White Paper

Geospatial Archive and Exploitation System

Chris Duhring, SGI

1.0 Introduction

This paper provides an overview of a spatial information analysis system that supports the storage, retrieval, and analysis of digital spatial information. This is a family of solutions that offers an information user a scalable approach for the storage and analysis of digital spatial information captured in any of several forms, including documents, photographs, maps, videos, and audio recordings. The spatial information analysis system provides information users with services for the discovery and retrieval of digital information and applications for the analysis of the digital information. These services allow a user to select appropriate information by browsing through the available information, by performing queries on attributes of the information, or by combining these methods. The spatial information analysis system enables the user's gathering of the information, either by directly accessing the information or by downloading or copying the information.

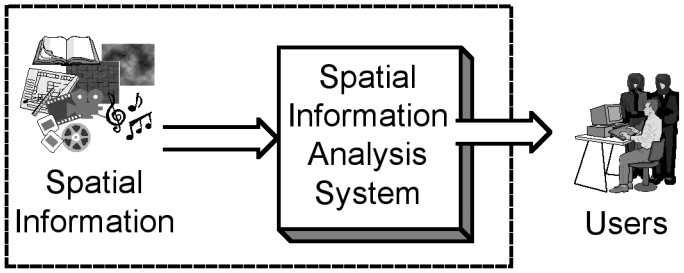

Fig. 1—Spatial Information Analysis System Overview

2.0 Capabilities

A spatial information analysis system provides users with the capability to select, retrieve, and analyze digital spatial information. Inherent in these capabilities are the services to ingest, store, and manage the information. These capabilities and services are provided in a scalable family of solutions that supports an increasing number of users and an expanding amount of information. The core components of the spatial information analysis system are shown in Fig. 2 and are configured to address a specific community's requirements. The core components include:

- Archive—The archive ingests imagery and geospatial information, stores the information, and serves the information out to the services for delivery to the users. The archive may manage the stored information on a combination of online, near-line, and offline systems in a hierarchical system or may be configured for entirely online storage. The type of storage and the interfaces and methods for receiving data into the archive are configured to match the local requirements.
- Catalog—The catalog extracts descriptive attributes from the information received into the archive and stores these attributes as metadata and preview images that support the user's selection of appropriate information through query and browse actions. In additional to a simple index, the catalog may also include an associative index oriented around user-specified topics, such as facilities or equipment. In this type of index, the user associates information with a specific topic, such as a facility, and enters key attributes to define the association. Using an associative index, spatial information can then be found in relationship to a topic, such as a specific facility. The information types that the catalog can ingest and the nature of the indices are configured to meet the users' needs.
- Catalog service—The user interfaces with the catalog service through a standard off-the-shelf Web browser. Through the Web browser, the user forms and submits queries or browses through indices and browses through the metadata displaying the information attributes as well as preview images of the information. The user selects the information for analysis and specifies delivery parameters such as direct download or media delivery, conversion and/or processing parameters such as geoextent extraction (clipping), or a combination (e.g., mosaicing, merging, blending). The catalog service interfaces with the catalog and the content service to satisfy user requests. The catalog service may be configured to support a specific look and feel and to provide extended capabilities such as user authentication.
- Content service—The content service delivers information as directed from the catalog service, supporting a combination of online and offline delivery mechanisms. The content service may be configured to support a variety of delivery methods, conversions, and preprocessing options.

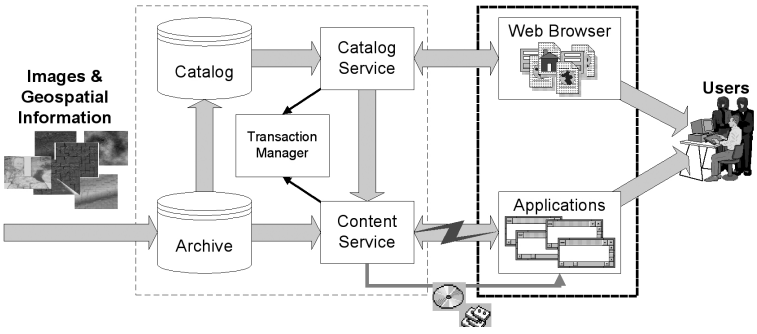

Fig. 2—Spatial Information Analysis System Components

- Applications—The applications provide users with the tools to analyze the imagery and geospatial information. The applications are selected and configured based on the types of information available and the objectives of the analysis.

- Transaction manager—The transaction manager is the foundation for coordinating and controlling the responses to each user request. In security-conscious environments, the transaction manager is the key to assuring that users are granted access only to authorized information. In charge-back or commerce situations, the transaction manager is the focal point for determining charges by user or organization. The transaction manager is configured according to the local access restriction and fee models.

A spatial information analysis system can provide users with a wide variety of capabilities in the archive, catalog, catalog service, content service, applications, and transaction manager components. The specific capabilities required by a group of users depend entirely on the mission and objectives. The following is a partial list of capabilities that can be provided in a spatial information analysis system.

2.1 Archive
- Support the hierarchical management of online, near-line, and offline storage
- Provide automated metadata (attribute) extraction

2.2 Catalog
- Support query and discovery by aspatial attributes (e.g., time, keywords, or text)
- Support query and discovery by spatial location (spatial extension to database)
- Support query and discovery by facility/equipment association (association folders)
  - —Facilities (targets) are defined by identifier, name, location, imaging requirement, exploitation requirement, report requirement, or distribution requirement
  - —Equipment is defined by identifier, resources and status code, information, or image ID(s)
- Support the display of preview/overview thumbnails
- Provide a user interface for manual population/update of information into catalog

2.3 Catalog Service
- Provide a user interface for object selection
  - —Indexed browse—provide hierarchical indices of information by user-selected attributes (e.g., date or identifier)
  - —Ad hoc query—provide free-text query capability of user-selected attributes
  - —Placename-based indices—provide hierarchical indices of information based on facility/equipment/placename association
  - —Map-based ad hoc query—provide graphical or textual methods for defining a spatial query
- Support a standing query capability for the automated identification of applicable information
  - —The user profile defines which user or group of users requires specific information on specific topics at regular intervals
  - —If the information is not in system on query, generate collection request
  - —If the information is in system, distribute to user or group of users or notify of availability
- Provide a published API for spatial object selection
  - —Name/identifier search
  - —Aspatial attribute queries
  - —Spatial location queries
  - —Placename association (gazetteer)
- Provide brokers for distributed access to other archives
- Provide metadata caching for improved performance

2.4 Content Service
- Provide automated preprocessing for retrieval/processing/streaming (RRDS)
- Provide a user interface for object retrieval
  - —Click-to-stream/download (direct access) delivery
  - —Order form for near-line or offline (asynchronous) delivery
    - Packaged download (tar, zip, etc.)
    - Packaged on media
- Published API for spatial object retrieval
  - —Streaming/direct access
  - —Direct download (as is)
- Tailoring
  - —Format conversion
  - —Cropping/clipping
  - —Geoextent extraction (chipping)
  - —Generalization/decimation
- Customization
  - —Rectification
  - —Projection
  - —Extraction (elevation, features, etc.)

2.5 Applications
- Retain working files/user preferences initialize each session
- Display color, false color, and grayscale imagery
- Pan, zoom, reduce, move, and rotate
- Display geographic location/jump to geographic location
- Contract adjustment, brightness, sharpness, dynamic range, and despeckling

- Edge or border enhancement, mirror, haze filtering, false color, and positive/negative flip
- Two-image fade-in, fade-out, flipping, split-screen display, stereo, and image difference
- Multispectral: band ratio, false color, NDVI vegetation index, and supervised and unsupervised classification
- Annotation: marks, points, arrows or orientation, circle, polygon, grid, and text
- Count each type of annotation and display count
- Drawing tools: examine, modify, recover, select font size, color, line style, emphasize, display/hide, select, and move
- Measurement: display location of point, measure distance between points, and calculate area within a circle or polygon
- Display pixel value, histogram, and spectral coding
- Overlay object files (vectors, shapes) on imagery according to geographic location
- Create, modify, delete vectors and overlay objects
- Create, store, examine vector, shape, annotation layers
- Drape imagery over DEM for 3D display
- Feature extraction, identification, and classification using monoscopic and stereoscopic images
- Report editing, modification, archive, and retrieval in English and local language
  - —Integrate textual description, image, shape objects, tables, etc.
  - —Catalog reports by geographic location, target description, country, data source, and date
  - —Reports from self-defined format, predefined tables, predefined military equipment report, and reports generated by database report
  - —Output in Postscript MS Office formats, TIFF, and HTML
  - —Import data in HTML

2.6 Transaction Manager

- E-commerce transaction tracking and billing
- Performance monitoring (user response times, background/automated processing, distributed queries, etc.)
- Prioritize user responses

3.0 Implementation

Each spatial information analysis system is implemented by selecting and integrating commercial off-the-shelf (COTS) components to meet the specific user's requirements. The integration of COTS components often requires customization of the components and the addition of customized middleware to implement a specific user's data flow.

3.1 Sample Configurations

The following are some sample implementations based on typical driving requirements.

Sample 1: Provide spatial imagery and information via a local network to existing and new IRIX®, Linux® , UNIX®, and Windows® workstations.

This sample configuration supports the archive, dissemination, and exploitation of spatial imagery and information with low to medium time criticality. Users access information only by georegion, publication date, or information type. The information is received from production systems electronically and/or on media. All users and their workstations are located on the same local area network, and the workstations are a mix of IRIX, Linux, other UNIX, and Windows platforms.

In this example (Fig. 3), the spatial imagery and information is received by the archive electronically or on

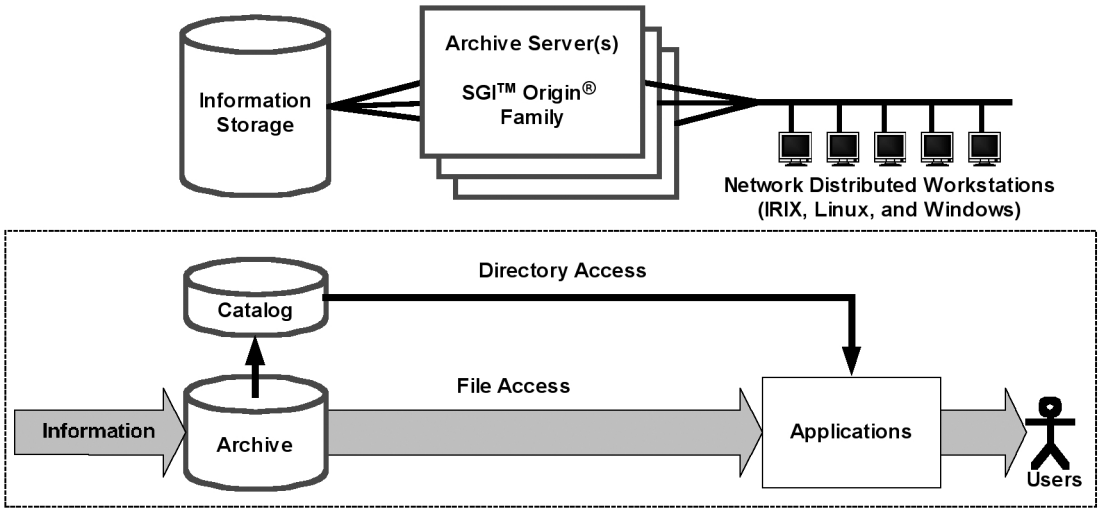

Fig. 3—Local Area Network Information Management

media, and the images and information are copied into the storage area. Indices are built in the catalog to allow cross-reference of the information, for example, by georegion, date imaged/published, and data type. In the simplest form, these indices can simply be well-ordered directory and file structures, where one structure actually holds the data and the other indices provide soft-link cross-references. For spatial information especially, but also for images, these file-based catalogs are sufficient for the user to locate applicable information.

Once the information is available in storage and the catalogs have been created, the users can access the information through exploitation applications already loaded and running on their workstations. Their access to a file-based structure is natural and already supported by the exploitation applications to locate the applicable information against any of the prebuilt directory structures. This type of solution supports local access to spatial imagery and information within a simple management structure that is easily administered and updated as information is received or updated.

Sample 2: Provide spatial imagery and information via a wide area network to existing and new IRIX, UNIX, Linux, and Windows workstations.

In sample configuration that addresses these requirements (Fig. 3), information is received electronically or on media. Because the users need to search for the information by any of several characteristics, the catalog applications extract metadata that describes the information in catalogs that will be searchable.

Users, through a combination of Web browsers and exploitation applications, search for the information through the services on the archive server. The users order the desired information, and the archive server packages and delivers the information as specified. This packaging may include tailoring the information to the user's specifications, including geoextent extraction, resampling/decimation or generalization/intensification, mosaicing or merging, projection or rendering, format conversion, compression/decompression, or other services. These tailoring services are dependent on the capabilities inherent in the information, matched to the processing capabilities of the archive server. The information, native or tailored, is delivered to the users' workstations as files and is loaded into the appropriate application. This transfer is typically implemented as FTP and the launch of helper applications common on most Web sites. This type of solution supports distributed access to information where time is not as critical and where a wide range of information types and applications must be supported.

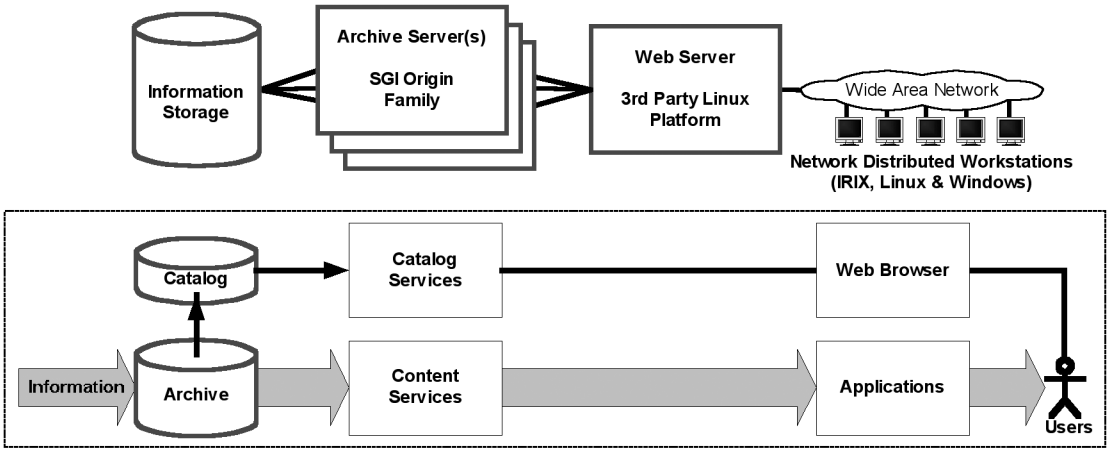

Fig. 4—Wide Area Distributed Information Management

In a second example, the users perform archive, dissemination, and exploitation with low time criticality. Users access information by any of several characteristics and may not limit information used to a single characteristic (random search). The information is received from production systems electronically and/or on media. The users and their IRIX, Linux, other UNIX, and Windows workstations are distributed across a wide area network.

Sample 3: Provide time-critical spatial imagery and information and direct-access applications to local users.

The third example supports a time-critical mission in which users must review updates to imagery or information within seconds or minutes of receipt. The local

users are notified of the arrival of information according to any of several characteristics. Some information types are supported as direct access by the applications (tiled imagery, feature, and point databases). There are also remote users, just as there were in the second sample, that require access to the information on a non-time-critical basis.

The example implementation (Fig. 5) supports both local and remote users. The time-critical nature of this problem requires an architecture in which the local exploitation is tightly coupled to the information through very high bandwidth communications. The SGI™ Geospatial Exploitation System (aka GroupStation) offers a unique solution to this problem. In this implementation, the information is received and ingested and catalogs are created to support search operations. The difference is that the information is preprocessed into the shared storage in a form that supports direct access by the user's exploitation applications. The user is notified that the information is available and it is queued for access, or the user searches the archive for the information needed and selects the information to be exploited. The application is then connected to the information and directly interacts with the information. No information is packaged or copied, and all applications have equal access to the information. This significantly improves delivery latencies and also supports exploitation collaboration.

All distributed users have access to the original information and to the information created by the local users. This is supported in the same way as in Sample 2. This sample implementation supports local time-critical access to information, collaborative exploitation, and dissemination of original and exploited information to distributed users.

3.2 Sample COTS Components

The following are some of the COTS components that can be combined to provide a spatial information analysis system. The table following this list compares the capabilities of each of these components.

- Autometric DataMaster™—DataMaster supports the ingest and processing of imagery. Originally developed to provide capabilities for the ingest of imagery into a shared archive, DataMaster provides strong capabilities for preprocessing and delivering imagery products. DataMaster also provides capabilities for the management, query, and retrieval of imagery. (Product description: *www.autometric.com/AUTO/PRODUCTS/DataMaster/*)
- Compusult Percipio—Software environment for the management of large disparate data sets suitable for data collection, maintenance, and distribution over wide area networks. Percipio is a part of the Web Enterprise Suite, which also includes the Map Manager built above ESRI's Internet Map Server

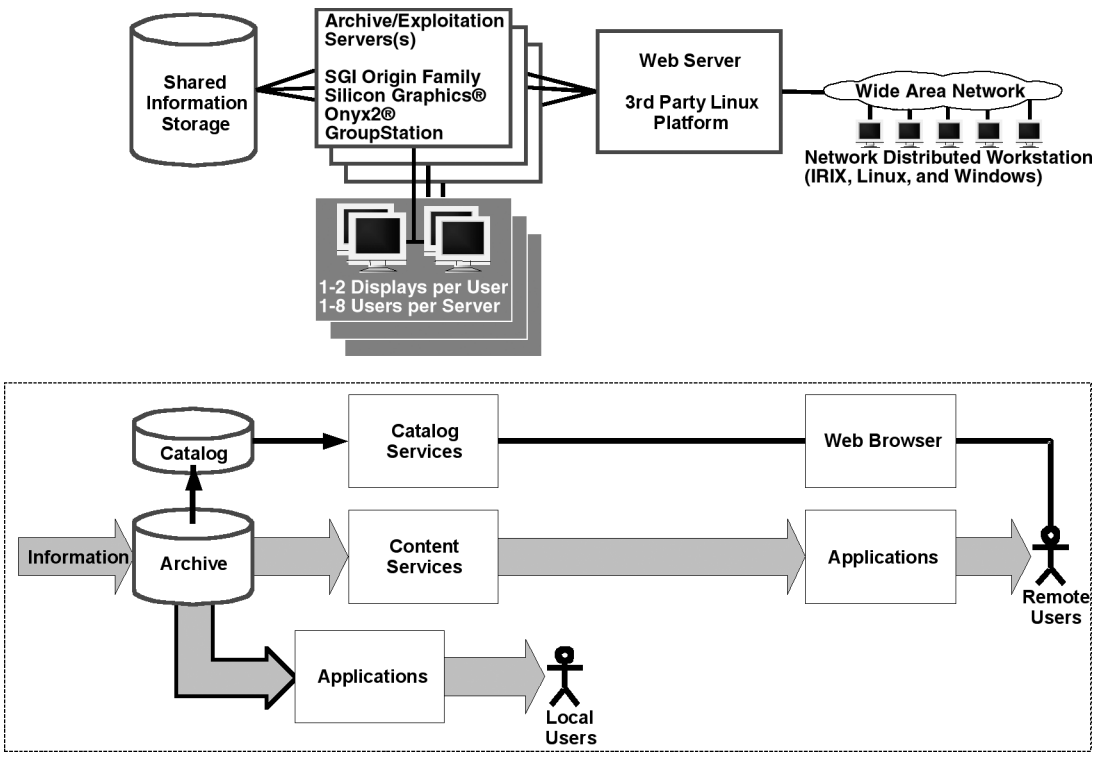

Fig. 5—Integrated Data and Applications with Dedicated Workstations

and the Meta Manager, which provides search and discovery tools for Z39.50 libraries. [Product description: *www.webenterprisesuite.com*]

- SGI™ Digital Library—User interface client and middleware that implements a digital asset manager above a COTS database. [Representative implementation: *http://tamudl.tamu.edu:5005/diglib*]
- Core Software Terrasoar™—User interface client and middleware that implements a geospatial data manager above a COTS database. [Product description: *www.coresw.com/software_frame.html*]
- ImageLinks Active Archive™—Active Archive is an imagery-management system that combines query, retrieval, and delivery capabilities with processing and customization services. [Product description: *www.imagelinks.com/pages/products/active.htm*]
- Sensor Systems RemoteView—RemoteView is an interactive application for viewing and reporting on government and commercial remote sensing data. RemoteView supports the import, processing, and analysis of remote sensing imagery. [Product description: *www.sensor.com/remoteview.html*]
- ESRI™ ArcView™/Image Analysis—The Image Analysis extension to ArcView provides geographic imaging tools in combination with the ArcView GIS and mapping tools. This enables users to use digital imagery for data visualization, data extraction/creation, and analysis. The Image Analysis extension was developed jointly by ESRI and ERDAS and provides a direct path to ERDAS Imagine for more sophisticated geographic imaging needs. [Product description: *www.esri.com/software/arcview/extensions/imageext.html*]
- ERDAS Imagine®—Imagine provides interactive desktop tools for remote sensing and complex image analysis. Imagine supports the import, processing, and analysis of remote sensing imagery in combination with topographic and cultural data. [Product description: *www.erdas.com/software/main.asp*]

Requirement	Autometric DataMaster[1]	Core Software Terrasoar[2]	Compusult Percipio[3]	SGI Digital Library	ImageLinks Active Archive[4]	Sensor Systems RemoteView[5]	ESRI ArcView/ Image Analysis[6]	ERDAS Imagine[7]
Archive								
·Support shared access to information by 1-N users	✔ (Heavy client)	✔	✔	✔				
·Support the hierarchical management of online, near-line, and offline storage	✔		✔	✔	✔			
·Provide automated metadata (attribute) extraction	✔		✔		✔			✔ (Local store)
Catalog								
·Support query and discovery by aspatial attributes (e.g., identifier, time, keywords, text)	✔	✔		✔				✔
·Support query and discovery by spatial location (spatial extension to database)	✔	✔		✔ (Future)	✔			✔
·Support query and discovery by facility/equipment association (association folders)	✔			✔				
—Facilities (targets) are defined by identifier, name, location, imaging requirement, exploitation requirement, report requirement, distribution requirement	Identifier only			✔				
—Equipment is defined by identifier, resources and status code, information, image ID(s)	Identifier only			✔				
·Support the display of preview/overview thumbnails	✔	✔		✔	✔			
·Provide a user interface for manual population/update of information into catalog		✔		✔				✔
·Provide a published API for the population of metadata (attributes)			✔					
Catalog Service								
·Provide a user interface for object selection	✔	✔		✔	✔			
—Indexed browse—provide hierarchical indices of information by user-selected attributes (e.g., date, identifier)				✔				
—Placename-based indices—provide hierarchical indices of information based on facility/equipment/placename association				✔	✔			
—Ad hoc query—provide free-text query capability of user-selected attributes	✔	✔		✔				✔
—Ad hoc gazetteer—provide placename-based ad hoc queries		✔						

Requirement	Autometric DataMaster[1]	Core Software Terrasoar[2]	Compusult Percipio[3]	SGI Digital Library	ImageLinks Active Archive[4]	Sensor Systems RemoteView[5]	ESRI ArcView/ Image Analysis[6]	ERDAS Imagine[7]
—Map-based ad hoc query—provide graphical or textual methods for defining a spatial query		✔		✔ (Future)	✔			✔
—Interactive overview of selected objects for applicability	✔	✔		✔	✔			
• Support a standing query capability for the automated identification of applicable information			✔					
—The user profile defines which user or group of users requires specific information on specific topics at regular intervals			✔					
—If the information is not in system on query, generate collection request								
—If the information is in system, distribute to user or group of users or notify of availability			✔					
• Provide a published API for spatial object selection	✔ (CLI)		✔	✔				
—Name/identifier search				✔				
—Aspatial attribute queries	✔ (CLI)			✔				
—Spatial location queries	✔ (CLI)			✔ (Future)				
—Placename association (gazetteer)				✔				
• Provide brokers for distributed access to other archives	✔	✔	✔		✔			
—Homogeneous catalogs	✔	✔			✔			
—Z39.50			✔					
—OGC Catalog Spec								
• Provide metadata caching for improved performance	✔							
Content Service								
• Provide automated preprocessing for retrieval/ processing/streaming	✔							
—RRDS	✔					✔		
—AOI geoextent extract	✔							
—Stereo pair matching	✔							
• Provide a user interface for object retrieval	✔	✔	✔	✔	✔			
—Click-to-stream/download (direct access) delivery			✔	✔				
—Order form for near-line or offline (asynchronous) delivery	✔	✔	✔		✔			

Requirement	Autometric DataMaster[1]	Core Software Terrasoar[2]	Compusult Percipio[3]	SGI Digital Library	ImageLinks Active Archive[4]	Sensor Systems RemoteView[5]	ESRI ArcView/ Image Analysis[6]	ERDAS Imagine[7]
• Packaged download (tar, zip, etc.)	✔		✔		✔			
• Packaged on media	✔		✔					
• Published API for spatial object retrieval	✔		✔					
—Streaming/direct access	✔		✔					
—Direct download (as is)	✔		✔					
• Automated tailoring	✔				✔			
—Format conversion	✔				✔			
—Cropping/clipping					✔			
—Geoextent extraction (chipping)	✔				✔			
—Generalization/decimation	✔				✔			
• Automated customization	✔				✔			
—Haze/contrast/brightness adjustment	✔				✔			
—Rotation	✔				✔			
—False color band selection	✔				✔			
—Planar rectification	✔				✔			
—Perspective transformation								
—Map projection					✔			
—Registration					✔			
—Mosaic					✔			
—Annotation and gridding					✔			
—Extraction (elevation, features, etc.)								
Applications								
• Retain working files/user preferences initialize each session					✔	✔	✔	✔
• Import/export national and commercial formats					✔	✔	✔	✔
• Interactive display						✔		✔
—Display color, false color, grayscale imagery						✔		✔
—Pan, zoom, reduce, move, rotate						✔		✔
—Display geographic location/jump to geographic location						✔		✔
—Contrast adjustment, brightness, sharpness, dynamic range, despeckling						✔	✔	✔

Requirement	Autometric DataMaster[1]	Core Software Terrasoar[2]	Compusult Percipio[3]	SGI Digital Library	ImageLinks Active Archive[4]	Sensor Systems RemoteView[5]	ESRI ArcView/ Image Analysis[6]	ERDAS Imagine[7]
—Edge or border enhancement, mirror, haze filtering, false color, positive/negative flip						✔	✔	✔
—Two-image fade-in, fade-out, flipping, split-screen display, stereo, image difference						✔		
• Interactive tailoring						✔	✔	✔
—Format conversion						✔	✔	✔
—Cropping/clipping						✔		✔
—Geoextent extraction (chipping)						✔		✔
—Generalization/decimation						✔		✔
• Interactive customization						✔	✔	✔
—False color band selection						✔	✔	✔
—Planar rectification						✔		✔
—Perspective transformation						✔		✔
—Map projection						✔	✔	✔
—Registration						✔	✔	✔
—Mosaic						✔	✔	✔
—Extraction (elevation, features, etc.)							✔	
• Analysis tools						✔	✔	✔
—Multispectral: band ratio, false color, NDVI vegetation index, supervised and unsupervised classification						✔	✔	✔
—Measurement: display location of point, measure distance between points, calculate area within a circle or polygon						✔		✔
—Display pixel value, histogram, spectral coding						✔		✔
—Drape imagery over DEM for 3D display						✔		✔
—Feature extraction, identification, and classification using monoscopic and stereoscopic images								
• Annotation						✔	✔	
—Annotation templates						✔		
—Annotation: marks, points, arrows or orientation, circle, polygon, grid, text						✔	✔	✔
—Count each type of annotation and display count						✔		✔ (Custom)

Requirement	Autometric DataMaster[1]	Core Software Terrasoar[2]	Compusult Percipio[3]	SGI Digital Library	ImageLinks Active Archive[4]	Sensor Systems RemoteView[5]	ESRI ArcView/ Image Analysis[6]	ERDAS Imagine[7]
—Drawing tools: examine, modify, recover, select font size, color, line style, emphasize, display/hide, select, move						✔	✔	✔
—Overlay object files (vectors, shapes) on imagery according to geographic location						✔	✔	✔
—Create, modify, delete vectors, and overlay objects						✔	✔	✔
—Create, store, examine vector, shape, annotation layers						✔	✔	✔
• Report editing, modification, archive, and retrieval in English and local language						Cut and paste into desktop tools	✔	Cut and paste into desktop tools
—Integrate textual description, image, shape, objects, tables, etc.							✔	
—Catalog reports by geographic location, target description, country, data source, date								
—Reports from self-defined format, predefined tables, predefined military equipment report, reports generated by database report							✔	
—Output in Postscript MS Office formats, TIFF, and HTML							✔	
—Import data in HTML								
Transaction Manager								
• Transaction tracking	✔	✔	✔	✔ (Future)				
• Transaction billing			✔					
• Performance monitoring (user response times, background/automated processing, distributed queries, etc.)								
• Prioritize user responses								
Platforms Supported								
• SGI IRIX	✔	✔		✔		✔		✔
• Linux	✔	✔			✔			
• Sun™ Solaris™	✔	✔	✔			✔		✔
• HP-UX		✔						
• IBM® AIX®								✔
• Compaq Tru64			✔					
• Windows NT® and Windows 2000		✔	✔			✔	✔	✔

[1] DataMaster User's Guide, Version 3.0
[2] Terrasoar 4.2 User Server Administration Guide; Data Server Administration Guide
[3] Percipio Overview; *www.compusult.nf.ca/percipio/html/overview.html*
[4] ImageLinks Active Archive™ Technology White Paper, August 2001
[5] Complete User Guide for RemoteView Professional for UNIX Version 2.30
[6] ArcView GIS Brochure; ArcView Image Analysis Extension (Version 1.1) Brochure
[7] ERDAS Field Guide, Fourth Edition, 1997; What's New—ERDAS Imagine v8.4, November 1999; What's New—ERDAS Imagine v8.5, June 2001

Corporate Office
1600 Amphitheatre Pkwy.
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America 1(800) 800-7441
Latin America (52) 5267-1387
Europe (44) 118.925.75.00
Japan (81) 3.5488.1811
Asia Pacific (65) 771.0290

3201 (12/01)

J13255