

Strukturierung von kontinuierlichen Rasterdaten mithilfe eines regionenbasierten Segmentierungsalgorithmus

Matthias Heinrich^a und Markus Möller^b

1. EINLEITUNG

Bodenkonturen als Träger von Bodeninformationen sind für eine Vielzahl planungsrelevanter Fragestellungen von besonderer Bedeutung (z.B. Bodenkartierung, Bodenfunktionsbewertung). Zur Abgrenzung von Bodenkonturen werden im zunehmenden Maße regionenbasierte Segmentierungsverfahren eingesetzt, die zu einer Strukturierung kontinuierlicher Flächendaten in homogene Flächenobjekte führen.

Das Ziel des Beitrages ist die Anpassung des frei verfügbaren Segmentierungsansatzes IVHG (Friedrich, 1996), dass speziell für bodenkundliche Fragestellungen entwickelt wurde.

1.1. Regionenbasierte Segmentierungsverfahren

Regionenbasierte Segmentierungsverfahren berücksichtigen bei der Objektbildung sowohl die n-Dimensionalität des Merkmalsraumes als auch die unmittelbare räumliche Nachbarschaft von Rasterzellen. Dabei wird typischerweise von einzelnen Rasterzellen (seed pixel) ausgegangen. Sie bilden die Kristallisationspunkte für ein iteratives Regionenwachstum, das solange andauert, bis bestimmte Abbruchbedingungen erfüllt sind. Als Abbruchkriterien dienen definierte Heterogenitätsbedingungen (z.B. Varianz) sowie Zwänge, die das Regionenwachstum stoppen (z.B. Anstoßen an eine andere Regionengrenze oder vom Nutzer vorgegebene Grenze) (Fortin et al., 2000; Muñoz et al., 2003) (Abb. 1).

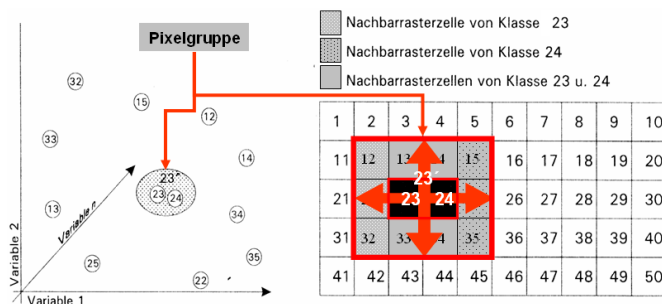


Abb. 1: Prinzip regionenbasierter Segmentierungsansätze (nach Friedrich, 1996, verändert). Die Rasterzelle 23 fungiert als Kristallisationspunkt, an die die Rasterzelle 24 gruppiert wird. Das neue Segment (23, 24) wird anschließend in das Segment 23 überführt, das als neues Kristallisationssegment für die Rasterzellen 12, 13, 14, 15, 22, 25, 32, 33, 34 und 35 dient.

^aWestfälische Hochschule Zwickau (FH), Fachbereich Physikalische Technik/Informatik, Fachgruppe Informatik, Dr.-Friedrichs-Ring. 2a, 08056 Zwickau, E-mail: matthias.heinrich@ufz.de, Telefon: ++49 (0) 341 2352586

^bUmweltforschungszentrum Leipzig-Halle, Department Angewandte Landschaftsökologie, Permoserstr. 15, 04318 Leipzig, E-mail: markus.moeller@ufz.de, Telefon: ++49 (0) 341 2352556

1.2. IVHG

Das Programm IVHG nutzt als Abbruchkriterium das Merkmal 'Regionenanzahl'. Das Abbruchkriterium kann durch die Gleichung (1) ausgedrückt werden. Da hier ein Bottom-Up-Ansatz vorliegt und somit zunächst jede Rasterzelle als Region betrachtet wird, werden hier die Anzahl der Regionen K zu der Anzahl der Rasterzellen M in Relation gesetzt. Dadurch ergibt sich der Generalisierungsgrad bezüglich des geometrischen Raums, der sich im Intervall $[0,1]$ abbilden läßt.

$$G_S = \frac{M - K}{M}, G_S \in [0, 1] \quad (1)$$

2. ANSATZ

Da G_S nicht die tatsächlichen Eigenschaften der Eingangsdaten berücksichtigt, wurde das folgende Maß entwickelt:

$$G_M = \frac{ed}{ed - \sum \sigma_d}, G_M \in [0, 1] \quad (2)$$

Das Hauptaugenmerk liegt hier darin, die Dateneigenschaften, insbesondere die globale Homogenität zu berücksichtigen. In das neu eingeführte Abbruchkriterium fließt deshalb die Ähnlichkeit der Rasterzellen im n-dimensionalen Merkmalsraum in Form der euklidischen Distanz ed (Abstand von Punkten im Raum) mit ein. Diese kann als Abstand bzw. Ähnlichkeit der Rasterzellen bzgl. ihrer Merkmale gesehen werden. Um sie als Abbruchbedingung zu verwenden, wird sie zur Standardabweichung σ der Eingangsdaten (Maß für die Heterogenität) ins Verhältnis gesetzt. Damit ist ein Maß für die Generalisierung hinsichtlich der Merkmale gegeben, welches sich ebenfalls im Intervall $[0,1]$ abbilden läßt.

Anhand der Gegenüberstellung des Verhaltens der beiden Generalisierungsgrade soll der Vorteil des neu eingeführten Maßes G_M dargelegt werden. In der Abbildung 2 wurden die Generalisierungsgrade auf die erzeugten Regionen aufgetragen:

- Für G_S ergibt sich demnach ein linearer Verlauf, der unabhängig von den Daten immer linear sein wird.
- G_M verhält sich dagegen eher entsprechend den Dateneigenschaften, da dieser die euklidische Distanz berücksichtigt. Der flache Verlauf bei einer hohen Regionenzahl deutet auf relativ homogene Daten hin, welche grundsätzlich zusammen gefasst werden können. Somit greift G_M erst da ein, wo die Homogenität signifikant steigt und die Regionen untereinander heterogener werden.

Die Verwendung des Intervalls $[0,1]$ hat den Vorteil, das die Freiheitsgrade für den Benutzer eingeschränkt

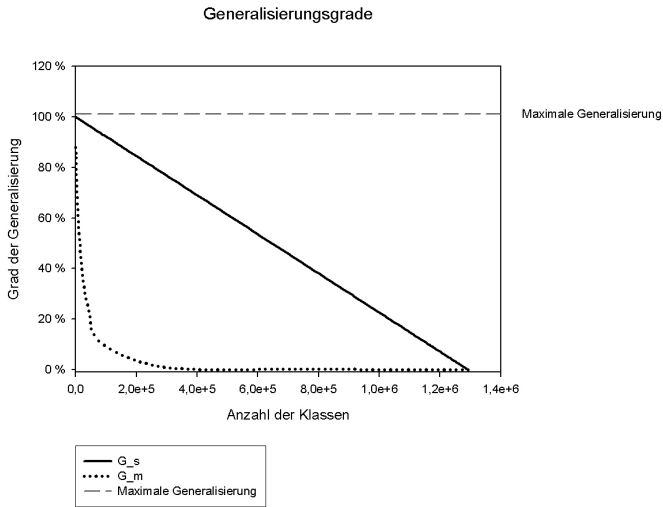


Abb. 2: Gegenüberstellung der Generalisierungsgrade

werden. Dadurch ist die Entscheidung des Nutzers hinsichtlich des Abbruchkriteriums datenunabhängig und damit vergleichbar. Zum anderen besteht dadurch die Möglichkeit, die Abbruchkriterien G_S und G_M zu kombinieren. Eine mögliche Kombination sieht wie folgt aus:

$$\begin{aligned}
 G &= w_S \times G_S + w_M \times G_M, & (3) \\
 w_S &\in [0, 1], \\
 w_M &\in [0, 1], \\
 w_S + w_M &= 1
 \end{aligned}$$

Dem Anwender steht es somit frei, mit Hilfe der Gewichte w_S und w_M das Regionenwachstum durch beide Generalisierungsgrade zu steuern.

3. AUSBLICK

Das Augenmerk lag bislang auf globalen Eigenschaften zur Steuerung des Regionenwachstums, allerdings ist deren Einfluss begrenzt. Es scheint daher sinnvoll, Objekteigenschaften (z.B. Homogenität, Form usw. der Regionen) als Steuerparameter zu verwenden. Weiterhin wird angestrebt, den Rechenaufwand zu minimieren (Speicherbedarf) und die Ergebnisausgabe nutzerfreundlicher zu gestalten.

Zukünftig soll das Programm auf kontinuierliche Rasterdaten (Fernerkundungs- und Satellitenbilder) insbesondere zur Detektion von Senkenbereichen und Acker schlägen angewendet werden.

Der Algorithmus wurde als C++ Programm nach der GNU GPL und somit als Open Source Projekt umgesetzt. Das Verfahren ist somit frei verfügbar. Außerdem bietet Open Source die Möglichkeit, das Verfahren kostengünstig weiter zu entwickeln. Er liegt als Quellcode vor und ist Betriebssystem-unabhängig.

Literatur

Fortin, M.-J., Olson, R., Ferson, S., Iverson, L., Hunsaker, C., Edwards, G., Levine, D., Butera, K., & Klemas, V. (2000). Issues related to the detection of boundaries. *Landscape Ecology*, 15, 453–466.

Friedrich, K. (1996). Digitale reliefgliederungsverfahren zur ableitung bodenkundlich relevanter flächeneinheiten. *Frankfurter Geowissenschaftliche Arbeiten, Serie D*, 21. Dissertation, Universität Frankfurt.

Muñoz, X., Freixenet, J., Cufi, X., & Martí, J. (2003). Strategies for image segmentation combining region and boundary information. *Pattern Recognition Letters*, 24, 375–392.