Global Grids and the Future of Geospatial Computing

Kevin M. Sahr
Department of Computer Science
Southern Oregon University

Southern
Terra
Cognita
Laboratory

The Situation

- Geospatial computing has achieved many impressive results
- But it now faces unprecedented challenges

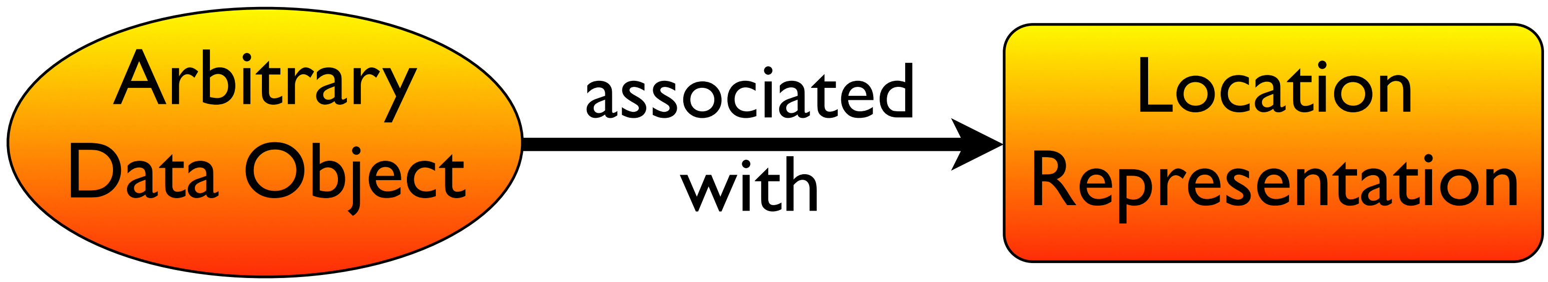
A Perfect Storm

- pressing global-scale problems like climate change
- vast quantities of geospatial “big data”
  - ✦ manipulate truly global datasets
- consumer-level 3D cloud-based visualization and analysis
- pervasive GPS-enabled mobile devices
  - ✦ real-time location-based services

A Pivotal Moment

- we can no longer expect to meet these challenges piecemeal
- a revolution in our fundamental approaches to geospatial computing is required
- the outlines of this revolution are now known and are indisputable

Geospatial Computing



Point Measurement	*point/vector*
Areal Measurement	*pixel*
Complex Agent	*polyline*
Event	*polygon*
etc.	*arbitrary set of above*

Traditional Location Representation

- attempt to mimic pre-computer approaches to location in
  - ✦ mapping
  - ✦ scientific analysis

Vector Location

Vectors of Real Numbers

3D ECEF coordinates
2D map projection space
polar latitude/longitude

➔ **Represented Using**

3- or 2-tuples of floating point values

Manipulated Using Synthetic and Analytic Geometry

➔ **Approximated Using Floating Point Operations**

Areal Units

Square Cells
photographic pixels
units of discrete analysis
& simulation

➔

Represented Using
square raster
indexed using 2-tuple of
integers
stored linearly

Traditional Problems

- traditional representations are familiar and convenient
- but it was quickly recognized that they are inefficient for many common operations
  - spatial queries
    - window, proximity, etc.
  - spatial predicates
    - intersection/collision, neighbors, etc.

Traditional Solution

- spatial databases/data structures
  - structured
    - attempt to preserve spatial structure
    - primary example: quadtree
  - semi-structured
    - attempt to optimize spatial access
    - examples: R-tree, ESRI SBN/SBX

Traditional Spatial DBs

- adds linear spatial index
  - correspond to buckets containing locations
    - more efficient spatial access
    - coarse filter for specific algorithms
- ***but*** retains underlying representation for
  - final stage of many algorithms
  - arbitrary spatial operations

Advent of "Digital Earth"

- would provide platform for:
  - ✦ simulation, visualization, & analysis
  - ✦ truly global data interaction
  - ✦ "mother of all (geospatial) databases"
- exemplifies the challenges facing geospatial computing

Abridged Early History

- 1962: *Geoscope* (R. Buckminster Fuller)
- early 1990's: variety of interactive 3D globes
  - ✦ SGI (⇒*GeoFusion*⇒*ESRI ArcGlobe)*
  - ✦ *Autometric, Inc. (⇒Boeing⇒Keyhole)*
  - ✦ 1993: US DoD: interactive globe with real-time distributed data display
- 1998: christened by Al Gore
- 2005: *Google Earth* (via *Keyhole*)

Google Earth Today

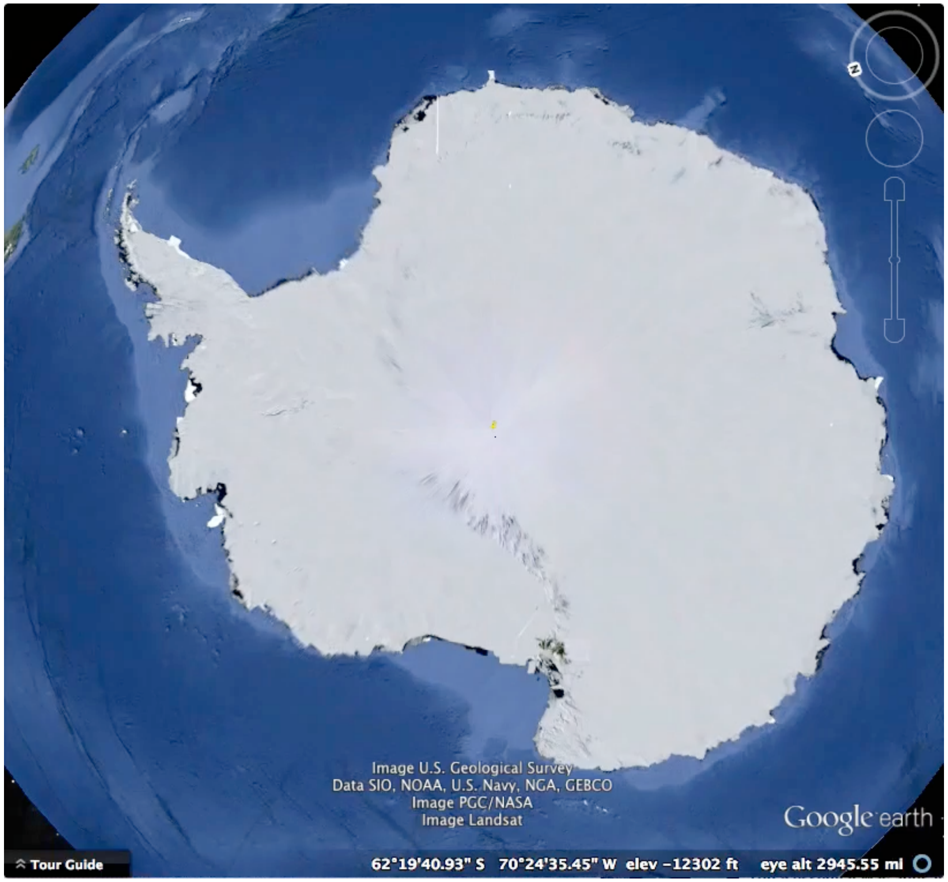
- Google Earth demonstrates the demand for, and potential of, the Digital Earth concept
- implements best practices (circa 2005) in:
  - ✦ geographic information systems (GIS)
  - ✦ computer cartography
  - ✦ computer graphics
  - ✦ spatial databases
  - ✦ web delivery of inter-operable data formats

Google Earth Today

- The Problem: *All of the fields of research that Google Earth implements are still in their relative infancies*
  - ✦ at most 50 years old
    - ✤ not very old for a science!
  - ✦ some (like the web) are much newer

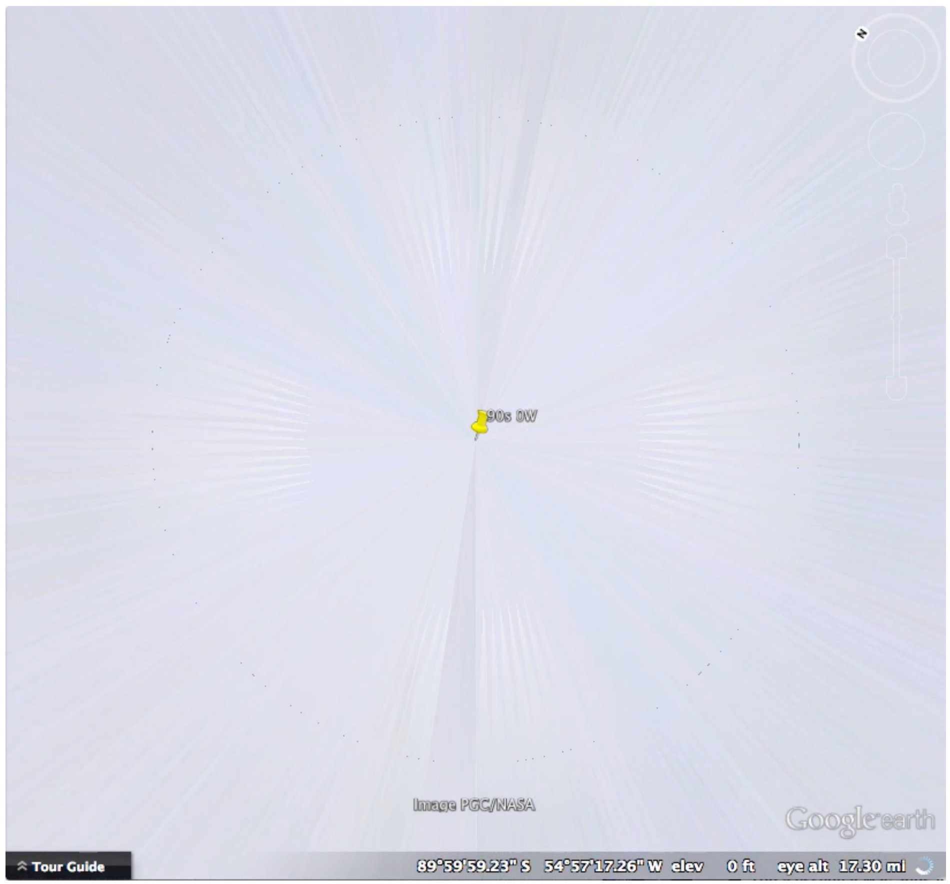
An Example

- Zoom in to the South Pole in Google Earth



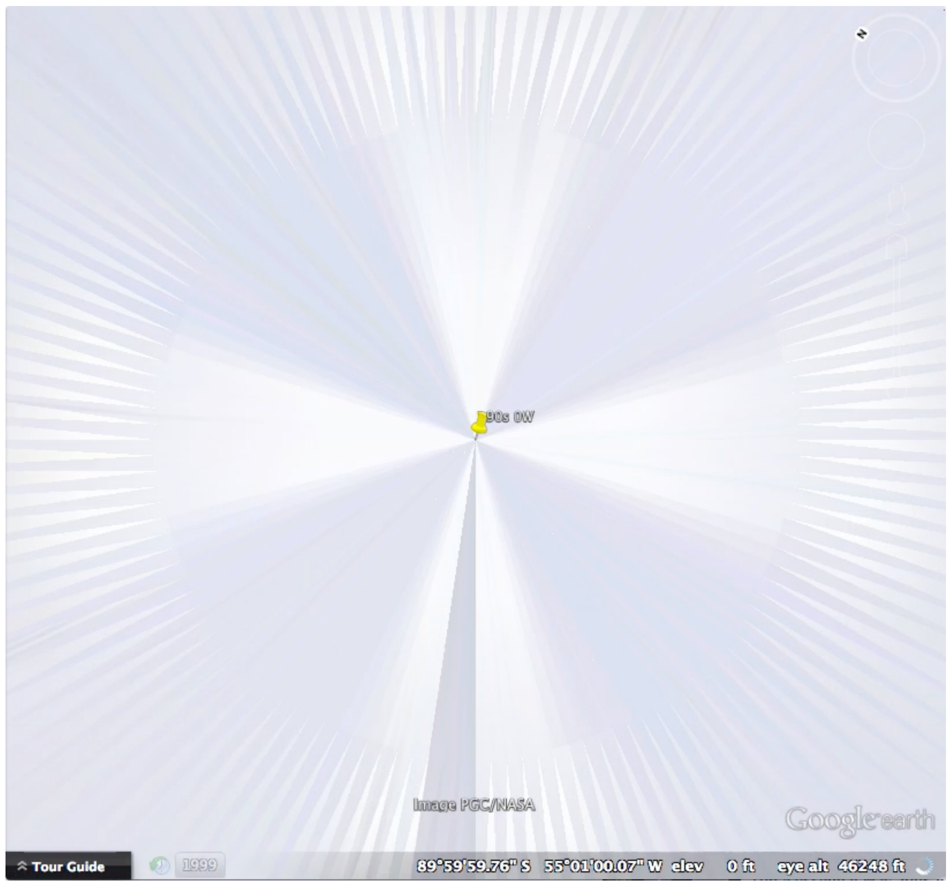
An Example

- Zoom in to the South Pole in Google Earth



An Example

- Zoom in to the South Pole in Google Earth

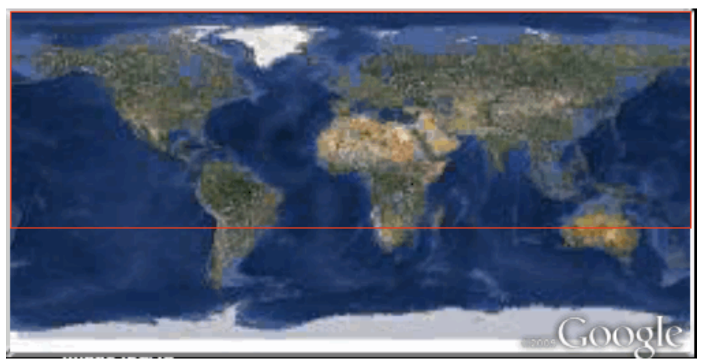


An Example

- Zoom in to the South Pole in Google Earth

Limitations

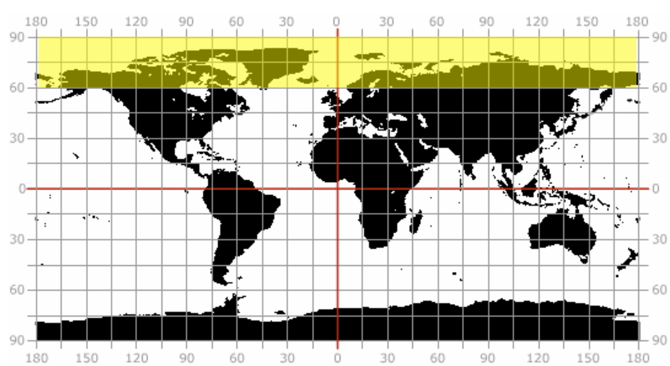
- Only the last stages of the graphics pipeline are truly 3D
- Google Earth uses traditional planar rasters for actual data representation

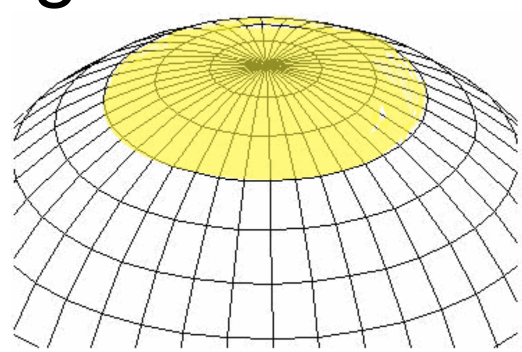


- texture-mapped onto 3D surface

Limitations

- The most common “global” representations treat the earth as a rectangle
  - ✦ latidude/longitude
  - ✦ cylindrical projections
- very inefficient
- contain topological singularities





A Deeper Issue

- vectors of real numbers
  - ✦ are continuous and infinite
  - ✦ allow for exact synthetic geometry
  - ✦ allow for exact analytic geometry

A Deeper Issue

- ***but*** tuples of floating point values
  - ✦ are discrete and finite
  - ✦ do not allow exact synthetic geometry
  - ✦ do not allow exact analytic geometry
- the simplest operations (e.g.,equality test) can result in profound semantic errors
- bounding the rounding error on individual operations can be difficult
  - ✦ on complex systems can be impossible

A Deeper Issue

- floating point values are no more “exact” than integer values
- given ***n*** bits, we can distinguish 2^{n} unique values
  - ✦ all other points must be mapped to these
- ***all*** computer representations of location — both raster ***and*** vector — are necessarily discrete

The Challenge

- in order to realize the full promise of digital earth applications we must find fundamental location representations that:
  - accurately reflect the topology of the earth
  - acknowledge the essential discreteness of location representation
  - are optimal in terms of
    - representational efficiency
    - semantic fidelity

Primary Spatial Key

- GIS introduced the powerful notion of a primary spatial key
  - ✦ singular location representation used system-wide
- ***but*** notion of primary spatial key is subverted in digital earth systems

Spatial Key Illusion

- data analyst sees two base 10 real numbers with well-defined precision
- computer stores two base 2 floating point values, often with loosely specified precision
- often accessed via a linear spatial index
- 3D graphics system requires 3D values
- primarily manipulated as linear array index

New Role for Traditional Representations

- traditional representations will continue to play an important role
- but we can no longer treat them as privileged primary spatial keys
- they can be generated as needed
  - ✦ or stored as a data attribute

Primary Spatial Index

- digital earth systems must replace the traditional notion of a primary spatial key
- new notion: primary spatial index
  - ✦ One Index to Rule Them All
    - ✤ unifies 2D and 3D system components
    - ✤ explicitly discrete
    - ✤ explicitly spherical

The Next Stage

The future is already here — it's just not very evenly distributed.

- William Gibson

Discrete Global Grid Systems (DGGS)

- regular, multi-precision partitions of the sphere into cells with linear indexes
- indexes can indicate either raster or vector location
  - these are equivalent from the perspective of the data structure
- hierarchical indexes enable coarse filter algorithms

Current Applications

- DGGSs have primarily been used in extending traditional grid applications to the globe
  - ✦ discrete simulation
  - ✦ distributed survey sampling
  - ✦ analysis of gridded data sets
- analysis and visualization of DGGS cells still takes place primarily using traditional 2D GIS applications

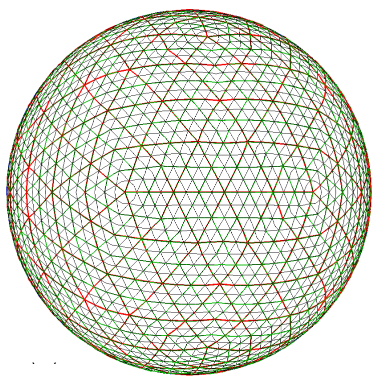
DGGS & Digital Earth

- DGGSs have seen limited use in digital earth applications like Google Earth
  - ✦ role limited to polyhedral meshes used to approximate sphere for 3D graphics
- recognized as replacement for traditional vector representations since 1970's (Dutton)
  - ✦ but so far not exploited

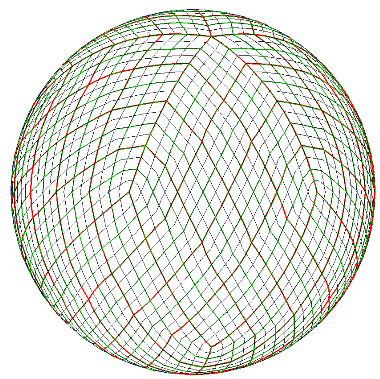
Cell Shape

- DGGSs have been developed with a number of different cell shapes, including:

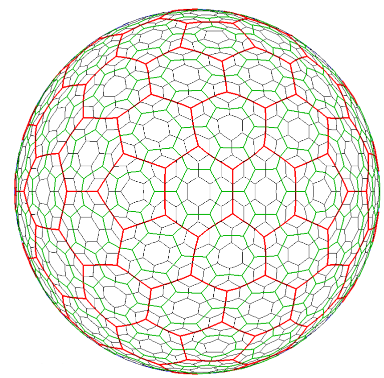
triangle



quadrilateral



hexagon



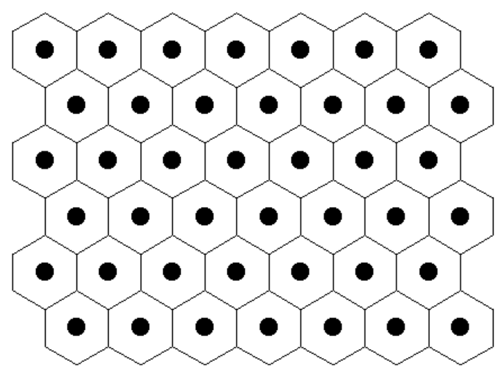
- any reasonable DGGS would provide a better basis for location representation
- but the evidence clearly supports one cell shape as superior

Raster Representation

- compared to square pixels, hexagon pixels
  - ✦ are 13% more efficient at sampling
  - ✦ 25% to 50% more efficient for common image processing algorithms
- the human eye uses a hexagonal arrangement of photoreceptors

Vector Representation

- the human brain represents location using a hexagonal arrangement of neurons
- quantization to the points of a hexagonal lattice is optimal using multiple formulations
    - ✦ least average quantization error
    - ✦ covering problem
    - ✦ packing problem



Conclusions

- achieving the full potential of geospatial systems requires a revolution in our fundamental approach to location representation
- hexagonal DGGSs provide the best known basis for the next generation of location representations
- work still needs to be done to explore and evaluate alternatives for hierarchical indexing

SOUTHERN
TERRA
COGNITA
LABORATORY

www.discreteglobalgrids.org