

# Kartographische Mustererkennung

## rasterorientierte Verfahren zur Erfassung von Geo-Informationen

**Doctoral Thesis**

**Author(s):**

Stengele, Roland Eduard

**Publication date:**

1995

**Permanent link:**

<https://doi.org/https://doi.org/10.3929/ethz-a-001442656>

**Rights / license:**

In Copyright - Non-Commercial Use Permitted

Diss. ETH Nr. 11099

# **Kartographische Mustererkennung**

**Rasterorientierte Verfahren zur  
Erfassung von Geo-Informationen**

**ABHANDLUNG**

zur Erlangung des Titels

**DOKTOR DER TECHNISCHEN WISSENSCHAFTEN**

der

**EIDGENÖSSISCHEN TECHNISCHEN HOCHSCHULE ZÜRICH**

vorgelegt von

**ROLAND E. STENGELE**

Dipl.-Ing., Universität Karlsruhe (TH)

geboren am 06. Dezember 1959

aus Deutschland

Angenommen auf Antrag von:

Referent Prof. Dr. A. Carosio, Institut für Geodäsie und Photogrammetrie, ETH Zürich

Korreferent Prof. E. Spiess, Institut für Kartographie, ETH Zürich

1995

## Zusammenfassung

---

Geo-Informationssysteme gelten als eine Schlüsseltechnologie der Zukunft und sind bereits heute in vielen Anwendungsbereichen ein unentbehrliches Instrument zur Erfassung, Verwaltung, Verarbeitung und Präsentation raumbezogener Daten. Die Erfassung raumbezogener Daten ist in aller Regel zeit-, personal- und damit kostenintensiv. Es liegt daher nahe, den Informationsgehalt der in grosser Vielfalt vorliegenden analogen Karten- und Planwerke für die digitale Technologie der Geo-Informationssysteme zu erschliessen. Insbesondere topographische Kartenwerke sind aufgrund ihrer neutralen Geometrie- und Sachaussage als Basisinformation für viele GIS-Anwendungen geeignet.

Die Scannertechnologie bietet die Grundlage für die automatische Erfassung analoger Speichermedien an. Eine über die Visualisierung hinausgehende Informationsverarbeitung erfordert die analytische Behandlung von Rasterbildern, um deren logische Bildinhalte zu extrahieren. Diese komplexe Aufgabe lässt sich automatisieren, indem Computerprogramme die menschliche Interpretationsfähigkeit bei der optischen Sinneswahrnehmung simulieren. Aktivitäten in diesem Gebiet konzentrieren sich auf die wissenschaftliche Disziplin der Mustererkennung, einem Teilgebiet der Künstlichen Intelligenz.

Die vorliegende Arbeit gibt einen Überblick über Techniken der Mustererkennung zur Interpretation kartographischer Vorlagen, nachfolgend als Kartographische Mustererkennung bezeichnet. Die Charakteristiken dieses Anwendungsgebietes und das Studium bekannter Arbeiten führten zur Erkenntnis, dass vermehrt rasterorientierte Ansätze in den Erkennungsprozess zu integrieren sind. Bisherige Ansätze gehen zumeist von einer Konversion des Rasterbildes in Vektorform aus und klassifizieren Muster aufgrund topologischer und metrischer Merkmale, die aus Vektordaten abgeleitet werden. Unvermeidbare Informationsverluste bei der Konversion wirken sich jedoch nachteilig auf den Musteranalyseprozess aus.

Der Vergleich zweier Rasterbilder wird in der Digitalen Bildverarbeitung durch Korrelationsmodelle realisiert. Beim speziellen Korrelationsverfahren des Template Matchings wird ein synthetisch erzeugter, in Rasterform vorliegender Vertreter einer Musterklasse in einem Rasterbild gesucht. Hierbei wird für jede Position ein Ähnlichkeitsmass zwischen diesem Template und dem von ihm überdeckten Teil des Bildes berechnet. Um das Potential dieses Verfahrens für die Kartographische Mustererkennung freizulegen, muss es in zweierlei Hinsicht optimiert werden:

- Höhere Erkennungsraten ergeben sich durch eine Ähnlichkeitsberechnung, die sich gegenüber kleinen Variationen der Muster einer Musterklasse (Grössenunterschiede, Verkantungen, Variation der Strichstärken etc.) robust verhält.
- Akzeptable Rechenzeiten werden nur durch eine streng hierarchische Strategie erzielt.

Von zentraler Bedeutung für beide Optimierungsziele ist ein gewichteter Ansatz zur Ähnlichkeitsberechnung. Der Ansatz basiert auf der Idee, dass die einzelnen Pixel eines Templates von unterschiedlicher Bedeutung für die Form eines Musters und damit für dessen Identifikation sind. Durch die Zuordnung unterschiedlicher Gewichte werden hoch-signifikante, signifikante und unbedeutende Pixel des Vorder- und Hintergrundes unterschieden. Diese Zuordnung kann ein Anwender durch unterschiedliche Farbgebung mit einem Rastereditor vornehmen. Schwer formulierbares Wissen über die spezielle Form eines Musters und über die feinen Unterschiede zu ähnlichen Mustern kann somit auf graphische Art und Weise sehr effizient in den Erkennungsprozess eingebracht werden. Die Bibliothek aller verfügbaren Templates stellt mehr als einen reinen Datenbestand dar, sondern repräsentiert Wissen in einer graphischen Darstellungsform. Dieses wissensbasierte Template Matching eignet sich beispielsweise zur Extraktion aller horizontal ausgerichteten Kartenschriften und in umfangreichen Stichproben wurde eine Erkennungsrate von 95.8% in der topographischen Karte 1:25'000 des Bundesamtes für Landestopographie erzielt. Auch beliebige kartographische Signaturen können zuverlässig lokalisiert werden, vor allem auch dann, wenn diese andere Graphikelemente berühren oder überlappen und keine isolierten Elemente der Kartengraphik sind.

Verschiedene lokale Rasteroperationen werden gezielt kombiniert, um die in schwarzem Vollton gehaltenen Gebäude- und Siedlungsflächen in einer topographischen Karte zu extrahieren. In einem typischen Kartenblatt im Massstab 1:25'000 können 98% der ca. 10'000 Gebäude automatisch erfasst werden.

Schwerpunkt eines Forschungsprojekts zur Kartographischen Mustererkennung war die Realisierung eines rasterorientierten Programmsystems, wobei einige spezielle Randbedingungen zu berücksichtigen waren. Anhand mehrerer potentieller Anwendungen im Umfeld von Geo-Informationssystemen und Kartographie wird das vielfältige Einsatzspektrum angedeutet: Datenakquisition zum Aufbau landesweiter Informationssysteme (z.B. Datenbank geographischer Namen, Projekt Vektor25 des Bundesamtes für Landestopographie), Unterstützung kartographischer Arbeitsabläufe (z.B. Nachführung, Schriftplatzierung) oder gezielte Datenerfassung für verschiedenste GIS-Anwendungen (z.B. Dokumentation des Landschaftswandels, Objektmodellierung zur Erzeugung von 3D-Landschaftsbildern).

## Abstract

---

Geographic Information Systems (GIS) is a key technology in future. Even today this technology is indispensable to a lot of applications for acquisition, management, processing and presentation of spatial data. Usually the acquisition of spatial data is time- and staff-consuming and thus an expensive process. Therefore it seems reasonable to open up the information of existing analogue maps and schemes available in great variety for modern GIS-technology. Due to their neutral geometry and thematic contents topographic maps are particularly suitable as basic information for many applications.

The scanner technology offers the basis for automating the acquisition of analogue storage media. Management of spatial data beyond the task of pure visualization requires analytic image processing to extract the logical features. This complex problem can be automated by simulating the capabilities of human perception with computers. These activities are focused on the discipline of pattern recognition, a branch of artificial intelligence.

This paper presents an overall view of pattern recognition techniques suitable for cartographic products, subsequently designated as cartographic pattern recognition. A study on the characteristics of this application and the investigations done so far resulted in the requirement to integrate more and more raster-based tools into the recognition-process. Previous approaches are mainly based on a raster-to-vector conversion of images and the patterns are classified by topological and metric features derived from vector-data. However, the analysis of patterns is affected adversely by unavoidable loss of information due to this conversion.

In the field of image processing two raster images are compared by correlation techniques. Template matching is a special correlation technique and describes the search for patterns in the raster image by calculating similarity of the covered image part and a synthetic prototype of a feature class at each position. This technique has to be optimized with regard to two aspects for exposing its potential for cartographic pattern recognition:

- Higher recognition rates are obtained by calculating similarity, which is robust against small variations of patterns belonging to the same feature class (different size, orientation, thickness of lines etc.).
- Acceptable calculating periods can only be achieved by a hierarchic strategy.

The measurement of similarity with different weights is of vital importance in regard to both objectives. The basis of this approach is the variable importance of individual pixels for the pattern shape and its identification. High-significant, significant and insignificant pixels of foreground and background are distinguished. The user can assign the weights to the pixels by editing different colours. In this way the acquisition of knowledge describing the characteristics of a

template and the small differences between similar templates can be done easily and very efficiently in a graphic manner.

The library of all available templates represents a graphic-designed knowledge base and is far more than a pure data base. For instance the method of knowledge-based template matching is suitable for recognition of map text in horizontal orientation. Extensive random samples proved recognition rates of 95.8% for topographic maps of the Swiss Federal Office of Topography on a scale of 1:25'000. Furthermore any cartographic symbols can be located reliable. Above all this is practicable even for symbols which are not isolated from the graphic context, but also touch or overlap other graphic elements.

Various local raster operators are specifically combined in order to extract the black-coloured signature for buildings in topographic maps. 98% of approximately 10'000 buildings can be recognized automatically in a typical map sheet on a scale of 1:25'000.

The development and implementation of a raster-based system for cartographic pattern recognition was a focal research topic. Some special requirements had to be considered. The variety of possible applications for GIS and cartography is indicated by means of several examples: Data acquisition for the establishment of nationwide spatial information systems (e.g. data base of geographic names, project Vector25 of the Swiss Federal Office of Topography), support for cartographic tasks (e.g. map revision, placement of map text) and acquisition of specific data for various GIS-projects (e.g. to document the change of landscape, modelling of objects for generating 3D-images of landscape).