An Open Source GIS Solution for Discovering Cambodia through Maps and Facts

Xianfeng Song, Yasuyuki Kono, and Mamoru Shibayama

Center for Southeast Asian Studies, Kyoto University, 46 Shimoadachi-cho, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan, tel. +81757537338, fax +81757537350, e-mail song@cseas.kyoto-u.ac.jp

1 Introduction

The open source software related projects encompass a broad range of GIS categories, such as spatial data conversion, cartographic project, spatial data infrastructure, spatial database engine, geo-processing, spatial data integration, and web mapping. Some projects resulted software perform extremely well, and their growing success has been promoting the use of open source solutions in geo-community.

The licenses for most open source software may be different, but all of them intend to guarantee the freedom to read, redistribute, modify and use the software freely [3]. This key statement plays an important role in breaking down barriers where the cost limits the use of public spatial data and the access to GIS tools. The application of open source alternatives to proprietary GIS software benefits users greatly, particularly those working in education, research, and in developing countries.

The deployment of open source does not mean a 'zero' cost. As most software relies on UNIX/Linux platform, users**,** though being allowed to save their cost in software, have to face many technical challenges. One big challenge is the installation process that is quite different from the Microsoft windows style. The typical procedure includes downloading the software source code and associate software components, configuring the desired feature, compiling and installing the application, and configuring again the installation for startup. These steps are straightforward and routine for system administrators and programmers, but foreign (and perhaps intimidating) to the users with little real programming experiences [1]. The other significant challenge is the thorough understanding of the related core GIS technology to design and implement an open source GIS application system. The richness of software offerings allows developers to build more complex applications, however, they needs to know well the logistics of orchestrating the interaction between several open source GIS software and to master the breadth of technical skills required to manage spatial information within distributed environment.

This paper illustrates an open source GIS solution to web-based map services by means of the development of an online prototype system mapping Environment Cambodia. To be open and interoperable, the mapping system was designed to be compatible with OGC WMS specification [6]. The data sources used almost cover the entire Cambodia, including the free copy of 1:1,000,000 Vector Map level 0, the open source Landsat MSS and TM/ETM scenes at Global Land Cover Facility, the 1:500,000 Cambodian geology maps, the 1:100,000 Cambodian topographic maps and Land Use maps.

The data preprocessing and the construction of web mapping system both used merely open source GIS software. The development of such an open source system clearly illustrates the logistics of open system architecture, the orchestration of the interaction

between open source software, the performance of integrated system, and a matter of concern to function limitations or bugs. As the geographic datasets used in this work contain a huge volume of collections, this, to some extent, is also an experiment that examines the capability of open source GIS solutions in building a large-scale GIS system.

The prototype system of Environment Cambodia is available at http://aris.cseas.kyoto-u.ac.jp/khmer/. It serves web users a guide system to discover Cambodia through maps and images, in which different scale of topographic maps, geology maps, land use maps, satellite images, population distribution, and administrative units can be found.

2 Open Source Web Mapping Solution to Environment Cambodia

The web mapping is growing in popularity within geo-informatics community. It serves web users an attractive geo-referenced map, instead of geographical data itself. In term of data presentation, web mapping uses symbology and typography to convey geographical information in a manner that is both appealing to the eye and easily understood. With respect to geo-processing, web mapping refers to the specific capabilities of GIS systems to handle a very large amount of spatial data, to converse numerous spatial data formats, to process complex cartographic projection, to support map symbol generation, and to annotate multi-language text. In the context of software, web mapping is an interactive Internet program that accepts requests from web users and correspondingly return feedbacks.

From the above, web mapping functions as a heavy geo-processing system, although it is not a full-featured GIS application system. The open source solution to web mapping as follows examines the capabilities of open source software to support complex tasks.

2.1 Prototype System

The Environment Cambodia system adopts a solution that has a lightweight client and a heavyweight back-end processing server, in order to speed up processing a huge volume of GIS datasets. The prototype system has a multiple-layer structure consisting of three parts – browser, map server and database server. Figure 1 briefly shows the system architecture.

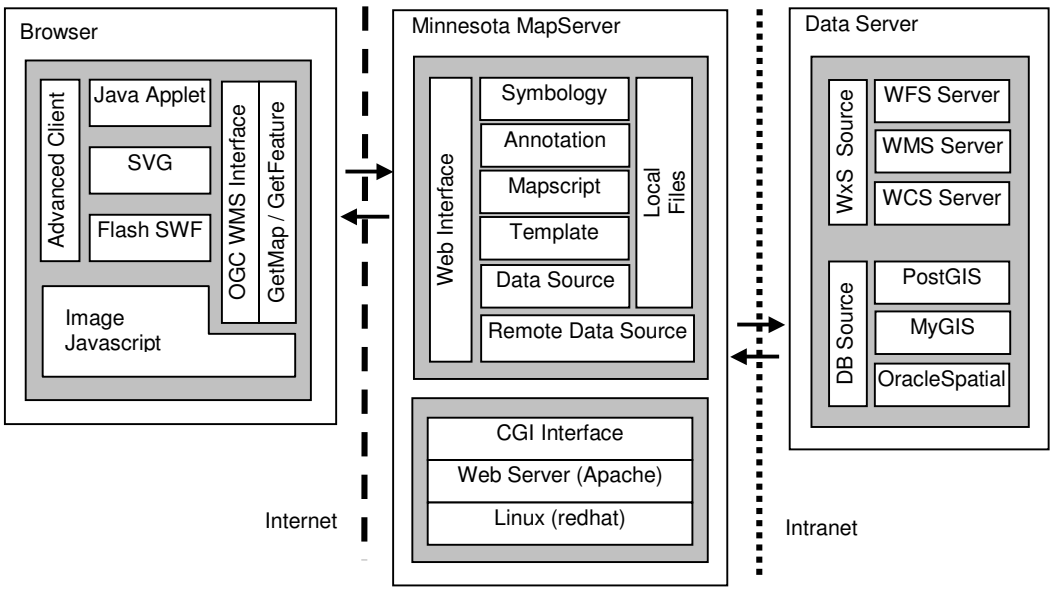

Figure 1: Open Source Prototype System for Web Mapping

The main component – map server for constructing our web mapping service is based on Minnesota Mapserver. It runs on the stable open source platforms – Redhat Linux OS and Apache Web Server, which are two popular and powerful platforms, providing a secure, efficient and extensible support to their child programs.

The database system we used to store the large volume of GIS dataset is the postgreSQL platform, whose spatial data engine – postGIS speeds up the access to spatial databases using spatial index technology.

The map clients in browser deploy multi-media graphic technologies: Java applet – jBox, SVG, and Flash SWF. The interoperation between mapclient and mapserver is implemented in both the Mapserver CGI interface and the OGC Web Map Service (WMS) Interface, but the advanced clients we programmed access map and retrieve attribute via WMS. All of them serve web users an impressive look-and-feel.

2.2 OGC WxS Interfaces and Spatial Web

OGC Implementation Specification – Web Map Service defines a set of open standard interfaces for web-based client/map-server interactions on basis of HTTP. It mandates communication parameters to access a WMS-compliant server and solves the interoperability problem among Web clients and servers thereafter. Thus, the advantage the web provides for broad access is taken again to enable the isolated spatial sources of this network to be connected.

The WMS specification work enables web-mapping applications to work cooperatively and to operate reciprocally without need to understand the details of the backend implementation. This greatly promotes the interoperation between web map servers on Internet, where each server might be attached to a huge GIS database. What is more, it also simplifies the construction of web mapping service, because open context and free exchange of spatial information guarantee a scaleable map system on Intranet.

The OGC WMS interface is not single, but there is Web Feature Server (WFS) for vector correspondingly. Moreover, the Web Coverage Server (WCS) is also under development. The irreversible movement to standards is leading to one big spatial web [2].

2.3 Map Server and Web Map Service

There are several open source web map servers with broad deployment and advanced functionality, for example, Minnesota Mapserver, GeoServer and Deegree. All of them have strong momentum and are able to be viable alternatives to proprietary map servers.

Minnesota Mapserver is built upon other open source projects and has the advantage of running on various platforms where most commercial systems won't or can't. The GeoServer project is a full transactional implementation of OGC WFS specification, with an integrated WMS. The Deegree project offers the substantial spatial-wares for building a Spatial Data Infrastructure. All parts of Deegree implements the OGC specifications, that makes the Deegree architecture extremely interoperable. Both GeoServer and Deegree is a java-based and XML/GML-oriented framework, they are thus platforms independent.

According to the declaration of supporting OGC WFS/WMS specification, all of three map servers are compatible with OGC WMS interface; the Mapserver only supports basic WFS functions, while GeoServer and Deegree fully implement OGC WFS interface, supporting additional transaction operations of GIS database – insert, update and delete. The Environment Cambodia currently focuses on the application of WMS on map browse

and geographic feature query. We selected Mapserver, other than GeoServer or Deegree, because it has an excellent performance of functionality and speed on processing large volume GIS datasets.

Figure 2 briefly shows the Mapserver architecture [4]. Although the server is primarily written in C/C++, it provides as well the scripting access to the Mapserver C API via popular script languages such as PHP, Perl, Python Java etc. This cross-platform development environment eases the development of spatially enabled Internet applications [4]. That also means there will be many choices of developing tools when extending the Mapserver-based system or cooperating it with other systems.

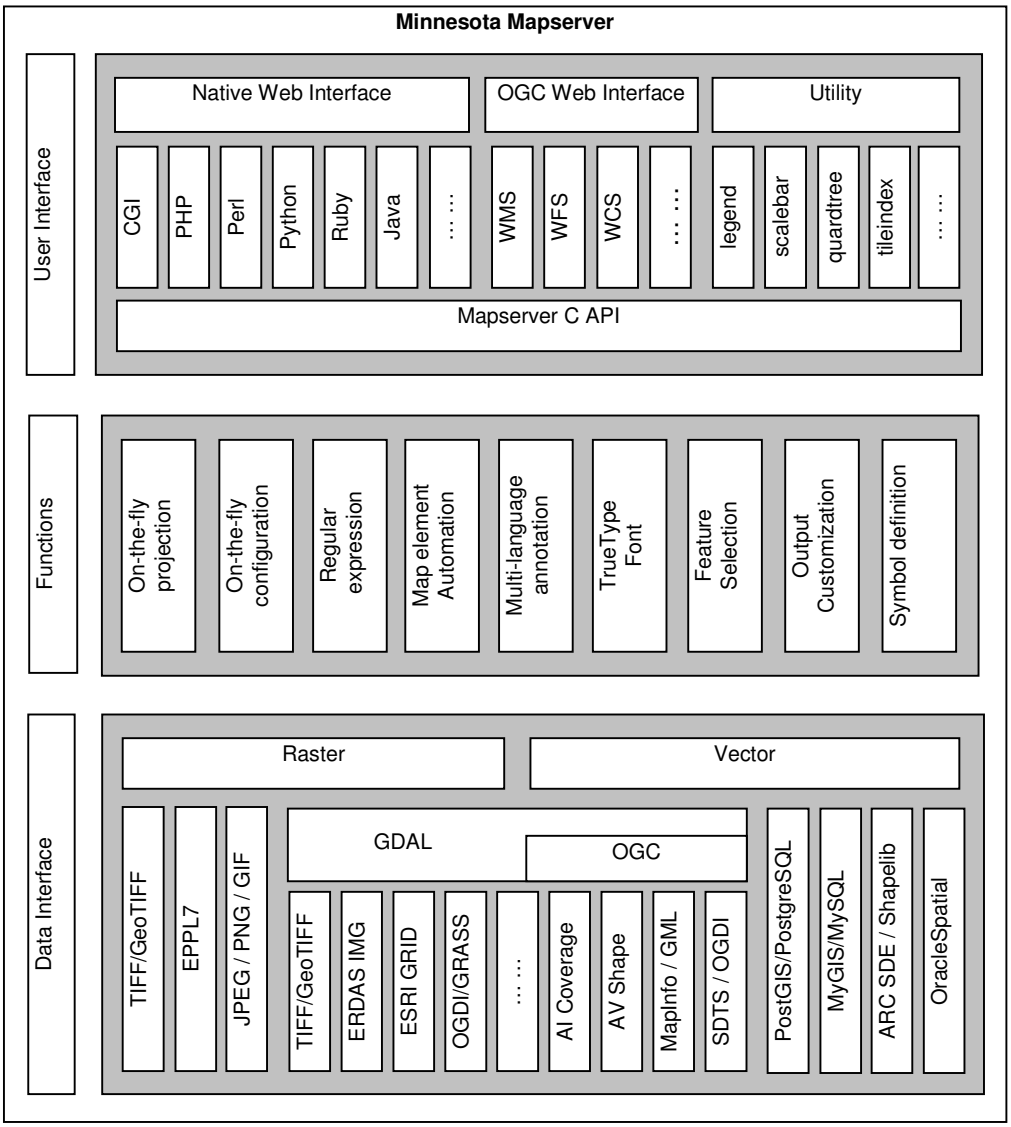

Figure 2: Minnesota Mapserver Components

The Environment Cambodia is a scaleable modular system as its interoperation solution is based on a complete industry standard OGC WxS family. In case of extension, for example, to support client-side modification of geographic features, the GeoServer or GeoServer with a full implementation of OGC WFS specification can be cooperated into the current system without change of existing parts.

3 Optimising Backend Geo-processing

3.1 Dataset Used

The data bases used in online mapping Cambodia are listed in Table 1. These datasets include raster and vector, and vary considerably in format. Among them, the Vmap0 (formerly known as Digital Chart of the World) and Landsat MSS/TM/ETM at Global Land Cover Facilities (GLCF) are open source databases. Other valuable collections are high quality digital data from Cambodia survey project that made use of a number of diverse data sets and types.

The Minnesota Mapserver can support all the spatial data formats listed in table 1 directly or via spatial components. Nevertheless, the data preprocessing still needs to improve the Mapserver performance, especially, in database access and satellite image visualization.

Data	Description	Format	Scale
Vmap0	Boundaries, Data Quality, Elevation, Hydrography, Industry, Physiography, Population, Transportation, Utilities, Vegetation.	VPF	1:1,000,000
Topographic map	Land use, river/stream, rail/road, contour, administrative units, infrastructure - temple/school/church/mosque/stupa/post office/hospital, historical site, airport and spot etc.	Arcinfo	1:100,000
Geology	Landforms, rock era and rock categories	Arcinfo	1:500,000
Administrative	Khet(province), Srok(district), khum(commune), cphum(village)	Arcinfo	1:100,000
Population	The village-level population census in 1994.	Arcinfo	
MSS/TM/ETM	MSS and TM - GeoTIFF format, ETM – HDF4 format.		

Table 1: The Databases for Mapping Cambodia

3.2 Spatial Indexing Large Volume GIS Database

Table 2 shows the process of data conversion. All the vector datasets were converted into the postGIS/postgreSQL database, instead of local files – VPF, Arcview Shape or Arcinfo Coverage, because the spatial index greatly speeds up GIS data access as the index on normal fields does.

We had an experiment in comparing the system performance where the mapserver was fed shape files and postGIS data source respectively. The tested dataset is Vmap0, a global database with four CDs. The Vmap0 was converted into Shape format and imported into postGIS respectively using OGR plus OGDI driver, and then, the quadtree spatial indexing for shape files and the spatial index (GiST) were built correspondingly.

The result showed that the use of spatial index of postGIS/postgreSQL was about 8-10 times faster than the quardtree-based spatial index of shape files. The examination also indicated the change of quadtree depth had no significantly improvement in performance**,** perhaps because the volume of Vmap0 shape files was too large, almost four GBs.

Input Format	Output Format	Free and Open Source Toolkits
VPF	PostGIS/PostgreSQL	OGR with OGDI vector support.
Arcinfo Coverage	PostGIS/PostgreSQL	OGR
HDF4	GeoTiff	GDAL with HDF4 support.
GeoTiff (Single band)	PPM (multi-bands)	GRASS: r.in.gdal imports single band GeoTiff first; and then r.out.ppm3 exports to a multi-bands PPM file.
PPM (multi-bands)	GeoTiff (multi-bands)	GDAL: gdal_translate converts a PPM file to a TIFF file. LibGeoTiff: geotifcp/listgeo writes geotags into the TIFF file.

Table 2: Data Preprocessing

3.3 Landsat Image Preprocessing

Image Composite

The Landsat satellite scene at GLCF is in either GeoTiff or HDF4 format, and each band is stored separately. To enable the Mapserver to render a pseudo-colour image to web users, all single image bands were merged into one multi-bands composite image and saved in GeoTiff format.

Table 2 lists raster data conversion in which the GDAL, GRASS and LibGeoTiff are main operation tools. The raster data processing is much more difficult because of the problem we faced in merging single bands. The HDF4 bands were converted into GeoTiff format, and then, single band GeoTiff files were supposed to be composited as follows:

GeoTiff ---(r.in.gdal)--->GRASS --- (r.out.erdas) ---> ERDAS --- (gdal_translate) ---> GeoTiff

However, the ERDAS (.lan) exported by current GRASS (version 5.3) was not compatible with GDAL, although the associate documents indicate they should be so with each other. We extensively tested a successful conversion as follows:

GeoTiff ---(r.in.gdal)--->GRASS --- (r.out.ppm3) ---> PPM --- (gdal_translate) ---> TIFF

GeoTiff ---(listgeo)--->GeoTag --- (geotifcp) ---> GeoTiff

This conversion requires all bands should have the same spatial extent and the same number of pixels in width and height. Otherwise, the bands would not match each other.

Tile Indexing

The Mapserver with GDAL support gives a high-performance read access to remote sensing images. The high performance is achieved particularly though GDAL library, GDAL utility and GDAL supported raster files.

The Cambodia territory occupies 18 Landsat MSS scenes. If mosaicing these images, the system performance would be significantly slowed down. Fortunately, the Mapserver uses 'TILEINDEX' to control displaying images that just overlap the current map windows. This tile index actually is a shape file that stores the name and map extent of each tile image files. In case of handling a very large volume scene such as Landsat ETM+, it is also convenient and has higher performance to split one scene into a number of small tiles.

On-the-Fly Raster Resampling

The on-the-fly resample process would also slow down the Mapserver performance, particularly in case of displaying multiple scenes simultaneously. There are two ways to overcome this problem. One is scale control and the other is to use the image with pre-built overviews. The former requires resampling a satellite image into multiple-resolution copies and uses MAXSCALE / MINSCALE in mapscript to control displaying the copy with an appropriate resolution. The later uses GDAL utility – gadaladdo to generate pre-built overviews for a satellite image and use GDAL in support of the Mapserver to access the image with pre-built overviews. In our work, we used the GeoTiff with built-in overviews to improve the performance of image rendering.

4 Advancing Map Access through Web Browser

Web users may use a simple web browser to access maps and query attributes if they are able to manually compose a HTTP request following the WMS specification. But in most cases, it is a WMS compliant client that automates calculating the HTTP request parameters and loading the resulting map. This map client is often a lightweight spatial-ware embedding in a web browser, thus the advantage of web browsers will be taken handily to expand the universe of accessing spatial data on Web.

Conducting a lightweight WMS client must revolve around the full compliance with its host system, great user interactivity, and impressive data representation. The development should concern two basic functions - auxiliary graphic user interfaces for interactively operating maps and automatic HTTP request processors corresponding to map operation.

In this study, we deployed three advanced multimedia clients to operate our OGC WMS compliant map server. The HTTP request processors are coded using client-side JavaScript, while the implementation of user interactivity and map display relies on Java applets, SVG and Flash SWF respectively. Each client gives a deep impression on enhancing user interactivity and map presentation.

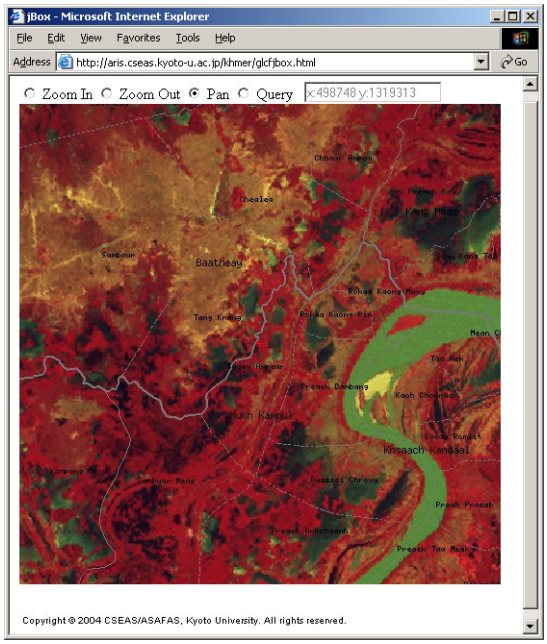

Figure 3a: Java Applet - jBox

4.1 Java Applet - jBox

The jBox (formerly called “Mapplet”) is an elegant open source Java applet. It is initially developed to enhance user interactivity of CGI Minnesota Mapserver. The jBox supports two image formats: JPEG and PNG. There are a number of interesting features supported in the current version, such as box zooming, panning, box querying, map image swapping, distance measuring and cursor coordinate dumping. In this practice, the jBox was configured as a smart WMS client, where the jBox serve users an interactive interface and JavaScript functions handle HTTP request. The parameter exchange between JavaScript and jBox is through the LiveConnect Java ARchives.

When using jBox as a OGC WMS compliant client, a particular attention should be paid to a bug in jBox’s setimage() method. That is the jBox displays map in a wrong position

when it operates maps in image swapping style. This is because the setimage() method forgets to set zero two variables ('ix' and 'iy') – the 'x' and 'y' direction's displacements between the event 'mousedown' and 'mouseup', although it already re-initialises the cursor positions of event 'mousedown' and 'mouseup'. Figure 3a is the bug fixed java applet - jBox.

Rosa is another Java applet from DM Solutions, providing the functionalities similar to jBox. It works with PHP script and can be scripted into a WMS client in same way.

4.2 Scalable Vector Graphics

SVG is XML graphics for the Web. It is a Web standard (W3C -World Wide Web Consortium Recommendation) that defines a language for describing two-dimensional graphics and graphical applications in XML (see W3C definition). SVG supports three types of graphic objects: vector shapes, images and text. The SVG process of graphic objects includes transformation among coordinate systems, rendering in CSS or XLS styling, dynamic access to Document Object Model (DOM) through events and scripting, and comprehensive support for SMIL animation. These features make SVG fulfil the very specific requirements for dynamic GIS mapping. Compared with traditional graphic formats, SVG is incredibly versatile and is revolutionizing how spatial data are graphically presented on the Internet.

SVG is an open source specification on basis of a collection of other successful open standards – XMS, PNG, DOM, CSS and SMIL. The strong industrial support makes SVG never tied to one particular implementation or vendor. Developer can use SVG as a common web-mapping platform upon which to build graphically rich GIS applications and user interface. Thus, programmers bypass web browser incompatibilities and missing functionality.

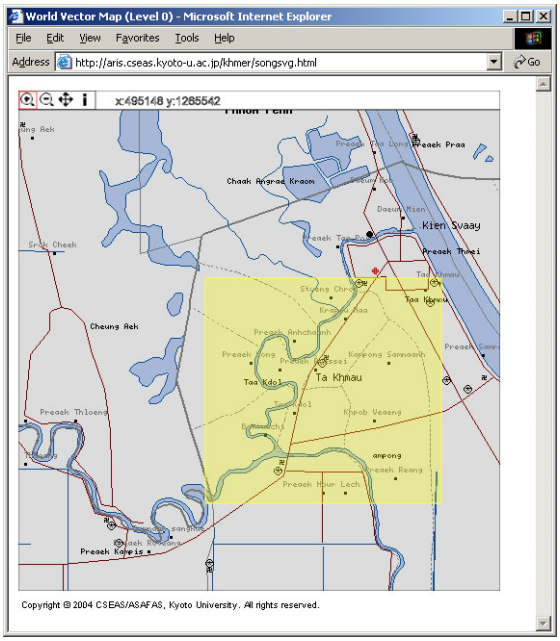

Figure 3b: SVG Client

Figure 3b shows the SVG-enabled WMS client we developed using SVG graphics and JavaScript. It interacts with the same remote WMS mapping server – Environment Cambodia as the above jBox does, while it enables enhanced graphics and sophisticated interactivity.

4.3 Flash SWF

The Macromedia Flash file format (SWF) specification is designed to deliver graphics, animation and sound over the Internet. It is licensed to Macromedia Flash, but there are no access or deployment fees required to use the specification. The SWF is a very efficient delivery format and is able to be rendered at a high quality very quickly. SWF visualization and sound efforts often come as a shock.

Flash SWF is still increasing in popularity. The Macromedia Flash compliant player may be the most pervasive graphic platform on the web**,** and it has become an industry standard to some extent. For these reasons, it is very attractive to use SWF as a platform

for web mapping. It is different from a textual format of SVG graphics, since the SWF is a binary file, rather than transparent. The generation of SWF files has to depend on some dedicated toolkits. Figure 3c shows a Flash SWF version of WMS client that was conducted using open source library – libMing and PHP script language. Its performance is extremely impressive.

The ActionScript is a powerful object-oriented programming language that is based on ECMAScript and JavaScript. It helps to create sophisticated interactive function. The libMing is a free/libre SWF output library that defines a set of Flash Movie objects. With PHP support, we can use Ming functions and ActionScript to create a WMS-compliant Flash client.

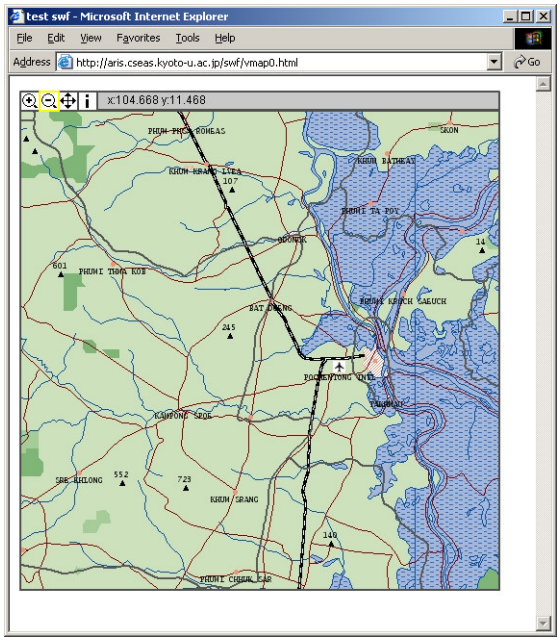

Figure 3c: Flash SWF Client

When using libMing-output Flash for interoperable web mapping, the web application developers have to be ware of some limitations. That is, libMing can only deal with JPEG baseline ("standard") at this point, no baseline optimised or progressive scan jpegs. Therefore, the output format of Minnesota Mapserver WMS server was configured to support the standard baseline JPEG in this system.

5 Conclusions

This work highlights an open source GIS solution to web mapping. Through developing such a system, we found: 1) The success that the prototype system served a huge volume of spatial datasets on the web indicated that the free and open source GIS software with a strong momentum could offer a technically competitive and open-source alternative to a close-source enterprise solution; 2) the OGC WxS specification not only allowed the system to reciprocate with other compliant Internet map services, but also guaranteed a scaleable modular system easily to be expanded on Intranet; 3) the promulgation of the multi-media technology, for example, Flash SWF, SVG graphics and Java applets, greatly expanded the universe of accessing map on web by attracting Internet users; and 4) the bug fix of java applet - jBox showed us once more that the open source freedom benefited system developers very greatly from solving program problems immediately.

The techniques, development experiences and system source code in this research work can be applied freely as a laboratory excise or a course curriculum. As Ramsey mentioned wherever people had problems to solve and a willingness to share their solutions with others, open source would continue to flourish [5]. We hope the dissemination of this development will contribute to the geo-community in case of deploying Open Source GIS software.

Acknowledgements

The present study is financially supported by a grant-in-aid from Japanese Geographic Data Center. The authors are also thankful to CSEAS Library, Kyoto University, for providing a number of high quality spatial datasets.

References

[1] Anderson, G., 2003. The Door Opens Open-Source GIS, *GEO World*, http://www.geoplace.com/gw/2003/0306/0306opn.asp

[2] McKee L., 2001. Web Mapping Guide - Technology Trends, *GEO Resources*, http://www.geoplace.com/gr/webmapping/technology.asp

[3] OSI - Open Source Initiative, http://www.opensource.org/licenses/

[4] Raghavan V., Santitamont P., Masumoto S., and Honda K., 2002. Implementing Web GIS Applications using Open Source Software, *Map Asia 2002*, http://www.gisdevelopment.net/technology/gis/techgi0062pf.htm

[5] Ramsey P., 2002. Open Source GIS Fights the Three-Horned Monster, *GEO World*, http://www.geoplace.com/gw/2002/0208/0208gis.asp

[6] OpenGIS® Specifications, http://www.opengis.org/