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 (high density areas)
2. Defining radial areas around each center point with equal area-size
3. Resize radial areas in accordance with their density
4. Scale with exponential decrease of the center’s effect with distance.
Proposed Approach – Angular Scale

1. Determining centers of interest (low density areas)
2. Defining angular segments around each center point
3. Resize angular areas in accordance with density.
4. Apply scaling for all points around all centers.
Framework: Pipeline of Separate Distortions

England & Wales – Population 2001 (estimate)
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.
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 \( P = \{ p_1; p_2; p_3; \ldots \} \), the centers \( C = \{ c_1; c_2; c_3; \ldots \} \), 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:

\[
a_j = \frac{2\pi}{w} \cdot \max_{p_i \in P} \left( d_{i,j}^2 \right),
\]

Radius of a bin \( k \):

\[
r_j^{(k)} = \sqrt{k \cdot a_j / (2\pi)}, \quad k \in \{1, 2, \ldots, w\}
\]

new Radius of a bin \( k \):

\[
r_j^{(k)} = \sqrt{\left( \frac{1}{|P|} \sum_{i} \left( p_i \cdot d_{i,j} \leq r_j^{(k)} \right) \right) \left( w \cdot a_j \right) / 2\pi}
\]

Relative Radius of point in the bin:

\[
r_{j}^{(b_{i,j})}
\]

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:

\[
\Delta r_j^{(k)} = r_j^{(k)} - r_j^{(k-1)} \quad \text{after:} \quad \Delta r_j^{(k)} = r_j^{(k)} - r_j^{(k-1)}
\]

new distance of point from center:

\[
d'_{i,j} = d_{i,j} + \left( r_{j}^{(b_{i,j})} - d_{i,j} \right) \cdot \Delta r_{j}^{(b_{i,j})} \cdot \frac{1}{e^{d_{i,j}}}
\]
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 $k$: $\alpha_j^{(k)}$

Relative Angle of point in the bin: $\Delta \alpha_j^{(b_{i,j})}$

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

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