LANDMAP: Serving Satellite imagery to the UK academic Community.

Kamie Kitmitto (1), Colin C. Venters (1), Matthew D. Cooper (1), Jan-Peter Muller (2), Jeremy G. Morley (2), Alec H. Walker (2), Sanjay Rana (2).

[1]Manchester Computing, The University of Manchester, Oxford Road, MANCHESTER M13 9PL

[2]Dept. of Geomatic Engineering, University College London Gower Street, LONDON WC1E 6BT
http://www.landmap.ac.uk

Abstract

The LANDMAP Project has created the first IfSAR DEM covering the entire British Isles. That DEM and associated orthorectified ERS images were used to produce a set of orthorectified images of the British Isles from LANDSAT, and SPOT. Additional merged and mosaiced images were also created with the three different satellite products. The dataset comprise images stored as files in the Geo-TIFF format. Muller et al. (this conference) discusses the data creation process.

These data-sets are freely available to the British Academic Community for teaching, learning and research purposes. We present here the methods of serving these data to the user community. Serving these data to users has been established through two different mechanisms:

1. An x-telnet service over Super-Janet
2. A full W3 web interface.

The two services are available from MIMAS (Manchester InforMation and Associated Services, www.mimas.ac.uk) and are open to all UK academics.

This paper will report on the LANDMAP online and W3 services. It will also report and expand on the different issues involved in the provision of online and W3 services. Also some results of our investigation into image web serving will be reported.

1 Introduction

The LANDMAP Project is in the process of generating a range of output data sets. These data range from final products, as listed in Table 1, and various intermediate products.

Table 1. Output datasets

Results of the two Kinematic GPS Campaigns.
A 30m Digital Elevation Model product
Ortho-rectified archive of ERS-SAR data
Ortho-rectified archive of Landsat data
Ortho rectified archive of SPOT 10m panchromatic data
Landsat – SAR fused archive
Landsat – SPOT fused archive
SPOT – SAR fused archive

Other papers at this conference (Muller et al.), (Dowman et al.), (Cross et al), (Morley et al.) have elaborated on the processes for the creation, and quality assurance of the IfSAR DEM and the associated orthorectified images of the UK

Providing on-line access to these data-sets requires tackling several issues involved in such a service. The following section will list and discuss these issues. The process of selecting the software for the W3 service is highlighted. The inclusion of an experimental content-based retrieval search feature for this project is outlined.

2 Issues for serving the imagery

This section will list and discuss the different issues involved in the provision of an on-line service of the data sets listed in Table 1. The focus of the service is ease of accessibility of the data to a wide variety of users.

2.1 Archive size

Although the price of storage medium has depreciated recently the amount of storage space available to provide a service will always be an issue for Satellite imagery. Before project LANDMAP started the archive consisted of a single coverage of LANDSAT and SPOT imagery from data acquired in he late 1980s and early 1990s. Recently the archive has been boosted by two tandem coverages of ERS data, and by the end of the project LANDSAT 7 data coverage and various additional ERS data will be added to the archive. Adding to the above intermediate products the amount of storage required will grow very rapidly. Current estimates indicate that over 1 Tbyte on- and off-line storage will be required for all these data products.

A rationalisation of the data available on line will have to be considered. In order to do this a series of assumptions will have to be made on the technical requirements of the audience.

2.2 Metadatabase

The size of, and the variety of, the data sets created require a robust search mechanism. In order to ensure proper access to LANDMAP data products, it was realised from the outset that a suitable metadata standard had to be adopted. A number of prominent metadata standards such as Dublin core, purl.org/dc/index.htm, Directory Interchange Format, gcmd.gsfc.nasa.gov/md/difguide/difman.html, and Content Standard for Digital Geospatial Metadata- CSGDM, www.fgdc.gov/metadata/contstan.html, were investigated for their suitability.

It was decided, on the basis of several considerations, that CSGDM will be adopted for LANDMAP because of its detailed structure (although perhaps overly complex for certain data) and extensive use in the GIS community, to name a few. Most ESRI products now provide and understand CSGDM standards. The set of descriptor elements to be used in the final metadatabase is currently being designed.

2.3 Intended Audience

Users of the Satellite archives in the past were usually either students with particular projects or experts from participating institutions who could deal with the complexities of raw data. Both types of users understood the amount of work involved in transforming from raw data into an orthorectified data product that could be integrated and used with other

data. The existence of only raw data has limited the access to and interest in satellite data severely to a small and select few. The work done by the LANDMAP project aims to make access to the data very simple. This will increase the interest and attract a wide spectrum of users. This could include casual users who only require a backdrop image that they can use in a newsletter, to users who have a need for satellite data, but without enough technical expertise or support to enable the exploitation of raw RS data. The needs of any of these constituencies were not fulfilled previously.

The experience of the Digimap project, digimap.edina.ac.uk, with the provision of high quality maps through the web indicates a huge interest from non-experts and casual users of the data. We anticipate that this kind of user will be the main users of the LANDMAP output datasets. Provision of the service will concentrate on meeting the access needs of intermediate and expert users. The funding and time limits on the LANDMAP project dictate that access needs of non-expert and casual users and their additional educational support will be tackled at a later stage of the MIMAS service provision.

Experts in remote sensing (RS) might require a more involved access mechanism, which may require the restoration of files from the off line archives. Although this could be automated we do not have the resources to provide this automation. Instructions on how to go about restoration will be available. A web-based metadatabase will also be made available for users to interrogate and locate a dataset.

This means that the mosaiced datasets will be available online. A limited number of un-mosaiced datasets will also be available online. The selection of these datasets has not been finalised yet.

2.4 Format of served data

Erdas ® Imagine format and Ground Receiving Station (GRS) formats have been used up till recently for the online service of the data. At the start of this project we decided to provide the data in GeoTiff format. The selection of the GeoTiff format was taken with the expectation that this would facilitate the import of the output datasets into most commercial software. The output DEM and the ERS data would be available in signed and unsigned 16bit format.

Table 2 GeoTiff handling capabilities of major Image Processing software

ArcInfo 8.01 (UNIX)	Problems with reading GeoTiff that are read by ArcView and ArcInfo 8.01 Desktop!
Erdas Imagine 8.4 (UNIX)	Exports unsigned 16 bit images Imports unsigned 16 bit images
ERMapper 5.5 (UNIX)	Only imports 8 bit GeoTiff images. Therefore not directly capable of importing 16 bit signed DEM files
PCI EAS/PACE 6.2 (UNIX)	PCI's TIFF support does not include 16-bit data, or any other data sizes other than 1-bit (bitmap) and 8-bit.
IDL 5.2 ENVI 3.2 (UNIX)	The WRITE_TIFF procedure can write TIFF files with 1 to 3 channels where each channel can have 8-bits, 16-bits, 32-bits, or be floating point.

However, our experience working with GeoTiff has led us to different conclusions. Table 2 shows some of the limitations encountered when using some of the commercial software with using the GeoTiff format.

A recent survey of Raster GIS packages in GeoEurope indicated that most packages could read GeoTiff or Tiff format though there is no indication of the 16bit capabilities (June 2000).

Despite the varied support for 16bit GeoTiff this format will still be supported as a download option as it is an emerging standard for the provision of cross platform raster products. The choice of GeoTiff as a format will affect the provision of the mulitspectral data. At this stage our preliminary thoughts are that each image would be split into three files in the following grouping

File 1	Bands 1,2 and 3
File 2	Bands 4, 5, and 7
File 3	Band 6
File 4	Pan (for Landsat 7)

In some cases we might consider that files are stored in Erdas ® Imagine file format.

2.5 Map projections and coordinate systems.

The data will be provided as a general rule in WGS84 ellipsoid in decimal degrees. This will facilitate the integration of these datasets with data collected using GPS receivers. The integration of the output datasets with other map data such as OS ® vector data, Post Code, and census digital boundaries means that some of the output datasets will be also available in Great Britain National Grid, in Meters. Instructions and software for the projection of datasets from one projection to the other will be available as part of the service.

3 Access mechanisms

Two access mechanisms are available for serving the LANDMAP data to users.

1. An x-telnet service over Super-Janet
2. A full W3 web interface.

The two services are available from MIMAS (Manchester InforMation and Associated Services, www.mimas.ac.uk) and are open to all UK academics.

The service is provided on a set of co-operating SUN computer systems:

- A Sun E6500 with 23 x 336MHz UltraSparc CPUs, 7Gigabytes of memory and a couple of Terabytes of local disc space of which LANDMAP uses 0.6 Terabyte.
- Tame (12 processors).
- Vault (2 processors). Vault is currently being upgraded.
- A number of NT servers.

An X-telnet service allows users to logon to a server over the network to gain access to the data and use one of several image processing systems to visualize, process, manipulate and extract data. The whole processing can be performed on the server or parts of the data can be downloaded for local manipulation. Software available at irwell.mimas.ac.uk includes Imagine, ERMapper 5.5, ENVI, IDL, and PCI. Many of these software packages come with additional modules such as IfSAR, Atmospheric Correction and Sub-pixel Classifier modules with limited licensing.

In addition a fully-fledged IfSAR processing system, PULSAR, is available for use on Irwell, which is the core of the National IfSAR Processing Facility. This facility will allow users to process Radar imagery from raw data to DEM products in the same manner as described by Muller (Muller et al.).

The W3 web interface is being developed to front-end the service. The system utilizes image web serving software that was selected as the result of an investigation into currently available software (see next section). The retrieval of places of interest will be done through the use of a gazetteer coupled with a metadatabase. The quick looks will be available for users to browse in JPEG format. However, the service that allows the download of actual data in GeoTiff format is restricted to UK academics.

4 Survey of GIS Image Web Servers

A survey was conducted into the feasibility of image web serving technology to provide an off-the-shelf solution for serving the data. The aim of the investigation was to identify and review image web serving technology that could be successfully utilized to serve the dataset via the Internet. The investigation primarily focused on functionality and performance. The survey identified several products that were specifically developed as solutions for serving GIS data across the Internet: ArcIMS (ESRI), Image Web Server (ERMapper), ION (Research Systems Inc.), MapGuide Server (Autodesk), Maptitude (Caliper), MrSID Image Server (LizardTech), TerraShare (Z/I Imaging Corporation), and sWeb Interface/Smartmap (Formida).

This section reports the preliminary findings of the survey into two of the main image web server vendors: ArcIMS (ESRI), and Image Web Server (ERMapper). This software packages are the two most likely to be used in the web serving service provision. In addition to their functionality and performance their cost is reasonable. MIMAS already has an ArcIMS license while Image Web Server is available through a CHEST license agreement, www.chest.ac.uk.

All software was installed separately on a Dell Precession Workstation 220 with the following hardware and software components:

- 533Mhz Intel Pentium III microprocessor with 256K Cache
- 128MB ECC RDRAM
- Windows NT 4.0 Server, Service Pack 6(a)
- Diamond nVidia TNT2 (2x AGP) Video Card 32MB RAM at 1600*1200 resolution in True Colour Mode.
- 10.2GB ATA-66 EIDE HD with a seek time of 7.200rpm

Before starting the performance tests the system was rebooted. This provided full system resources to each product with an empty cache and clean paging file. The results of the performance tests have not yet been analysed sufficiently to be included in this paper. However, preliminary results suggest that the overall performance fall somewhere short of the performance levels claimed by the vendors. It is currently unclear why performance levels were affected.

4.1 ArcIMS, ESRI

ArcIMS was developed by ESRI, www.esri.com. This solution provides an underpinning for disseminating high-end geographic information data services via the Internet. The ArcIMS

server technology is part of a multi-tier architecture, which consists of clients, services, and data management integrated into present Web server technology. The architecture provides several features for supporting both client and server-side operations including:

- Integrating local and Internet data, which allows users to combine data and information, accessed via the Internet with local data for display, query, and analysis in their Web browser.
- Wizards, templates and development tools for creating, designing, and managing Web sites.
- HTML/DHTML and Java clients for local geo-processing.
- Scalable server architecture - for publishing GIS services from a single server or distributed services across multiple servers.
- Dynamically edit or annotate maps.
- Cartographic rendering.
- Collections of geographic data.

ArcIMS supports several industry standard raster formats including ADRG, BIL, BIP, BMP, BSQ, CADRG, CIB, ERDAS GIS, ERDAS LAN, GeoTiff, GIF, IMPELL, JPEG, MrSID, NITF, SUN Raster, TIFF, ASRP, ERDAS IMAGINE, USRP (Windows NT only), GRID, ArcView Image Catalogs (DBF files). Vector formats include ESRI shapefile, ArcSDE layers (3.x), ArcSDE layers from ArcSDE 3.02 and 8.02.

There are three O/S distributions for ArcIMS including Windows NT 4.0 Intel (SP4, 5, 6a), Solaris 2.6 and 7 with several future releases in the pipeline for Digital UNIX 4.0D, HP-UX 11.0, IBM AIX 4.3.2, and SGI IRIX 6.5. ArcIMS supports several Web server solutions including Netscape Enterprise Server 3.6, Internet Information Server 4.0, Apache 1.3.9 and 1.3.11, IPlanet Web Server, Enterprise Edition 4.0 and Web Sphere 3.0.2.

The possibility of deploying ArcIMS solutions over a combination of UNIX and NT machines presents a good combination for big data centers like MIMAS. This allows the investment in big UNIX servers to be used efficiently.

4.2 Image Web Server, ERMapper

Image Web Server is developed by ERMapper, www.ermapper.com. The Image Web Server is incorporated into existing Web server technology. The solution was specifically developed to serve GB or TB images to any application, including Web browsers, MS Office, GIS and CAD applications. The Image Web Server utilizes a wavelet compression technology, ECW (Enhanced Wavelet Compression), to minimize the size of images distributed without any significant degradation of quality. Data is transferred across the network using a new transfer protocol, ECW Protocol (ECWP), and is identified by the prefix ecwp://. The viewing and decompression of image data is a client-side function, which reduces the load on the Image Web Server. The client automatically caches portions of the image being viewed to give fast access. The client and server exchange information about the image as the user roams and zooms within the data. This process is de-coupled from the iDWT process that constructs the view user see, and these are in turn de-coupled from the information being requested from the server.

Users can access the data via Internet Explorer and Netscape web browsers, or from within applications like ERMapper, MapInfo, ArcView, and Microsoft Office tools. PCI Geomatics indicated that they are going to provide support for the ECW wavelet compression. Plug-ins are required to take advantage of the technology within third-party client applications. These are free and can be downloaded and installed with ease.

The Image Web Server comes in several distributions including Enterprise edition, Corporate edition, Office edition, ERMapper edition and a Free edition. Image Web Server is an MS Windows based solution with distributions for the MS Windows family, WIN95/98/NT/2000. Image Web Server requires that either Personal Web Server (95/98) or Personal Web Server/Internet Information Server (NT, Service Pack 5) are installed.

The dependence of this product as a MS Windows only solution poses questions on the possibility of service deployment of this product. Heavy investment in UNIX servers at MIMAS precludes the deployment of a major NT service without additional funding, which is not available at the moment.

5 Content-Based Image Retrieval

Digitisation has provided the catalyst for the unparalleled swell in the volume of digital image collections with image capture, production and use now routinely occurring across a broad and diverse range of disciplines and subject fields. The potential application areas of computer-based image retrieval systems are numerous and diverse ranging from fabric and fashion design to geographic and medical-image based systems. The successful management of any information resource involves ensuring that the content is exploited efficiently, effectively, and is easily accessible by those users whose information needs it is intended to satisfy [Van Rijsbergen, 1979; Salton and McGill, 1983]. To extend the basic search facilities of the LANDMAP web service content-based image retrieval search capabilities are to be added as an experimental feature.

Content-based image retrieval provides an alternative approach to the traditional information retrieval paradigm. It is proposed as a natural extension of the advances made in the fields of computer vision and pattern recognition. The origins of the method have been attributed to the early experiments conducted by Kato [Kato, 1992] into the automatic retrieval of images by colour and shape feature [Eakins and Graham, 1999]. Content-based image retrieval involves a direct matching operation between the feature vectors of the query image and a database of stored images. The three most common image matching feature characteristics are colour (global and local), shape and texture (Ruiet al. 1999). The process involves computing a feature vector for the unique characteristics of the image. Similarity is computed by comparing the feature vectors of the images. The result of this process is a quantified similarity score that measures the visual distance between the two images represented by the feature vectors. This approach permits the objective analysis of pixel distribution and the automatic extraction of measurements from raw sensory data (Del Bimbo, 1999). This technique potentially creates new and exciting opportunities for retrieving the data. Nevertheless, it should be noted that there are still a number of significant research issues to be addressed if this technique is to prove successful beyond the research laboratory.

6 Conclusion

The provision of online orthorectified satellite imagery in a way accessible to a wide spectrum of users involves in-depth evaluation of a series of important considerations. The proposed shape of this service is an initial position based on past expertise in serving digital map data and current research. A successful service is one that responds to the reaction of the user to the service and responds to the changing needs of this user. In order to achieve this, the LANDMAP project and MIMAS are aiming to provide the best service possible,

fulfilling the discussed criteria, and as such are open to suggestions and proposals to introduce to the service.

7 Acknowledgment

The LANDMAP project is being funded by the UK Higher Education Funding Council of England (HEFCE), Higher Education Funding Council For Wales, Scottish Higher Education Funding Council (SHEFC), and Department of Education for Northern Ireland Joint Information Systems Committee (JISC).

8 References

CROSS P.A, Barnes J., Walker A.H., Muller J.P, Kitmitto K. and Morley J.G, GPS validation of IfSAR digital elevation models from LANDMAP. Remote Sensing Society 2000 Adding value to Remote sensing data 26th Annual Conference of the Remote Sensing Society 12th-14th September 2000 Leicester.

DEL BIMBO, A., Visual Information Retrieval. Morgan Kaufmann. 1999.

DOWMAN I.J., Kim J.R., Morley J.G., Chugani, K., Kitmitto K - Automated registration for data fusion and planimetric quality assurance. Remote Sensing Society 2000 Adding value to Remote sensing data 26th Annual Conference of the Remote Sensing Society 12th-14th September 2000 Leicester.

EAKINS, J. P., and Graham, M. E., Content-Based Image Retrieval: A Report to the JISC Technology Applications Programme. Institute for Image Data Research: Newcastle-Upon-Tyne. 1999.

KATO, T. Database Architecture for Content-Based Image Retrieval. In: Proceedings of Society of the Photo-Optical Instrumentation Engineers: Image Storage and Retrieval, 1662. 1992. San Jose, California, USA, SPIE, 1992.

MULLER J.P. Morley J.G, Walker A.W., Mitchell K., Kim J.R, Cross P.A., Barnes J., Dowman I.J., Smith A., Chugani K., and Neil Quarmby N., and Kitmitto K.. The LANDMAP project for the automated creation and validation of multi-resolution orthorectified satellite image products and a 1” DEM of the British Isles from ERS tandem SAR interferometry.

MORLEY J.G, Walker A.W, Kitmitto K., and Muller J.P, GIS techniques employed for the LANDMAP project. Remote Sensing Society 2000 Adding value to Remote sensing data 26th Annual Conference of the Remote Sensing Society 12th-14th September 2000 Leicester.
Rui, Y. and Huang, T. S. Chang, S-F., Image Retrieval: Current Techniques, Promising Directions, and Open Issues. Journal of Visual Communication and Image Representation, 10, (1), 1999, pp. 39-62.

SALTON, G. and McGill, M. J., Introduction to Modern Information Retrieval, McGraw-Hill Book Company, 1983.

VAN RIJSBERGEN, C. J., Information Retrieval, Butterworths, 1979.