Segmentierung mehrkanaliger Datenaufkommen am Beispiel einer Verarbeitungskette für die Fernerkundung

T. Koch, R. Nestler, K.-H. Franke

Zentrum für Bild- und Signalverarbeitung e.V. Ilmenau (ZBS)
Werner-von-Siemens-Str. 10, D-98693 Ilmenau
eMail:
torsten.koch@zbs-ilmenau.de
rico.nestler@zbs-ilmenau.de
karl-heinz.franke@zbs-ilmenau.de
URL: http://www.zbs-ilmenau.de

Zusammenfassung Die satellitenbildgestützte Erfassung der Landoberfläche zur Ableitung und Kartierung verschiedener Eigenschaften (z.B. Landnutzung, Biomasse, Feldfruchtarten) oder das Monitoring der Landbedeckung wird seit Langem betrieben. Aus der synergetischen Nutzung von Bilddaten unterschiedlicher Sensorprinzipien und der gleichzeitigen Berücksichtigung mehrerer Aufnahmezeitpunkte können dabei deutliche Vorteile gezogen werden. Wesentlich für die Akzeptanz von Verarbeitungslösungen sind neben der erzielten Prozessierungsgüte im Endprodukt die effiziente und weitestgehend automatische Prozessierung des dabei anfallenden umfangreichen Datenmaterials.
Die im Beitrag vorgestellte Lösung zur Ableitung vollständig segmentierter Szenenabbilder aus multimodalen und multitemporalen Datenaufkommen schöpft das den Daten innewohnende szenenbeschreibende Informationspotential optimal aus. Das Verfahren arbeitet automatisch und kann beliebige multikanalige Daten simultan und effizient verarbeiten. Das Ergebnis bildet die Grundlage für objektorientierte Szenenbildanalysen, die durch die Berücksichtigung integraler Objekteigenschaften robuster gegenüber sensorspezifischen Messunsicherheiten, wie zum Beispiel Speckle, sind als bekannte Lösungen. Der Algorithmus ist leicht konfigurierbar und so an die Bedürfnisse unterschiedlicher Anwendungsfelder anpassbar.
Das vorgestellte Verfahren ist Bestandteil einer komplexen automatischen Verarbeitungskette zur Verarbeitung von Satellitenbilddaten, die weitere Module, wie sensorübergreifende Vorverarbeitung, Registrierung und Objektklassifikation, beinhaltet. Im Rahmen der objektorientierten Klassifikation erfolgt die semantische Zuordnung der Szenenregionen entsprechend der Zielstellungen der Prozessierung.

1 Einleitung

Die Erzeugung von zusammenhängenden Regionen durch das Zusammenfassen von Pixeln ist ein wesentlicher Verarbeitungsschritt vor der nachfolgenden anwendungsbezogenen Bildanalyse. Im Fall mehrkanaliger Bilddaten wird der Zusammenhang über mehrdimensionale Homogenitätskriterien definiert, die darüber hinaus die unterschiedliche Aussagekraft der einbezogenen Datenkanäle berücksichtigen müssen. Dies ist insbesondere im Fall multimodaler oder multitemporaler Daten aufgrund der unterschiedlichen Szenenrepräsentation und abgebildeten Szeneninhalte von Bedeutung. Sollen im Rahmen einer ebenfalls objektorientierten Klassifikation, wie sie im Bereich der Fernerkundung heutzutage Stand der Technik ist [Lucht2009], segmentbasierte Attribute und Relationen verwendet werden, ist ein Segmentierungsergebnis erforderlich, das vollständig, semantisch exakt und möglichst nicht übersegmentiert ist. Die bekannten Lösungen zur regionenorientierten Segmentierung, wie zum Bei-

spiel auf Basis des Region Growing, erfordern zum Erreichen dieser Zielstellungen die manuelle Festlegung einer Vielzahl von Prozessparametern.
Das nachfolgend vorgestellte mehrkanalige Segmentierungsverfahren arbeitet hingegen automatisch und realisiert ein bezüglich der obengenannten Kriterien nahezu optimales Ergebnis. Dabei können Bilddaten beliebiger Bildsensoren und Aufnahmezeitpunkte verarbeitet werden, sofern sich zwischen diesen ein semantischer Bezug herstellen lässt. Am Beispiel der Prozessierung von Satellitenbildern zur Bestimmung der Landbedeckung wird gezeigt, wie sich das Verfahren in eine komplexe automatische Verarbeitungskette einfügt. Die spektral und lateral hochaufgelöste Erfassung der Landoberfläche mit hoher zeitlicher Wiederholrate stellt besondere Anforderungen an die zur Verarbeitung einzusetzenden Methoden. Die Bewertung der Güte des Segmentierungsergebnisses erfolgt hier auf der Ebene der generierten Landnutzungskarten.

2 Multikanalige Segmentierung

Die algorithmische Grundlage des regionenbasierten Segmentierungsverfahrens bildet die Kombination eines *Region-Growing*- und *Region-Merging*-Verfahrens. Durch das Region-Growing wird eine Start-Segmentierung für das anschließende hierarchische Region-Merging erzeugt, welches die initial vorhandene Übersegmentierung in mehreren Detailstufen abbaut. Die Verfügbarkeit mehrer Detailebenen bietet für die nachfolgende objektorientierte Bildanalyse optimale Voraussetzungen. Der grundsätzliche Ablauf des Segmentierungsvorganges ist in Abbildung 1 dargestellt.

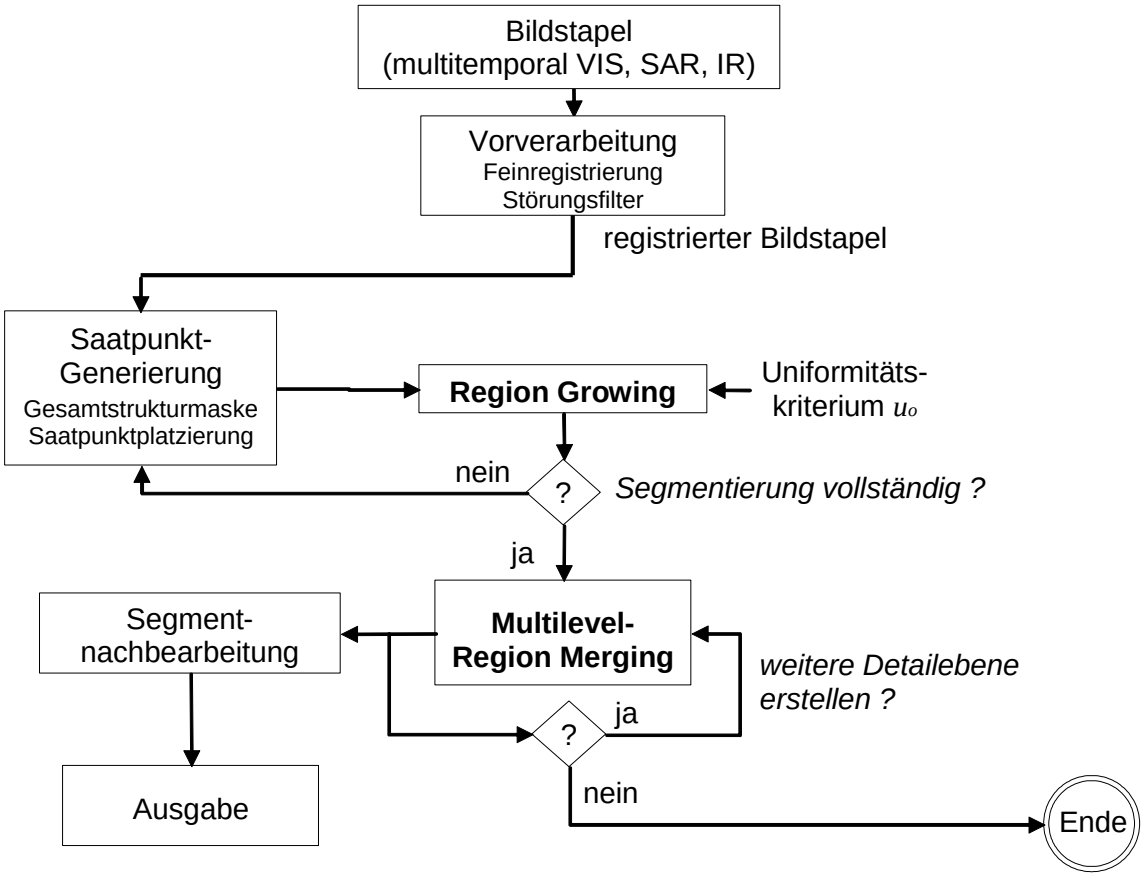

Abbildung 1: Schematischer Ablauf des Segmentierungsvorgangs

Die Regionenzugehörigkeit wird durch ein gewichtetes Homogenitätskriterium definiert, das ein euklidisches Abstandsmaß im mehrdimensionalen Merkmalsraum des registrierten multimodalen oder multitemporalen Bildstapels darstellt.
Bei der Verwendung der Bilddaten verschiedener Sensorprinzipien liegen üblicherweise unterschiedliche Auflösungen und Szenenausschnitte vor. Demzufolge ist die exakte Überführung der angebotenen Szenenbildrepräsentationen auf eine gemeinsame Referenz im Rahmen der Vorverarbeitung erforderlich.
Alle Kanäle werden während des Region-Growings simultan verarbeitet. Es erfolgt demnach keine Verschneidung von einzelnen (z.B. sensorspezifisch) prozessierten Kanälen. Dadurch können Informationslücken oder -defizite in einzelnen Datenkanälen durch die übrigen Kanäle ausgeglichen werden. Zu diesem Zweck wird der Kanalbeitrag zum Homogenitätskriterium gewichtet.
Als Ergebnis der Segmentierung entsteht ein Stapel aus Segmentbildern, wobei jede Detailebene durch das Zusammenfassen von Segmenten der vorhergehenden Ebene gebildet wird.
Der Segmentierungsprozess ist im Gegensatz zu bekannten Region-Growing-Verfahren mit zufällig gesetzten Saatpunkten determiniert. Damit wird die Vergleichbarkeit der Segmentierungsergebnisse bei Realisierung weiter gehender Verarbeitungsketten sichergestellt.
Um eine möglichst optimale Start-Segmentierung durch Region-Growing zu erzielen, sollten typische Störungen (Messunsicherheiten) im Vorfeld strukturerhaltend reduziert werden. Die Auswahl des einzusetzenden skalaren oder auch vektoriellen Filterverfahrens muss nach initialer Datenanalyse anwendungsbezogen erfolgen. Neben breit etablierten Filtern stehen auch neuere Verfahren, wie z.B. Anisotroper Diffusionsfilter (AD), Geometrisches Filter und Vektor-Median-Filter zur Verfügung. Im Anwendungsfall Fernerkundung bewähren sich Filter nach dem Prinzip der anisotropen Diffusion für Radar- und optische Daten besonders. In Abbildung 2 ist das Ergebnis der strukturerhaltenden Störungsreduktion am Beispiel von Radardaten dargestellt.

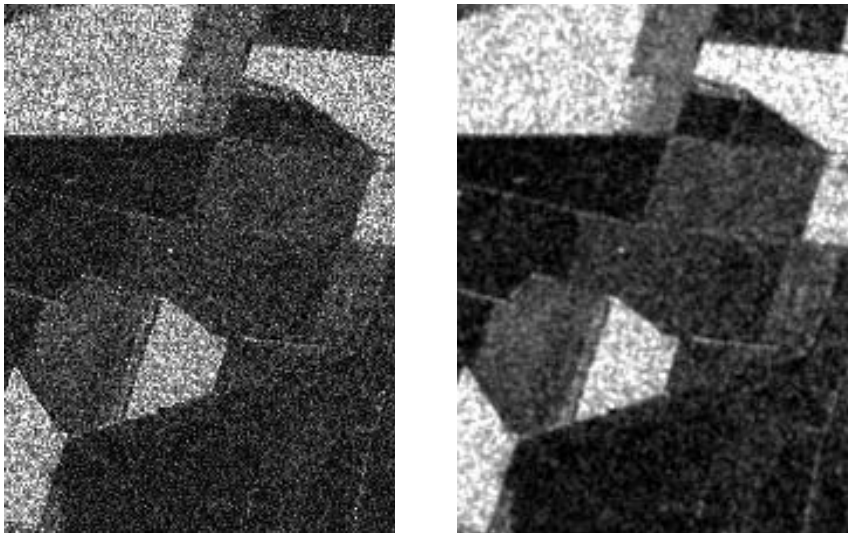

Abbildung 2: Ausschnitt aus einer Radarszene (©DLR, 2009) mit starkem Speckle-Einfluss, links vor und rechts nach Vorverarbeitung mit Diffusions-Filter

Der oben erwähnte determinierte Charakter der Segmentierung ergibt sich dadurch, dass die Saatpunkte im hier vorgestellten Verfahren nicht zufällig, sondern automatisch szenenangepasst, also determiniert, an optimale Positionen gesetzt werden.
Hierfür wird grundsätzlich von der Annahme ausgegangen, dass diese Positionen sich durch ein aus den Eingangskanälen globales, fusioniertes Streuungsminimum auszeichnen. Dadurch wird ausgeschlossen, dass Saatpunkte auf Regionengrenzen platziert werden und der Wachstumsprozess in glatten Regionen begonnen wird.

Mit diesem Ziel werden zu allen Eingangskanälen Streuungsbilder (laterale Varianz) berechnet und anschließend fusioniert (semantisch gewichtet akkumuliert). Dieses fusionierte Varianzbild stellt die Grundlage der nachfolgenden Saatpunktplatzierung und Verfahrensparametrierung dar.

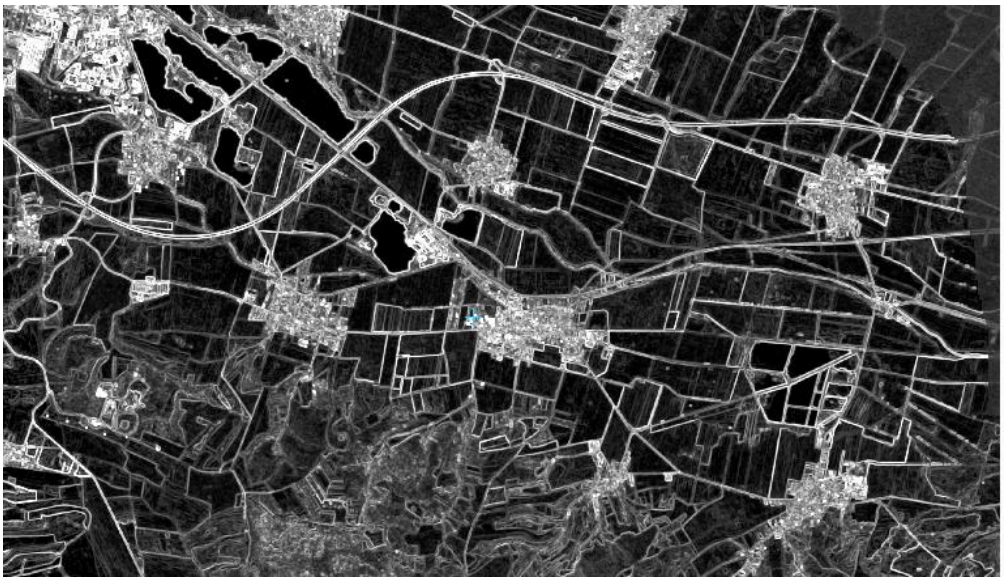

Abbildung 3: Fusioniertes Varianzbild aus optischen- und Radar-Kanälen

Zur *Saatpunktplatzierung* werden alle Pixel des fusionierten Varianzbildes mit $\sigma_{Thresh}=f(u_0)$ (u_0 - initiales Uniformitätskriterium) ausmaskiert und der verbleibende Rest mit einem Zeilen-Koinzidenz-Verfahren zu geschlossenen Regionen zusammengefasst. Die Saatpunkte werden auf die Positionen des kleinsten Streuungswertes der jeweiligen Regionen gesetzt. Werden weniger als *k* (proportional zur Bildgröße) Saatpunkte gefunden, muss der Schwellwert σ_{Thresh} um $d\sigma$ erhöht und der Vorgang wiederholt werden. Der Parameter $d\sigma$ ergibt sich aus der ermittelten Variationsbreite des fusionierten Varianzbildes (Abbildung 3).
Ausgehend von den ermittelten Saatpunkten wird unter Berücksichtigung der gegebenen Uniformitätskriterien ein simultanes Wachstum aller Saatpunkt-Regionen durchgeführt. Hierbei werden neue Pixel einer Saatpunktregion nur dann zugeschlagen, wenn deren euklidischer Abstand im Merkmalsraum zum jeweiligen Regionenmittelwert kleiner oder gleich dem gewählten Uniformitätskriterium u_0 ist.
Durch wiederholte Iteration zwischen *Saatpunktgenerierung* und *Region-Growing* wird die gesamte Szene determiniert, vollständig segmentiert.
Nach Abschluss des Region-Growings ist die Segmentierung durch die mehrkanalige Identifikation und Berücksichtung von Regionengrenzen zwar sematisch exakt, jedoch noch nicht optimal. Insbesondere ist oft eine deutliche Übersegmentierung zu verzeichnen, die zur Gewährleistung semantisch korrekter Endergebnisse in dieser Verarbeitungsphase toleriert werden muss.
Da sich Objekte aus mehreren simultan entstandenen Nachbarregionen zusammensetzen können, wird ein nachträgliches Zusammenfassen dieser Regionen (Merging) notwendig. Dieser Vorgang findet im Anschluss an das Regionen-Wachstum statt und wird in mehreren Zyklen durchgeführt. Dadurch entsteht eine geordnete Menge von sukzessive gröber segmentierten Ergebnissen (Segment-Level). Das Generieren und Vorhalten unterschiedlichen Segmentlevel ist sinnvoll, um unterschiedliche Analyseziele zu verfolgen.
In einem Region-Merging-Schritt werden nur dann Segmente miteinander verschmolzen, wenn der euklidische Abstand ihrer mittleren Merkmalsvektoren kleiner als das levelspezifi-

sche Uniformitätskriterium u_L ist. Dieses wird von einem Region-Merging-Schritt zum Nächsten, wie in Gleichung 1 dargestellt, modifiziert.

$$u_L = u_0 \cdot k_L \qquad \begin{array}{c} mit \\ k_0 = 1 \\ L \in [0, L_{Max} - 1] \subset N \end{array}$$

Gleichung 1: Ansatz zur Aufweichung des Uniformitätskriteriums u_L; L_{Max} = Anzahl der Segmentierungslevel, k_L=konst., u_0=initiales Uniformitätskriterium

Bevor ein Segment-Level ausgegeben wird, erfolgt eine *Segmentnachbearbeitung*, um Kleinstsegmente insbesondere an Regionengrenzen zu eliminieren. Dabei werden Kleinstsegmente durch Aufweichung des Uniformitätskriteriums einem angrenzenden „uniformen" Nachbarn zugeordnet.

Im Zuge des Region-Merging kann es erforderlich sein, Untersegmentierungen durch große Segmente mit kleiner gemeinsamer Berandung zu verhindern. Ein zusätzlich eingeführtes, formbasiertes Kriterium p_{Shape} soll dies verhindern. Ein korrektes Verschmelzen wird umso wahrscheinlicher, je größer der Anteil der gemeinsamen Grenzpixel l im Verhältnis zum Umfang eines potentiell aus dem Merging beider Segmente neu entstehenden Segmentes ist (Gleichung 2).

$$p_{shape}(a,b) = \frac{l(a,b)}{u(a) + u(b) - 2l(a,b)} \in [0;1] \subset R$$

Gleichung 2: Formmerkmal p_{shape}; a,b= Segmente, $u(a)$=Berandung von Segment a [in Pixel], $l(a,b)$=Anzahl der gemeinsamen Grenzpixel von Segment a und b

In Abbildung 4 ist die Wirkung des formbasierten Kriteriums p_{Shape}, dargestellt. Es wird deutlich, dass hinsichtlich der Anzahl der erzeugten Segmente in den einzelnen Region-Merging-Detailstufen eine nahezu asymptotische Annäherung an ein stabiles Endergebnis statt findet. Vor allem in Hinblick auf eine nachfolgende segmentbasierte Klassifikation ist dies von großem Vorteil. Durch das hier vorschlagene Prinzip des Region-Mergings in Kombination mit der formbasierten Segmentnachbearbeitung werden aus dem Ergebnis des Region Growings die gewünscht optimalen Segmentierungsergebnisse realisiert.

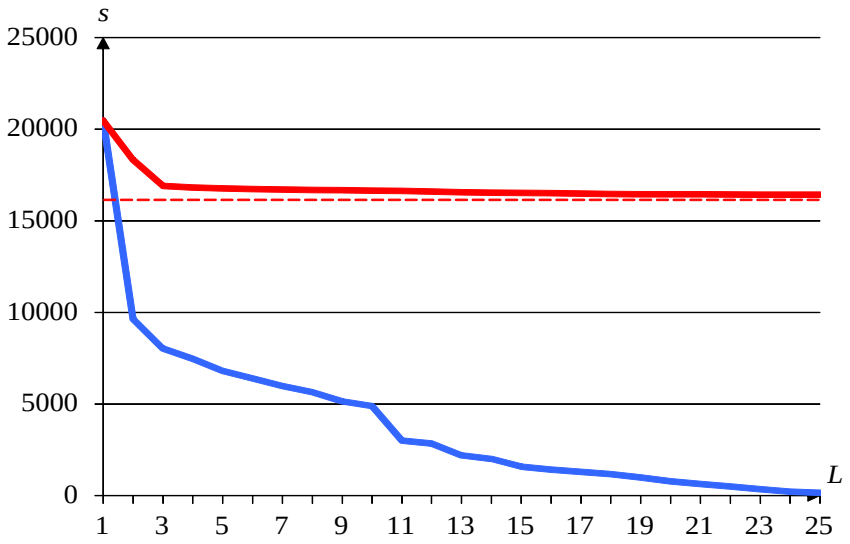

Abbildung 4: Segmentanzahl s in Abhängigkeit vom Detail-Level L (rot: mit Berücksichtigung von des formbasierten Entscheiders p_{Shape}; blau: ohne Berücksichtigung von p_{Shape})

3 Anwendungsszenario bildgestützte Fernerkundung

Im Rahmen des Projektes Enviland-2[1] wurde eine Verarbeitungspipeline für die Prozessierung von mehrkanaligen, multimodalen Satellitenbildmaterial entwickelt. Dies betrifft vor allem die Kantenextraktion, das Matching, die Segmentierung sowie die Klassifikation (Abbildung 5). Wesentliche Besonderheit und Neuheit der damit realisierten Lösung sind die weitestgehend automatische Ableitung von Prozessierungsendprodukten hoher Güte. Dies wird durch die durchgängig synergetische Datenprozessierung in allen Prozessstufen und den Einsatz segment- bzw. objektorientierter Klassifikatorensembles realisiert. Durch simultane Verarbeitung von Bildmaterial unterschiedlichen sensorischen und zeitlichen Ursprungs können Informationsdefizite einzelner Kanäle, z.B. durch Wolkenbedeckung im optischen Bereich, durch den Informationsgehalt anderer Kanäle (z.B. Radar-Kanäle) ausgeglichen werden. Einen besonders hohen Stellenwert besitzt hierbei die Segmentierung als Vorstufe der anwendungsbezogenen Bildanalyse.

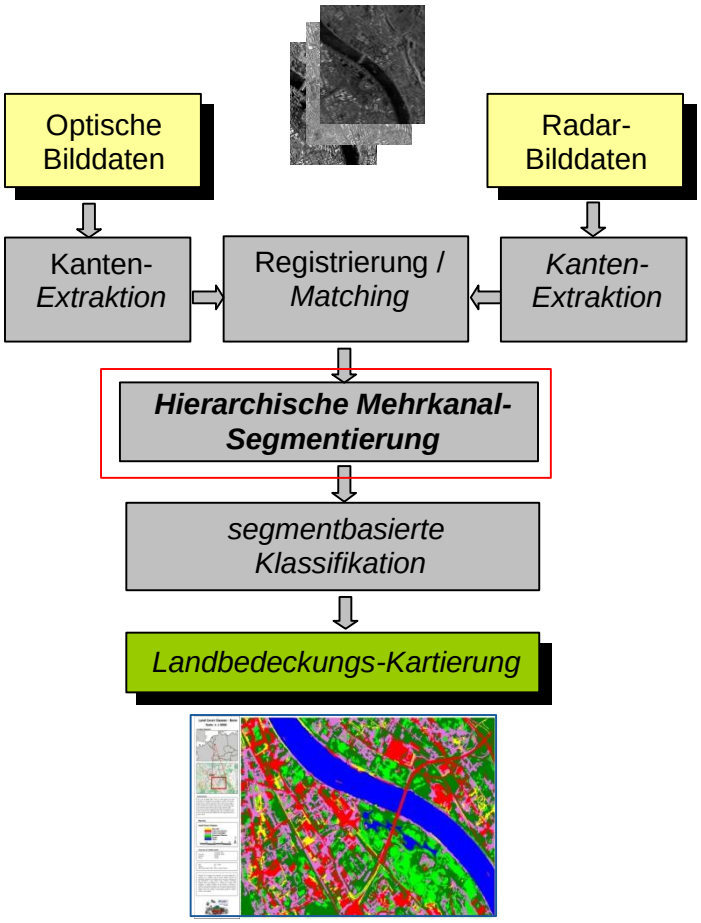

Abbildung 5: Schlüsselstellung der hierarchischen Mehrkanal-Segmentierung in der automatischen Prozesskette zur Prozessierung multimodaler / multitemporaler Satellitendaten (gelb: Input-Daten; grün: Prozessierungsendprodukt)

Im hier dargestellten Testfall wurden insgesamt 17 Kanäle einer Szene aus der Region Nordhausen zu verschiedener Beobachtungszeitpunkten im Jahr 2009 verarbeitet. Der multitemporale Datensatz besteht aus 5 optischen Kanälen (Abbildung 6) und 6 Radar-Kanälen mit jeweils 2 Polarisationen (Abbildung 7).

[1] Förderprojekt des BMWi, Nähere Informationen siehe Danksagung.

Die Verarbeitungseinheiten Kantenextraktion und Registrierung/Image-Matching stellen sicher, dass die Eingangsbilder trotz prinzipbedingt stark unterschiedlicher Repräsentation des Szeneninhaltes pixelgenau registriert sind [Koch2011]. Der so vorprozessierte Bildstapel wird im Rahmen der *Segmentierung* zu einem Stapel von Segmentbildern in unterschiedlichen Detailstufen verarbeitet.

In Abbildung 8 zeigt das Ergebnis der Segmentierung einer strukturbasiert optimalen Detailstufe. Es sind deutlich Feldschläge zu erkennen, die größtenteils als einzelnes Segment erscheinen. Wesentlich für den nachgelagerten Klassifikationsprozess ist die semantische Korrektheit des Segmentierungsergebnisses. Der subjektive Eindruck zeigt eine sehr gute Übereinstimmung. Wie zu erwarten, zerfallen urbane Regionen in eine Vielzahl kleinerer Segmente (rote Markierung in den Abbildungen).

Abbildung 6: RGB-Komposit der Kanäle 1-3 einer Rapid-Eye-Szene (urbane Vergleichsregion rot markiert; Bereitstellung durch DLR 2009)

Abbildung 7: HH-Polarisation einer Terra-SAR-X-Szene © DLR 2009 (urbane Vergleichsregion rot markiert)

Abbildung 8: Segmentierungs-Ergebnis (urbane Vergleichsregion rot markiert)

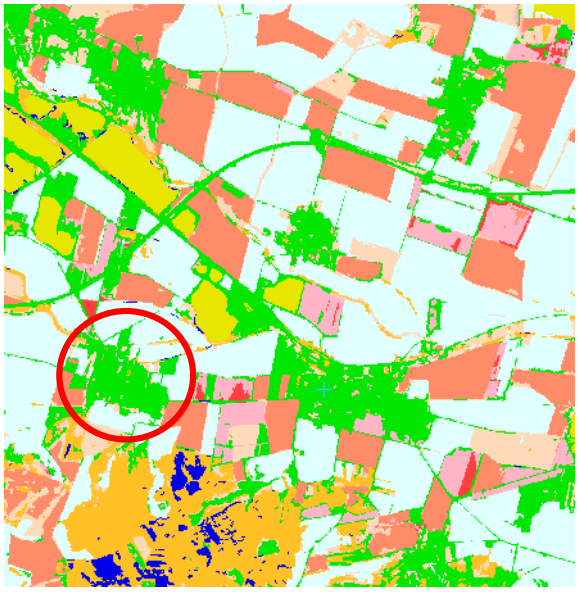

Abbildung 9: Segmentbasiertes Klassifikations-Ergebnis (SVM) (urbane Vergleichsregion rot markiert)

Die Laufzeit des Segmentierungsvorganges betrug 23 min (davon 65,25% Vorverarbeitung, 30% Region-Growing, 4,3% Region-Merging; prozessierte Pixel=17x2565x2958 Pixel; PC-System: WinXP64, QuadCore, 2,83GHz).
Die Bewertung der Segmentierungsgüte erfolgt hier auf der Ebene des Prozessierungsendproduktes „Landnutzungskartierung“ (Abbildung 9). Der zur überwachten segmentbasierten Klassifikation verwendete Merkmalssatz besteht zunächst aus einfachen Merkmalen der Segmentstatistik 1. Ordnung. Die Landnutzungsklassifikation mittels verschiedener etablierter Klassifikatoransätze (Support-Vector-Machine SVM, kNearestNeighbour kNN, RandomForest RF) ergibt für ein definiertes Szenario mit 9 Landnutzungsklassen die in Tabelle 1 genannten Gütewerte. In diesem Anwendungsfeld sind die Qualitätsmaße Overall-Accuracy (OA), für die Gesamtgüte der Klassifikatorentscheidung, und Kappa-Accuracy (KHAT), für die Determiniertheit der Klassifikatorentscheidung, etabliert [Tso09]. Um eine Vergleichbarkeit zwischen den Gütekennwerten der pixelbasierten (das heißt nicht segmentierten) und der segmentbasierten Klassifikation zu erreichen, wird der klassifizierte Flächenanteil der Segmente bei der Bewertung berücksichtigt.

Klassifikator	**seg**mentbasiert / **pix**elbasiert	**OA** [%]	**KHAT** [%]
SVM	**seg**	**82,96**	**79,06**
	pix	78,55	73,66
kNN	**seg**	**81,89**	**77,44**
	pix	76,44	71,15
RF	**seg**	**78,26**	**72,81**
	pix	77,69	72,63

Tabelle 1: Gegenüberstellung der Güten von segment- und pixelbasierter Klassifikation

Es wird deutlich, dass die segmentbasierte Klassifikation hinsichtlich ihrer Gütewerte bei allen getesteten Klassifikationsverfahren grundsätzlich besser als die pixelbasierte Klassifikation abschneidet.
Zum Rückschluss auf die Segmentierungüte über die erreichte Klassifikationsgüte dient der nachfolgende Falsifikations-Test. Hierbei wird ein semantisch nicht korrektes Segmentbild, das hinsichtlich der Segmentgrenzen nicht mit den Szenenobjekten korreliert, für eine erneute segmentbasierte Klassifikation herangezogen. Die Ergebnisse des Falsifikationstests sind in Tabelle 2 aufgeführt.

Klassifikator	**seg**mentbasiert	**OA** [%]	**KHAT** [%]
SVM	seg	63,05	54,74
kNN	seg	62,46	53,79
RF	seg	63,53	55,33

Tabelle 2: Ergebnisse des Falsifikations-Tests

Die erzielten Gütewerte sind deutlich niedriger als im vorhergehenden Experiment mit optimaler Segmentierung. Die Ursache liegt im Wesentlichen in der ungenaueren Klassenbeschreibung der resultierenden Segmentmischpixel. Es ist daher davon auszugehen, dass die Qualität der Segmentierung signifikanten Einfluss auf die Güte einer darauf aufsetzenden Klassifikation hat. Während der Umfang der Übersegmentierung für den verwendeten integralen Merkmalssatz nur eine geringe Bedeutung für die Klassifikationsgüte besitzt, ist bei Anwendung von erweiterten Segmentmerkmalen, z.B. zur Charakterisierung von Form und Berandung, diese zu berücksichtigen. Mit den vorliegenden nur gering übersegmentierten Segmentierungsergebnissen ist diese Erweiterung des Merkmalssatzes in jedem Fall möglich.

4 Zusammenfassung und Ausblick

Im vorliegenden Beitrag wurde ein automatisches mehrkanaliges Segmentierungsverfahren vorgestellt, das beliebig dimensionale bildhafte Eingangsdaten vollständig und semantisch korrekt segmentieren kann. Das Verfahren arbeitet im Gegensatz zu den bekannten Ansätzen determiniert. Es ist in der Lage große Datenmengen zeiteffizient zu verarbeiten. Aufgrund der nicht begrenzten Kanalanzahl und Bildgröße ist das vorgestellte Segmentierungsverfahren auch für Multi- oder Hyperspektralanwendungen im industriellen Kontext einsetzbar.
Obwohl im dargestellten Anwendungsfall das Potential des segmentbasierten Vorgehens noch nicht voll ausgeschöpft wurden, konnte dennoch gezeigt werden, dass die segmentbasierte Klassifikation Vorteile gegenüber der bislang zumeist nur pixelbasierten Klassifikation besitzt. Weitere Effekte lassen sich über die Erweiterung des Merkmalssatzes über abgeleitete Formmerkmale sowie Nachbarschaftsrelationen zwischen Segmenten erzielen.

5 Danksagung

Das ENVILAND-2 Projekt (Laufzeit 2009-2012) wurde durch die Raumfahrt-Agentur des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR) mit Mitteln des Bundesministeriums für Wirtschaft und Technologie aufgrund eines Beschlusses des Deutschen Bundestages unter den Förderkennzeichen 50EE0844 - 50EE0847 gefördert.
Die Bereitstellung der dem Ergebnis zugrundeliegenden RapidEye Daten wurde im Auftrag des Deutschen Zentrums für Luft- und Raumfahrt e. V. mit Mitteln des Bundesministeriums für Wirtschaft und Technologie sichergestellt. Die Verantwortung für das Ergebnis liegt beim Datennutzer. Wir danken dem DLR für die Bereitstellung von Daten aus dem RapidEye Science Archive und für die Bereitstellung von TerraSAR-X-Daten.

6 Literatur

[Quegan01] Quegan, S.: „Filtering of Multichannel SAR Images“, IEEE Transactions on Geoscience and Remote Sensing, Vol. 39, No. 11, Nov. 2001

[Koch2009] Koch, T.; Kubertschak, T.; Franke, K.-H.: "Aspekte der Segmentierung von Bildmaterial aus multispektralen Merkmalsräumen am Beispiel von Satellitendaten", Beitrag zum 15. Workshop Farbbildverarbeitung 2009, Berlin, 08.-09. Oktober 2009

[Lucht2009] Lucht, C.; Nestler, R.; Franke, K.-H.: "Landnutzungsklassifikation auf multisensorischen und multispektralen Bilddaten - Anforderungen und Lösungsansätze", Beitrag zum 15. Workshop Farbbildverarbeitung 2009, Berlin, 08.-09. Oktober 2009

[Koch2011] Koch, T.; Nestler, R.; Franke, K.-H.: "Verfahren zur Registrierung von mehrkanaligen Bildmaterial unterschiedlicher Modalität", Beitrag zum 17. Workshop Farbbildverarbeitung 2011, Konstanz , Markus Schnitzlein, Chromasens (Hrsg.), ISBN: 978-3-00-035834-0

[Tso09] Tso, B.; Mather, P.: Classification methods for remotely sensed data; CRC Press, Taylor & Francis Group, LLC; Boca Raton; ISBN 978-1-4200-9072-7; 2009