An Integrated Global GIS and Visual Simulation System

Peter Lindstrom   David Koller   William Ribarsky
Larry F. Hodges   Augusto Op den Bosch   Nick Faust

Graphics, Visualization, & Usability Center
Georgia Institute of Technology

Abstract

This paper reports on an integrated visual simulation system supporting visualization of global multiresolution terrain elevation and imagery data, static and dynamic 3D objects with multiple levels of detail, non-protrusive features such as roads and rivers, distributed simulation and real-time sensor input, and an embedded geographic information system. The requirements of real-time rendering, very large datasets, and heterogeneous detail management strongly affect the structure of this system. Use of hierarchical spatial data structures and multiple coordinate systems allow for visualization and manipulation of huge terrain datasets spanning the entire surface of the Earth at resolutions well below one meter. The multithreaded nature of the system supports multiple windows with independent, stereoscopic views. The system is portable, built on OpenGL, POSIX threads, and X11/Motif windowed interface. It has been tested and evaluated in the field with a variety of terrain data, updates due to real-time sensor input, and display of networked DIS simulations.

1   INTRODUCTION

This paper reports on significant progress in our efforts to design and construct a real-time visual simulation and geographic information visualization system, named VGIS (Virtual Geographic Information System) [17, 30]. VGIS supports the accurate depiction of terrain elevation and imagery, in addition to features such as ground cover and trees, moving vehicles, buildings and other static objects, roads, and atmospheric effects. Thus an entire environment composed of heterogeneous parts must be simulated and integrated at rendering time. The system must be set up to efficiently manage this integration and, ultimately, to manage dynamically the complexity of each part with respect to the others in order to both conform to a strict time budget and to present the most telling details. Integration of all this with a GIS database is important because many applications require access to geographically located information (e.g. building names, contents, and even floor plans). The GIS data can also be handled by the VGIS data managers and threaded through the real-time renderer for visualization in the 3D display environment.

The visual simulation system described above implies very large, even huge amounts of data. Automatic paging and caching techniques handling heterogeneous data from the different parts of the system must be in place. If, for example, the system is to visualize urban scenes, it must manage hundreds to thousands of buildings, plus their textures, and also street layouts. For flexibility the terrain visualization sub-system should handle terrain from any part of the world and integrate these terrains into a common coordinate system without seams or gaps (e.g. between levels of detail or due to multiple coordinate systems). All this should be in a hierarchical organization structure so that the terrain detail can be continuously adapted based on user viewpoint and scene content. Yet the hierarchy must be flexible so that detail can be added or deleted as needed. Such flexibility is quite important due to database size as the global datasets used with VGIS often require ten or more gigabytes.

In this paper we describe a visual simulation system that provides a structure supporting all the parts described above. We also discuss in detail our implementations for some of these parts, concentrating especially on global terrain visualization. The structure is in a multithreaded form to facilitate balanced and separable management of the system parts. It is also quite portable, due to standard libraries such as *Pthreads* and *OpenGL*, and has been ported to multiple workstation environments, including SGI and Sun platforms. We are now working on a PC version using Windows NT. The system has wide applicability, having been used for battlefield visualizations, tactical planning, and complex urban visualizations.

2   RELATED WORK

VGIS takes advantage of advances in a number of areas to create an integrated global geographic information visualization system. A large body of previous work has addressed issues in modeling, representing, and manipulating spatial data for geographic information systems. Applying 3D visualization techniques to global and spatial data has recently enjoyed increasing attention [12, 22].

VGIS manages its huge, complex terrain and GIS datasets in an efficient manner by using hierarchical spatial data structures. A number of such data structures have been adopted in GIS systems and other spatial databases. Samet [25] describes the quadtree, a fundamental data structure in the VGIS system. Representing spatial datasets that span the entire globe requires special data structures which take into account the curvature of the Earth and allow for efficient searching and rendering operations on the large amounts of data. Fekete [8] describes sphere quadtrees, a spatial data structure applicable to global representations of the Earth. Other researchers have proposed similar hierarchical spatial data structures which have been demonstrated to be useful for global geographic information systems [13].

VGIS also relies on multiresolution techniques to allow truly interactive visualization with its geometrically complex terrains. A large number of researchers have developed multiresolution representations and rendering techniques for large, complex terrains and height fields using polygonal meshes, as VGIS does. These algorithms attempt to represent surfaces within a given number of vertices, or within a given geometric error metric, or in a manner that preserves application specific critical features of the surface. Uniform grid methods or irregular triangulations are employed to represent the surfaces, and techniques including hierarchical subdivisions and decimations of the mesh are used for simplification and creation of multiresolution representations.

Much of the previous work on polygonalization of terrain-like surfaces has concentrated on triangulated irregular networks

(TINs). A number of different approaches have been developed to create TINs from height fields using Delaunay and other triangulations [10, 11, 27], and hierarchical triangulation representations have been proposed that lend themselves to usage in multiresolution level of detail algorithms [4, 5, 26]. Regular grid surface polygonalizations have also been implemented as terrain and general surface approximations [3]. Such a gridded terrain surface representation is used in VGIS and is described in [20]. Other surface approximation representations include techniques such as wavelet transforms [14] and methods that meet application specific criteria, such as preserving important terrain features [6, 10, 28].

VGIS uses an approach which treats the terrain as a single connected surface for rendering, using "continuous" level of detail representations for the terrain geometry. Similar methods for such "continuous" or "view-dependent" level of detail rendering for terrains and other surfaces are described in [3, 9, 20, 31, 32].

A number of visualization systems have been implemented which integrate 3D visualization techniques with large spatial geographic information and terrain data. Some systems stress accurate rendering of global images, or accurate modeling of environmental processes, often sacrificing interactivity of the system [21, 24]. Other systems emphasize tight integration of the 3D visualization with the powerful spatial analysis capabilities of GIS [7].

Systems such as VGIS place a high priority on real-time, highly interactive 3D visualizations of the spatial data. Maintaining truly real-time update rates in the face of large, complex datasets requires special techniques and time-critical visualization system designs. Bryson and Johan [1] discuss some issues particular to such time-critical computations in visualization environments. Software toolkits such as IRIS Performer [23] provide an architectural framework similar to that of VGIS which provides support for efficient rendering and simulation operations on high-end computer graphics workstations.

Interactive 3D visualization systems for visual simulation and training address many of the same technical problems as VGIS, including real-time rendering of large terrain databases. Flight simulator systems first implemented sophisticated image generation techniques to allow efficient rendering and visualization of complex spatial databases [33]. Recent developments have made interactive 3D graphics rendering of such databases possible using off-the-shelf graphics workstations. NASA's Virtual Planetary Exploration project [16] supported virtual exploration of planetary terrains such as Mars at interactive frame rates. SRI's TerraVision system also allows interactive viewing of 3D landscapes, using similar terrain level of detail and paging techniques as VGIS to allow very large, high-resolution geo-specific terrain datasets to be visualized. The T_Vision research project [15] provides a distributed virtual globe as a multimedia interface for visualizing geographic data. Other 3D terrain visualization systems use parallel architectures to render complex datasets while maintaining interactive performance for visual simulation and planetary visualization applications [2, 19].

3 SYSTEM DESIGN AND REQUIREMENTS

The rationale behind many of the design decisions made in the development of VGIS was driven by both hardware constraints and user requirements. Typical hardware constraints include available main memory, available texture memory, available precision in geometric calculations, rendering speed, disk transfer speed, etc., while the user requirements include interactivity, off-line processing time, display accuracy, data registration, as well as flexibility, extensibility, and portability from both the end user's and developer's perspectives.

Due to the large datasets that VGIS typically works with, most of the data must be put in secondary storage, and a paging scheme is used to bring in and cache the appropriate data for display. Level of detail techniques are applied to both geometry and imagery to further limit the amount of data stored in main memory and texture memory. These techniques are based on image quality metrics whose parameters can be manipulated interactively by the user to obtain a desirable balance between rendering speed and scene fidelity.

The user is given the ability to visualize the scene through multiple *views* which map onto separate windows. Each such view may display the scene from a different viewpoint, e.g. as a 3D immersive view or a 2D overview map, or may display different aspects of the same scene, e.g. as phototextured terrain or as a contour map with surface features such as roads and rivers turned on. In order to conserve memory, view-independent data is shared among the views and is accessed from a single primary cache.

To further obtain a high degree of interactivity, the system is broken down into a number of asynchronous threads that are prioritized according to their relevance to the final display update rate, which is one of the most important constraints in the system. For example, a dedicated render thread is used whose single task it is to update one or more views at the highest possible rate; level of detail (LOD) management is distributed over several threads according to the data they operate on, which generally update the scenes at a rate lower than the rendering rate; while a number of server threads execute only when data requests are made. This fine-grained subdivision of tasks ensures a high degree of CPU utilization and eliminates the bottlenecks often associated with blocking system calls (e.g. disk I/O, input device polling) in the real-time components of the system.

To accommodate data paging, level of detail management, and view culling, a quadtree data structure [25] is used to spatially subdivide and organize the terrain raster data. The globe is subdivided into a small number of pre-determined areas, each corresponding to a separate quadtree. The tiles associated with the quadtree nodes, or *quadnodes*, are stored in a file format that closely matches the internal representation, which allows for good paging performance. Rather than using a single global coordinate system, a large number of local coordinate systems is used. This is necessary as the precision afforded by current graphics hardware is insufficient for representing detailed geometry distributed over a large volume in a single coordinate system. Different branches within the quadtrees are assigned to different local coordinate systems, which are centered such that precision is maximized, and oriented to locally preserve natural directions such as "up", which can be exploited by the terrain geometry level of detail algorithm.

VGIS has been designed to be portable across a spectrum of different platforms, but is primarily targeted towards high end graphics architectures such as Silicon Graphics workstations. Certain platform specific extensions are optionally incorporated at compile time to maximize the system performance. The system is layered on top of portable standardized libraries such as OpenGL and POSIX threads, and provides a level of indirection for interfacing with the windowing subsystem, e.g. X11/Motif for Unix platforms.

4 SYSTEM OVERVIEW

The VGIS system consists of a dataset pre-processing component, as well as the actual run-time visualization display system. Optionally, external processes may provide real-time or remote data input. Due to the vast volumes of data that make up the global terrain database, the data is prepared off-line and structured in a form that closely matches the internal representation so that it can be paged into main memory efficiently. This form of pre-processing includes data format conversion, reprojection, resampling, data analysis, and data synthesis. In order to maximize the run-time performance,

non-interactive processes such as simulation, real-time data acquisition, and remote queries can be separated from the run-time system and be executed as remote processes. A number of server threads run in the back end of the run-time system that communicate with the external processes and fetch data from local disk. In addition, these threads perform intermediate tasks such as data transformations (e.g. conversion of height field data to Cartesian coordinates), data synthesis and initialization (e.g. generation of LOD parameters, synthesis of different image types), and repackaging of the data into the internal representation whenever necessary (e.g. parsing of DIS packets and GIS queries). Such processing is sometimes necessary to limit the amount of data stored on disk, and also allows VGIS to interface with external modules such as ARC/INFO GIS servers and ModSAF DIS simulators.

The servers are broken down by data type (e.g. terrain, symbology, GIS) and are run as independent threads. Each server handles requests from a number of clients that manage the corresponding data type. As mentioned in the previous section, VGIS supports the concept of multiple, independent views, each corresponding to a window on the screen. For example, one view could display the user riding a moving vehicle in three dimensions, another may be a "God's eye view" looking down upon the vehicle, while a third view could be a 2D overview map of a larger area. For each view, there is a *view module* that contains data managers for each data type. For example, there is a terrain manager that makes data requests to the terrain server and handles level of detail management, surface queries, terrain rendering, etc. Similarly, there is an object manager that manages animation and display of 3D symbology and protrusive features. The actual on-screen rendering is done by a single dedicated render thread. This thread runs asynchronously from all other threads to provide maximum rendering performance and interactivity. Within each view module, the scenes that are to be rendered are prepared and buffered in data structures similar to OpenGL display lists. These display lists are then sent to the render thread for display. The display lists are typically updated less frequently than the scenes are rendered, thus allowing the renderer to reuse a display list over several frames.

To minimize the response time, the renderer fetches copies of the most recent view parameters (i.e. position, orientation, field of view, etc.) for each view at the beginning of each frame. The view parameters are updated by a single user interface (UI) thread, which also acts as an overall manager of the system. Whenever a UI event is generated, e.g. from an input device or from remote commands, the corresponding view parameters are updated for that view.

An overview of the VGIS architecture is given by Figure 1. The following sections describe each of the system components in more detail.

4.1 Pre-Processing

Data processing in the VGIS system comes in three forms: off-line pre-processing, which is done once per dataset; intermediate, on-line processing, which is performed during the data paging stage; and real-time, on-line processing, which is typically done once per frame. The purpose of the pre-processor, or *dataset builder*, is to gather the different types of data and transform them into a format that is readable by and quickly accessible to the run-time system. These tasks are inherently compute intensive and cannot be performed by the run-time system at high enough rates. Additionally, the pre-processing only has to be done once for a given dataset, so it makes sense to pay the penalty of assembling a dataset up front.

While the pre-processor may handle a large variety of different data types, we will limit this discussion to the processing of static terrain data. Some of the basic data types within this domain include terrain geometry (e.g. "raw" elevation, elevation corrected for non-protrusive surface features such as roads and rivers, as well
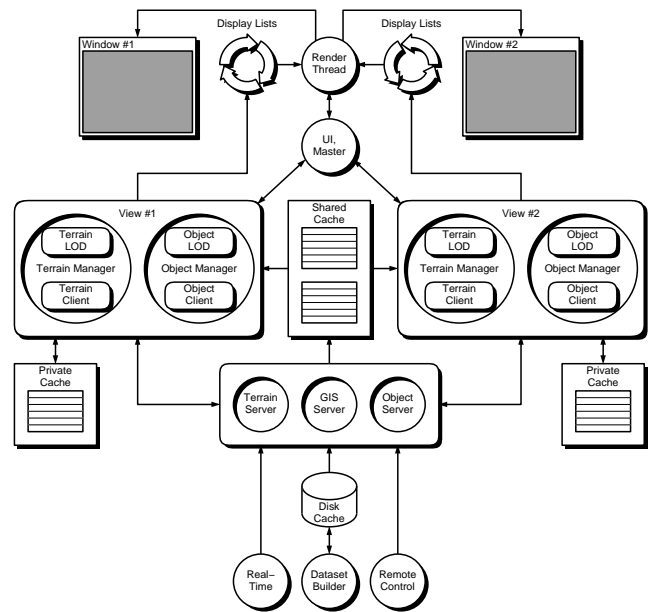

Figure 1: Overview of the VGIS system architecture. This figure illustrates two independent views but can be generalized to an arbitrary number. Modules are represented by rounded boxes, processes and threads by circles, data structures by rectangular boxes, and data flow and communication by arrows.

as the surface features themselves), imagery (e.g. phototexture, Arc Digitized Raster Graphics—or ADRG, pre-shaded relief maps, contour maps, etc.), and surface properties (e.g. surface roughness parameters, bounding volumes). The source data may come in a variety of file formats, which the pre-processor must translate to a single common format. In addition, for each data type, the source data may consist of multiple, variable resolution, possibly overlapping/nested datasets. The pre-processor has the ability to layer these and composite them into a single dataset if so desired.

As mentioned above, VGIS utilizes a quadtree structure for organizing multi-resolution terrain data in a hierarchical manner. Each node in a quadtree identifies a fixed, square area (in geodetic lat/lon coordinates) at a given discrete resolution. Additional constraints force the dimensions of the raster tiles associated with the nodes to be powers of two,[1] and the post spacing is successively doubled on each consecutive level up the tree. To facilitate the use of this data structure, each terrain dataset must be resampled to one of the discrete pre-determined resolutions. To accommodate fast paging, the data is also resampled and generated for all the internal nodes of each quadtree, resulting in what is commonly referred to as a *pyramid structure*.

The pre-processing of the data is done entirely in geodetic coordinates, assuming a pre-determined geodetic datum.[2] This ensures good registration between different source datasets and avoids the discontinuity problems often associated with flat-projected data. Both the height field and imagery are represented as regular grids. Hence, we do not require additional compute time to triangulate the height field—the different discrete levels of detail of the geometry are rather represented by the pyramidal regular grid structure, and further refinement is employed on-the-fly by the run-time system. See [20] for a detailed discussion of the height field level of detail algorithm. To further improve dataset registration, *alpha chan-*

---

[1] The elevation height field nodes have raster dimensions $2^n + 1$ as their boundary rows and columns overlap.

[2] In the current implementation, the WGS-84 datum is used.

nels can be added to the source data (both imagery and elevation) to allow blending/smoothing across dataset boundaries as well as masking of the data, e.g. for irregularly (non-rectangular) shaped datasets or for regions where data is missing.

The dataset builder has been designed to be very flexible and extensible in terms of incorporating many different data types, and is structured in a modularized manner to allow for fine-grained parallel execution. The different data types are declared by the user using a class hierarchy language. For each class, a number of methods/modules are defined that are used for reading, writing, and processing the data, with bindings to implementations of the methods. Class inheritance can be exploited to avoid redundant methods and class members. The modules are then connected in a data flow network similar to the structure provided by visualization systems such as *AVS* and *Data Explorer*, that is the output of one module is piped to the input of another set of modules. A set of operators are provided for expressing which modules are allowed to run in parallel, and which parts must execute in series. The output data must be classified into a few number of meta-types, such as geometry and imagery, that the run-time system understands how to manage and display. A repository of generic and specialized modules is maintained which can be supplemented to handle near arbitrary data types, both by the pre-processor and the run-time system. The dataset building process is entirely automated, making it very easy to insert and remove data. This is important for facilitating real-time acquisition and integration of data such as up-to-date satellite imagery.

4.2   Run-Time System

The run-time part of VGIS has been designed to support highly interactive frame rates. As such, it relies heavily on a multithreaded, fine-grained task distribution. Since the display update rate is often of higher importance than the scene update rate, including level of detail selection and animation, more resources are allocated towards satisfying a minimum frame rate. In [20], we propose a terrain geometry level of detail algorithm that generates "continuous" levels of detail on-the-fly. While being highly efficient in selecting the vertices and triangles that make up the terrain surface for a given view, the recursive traversal of the terrain data structures for triangle stripping is a bottleneck that limits the rendering rate to approximately 10 frames per second for pixel-accurate full-screen views. As a lot of frame-to-frame coherence is evident in the resulting triangle strips, it is often satisfactory to perform the LOD management less frequently, say 1–5 times per second, and trade the scene update rate for higher display rates. We accomplish this by decoupling the two tasks and separate them into two different, asynchronous threads. This scheme allows the scene to be redisplayed (with potential changes to the view between frames) until new parts of the scene have been generated and submitted by a scene manager. In addition, we further segregate different data types, recognizing that different data products may require different display update rates. For example, animated vehicles may be updated at a higher rate than the terrain geometry.

By and large, our approach has been to identify the major bottlenecks, such as blocking system calls, and isolate them from the time-critical components. This means offloading the I/O intensive, high latency data paging from the scene managers, and feeding the render thread with "ready to render" display lists. The resulting architecture forms a hierarchy of successively lower priority tasks. Each of these tasks is discussed in the following sections.

4.2.1   Paging and Intermediate Processing

The server layer in VGIS provides an interface to accessing external data and transforming it into the internal data structures employed by the display subsystem. Since the latencies involved in disk transfers and inter-thread communication are significant compared to display update times, the servers are decoupled and run asynchronously from the more time-critical pieces of VGIS. However, since the requirements on bandwidth between servers and clients are quite high, we chose to integrate the servers with the back end of VGIS and have them communicate with the clients via shared memory. In fact, the servers have direct access to the internal caches of VGIS, limiting the communication to small request messages and acknowledgements (see Figure 1). For less time-critical services (e.g. GIS queries), another (external) layer of servers can be built on top of VGIS. An example of such an external server is the ModSAF battlefield simulator which sends DIS packets over a Unix socket (possibly from another machine) to the internal object server, which in turn interprets the packets and reformats them for internal storage. Thus, the VGIS servers have two tasks: "data paging" and "intermediate processing," or preparation and initialization of the paged data. In this section, we will focus on the most bandwidth-consuming service; terrain paging. Other servers operate in a similar manner but under different conditions, and are typically not limited by bandwidth.

For each view, there is a terrain manager thread, part of which is a client module. The client module is responsible for making data requests to the terrain server whenever data of some type and resolution is needed for a particular area, and taking the appropriate actions upon notification by the terrain server that the request has been serviced. When data is needed for a node in a quadtree, the client allocates space for the data within a shared cache and sends a message via a shared memory priority queue to the server. Message priorities in this queue are changed dynamically according to the importance of the associated request as determined by the level of detail manager. Thus, requests that gradually become less important, or even obsolete, sift towards the end of the queue and get serviced only when no higher priority requests remain in the queue. This is important as the paging rate, during short bursts of requests, is typically much lower than the request rate. The server dequeues the highest priority request and either reads the data from disk if it exists, or synthesizes the data from other sources (or possibly a combination of both). After transferring the data from disk, the server may have to do additional processing. In the case of elevation data, the server reads a height field raster from disk, and then proceeds to transform the height field lat/lon/height coordinates into an array of Cartesian vertices. LOD state and other parameters are also generated and initialized before the request is completed, and the terrain client making the request is notified by returning an acknowledgement message.

Modules can be attached to the servers to handle paging of "arbitrary" data types. A "read module" for each data subtype must be registered with its corresponding server, and the core of the server is merely a dispatcher of jobs to be executed by these modules. These independent modules can be made to run in parallel to further increase the throughput. In this fashion, a number of data types can be synthesized from existing data and be easily integrated with VGIS without having to restructure the server. For example, a module can be added that synthesizes contour map images on-the-fly from existing elevation data. In this particular case, the elevation data may not even have to be read in if it already resides in memory since the servers have access to the shared memory cache.

4.2.2   Internal Representation and Caching

In the previous sections, we briefly mentioned the quadtree and shared cache data structures. These two structures constitute the basic components of the internal terrain representation. Rather than having a single quadtree represent the globe, we chose to subdivide

the globe into a small number of quadrants.[3] Each such quadrant is further subdivided and organized by a hierarchical quadtree structure. A node in a quadtree corresponds to a raster tile of fixed dimensions and lat/lon resolution according to the level on which it appears in the quadtree. Quadnodes are identified by "quadcodes," which are constructed in a manner similar to the indices of array representations of binary trees, that is the children of a node with quadcode q are identified by $4q + 1$ through $4q + 4$. In addition, the quadcode contains a quadtree identifier which allows each quadcode to uniquely identify an area on the globe.

In order to conserve memory, the static, view-independent data associated with a node is stored in a shared cache. This allows multiple terrain managers to access the same data without having to replicate it. The shared cache itself is implemented as a set of hash tables, one for each data type (e.g. elevation, phototexture, ADRG), which have enough slots to hold all the quadnodes that exist in the dataset. These slots are initially empty, and are filled with terrain data whenever a request is processed by the terrain server. If a node is no longer needed by any of the terrain managers, the space for it is deallocated. The quadcodes are used as hash keys for accessing nodes in the hash table. Since the hash table slots are initialized at startup, the terrain managers know what nodes exist externally such that no invalid data requests are made to the server. To avoid race conditions on the hash tables, each table is supplemented with a lock or a semaphore. Note that the hash tables need to be locked only on transitions such as when space for a node needs to be allocated or freed (ensuring that multiple terrain managers do not allocate or free a node's data simultaneously), or whenever a manager begins or ceases to reference a node. Reference counts are maintained so that the managers know when to allocate and free nodes.

The actual quadtree structure and any view-dependent parameters—such as vertex LOD state information—are stored in a view-private cache (Figure 1). The quadtree data structure is mostly a skeleton that indicates the presence of nodes, while the shared cache holds most of the actual terrain data. This scheme can be extended to data other than terrain, e.g. for moving vehicles and stationary objects, where pointers are used to reference the appropriate level of detail model within the shared cache.

In addition to spatially organizing the data, the quadtrees also define the boundaries of the aforementioned local coordinate systems. If a single, geocentric coordinate system were used, and assuming 32-bit single precision floating point is used to describe geometrical objects,[4] the highest attainable accuracy on the surface of the Earth is half a meter. Clearly, this is not sufficient to distinguish features with details as small as a few centimeters, e.g. the treads on a tank. As a matter of fact, many of the terrain datasets that have been used in VGIS have 10 centimeter resolution. This lack in precision results in "wobbling" as the vertices of the geometry are snapped to discrete positions. To overcome this problem, we define a number of local coordinate systems over the globe, which have their origins displaced to the (oblate) spheroid surface that defines the Earth sea-level. The origins of the top-level coordinate systems are placed at the geographic centers (i.e. the mean of the boundary longitudes and latitudes) of the quadtree roots. While the centroid of the terrain surface within a given quadrant would result in a better choice of origin in terms of average precision, we decided for simplicity to opt for the geographic center, noting that the two are very close in most cases. The z axis of each coordinate system is defined as the outward normal of the surface at the origin, while the y axis is parallel to the intersection of the tangent plane at the origin and the plane described by the North and South poles and the origin, that is the y axis is orthogonal to the z axis and points due

North. The x axis is simply the cross product of the y and z axes, i.e. $\hat{x} = \hat{y} \times \hat{z}$, and the three axes form an orthonormal basis. This choice of orientation is very natural as it allows us to approximate the "up" vector by the local z axis, which further lets us treat the height field as a flat-projected surface with little error. Hence, the height field LOD algorithm, which is based on vertical error in the triangulation, does not have to be modified significantly to take the curvature of the Earth into account. The delta values (see [20]) are however computed in Cartesian rather than geodetic coordinates to avoid over-simplification of "flat" areas such as oceans. Figure 2 illustrates the local coordinate systems for a few quadrants.
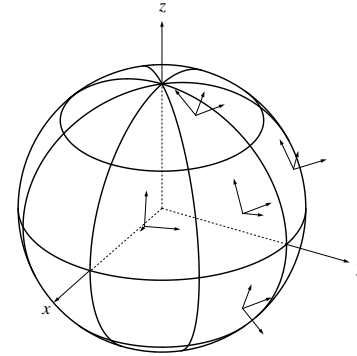

Figure 2: Local coordinate systems for the quadtree roots. The labeled axes correspond to the conventional Earth Centered, Earth Fixed Cartesian XYZ global coordinate system.

Using the above scheme, the resulting worst case precision for a $45° \times 45°$ quadrant is 25 cm—not significantly better than the geocentric case. We could optionally use a finer subdivision with a larger number of quadrants to obtain the required precision. However, this would result in a larger number of quadtrees, which is undesirable since the lowest resolution data that can be displayed is defined by the areal extent of the quadtree roots. Hence, too much data would be needed to display the lowest resolution version of the globe. Instead, we define additional coordinate systems within each quadtree. In the current implementation, we have added 256×256 coordinate systems within each quadtree, resulting in a 1 mm worst case precision.

4.2.3 Terrain Level of Detail

VGIS applies level of detail techniques to simplify both geometry and texture detail. This is necessary to maintain interactive frame rates as the global views typically contain millions, or even billions, of surface polygons, with gigabytes worth of imagery. Both of these techniques guarantee to meet an upper bound on the screen space simplification error, and the error bounds can be manipulated interactively by the user until sufficiently high frame rates are obtained.

The terrain geometry LOD algorithm is presented in [20], but we will give a brief overview of it here. The algorithm proceeds in two stages; a coarse-grained simplification in which quadnodes are selected at the appropriate resolution, followed by a fine-grained simplification in which individual vertices within each node are decimated. The decision as to whether a vertex can be removed is based on the screen space distance a vertex travels from its original position to the resulting surface if it were to be removed. The corresponding world space distance is referred to as the vertex's "delta value". If this distance is smaller than a screen space threshold, the vertex is decimated and further simplification of nearby vertices is considered recursively. The first stage—coarse-grained

---

[3]In the current implementation, thirty-two $45° \times 45°$ quadrants are used.

[4]This is typically the highest precision available in current graphics hardware.

simplification—entails traversing the quadtrees, evaluating the resolution required for each area, and making requests for height field data whenever necessary. The delta values are computed by the terrain server on-the-fly, which is an example of an intermediate processing task (see Section 4.2.1). After the final set of vertices have been selected, the LOD manager produces a single triangle strip for each quadnode, which is simply a list of alternating vertices and texture coordinates enclosed by begin/end triangle strip commands. These display lists emulate the OpenGL counterpart, but are not the same. The scene managers cannot utilize OpenGL display lists directly as only one thread at a time is allowed to issue graphics commands within an OpenGL context. Thus, synchronization is required, potentially forcing other threads that are ready to render to stall. Color Plate 6 shows a terrain surface tessellation resulting from application of the continuous level of detail algorithm. The pixel threshold in this view is one pixel. Note that the shading has been applied to the geometry before simplification. See Section 4.2.5 for further discussion on rendering.

The texture level of detail management bears some similarities to the coarse-grained geometry simplification. Rather than projecting the largest delta value to the screen, the largest (in screen space) *texel* is projected and compared to a threshold. The largest texel is found in a way similar to the determination of the largest delta value projection, i.e. it involves minimizing the viewpoint-to-texel distance while simultaneously minimizing the angle the texel normal makes with the viewpoint-to-texel vector. We here assume that all texels lie in a horizontal plane, but may be distributed in elevation within the bounding box of a node. As a result, low detail is used when the surface is viewed from the side, while relatively higher detail is required for top-down views of the terrain.

4.2.4 Object Management

The object manager handles the display of static and moving protrusive objects. For the latter it also handles animation. Since hundreds to thousands of objects may appear in certain scenarios, the object manager must accept and handle multiple levels of detail. Individual objects will have levels of detail, and there will also be grouping strategies. We have reported previously on hierarchical grouping for sets of moving units [30]. In general moving and static objects should be managed differently since they will have different grouping strategies and different ways to handle object detail (e.g. building textures). This is our eventual goal for the object manager. For the rest of this section, we will discuss how we handle moving objects.

Moving units can be simulated with ModSAF and transmitted to VGIS using a a DIS protocol. This is handled by a thread that keeps up with the events generated by ModSAF and performs dead reckoning when intermediate states need to be interpolated. VGIS has also accepted real-time position information for vehicles via the DIS interface. Color Plate 5 illustrates a number of tanks that have been instantiated in VGIS as a result of DIS packets received from ModSAF.

Although ModSAF can provide hierarchical information, it is often unavailable in other situations (e.g. in a battlefield scenario involving enemy units). For these situations we have developed dynamic clustering methods. Even though the clustering analysis can be handled asynchronously with the render thread in the VGIS framework, it still must be fast. To achieve this we use an algorithm that clusters units that are located within a threshold distance from an existing cluster. Units that satisfy this criterion are added to the cluster. The ones that do not meet this requirement are used to generate new clusters. At the end clusters that are too close to each other are combined. In most cases this analysis can be completed at interactive rates.

The cluster information is used to extract hierarchical information from the moving units by treating the resulting clusters as entities and clustering them at a higher level. This information is used by the LOD manager to replace groups of units with corresponding symbols based on viewing distance and total number of moving objects.

Since there will usually be semantic information such as object type and behavior associated with the clustered objects, the clusters can be evaluated for importance. Stytz et al. [29] have shown this is possible using fuzzy logic methods. We are applying and extending these methods to the dynamic clusters within VGIS. The Stytz approach sets static watch areas and evaluates only units that enter these areas whereas our approach handles dynamic clusters wherever they occur. The fuzzy logic techniques can handle large amounts of semantic information and transform it into relative importance weightings. These weightings can be used for situation awareness (highlighting those clusters that are especially important) and for detail management (more important groupings would show finer detail).

4.2.5 Rendering

Monoscopic and stereo rendering in VGIS is handled by a single thread, which renders into one or more windows. The scene managers communicate with the render thread via buffering of graphics commands that are encapsulated in display lists. For each connection with the render thread, there is a buffer of three dynamically growing/shrinking display lists, called a "triple buffer". One of the three display lists corresponds to what the renderer is currently drawing, a second display list is used by the scene manager to buffer graphics commands, while the third display list contains data that is ready to be displayed. This scheme allows both the renderer and the scene managers to run simultaneously without having to be synchronized. Consider, for example, if only two display lists per connection were used. In such a case, the scene manager, upon submitting a display list, would have to wait for the renderer to finish rendering a frame before the lists could be swapped. When triple buffers are used, the scene managers have to stall only when they produce scenes faster than the renderer can process them.

At the beginning of each frame, the renderer fetches the most current view parameters from the user interface thread. Typically, the UI thread runs at least as fast as the renderer. For each window, the renderer then fetches the most recently updated associated display list, and begins parsing it. For stereo views, the same display list is used twice, with different view parameters for each eye. Nearly all of the commands and parameters in the display list map directly to OpenGL function calls. However, there are a few commands that have special meanings, which will be described below.

Because the scene managers and the renderer work asynchronously, there is no single consistent set of view parameters among them.[5] They all fetch the most current parameters from the UI thread at the beginning of each frame/pass. Hence, view culling cannot be performed by the scene managers since the field of view may change after a scene has been submitted, at which point parts of the scene may be absent. Since view culling can drastically improve rendering rates, VGIS assigns this task to the renderer. The scene managers can insert `cull-begin/end` commands accompanied by bounding boxes that completely enclose the object in question. Whenever the renderer encounters a bounding box, it tests the box for intersection with its copy of the view volume, and skips the commands up till the corresponding `cull-end` command if the two do not intersect. Cull commands can be nested to quickly eliminate large parts of the scene.

The terrain surface is represented as a number of quadnodes in possibly different coordinate systems (see Section 4.2.2). Whenever a transition from one coordinate system to another occurs, the

---

[5]The view parameters between threads exhibit a lot of coherence, however, since they are updated at nearly the same rate.

scene manager must instruct the renderer to switch to the new co-ordinate system. This implies changing the current modelview matrix, which is the product of the local to global coordinate transformation and the world to eye coordinate transformation. Thus, there is a command for loading the current modelview matrix, as well as the regular OpenGL matrix multiply command. All matrix transformations are done in double precision to reduce roundoff errors. Many implementations of OpenGL do not perform matrix multiplications in double precision, so these have to be carried out manually, and the result is then loaded onto the OpenGL matrix stack.

Texture data is not passed to the renderer explicitly, but is rather referenced via a pointer to a shared memory structure. Whenever a `use-texture` command is encountered, the renderer determines whether the texture has already been defined. If not, a texture identifier is allocated and the texture is defined using `glTexImage`. If a texture is no longer needed, the last scene manager to use it sends a `delete-texture` command, which allows the renderer to deallocate the texture ID for later reuse. A number of policies are enforced such that texture allocation and deallocation are synchronized among threads, ensuring that the thread deallocating the image data is the only remaining thread with a reference to the texture.

Many terrain rendering systems approach the problem of rendering non-protrusive features such as roads and rivers by constraining the surface triangulation to follow the terrain features. This is not a viable alternative in VGIS as the surface triangulation is regular. Instead, we exploit the automatic texture coordinate generation and clipping facilities supported by OpenGL. As a pre-processing step, 2D vector data that describes the features is converted to a set of adjoining quadrilaterals (assuming some feature width). For features that use texturing, two adjacent edges of each quadrilateral in conjunction with the vertical define two reference planes from which texture coordinates are determined using the `glTexGen` function. A single, high resolution image such as a road texture can be repeated across the quadrilateral, whose boundary is defined using OpenGL's `glClipPlane` feature. For each quadrilateral, the terrain surface triangles that intersect it in two dimensions are identified using simple line intersection tests in the *x-y* plane. This culling task is aided by the hierarchical structures associated with the quadtree and the recursively defined triangle mesh. The resulting triangles are then rendered on top of the already drawn terrain surface.

4.2.6   User Interface and Navigation

The user interface thread handles input events from steering devices and interface widgets such as menus and sliders, performs navigation, and also acts as the overall manager of the run-time system. The decoupling of the user interface from other threads ensures high responsiveness to user input, while letting the most time-critical threads bypass expensive systems calls such as device polling. The user interface is otherwise mostly callback driven and is based on the X11/Motif mechanisms for event handling in the Unix version. Menus, sliders, and forms are provided to the user for interacting with and navigating the environment, with real-time responses to user input. Several input devices, such as mouse, space-ball, and 3D position and orientation tracking are supported. Two modes of navigation are currently in use: orbital mode, in which the user manipulates the globe via rotations and zooming, and free flight mode, which provides a six degree of freedom navigation interface. For each view, the UI thread maintains a master copy of the view parameters, which are fetched frequently by the renderer and scene managers. The UI thread additionally oversees system-wide tasks such as system initialization, resource allocation (e.g. texture memory, polygon budgets, thread and message queue creation,

CPU time, etc.), and coordinates both internal threads and external connections.

Due to the large range in geographic scale that VGIS operates over, several navigation and view parameters must be set according to the scale of the environment that is visualized. For example, the flight speed should be relatively large when the entire globe is visible to the user, while it must be reduced as the user approaches the surface. Similarly, the angular velocity in orbital viewing mode should preferably be mapped such that the arc length of a rotation in screen space is proportional to the mouse cursor's movement on the screen, i.e. the angular velocity should increase with altitude, providing the user with a sense of grabbing and panning the terrain. For stereo views, parameters such as focal length and eye-to-eye distance must be adjusted appropriately to enhance the stereo effect as perceived by the user. From a rendering standpoint, the two most important view-dependent parameters are the placements of the near and far clipping planes. In order to maximize z buffer resolution, the far plane distance is made a small as possible without clipping away any visible parts of the terrain. By using the curvature of the Earth and an upper bound on elevation, the appropriate far plane distance is recomputed whenever the user moves. The near plane distance is then expressed as a fixed fraction of the far plane distance, and most other view-dependent parameters are similarly kept proportional to the latter.

5   RESULTS

VGIS has been extensively used both by itself and as the central component of the Battlefield Planning and Visualization (BPV) system produced by the Army's CECOM organization. It has been fielded for military exercises at Ft. Bragg, Ft. Hood, and the National Training Center at Ft. Irwin among other places. Recently it was used in the Force 21 Army Warfighter Experiment at NTC, the biggest experiment of its kind in 20 years. In all these cases VGIS was operated by non-expert military personnel (usually participants in the exercises). This has generated a wealth of feedback about the navigation interface, the tools provided, the quality of the visualizations, the rendering speed, and the ability to add new data to the database. Many of these evaluations were used to improve the VGIS design and capabilities. VGIS has also been used for strategic and tactical planning for military hotspots and potential hotspots. Often very large amounts of new terrain data must be quickly prepared and inserted for these activities. Since the data is often classified, this work must be done without direct or indirect help of the VGIS team.

To evaluate the performance of VGIS, we have built a dataset consisting of elevation and imagery data at multiple resolutions. Color Plates 1–4 illustrate views along a continuous flight path from outer space into downtown Atlanta. A 3.5 arcminute (approximately 8 km) resolution global dataset constitutes the base layer, with additional higher resolution data inset in select areas: a 1 km resolution dataset for North America, a 100 meter resolution dataset for the state of Georgia, the metropolitan Atlanta area at 10 meter resolution, and finally higher resolution hotspots for the downtown area (1 m) as well as the Georgia Tech campus (0.4 m). Both elevation and image data exist for these areas, except for the last two for which only imagery was obtained. In addition, 471 textured 3D building models have been inserted into their geographically correct locations in the downtown area. The total size of this dataset exceeds 900 MB. Despite being provided by a number of different sources, we observed exceptionally good registration between the datasets in both the elevation and imagery data, requiring no manual labor to rectify or correlate the data.

The timings for the performance evaluation were obtained on an SGI Onyx InfiniteReality workstation with a single 194 MHz R10000 processor and 16 MB of texture memory. The pre-

processing stage took approximately 30 minutes from start to finish, and required no user interaction other than the initial specification of what datasets to include. The run-time performance was measured in terms of rendered frames per second, which ranged from 20–30 for a 1024×768 pixel window, with geometric and image accuracies of two pixels. The scene update rates (level of detail evaluations etc.) for the terrain varied between 1 and 20 Hz.

6 CONCLUSION AND FUTURE WORK

We have designed and implemented a real-time 3D visualization system, VGIS, which integrates visual simulation techniques for interactive rendering and management of huge graphical databases with query and manipulation capabilities for spatial geographic data. The system allows visualization of terrain and other data types over the entire surface of the Earth, and manages very high resolution datasets at real-time rates, by taking advantage of hierarchical, multiresolution spatial data structures and asynchronous multithreading. The system has met its original requirements for high interactivity and the ability to handle huge databases, and has proved useful for military planning and visualization tasks in Army exercises.

VGIS continues to be improved and extended to meet new application demands. One issue we plan to address is acquisition and integration of terrain data "on-the-fly," modifying the current terrain data pre-processing stage to allow fast changes to the underlying terrain and GIS source data. This will allow for dynamic deformations of the terrain surface, as well as dramatic decreases in the time required to visualize newly obtained terrain data. Other areas of current and future work include further development of a level of detail framework integrating the large variety of data types, and improving the user interface functionality and system support for users in making sense of the huge amounts of information VGIS is capable of displaying.

Acknowledgement

References

[1] BRYSON, S. T. and JOHAN, S. Time Management, Simultaneity and Time-Critical Computation in Interactive Unsteady Visualization Environments. In *Proceedings of Visualization '96*, October 1996, pp. 255–261.

[2] COHEN-OR, D., RICH, E., LERNER, U. and SHENKAR, V. A Real-Time Photo-Realistic Visual Flythrough. *IEEE Transactions on Visualization and Computer Graphics* 2(3), September 1996, pp. 255–265.

[3] COSMAN, M. A., MATHISEN, A. E., and ROBINSON, J. A. A New Visual System to Support Advanced Requirements. In *Proceedings, IMAGE V Conference*, June 1990, pp. 370–380.

[4] DE BERG, M. and DOBRINDT, K. T. G. On Levels of Detail in Terrains. In *11th ACM Symposium on Computational Geometry*, June 1995.

[5] DE FLORIANI, L. and PUPPO, E. Hierarchical Triangulation for Multiresolution Surface Description. *ACM Transactions on Graphics* 14(4), October 1995, pp. 363–411.

[6] DOUGLAS, D. H. Experiments to Locate Ridges and Channels to Create a New Type of Digital Elevation Model. *Cartographica* 23(4), 1986, pp. 29–61.

[7] ERVIN, S. M. Landscape Visualization with Emaps. *IEEE Computer Graphics and Applications* 13(2), 1993, pp. 28–33.

[8] FEKETE, G. Rendering and Managing Spherical Data with Sphere Quadtrees. In *Proceedings of Visualization '90*, 1990.

[9] FERGUSON, R. L., ECONOMY, R., KELLY, W. A., and RAMOS, P. P. Continuous Terrain Level of Detail for Visual Simulation. In *Proceedings, IMAGE V Conference*, June 1990, pp. 144–151.

[10] FOWLER, R. J. and LITTLE, J. J. Automatic Extraction of Irregular Network Digital Terrain Models. Proceedings of SIGGRAPH 79. In *Computer Graphics* 13(2) (August 1979), pp. 199–207.

[11] GARLAND, M. and HECKBERT, P. S. Fast Polygonal Approximation of Terrains and Height Fields. Technical Report CMU-CS-95-181, CS Dept., Carnegie Mellon U., 1995.

[12] GIERTSEN, C. and LUCAS, A. 3D Visualization for 2D GIS: an Analysis of the Users' Needs and a Review of Techniques. In *Proceedings of Eurographics'94*, 1994, pp. C-1–C-12.

[13] GOODCHILD, M. F. and SHIREN, Y. A Hierarchical Spatial Data Structure for Global Geographic Information Systems. *CVGIP: Graphical Models and Image Processing* 54(1), 1992, pp. 31–44.

[14] GROSS, M. H., GATTI, R., and STAADT, O. Fast Multiresolution Surface Meshing. In *Proceedings of Visualization '95*, October 1995, pp. 135–142.

[15] GRUENEIS, G., MAYER, P., SAUTER, J., and SCHMIDT, A. T_Vision. In *Visual Proceedings of SIGGRAPH 95*, August 1995, p. 134.

[16] HITCHNER, L. E. Virtual Planetary Exploration: A Very Large Virtual Environment. ACM SIGGRAPH 92 Tutorial on Implementing Immersive Virtual Environments, 1992.

[17] KOLLER, D., LINDSTROM, P., RIBARSKY, W., HODGES, L. F., FAUST, N., and TURNER, G. Virtual GIS: A Real-Time 3D Geographic Information System. In *Proceedings of Visualization'95*, October 1995, pp. 94–100.

[18] LECLERC, Y. G. and LAU, S. Q. TerraVision: A Terrain Visualization System. SRI International Technical Note No. 540, April 1994.

[19] LI, P. P., DUQUETTE, W. H., and CURKENDALL, D. W. RIVA: A Versatile Parallel Rendering System for Interactive Scientific Visualization. *IEEE Transactions on Visualization and Computer Graphics* 2(3), September 1996, pp. 186–201.

[20] LINDSTROM, P., KOLLER, D., RIBARSKY, W., HODGES, L. F., FAUST, N., and TURNER, G. A. Real-Time, Continuous Level of Detail Rendering of Height Fields. Proceedings of SIGGRAPH 96. In *Computer Graphics* Proceedings, Annual Conference Series, 1996, ACM SIGGRAPH, pp. 109–118.

[21] NISHITA, T., SIRAI, T., TADAMURA, K., and NAKAMAE, E. Display of the Earth Taking into Account Atmospheric Scattering. Proceedings of SIGGRAPH 93. In *Computer Graphics* Proceedings, Annual Conference Series, 1993, ACM SIGGRAPH, pp. 175–182.

[22] RHYNE, T. M., IVEY, W., KNAPP, L., KOCHEVAR, P., and MACE, T. Visualization and Geographic Information System Integration: What are the needs and requirements, if any? In *Proceedings of Visualization '94*, 1994, pp. 400–403.

[23] ROHLF, J. and HELMAN, J. IRIS Performer: A High Performance Multiprocessing Toolkit for Real-Time 3D Graphics. Proceedings of SIGGRAPH 94. In *Computer Graphics* Proceedings, Annual Conference Series, 1994, ACM SIGGRAPH, pp. 381–394.

[24] ROST, R., DOZIER, J., HIBBARD, B., KOCHEVAR, P., TREINISH, L., and VON SANT, T. ACM SIGGRAPH 93 Course Notes 71: Visualizing Planet Earth, 1993.

[25] SAMET, H. The Quadtree and Related Hierarchical Data Structures. *ACM Computing Surveys* 16(2), June 1984, pp. 187–260.

[26] SCARLATOS, L. L. A Refined Triangulation Hierarchy for Multiple Levels of Terrain Detail. In *Proceedings, IMAGE V Conference*, June 1990, pp. 114–122.

[27] SCHRODER, F. and ROSSBACH, P. Managing the Complexity of Digital Terrain Models. *Computers & Graphics* 18(6), 1994, pp. 775–783.

[28] SOUTHARD, D. A. Piecewise Planar Surface Models from Sampled Data. *Scientific Visualization of Physical Phenomena*, June 1991, pp. 667–680.

[29] STYTZ, M., BLOCK, E., and SOLTZ, B. Providing Situation Awareness Assistance to Users of Large-Scale, Dynamic, Complex Virtual Environments. *Presence*. Vol 2. No 4, Fall 1993, pp. 297–313.

[30] TURNER, G., HAUS, J., NEWTON, G., RIBARSKY, W., FAUST, N., and HODGES, L. F. 4D Symbology for Sensing and Simulation. In *Proceedings of the SPIE—The International Society for Optical Engineering* 2740, April 1996, pp. 31–41.

[31] WILLIS, L. R., JONES, M. T., and ZHAO, J. A Method for Continuous Adaptive Terrain. In *Proceedings of the 1996 IMAGE Conference*, June 1996.

[32] XIA, J. C. and VARSHNEY, A. Dynamic View-Dependent Simplification for Polygonal Models. In *Proceedings of Visualization '96*, 1996, pp. 327–334.

[33] YAN, J. K. Advances in Computer-Generated Imagery for Flight Simulation. *IEEE Computer Graphics and Applications* 5(8), 1985, pp. 37–51.
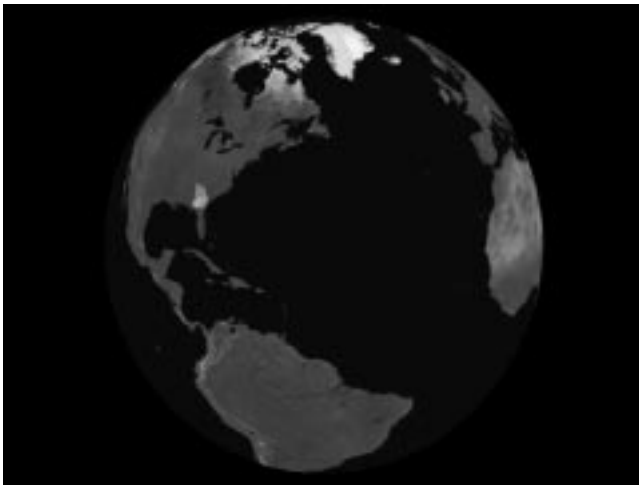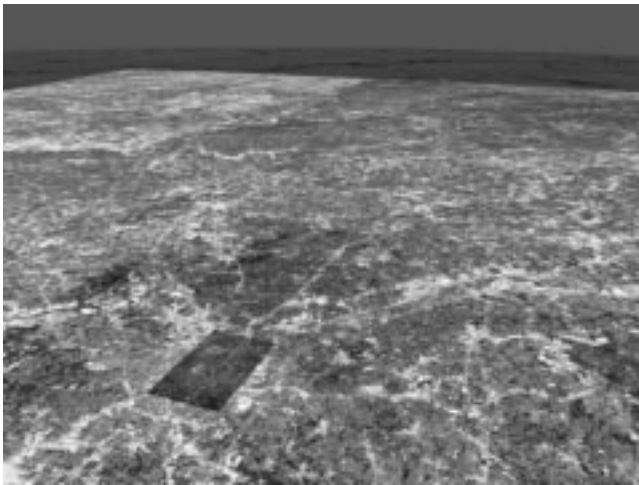
1. Visualization of the global dataset (8 km resolution) from an altitude of 12,000 km.
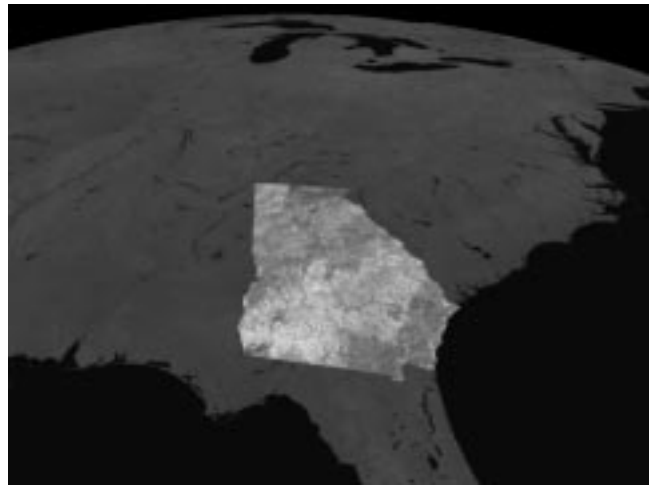
2. View approaching the state of Georgia inset data (100 m resolution).
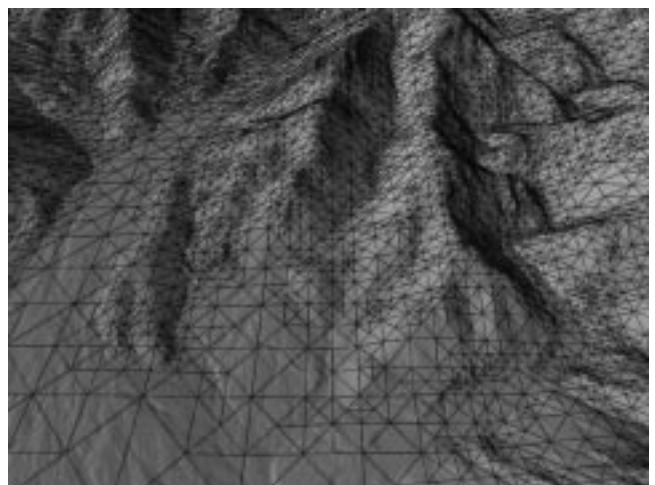

3. Four data insets of varying resolutions (1, 10, 100, and 1,000 m).

4. View of downtown Atlanta data with 3D building models.

5. DIS simulation of tanks engaged in battle at Ft. Irwin, CA.

6. Continuous level of detail terrain surface tessellation.