An analyzing sampled terrain volumetrically with regard to error and geologic variation

Thomas Butkiewicz  Remco Chang  Zachary Wartell  William Ribarsky
UNC Charlotte  UNC Charlotte  UNC Charlotte  UNC Charlotte

ABSTRACT
Most terrain models are created based on a sampling of real-world terrain, and are represented using linearly-interpolated surfaces such as triangulated irregular networks or digital elevation models. The existing methods for the creation of such models and representations of real-world terrain lack a crucial analytical consideration of factors such as the errors introduced during sampling and geological variations between sample points. We present a volumetric representation of real-world terrain in which the volume encapsulates both sampling errors and geological variations and dynamically changes size based on such errors and variations. We define this volume using a hierarchical voxel grid, and demonstrate that when used within simplification and line-of-sight applications, the generated simplified model or calculated visibility information is guaranteed to be within a user-defined confidence level of the real-world terrain.

CR Categories and Subject Descriptors: I.3.m Computer Graphics, Miscellaneous.
Additional Keywords: Terrain, Visualization, Analytics, Geology, Sampling, LIDAR

1 INTRODUCTION
Numerous applications utilize models generated from sampling of real-world terrain. Triangulated irregular networks (TIN) created from digital elevation models (DEM) and other traditional techniques for creating these terrain models, often for the purposes of visualization, lack an analytical understanding of the first steps in the process, namely, producing a model from actual terrain. Yet, this understanding is of vital importance for line-of-sight, trafficability, and other terrain analysis applications. In line-of-sight, for example, the confidence measure for the spatial analysis of where, from a given vantage point, objects would be hidden or visible is much more important than the perceptual goodness of the terrain model for visualization, and this confidence measure must be derived from an analytical understanding of the relation between the sampled model and the actual terrain. In this paper we combine very efficient techniques from visualization and computer graphics with analysis methods to achieve this understanding. We then apply this visual analytics approach to terrain analysis applications, detailing a line-of-sight application and discussing other applications.

Our approach considers not only the sample points, but the sampling methods used to collect them. The geologic qualities of the land are also taken into account when calculating the variations possible between sampling points.

We intelligently calculate a volumetric representation for our terrain model, based on both the sampling techniques and geologic properties of the region. This volume encapsulates the family of all possible surfaces which could have produced the sample data.

We directly utilize this volume for calculations in our applications. For example it can be used along with existing techniques to bound simplification of a TIN of the surface, allowing for a low triangle count mesh that is guaranteed to be within a confidence level of the real-world terrain. (The approach could be used for other mesh representations as well.)

Similarly, in terrain analysis applications, the particular analysis can be applied directly to the bounding volume of the terrain. For example, in line-of-sight, the visibility calculations can be applied directly to the volume, producing a result with a conservative bound on errors and uncertainties introduced during sampling, accounting for possible variations between sample points either locally or globally.

Terrain analysis has the following 3 general challenges: being able to quickly and accurately analyze large areas of densely sampled terrain; being able to sustain measures of error and uncertainty appropriate to the application throughout the analysis process; and providing an analytical framework for understanding the behavior of the particular terrain models used. This paper addresses the first two challenges and provides a foundation for considering the third.

In this work we have produced the following novel capabilities and significant results:
- A new, general approach for carrying forward bounded error and uncertainty measures throughout an analysis process, regardless of the terrain model.
- An effective and efficient method to drive feature-based terrain simplification for an area of focus under the constraint of a given confidence measure.
- A non-uniform, distance-based voxelization with a compact and efficient hierarchical structure.
- A fast line-of-sight algorithm that operates within the hierarchical voxelization and without regard to a particular mesh.
- A method for creating a simplified TIN guaranteed to represent the terrain within a given user-defined confidence level.

2 RELATED WORK
Several areas of related work are relevant to this paper: terrain and mesh rendering in interactive 3D computer graphics, modeling uncertainty in terrain GIS, and methods for implicit surfaces found in computer graphics, computational physics, and related fields. Multi-resolution terrain and mesh rendering in interactive 3D computers graphics is a mature field [12] with seminal works such as [6][8][11] among many others. Advances in efficient
rendering of terrain systems continue to be made [9]. This paper extends work by Garland [6] and Zelinka [5] by adding probabilistic notions to the mesh vertex locations. This concept is not found in other rendering literature.

The GIS community continues to investigate approaches to modeling uncertainty in terrain and its effect on common GIS computations such as line-of-sight computations, hydrology simulations, etc. Santos et al [14] present a method for incorporating uncertainty in terrain modeling by expressing elevations as fuzzy numbers and they construct surfaces that incorporate the uncertainty. They generalize some classic interpolators (linear versus splines, etc.) and compare them qualitatively. Given a set of fuzzy sample elevation points \((x_i, z_i)\) where \(x_i\) is a point on the 2D plane and \(z_i\) is a height value represented as a fuzzy number. Santos et al use Matlab to numerically solve for a triple of surfaces that represent an interpolation of the fuzzy points. The upper surface is an upper fuzzy boundary, the lower surface is the lower fuzzy boundary, and the middle surface is the modal surface. This work is closely related to our work and is similar in spirit to simplification envelopes [21]. However, as pointed out by Zelinka permission grids appear to be faster and simpler to compute with than simplification envelopes. Further, Santos et al is limited to height fields whereas permission grids are not. Anile et al [1] propose the use of fuzzy terrain models to incorporate uncertainty and unpredictable variability in the landscape due to factors such as vegetation, presence of unmapped human artifacts, etc. They incorporate this fuzzy terrain model into line of sight computation over DEMs using discrete lines-of-sight. Their work is basically an extension of Cohen-Or and Shaked [3]. Again, our work is more general since it is not limited to height fields.

Finally, our work has some relation to methods of implicit surfaces. This work is found in computer graphics and computational physics. Blinn [2] is perhaps the first to introduce implicit surface methods from computational physics to the computer graphics community. Implicit surface techniques can be used for interpolating and approximating surfaces from polygon soup [15], while specifying an error tolerance within which the implicit surface should lie relative to the original data. The latter general mesh approach along with level set methods [18] can certainly be applied to the domain of terrain visualization and analysis. However, Permission Grid based approaches appear computationally simpler and faster, especially if one desires to carry the errors in the original sample points into a representation for the errors in the interpolated mesh.

3 SOURCES OF ERROR

Terrain data collected through sampling inherently contains errors and uncertainty [4][13]. Traditional methods address this uncertainty by linearly interpolating between the sample points (such as triangle faces in a TIN or regions between topographic contour lines). In this section, we address sampling errors and geologic variations through an analytical understanding of the errors, and incorporate these errors into our model based on their characteristics.

3.1 Sampling Errors

Many different methods for collecting terrestrial sample points exist, each with their own characteristic output as well as idiosyncrasies which must be accounted for when generating a terrain model from the samples.

An illustrative example is the use of LIDAR. These systems make distance measurements from aircraft to the ground using a pulsed laser. The vertical component of each sample point, being a measurement of the speed of the beam’s return, is highly accurate. However, the position/horizontal component of each sample point is generated from a GPS system which is limited in its accuracy and update rate. To compensate for the update rate, an inertial navigation unit (INU) is used to estimate the aircraft’s location between GPS updates, which introduces additional error into the reported positions of samples [4][13].

The errors introduced from these LIDAR systems become more complex when applying the system to actual terrain. As the slope of the sampled surface increases, the positional error begins to substantially affect the vertical components. Figure 1 illustrates this behavior where positional error affects vertical accuracy based on the terrain’s slope. Notice that because the positional data is acquired through the less accurate GPS and INU, while the vertical or “height” information is acquired using an accurate laser, the combined position from these two instruments can result in a change from point A (where the laser actually struck the surface) to point B (due to the error in the GPS and INU).

$$Error_{Total\ Vertical} = Error_{Vertical} + (\tan(slope) * Error_{Positional}) \quad (3.1)$$

We can then determine a conservative bound on the total error introduced by the sampling process:

$$Error_{Sampling} = \sqrt{(Error_{Total\ Vertical})^2 + Error_{Positional}^2} \quad (3.2)$$

It is crucial to consider these non-uniform sampling errors for applications that depend on knowing where the actual terrain could reside, such as line-of-sight or trafficability, and where one desires an accurate estimation of the certainty of calculations.

3.2 Geologic Variations

Traditionally the areas between sample points are linearly interpolated from the surrounding points (such as the faces in a TIN or the regions between topographic contour lines). This method can incorrectly treat the terrain as a uniform surface constrained only by the sample points. Real terrain behaves erratically, and the data points are only a sampling of a varied terrain. One must instead examine the geologic characteristics of the region and judge to what extent the terrain may diverge from its predicted location based on its distance from a known measurement.
Consider a terrain sampled at 5 meter intervals. If the sampled terrain is a prairie/grassland such as that depicted in Figure 2(a), we would say that the possible vertical difference between two sample points (geologic variation) would be very small, say 0.5 meters. However if the sampled terrain is craggy and prone to unpredictable protrusions, as is shown in Figure 2(b), the geologic variation between two sample points could easily be as much as 4 meters.

These land characteristics might be manually selected from a list of terrain types, entered parametrically for a specific area, or generated automatically from inspecting the variation of the region’s height measurements. Analyzing the terrain to automatically generate these geologic properties is beyond the scope of this paper, but we refer to Santos et al [14], which utilizes fuzzy sets to approximate these properties. We choose to implement a distance-based formula representing geologic variation as a Gaussian distribution about each sample point. Doing so allows us to easily associate a confidence level with distances from sample points based on a user-defined value for the terrain’s geologic variation (\( \Delta \)). We define the maximum geologic error as:

\[ \text{Error}_{\text{Geologic}} = (1 - P(d)) \times \Delta \quad (3.3) \]

Given a point on the terrain that is at distance \( d \) from a sample point, \( P(d) \) is the probability that this point actually lies on the interpolated surface between the sample points. (See Figure 3). \( P(d) \) is further defined as:

\[ P(d) = \text{Normalized} \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} f(\Delta)^2 - \frac{d^2}{\Delta^2}} \right) \quad (3.4) \]

Where \( f(\Delta) \) is a function relating the maximum rise value for the geologic variation to sharpness of the normal distribution around the sample points. \( f(\Delta) = \Delta^2 \) provides a simple but stable relationship as outlined in Figure 4. By normalizing the function we ensure that for a distance \( d \), \( P(d) \) approaches 1 as \( d \) approaches 0, and \( P(d) \) approaches 0 as \( d \) approaches infinity.

As can be seen in Figure 4, if the terrain is craggy (if \( \Delta \) is large), the confidence level in the interpolated surface drops exponentially. However, if the terrain is mostly flat, the confidence level decreases more slowly. In all cases, the maximum geologic error approaches \( \Delta \) as the distance (denoted as \( d \)) to the nearest sample point increases.

As illustrated in Figure 5, the sampling error from Equation 3.2 and the geologic error from Equation 3.3 can be summed together to find the total error:

\[ \text{Error}_{\text{Total}} = \text{Error}_{\text{Sampling}} + \text{Error}_{\text{Geologic}} \quad (3.5) \]

Figure 5. The total error is a combination of the error resulting from the sampling processes and error due to geologic variations. The areas underneath the curves depict the bounded error volume around the surface.

4 IMPLEMENTATION

In order to accommodate both sampling error and geologic variations, we adopt a volumetric representation for the terrain model. The volume bounds the sampling error and geologic variation around the model and represents all possible surfaces of the terrain. The volume is created as a hierarchically discretized voxel grid, similar to the Permission Grids proposed by Zelinka et al [5].

4.1 Sample Data

Our input is formatted as a triangulated irregular network (TIN). We use a TIN because of the relative ease of creating it from digital elevation models (DEM) and other sources of terrain data. TINs preserve the areas of sparse sampling as such, while DEMs...
usually have these areas filled in with interpolated values which, in the final DEM image are not distinguishable from actual sample points. The provided connectivity information is also useful for dealing with irregularly sampled data such as that resulting from LIDAR. We consider the vertices to be the set of sample points and the faces to be a linearly interpolated guess as to the surface in between.

4.2 Grid Creation

We model and store our surface conceptually as a simple 3D voxel-based volume slightly larger than the bounding box of the terrain. This approach is based upon the concept of Permission Grids [5], which Zelinka uses to define a volume around a surface to bound error during mesh simplification.

The volume’s resolution and size must first be chosen. The two specified parameters that guide its creation are a confidence value \( \varepsilon \) value and a precision value \( \alpha \). The value \( \varepsilon \) is the positional error amount for the desired confidence level. (The ratio between positional error and elevation error is provided during voxelization.) For example if the sampling system gives positions accurate within 1 meter 95% of the time, then \( \varepsilon = 1m \) would result in a 95% confidence level. The precision value \( \alpha \) determines how many voxels are used for that particular \( \varepsilon \). (Voxel size = \( \varepsilon / \alpha \)) Zelinka shows that the minimum effective value for \( \alpha \) is \( \sqrt{3} \) [5].

The volume, while conceptually a simple 3D array of voxels, is implemented in a hierarchical data structure similar to an octree. During the voxelization phase, the volume recursively subdivides itself until it reaches the necessary voxel size as determined by \( \varepsilon \) and \( \alpha \). This hierarchical structure allows the creation and use of Permission Grids with minimum storage and memory requirements.

A major benefit of using this discretized volume for our models is the ease of splitting the model into smaller models. This allows calculations on large portions of terrain to be done in parallel, and also facilitates easier storage and level of detail. Additionally, once a volume has been calculated it is easily saved for future reuse.

4.3 Voxelization

In their original form, Permission Grids create a discretized volume guaranteed to be entirely within a constant distance of the mesh. However, by modifying the existing voxelization algorithms for faces and edges we create volumes around the surface using adaptive distances based on the sampling density and geologic qualities. The volume is then distorted in a manner consistent with the sampling errors (see Figure 5).

We voxelize our TIN one face at a time. For each face, we first determine a bounding box of voxels around the triangle. Based on the size of the triangle, we expand the bounding box to be large enough so as to accommodate any possible sampling and geologic errors (see Equation 3.5).

Each voxel in this box is then tested individually to determine if it should be occupied or empty in the final volume. The process of determining each voxel’s status begins by first evaluating if the voxel is closest to an edge or the face. If a voxel is within an acceptable distance (as defined by Equation 3.5) from its closest component, it is considered to be occupied. Note that the acceptable distance is adaptively calculated based on the behavior of sampling errors and local geologic properties.

Figure 6. Given a voxel \((v)\), figures (a) and (b) demonstrate the cases where the voxel is closest to an edge or a face respectively. In both cases, the voxel’s position is projected onto its closest component (the projected position is denoted as \(p\)) and the distance \((d)\) to the nearest sample point \((s)\) is determined.

Possible cases are shown in Figure 6. For the case where a voxel is closest to an edge (see Figure 6(a)), we determine the projected point \((p)\) of the voxel onto that edge. We then find the distance \((d)\) from point \((p)\) to the nearest end (sample) point of the edge. The value \((d)\) is then used in Equation 3.5 to obtain the total error from the interpolated surface at the point \((p)\). If the total error given is less than the length of the vector from the voxel to the edge, then the voxel is considered to be occupied (see Figure 7).

Similarly, for the case where a voxel is determined to be closest to the face itself (see Figure 6(b)), we find the nearest point on the face to that voxel. Once again we find the distance from that point to the nearest sample point and use that value in the same fashion as the edge case.

As illustrated in Figure 7, if the resulting total error is more than the distance from the voxel to its nearest component, then the voxel is occupied in our model as it must be considered a possible location of the surface.

Figure 7. The volume surrounding the surface is discretized and any voxels entirely within these error bounds (see Figure 5) around the surface are occupied in our model and considered a portion of the possible terrain.

Figure 8 shows the effect of our volumetric approach given an arbitrary equilateral triangle for a range of values for \(\Delta\). The blue lines represent the voxel grids, which in turn represents the bounding volume. Notice the increasing semi-circles emanating from the sample points; this is consistent with our probabilistic model shown in Figure 3. As \(\Delta\) increases, the size of the thickness of the volume perpendicular to the face increases accordingly. The bounds surrounding the edges can be seen behaving similarly to the lines in Figure 4.
Figure 8. (a) Top down and side view of the volumetric result (shown in blue) of the voxelization process conducted on an arbitrary equilateral triangle. (b) Results for different geologic variation ($\Delta$) values. Notice in all instances the thickness increases farther from sample points.

Figure 9. Above: Original TIN generated from a DEM with samples every 10 meters. Insert shows rendered appearance of the original TIN. Below: Simplified TIN and insert showing rendered appearance of simplified TIN. Notice that not only does the simplified TIN appear visually consistent with the original model, but it is also guaranteed to be within a confidence level of the real-world terrain.

5 APPLICATIONS

There are many applications in terrain analysis that do not depend on visual goodness, but instead require a more precise, analytical representation of the model. Our volumetric representation of the terrain guarantees that all surfaces within the volume are mathematically bounded to a given confidence level based on the sampling error and geologic variations. Through our algorithm, we demonstrate that we can simplify a TIN such that the resulting simplified TIN has fewer triangles, but is still guaranteed to be within the desired confidence level from the actual terrain in the real-world. Similarly, our line-of-sight application gives precise bounded probabilities to objects visible and invisible from a vantage view volume.

5.1 Generating Simplified TINs

Many commercial and free applications exist that can convert widely available digital elevation models to TINs. LIDAR scanning data is often returned in this format as well. However, the resulting meshes are usually unnecessarily dense, which leads most users to simplify these meshes into a manageable size. Unfortunately, most terrain simplification algorithms available focus on creating visually appealing models, but not on the validity of the model itself.

Our algorithm efficiently creates a volumetric representation of the terrain taking sampling errors and geologic variations into account. Any surface that falls within this volume is mathematically proven to be a valid model. We adopt an iterative face decimation technique (QSim[6]) and simplify the model one face at a time until any further decimation would cause the resulting model to fall outside of our volumetric boundary.

For applications that require precise analytics of the terrain model, the resulting TINs produced by our algorithm give a mathematical guarantee on confidence (precision). Figure 9 illustrates the result of simplification of a terrain within our bounding volume. Not only does the appearance of the simplified model remain visually consistent with the original model, but it provides a guarantee that it is within a level of confidence from the real-world terrain.

5.2 Line of Sight / Visibility

Unlike most traditional line-of-sight or visibility algorithms (see [19] for a survey of some of these algorithms), the goal of generating a volumetric encapsulation models is to place a conservative bound on the possible real-world surface that could have produced the sampling data. This is important for all terrain analyses and, in particular, for applications concerned with visibility calculations. We demonstrate that our algorithm integrates easily and efficiently within a line-of-sight application for determining areas of invisibility to military sentry units. Unlike most existing line-of-sight applications where visibility is considered to be either visible or invisible, our application gives confidence levels to the visibility. Due to the errors and geologic variations in the sampled data, such confidence levels are a more realistic representation.

We conduct our visibility calculations entirely within the hierarchical data structure, without regard to any triangle mesh.
We store our volumetric visibility output in a separate, identical structure.

By volumetrically calculating line-of-sight we gain the ability to specify volumetric patrol areas (volumes) instead of exact eye points similar to [20] (the view volume appears as red cubes in Figure 10). The volumetric output allows us to consider the hiding of larger extended units such as heavy artillery. By varying our error tolerance during voxelization, we can produce volumes that provide any desired certainty measure. For large and complex terrain models, the line-of-sight calculations can be done in parallel as a speed-up due to the independent structure of our voxel grid.

First, voxels that represent the view volumes are determined. A standard 3D DDA algorithm [7] is used to traverse the voxels along rays cast from that eye point outward in all directions to the edges of the bounding box. All occupied voxels are considered as occluders, and voxels with an occluder between them and the view voxel are considered to be occluded.

Voxels found to be invisible from the view volumes (shown in red) are displayed as a transparent fog over textured TINs simplified as described in section 5.1. Displaying invisible regions as volumes instead of areas permits the qualification of hiding for units of substantial height.

As shown in Figure 10, the resulting volume, voxelized in all invisible areas, is then displayed transparently accompanied by the simplified terrain mesh. The resulting visibility calculation can then be interactively viewed providing the user an idea of the potential “fog of war” remaining after placing that Sentry unit.

If desired, additional eye points (or view volumes) can be added and their visibilities calculated and subtracted from the original invisibility volume. In this way the user can interactively place units to minimize the unobserved areas of the map.

6 Future Work and Discussion

Although representing geologic errors as Gaussian distributions provides an intuitive formulation for terrain, we believe that better models can be created to accommodate both terrain and urban models. Urban models differ drastically from terrain in that variations in urban models could be abrupt whereas variations in terrain are more natural and gradual. For example, in sampling urban models, one sample point could fall on the ground, while the next is on the top of a skyscraper. We theorize that formulation for such drastic variations could better be modeled using step-wise functions instead of Gaussian distributions.

For the line-of-sight application, we would like to make use of non-boolean occupancy of the voxels. Such voxels would no longer be outright occluders, but occlude adaptively based on desired confidence levels.

Strengthening the line-of-sight algorithm to detect when rays pass through, then below the surface, and emerge once again, could provide output with core volumes guaranteed to be invisible with 100% certainty. Adaptively subdividing the voxels during the visibility tests could also lead to a finer level of detail in the output.

7 Conclusion

We combine techniques from visualization and computer graphics with analysis methods to achieve an understanding of the errors involved with sampling a terrain and generating a model from said samples. We apply our approach to terrain analysis in the forms of terrain simplification and line-of-sight determination.

We demonstrate that by using a volumetric representation for terrain models, we can encapsulate all possible surfaces for the sample points given sampling errors and geologic variations. We further demonstrate the effectiveness of this volume representation when used along with existing techniques as well as new analytical methods such that all analysis is performed entirely within the voxel-based data structure. Using our technique, we can generate simplified terrain models, or determine visibility on the terrain with an associated confidence level given the sampling errors and geologic variations.

Acknowledgment

This work is supported by the Army Research Office under contract no. W911NF-05-1-0232.

Reference


