AN INTEROPERABLE WEB-BASED CLASSIFICATION SERVICE FOR REMOTE SENSING DATA

Meixia Deng, Yang Liu, Liping Di
George Mason University

ABSTRACT

Most users of remote sensing data want to get the ready-to-use information that tailors to their individual needs, rather than the raw data quickly and easily. One of the promising approaches for fulfilling users' such requirements is from the Web through dynamically chaining data with services and then executing the service chain under the Web service environment. One of the key components for success of such an approach is to have large numbers of standard-based interoperable services available over the Web for dynamically constructing service chains. This paper presents an experiment for building a web-based image classification service that provides both supervised and unsupervised classification algorithms. The service is implemented in Java with both http POST and SOAP binding. The service takes the coverage data as input and produces the classification coverage. The Open GIS Consortium (OGC) Web Coverage Service (WCS) interface is provided for the service. The service is described in the web services description language (WSDL) and registered via OGC web registration service (WRS). This implementation provides us the experience for building other service modules to be part of a large intelligent web-based interoperable system for geospatial information and knowledge services.

KEY WORDS: service chaining, interoperable, coverage, supervised and unsupervised classification, WCS, WSDL, WRS

1. INTRODUCTION

NASA and other space agencies have collected huge amount of Earth remote sensing data. Those data are valuable not only to scientific research but also to socio-economic activities and education. Currently those data are available to users in raw forms and users need to spend significant amount of time and resources to process the data for deriving useful information. With advancement in both computer hardware and software, more and more data providers are putting data on-line for users to direct access the data through the standard interfaces, such as Open GIS Consortium's (OGC) Web Coverage Service (WCS) [Evans, 2002] and Web Feature Service (WFS) [Vretanos, 2002]. The research and standards on semantic web and web service technologies make the coupling of diverse data with services, and services with services possible. It is envisioned that in the near future users of remote sensing data will be able to get the requested information that tailors to their individual needs, rather than the raw data, from the Web through dynamically chaining data with services and executing service chains over the Web. One of the keys for realizing such a vision is to have large numbers of standard-based interoperable services available over the Web for dynamically constructing service chains for fulfilling a range of user requests. This paper describes a web-based coverage-data classification service based on OGC and World-Wide Web Consortium (W3C) service standards [Lieberman, 2003]. The classification algorithms available in the service include both supervised maximum likelihood and unsupervised cluster classifiers. The service takes the multi-spectral remote sensing coverage data as

input and produces the result in the classified coverage. The OGC Web Coverage Service (WCS) interface is provided for the service. The service is described in the Web Services Description Language (WSDL) and registered via OGC (Open GIS Consortium) Web Registration Service (WRS). All users in the web community can instantly and dynamically use this service under authorization. The successful implementation of the web-based coverage classification service provide us the experience for building other service modules to be part of a large intelligent web-based interoperable system for geospatial information and knowledge services.

2. Supervised & Unsupervised Classification

Image classification is one of the most popular methods of information extraction. Classically, there are two major approaches for classification, supervised and unsupervised. A variety of algorithms are available for both approaches [Jensen, 1996]. In a supervised classification, we need to have data with known classes in advance, commonly referred to as training sites, to train the classification algorithm. Statistical parameters (means, standard deviations, covariance matrices, correlation matrices, etc.) are calculated for each training site. After the training, the classifier is called the trained classifier. Then, the classifier will be used to classify images. Each pixel in the to-be-classified image is then evaluated and assigned to the class of which it has the highest likelihood of being a member. In an unsupervised classification, the computer uses algorithms to automatically determine the similarity of the spectral signatures and group pixels with similar spectral characteristics, based on some known statistically determined criteria that are coded into the software. Unsupervised classifications require a minimum user interaction (e.g., the user only provides the minimum and maximum number of classes that he is expecting from the classification service).

3. OGC Classification Service

The OGC Classification Service uses the getCoverage request of the Web Coverage Service. The Classification Service is defined here using stateless request/response [Percivall & Smits, 2001], which was based on the following premises:

- The input and output of a Classification Service are Coverages, therefore Coverage services should be reused in defining the classification service.
- Stateless services using a request/response operations are desired, therefore the unsupervised classification is defined as a single stateless operation and supervised classification is defined with two stateless operations: a training operation and a coverage request operation.
- The approach to reuse of interfaces and chaining of services employed by the Coverage Portrayal Service is used here, i.e., request to a CPS is a WMS getMap request with the added SLD information of styling parameters and the URL of the coverage to be portrayed.

In the last several years, OGC has been working on technologies for enabling interoperable geospatial web services. Coverages are the topic of several existing OGC technologies, including (1) OGC Abstract Specification – Topic 6: Coverages (Topic 6 contains ISO 19123 and additional material), (2) OGC Grid Coverages Implementation Specification, (3) OGC Web Coverage Service Implementation Specification.

An approach to chaining of services in OGC and the reuse of interfaces is found in the use of a Styled Layer Descriptor in the Web Mapping Service used by a Coverage Portrayal Service. These topics are described in the following specifications as, (1)OGC Web Mapping Service Implementation Specification, Version 1.1.1, (2) OGC Styled Layer Descriptor Discussion Paper, (3) OGC Coverage Portrayal Service Interoperability Program Report.

OGC is developing a service model to support chaining of geographic services. The relevant OGC specifications are as follows, (1) OGC Abstract Specification – Topic 12, OGC Service Architecture, (2) OGC Web Services Architecture Interoperability Program Report. All above-mentioned specifications can be found in the OGC web site.

4. A PROPOSED CLASSIFICATION SERVICE STRUCTURE

Based on OGC and W3C service standards, we propose a web-based classification service structure as Figure 1.

The bottom part is **Image Classification Core Class**. The basic classification algorithm is written in C. A Java class – ImageClassifier – provides java access to those methods through the Java Native Interface. Image Classification Core Class provides 3 functions, supervisedClassify, unsupervisedClassify and training. This Core Class will be invoked by the **Web Image Classification Service (WICS)**.

WICS consists of two different implementations, the HTTP GET & POST version and the SOAP version. Both versions provide three operations, GetCapabilities, GetCoverage, TrainClassifier (supervised classification). Both versions accept the same parameters. The only difference is the protocol of request and response. The HTTP version is very similar to OGC WCS.

The **WICS web interface** is a web application, including a set of html pages and a Java servlet. To the **WICS**, this web interface is a client. User can use these web pages to submit a WICS request to the service (HTTP version) and specify whether he wants the raw data back or a picture result. The web interface will generate pictures for user if requested. Also user can use other client tools to connect to both versions of service.

5. The WSDL Descriptions of the Web Image Classification Service

As we mentioned above, we have two binding methods, http GET & POST and SOAP, for the service. They both take the same parameters and provide the same operations. The Web Service Description Language (WSDL) is used to describe our Web Image Classification Service as the following:

```
- <definitions name="WICS"
    xmlns="http://schemas.xmlsoap.org/wsdl/"
    xmlns:http="http://schemas.xmlsoap.org/wsdl/http/"
    xmlns:mime="http://schemas.xmlsoap.org/wsdl/mime/"
    xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
    xmlns:wcs="http://www.opengis.net/wcs"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:wics="http://laits.gmu.edu/ImageClassificationService"
    targetNamespace="http://laits.gmu.edu/ImageClassificationServi
    ce" xsi:schemaLocation="http://schemas.xmlsoap.org/wsdl/wsdl-
    1.11.xsd">
```

```
- <!--
 Import of useful WCS types
  -->
- <!--
 Definition of abstract messages for WCS
  -->
+ <message name="WCSExceptionMessage">
+ <message name="GetCapabilitiesRequest">
+ <message name="GetCapabilitiesResponse">
+ <message name="GetCoverageSupRequest">
+ <message name="GetCoverageSupResponse">
+ <message name="GetCoverageUnsupRequest">
+ <message name="GetCoverageUnsupResponse">
+ <message name="TrainClassifierRequest">
- <message name="TrainClassifierResponse">
     <part name="TrainingParameters" type="xsd:string" />
  </message>
- <!--
 PortType Definitions for WCS
  -->
+ <portType name="ImageClassificationPortType">
- <!--
 WCS Instance specific part of WSDL
  -->
- <!--
 GET binding
  -->
+ <binding name="ImageClassificationGETBinding"
       type="tns:ImageClassificationPortType">
- <!--
 POST binding
  -->
+ <binding name="ImageClassificationPOSTBinding"
       type="tns:ImageClassificationPortType">
- <!--
 SOAP binding
  -->
+ <binding name="ImageClassificationSOAPBinding"
       type="tns:ImageClassificationPortType">
- <!--
 Service
  -->
- <service name="ImageClassificationService">
    - <!--
     GET
      -->
```

```
+ <port name="ImageClassificationGETPort"
      binding="tns:ImageClassificationGETBinding">
- <!--
  POST
   -->
+ <port name="ImageClassificationPOSTPort"
      binding="tns:ImageClassificationPOSTBinding">
- <!--
  SOAP
   -->
+ <port name="ImageClassificationSOAPPort"
      binding="tns:ImageClassificationSOAPBinding">
  </service>
</definitions>
```

With this WSDL part, the Web Image Classification Service can be registered and published through Web Registry Service.

6. Implementation and Demonstration

To demonstrate our ideas, we have built the service with Java and C codes which can deal with the binary image data. Figure 2 to 4 are the screen shots of the web interface accessing the classification service. Users can request the capabilities file by clicking the link of *GetCapabilities*. They can also select the classification algorithm (supervised or unsupervised) by clicking the respective link (Figure 3 and 4). For supervised classification, *TrainClassifier* should be done first. The interface for the *TrainClassifier* is illustrated in Figure 2. The training datasets are provided by URLs and the parameters such as number of rows, number of columns, number of bands, and number of clusters should be provided for training. After training, users then can provide the data through an URL for the supervised classification. The web client for the supervised classification is shown in Figure 3. For unsupervised classification, the number of clusters is provided as the expected minimum clusters and maximum clusters, which is different from the supervised classification, as shown in Figure 4.

We use a subset of 5-band Landsat data as example to test our service. The test image is 1500X1500 in size. Figure 5 and 6 are two of the five-band images before classification. Figure 7 is the classified image with the unsupervised classification.

7. Conclusions and Future Work

In this work, we successfully built an experimental interoperable web-based image classification service based OGC and W3C service standards. In the future, more work will be done to make this service operational in many ways, such as providing better user interaction interface, accepting more data formats, tuning up the classification methods, and speeding up the processing.

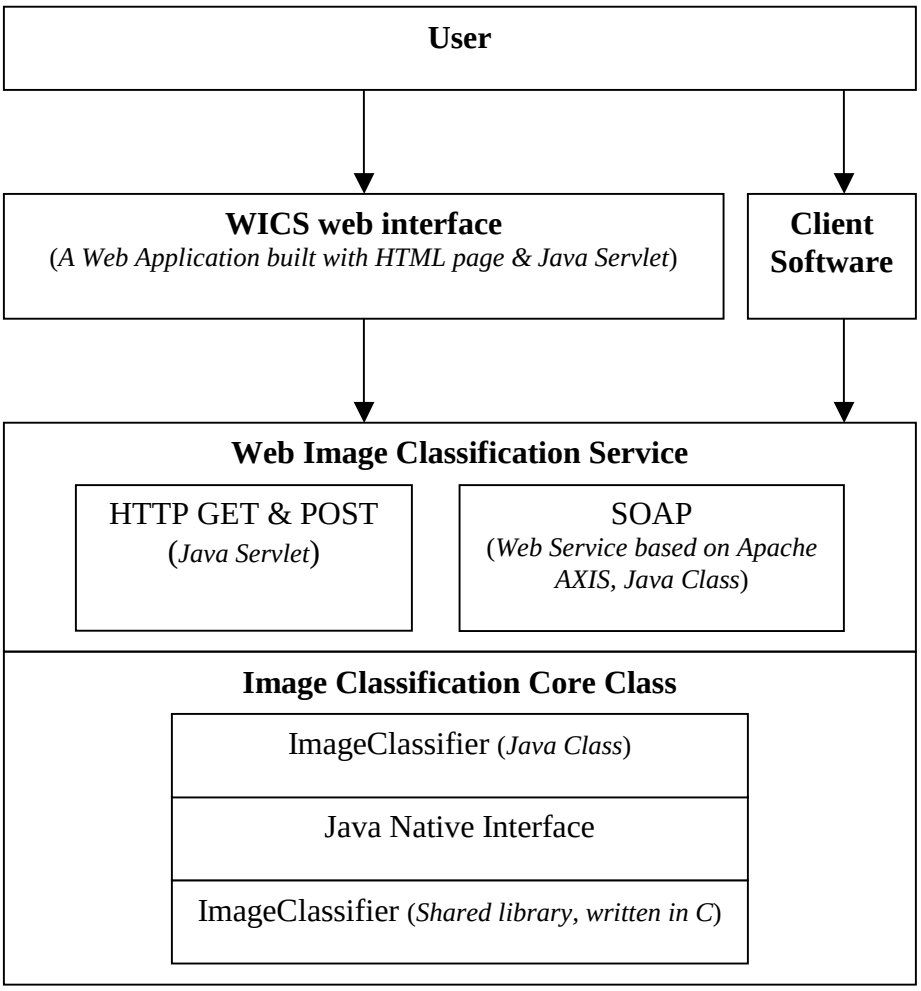

Figure 1 The structure of Web Image Classification Service

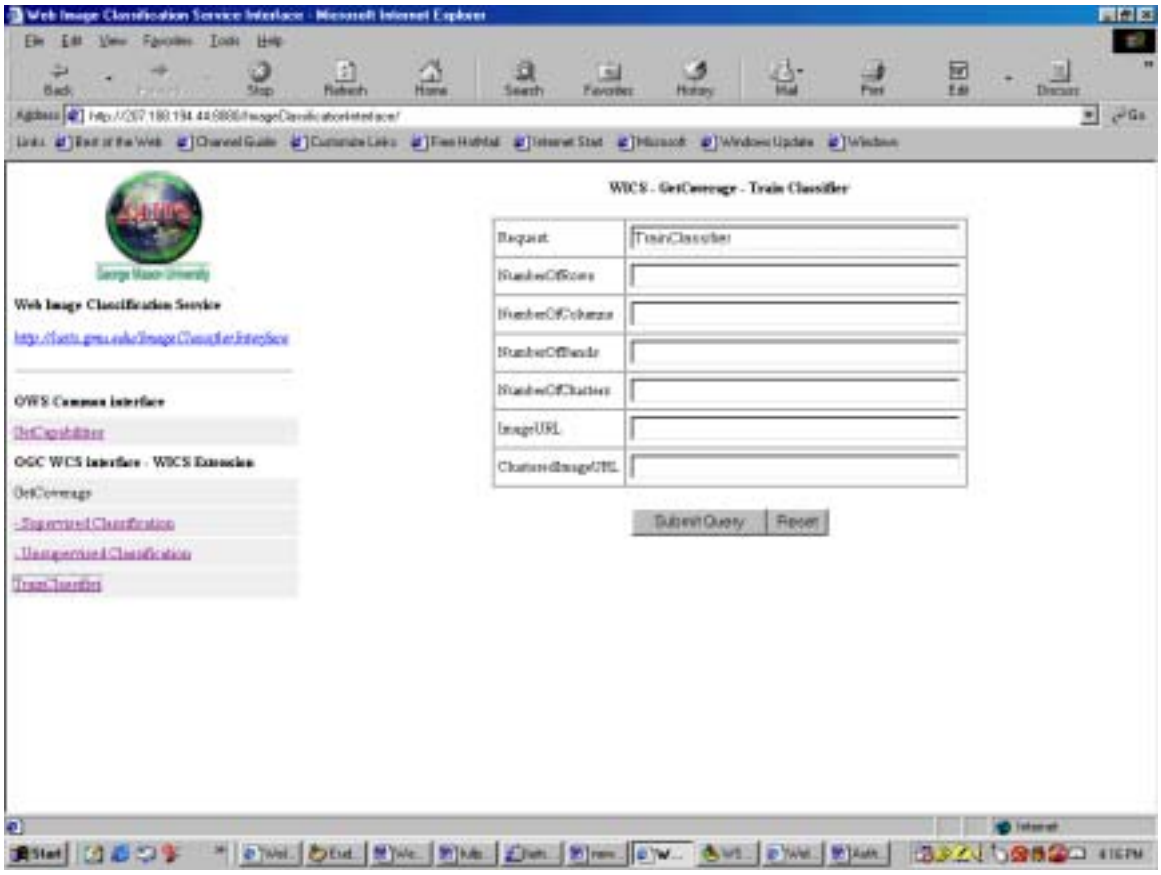

Figure 2. The Training Classifier for WICS

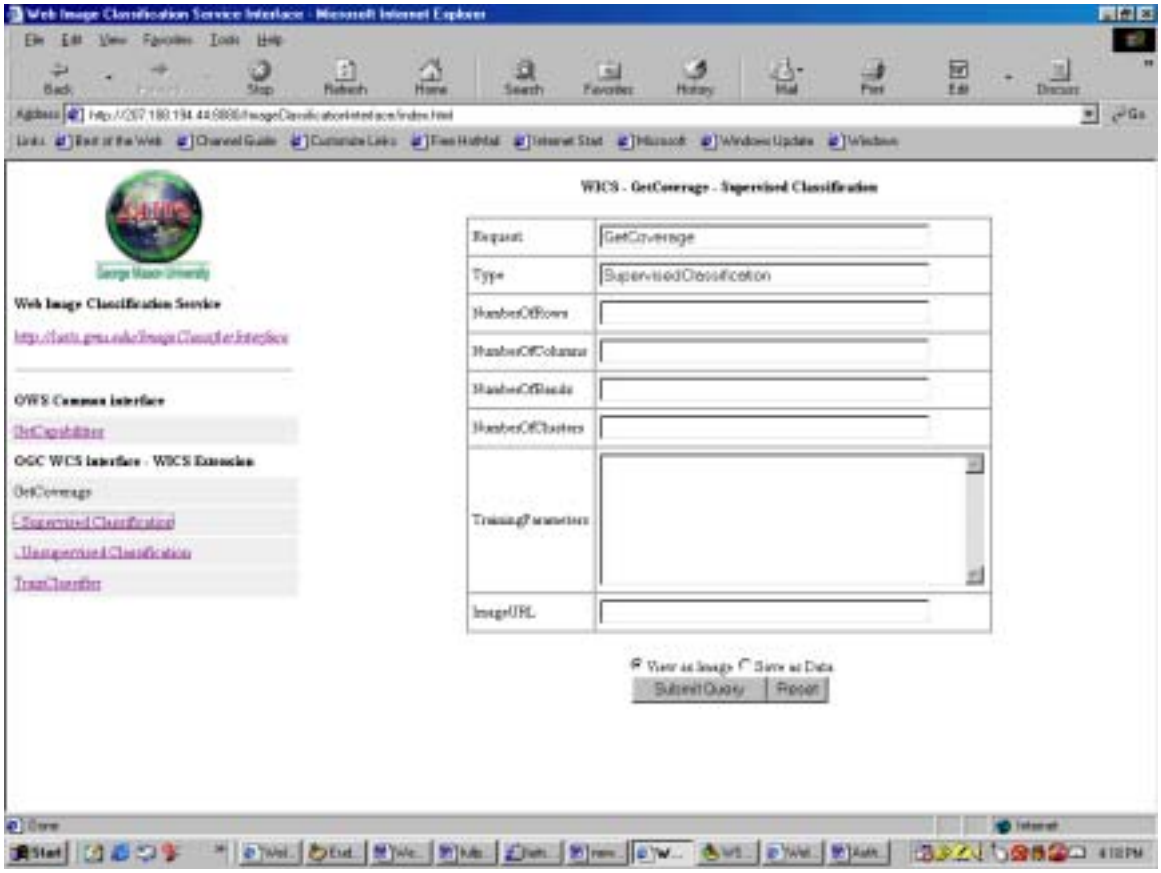

Figure 3 The Supervised Classification of WICS

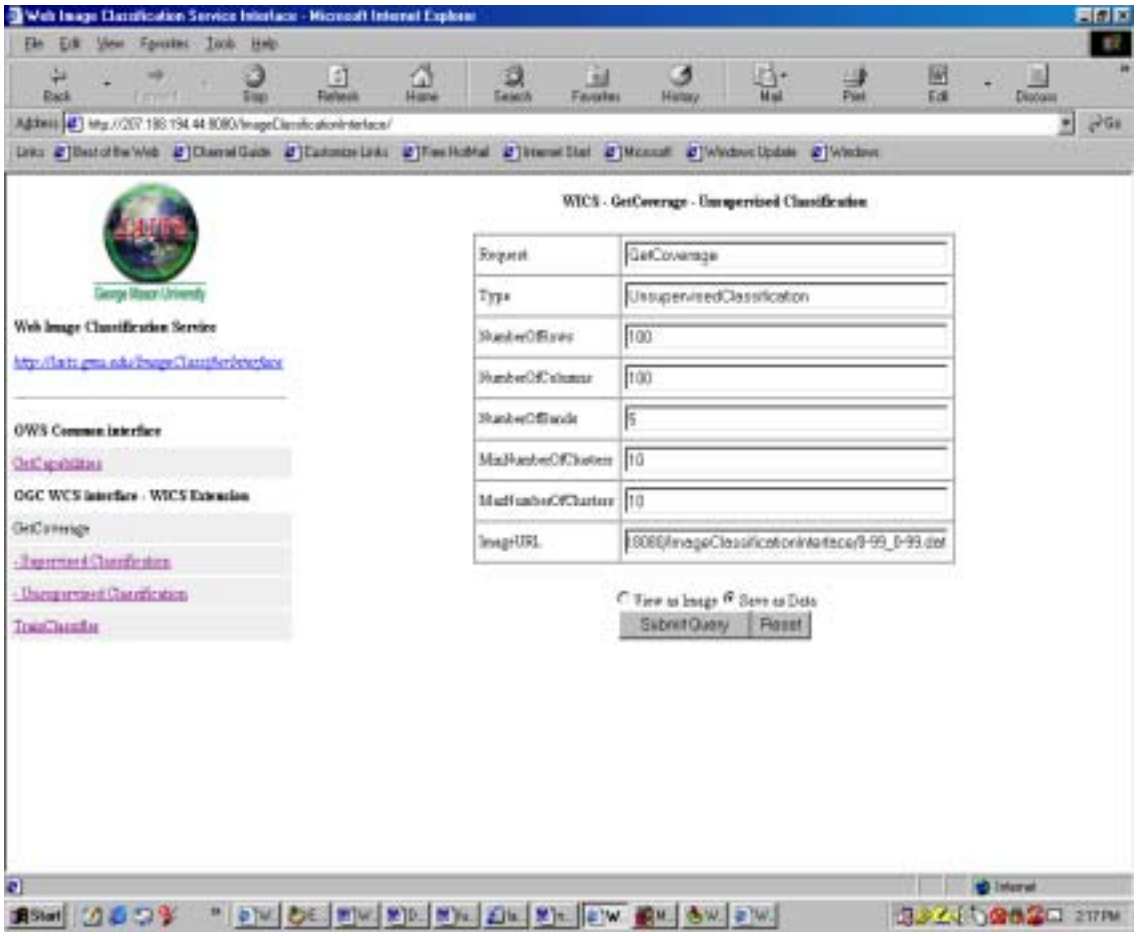

Figure 4 The Unsupervised Classification for WICS

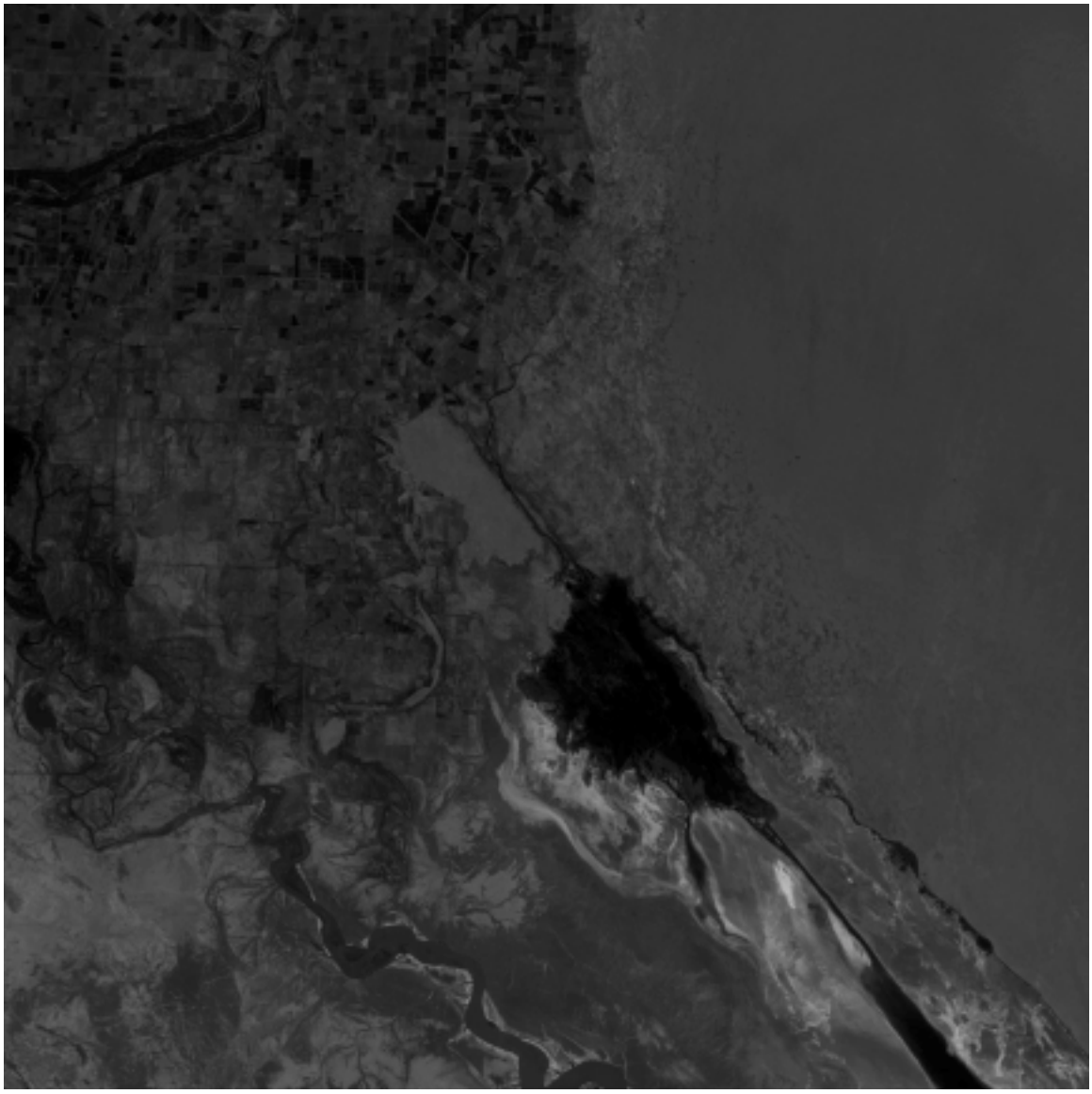

Figure 5 The raw image of one band data

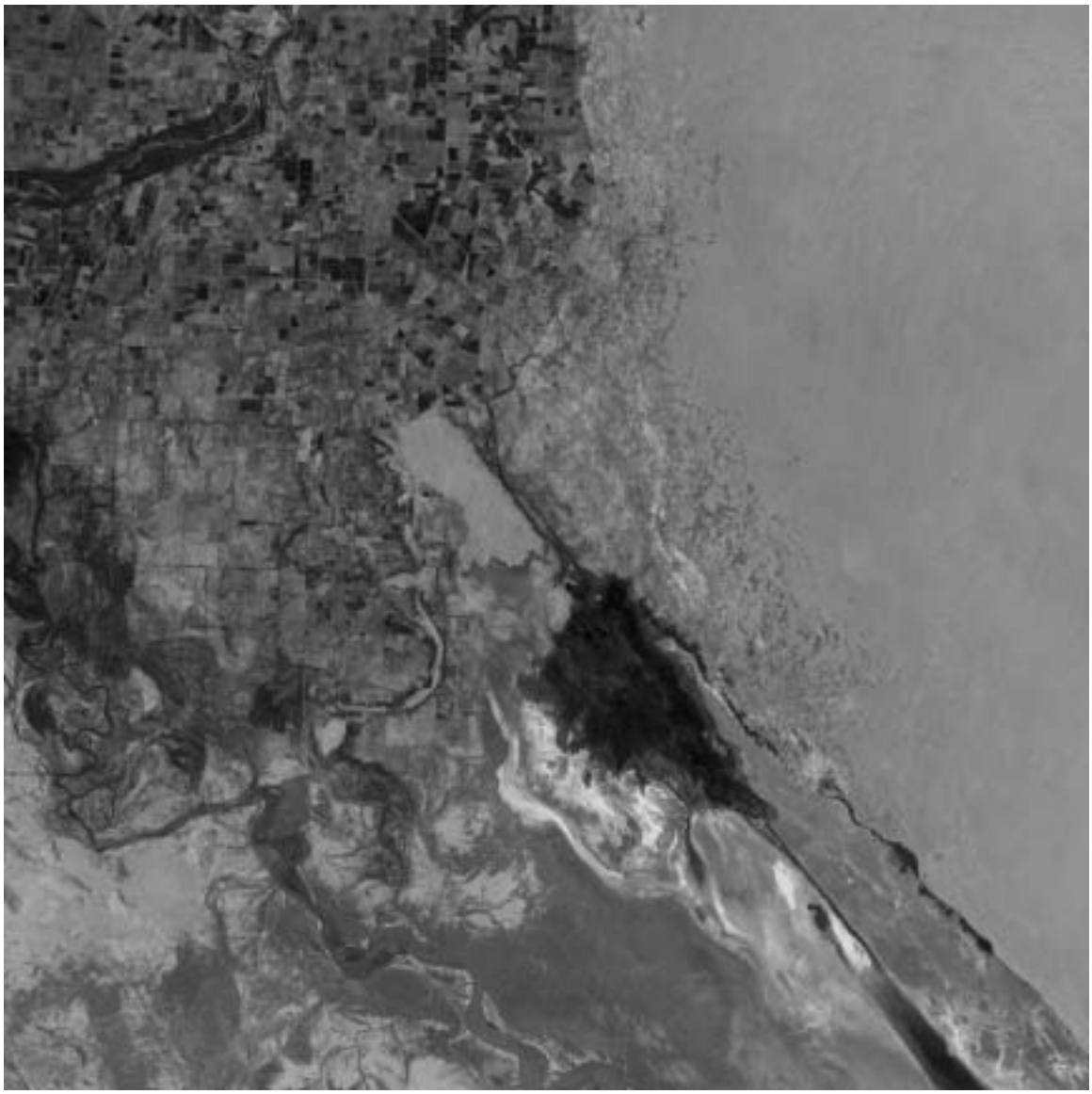

Figure 6 The raw image of another band data

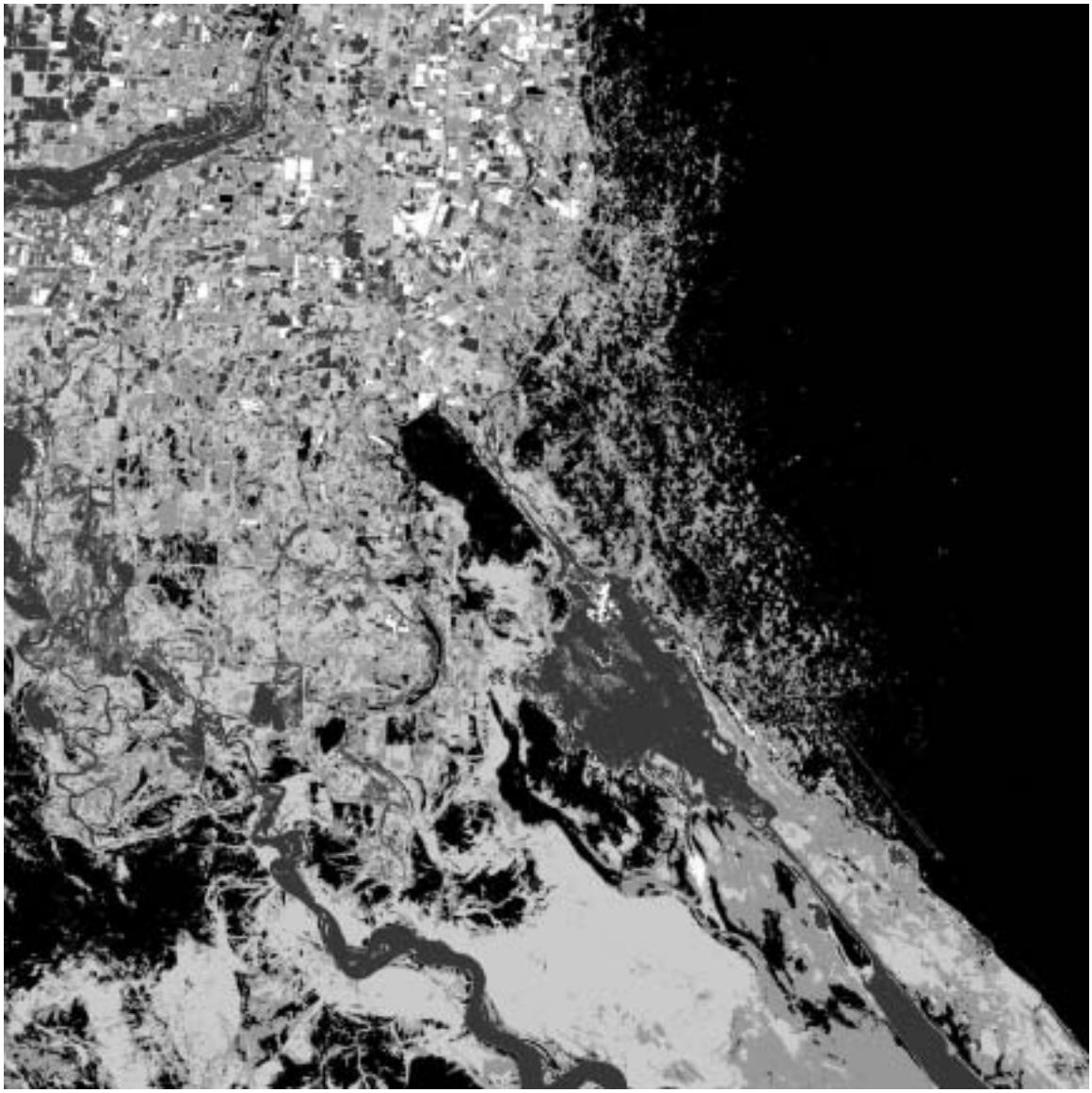

Figure 7 The classified image

REFERENCE

Evans, J. eds, 2002. Web Coverage Service Implementation Specification, OpenGIS® Project Document 02-024, http://www.opengis.org/techno/discussions/02-024.pdf.

Jensen, J.R. (1996), *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice Hall

Lieberman, J., editor, 2003. OpenGIS Web Services Architecture. OGC 03-025. http://www.opengis.org/techno/discussions/03-025.pdf

OGC, 2003. OGC homepage, http://www.opengis.org

Percivall, G. and Smits, P. (2001), *OGC Classification Service*, OpenGIS Project Document 02-012.

Vretanos, P. A. editor, 2002. Web Feature Service Implementation Specification, Version 1.0.0 OGC 02-058. http://www.opengis.org/techno/specs/02-058.pdf.