modell erfolgt keine additive Farbmischung, sondern es werden die Farbschattierung (*Hue*) von Rot, Gelb, Grün, Cyan, Blau, Magenta, die Intensität (*Lightening* bzw. *Intesity*) und die Farbsättigung (*Saturation*) dargestellt (vgl. dazu LILLESAND/KIEFER, S. 579 ff, LÖFFLER, S.71). Besser zu differenzierende *composit*-Darstellungen sind möglich, wenn vor der Bildkombination folgende Schritte ausgeführt werden:

- die Komponenten von **RGB** werden in den Farbraum **HLS** transformiert,
- der Kontrast eines der neu erzeugten Bilder (z.B. des Bildes, das die Farbsättigung der Komponenten wiedergeben soll) wird verstärkt,
- Rücktransformation der Datensätze nach **RGB**.

Für die Zusammenführung (*merging*) der Informationen des geometrisch höher aufgelösten panchromatischen **SPOT**-Kanals mit den multispektralen Datensätzen ist die **RGB/HLS**-Transformation erforderlich. Zunächst werden die multispektralen Daten von **RGB** nach **HLS** umgewandelt, bei der Rücktransformation werden die Schwarzweiß-Informationen des panchromatischen Bildes genutzt, indem sie als Intensitätskomponente (**L**) eingeführt werden (Arbeitsschritte Abb. 5-10).

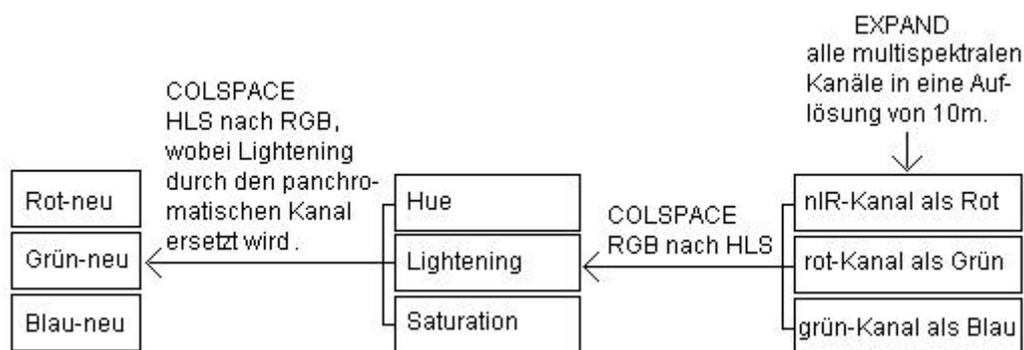


Abb. 5-10 Verbindung von panchromatischen und multispektralen Informationen im IDRISI-System.

Das auf diesem Wege erzeugte Image mit Informationen aus dem Spektrum des sichtbaren Rot wurde als fester Bestandteil in die weiteren Arbeitsschritte einbezogen und ersetzte in den meisten Fällen das Original-Image aus diesem Spektralbereich. Die größere Informationsdichte wird im Vergleich des Originals (Abb. 5-11) mit dem neuen Bild (Abb. 5-12) sichtbar.

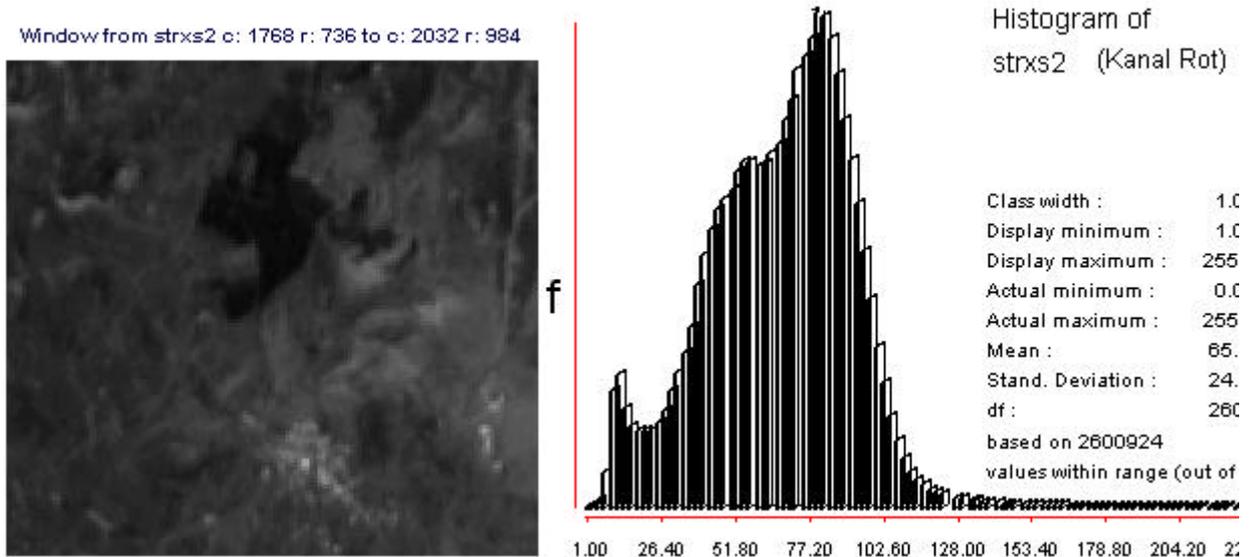


Abb. 5-11 Ausschnitt und Histogramm des gestreckten Original-Image aus dem Spektrum des sichtbaren Rot.

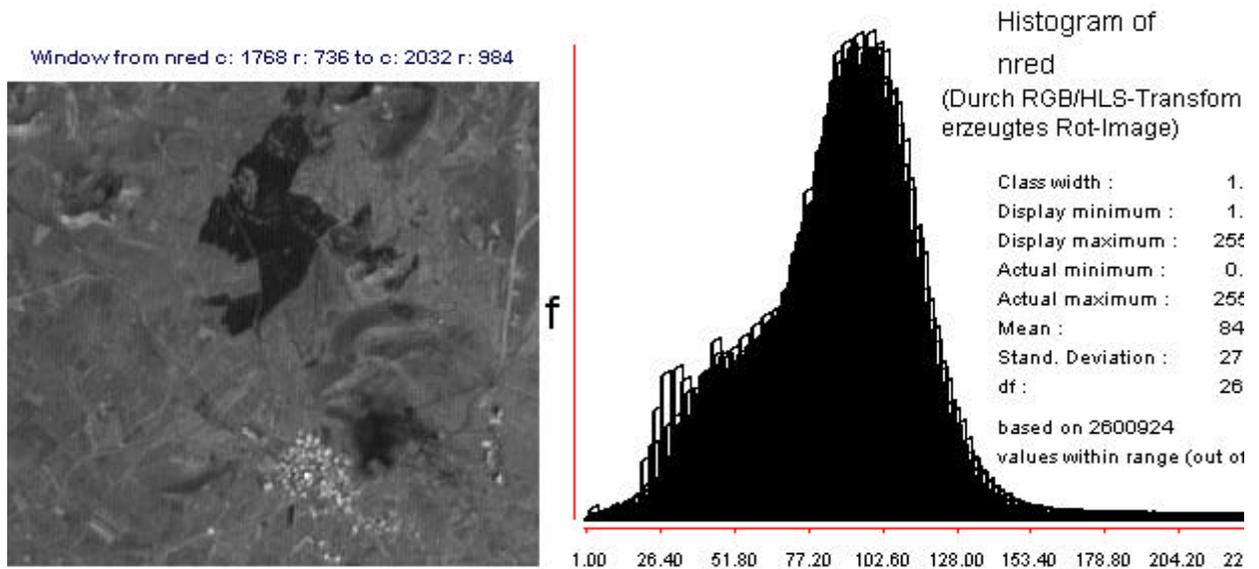


Abb. 5-12 Ausschnitt und Histogramm des durch RGB/HLS-Transformation neu gewonnenen Image mit Informationen aus dem Spektrum des sichtbaren Rot.

5.1.3.2 VEGETATIONS-INDICES

Arithmetische Transformationen werden auch angewandt, um das Vorhandensein und den Grad grüner Vegetation darzustellen, wobei vor allem ihre in Kapitel 3.2.2.2 beschriebenen unterschiedlichen Reflexionseigenschaften im sichtbaren Rot und im nahen Infrarot genutzt werden (vgl. dazu auch Abb. 5-13). Die Verfahren reichen von einer einfachen Ratio-Bildung zwischen diesen beiden Kanälen bis zu aufwendigeren Berechnungen, in denen auch die Reflexion unbewachsener Böden einbezogen wird (vgl. dazu v.a. ACHEN 1993 und EASTMAN 1997b, S. 13-1 ff).

Ratio-Images können ebenfalls benutzt werden, um die Reflexionsunterschiede gleicher Arten von Bodenbedeckung zu minimieren, die insbesondere durch starke Reliefunterschiede und die daraus resultierende Beschattung von Teilen des Geländes bei gleichzeitig niedrigem Sonnenstand zum Aufnahmezeitpunkt hervorgerufen werden. *A major advantage of ratio images is that they convey the spectral or color characteristics of image features, regardless of variations in scene illumination conditions. (...) The DN's observed for each cover type are substantially lower in the shadowed area than in the sunlit area. However, the ratio values are nearly identical, irrespective the illumination conditions. Hence, a ratioed image of the scene effectively compensates for the brightness variation caused by the varying topography and emphasizes the color content of the data* (LILLESAND/KIEFER, S. 567). Andererseits ist die Einbeziehung eines Ratio-Image (Rot/nIR) in ein *composit* nicht unproblematisch. (...) *While it normalizes the shadowed slopes and improve discrimination of landcover in these areas, the method tends to exaggerate biomass everywhere else. In the process, we can loose discrimination between different land cover types as the range of reflectances in the more illuminated areas decreases* (MILETTE, S. 7). Um die negativen Einflüsse von beschatteten Gebieten auszuschließen, wird empfohlen, in stark reliefierten Gebieten auf Daten zu verzichten, zu deren Aufnahmezeitpunkt der Sonneneinfallswinkel unter 45° liegt (MILETTE, S. 1). Durch die starken Reliefänderungen sind im Bereich des Risco/Golfo in den Datensätzen zwar Schatteneinflüsse zu beobachten, weil aber der Einfallswinkel der Sonne zum Aufnahmezeitpunkt relativ hoch war (55.9°, vgl. Kapitel 4.2.2.2), wurde auf die Einbeziehung von Ratio-Images in die Analyse verzichtet.

Gleichwohl war geplant, Dateninformationen über Deckungsgrad und „greenness“ der Vegetation in die Bewertung über die Richtigkeit der späteren Vegetationsklassen einzubeziehen. Deshalb wurden einige der vorgeschlagenen Vegetations-Indices (VIs) erstellt und miteinander verglichen. Weil VIs von vielen Datenanalytikern immer wieder neu entwickelt und an spezielle Untersuchungsgebiete angepaßt wurden, gibt es eine Vielzahl der Ermittlungsmethoden. Allein das IDRISI-Modul VEGINDEX bietet 19 verschiedene Modelle, deren Grundlage immer die Images aus dem Spektrum des roten Lichtes und des nahen Infrarot sind. Das Original-Image des roten Kanals (xs2) wurde hier, der größeren räumlichen und spektralen Auflösung wegen, durch das neu gewonnene nred-Image ersetzt. Bei der Farbraumtransformation wird kein Image mit Informationen des niR-Spektrums gewonnen, weshalb nur das Original-Image (Kanal xs3) benutzt werden kann. Die räumliche Auflösung der Pixel (20 m * 20 m) muß an die Auflösung des erzeugten Image nred (10 m * 10 m) angepaßt werden.

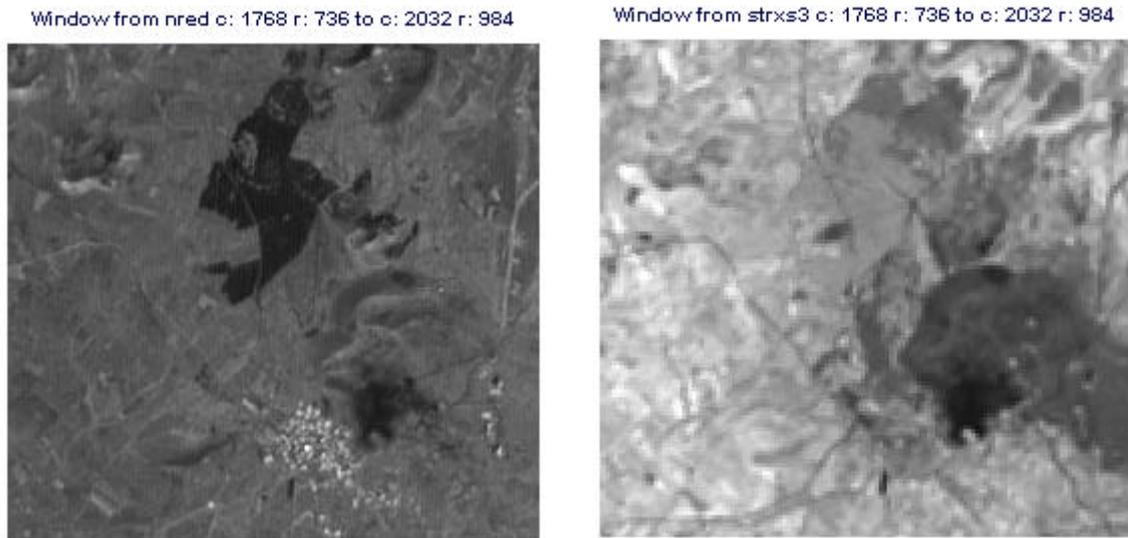


Abb. 5-13 Links: Ausschnitt aus dem neu gewonnenen **nred**, rechts: aus dem **nIR**-Kanal (xs3). In beiden Ausschnitten werden stark reflektierende Flächen durch helle, schwach reflektierende durch dunkle Bildpixel wiedergegeben (hohe resp. niedrige **DN**s). Die Waldfläche läßt erkennen, daß sie im Vergleich zu den Wiesen (linke Bildhälften) mehr sichtbares Rot absorbiert und gleichzeitig weniger **nIR** reflektiert. Der vegetationslose Aschekegel des Chamuscada rechts oberhalb des Dorfes ist in beiden Kanälen durch niedrige Grauwerte gekennzeichnet.

Nachfolgend werden drei Modelle zur Ermittlung vorhandener grüner Vegetation dargestellt und erläutert. Die Darstellung der Bildausschnitte erfolgt mit einer bipolaren kontinuierlichen Farbpalette. Vorhandene Vegetation wird in abnehmender Menge von Dunkel- bis Hellgrün dargestellt, vegetationsarme Flächen werden in zunehmender Menge (Zunahme der Vegetationsfreiheit) von Hellgrün bis Rotbraun dargestellt. Dichter Bewuchs aus Pflanzen mit chlorophyllhaltigen Blättern entspräche in der Darstellung daher einem dunklen Grün, vegetationsfreie Fläche dem Rotbraun.

5.1.3.2.1 NDVI (NORMALIZED DIFFERENCE VEGETATION INDEX)

Die Ermittlung des **NDVI** gehört, wie die einfache Ratio-Bildung, zu den sogenannten *slope-based VIs*. Diese Verfahren werden „nur“ zweidimensional durch das Verhältnis der beiden Kanäle berechnet. Das am häufigsten angewandte und in vegetationsreichen Gebieten häufig ausreichende Verfahren ist die Ermittlung des **NDVI** mit folgender Formel:

$$\mathbf{NDVI} = \frac{(nIR - Rot)}{(nIR + Rot)}$$

Errechnet werden damit Werte in reellen Zahlen von -1 (vegetationslos) bis +1.

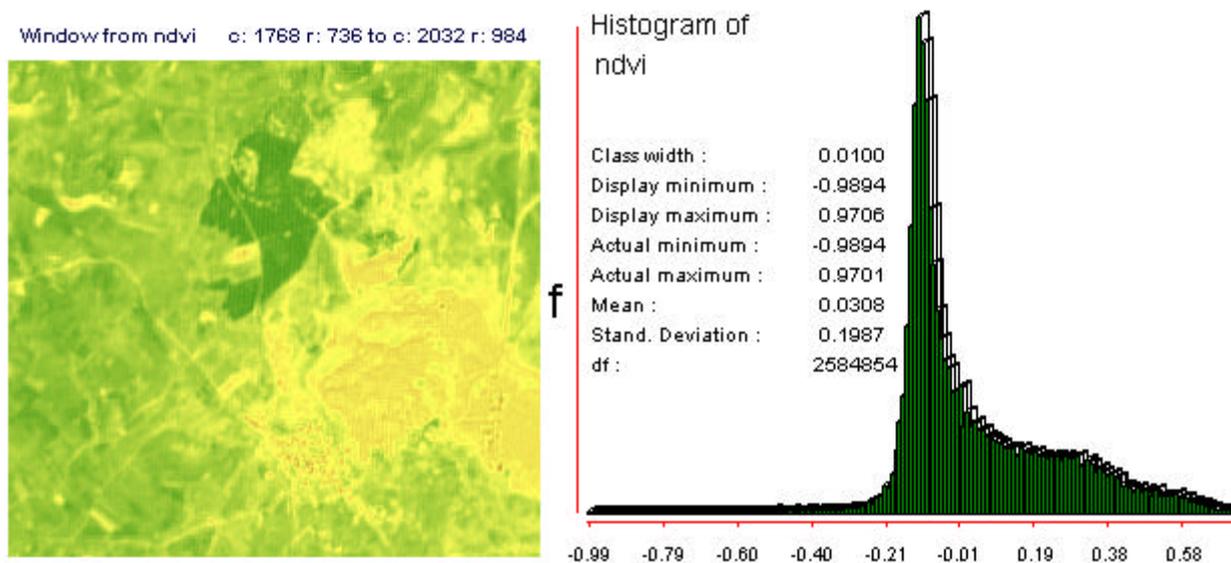


Abb. 5-14 Ausschnitt und Histogramm aus dem **NDVI**-Image.

In Abb. 5-14 wird bereits deutlich, daß mit dem **NDVI** in bestimmten Regionen die Vitalität der Vegetation nur unzureichend beschrieben wird. Die Pinienwaldfläche, in deren Nadeln vitale Prozesse weit langsamer verlaufen als in den schnellwüchsigen Gräsern und Kräutern der umgebenden Wiesen, werden mit einem höheren **VI** wiedergegeben, während der völlig vegetationslose Aschekegel überwiegend Werte nahe Null aufweist. Zur Vermeidung solcher Phänomene wurde eine Reihe von Modifikationen entwickelt. Durch einige davon sollen u.a. die Negativwerte entfallen oder die *overestimation of greenness* (EASTMAN 1997b, S. 13-3) reduziert werden. Der Eindruck von „viel Grün“ entsteht zwangsläufig in Gebieten, in denen große Anteile vegetationsloser oder vegetationsarmer Flächen vorhanden sind und in denen die Reflexionseigenschaften der unterschiedlichen Böden eine Rolle spielen³.

5.1.3.2.2 **DVI (DIFFERENCE VEGETATION INDEX)**

Der **DVI** gehört zu den sogenannten *distance based VIs*, bei denen als dritte Dimension der Berechnung die Bodenreflexion einbezogen wird. *This group of vegetation indices is based on the Perpendicular Vegetation Index (PVI). The main objective of these VIs is to cancel the effect of the background soil brightness and generate an image that only highlights the vegetation signal. The procedure is based on the soil line concept, a regression line that is obtained through linear regression of the near infrared band against the red band* (EASTMAN 1997b, S. 13-4 ff, dort auch die Formeln zur Berechnung dieser Indices). Die Arbeitsschritte bis zur Fertigstellung des **DVI**-Image sind in Abb. 5-15 dargestellt.

³ So absorbieren z.B. ansonsten gleiche Bodenarten mit höherem Feuchtegrad mehr Strahlung im roten Spektrum, was bei gleicher Vegetationsart und -dichte zu unterschiedlichen **VI** führt. Vegetationsfreie Böden, deren Reflexionseigenschaften sowohl im roten wie im niR-Spektrum in niedrigen Bereichen liegen, werden sehr nahe an Null wiedergegeben und täuschen damit vorhandene Vegetation vor.

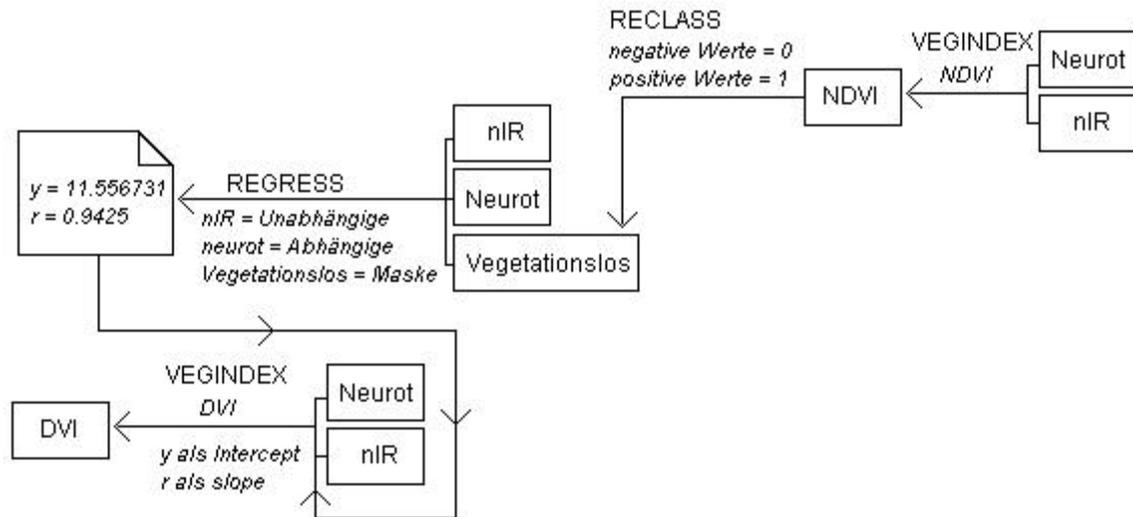


Abb. 5-15 Arbeitsschritte zur Berechnung des Vektors (r), der die Bodeninformation darstellt (*soil line* bzw. *slope*), und dem Schnittpunkt, an dem jede Pixelinformation diese Linie senkrecht schneidet (y), bis zur Fertigstellung des **DVI**-Image.

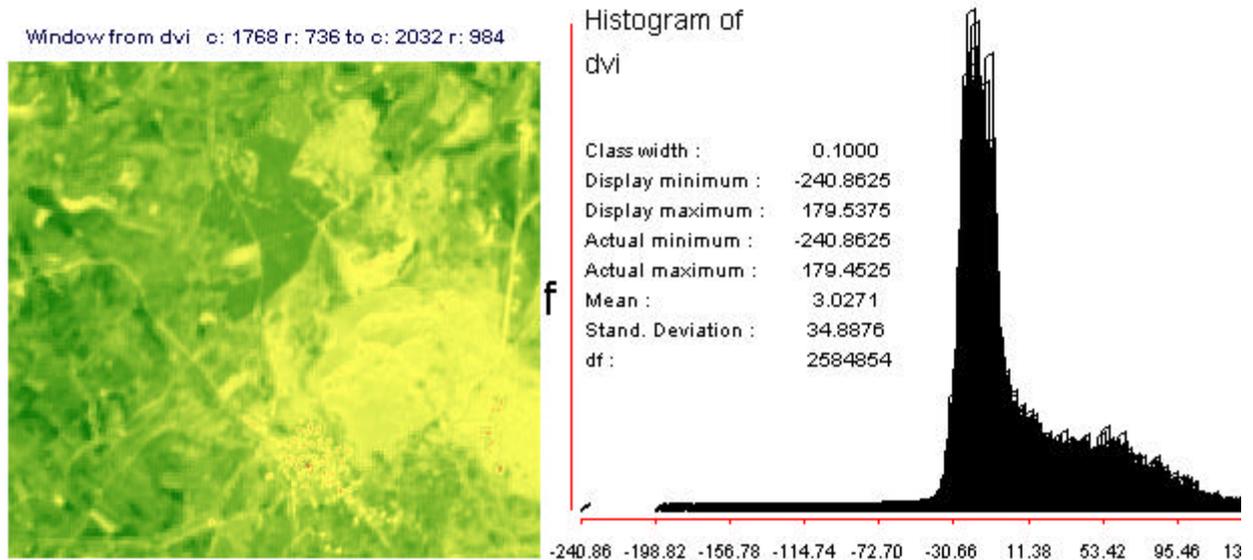


Abb. 5-16 Ausschnitt und Histogramm des **DVI**-Image.

Der höhere Anteil an chlorophyllreicher Blattmasse dichter Wiesen im Vergleich zum Pinienvald wird deutlicher, auf dem Vulkankegel wird jedoch noch immer eine zu große Menge Vegetationsbedeckung simuliert (Abb. 5-16). Alle erzeugten Bilddarstellungen der Vegetations-Indices entsprachen nur unzureichend den bisherigen Wahrnehmungen der Autorin vom Anteil grüner Vegetation auf der Insel. Am ehesten eignete sich zum visuellen Vergleich die Darstellung des **DVI**-Image. Auch später im Gelände konnte der **DVI** als Vergleichsmedium eingesetzt werden. Als Beispiel seien hier die Weidegebiete in der Umgebung des Dorfes San Andrés angeführt. Sie unterscheiden sich weder in ihrer Vegetationszusammensetzung noch ist der Boden verschieden, der einzige Unterschied ist die Dichte des Bewuchses durch eine leichte Änderung der Bodenfeuchte. Mit der automatischen Klassifizierung wurden hier zwei getrennte Klassen ausgewiesen (Klasse 15 und Klasse

22). Auf den im Gelände als „dichte Wiese“ identifizierten Flächen lagen die Werte des **DVI** immer höher als auf den weniger dicht bewachsenen Arealen. Dessenungeachtet sollte eine Darstellung erreicht werden, mit der die bei der Geländearbeit festgestellten Unterschiede besser zur Geltung gebracht werden konnten. Nachträglich wurde deshalb mit einem Transformations-Modell experimentiert, dessen Anwendung bei der Verarbeitung von **SPOT**-Datensätzen bisher in der Literatur, die der Verfasserin zugänglich war, nicht beschrieben wurde.

5.1.3.2.3 THE GREEN VEGETATION INDEX OF TASSELED CAP (GVI)

Bei diesem Verfahren handelt es sich um eine orthogonale Transformation auf der Basis mehrerer Spektralkanäle, ähnlich der **PCA**.

The Green Vegetation Index of the Tasseled Cap (GVI) is the second of the four new bands that Kauth and Thomas (1976) extracted from raw MSS Images. The GVI provides global coefficients that are used to weight the original MSS digital counts to generate the new transformed bands. (...) The expression of the green vegetation index band, GVI is written:

$$GVI = [(-0.386MSS4)+(-0.562MSS5)+(0.600MSS6)+(0.491MSS7)]$$

The negative weights of the GVI on the visible bands tend to minimize the effects of the background soil while its positive weights on the near infrared bands emphasize a green vegetation signal (EASTMAN 1997b, S. 13-11).

Die Daten von Landsat-**MSS** stehen in vier Spektralkanälen mit einer räumlichen Auflösung von ca. 80 m zur Verfügung. Die spektralen Informationen im Vergleich zu den **SPOT**-Daten sind in Abb. 5-17 dargestellt.

The so called „tasseled cap“ transformation (...) is a means for highlighting the most important (spectrally observable) phenomena of crop development in a way that allows discrimination of specific crops, and crops from other vegetative cover, in Landsat multitemporal imagery. Its basis lies in an observation of crop trajectories in band 6 versus band 5, and band 6 versus band 4 subspaces (RICHARDS, S. 148, auf den folgenden Seiten auch die ausführliche Beschreibung des Verfahrens). Mit der Transformation werden aus den vier Original-Images vier neue Komponenten in folgender Reihenfolge erstellt:

- Das Image *soil brightness* erklärt die Unterschiede der Bodenstrahlung,
- mit *green-stuff* wird der Grad chlorophyllhaltiger Vegetation wiedergegeben,
- absterbende bzw. trockene Vegetation wird durch das Image *yellow-stuff* dargestellt,
- eine vierte Komponente *no such* enthält die Informationen, die weder mit der Bodenstrahlung noch mit dem Anteil grüner resp. vertrocknender Vegetation im Zusammenhang stehen. Sie repräsentiert z.B. durch Wasserflächen oder Dunst hervorgerufene Reflexion.

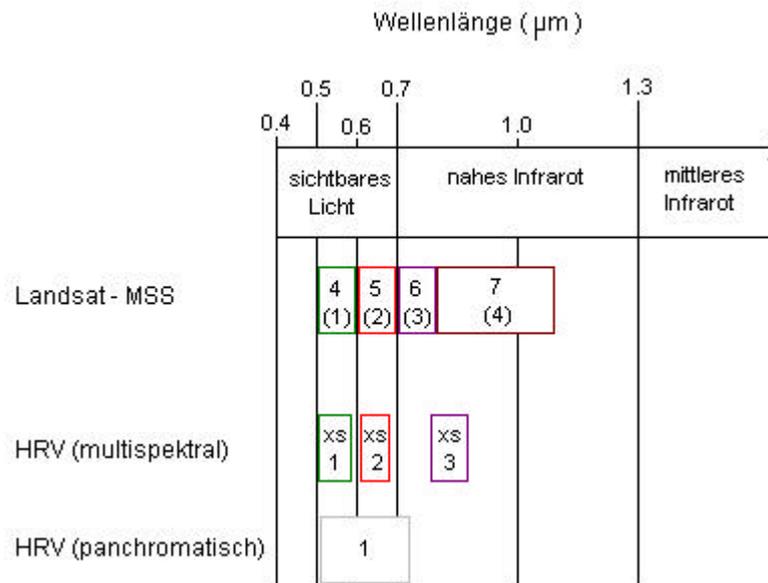


Abb. 5-17 Vergleich der spektralen Aufnahmebereiche zwischen Landsat-MSS und SPOT. Die Spektralbereiche umfassen: MSS4 0.5-0.6, MSS5 0.6-0.7, MSS6 0.7-0.8, MSS7 0.8-1.1, SPOT xs1 0.5-0.59, xs2 0.61-0.68, xs3 0.79-0.89, panchromatisch 0.51-0.73.

In Abb. 5-17 ist erkennbar, daß sich die spektrale Auflösung der Kanäle aus dem grünen und dem roten Spektrum von Landsat-MSS und SPOT relativ ähnlich sind. SPOT-Kanal xs3 enthält den Anfangsbereich der Strahlung aus MSS7, während der panchromatische Kanal auch Teile von MSS6 wiedergibt. Im IDRISI-System wird die Transformation mit dem Modul TASSCAP berechnet. Die Arbeitsroutine erfordert die Eingabe der MSS-Bänder in aufsteigender Reihenfolge. Weil im panchromatischen SPOT-Kanal auch geringe Anteile des nahen Infrarot vorhanden sind (Teile von MSS6), wurde mit den vorhandenen Daten zunächst die Eingabereihenfolge xs1, xs2, panchromatisch, xs3 ausprobiert. Das daraus entstandene Bild des *green stuff* stellte im Vergleich zu den bis dahin durchgeführten Versuchen eine entscheidende optische Verbesserung dar (Abb. 5-18). Vegetationslose Gebiete wurden deutlich verschieden von der Vegetation dargestellt, gleichzeitig wurden jedoch noch einzelne, als vegetationslos bekannte Areale als begrünt ausgewiesen. In einem Vergleich der Mittelwerte (Abb. 5-19) verschiedener Vegetationsklassen auf der Grauwertskala unterschiedlicher VIs (Erläuterungen dazu in Kapitel 7.1) wurde dann auch deutlich, daß auch in völlig vegetationslosen Gebieten noch grüne Vegetation vorgetäuscht wurde.

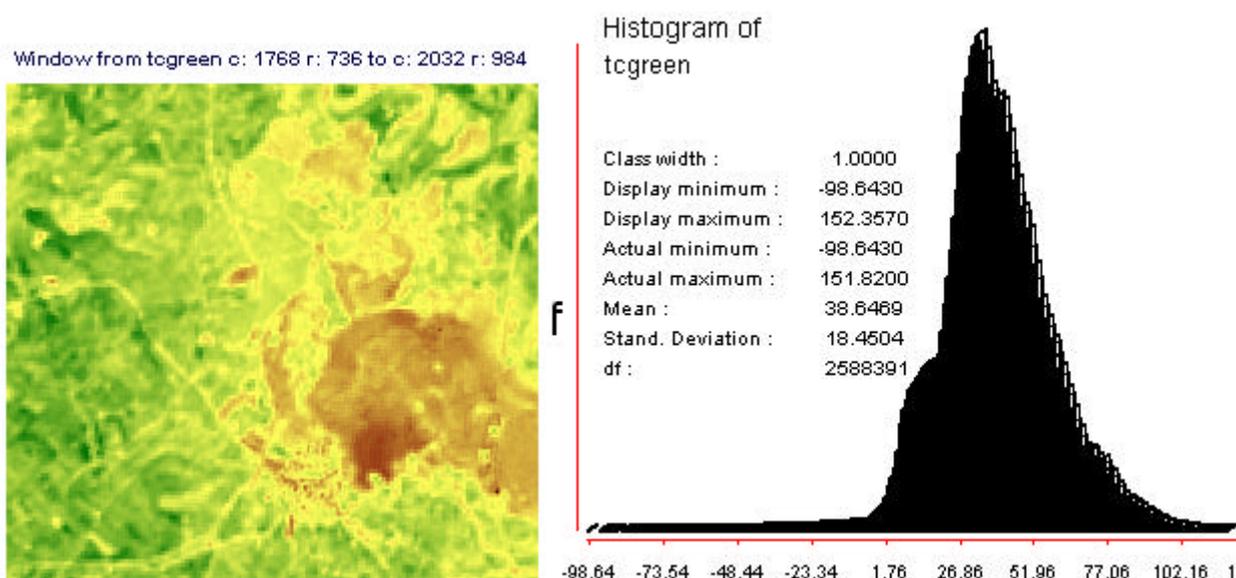


Abb. 5-18 Ausschnitt aus dem *green stuff*-Image der *tasseled cap*-Transformation bei einer Berechnungsreihenfolge Band1 = HRV xs1, Band2 = HRV xs2, Band3 = HRV panchromatisch, Band4 = HRV xs3.

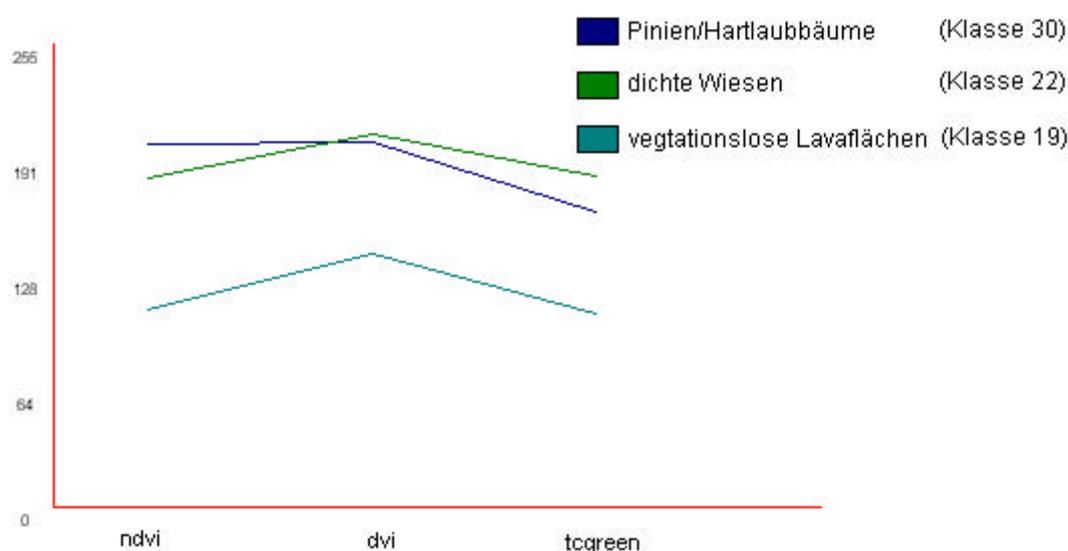


Abb. 5-19 Vergleich der Mittelwerte von **NDVI**, **DVI** und des **GVI** aus der in Abb. 5-18 angegebenen Eingabefolge. Die ermittelten reellen Werte der **VI**s wurden für diese Vergleiche linear gestreckt. Die drei dargestellten Vegetationsklassen entstammen der in Kapitel 5.2.3 beschriebenen Klassifizierung und ihrer spätere Zuordnung. Im Vergleich ist die eingangs beschriebene Darstellungsverbesserung durch die Einführung des *soil line concepts* in den **DVI** ebenfalls erkennbar. Bestände von Pinien und Hartlaubbaeumen werden, der Realität entsprechend, als weniger chlorophyllhaltig ausgewiesen als die Wiesenaeaelchen.

Aus den oben genannten Gründen wurde ein zweiter Versuch mit der Eingabereihenfolge HRV xs1, HRV xs2, **nred**, HRV xs3, vorgenommen. Der Bildausschnitt in Abb. 5-20 wirkt zwar etwas verschwommener als im vorhergehenden Versuch, jedoch sind nun auch kleinere vegetationslose Areale als solche erkennbar.

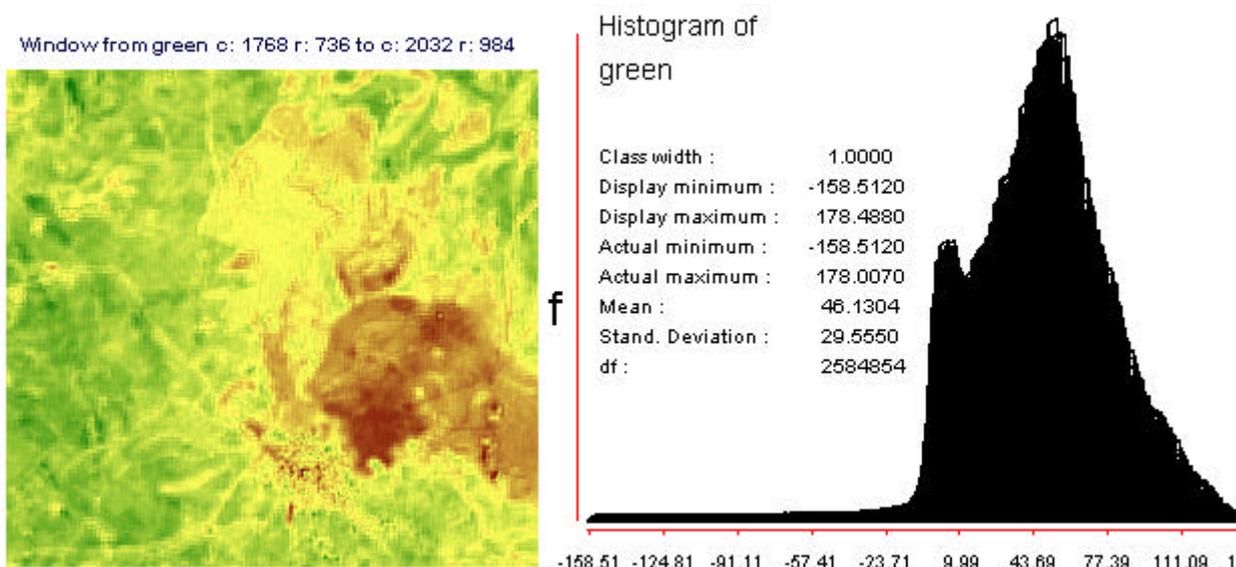


Abb. 5-20 Ausschnitt aus dem **GVI**-Image der Eingabefolge HRV xs1, HRV xs2, nred, HRV xs3,

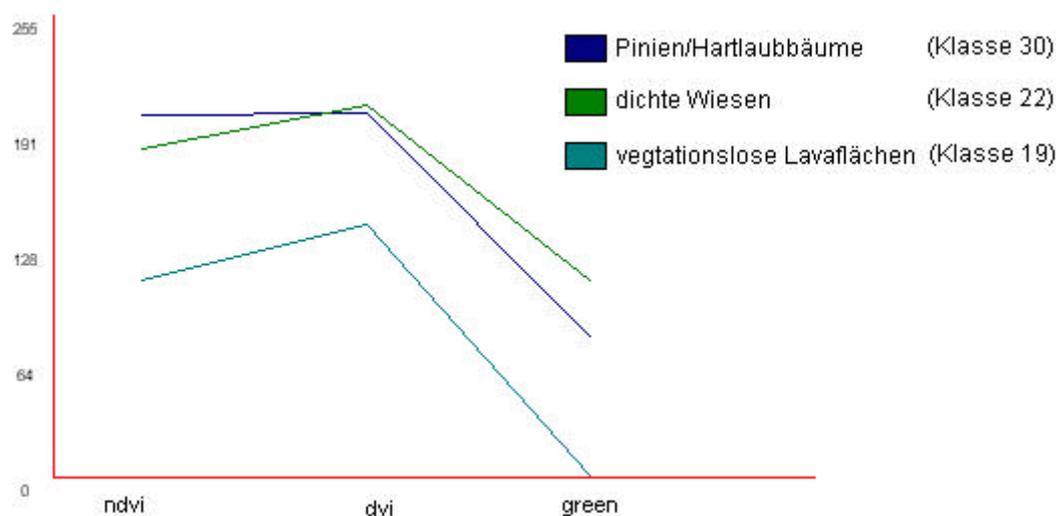


Abb. 5-21 Vergleich der Mittelwerte von **NDVI**, **DVI** und des **GVI** aus der in Abb. 5-20 angegebenen Eingabefolge. Vegetationslose Flaechen werden hier nach der linearen Streckung in **DN**s von 1-255 auch durch den Index 1 ausgewiesen.

Das Histogramm in Abb. 5-20 macht deutlich, daß die vorkommenden Werte auf eine größere Skalenbreite verteilt sind (-158.51 bis 178.00 gegenüber -98.64 bis 152.82), was auch das „verschwommene“ Erscheinungsbild erklärt. Die Negativwerte, welche wie bei den übrigen **VI**s die Abnahme der Vegetation kennzeichnen, sind über einen ca. 60 % breiteren Skalenraum verteilt. Dadurch werden Gebiete mit geringem Deckungsgrad stärker differenziert. Die Skalenbreite positiver Werte (zunehmende Vegetationsbedeckung) liegt nur um 17 % höher, wodurch bei einem direkten Vergleich der Bildausschnitte der Eindruck von „weniger“ grüner Vegetation z.B. auf den Wiesenflaechen entsteht. Durch die differenziertere Darstellung ist die leichte Überbetonung der *greenness* aus dem ersten Arbeitsschritt fast ausgeglichen, wie auch am Vergleich der arithmetischen Mittelwerte in Abb. 5-21 erkennbar

ist. Die Bilddarstellung entspricht dem Vitalitätsgrad der Frühlingsvegetation und der Bodenbedeckung der Insel.

Es muß an dieser Stelle nochmals betont werden, daß es sich bei der Ermittlung des **GVI** auf der Grundlage der **SPOT**-Datensätze **nicht** um ein neues, mathematisch ausgearbeitetes und bewiesenes Anwendungsverfahren für die Transformation von **SPOT**-Daten handelt. Es wurde hier lediglich mit einer für Landsat-Daten entwickelten Transformation experimentiert. Untersuchungen und Berechnungen ob die Transformation immer auch zuverlässig bei der Analyse von **SPOT**-Daten angewendet werden kann, müssen späteren Untersuchungen vorbehalten bleiben. Der ermittelte **GVI** stellte jedoch den Grünanteil der Vegetation El Hierros am deutlichsten dar (Gleiches gilt auch für das Image *yellow stuff*, welches quasi als Negativbild den Anteil vertrocknender bzw. chlorophyllarmer Vegetation erklärt. Mit diesen Informationen wurde hier nicht experimentiert.) Bei der Bestimmung der dreißig Bodenbedeckungsklassen, die durch automatische Klassifizierung erstellt wurden, waren die Klassenunterschiede im Gelände teilweise nur durch einen geringwertig höheren Anteil einjähriger Gräser und Kräuter zu erklären. Diese Beobachtungen konnten mit Hilfe des **GVI**, der erst nach der Geländearbeit ermittelt wurde, bestätigt werden. Auch die geplante Zusammenfassung wenig differenzierter Klassen in größere Einheiten ließ sich mit Hilfe des **GVI** absichern.

5.1.4 KOMBINATION UNTERSCHIEDLICHER SPEKTRALINFORMATIONEN (COLOR COMPOSIT)

Mit dem *color composit* werden drei verschiedene Images, deren Datenstruktur im Bytemodus (Zahlen zwischen 0 und 255) vorliegen muß, zu einem neuen Bild verbunden. Ein *color composit* ist nicht nur mit Original-Images möglich, sondern es können auch bearbeitete Datensätze eingeführt werden, wie z.B. Images, deren Inhalt Ergebnisse einer Ratio-Bildung oder **VI**s (nach einer Streckung der reellen Zahlen in Bytes) sind. Durch den verwendeten Algorithmus ist es möglich, den Dateninhalt der verwendeten Images zu gewichten und damit zu unterschiedlichen Farbdarstellungen zu gelangen. Unterschiedliche inhaltliche Aspekte können auf diese Weise verstärkt wiedergegeben werden. Im **IDRISI**-System erfolgt die Zuordnung der Datensätze in der Farbreihenfolge **Blau-Grün-Rot**. Zunächst wird jedes Eingabe-Image in nur noch sechs *levels* (d.h. in Werte von 0-5) gestreckt, wobei der *stretch*-Modus vom Bearbeiter bestimmt werden kann. Die frei wählbare Kombination der Images wird nach folgendem Algorithmus zu einem neuen Bild vereint:

$$\text{Composit} = \text{Blau} + (\text{Grün} * 6) + (\text{Rot} * 36)$$

Mit dieser Berechnung sind Ausgabe-Images in Pixelwerten von 0-215 möglich. Dem jeweils zuletzt eingegebenen Image wird das größte Gewicht beigemessen, damit wird die Information aus diesem Bild zum bestimmenden Inhaltsfaktor des Ausgabe-Image. Zur Darstellung der *color composites* ist im **IDRISI**-System eine eigene Farbpalette (*false color composit palette*) vorhanden, die aber in erster Linie zur Darstellung von *color composites* auf der Grundlage von Landsat-**TM**-Daten entwickelt wurde. Bei der Verwendung von **SPOT**-Daten ergeben sich deshalb bei der Bildschirmanzeige von Normal- oder Falschfarbenbildern (zur Eingabereihenfolge siehe Kapitel 5.1.3.1) leichte Farbveränderungen. In Verbindung mit den bisher in diesem Kapitel beschriebenen Bildverbesserungsmethoden ist eine Vielzahl von Kombinationsmöglichkeiten gegeben. Die *color composites* dienen in erster Linie der visuellen Darstellung bestimmter Sachverhalte. Sie dienen in unterschiedlichen Kombinationen als Basis für die Digitalisierung geeigneter Trainingsflächen zur überwachten Klassifizierung einzelner Bodenbedeckungsklassen und sind im **IDRISI**-System Voraussetzung für die automatisierte Klassenbildung.

Composit von blue, xs1, xs2. 0.5% Sat.

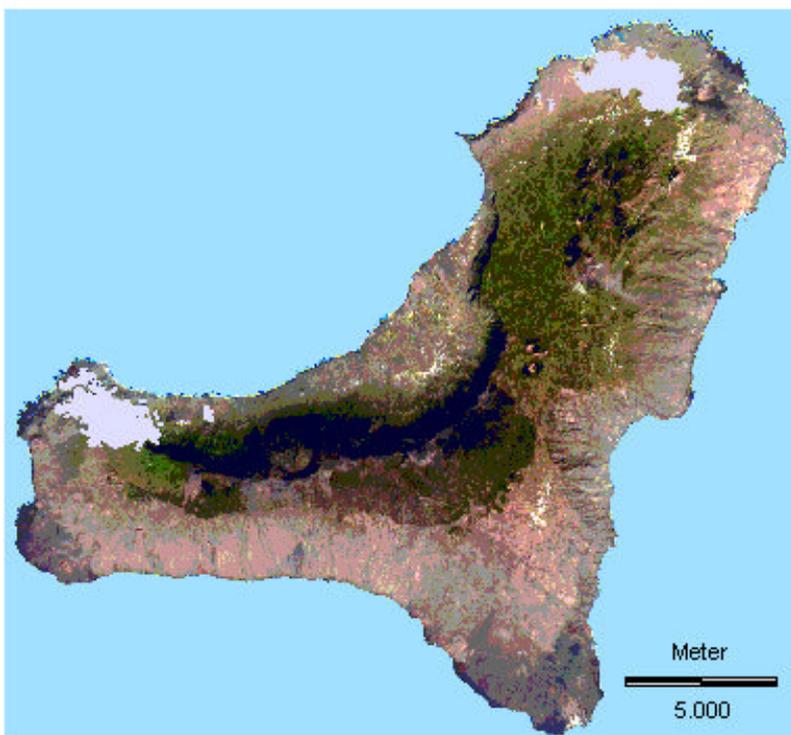


Abb. 5-22 Simulierte Normalfarbendarstellung. *color composite* aus B,G,R: durch Farbraumtransformation gewonnener Kanal Blau, HRV xs1 (Grün), HRV xs2 (Rot). Der Hintergrund und die dunkle Waldsignatur wurden nachträglich bearbeitet.

Für die automatische Bildung unterschiedlicher Vegetationsklassen ist diese Kombination nicht geeignet, weil sämtliche Informationen aus dem zwingend erforderlichen Bereich des nahen Infrarot fehlen. In Abb. 5-23 wird das notwendige Falschfarben-*Composit* vorgestellt, das als Grundlage für die weiteren Arbeitsschritte diene. Vorhandene grüne Vegetation wird hier in verschiedenen Rotabstufungen dargestellt, gering bewachsene bzw. überwiegend trockene bis vegetationsfreie Flächen in Übergängen von Graugrün bis Dunkelblau.

Composit von strxs1, nred, strxs3

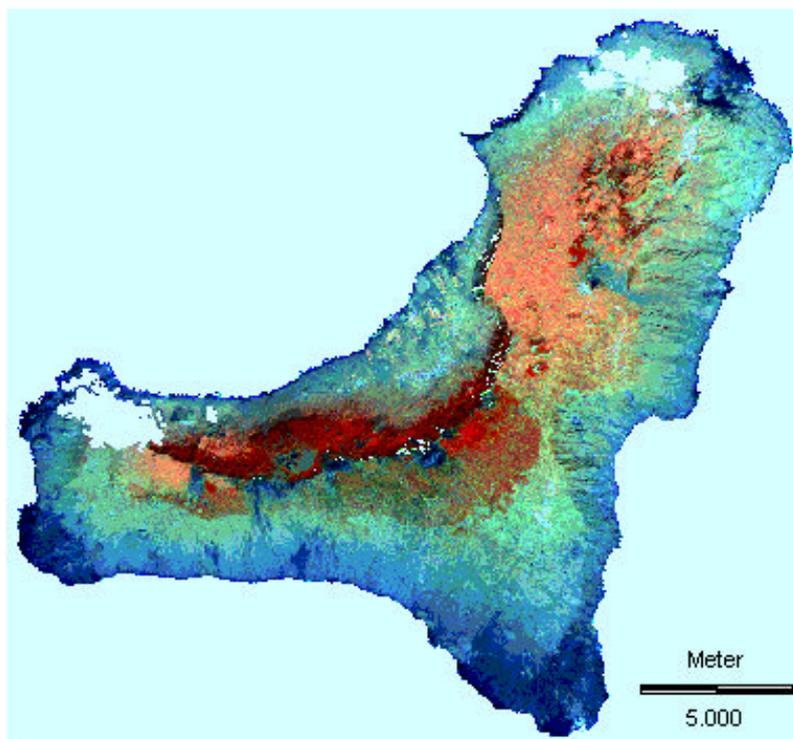


Abb. 5-23 Falschfarben-*composit* aus B,G,R: HRV xs1 (Grün), durch Transformation gewonnenes Rot (nred), HRV xs3 (nIR).

Aus der großen Menge herstellbarer *color composites* werden hier zwei Möglichkeiten vorgestellt. Abb. 5-22 zeigt die Simulation eines Image in „Echtfarben“. Weil in den **SPOT**-Datensätzen der für das *color composite* notwendige Kanal mit Informationen aus dem sichtbaren Blau fehlt, ist diese Darstellung erst nach „Schaffung“ dieses Spektrums durch die beschriebene **HLS/RGB**-Transformation möglich.

5.2 KLASSIFIZIERUNG DER DATEN

Die bisher beschriebenen Verfahren dienen der Verbesserung der Bildinterpretation durch den Betrachter, gleichzeitig sind sie erforderlich, um die digitale Bildauswertung zu steuern und zu beurteilen. *Einfache Überlegungen zeigen, daß unser menschliches Sehvermögen gewisse in Bild-daten vorliegende Informationen - z.B. Grauwerte oder Flächengrößen - nur sehr ungenau wahrnehmen kann und dabei sogar vielfachen optischen Täuschungen unterliegt. Andererseits ist seine Leistungsfähigkeit enorm, wenn es um die Wahrnehmung von Bildstrukturen, Formen, Texturen usw., um das Erkennen von Objekten oder um die Analyse von Zusammenhängen geht. Beim Computer ist es genau umgekehrt. Er kann mit hoher Präzision Grauwerte analysieren oder Flächen ermitteln, ohne dabei optischen Täuschungen zu unterliegen* (ALBERTZ, S. 139). Die Zusammenführung der gespeicherten Informationen gleich mehrerer multispektraler Datensätze zu verschiedenen Objektklassen ist dem menschlichen Auge nicht möglich. Diese Art der digitalen Bildauswertung (engl. *image classification* bzw. *multispektral classification*) ist eines der Hauptanwendungsgebiete rasterbasierter Geographischer Informationssysteme, während die Beurteilung über Richtigkeit und Zusammenhänge der ermittelten Objektklassen weiterhin dem Bearbeiter obliegen. Für die Klassifizierung stehen verschiedene Verfahren zur Verfügung.

5.2.1 UNÜBERWACHTE KLASSIFIZIERUNG

Die *unsupervised classification* erfolgt mittels einer Clusteranalyse, die eine vorherige Geländekenntnis des Bearbeiters nicht erfordert. Im Modul CLUSTER des IDRISI-Systems wird dazu eine *histogram peak*-Technik benutzt. Ein *histogram peak* ist in einem eindimensionalen diskreten Histogramm als ein Wert definiert, der mit größerer Häufigkeit (f) auftritt als die benachbarten Werte zu beiden Seiten. Durch die Clusterung werden nach der Ermittlung der *peaks* die dazwischenliegenden Werte dem jeweils am nächsten liegenden *peak* zugeordnet. (Zu verschiedenen Cluster-Techniken v.a. RICHARDS, S. 229 ff; LILLESAND/KIEFER, S. 604 ff; ALBERTZ, S.143 ff). Weil im IDRISI-System die Clusterung auf dem aus drei Kanälen ermittelten *color composit* erfolgt, wird für die Clusterung ein dreidimensionales Histogramm verwendet, das auch die diagonalen Nachbarn einschließt (EASTMAN 1997b, S. 11-21). Je differenzierter die Darstellung der Daten in dem verwendeten *color composit* ist, desto größer ist auch die Anzahl der möglichen zu bildenden Cluster (Abb. 5-24 und Abb. 5-25). D.h. durch sorgfältiges *image enhancement* und gezielte Auswahl des *color composit* kann auch bei einer unüberwachten Klassifizierung eine relativ genaue Zuordnung der Klassen erreicht werden. Nach der Clusterung ist es die Aufgabe des Bearbeiters, die unterschiedlichen Cluster zu definieren.

Zunächst wurde die Clusterung auf der Grundlage eines *color composit* der Original-Images durchgeführt. Gewählt wurde eine Feinclusterung mit der Option „*drop least significant clusters (<1 % of total area)*“. Durch dieses Verfahren wurden automatisch 19 Cluster erzeugt (Abb. 5-24). Das Ergebnis war nicht zufriedenstellend.

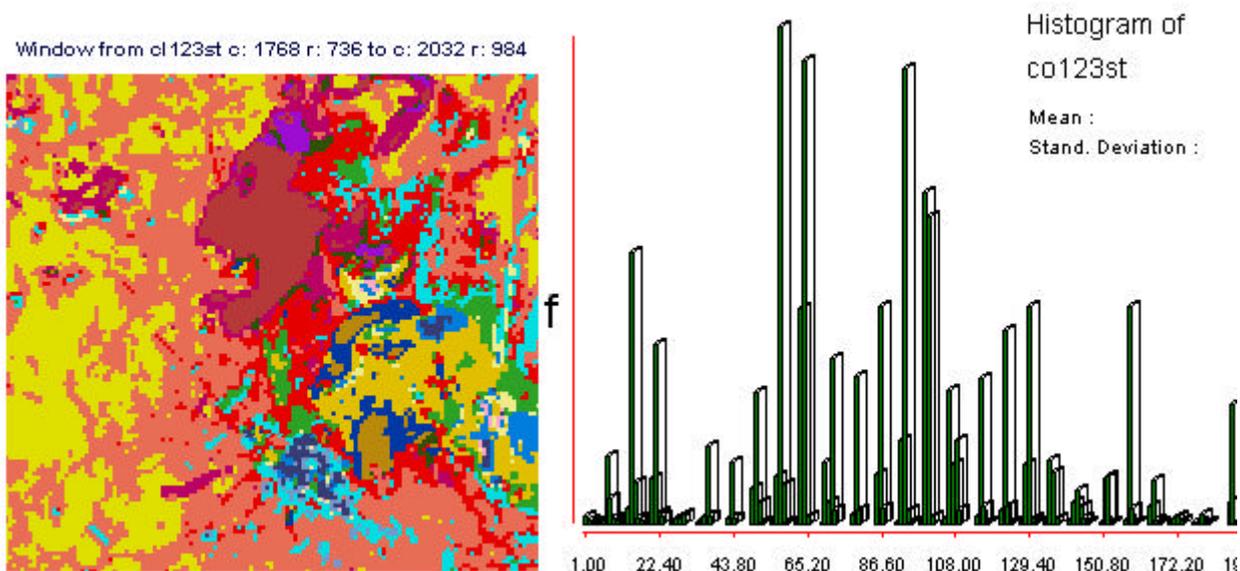


Abb. 5-24 Links: Ausschnitt einer Clusterung auf der Grundlage des *color composit* der bearbeiteten Kanäle *xs1*, *xs2*, *xs3*. Rechts: Dreidimensionales Histogramm des verwendeten *color composit*.

Nach der Einbeziehung des Image **nred** anstatt des Original-Image aus dem roten Spektrums wurden bei gleichem Verfahren 24 Cluster gebildet (Abb. 5-25).

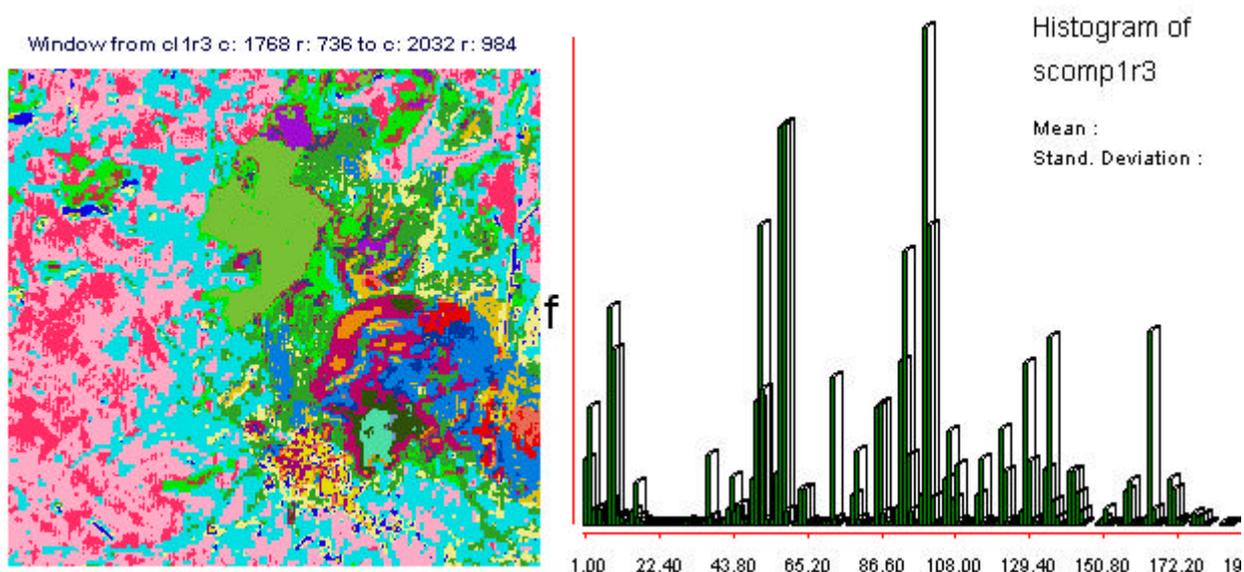


Abb. 5-25 Ausschnitt der Clusterung auf der Grundlage des *color composit* der Kanäle *xs1*, **nred**, *xs3* mit dessen Histogramm. Bei sonst gleichem Vorgehen wie in Abb. 5-24 beschrieben, wurden 24 Klassen (inkl. Hintergrund) gebildet.

Dieses Image stellte das Gelände von El Hierro, wo visuell nur eine unzureichende Differenzierung der Objektflächen möglich ist, bereits sehr differenziert dar. Viele der gebildeten Cluster ließen sich als der Autorin bekannte Vegetationseinheiten identifizieren. Diese Clusterung hätte eine gute Arbeitsgrundlage zur nachfolgenden Bestimmung der Objektklas-

sen im Gelände darstellen können. Dennoch kann nicht außer acht gelassen werden, daß es sich hier nur um statistische Nachbarschaftsbeziehungen spektral trennbarer Klassen handelt. Eine Wahrscheinlichkeitsberechnung, ob einzelne Bildelemente tatsächlich der Klasse zugehören, der sie zugeordnet werden, erfolgt nicht. Dies ist den Methoden der *supervised classification* vorbehalten. Es können jedoch (...) bestimmte spektrale Klassen innerhalb des Datensatzes erfaßt werden, die bei einer überwachten Klassifizierung nicht erkannt worden wären, da diese durch die Auswahl der Trainingsgebiete eine starke visuelle Komponente enthält. Eine nicht zu unterschätzende Fähigkeit des Verfahrens ist auch die schnelle automatische Identifizierung zahlreicher spektraler Klassen. Aus diesen Gründen wird die unüberwachte Klassifizierung oft der überwachten vorgeschaltet; sie läßt mitunter bereits im Voraus erkennen, wie gut die spektrale Trennbarkeit einzelner Erscheinungen ist und ob eine bestimmte, vom Interpreten ins Auge gefaßte Klasse spektral gut erkennbar ist oder nicht (LÖFFLER, S. 193). Nach der Geländearbeit wurde zu Vergleichszwecken eine Feinclusterung des *color composit* der Images xs1, nred, xs3 in 20 Klassen vorgenommen, das Ergebnis wird in Kapitel 8.1 diskutiert

5.2.2 ÜBERWACHTE KLASSIFIZIERUNG

Alle Methoden der überwachten Klassifizierung gehen von zwei Grundvoraussetzungen aus:

- daß dem Bearbeiter das Gelände und die Art seiner Bodenbedeckung bekannt sind, entweder durch eigene Erfahrung oder durch den Vergleich mit vorhandenen Landnutzungs- oder Vegetationskarten,
- daß genug einzelne Klassen in einer für die Art der Klassifizierung ausreichenden Pixelmenge⁴ sicher unterschieden und daraus Trainingsflächen digitalisiert werden können.

Weil keine ausreichend genaue Karte der Insel El Hierro existiert und die Vegetationseinheiten sehr kleinräumig sind, mußte die *supervised classification* ausgeschlossen werden. Es wurde bereits mehrfach darauf hingewiesen, daß bei der visuellen Betrachtung der Daten trotz der Geländekenntnis nicht mehr als maximal fünf Klassen sicher zugeordnet werden konnten, was für die Erstellung einer Vegetationskarte eindeutig zu wenig ist. Zudem sollte mit dieser Arbeit die Anwendbarkeit automatisierter Verfahren überprüft werden. Nach der oben beschriebenen Clusterung hätte der nach LÖFFLER zitierte Weg eingeschlagen werden können, mit den Vegetationsstrukturen korrelierende Cluster hätten später als Trainingsgebiete bestimmt und danach einer überwachten Klassifikation unterzogen werden können. Das **IDRISI**-System bietet aber zwei hybride Modelle der halbüberwachten Klassifizierung an. Benutzt wurde das Modul ISOCCLUS, das mit einer Methode der *supervised*

⁴ Als Faustregel für das Minimum gilt: Anzahl der in die Klassifizierung einbezogenen Kanäle x 10. Mit dem Minimum wird aber nur eine geringe Sicherheit der Klassifizierung erreicht.

classification, dem Verfahren der *maximum-likelihood*, verknüpft ist. Dieses Verfahren sei daher kurz beschrieben. Bei *maximum-likelihood* gilt die Annahme, daß die Reflexionswerte bestimmter Objektklassen nicht willkürlich gestreut werden, sondern sich in der statistischen Darstellung in Form einer Gauß'schen Normalverteilung um einen Mittelpunkt verteilen. Die spektrale Mischung digitalisierter Trainingsgebiete wird als Mittelpunkt bestimmter Klassen definiert, denen jedes einzelne Pixel nach der Kovarianz seiner spektralen Eigenschaften in den verschiedenen Kanälen mit größter Wahrscheinlichkeit zugeordnet werden kann (RICHARDS, S. 182 ff; LILLESAND/KIEFER, S. 594 ff; ALBERTZ, S. 192, LÖFFLER S. 144⁵). Dabei entstehen

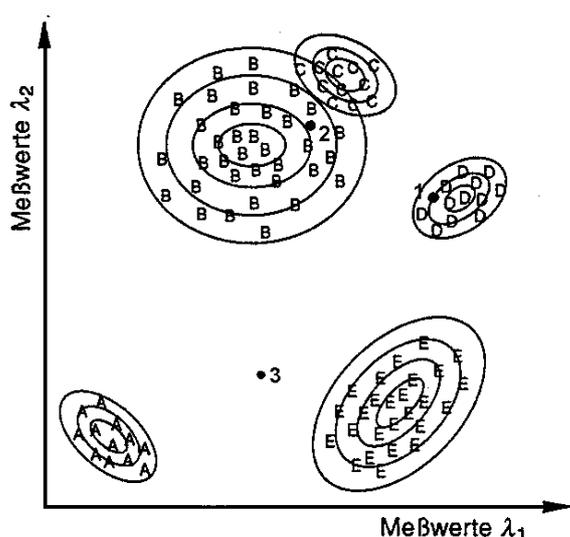


Abb. 5-26 Linien gleicher Dichte (Wahrscheinlichkeit) um einen Mittelpunkt, hier als zweidimensionale Abbildung. Aus ALBERTZ, S.144.

in einem dreidimensionalen Histogramm Linien gleicher Dichte um den Mittelpunkt, ähnlich den Höhenlinien um einen Hügel. Ein vorgegebener Prozentsatz schwer zuzuordnender Pixel kann von der Klassifizierung ausgeschlossen und in einer Klasse „nicht zuzuordnen“ ausgewiesen werden. Das Verfahren ist sehr zeitaufwendig, weil es die spektralen Meßdaten jedes für die Klassifizierung benutzen Spektralkanals einzeln und in der Beziehung zu jedem anderen berechnet, gilt aber bei einer größeren Anzahl von Trainingsgebieten für jede zu bildende Klasse als sicherstes.

5.2.3 INTERAKTIV ÜBERWACHTE AUTOMATISCHE KLASSIFIZIERUNG

Das Modul ISOCLUST verbindet die beschriebene *histogram peak*-Technik der Clusterbildung mit einem Algorithmus, der auf einer mehrfachen Wiederholung der *maximum-likelihood*-Berechnung beruht. Es handelt sich um den von BALL UND HALL⁶ entwickelten ISODATA-Algorithmus (ausführliche Beschreibung des Verfahrens bei RICHARDS, S. 231 ff). Dabei werden die zunächst durch das Clusterverfahren gebildeten „zufälligen“ Objektklassen an Stelle manuell digitalisierter Trainingsgebiete für die *maximum-likelihood*-Klassifizierung benutzt. Nach dieser Prozedur wird der Vorgang wiederholt, wobei diesmal diejenigen Klassen als Trainingsgebiete benutzt werden, die durch das *maximum-likelihood*-Verfahren entstanden sind. Dieser Vorgang wird mehrfach wiederholt. Das Ziel ist, die Summe der qua-

⁵Bei LILLESAND/KIEFER ist auch der Versuch einer dreidimensionalen Darstellung abgebildet, bei RICHARDS sind die mathematischen Formeln des komplexen Verfahrens zu finden.

⁶Literaturangabe bei EASTMAN 1997b und RICHARDS: G.H. BALL U. D.J HALL, 1965; A Novel Method of Data Analysis and Pattern Classification. Stanford Research Institute, Menlo Park, California. Der Artikel war der Autorin leider nicht zugänglich.

dratischen Abweichungen aller Pixel vom Mittelpunkt der Klasse, der sie zugeordnet wurden, zu reduzieren. *The iterative optimization clustering procedure, also called the migrating means technique, is essentially the isodata algorithm (...). It is based upon estimating some reasonable assignment of the pixel vectors into candidate clusters and then move them from one cluster to another in such a way that the SSE (sum of squared error) measure of the proceeding section is reduced (RICHARDS, S. 231).*

Die ISOCLUST-Routine wird in folgenden Schritten durchgeführt:

- Erstellung eines geeigneten *color composit* durch den Bearbeiter;
 - Auswahl der Datensätze, die in die Klassifizierung einbezogen werden sollen, und deren Reihenfolge in der Berechnung durch den Bearbeiter;
 - das *color composit* erfährt automatisch eine Feinclusterung mit der Option „*retain all clusters*“;
 - es wird ein Histogramm des erzeugten Datensatzes dargestellt, an Hand dessen der Bearbeiter entscheidet, wie viele Klassen gebildet werden sollen;
 - der Bearbeiter bestimmt die Anzahl der Wiederholungen der Klassifizierung;
 - automatisch werden aus der Anzahl der ausgewählten Klassen des ersten Clusters *signature*-Dateien erzeugt (jede enthält **alle** Polygone **eines** Clusters, stellt also die einzelnen „Trainingsgebiete“ für diese zu bildende Klasse dar), anschließend läuft die *maximum-likelihood*-Routine ab;
 - Wiederholung des zuletzt genannten Vorganges so oft, wie vom Bearbeiter vorgegeben.
- Im konkreten Fall heißt das:
- Erstellung des *color composit* mit den wie beschrieben verbesserten Kanälen **HRV xs1**, dem n(eugewonnen Kanal)**red, HRV xs3**;
 - Auswahl und Reihenfolge der Kanäle für die ISOCLUST-Routine. Gewählt wurde **HRV xs1, HRV xs2, HRV xs3, nred**;
 - nach der automatischen Clusterung erfolgte die Auswahl der Menge der zu bildenden Klassen aus den von ISOCLUST hergestellten temporären Clustern;
 - es wurden 5 Wiederholungen der *maximum-likelihood*-Klassifizierung gewählt.

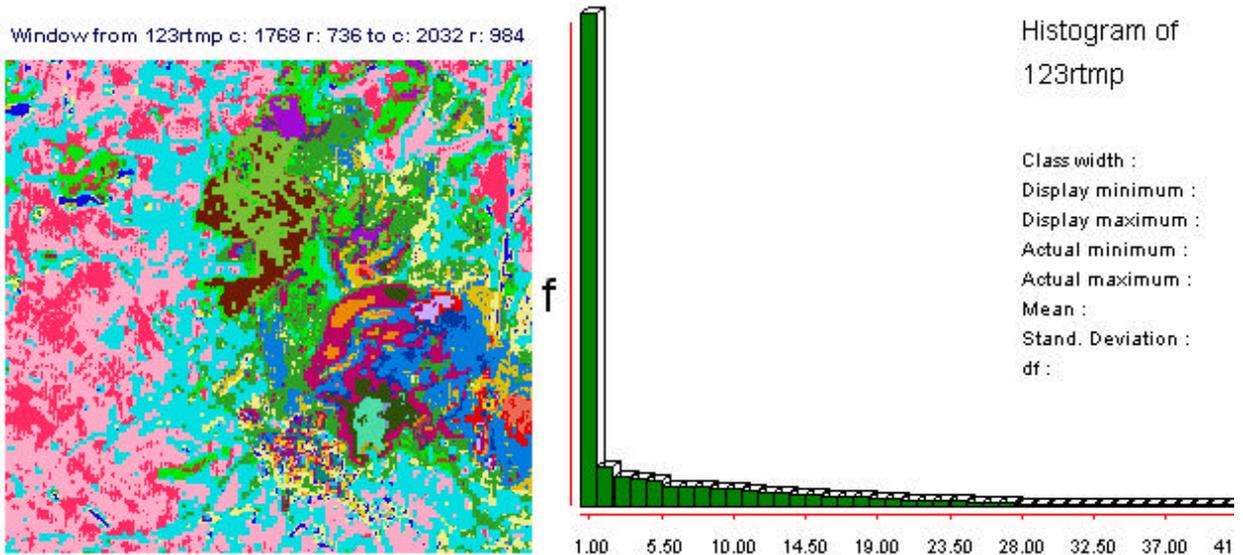
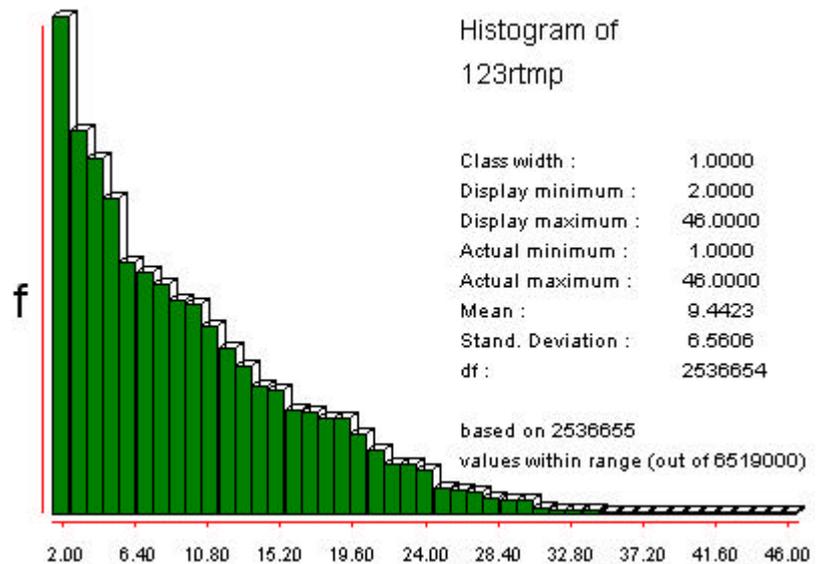


Abb. 5-27 Ausschnitt des aus dem als *seeding image* verwendeten *color composit* mit nebenstehendem automatisch erzeugtem Histogramm.

Bei der temporären Feinclusterung wurden 46 Cluster gebildet. Weil das gesamte *color composit*, einschließlich des Hintergrundes, in die Clusterbildung einbezogen wird und die Numerierung der Cluster abnehmend nach der Größe ihrer Gesamtfläche stattfindet, stellte das Cluster 1 auch den Hintergrund dar, der in allen Eingabe-Images den



Histogramm 5-3 Histogramm des Clusters aus Abb. 5-27 unter Auslassung des ersten Clusters, das überwiegend den Bildhintergrund enthält.

Pixelwert 0 hatte. Völlig vegetationslose Flächen aus schwarzer Asche oder Lava wurden in dieses Cluster einbezogen. Erkennbar ist dies im Ausschnitt in Abb. 5-27 an dem hier türkisfarbig dargestellten Aschekegel des schon bekannten Chamuscada. Das automatisch erstellte graphische Histogramm war daher wegen der Überbetonung des Hintergrundes nicht zu nutzen (ebenfalls Abb. 5-27). An dieser Stelle läßt die Arbeitsroutine aber die Erstellung weiterer Histogramme zu. Auf Grund des tabellarischen Histogramms (siehe Histogramm 5-4) wurden inklusive des Hintergrundes 31 Klassen ausgewählt. Verzichtet wurde dabei auf die Klassifizierung von 0,07 % (4784) der Pixelmenge des gesamten Image. Weil Hintergrund und Wolken aber schon 60,29 % des gesamten Image darstellen, bedeutet dies, daß auf 0,17 % aller auf der Insel vorkommenden möglichen Spektralkombi-

nationen verzichtet wurde. Die Bildpixel dieser Klassen wurden denjenigen zugeordnet, deren spektraler Mischung sie am ähnlichsten sind.

Als irrelevant beurteilt wurden die Cluster 32 bis 46. Vor einer solchen Entscheidung muß allerdings eine visuelle Kontrolle des Image stattfinden. Das Cluster 32 beispielsweise besteht aus 1635 Bildpixeln, was bei einer Auflösung von 10 m *10 m eine Größe von 16.35 ha bedeutet. Hätte es sich dabei um **eine** Fläche gehandelt, wäre darauf nicht zu verzichten gewesen. Es stellte sich aber bei der Überprüfung des Image heraus, daß bereits ab Cluster 26 die Bildpunkte z.T. als Einzelpixel über das gesamte Image verteilt waren. Die Anzahl der verbleibenden 31 Klassen wurde beibehalten (Gesamtansicht in Karte 1 im Kartenanhang, Ausschnitt in Abb. 5-28. Damit wurde verhindert, daß in ihrem Spektralverhalten ähnliche Klassen zu einer Klasse zusammengefaßt werden. Weil das Spektralverhalten von Vegetation nicht das einzige Kriterium ihrer Unterscheidung sein kann, blieb somit die Entscheidung vorbehalten, auch spektral sehr ähnliche Klassen (ihrer ökologischen Wichtigkeit wegen) als getrennte Klassen aufrecht zu erhalten. Nach der Geländearbeit und dem Abschluß einer Zusammenfassung in 19 Arten der Bodenbedeckung wurde die Isoclust-Routine wiederholt, wobei diesmal die Erzeugung von neunzehn Spektralklassen (plus Hintergrund) gewählt wurde. Die Gegenüberstellung dieser automatischen Klassifizierung mit der eigenen Zusammenfassung wird ebenfalls in Kapitel 8.1 erörtert.

Histogram of.123rtmp

Class	Lower Li-	Upper	Frequency	Prop	Cum.	Cum. Prop
0	1.0000	1.9999	3982345	0.6109	3982345	0.6109
1	2.0000	2.9999	299052	0.0459	4281397	0.6568
2	3.0000	3.9999	230052	0.0353	4511449	0.6920
3	4.0000	4.9999	212951	0.0327	4724400	0.7247
4	5.0000	5.9999	188777	0.0290	4913177	0.7537
5	6.0000	6.9999	150548	0.0231	5063725	0.7768
6	7.0000	7.9999	144808	0.0222	5208533	0.7990
7	8.0000	8.9999	137493	0.0211	5346026	0.8201
8	9.0000	9.9999	127557	0.0196	5473583	0.8396
9	10.0000	10.9999	125074	0.0192	5598657	0.8588
10	11.0000	11.9999	112097	0.0172	5710754	0.8760
11	12.0000	12.9999	98389	0.0151	5809143	0.8911
12	13.0000	13.9999	88427	0.0136	5897570	0.9047
13	14.0000	14.9999	76084	0.0117	5973654	0.9163
14	15.0000	15.9999	74136	0.0114	6047790	0.9277
15	16.0000	16.9999	61092	0.0094	6108882	0.9371
16	17.0000	17.9999	59851	0.0092	6168733	0.9463
17	18.0000	18.9999	57172	0.0088	6225905	0.9550
18	19.0000	19.9999	56106	0.0086	6282011	0.9636
19	20.0000	20.9999	47336	0.0073	6329347	0.9709
20	21.0000	21.9999	37677	0.0058	6367024	0.9767
21	22.0000	22.9999	29502	0.0045	6396526	0.9812
22	23.0000	23.9999	29455	0.0045	6425981	0.9857
23	24.0000	24.9999	25041	0.0038	6451022	0.9896
24	25.0000	25.9999	14064	0.0022	6465086	0.9917
25	26.0000	26.9999	12854	0.0020	6477940	0.9937
26	27.0000	27.9999	11538	0.0018	6489478	0.9955
27	28.0000	28.9999	7858	0.0012	6497336	0.9967
28	29.0000	29.9999	7160	0.0011	6504496	0.9978
29	30.0000	30.9999	6928	0.0011	6511424	0.9988
30	31.0000	31.9999	2792	0.0004	6514216	0.9993
31	32.0000	32.9999	1635	0.0003	6515851	0.9995
32	33.0000	33.9999	762	0.0001	6516613	0.9996
33	34.0000	34.9999	663	0.0001	6517276	0.9997
34	35.0000	35.9999	518	0.0001	6517794	0.9998
35	36.0000	36.9999	408	0.0001	6518202	0.9999
36	37.0000	37.9999	408	0.0001	6518610	0.9999
37	38.0000	38.9999	175	0.0000	6518785	1.0000
38	39.0000	39.9999	86	0.0000	6518871	1.0000
39	40.0000	40.9999	54	0.0000	6518925	1.0000
40	41.0000	41.9999	28	0.0000	6518953	1.0000
41	42.0000	42.9999	25	0.0000	6518978	1.0000
42	43.0000	43.9999	11	0.0000	6518989	1.0000
43	44.0000	44.9999	7	0.0000	6518996	1.0000
44	45.0000	45.9999	3	0.0000	6518999	1.0000
45	46.0000	46.9999	1	0.0000	6519000	1.0000

Class width = 1.0000
 Display minimum = 1.0000
 Display maximum = 46.0000
 Actual minimum = 1.0000
 Actual maximum = 46.0000
 Mean = 4.0453
 Stand. Deviation = 5.7988
 df = 6518999

Histogramm 5-4 Tabellarisches Histogramm der temporären Clusterung von ISOCLUST.

Das Ausgabe-Image der Klassifizierung in 31 Klassen entsprach in vielen Einzelheiten den Kenntnissen über bestimmte Vegetationsstandorte. Es war deshalb für die Zuordnung der Klassen im Gelände außerordentlich gut geeignet. Das Histogramm in Abb. 5-28 verdeutlicht im Vergleich mit dem Histogramm in Abb. 5-27, welchen Klassen die Pixel der für nicht relevant gehaltenen Cluster zugeordnet wurden. Durch die Beibehaltung der Numerierungsfolge der ersten 31 Klassen sind die Flächen nun nicht mehr nach ihrer Größenabnahme geordnet dargestellt. Es ist erkennbar, daß besonders viele Bildpunkte unter die Klasse 4 subsumiert wurden, im Gelände wurde diese Klasse als trockener Matorral mit einer Vegetationsbedeckung von ca. 50 % definiert. Sie stellt damit die flächenmäßig größte Vegetationseinheit dar. Im Histogramm wird wieder auf die Darstellung der 1. Klasse verzichtet, die diesmal ausschließlich den Bildhintergrund repräsentiert.

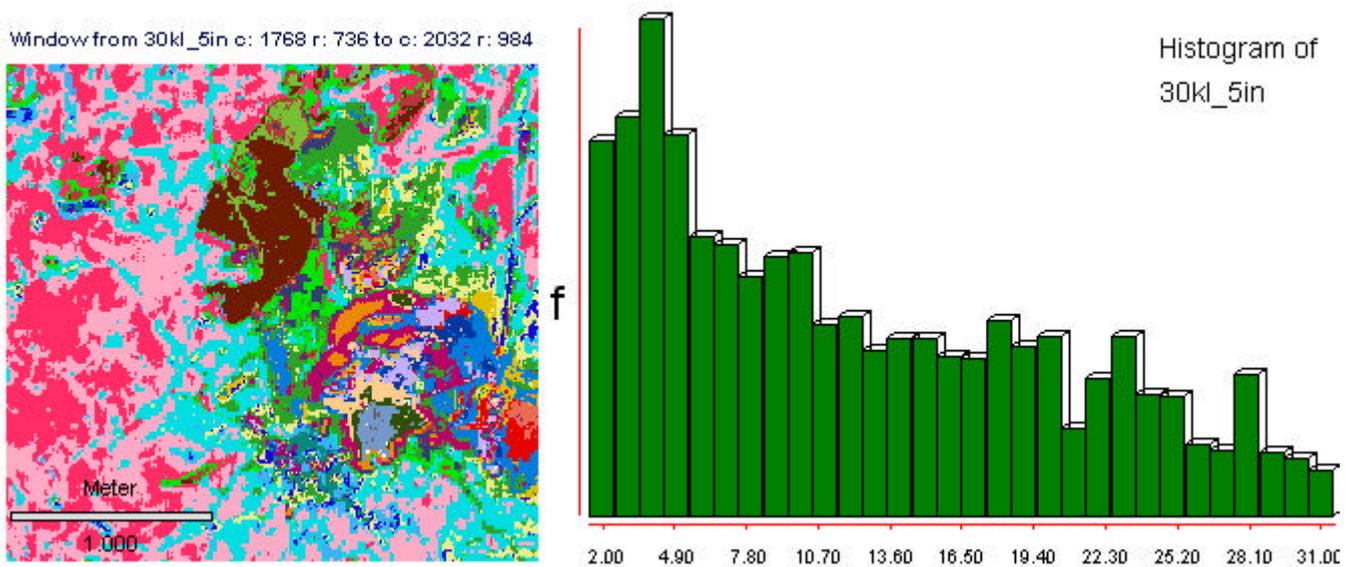


Abb. 5-28 Ausschnitt des mit ISOCLUST erzeugten Image mit 31 Klassen und zugehörigem Histogramm.