

Selbstorganisierende Systeme – 4.Übung

Christian Gruber, 0625102 und Johannes Reiter, 0625101

Automatische Analyse von Metro Maps

1. How-To

Wir haben unsere Aufgabe gleich direkt in den Source-Klassen der SOM Toolbox implementieren können, wodurch sich für das Starten des Programms nichts geändert hat.

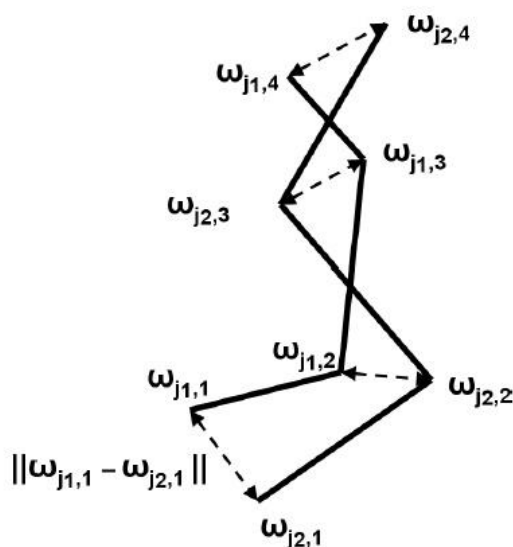
Es bleibt bei den Standardparametern: `./somtoolbox.sh SOMViewer -u /path/to/file.unit.gz -w /path/to/file.wgt.gz --dw /path/to/file.dwm.gz`

2. Dokumentation

Unser Modul zur automatischen Analyse und Interpretation von Metro Map Visualisierungen umfasst das Erkennen von Korrelationen zwischen Attributen, die Ausgabe vom Verlauf von Attributen und die Bestimmung der Extrema von Attributen. Diese drei Erweiterungen erlauben es weitere Rückschlüsse auf die Attribute und somit auch auf die zugrundeliegenden Daten zu ziehen.

2.1. Korrelation von Attributen

Für die Berechnung der Korrelation von zwei Attributen haben wir uns an das Paper „The Metro Visualisation of Component Planes for Self-Organising Maps“ von Robert Neumayer, Rudolf Mayer, Georg Pözlbauer und Andreas Rauber gehalten. Die folgende Grafik, die aus diesem Paper entstammt, zeigt schon gut, nach welchem Prinzip hier vorgegangen wird.



Zuerst berechnet man sich die Summe der Distanzen zwischen den einzelnen Bins, wobei man hierbei an beiden Enden einmal starten muss, um dann das Minimum verwenden zu können. Die erhaltene Distanz der beiden Metro Lines vergleichen wir nun mit zwei verschiedenen Limits (c_{weak} , c_{strong}), die wir nach unserer Evaluierung als sinnvolle Grenze für eine schwache bzw. eine starke Korrelation erachtet haben. Dadurch kann unsere automatische Metro Map Analyse auch den Unterschied von eher schwach oder eher stark korrelierenden Attributen erkennen.

$$c_{weak} = \frac{map_{height} + map_{length}}{2} * |bins| * 0.5$$

$$c_{strong} = \frac{map_{height} + map_{length}}{2} * |bins| * 0.33$$

Falls nun die zuvor berechnete Distanz zwischen den beiden Metro Lines kleiner als c_{strong} ist, bezeichnen wir die beiden Attribute als stark korrelierend und falls sie kleiner als c_{weak} ist, als schwach korrelierend. Nachdem in die Berechnung der Limits auch die Map-Größe und die Anzahl der Bins eingehen, sollten sie weitgehend unabhängig von deren Änderungen sein.

2.2. Verlauf von Attributen

2.3. Extrema von Attributen

3. Evaluierung

