



Doctoral Thesis

## Stochastic modeling of image content in remote sensing image archives

**Author(s):**

Schröder-Brzosniowsky, Michael

**Publication Date:**

2000

**Permanent Link:**

<https://doi.org/10.3929/ethz-a-003851205> →

**Rights / License:**

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

Diss. ETH No. 13458

# Stochastic Modeling of Image Content in Remote Sensing Image Archives

A dissertation submitted to the  
**SWISS FEDERAL INSTITUTE OF TECHNOLOGY**

for the degree of  
Doctor of Technical Sciences

presented by

**Michael Schröder-Brzosniowsky**

Dipl. Phys., University of Ulm

born March 21, 1971

nationality: German

accepted on the recommendation of  
Prof. Dr. Luc Van Gool, examiner  
Prof. Dr. Mihai Datcu, co-examiner  
Prof. Dr. Wolfgang Förstner, co-examiner

Zurich, February 3, 2000

# Abstract

In this dissertation, a concept of stochastic modeling of image content in remote sensing image data is presented. The main application of this concept is to provide a tool for content-based query and retrieval from large archives—a tool that becomes more and more necessary with the continuously increasing volume of remote sensing and other multimedia image data. The concept is organized in such a way that the (time consuming) process of information extraction is completely application-free and attempts to capture characteristic structures in the data in terms of a family of signal models. This un-supervised image content description is subsequently used as the basis for the fast definition of user-specific cover-types and the easy incorporation of applications. Furthermore, the fast generation of thematic maps in terms of user-specific cover-types is possible using the un-supervised content description.

The actual process of information extraction from the data is based on the repeated application of stochastic modeling in a Bayesian framework. The extracted content information is arranged in a hierarchical scheme according to the different levels of semantic abstraction: the image data (level 0), image features (level 1), meta features (level 2), image classification (level 3), and user-specific semantics (level 4). Between each level of this hierarchy, stochastic modeling using parametric data models is applied.

Three main steps of information extraction can be defined that are all modeled in the same Bayesian manner: First, spatial structures in the image data itself are modeled using Gibbs random field models (levels 1 and 2). These models are defined via parametrized energy functions that describe different spatial structures through their mathematical form, the selected neighborhood pixels, and the values of their parameters. In the presented Bayesian modeling, the most evident model with its most evident parameters and complexity can be selected, thus providing an “optimal” description of the image data.

Second, characteristic signal classes in the feature space of the image data are found via un-supervised classification (level 3). Two probabilistic algorithms are discussed in more detail that are capable of providing a classification in terms of an a priori unknown number of classes: “clustering by melting” and “Bayesian classification”. These algorithms have significantly larger computational demands than standard pattern recognition algorithms, but are able to automatically detect the number of classes and to describe highly un-balanced clusters. One of the two al-

gorithms, Bayesian classification, is additionally able to model overlapping clusters, and is invariant with respect to a re-scaling of the parameter space. Furthermore, the concept of across-image segmentation and two concepts for assessing image complexity are presented.

Third, user-specific interests are linked to the un-supervised image content description using Bayesian networks (level 4). Application-specific cover-type labels are probabilistically defined using the signal classes derived from different models. Based on user-provided examples of a particular cover-type label, the probabilities of the Bayesian network are trained or, for an already existing label, updated. This results in a very convenient way of sequential interaction between the database system and the user.

Due to the stochastic nature of the cover-type definitions, the retrieval of images of interests from an archive of images is also possible in a probabilistic way. Thereby, images can be retrieved according to coverage, posterior probability, and separability. The latter, the “separability”, encodes the uncertainties resulting from limited training data and enables the user to continue defining his personal cover-type on images on which the posterior probability is not yet precisely defined. The size of the content-based, probabilistic search index is less than 0.1% of the overall image data size, resulting in a very fast search and making this concept applicable to very large archives.

The concept of this dissertation and the different steps of Bayesian inference have been successfully implemented in the ETHZ/DLR MMDEMO system that provides content-based access to a test archive of about 360 Landsat TM, X-SAR, and aerial images. The complete demonstration system, including interactive cover-type definition and probabilistic retrieval, can be accessed via the Internet [<http://www.vision.ee.ethz.ch/~rsia/mmdemo>].

# Zusammenfassung

Diese Dissertation stellt ein Konzept für die stochastische Modellierung von Bildinhalten in Fernerkundungsdaten vor. Hauptanwendung ist dabei, inhaltsorientierte Techniken für die Abfrage von großen Fernerkundungsarchiven zur Verfügung zu stellen. Derartige inhaltsorientierte Abfragetechniken werden aufgrund der dramatisch anwachsenden Bilddatenbestände in Fernerkundungs- und anderweitigen Multimediadatenbanken immer dringender benötigt. Das Konzept dieser Arbeit ist derart aufgebaut, dass der (zeitaufwendige) Prozess der Informationsextraktion aus den Daten nicht auf eine bestimmte Anwendung hin ausgerichtet ist. Stattdessen wird versucht, unüberwacht mit einer Reihe von Datenmodellen charakteristische Strukturen in den Daten zu erkennen. Aufbauend auf dieser Beschreibung können mit relativ wenig Rechenaufwand benutzerspezifische Bodenbedeckungstypen definiert und bestimmte Anwendungen integriert werden. Zudem können vergleichsweise schnell thematische Karten mit den Benutzerinteressen erstellt werden.

Als zentrales Element kommt bei der Informationsextraktion Bayessche Statistik zur Anwendung. Dabei wird die extrahierte Information über den Bildinhalt in einem hierarchischen Schema angeordnet, welches die verschiedenen Abstraktionsniveaus widerspiegelt: die Bilddaten (Niveau 0), Bildmerkmale (Niveau 1), Metamerkmale (Niveau 2), Bildklassifikation (Niveau 3) und die benutzerspezifische Interpretation (Niveau 4). Zwischen diesen Niveaus kommt jeweils stochastische Modellierung mit parametrischen Modellen zur Anwendung.

Drei wesentliche Schritte der Informationsextraktion mit ähnlicher Bayesscher Beschreibung können identifiziert werden. Erstens, die Beschreibung von räumlichen Strukturen in den Bilddaten mittels Gibbsschen Zufallfeldern (Niveaus 1 und 2). Diese Modelle beschreiben die Daten mittels einer parametrischen Energiefunktion, die durch ihre Form, die ausgewählten Nachbapixel und die Werte ihrer Parameter verschiedene räumliche Strukturen beschreiben kann. Mittels der vorgestellten Bayesschen Modellierung können die für gegebene Daten optimal passenden Charakteristika bestimmt werden.

Zweitens, in einem Schritt unüberwachter Klassifikation werden charakteristische Signalklassen bestimmt. Dabei werden zwei probabilistische Algorithmen näher erläutert, die in der Lage sind, eine Klassifikation mit von vorne herein unbekannter Anzahl Klassen durchzuführen: „Clustern durch Schmelzen“ (englisch “clustering by melting”) und „Bayessche Klassifikation“. Im Vergleich zu Standardverfahren

der Mustererkennung sind diese Algorithmen vom Rechenaufwand viel aufwendiger, bieten jedoch eine bessere Beschreibung der charakteristischen Signalklassen in den Daten. Ausgehend von der Klassifikation werden auch die „Klassifikation über alle Bilder“ und zwei Maße für die Bildkomplexität vorgestellt.

Drittens, als letzter Schritt werden die Benutzerinteressen mittels Bayesscher Netzwerke mit der unüberwachten Bildbeschreibung verknüpft. Aufbauend auf den Signalklassen verschiedener Modelle können anwendungsspezifische Bodenbedeckungstypen stochastisch definiert werden. Mittels Benutzerbeispielen werden dabei die Wahrscheinlichkeiten eines bestimmten Types gelernt bzw., falls der Begriff bereits existiert, verfeinert. Dies resultiert in einem sehr intuitiven und benutzerfreundlichen Lernverfahren.

Da die Bodenbedeckungstypen stochastisch definiert werden, ist es auch möglich, die Menge aller Bilder stochastisch nach einem Typ abzufragen. Dabei können Bilder nach Bedeckung, nach Wahrscheinlichkeit oder nach Separabilität abgefragt. Letztere repräsentiert dabei die Unsicherheit, die aus der geringen Anzahl von Benutzerbeispielen resultiert. Die Separabilität unterstützt den Benutzer auch bei der Definition von Interessen, indem sie auf Bilder verweist, auf denen der Begriff noch nicht eindeutig definiert ist. Die Größe des inhaltsorientierten, probabilistischen Indizes ist kleiner als 0.1% der Bilddaten, was eine schnelle Suche und die Anwendung auf sehr große Archive ermöglicht.

Das Konzept dieser Dissertation und die verschiedenen Schritte Bayesscher Folgerungen wurden im ETHZ/DLR MMDEMO System erfolgreich implementiert. Dieses Demonstrationssystem bietet inhaltsorientierten Zugang in der hier beschriebenen Weise zu insgesamt 360 Landsat TM, X-SAR und Luftbildern. Es ist über das Internet für jedermann zugänglich [<http://www.vision.ee.ethz.ch/~rsia/mmdemo>].