

Aufgabenorientierte Datenklassifikation in Choroplethen-Karten für die Erhaltung lokaler Extremwerte

Juiwen Chang & Jochen Schiewe

HafenCity Universität Hamburg, Labor für Geoinformatik und Geovisualisierung (g2lab)

{juiwen.chang, jochen.schiewe}@hcu-hamburg.de

Zusammenfassung

In thematischen Karten werden die Werte einzelner Regionen zur besseren Übersicht oft in verschiedene Klassen eingeteilt und entsprechend farblich kodiert, um einen besseren Überblick zu erhalten. Der visuelle Eindruck solcher Choroplethen-Karten wird dabei durch die gegebene Werteverteilung, die Methode zur Datenklassifikation und die Farbwahl bestimmt. Konventionelle Klassifizierungsmethoden funktionieren datenbasiert und beachten dabei nicht den räumlichen Kontext der Daten. Beispielsweise geht die Information über einen lokalen Extremwert verloren, wenn er auch nur teilweise in die gleiche Kategorie wie seine Umgebung eingeteilt wird. Anstelle der konventionellen datenbasierten Ansätze verfolgen wir nun den umgekehrten Weg - eine aufgabenorientierte Klassifikation.

Beschreibung

Für die Erhaltung von Polygonen, die lokale Extremwerte darstellen, wurde folgendes Verfahren implementiert und getestet.

1. Identifizierung lokaler Extremwerte: Ein lokaler Extremwert ist als Polygon definiert, welches größere (oder kleinere) Werte im Vergleich zu all seinen Nachbar-Polygonen aufweist. Als Nachbarn werden die Polygone definiert, die eine gemeinsame Grenze oder einen gemeinsamen Eckpunkt aufweisen. Um die Suche nach Nachbarn zu beschleunigen, wird ein räumlicher Index basierend auf einem *R-Tree* verwendet. Unter all den Nachbarn eines lokalen Extremwert-Polygons wird dasjenige ausgewählt und gespeichert, welches die minimale absolute Differenz in den Attributwerten aufweist.

2. Sortierung aller Extremwert-Polygone entsprechend ihrer Signifikanz: Die sich aus dem ersten Schritt ergebenden Werte-Intervalle (abgegrenzt durch die Attributwerte des Extremwert-Polygons und dem numerisch nächsten Nachbarn) werden nach ihrer Intervall-Breite sortiert.

3. Bestimmung der Klassenbegrenzung: Um optimale Schnittpunkte zwischen den Intervallsegmenten zu erhalten, wird ein *Plane Sweep Algorithmus* angewendet: Eine vertikale Linie wird vom globalen Minimum zum globalen Maximum geschoben, dabei wird die Anzahl der geschnittenen Intervalle gezählt. Der Wert, der die maximale Anzahl von Schnittpunkten aufweist, definiert die erste Klassenbegrenzung. Nach Ausschluss der Segmente, die hiermit geschnitten werden, wird das Verfahren wiederholt, um auf die gleiche Weise die zweite und alle weiteren Klassen zu erhalten, bis eine vorgegebene Anzahl an Klassen oder alle Intervalle erreicht wurden.

Offensichtlich sind alle Datenklassifizierungsmethoden in der Lage, alle Extremwert-Polygone mit einer großen Anzahl an Klassen zu erfassen (im Extremfall ist die Anzahl der Klassen identisch mit der Anzahl der Extremwerte). Erste empirische Tests haben gezeigt, dass unser Ansatz jedoch eine signifikant geringere Anzahl an Klassen benötigt, um das gewünschte Ziel der Erhaltung von Extremwert-Polygonen zu erreichen.