A Paradigm Shift in Geospatial Image Access

PETER BECKER, Munich

ABSTRACT

With the explosion in the quantities of image data, conventional solutions for how to manage, process and distribute geospatial imagery have broken down. This has resulted in huge quantities of geospatial image data being available but not accessible, and new imagery not being fully exploited. With the value of imagery dependant on its accessibility, there is a great need for solutions that solve the many issues related to image accessibility. The unique aspects of geospatial imagery are discussed and the conventional geospatial process flow is investigated to identify the bottlenecks. After a review of the current issues, it can be identified that these can be resolved by the creation of **an** Image Processing Server that combines the stages of image processing and distribution. This has been implemented in the successful PromptServer product that additionally enables the creation of new application that further increase the usage and value of geospatial imagery.

1. THE GROWTH OF IMAGERY

Images are a natural way to store and transmit information and are a critical component in all our lives. This can be quantified by the fact that about 80% of our brain is devoted to vision. The volumes of geospatial imagery are increasing at a rate of about 30% per year. An increasing number of Satellites coupled with greater resolutions and an emerging acceptance of digital aerial cameras is driving this phenomena. In the amateur photography market the very rapid dominance of digital over analog film gives an indication as to the rapid changes we can expect for our aerial photography market.

2. THE IMPORTANCE OF IMAGERY

Imagery has always been a critical component of mapping and over 90% of all vector map data is extracted from imagery. Remote sensing, for many years had been dominated by the analytical processes of image analysis for the extraction of quantitative statistics and the interpretation by training operators for the extraction of features. With the increase in availability of imagery, such analysis and interpretation will continue to increase to provide the features for inclusion into databases required for further GIS analysis.

The very high resolution of today's commercially available satellite images and aerial imagery permits also greater detail identification. This has extended the use of imagery beyond the domain of remote sensing classification and statistical analysis into general applications where users identify features of interest. Accurate orthoimagery has become common place and being used as a backdrop to GIS and CAD systems. End users now make their own feature identification and measurements when vector data is not available. With the costs and time of acquiring imagery generally lower then collection of equivalent vector datasets, users often also use imagery to check the reliability of any existing vectors. The demand for geo-referenced imagery therefore rapidly increasing.

3. GEOIMAGERY SOURCES

A tremendous amount of satellite and aerial imagery data is already in existence. There are hundreds of thousands of TM, ETM, SPOT and ASTER scenes; not to speak of the unfathomable volumes of aerial photography and military datasets. The vast majority of this data has never been accessed at full resolution, or if so, only in a pinpoint manner.

The vast archives of existing aerial imagery and paper maps provide knowledge about our history and enable us to plan better for the future. Over time, these assets will be scanned to extend the existing digital archives.

A huge amount of image data - past, present and future - is acquired speculatively. Some of it is taken simply to take advantage of installed and operational sensors, while others to provide up-to-data data that is accessed only when the need arises. With the lowering of cost of data storage such data sets will further increase.

With the number of digital aerial cameras and satellites rapidly increasing, multispectral and hyperspectral image acquisition will become the norm and produce unprecedented large data volumes. We will see not only an increase in the number of images as such, but also in bit depth and number of bands ranging far beyond the 8bit multispectral imagery we are currently accustomed to.

Over the last few years, we have witnessed a staggering increase in the volumes of digital geospatial image data, yet the methods for image management, processing and distribution have not changed significantly.

4. THE IMAGE ACCESSIBILITY PROBLEM

Geospatial imagery is a vastly under utilized asset. The explosion in the volumes of geospatial imagery has caused significant problems for organizations associated with the collection, processing, dissemination and use of geospatial imagery. With current methodologies, managing, processing and distributing imagery is becoming untenable.

The vast majority of “available” digital imagery is not accessible to the end user. Imagery is stored on tapes or disks, and on propriety systems with formats that are difficult to access. Data in most organizations is spread over multiple platforms, across large networks allowing for only clumsy, fragmented data management. This imposed, but never desired data shifting severely handicaps the efficient use of imagery and production flows. Image Servers have been developed and deployed to assist in the distribution of pre-processed and mosaiced imagery, but this currently represents a small fraction of the imagery available.

The key issue in maximizing the value of geospatial imagery is to provide 'Accessibility'. It is providing access to the imagery, which increases use and demand which intern creates value and profit.

5. THE UNIQUENESS OF IMAGERY

Geospatial imagery can be differentiated from other types of geospatial data such as vectors and attributes in three distinct ways. Understanding this uniqueness is important when you consider the required solutions:

Volume - The data volumes are many magnitudes larger then comparable vector data and the rate of increase of image data will far outpace the increase of vector data. These data volumes have for many years limited not only the ability to store and access the imagery, but resulted in long processing times. It is apparent therefore that the methods to store and manage imagery will need to be different from other data.

Value - The value of imagery changes quickly with time. It rapidly de-values and then over time may increases again in value. A recent image of an area has highest value, but as time passes and rapid changes take place on the ground, the value of the image quickly decreases. Over time, the value of the imagery may increase as its historical value is realized. To maximize the value of imagery the latency or delay between image acquisition and usage must be reduced to a minimum.

Fixed -The raw image data never changes. Being a snapshot is actually one of its greatest assets. Once imaged, the data is fixed and unlike vector and attribute data, it is not modified. What is modified is how the imagery is geometrically and radiometrically enhanced and processed to create a product. Over time the georeferencing may be enhanced by the addition of control, DTM or better sensor models, while radiometric enhancement may be changed or improved for different application.

6. THE GEOSPATIAL IMAGERY PROCESS CHAIN

Geospatial image processing can be classified into the following stages:

- Data acquisition - During which the imagery and control is acquired and stored. Included is the direct digital imaging or scanning of analogue media.
- Parameter determination - During which the parameters that affect the geometric and radiometric processing of the imagery are determined. This includes the determination of the parameters for image trend removal, aerial triangulation, DTM extraction and mosaic line determination.
- Image Processing - During which the imagery is processed to create a defined product based on the parameters.
- Distribution – The vehicle by which the product is distributed to the end users.
- Interface - the stage, in which imagery is viewed, analysed or interpreted.

In many systems, the stages of Parameter Determination and Image Processing appear in a loop, such as orthorectification followed by trend removal, followed by mosaicing. Additionally the interface stage is often used to make measurements that affect the parameters which are feedback for the creation of revised products. Traditionally these steps result in multiple intermediate products further increasing data volumes.

Over the years, issues associated with each of these stages have been tackled in isolation.

- Data acquisition systems have improved over time; the development of better scanners and digital cameras being the most obvious.

- Better methods have emerged for parameter determination, such as Automated Bundle Block Adjustment or Automatic Mosaic Line determination.
- Image processing algorithms have improved as well as computer speed and capacity.
- Image Servers have been developed to facilitate client/server based distribution of the imagery.
- Interfaces have improved and become richer, with 3D and Stereo now becoming more prevalent.

These independent improvements help to increase the availability of imagery, but they are incremental and do not address the key issues of geospatial imagery.

To address the problems associated with huge data volumes, industry has introduced larger data storage media and improved compression techniques. The rate of technological development in data storage, processing and transmission are quite staggering. Computing power and data storage capacities double about every 18 months. There are predictions that deployed bandwidth will grow 300% each year. In comparison, data compression algorithms have improved very slowly. Wavelet (e.g. JPEG2000) compression has increased by a factor of about 3 over JPEG in the last 10years (10% per year). The compression of imagery does decrease data volumes, but degrades the quality of the imagery, increases the processing time and creates numerous problems for the further analysis and processing of the imagery.

The volumes of geospatial imagery may be very large, but they are also compounded due to redundancy in multiple forms of the same image. Conventional image processing techniques create products as the result of linear workflows. Imagery enters the workflow, is processed and outputs the processes often to forms input to the next stage of production. Any changes in parameters result in partial or complete reprocessing of the complete data set which takes time and creates many semi redundant data sets. This compounds the dilemma of data volumes and increased processing time, further devaluing the imagery.

The increase in the number of bands and bit depths, allows for the production of more diverse products, with different users requiring different band combinations or radiometric and geometric enhancement. The conventional image processing procedures required for standard deliveries can not cope with these new requirements as each product adds to the volume levels, and is further hampered by preprocessing and distribution. Neither will faster processors, larger memories and speedier data transmission or compression provide the answer. With conventional procedures, user demands and expectation will continue to outdo the systems capabilities.

7. THE PROMPT IMAGE PROCESSING SERVER

It can be identified that the accessibility issue is caused primarily due to the separation of the image processing and distribution stages in the imagery chain. This separation issue can be solved by using an Image Processing Server that combines the image processing and distribution stages.
Utilizing an Image Processing Server provides huge advantages as it resolves the majority of the imagery issues, making the huge quantities of imagery accessible to large numbers of uses in multiple environments.

An image processing server works by enabling the creation of a catalog that stores image meta data as well as processing parameters in a database. Voluminous imagery can be stored in its raw (or preprocessed) form. The administrator of the catalog is able to create different catalogs from the same data for different applications and update or append these as and when required. To the end

users each catalog appears as a virtual product which is accessed using client/server computing. Requests for imagery covering a required extents and ground sample distance are transmitted by the client (interface) application to the server. The server extracts the required imagery, processes it 'on-the-fly' according to the parameters in the database, and returns the imagery to the user. Users can also control parameters on the server that affect the processing such as the compression factor for data transmission or sub-sampling. With optimized algorithms and today's processing capabilities, such processes can be achieved with sub-second response times, and future technological advances will further increase speed.

PromptServer has been developed as an Image Processing Server that fully implements the concepts of an Image Processing Server and so changes the way imagery is managed, processed and distributed. As PromptServer does not require the creation of one large file or database that stores a pre-processed product, the implementation possibilities are extensive and the solution provides huge advantages to organizations that have or need access to geospatial imagery. Additionally PromptServer opens up a number of new methods and applications for geospatial imagery.

One concept that can now be implemented is that of ‘Graded Data’. In a conventional process flow, a product is created to a specified grade. To get the image product to the required grade may take a lot of time, during which the end users has no access to the image. After production and QC of the product, the user is provided access, by which time much of the value of the imagery will have gone. PromptServer enables the imagery to be published as soon as it is available in a digital form. Initially the data may be only enhanced and geo-referenced rudimentarily, but is still accessible for some applications. As the processing progresses and parameters for enhancement and geo-referencing are refined, the grade of the product increases until such time that the required grade is achieved. This provides access to the imagery quickly to maximize value, while providing better grades as they become available. This has a lot of implications on the time scales and delivery mechanism for large engineering projects. There are also many applications for Graded Data such as disaster relief and emergency response.

PromptServer works directly with a large range of image formats, removing costly data ingest and reformatting. This has certain advantages including the ability to quickly implement the system within organizations, as it requires no restructuring or reformatting of the existing data. As new data becomes available, it can be added to the catalog increasing the extent or replacing areas with up-to-date imagery.

A large range of different raster data types are supported ranging from 1bit to 32bits per channel in varying formats with and without lossy and lossless compressions. Catalogs consisting of scanned cadastral documents, to satellite imagery through to detailed terrain models can be handled. When working with the raw data, uncompressed image formats are preferred such that data can be analyzed with no data loss.

As no special formats are required is it fully compliant with legacy production environments. Such systems are often used for parameter determination such at AT. PromptServer can also be used directly in a production environment to output processed products such as tiled orthoimages for writing to CDs etc or ingest into conventional image servers.

In standard image processing procedures were a single coverage is created, the overlapping data is destroyed in the mosaicing process. With PromptServer the integrity of overlapping imagery can be retained and resolved in different ways. The images can be mosaiced using seamlines that are generated automatically or manually to provide seamless imagery. Alternatively images can be

merged in user specified orders, for example specifically dated or higher accuracy imagery displayed on the top. When multiple similar overlapping images, (a common occurrence with aerial imagery) are available the merge order can also be controlled by the location of the view extent such that the most nadir image of an area can be displayed when a user zooms to a specific location. Control can be provided to enable the user to change between the multiple views of the same location. In large scale photography of urban areas this enables the observation of different sides of buildings, thereby solving the problems of occlusions. This technique substantially increases the value of imagery for applications such as emergency services or utility companies that need to observe features on the sides of the building or close to their base.

The bit depth of remote sensing sensors is steadily increasing. While older satellites and scanners had a bit depth of 8 or less, newer scanners and satellites have 11-14bits/channel. This increase in image quantization enables newer digital procedures to extract more detail. Many applications though can only handle 8bit/channel imagery and our own eyes can generally only see about 30 different shades of a color, so a conversion of the data is required. PromptServer provides a number of features for the on-the-fly enhancement of the imagery to create 8bit/channel image products that provide good overall contrast or more details in shadow or highlight areas. The iCOMP feature provides advanced radiometric corrections by adaptive localized image enhancement and color balancing. This maximizes the visual image content and removes trends in the imagery caused by light fall off, vignetting effects, haze or different angles of incident of light. The result is high quality detailed imagery.

Pan sharpening is a process of merging the spectral resolution of one image with the spatial resolution of another. PromptServer contains a specially developed proprietary pan sharpening algorithm that is optimized for lightning-fast processing, enabling it to process the imagery on-the-fly. It works on high bit depths, enabling 16bit/channel imagery to be pan sharpened to extract detail within the full dynamic range of the imagery. On-the-fly pan sharpening enables multiple band combinations to be processed while reducing the data storage volume. For example from a single QB scene consisting of 4spectral bands with a 2.4m GSD and a pan band with 0.6m GSD, a range of images can be created, including: 0.6m Natural Color, 0.6m Infrared or 2.4m NDVI classifications, while not increasing data volume.

Different method for on-the-fly georeferencing of the imagery are available, from simple affine or polynomial warping to detailed correction grids and orhtorectification. The more precise orthorectification includes the use of digital elevation image in the computations to improve the accuracy especially in areas of undulating terrain. The on-the-fly orthorectification is a feature to enable the creation of Graded Data products based on different camera orientation and terrain data.

Compression is most important not for storage but for transmission of data over limited bandwidth networks. For users requiring access to imagery over low bandwidth networks, PromptServer supports dynamic compression, independent of whether the data is stored compressed or uncompressed. After processing the data and prior to transmission, PromptServer can compress the data using a range of codec's. The format and compression factor can be changed by the users. This enables the user to set high compression values while browsing the imagery and then turn down or off the compression for detailed analysis.

No special viewer is required for PromptServer. Extensions and PlugIns exist for all major CAD and GIS applications so users can quickly add imagery to their existing applications. Enterprises with different CAD and GIS system can then all share the same imagery. An OGC compliant WMS also enables access to imagery over standard web-browser interfaces.

A whole range of additional features are included, such server side caching and a gateway to enable servers to run in parallel for scalability. The PromptPlotServer also enables the creation of high resolution large format image plots without the requiring the client application to load large quantities of images or create large plot files.

8. THE IMAGE ACCESSIBILITY SOLUTION

PromptServer takes advantage of the three unique features of imagery and resolves issues related to geospatial imagery in a single step. The advantages of such a solution are manifold and include but are not restricted to:

- Providing near instant access to unlimited large image data
- Processing images directly to the end-user, with no need for image recording. Any view required in the future may be simply re-created. Data volumes are therefore substantially reduced as data redundancy is removed.
- Graded data products can be produced. These enable preliminary data to be provided to the end user as soon as the imagery is available in digital form and possibly increasing the grade over time.
- Data revision can be performed easily. As newer data sets become available they can be seamlessly added to the catalog without regeneration and distribution of a complete product.
- Image quality is improved as the data stored on the server is processed in a single step with no intermediate images and no requirement for compression.
- Meta data is provided to the end users on all imagery, parameters and processes used. This information has high value to many users and enables the product to be verified. This maintains data integrity avoiding contamination through unknown intermediate processing of unknown quality and so maintains the legal status of image content.
- The processing capabilities are expandable to new image processing techniques therefore allowing easy integration into future applications or proprietary data processing.
- Direct integration into most CAD/GIS applications provides for enterprise applications as well as OGC WMS compliance
- Data access can be controlled while data access logs can be used to analyze data usage.
- The maintenance of primary data permits temporal and comparative image processing such as change detection as well as providing data for applications such as stereo display.

9. CONCLUSION

The many issues faced by most organizations currently with large quantities of imagery is a result of the separation of image processing and distribution stages. An Image Processing Server methodology provides substantially improved accessibility to imagery. It resolves a large number of the issues for geospatial imagery as well as opening up a range of new possibilities. Looking at the process chain for geospatial image processing, we expect to see in the future a clearer separation between the processes of parameter determination and image processing, and a merge between image processing and distribution.

10. REFERENCES

Gary, J. (2000): Computer Technology Forecast for Virtual Observatories. Technical Report MSR-TR-2000-102. Microsoft Research.