Multi-Resolution Deformation in Out-of-Core Terrain Rendering

William E. Brandstetter III, Joseph D. Mahsman, Cody J. White, Sergiu M. Dascalu, Frederick C. Harris, Jr.

Computer Science and Engr. Computer Science and Engr.
University of Nevada, Reno Desert Research Institute
Reno, NV 89557 Reno, NV 89512
{brandste,mahsman,cjwhite,dascalus,Fred.Harris}@cse.unr.edu

Abstract
Large scale terrain rendering in real-time is a well known problem across the computer graphics community which has garnered many solutions relying on dynamic level of detail changes to the terrain. These algorithms typically fit into two categories: in-core and out-of-core. Out-of-core algorithms usually require data to remain static, thus disallowing terrain modification whereas in-core algorithms allow for deformation, but usually require updating of modified data through a data hierarchy which can potentially be a slow process. We present a solution for out-of-core deformable terrain rendering in real-time.

Keywords: level-of-detail (LOD), deformable

1 Introduction
Terrain rendering is a highly researched area due to demand from the military, scientific visualization, and computer gaming communities. Even as advances in graphics hardware continue to be released, these applications will always push the current technology to the limit such that a brute force method will never be practical. Level-of-detail (LOD) rendering algorithms are one of the applications which continue to be developed to give the best visual representation of large-scale landscapes in real-time.

The size of datasets is one of the major problems in terrain rendering. First, brute force rendering is not an option when dealing with large datasets, so a LOD approach needs to be taken. Second, a large heightmap takes up quite a bit of memory and thus out-of-core (outside of system memory) rendering needs to be supported. The most common approach is to extract a good view-dependent approximation of the mesh in real-time. This is accomplished by storing data in a specific hierarchical structure, in which terrain can usually be categorized. Terrain can be represented in many different data structures such as a triangulated irregular mesh (TIN) [7], which gives the best approximation, a regular grid, which uses somewhat more triangles to represent a surface, quadtrees [13], binary triangle trees [5], or directed acyclic graphs [9].

Refinement may take place on a per-triangle basis, or tessellate aggregates of polygons. Some existing algorithms refine the terrain every frame, having a “split-only” approach. Others may merge and split from previous frames’ work. Refinement can be accomplished using a nested-error bound metric, or as in [10] solely the viewing position. Some terrain algorithms only support in-core (inside of system memory) [5], while others support out-of-core rendering [9] and dynamic addition of procedural detail [11].

Since terrain data can consume such a large memory footprint, out-of-core algorithms often limit their datasets to be static. Large amounts of terrain data are usually processed in a way that leaves the geometry optimal for video hardware and is not expected to change. Dealing with an out-of-core terrain system that handles dynamic updates of its height values is not trivial. For the most part, the areas of the mesh that need to be rendered stay in memory, while areas that aren’t visible can be discarded to the hard-drive until needed. With deformable terrain, updates to the mesh could be made outside the viewing frustum, in which case those areas would need to be loaded, updated, and cached back to disk. If a hierarchy of LOD mesh representations were preprocessed, then updated data may need to be propagated up through the tree or reprocessed altogether. The idea of dealing with large amounts of data in a dynamic terrain algorithm can quickly become unmanageable, thus when combined with the first problem, a second problem of terrain rendering is presented: dynamic terrain. Therefore, presented here is an out-of-core terrain algorithm which supports dynamic updates to the heightfield in real-time, allowing for deformable terrain.

2 Selected Previous Work
ROAM. ROAMing Terrain: Real-time Optimally Adapting Meshes [5] is a well known level of detail al-
formation could be supported, but geomipmaps would have to be recreated and geometrical errors recalculated, which could hinder real-time deformation. The downside is that the number of geomipmaps increases quadratically \((N^2)\) based on the size of the terrain; therefore, possibly resulting in slow computation and rendering.

**Chunked LOD.** At SIGGRAPH’02 Ulrich presented a hardware friendly algorithm based on the concept of a chunked quadtree, which is described in [15]. This algorithm, also referred to as Chunked LOD, is somewhat similar to GeoMipMapping; however, it scales much better due to the quadtree structure. There is often confusion of the differences between Chunked LOD and GeoMipmapping since the algorithms are similar. However, Chunked LOD exploits a quadtree data structure of mipmapped geometry. Therefore the number of rendered nodes does not quadratically increase due to the size of the terrain.

A requirement of this algorithm is to have a view-dependent LOD algorithm that refines aggregates of polygons, instead of individual polygons. As ROAM tessellates down to a single triangle, Chunked LOD refines chunks of geometry that have been preprocessed using a view-independent metric. Since chunks are stored in a quadtree, the root node is stored as a very low polygon representation of the entire terrain. Every node can be split recursively into four children, where each child represents a quadrant of the terrain at higher detail than its parent. Every node is referred to as a chunk, and can be rendered independent of any other node in the quadtree. Having such a feature allows for easy out-of-core support.

The Chunked LOD quadtree structure is one of the best known hardware friendly LOD algorithms since it can be utilized for very large out-of-core terrain. Adding detail resolution requires extending the chunked quadtree, which could be easily done in a preprocessing step. However, deformation isn’t trivial since the algorithm requires a static mesh; if any height samples were changed, it would require reprocessing the entire quadtree, which is unacceptable for real-time deformation.

### 3 Proposed Approach

The following sections present our out-of-core deformable terrain algorithm for preprocessing and rendering of large-scale terrain datasets. This algorithm is described in greater detail in [3]. We start with an overview of the hierarchical representation of the terrain data and then describe the runtime algorithm for mesh refinement, rendering, memory management, and deformation. Figure 1 illustrates the flow of data through the system from program initialization to rendering.
3.1 Hierarchical Representation

The hierarchical representation of the original mesh is built during a preprocessing step. For a \( n \times n \) input mesh, a quadtree is used to organize the data such that the root node defines a low-detail representation of the entire mesh. Each subsequent child node contains more detail at the scale of one quarter of its parent’s mesh, while the leaf nodes constitute the original mesh. Every node is of size \( m \times m \), and therefore each node uses the same amount of vertices. The dimensions \( n \) of the input mesh and \( m \) of the nodes must be one greater than a power of two to allow for optimization of the construction and refinement algorithms.

The quadtree is constructed using a simplification process similar to [8]. First, the input mesh is partitioned into leaf nodes of size \( m \times m \), where each node overlaps neighboring nodes by one row and one column. Nodes are combined into \( 2 \times 2 \) blocks and upsamped by removing every other row and column vertex. This is repeated recursively until \( 2 \times 2 \) blocks can no longer be made. Each node is given a bounding box that encapsulates the entire mesh, as shown in Figure 2.

The process of removing every other row and column vertex when creating parent nodes implies that the data for each node, except for the root, comprises its parent’s data (the excluded rows and columns) and its own data. During terrain deformation, this property obviates the need to propagate changes through the tree. In addition to a node’s individual data, it contains pointers to its parent’s data. To guarantee this property holds true, when a node is loaded into memory all of its ancestors must be in memory as well. The memory layout for any given node is shown in Figure 3.

For example, the bottom left vertex of an underlying heightfield belongs to the root node. Child nodes receive a pointer to this vertex in order to access it. This is similar to the wavelet compression scheme from [2]. However, we do not encode the child data within the parent’s node. Instead the individual data for each node is stored in a file that can be loaded on demand. This eliminates the need to decode node information at runtime and allows for deformation without encoding new vertices into the quadtree. In order to query a node’s data, it must simply dereference the vertices it points to.
3.2 Runtime Algorithm

3.2.1 Mesh Refinement

The goal of any terrain rendering algorithm is to quickly create the best approximation mesh for each frame. Our approach uses a split-only top-down refinement. Previous algorithms use properties of the underlying geometry (e.g. nested error bounds) during refinement as in [15] and [4]. Deformation of the terrain requires recalculation and propagation of these properties throughout the tree. We take an approach similar to [10] and use only the view position and frustum as refinement criteria. Although this approach looks awkward for high-frequency data (e.g. a steep mountain consisting of a few vertices), natural terrain datasets often feature a smooth gradient.

Refinement begins at the root node and proceeds recursively for each child node. A breadth-first traversal is required for linking neighboring nodes. For every node, if the node’s bounding box is inside the view frustum and the center of the bounding box is closer than a predefined threshold, the node is refined by traversing its four children, otherwise it is prepared for rendering. A threshold should be chosen such that a nested regular grid surrounds the viewer. Since no other metrics are taken into account during refinement, this will yield the best visual fidelity. Note that the LOD of neighboring nodes is never limited, as in a restricted quadtree where nodes are forced to split based on the level of its neighbors as in [12].

3.2.2 Neighboring Nodes

Smooth transitions between nodes of different LOD must be rendered correctly, otherwise seams will be visible due to gaps in the rendered mesh or inconsistent shading from incorrect normal calculations. Also, since each neighbor holds its own copy of edge vertices, care must be taken while deforming edges or edge boundaries. To handle these variations, a node must be aware of its neighbors. Since quadtree refinement isn’t restricted, the difference between two nodes may be one level or more. Since a node’s LOD may change from frame to frame, neighboring links are recreated during refinement.

When linking nodes together, a node is only allowed to point to a neighbor of equal level or higher. Enforcing this rule allows each node to store no more than four neighbor references. When a node is split during refinement, the node is responsible for updating its children with the correct neighborhood information. This cannot be accomplished with a depth-first traversal, commonly used in LOD algorithms. Instead, a breadth-first traversal is performed.

Neighbor links play an important role for correct normal calculation. Normals are needed to simulate a realistic lighting model, and can also be used for collision response. The biggest problem of normal calculation presents itself on the seams of terrain patches. Vertices on an edge need the height values of neighboring nodes.

The most common approach to calculate a normal is to compute a normal for each vertex in the heightfield by taking the average normal of all faces that contain the vertex [16]. This process consists of several costly mathematical operations, such as square roots. Several optimizations can be made by exploiting properties of the heightfield. The method we use is described in [14], which only requires the four neighboring heightsamples of a vertex. In order to create a smooth transition across a patch seam, neighboring vertices must be queried and the computed normal is then stored for each edge.

3.2.3 Detail Addition

To improve the appearance of the terrain without increasing the size of the data on disk, procedural detail is added at runtime for leaf nodes that meet the refinement criteria. The detail is added to the hierarchy in the form of new leaf nodes that extend the quadtree until a user-specified level is reached. When creating a new node, a reverse process of adding rows and columns is performed and the new node is linked to its parent, which was previously a leaf node. The new vertices are then assigned procedural data.

Linear interpolation is not sufficient for creating additional detail because the resulting data is uniform. Instead, fractals are used to give the data a non-uniform appearance. Each interpolated vertex is shifted a random amount such that it stays within the bounds of the surrounding vertices. Since detail addition is subtle, the process does not need to be deterministic, therefore detail can be randomized each time it is created.

3.2.4 Rendering

The result of refinement is a list of patches to be rendered. Before rendering, indices can be recalculated for nodes whose neighbor’s LOD have changed and normals can be recalculated if deformation had occurred. Each node must then dereference its pointer data to create a vertex list. Finally, each node can be transformed into world space and the data sent across the bus to be rendered. The rendering process is decoupled from the updating and disk I/O methods, allowing for smooth loads of data and no hiccups in the system regardless of how fast the viewer is moving around the terrain.

Stitching is accomplished by having the finer detail node omit vertices on its edge to match that of its coarser neighbor. This is done by rendering degenerate triangles. Geometrical skirts [15] were not chosen since
the size of the skirt may change after deformation. Recalculation of the skirt can become tedious and slow. Figure 4 illustrates the removal of T-junctions by utilizing degenerate triangles.

![Figure 4: T-Junctions appear at the neighboring nodes of different levels of detail.](image)

(a) T-Junction (b) Degenerate Triangles

3.3 Memory Management

Our algorithm supports large datasets stored out of core, i.e. data that resides outside of main memory [11]. A separate loading and caching thread is fed patches to load or write to disk. The patches to load are based on refinement, while the patches to write are based on a least recently used (LRU) algorithm.

During refinement, if a parent cannot be split because the data for its children is not in-core, a request for the data is made to the loading and caching thread. The system never stops to wait for data to load; until the data for the children is loaded, the parent’s data is rendered.

When the memory footprint exceeds a predefined threshold, LRU patches are discarded to disk until the used memory falls below the threshold. Each node is given a timestamp representing the last time the node was either rendered or deformed. A priority queue is used to efficiently determine which nodes should be discarded. Since every node relies on its parent for some of its data, discarding a parent to disk will invalidate memory references for its children, therefore only leaf nodes of the currently refined mesh are considered for caching.

Depending on the actions of the user (fast movement or several deformations) and the current memory footprint, nodes may require continuous allocation and deallocation. Instead of using operators such as new or delete which are notoriously slow for small and frequent allocations, a freelist is used as in [6].

3.4 Deformation

Real-time modifications are applied to the terrain by refining the currently active mesh based on a rectangular selection of the terrain, called a brush, in addition to the view-dependent refinement criteria described earlier. The vertex data for each refined node is modified to fit the brush specification.

A brush defines the rectangular extent (defined by position, width, and height) and the resolution of deformation (defined by a level in the hierarchy, which may not exist). In addition, a brush holds an array of pointers to vertices in the terrain, allowing deformations to cross node boundaries. Nodes that intersect the brush are selected during refinement and vertices from each node are given to the brush. Dereferencing the brush gives access to vertex data which can be overwritten with new data. Since vertices on edges are duplicated for each patch, care must be taken for deformations across boundaries by syncing adjacent vertices. This is done in a pre-rendering step that compares dirty flags of neighboring nodes in the quadtree.

Refinement is based on brush extent and resolution as well as view-dependent criteria. Therefore, a node may be refined even though it is not sufficiently close to the viewer or inside the view frustum. Depending on the resolution of the brush, data for nodes deep into the hierarchy may be requested for loading. Only when all of the data required by the brush’s resolution has been loaded can deformation be applied.

As described in Section 3.2.3, procedural detail is added for leaf nodes that meet the view-dependent refinement criteria. If a brush alters a node with procedural data, disk space is allocated for the node and it is allowed to be discarded to disk by the memory manager.

When a node is chosen for rendering, it is possible that an ancestor has previously been deformed. Time stamps are compared, and if a node’s last modification is older than its parent’s, its data is adapted to the parent mesh by creating procedural detail.

3.5 Texturing

Textures are processed similarly to the terrain data. A large texture can be cut into user-defined partitions and merged into $2 \times 2$ blocks before being mipmapped. This process continues until an entire quadtree is built over the original texture data.

Nodes in the terrain quadtree directly map to nodes in the texture quadtree. However, with modifications to the terrain quadtree (deformation and procedural detail), it becomes impractical to create a texture quadtree of the same depth. If a node in the terrain quadtree cannot be mapped directly to a node in the texture quadtree, the parent’s texture and texture coordinates are used. When a node is being loaded or deleted it can also load or delete its texture.

Just as the terrain quadtree presented issues at seams, so does the texture quadtree. This is due to the kind of texture filtering used to generate the texture quadtree. Although no seams are visible with nearest filtering, this type of filtering is not visually appealing. With linear filtering, seams appear at the texture edges because the edge texels are not being blended with the
correct neighbor texel. This is solved by overlapping adjacent textures during texture quadtree construction such that neighboring nodes have exact texels on shared edges. Clamping the texture edges during rendering causes these texels to blend, removing the seam. There is no perfect solution, and the amount of pixels to overlap can vary.

4 Results

![Screenshots from the visualization.](image)

The following tests were performed on a machine with an Intel Core2 Quad Q9450 processor with 8GB of DDR2 RAM and a NVIDIA GeForce GTX 275 under Windows XP Service Pack 2, which can only utilize 3.5GB of RAM.

The data used for these results was obtained from [1], which holds 10-meter elevation data of the big island of Hawaii along with a $4096 \times 4096$ texture. The terrain file has dimensions of $8193 \times 8193$ and was already in binary terrain (.bt) format. It was first converted into the internal .ter file and the texture image (.jpg) was converted into a .tex file. Since these files were so small, the preprocessing took less than five minutes.

With this application, the user is able to move around the scene via keyboard and mouse input. By clicking and dragging the mouse, the user can select a single axis-aligned brush, and change the resolution of that brush via keyboard input. Once a brush is selected with the desired resolution, the user may create a hill or crater by raising or lowering the terrain. Any changes to the terrain are automatically saved to the .ter file and will be loaded back in when the application restarts. Figure 5 shows screen shots of this application.

In order to determine how well this algorithm runs, we ran various operations on it as illustrated in Table 1 with a frame buffer size of $1024 \times 768$. The file tested was a ten meter resolution digital elevation map (DEM) of Hawaii of raster size $8193 \times 8193$ which can be freely downloaded over the Internet [1]. The first test was to simply move over the terrain with no deformation occurring. This tested the LOD refinement algorithm used to render the terrain in real-time. The next section of results in the table show the speeds of deformation of the terrain in terms of frames-per-second. Using different brush sizes, we deformed the terrain over the same part of the dataset. For all of the brush sizes used, the algorithm demonstrated interactive framerates. The largest brush size used exhibited a relatively low framerate due to the increased amount of refining of the mesh down to the deepest parts of the terrain hierarchy, which can be considered a worst-case scenario.

<table>
<thead>
<tr>
<th>Operation</th>
<th>FPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbitrarily moving over the dataset</td>
<td>48.41</td>
</tr>
<tr>
<td>Deformation with brush of size $32 \times 32$</td>
<td>38.37</td>
</tr>
<tr>
<td>Deformation with brush of size $64 \times 64$</td>
<td>23.54</td>
</tr>
<tr>
<td>Deformation with brush of size $128 \times 128$</td>
<td>10.96</td>
</tr>
</tbody>
</table>

Table 1: Average frames per second for four brushes.

Another application of the algorithm has been used for tire track deformation from a military vehicle navigating the terrain in a dataset from Yuma Proving Ground, an Army installation in Arizona.

5 Conclusion

We have presented a complete LOD terrain algorithm including the major features of deformation and out-of-core rendering. Refinement is not only based upon the viewing frustum, but also takes into account the selected deformation brushes. This allows data that is not being viewed to remain in memory and be subject to deformation. Previous methods that allow out-of-core rendering usually preprocess the geometry into a triangulated irregular mesh for optimal polygon throughput, and require that the terrain mesh remain static. Other in-core algorithms support changes to the underlying heightmap, but need to recalculate and propagate nested error-bounds through a hierarchical structure. Our approach eliminates the need for any geometry tessellation or propagation after a modification to the terrain heightmap. By exploiting the features of a regular grid, x and z coordinates will never change requiring only updates to the y coordinate (height offset). The quadtree structure exploits a child-parent relationship in which child nodes actually point to their parent’s data. In this way, when the data of children nodes are modified, the pointer actually dereferences some parent data completely eliminating any propagation back up through the quadtree. The need for nested error-bounds is also eliminated by depending solely on the view position for refinement.
Even though this results in a less accurate refinement, the tessellation is tolerable and the tradeoff of propagation removal is well worth it.

Deformation is allowed to be done at any resolution within the extended quadtree. The quadtree may be extended to a user specified resolution by scaling up the original terrain and adding procedural fractal detail to the leaf nodes. These extra nodes are created on the fly in real-time and only need to be saved to disk if deformed. Since detail addition is so subtle, the extra nodes do not need to be spatially deterministic and can be randomly created each time. By comparing the time stamp of a nodes parent, data may procedurally adapt to a low resolution modification using this same method to create detail.

The terrain is represented as a heightmap, precluding such features as caves and overhangs. The dimensions of the input heightmap are required to be $2^N + 1$ on each side to allow for optimizations. Additionally, the preprocessing step to build the terrain hierarchy is non-trivial for large datasets.

Along with our algorithm, we have presented support for large texture maps, fast normal calculation, and dealing with large world coordinate and depth buffer precision.

## 6 Future Work

For simplicity, not all optimizations were used when implementing this algorithm. It would be possible, with some effort, to port the entire algorithm to the GPU. Terrain data would reside completely in video memory in the form of a texture, and a quadtree structure could be mimicked via indices to a memory location. Vertex lists can easily be generated due to the regular grid layout, and indices could properly be generated with triangles in a vertex shader.

Finally, the algorithm could be modified for rendering terrain at a planetary scale, which would require a specialized acceleration structure for ellipsoidal geometry.

## Acknowledgements

This work is funded by NASA EPSCoR, grant #NSHE 08-51, and Nevada NASA EPSCoR, grants #NSHE 08-52, NSHE 09-41, NSHE 10-69. In addition, this work is partially funded by the CAVE Project (ARO# N61339-04-C-0072) at the Desert Research Institute.

## References


