On the Optimal Representation of Vector Location using Fixed-Width Multi-Precision Quantizers

Kevin M. Sahr
Department of Computer Science
Southern Oregon University

SOUTHERN TERRA COGNITA LABORATORY

FWFP

- most common representation for vector location is as a tuple of floating point values with a fixed number of bits
  - Fixed-Width Floating Point (FWFP)
- but FWFP values are discrete and finite approximations to continuous and infinite real numbers
- is there a better representation for vector location?

Discrete Representation

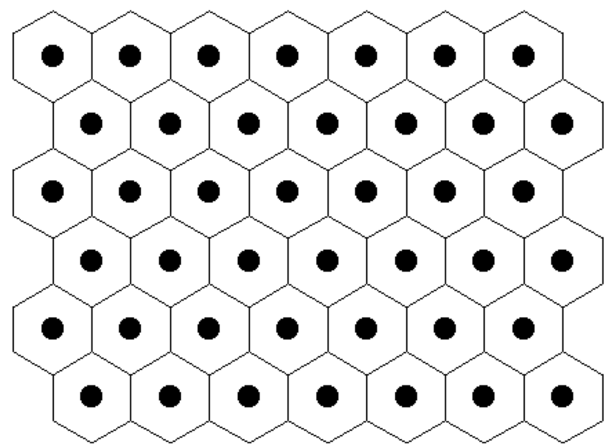
- given a fixed number of bits ***n***, we can distinguish at most 2^n unique values
  - ✦ all other points must be mapped to these values
  - ✦ form 2^n equivalence classes with respect to location

Questions

- given ***n*** bits for a fixed-width representation of vector location
  - ✦ what is the most efficient arrangement of the $2^{\boldsymbol{n}}$ fixed points?
  - ✦ what properties of the real numbers do we want our representation to maintain?
  - ✦ are there representations that can maintain those properties?

Efficient Quantization

- the points of a hexagonal lattice provide optimal quantization using multiple formulations
  - ✦ least average quantization error
  - ✦ covering problem
  - ✦ packing problem



Exact Representation

- the real numbers are capable of exact representation
  - representing values with infinite precision and zero location uncertainty
- example: north pole is at exactly 90° north latitude
  - adding more digits does not increase the precision of that representation

Exact Representation

- FWFP representations are incapable of exact representation (without metadata)

Single Precision

- FWFP representations encode quantization at a single precision
- example: quantize the real number 99.67628
  - ✦ at whole unit precision: 100
  - ✦ at 1/10 unit precision: 99.7
- truncating digits does not yield a valid coarser resolution quantization
  - ✦ values must be rounded
  - ✦ cannot be progressively transmitted

Balanced Ternary

- there are representations of real numbers that encode multi-precision quantization
- one is balanced ternary (Lalanne, 1840)
- radix-3 system using digits -1, 0, and 1
- truncation and rounding are equivalent
- can be progressively transmitted

The Goal

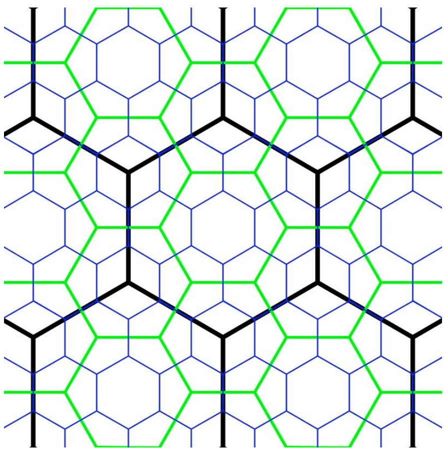
- create a vector location representation that
  - ✦ uses explicitly discrete integer indexes
  - ✦ uses optimal hexagonal lattice points
  - ✦ capable of exact representation
  - ✦ capable of multi-precision quantization

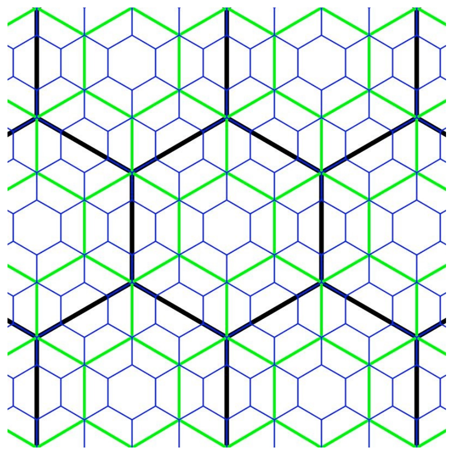
Central Place Apertures

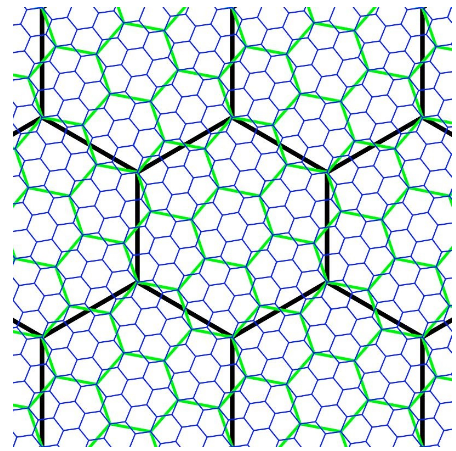
- multi-precision hexagon lattices can be created with an infinite number of apertures
    - aperture: ratio of cell areas between precisions
- research has focused on the Central Place (Christaller, 1966) apertures 3, 4, and 7

Central Place Apertures

aperture 3 aperture 4 aperture 7





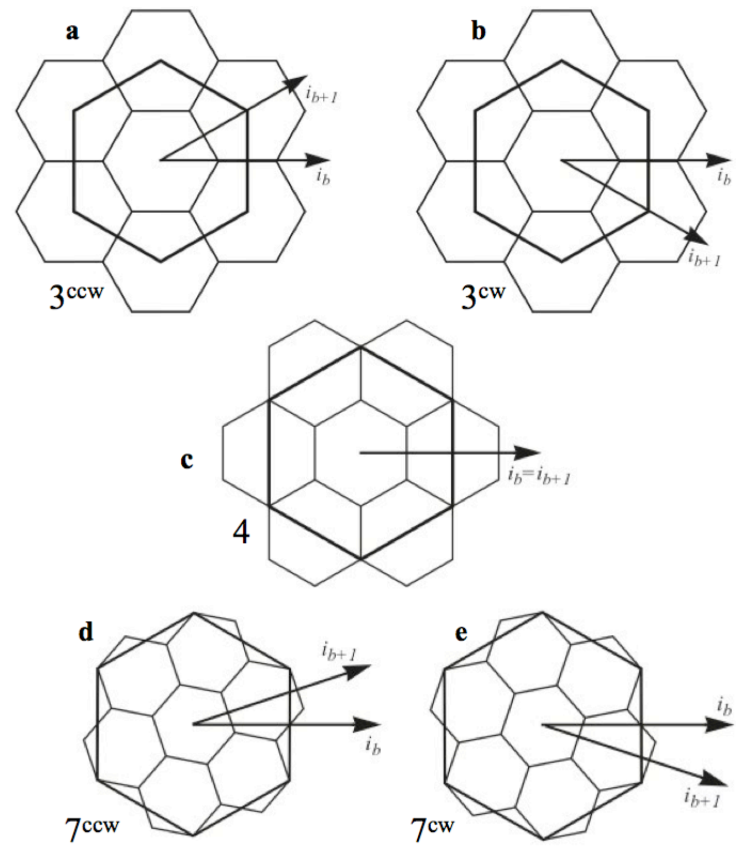


Prefix Codes

- hierarchical prefix codes have many advantages for hierarchical spatial indexes
    - each digit in index corresponds to a single precision in the representation
        - provides locality preserving total ordering
        - implicitly encodes precision
        - provides efficient generalization via truncation

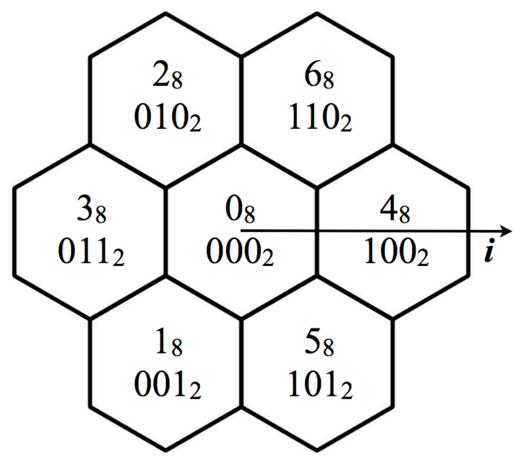
Central Place

- note that in central place apertures each cell naturally has 7 finer precision potential indexing children



GBT

- Generalized Balanced Ternary (GBT) (Gibson & Lucas, 1982) is an extension of 1D balanced ternary to the 2D aperture 7 case
- indexing children add the appropriate digit to their parent's index



CPI

- we can apply this arrangement to the aperture 3 and 4 cases
- we call the result Central Place Indexing (CPI)
  - ✦ provides uniform indexing for all 3 apertures
  - ✦ allows for indexing mixed-aperture precision sequences

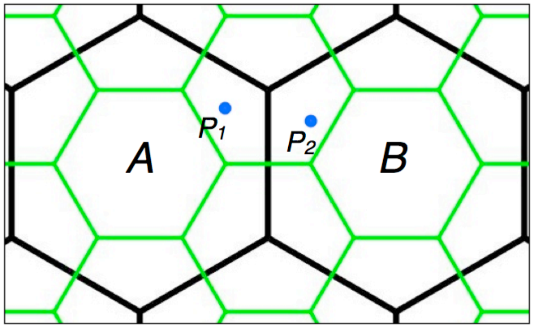
CPI Algorithms

- we have implemented planar CPI algorithms for
  - ✦ forward & inverse quantization
  - ✦ addition/translation
  - ✦ subtraction
  - ✦ metric distance
- implemented using efficient per-digit table lookups

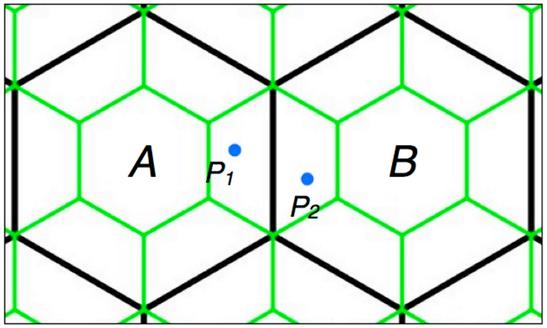
Multi-Precision Quant.

- apertures 3 and 4 CPI indexing encodes true multi-precision quantization
  - ✦ point quantization is performed at each successive precision

aperture 3



aperture 4



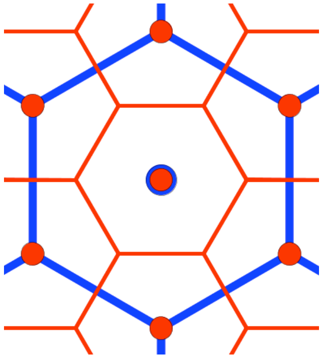
Exact Representation

- using 3 bits per digit, CPI children use digits 0_8 (000_2) through 6_8 (110_2)
- leaves digit 7_8 (111_2) for other purposes
- we can use this digit to indicate exact representation
  - ✦ indicates that any further digits would be 0
  - ✦ provides infinite precision with finite representation

Exact Representation

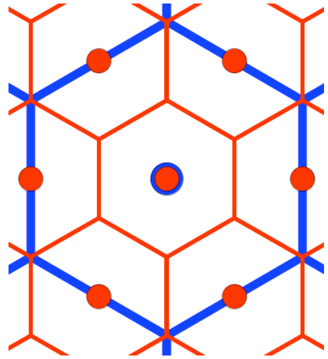
- for all three apertures, the center point of a cell with index ***A*** is exactly described by ***A7***
- depending on aperture, the cells ***A17***, ***A27***, ..., ***A67*** would exactly describe the cell:

aperture 3



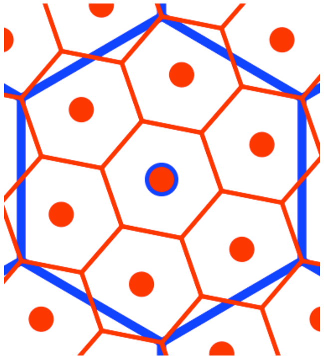
vertices

aperture 4



edge midpoints

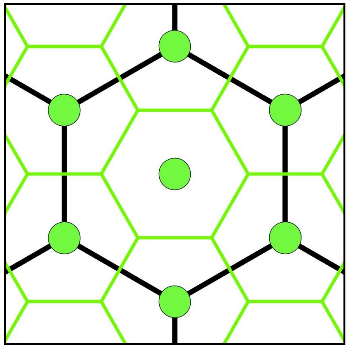
aperture 7

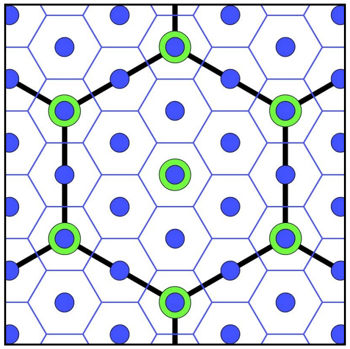


interior points

Self-Describing

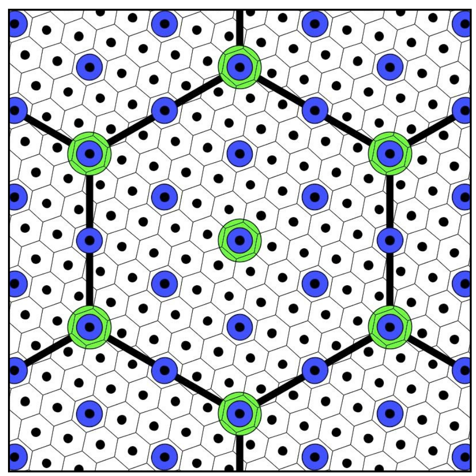
- by adding an aperture 3 and 4 precision to the finest precision of any CPI system we create a system that can exactly represent all cell center points, vertices, and edge midpoints at all system precisions
  - ✦ we call such a system self-describing





347-Suffix System

- adding an additional aperture 7 grid allows sub-frequency addressing of the internal cell regions
  - ✦ we call this a 347-suffix system

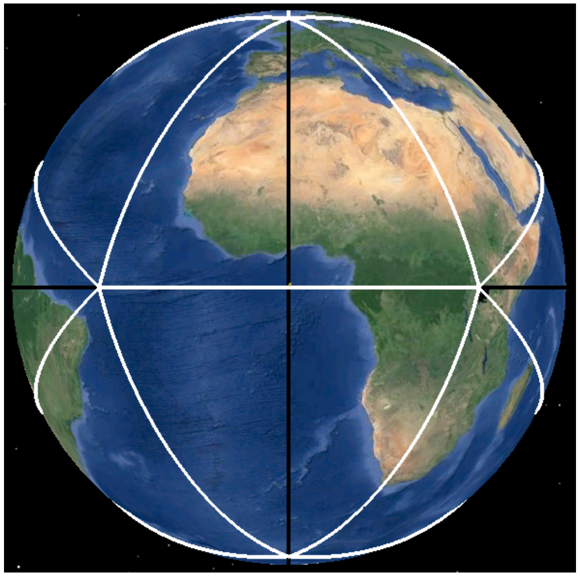


The Sphere

- we can extend any CPI system to the spherical icosahedron
- note that cells centered on the icosahedral vertices are pentagons
  - ✦ we can apply CPI indexing to them by deleting one of the non-centroid indexing sub-sequences
- we can now use what we’ve learned to design an optimal geospatial vector location system

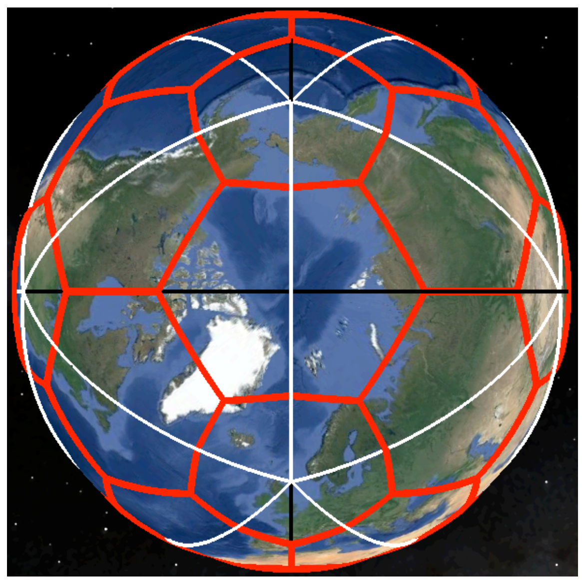
Orientation

- we align the icosahedron symmetrically about the equator and prime meridian using its octahedral symmetry



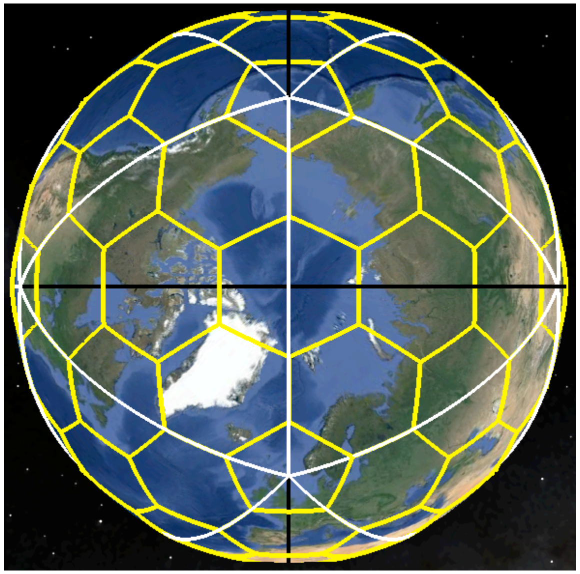
First Precision

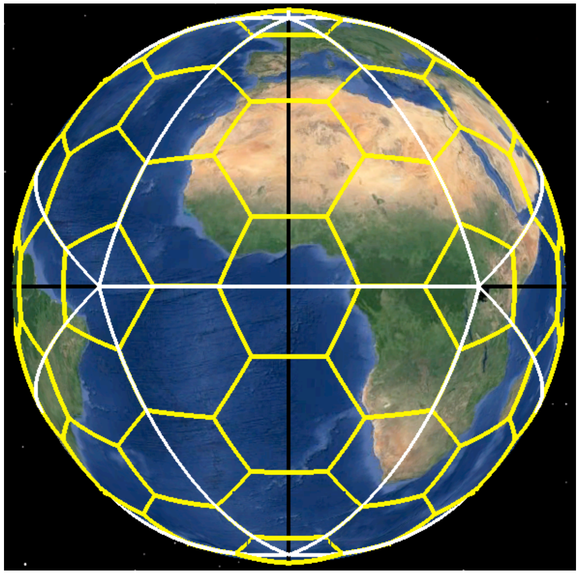
- we choose aperture 4 for the first precision
    - ✦ centers cells on the octahedral vertices



Second Precision

- we choose aperture 3 for the second precision
  - ✦ centers cells on the icosahedron face midpoints





Base Tiles

- we choose these cells as the 122 base tiles for our CPI system
  - ✦ number them 0 (North Pole) to 121 (South Pole)

CPI43 Fixed Width System

- we encode the base tile number in 7 bits
- we use the last bit in this byte to indicate whether or not the base tile value is exactly encoded
  - ✦ we call this a CPI43 Fixed Width System
  - ✦ example: 00000001_2 exactly represents the North Pole in all CPI43 Fixed Width Systems

Additional Precisions

- we can now add additional CPI precisions as appropriate for our application
- aperture 3 and 4 precisions provide multi-precision quantization
- aperture 7 precisions can be added first, to quickly zoom to the desired spatial resolution
- any system can be made self-describing by adding precisions to make it a 347-suffix system

Conclusions

- we have presented an optimal global grid that provides fixed width vector location representation that meets all of our desired properties
  - ✦ integer hierarchical prefix code indexes
  - ✦ hexagonal fixed point arrangement
  - ✦ exact representation
  - ✦ multi-precision quantization

SOUTHERN
TERRA
COGNITA
LABORATORY

www.discreteglobalgrids.org