Visual Analytics Using Density Equalizing Geographic Distortion

Peter Bak

bak@dbvis.inf.uni-konstanz.de

Park City / Utah, 23.09.2008

P. Bak*, D.A. Keim, M. Schaefer, A. Stoffel, I. Omer

Why density equalization?

Visualizing the flood of large-scale geographic data can not cope with available screen space.

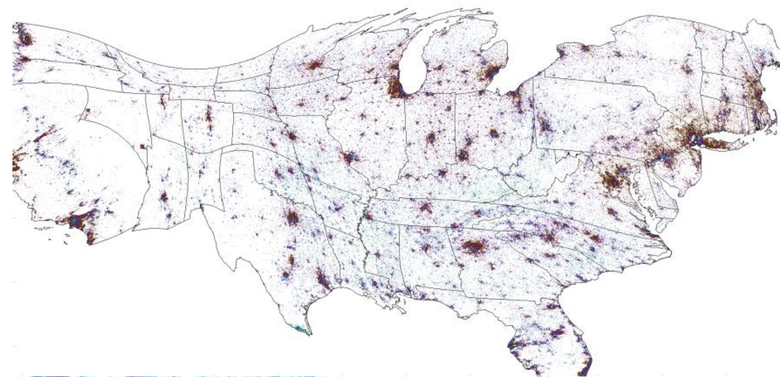
- Dens areas are hardly visible and accessible.
- Sparse areas waste space and bias perception.
- Analytic tasks – clustering, correlations, comparisons – can not be carried out.

☞ create density-equalized maps
while preserve recognizable features
and neighborhoods in the visualization

Related Approaches

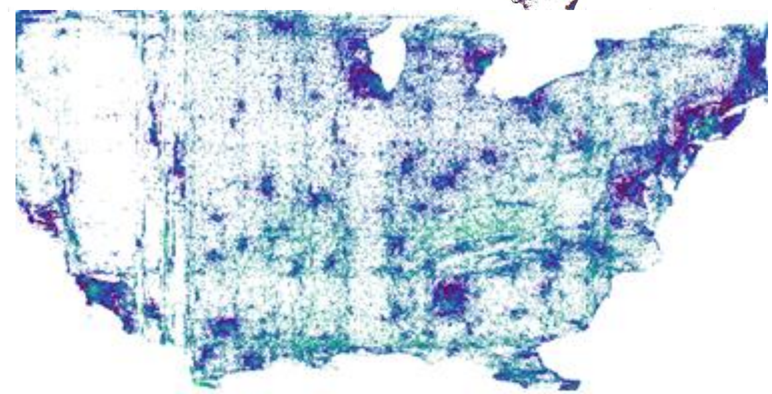
Cartograms:

- ✓ Applicable for polygons
- ✗ Point data indirectly distortable
- ✗ Highly dependent on data resolution



HistoScale:

- ✓ Applicable to point data
- ✓ Good for reducing noise / outliers
- ✗ Creates artifacts (Euclidian dim.)

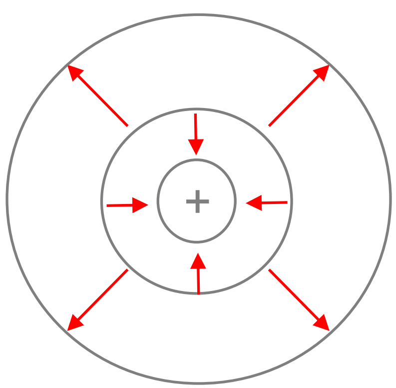


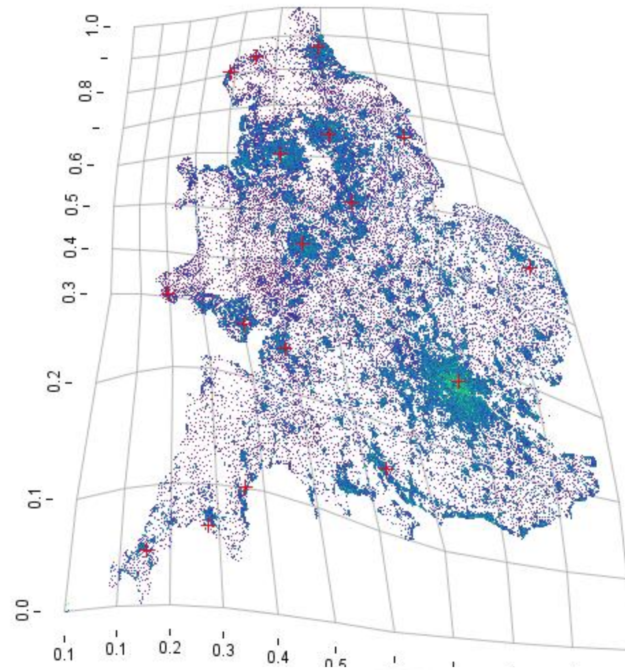
Pixel Placement:

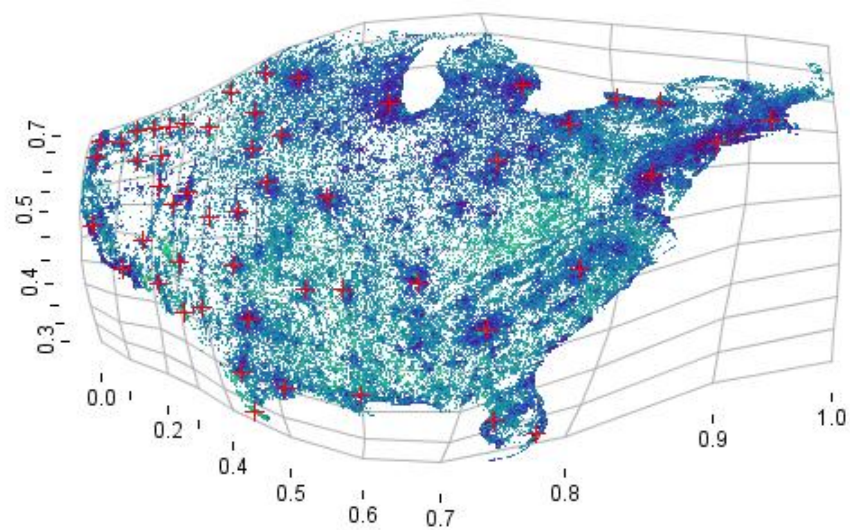
- ✓ Applicable to polygons and points
- ✓ Avoids overlap
- ✓ Equalized density
- ✗ Discontinuous neighborhood

Proposed Approach – Radial Scale

1. Determining centers of interest (<u>high density areas</u>)
2. Defining <u>radial areas</u> around each center point with equal area-size
3. <u>Resize</u> radial areas in accordance <u>with their density</u>
4. <u>Scale with</u> *exponential decrease* of the center's effect with distance.

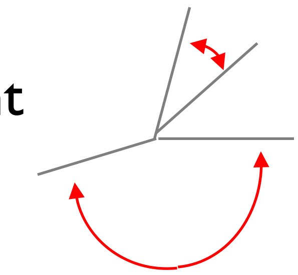


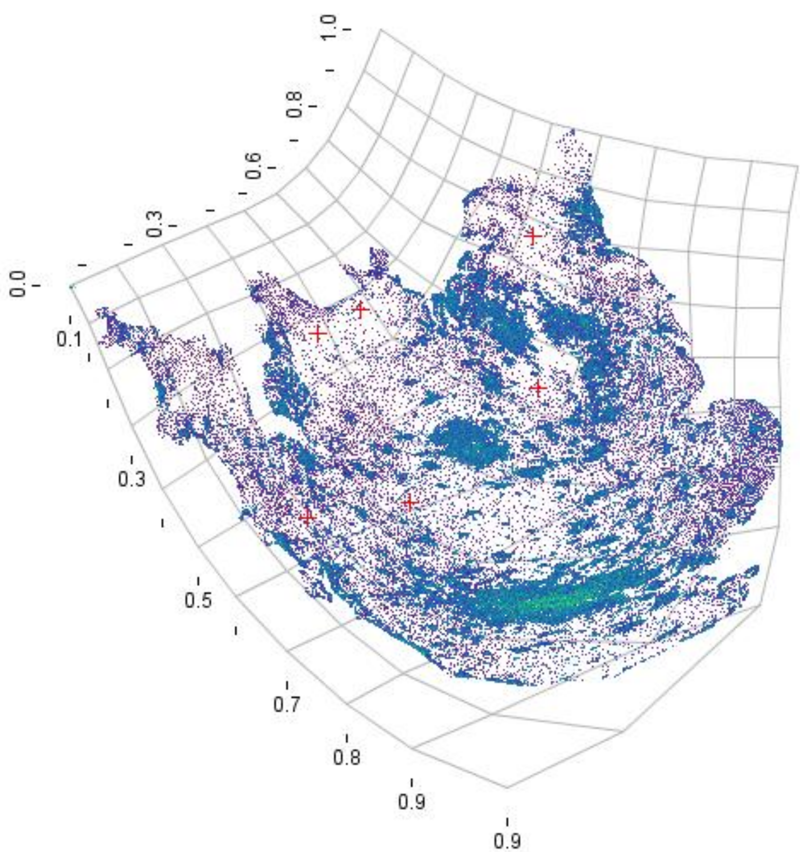


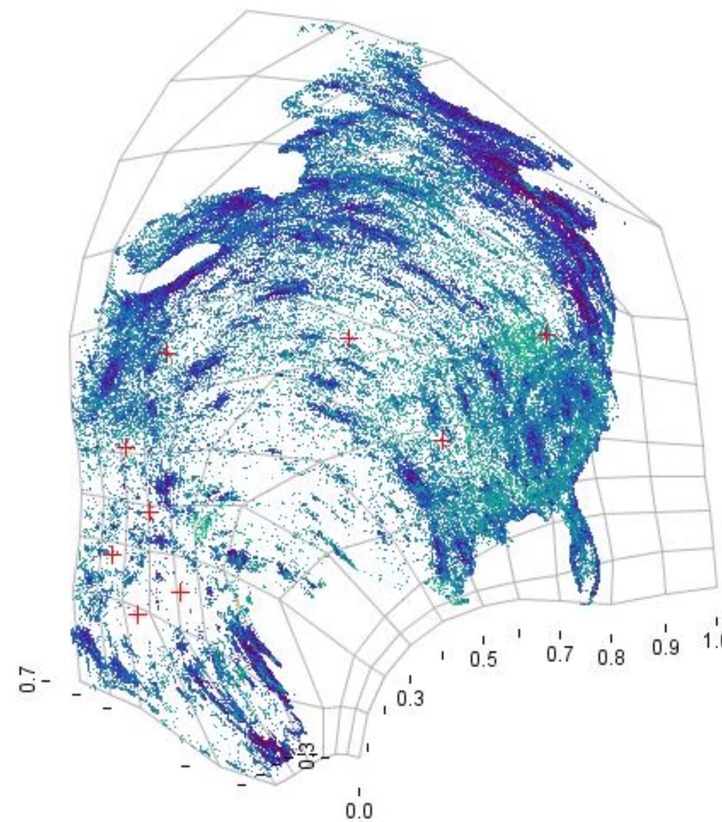


Proposed Approach – Angular Scale

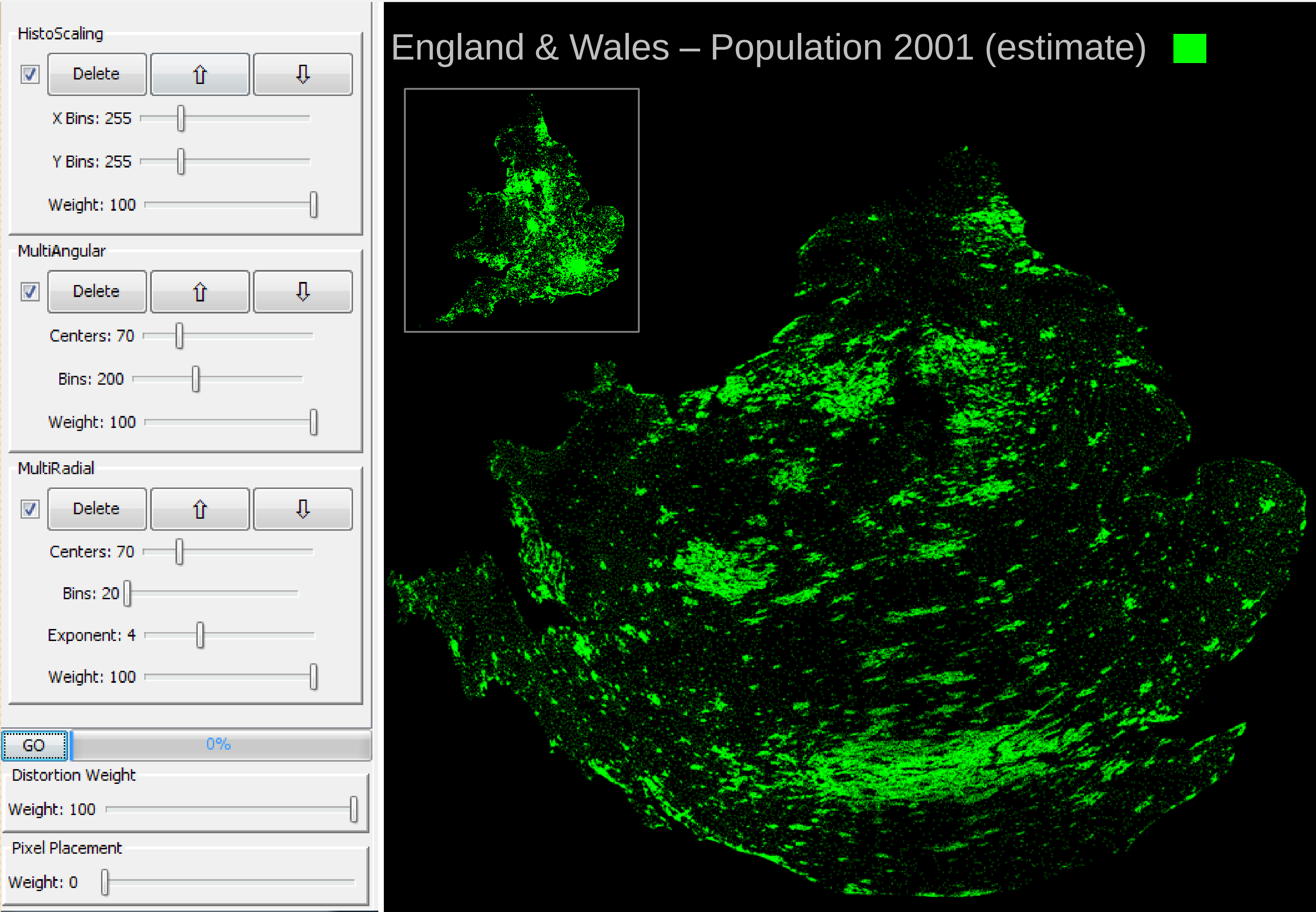
1. Determining centers of interest (<u>low density areas</u>)
2. Defining <u>angular</u> segments around each center point
3. Resize angular areas in accordance with <u>density.</u>
4. Apply <u>scaling</u> for all points around all centers.



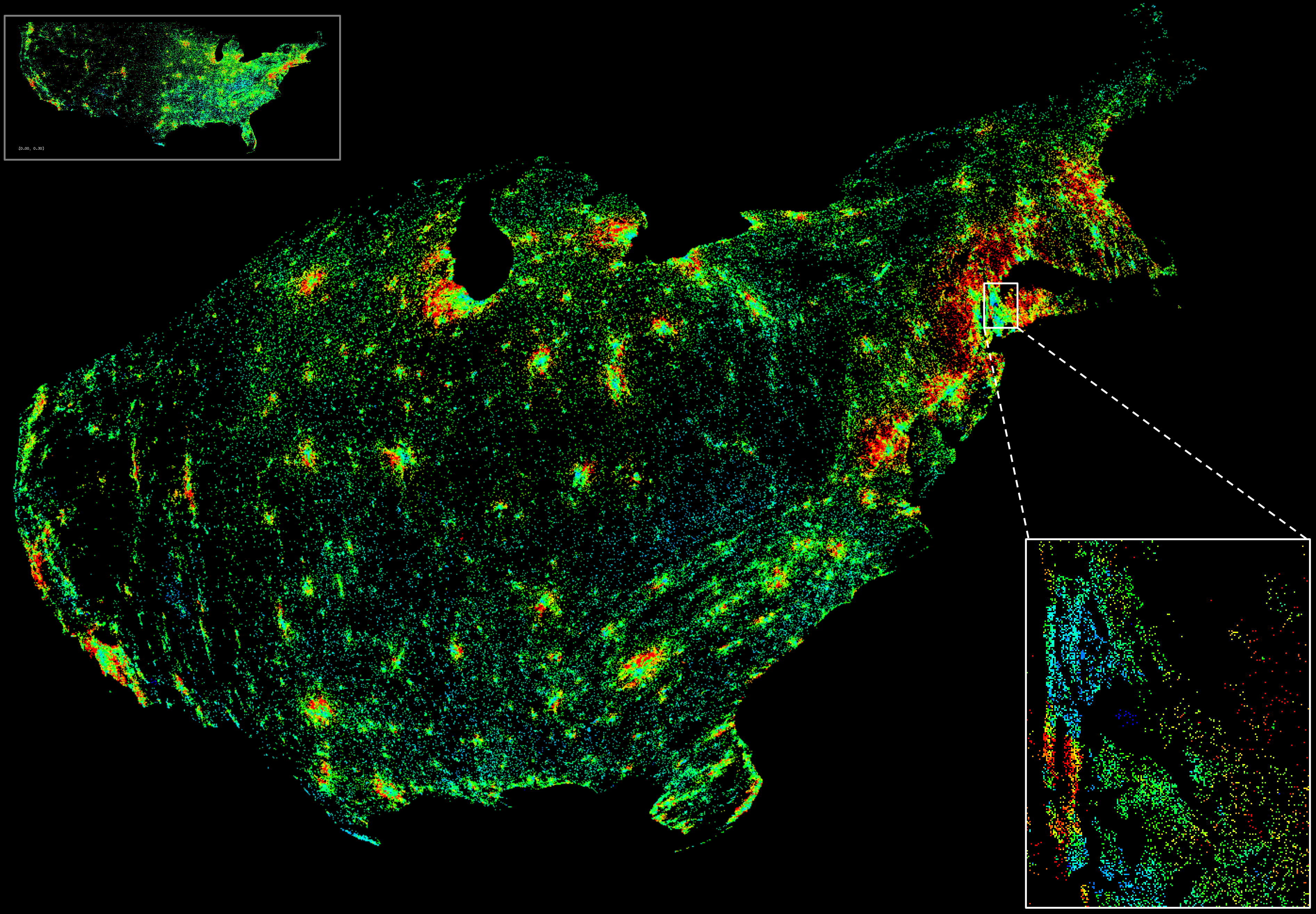




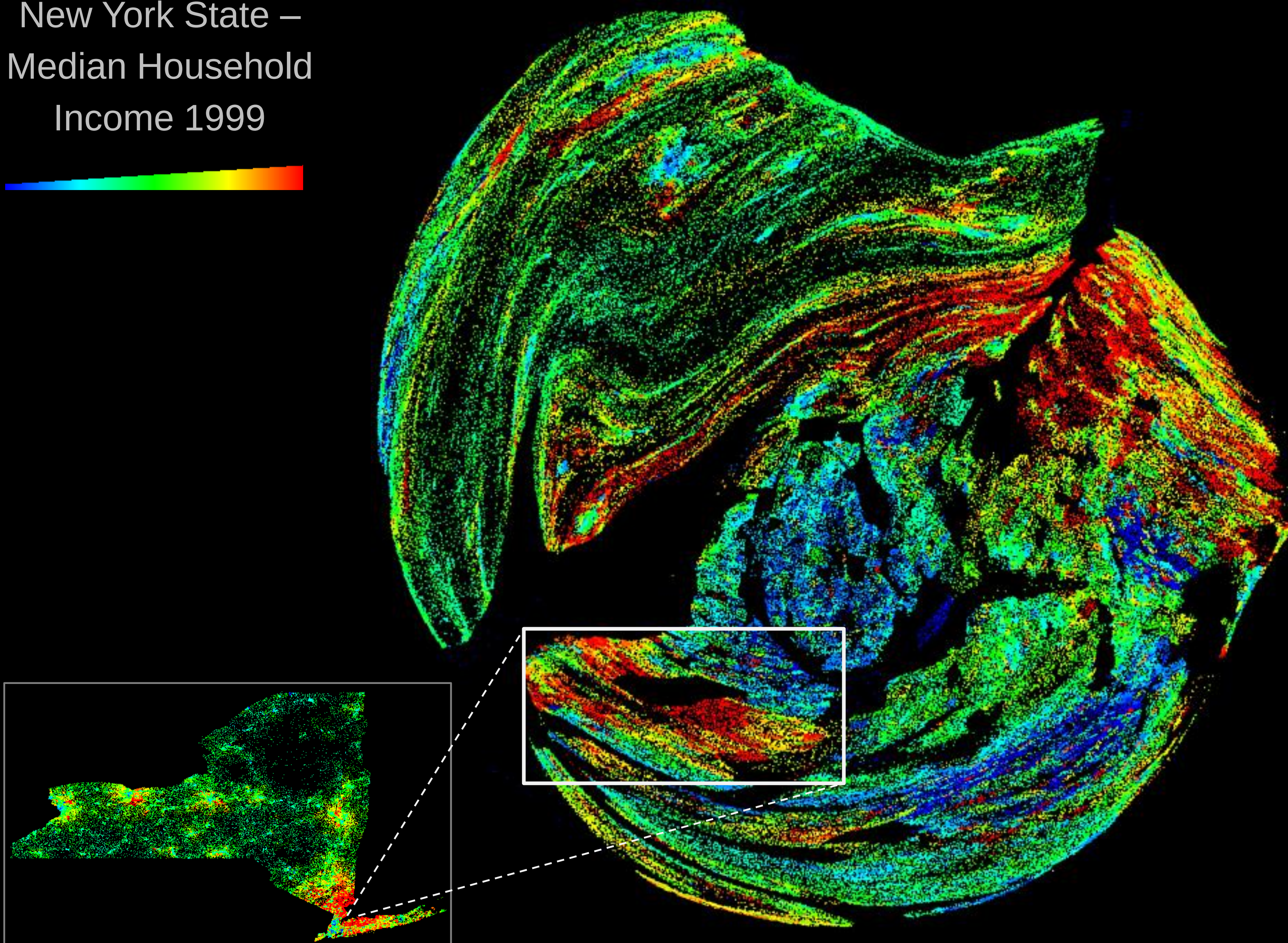
Framework: Pipeline of Separate Distortions



US – Median Household Income 1999

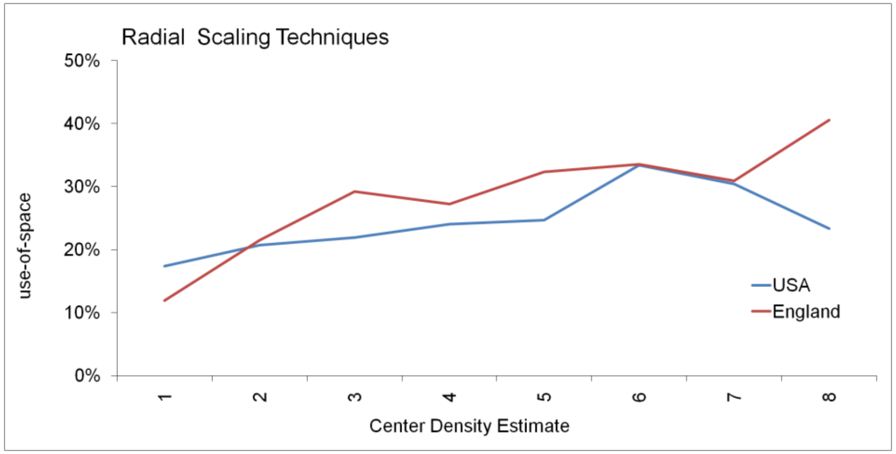


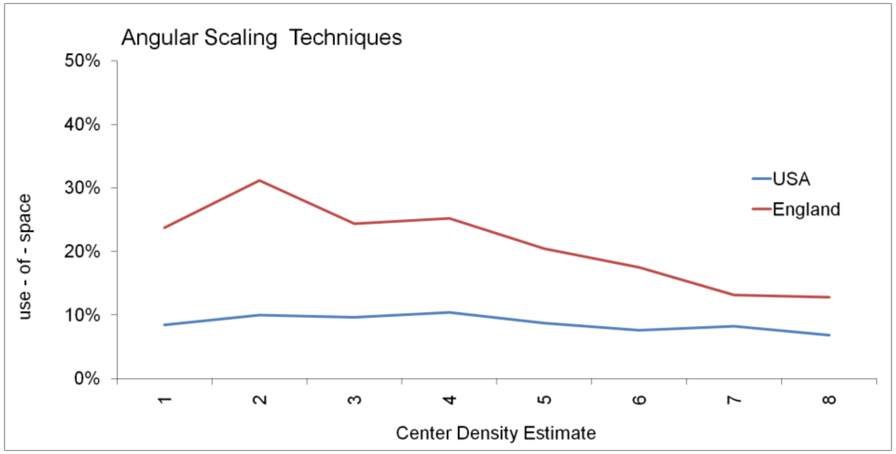
New York State – Median Household Income 1999



Use-of-Space Estimate

Estimate the amount of efficiently used screen-space:





Similar Patterns for both datasets:

- Performance improves with more centers, just until a certain point, after which it decreases.

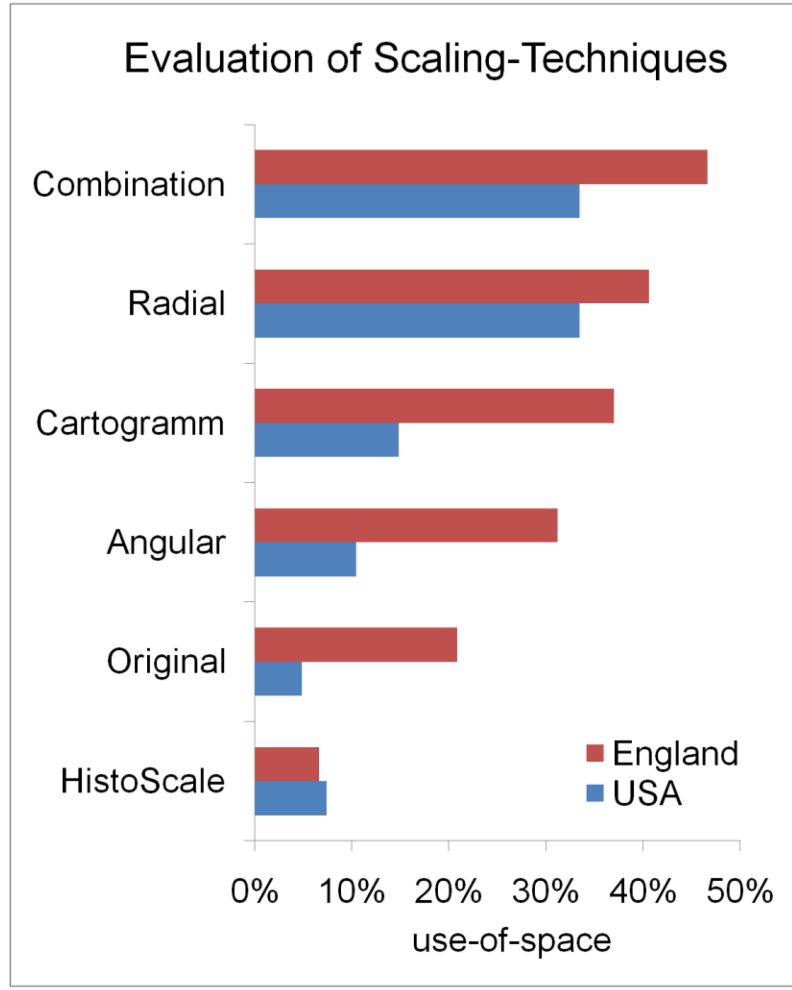
Data Dependency:

- England data shows a clear peak of performance.
- USA showed no effect of number of center points on performance.

Discussion

Advantages depend on:

- ✓ Data Properties:
  - ➢ Distribution in space,
  - ➢ Size of dataset
  - ➢ Type of data
- ✓ Task of the Distortion:
  - ➢ Highlighting,
  - ➢ Area-relations,
  - ➢ Density-equalization
- ✓ Type of evaluation measures:
  - ➢ Use – of – space
  - ➢ Shape-keeping
  - ➢ Information Visualization



Techniques should adjust to properties of the data.

User-involvement improves performance

THANK YOU FOR YOUR ATTENTION

Computation – Radial Scale

The input for the distortions are the data-points P = {p1; p2; p3; …}, the centers C = {c1;c2;c3; …}, and the number of bins to use *w*.
With d_{ij} as the distances between point p_i and the center c_j.

Area of one bin: $$a_j = 2\pi / w \cdot \max_{p_i \in P} \left(d_{i,j}^2 \right),$$

Radius of a bin *k*: $$r_j^{(k)} = \sqrt{k \cdot a_j / (2\pi)},\ k \in \{1, 2, \ldots, w\}$$

new Radius of a bin *k*: $$r'^{(k)}_j = \sqrt{\frac{\left|\left\{ p_i \,|\, d_{i,j} \le r_j^{(k)} \right\}\right|}{|P|} \frac{(w \cdot a_j)}{2\pi}}$$

Relative Radius of point in the bin: $$r'^{(b_{i,j})}_j$$

Bin number of point and center: $$b_{i,j} = \left\lfloor \frac{2\pi d_{i,j}^2}{a_j} \right\rfloor + 1$$

With of a bin *k* before distortion: $$\Delta r_j^{(k)} = r_j^{(k)} - r_j^{(k-1)}$$ after: $$\Delta r'^{(k)}_j = r'^{(k)}_j - r'^{(k-1)}_j$$

new distance of point from center: $$d'_{i,j} = d_{i,j} + \left(r'^{(b_{i,j})}_j - \frac{r_j^{(b_{i,j})} - d_{i,j}}{\Delta r_j^{(b_{i,j})}} \cdot \Delta r'^{(b_{i,j})}_j - d_{i,j} \right) \frac{1}{e^{d_{i,j}}}$$

Computation – Angular Scale

The input is the point set P, the centers C, and the number of bins to use *w*. With d_{ij} as the distances between point p_i and the center c_j.

Angle of one point for center: $\phi_{i,j} \in [0, 2\pi)$

Bin number of point for center: $b_{i,j}$

Scaled Angle of bin *k*: $a'^{(k)}_j$

(Bin number and scaled angle: analog to Radial Scale)

Relative Angle of point in the bin: $a'^{(b_{i,j})}_j$

new angle of point from center: $\phi'_{i,j} = \phi_{i,j} + \left(a'^{(b_{i,j})}_j - \left(b_{i,j} - \frac{\phi \cdot w}{2\pi} \right) \cdot \Delta a'^{(k)}_j - \phi_{i,j} \right)$

Final location of a point as vector sum: $p'_i = p_i + \sum_{j=1}^{|C|} (p'_{i,j} - p_i)$