Data Visualization (DSC 530/CIS 568)

Geographic Visualization

Dr. David Koop
Colormap

- A colormap specifies a mapping between colors and data values
- Colormap should follow the expressiveness principle
- Types of colormaps:

**Binary**

- y
- n

**Diverging**

- -1
- 0
- +1

**Categorical**

- T
- F
- A

**Sequential**

- 3
- 2
- 1

[Munzner (ill. Maguire), 2014]
Continuous Colormap

US EPA Regional Oxidant Model -- Midwest
Ozone (ppbv): June 26, 1987, 18:00

[Bergman et al., 1995]
Segmented Colormap

US EPA Regional Oxidant Model -- Midwest
Ozone (ppbv): June 26, 1987, 18:00

[Bergman et al., 1995]
Continuous vs. Segmented

![Continuous vs. Segmented](image1.png)

**Lookup Task (Lower)**

![Plot of Absolute Mean Error (m) vs. Bin Size](image2.png)

[Padilla et al., 2017]
Don't Use Rainbow Colormaps

Which has a discontinuity?

[M. Bussonnier]
Other Colormaps Work Better

Which has a discontinuity?

[M. Bussonnier]
Artifacts from Rainbow Colormaps

[Borland & Taylor, 2007]
"Get It Right in Black and White" - M. Stone

Matlab jet colormap

Matlab jet colormap (B&W)

[S. Eddins (Matlab Blog), 2014]
"Get It Right in Black and White" - M. Stone

Matlab parula colormap

Matlab parula colormap (B&W)

[S. Eddins (Matlab Blog), 2014]
D3's color scales

- [https://github.com/d3/d3-scale-chromatic](https://github.com/d3/d3-scale-chromatic)
- In v5, included in default bundle (no separate import)
- D3's built-in color scales
- Derived from ColorBrewer
- Sequential and diverging scales created using interpolation
- Hue *can* change, but be careful
- Color ramp [M. Bostock]
Bivariate Colormaps

[Image of bivariate colormaps with categories and sequences]

[Munzner (ill. Maguire), 2014]
Bivariate Colormap (Uncertainty → Saturation)

[Correll et al., 2018]
Value-Suppressing Uncertainty Palette

[Correll et al., 2018]
Assignment 3

• Due next Thursday (3/28)
• Geographic Visualization and Colormaps: D3 Map Example
• Uses data about affordable housing in Massachusetts
Projects

• **DO NOT PLAGIARIZE!**

• Many ideas are relatively boring (one very creative idea)

• Examples of creative ideas:
  - [http://mbtaviz.github.io](http://mbtaviz.github.io)

• Need to think about data attribute types

• Tasks are the "why" of the visualization
  - Trends need to be better defined. Why do we care?
  - Think about the justification for the task

• "What" and "Why" come before "How"
Geographic Data

• Spatial data (have positions)

• Cartography: the science of drawing maps
  - Lots of history and well-established procedures
  - May also have non-spatial attributes associated with items
  - Thematic cartography: integrate these non-spatial attributes (e.g. population, life expectancy, etc.)

• Goals:
  - Respect cartographic principles
  - Understand data with geographic references with the visualization principles
Map Projection

[P. Foresman, Wikimedia]
Flattening the Sphere?

Central Meridian (selected by mapmaker)

Great distortion at high latitudes
Examples of two rhumb lines (direction true between any two points)
Equator touches cylinder if cylinder is tangent
Reasonably true shapes and distances within 15 degrees of Equator

[USGS Map Projections]
Lambert Conformal Conic Projection

Two standard parallels
(selected by mapmaker)

Large-scale map sheets can be joined at edges if they have the same standard parallels and scales.
Standard Projections

Regular Cylindrical

Regular Conic

Transverse Cylindrical

Polar Azimuthal (plane)

[J. P. Snyder, USGS]
Map Projections

What your favorite map projection says about you
Projection Classification

![Projection Classification Diagram](image)

[J. van Wijk, 2008]
Myriahedral Projections

After labelling edges as folds and cuts, we obtain two spanning trees of graphs are a well-understood that fold-over is rare by observing that the following almost never lead to fold-overs, and we do not explicitly test on this. The problem of fold-overs is complex, we found empirically that the schemes we use in the folding out mesh should not only be planar, it should be such that the foldout does not suffer from fold-overs. The folded out mesh should not only be planar, it should be such that the foldout does not suffer from fold-overs. The folded out mesh should not only be planar, it should be such that the foldout does not suffer from fold-overs.

Foldout

The term spanning tree suggests a solution for labelling. Consider two neighbouring triangles these faces are never split up. Unfolding is done by first by inserting interior edges with very high weights, such that the new angle is not a right angle. All images were produced with a custom developed, supersampling per pixel with a jittered grid was used, gnomonic projection here. Rendering maps for presentation purposes requires control the projection in the interior. We use a simple spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The geography of the earth (or whatever image on a spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The use of an arbitrary supersampling per pixel with a jittered grid was used, gnomonic projection here. Rendering maps for presentation purposes requires control the projection in the interior. We use a simple spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

In the following sections, we discuss various choices for the algorithm to produce a myriahedral projection is now as follows: All vertices should have one or more adjacent faces. Consider two neighbouring triangles these faces are never split up. Unfolding is done by first by inserting interior edges with very high weights, such that the new angle is not a right angle. All images were produced with a custom developed, supersampling per pixel with a jittered grid was used, gnomonic projection here. Rendering maps for presentation purposes requires control the projection in the interior. We use a simple spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The geography of the earth (or whatever image on a spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The use of an arbitrary supersampling per pixel with a jittered grid was used, gnomonic projection here. Rendering maps for presentation purposes requires control the projection in the interior. We use a simple spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The geography of the earth (or whatever image on a spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The use of an arbitrary supersampling per pixel with a jittered grid was used, gnomonic projection here. Rendering maps for presentation purposes requires control the projection in the interior. We use a simple spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The geography of the earth (or whatever image on a spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/

The use of an arbitrary supersampling per pixel with a jittered grid was used, gnomonic projection here. Rendering maps for presentation purposes requires control the projection in the interior. We use a simple spherical surface has to be displayed) is mapped as a texture vanwijk/myriahedral). (examples are shown in http://www.win.tue.nl/
Cut along parallels or meridians (graticules)

The simplest way to define a mesh is to use the graticule itself, and to cut along parallels or meridians. The results can be used as an introduction to map projection. A weight for edges, using the value of $w$ and $l$ of the midpoint of an edge, can be defined as $w(l) = \frac{W_w}{w_j} w_j \left\{ \frac{w_0}{z} + \frac{l_{\min}}{l_j} \right\}$, where $W_w$ and $W_l$ are overall scaling factors, and $w_0$ and $l_0$ denote where a maximal strength is desired. Different values for these lead to a number of familiar looking projections (Figure 3). The use of a high value for $W_w$ gives cuts along meridians. Dependent on the value of $w_0$ a cylindrical projection ($0 \theta$, equator), an azimuthal projection ($90 \theta$, North pole), or a conical projection ($25 \theta$) is obtained when the meridian strips are unfolded. Use of a negative value for $W_w$ gives two hemispheres, each with an azimuthal projection. The meridian at which to be centred can be controlled by using a low value for $W_l$ and a suitable value for $l_0$. The use of a high value for $W_l$ gives cuts along parallels. Unfolding these parallels gives a result resembling the polyconic projection of Hassler (1820).

The relation between a spatially varying weight $w$ and the decision where to cut and fold can be understood by considering Prim's algorithm. Suppose, without loss of generality, that we start at a maximum of $w$ and proceed to attach the edges with the highest weight. At some point, edges at the boundary will have approximately the same weight and, after a number of additions, a ring of faces is added, with cuts in between neighbouring faces in this ring. Hence, edges aligned with contours of $w$ typically turn into folds, whereas edges aligned with gradients of $w$ turn into cuts.

Each strip is almost free of angular or area distortion, however, a large number of interrupts occur with varying widths. These gaps visualise, just like the Tissot indicatrix, the distortion that occurs when a non-interrupted map is used, and can be used to explain the basic problem of map projection. If we want to close these gaps, the strips must be broadened. However, to maintain an equal area, they have to be shortened, and to maintain the same aspect ratio they have to be lengthened, which is not possible simultaneously. Also, it is clearly visible that mapping a point (such as a pole) to a line leads to a strong distortion. When the number of strips is increased, the gaps are less visible, and the distortion is shown via the transparency of the map (Figure 4).

Figure 3. Graticular projections, derived from a 5$\theta$ graticule. 2592 polygons: a) cylindrical; b) conical; c) azimuthal; d) azimuthal, two hemispheres; e) polyconical

[J. van Wijk, 2008]
Subdividing regular polyhedra

For the graticular projections, thin strips of faces are attached to one single strip or face. This is a degenerated tree structure. In this section, we consider what results are obtained when a more balanced pattern is used. To this end, we start with Platonic solids for the projection of the globe, and recursively subdivide the polygons of these solids. This approach has been used before for encoding and handling geospatial data (Dutton, 1996).

At each level \( i \), each edge is split and the new centres, halfway on the greater circle connecting the original endpoints, are connected. As a result, for instance each triangle is replaced at each level by four smaller triangles. Other subdivision schemes can also be used, for instance triangles can be subdivided into nine smaller ones.

The edge weights are set as follows. We associate with each edge three numbers \( w_0 \), \( w_1 \), and \( w_c \), where the first two correspond with the endpoints and the latter with the centre position. For new edges, \( w_0^r \), \( w_1^r \), and \( w_c^r \). If an edge \( e \) is split into two edges \( e' \) and \( e'' \), we use linear interpolation for the new values.

As a result, the weights are highest close to the centre of original edges. Finally, we use \( w_c \) as the edge weight for the edges of the final mesh, plus a graticule weight \( w_1 \) with small values for \( W_1 \) and \( W_1 \) to select the aspect.

The resulting unfolded maps are, at first sight, somewhat surprising (Figure 5). One would expect to see interesting fractal shapes, however, at the second level of subdivision the gaps are already almost invisible (Figure 6). Indeed, the structure of the cuts is self-similar, however, for higher levels of subdivision and smaller triangles, the surface of the sphere quickly approaches a plane, which has Hausdorff dimension 2. Only when areas would be removed, such as the centre triangles in the Sierpinski triangle, a fractal shape would be obtained.

As a step aside, fractal surfaces and foldouts do not match well either. Unfolding, for instance, a recursively subdivided surface with displaced midpoints leads to a large number of fold-overs (Figure 7).

As another step aside, let us consider optimal mapping on Platonic solids. We consider a map optimal when the cuts do not cross continents. To find such mappings, we assign to each edge a weight proportional to the amount of land cut, computed by sampling the edges at a number of positions (here we used 25) and looking up if land or sea is covered in a texture map of the earth. Next, the map is unfolded using the standard method and the sum of weights of cut edges is determined. This procedure is repeated for a large number of orientations of the mesh, searching for a minimal value. We used a sequence of three rotations to vary the orientation of the mesh, and used steps of 1 degree per rotation. Results are shown in Figure 8.
Geographically-aligned

aligned with continents

continents and oceans separated

north-up, south-down

north-up, south-down, graticular mesh

[J. van Wijk, 2008]
Australia-centric

[J. van Wijk, 2008]
## Search Tasks

<table>
<thead>
<tr>
<th>Target known</th>
<th>Target unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location known</td>
<td>Location unknown</td>
</tr>
<tr>
<td><strong>Lookup</strong></td>
<td><strong>Locate</strong></td>
</tr>
<tr>
<td><strong>Browse</strong></td>
<td><strong>Explore</strong></td>
</tr>
</tbody>
</table>

[Munzner (ill. Maguire), 2014]
Route Maps

Figure 1: Three route maps for the same route rendered by (left) a standard computer-mapping system, (middle) a person, and (right) LineDrive, our route map rendering system.

The standard computer-generated map is difficult to use because its large, constant scale factor causes the short roads to vanish and because it is cluttered with extraneous details such as city names, parks, and roads that are far away from the route. Both the handdrawn map and the LineDrive map exaggerate the lengths of the short roads to ensure their visibility while maintaining a simple, clean design that emphasizes the most essential information for following the route. Note that the handdrawn map was created without seeing either the standard computer-generated map or the LineDrive map. (Handdrawn map courtesy of Mia Trachinger.)

Abstract

Route maps, which depict a path from one location to another, have emerged as one of the most popular applications on the Web. Current computer-generated route maps, however, are often very difficult to use. In this paper we present a set of cartographic generalization techniques specifically designed to improve the usability of route maps. Our generalization techniques are based both on cognitive psychology research studying how route maps are used and on an analysis of the generalizations commonly found in handdrawn route maps. We describe algorithmic implementations of these generalization techniques within LineDrive, a real-time system for automatically designing and rendering route maps. Feedback from over 2200 users indicates that almost all believe LineDrive maps are preferable to using standard computer-generated route maps alone.

Keywords: Information Visualization, Non-Realistic Rendering, WWW Applications, Human Factors

1 Introduction

Route maps, which depict a path from one location to another, are one of the most common forms of graphic communication. Although creating a route map may seem to be a straightforward task, the underlying design of most route maps is quite complex. Mapmakers use a variety of cartographic generalization techniques including distortion, simplification, and abstraction to improve the clarity of the map and to emphasize the most important information [16, 21]. This type of generalization, performed either consciously or sub-consciously, is prevalent both in quickly sketched maps and in professionally designed route maps that appear in print advertisements, invitations, and subway schedules [25, 13].

Recently, route maps in the form of driving directions have become widely available through the Web. In contrast to handdesigned route maps, these computer-generated route maps are often more precise and contain more information. Yet these maps are more difficult to use. The main shortcoming of current systems for automatically generating route maps is that they do not distinguish between essential and extraneous information, and as a result, cannot apply the generalizations used in hand-designed maps to emphasize the information needed to follow the route.

Figure 1 shows several problems arising from the lack of differentiation between necessary and unnecessary information. The primary problem is that current computer-mapping systems maintain a constant scale factor for the entire map. For many routes, the lengths of roads can vary over several orders of magnitude, from tens of feet within a neighborhood to hundreds of miles along a highway. When a constant scale factor is used for these routes, it forces the shorter roads to shrink to a point and essentially vanish. This can be particularly problematic near the origin and destination of the route where many quick turns are often required to enter or exit a neighborhood. Even though precisely scaled roads might help navigators judge how far they must travel along a road, it is far more important that all roads and turning points are visible. Handdrawn maps make this distinction and exaggerate the lengths of shorter roads to ensure they are visible.

Another problem with computer-generated maps is that they are often cluttered with information irrelevant to navigation. This extraneous information, such as the names and locations of cities, parks, and roads far away from the route, often hides or masks information that is essential for following the route. The clutter makes the maps very difficult to read, especially while driving. Handdrawn maps usually include only the most essential information and are very simple and clean. This can be seen in figure 1 (middle) where even the shape of the roads has been distorted and simplified to improve the readability of the map. Furthermore, distorting [Agrawala & Stolte, 2001]
Locate
Adding Data

• Discrete: a value is associated with a specific position
  - Size
  - Color Hue
  - Charts

• Continuous: each spatial position has a value (fields)
  - Heatmap
  - Isolines
Discrete Categorical Attribute: Shape

[Acadia NP, National Park Service]
Discrete Categorical Attribute: Shape
Discrete Quantitative Attribute: Color Saturation
Discrete Quantitative Attribute: Size
Discrete Quantitative Attributes: Bar Chart

Railway Network Development and Bar Chart of Province Population in Turkey

[http://mis4gis.com/hgistr.org/]
Continuous Quantitative Attribute: Color Hue

[http://tampaseo.com/2012/02/websites-heat-mapping-users/]
Time as the attribute

[Map showing predicted tsunami wave heights with different time markers (1 hour, 3 hours, 6 hours, 9 hours, 12 hours, 15 hours, 18 hours, 21 hours)].

[NYTimes]
Isolines
Isolines

- Scalar fields:
  - value at each location
  - sampled on grids
- Isolines use **derived data** from the scalar field
  - Interpret field as representing continuous values
  - Derived data is **geometry**: new lines that represent the same attribute value

- Scalability: dozens of levels
- Other encodings?
Choropleth (Two Hues)

[M. Ericson, New York Times]
Choropleth Map

• Data: geographic geometry data & one quantitative attribute per region
• Tasks: trends, patterns, comparisons
• How: area marks from given geometry, color hue/saturation/luminance
• Scalability: thousands of regions

• Design choices:
  - Colormap
  - Region boundaries (level of summarization)
Choropleth (Two Hues)

[M. Ericson, New York Times]
Problem?

2008 Popular Vote

Obama: 68 million
McCain: 59 million

[M. Ericson, New York Times]
Problem?

2008 Popular Vote

- Obama: 68 million
- McCain: 59 million

Amount of red and blue shown on map

- Obama: 850,000 mi$^2$
- McCain: 2,150,000 mi$^2$

[M. Ericson, New York Times]
Adding Saturation

[Washington Post, 2018]
Aggregation: 2016 Election by Precinct

[Interactive Version, NYTimes]  
[R. Rohla and Washington Post, 2018]
Aggregation: 2016 Election by State

[D. Koop, DSC 530, Spring 2019]

[Washington Post, 2018]
Aggregation: 2016 Election by Country

[Image of a map with color coding for election results: Clinton +50-100, +15-50, +2.1-15, +0-2.1, Trump +0-2.1, +2.1-15, +15-50, +50-100.]

[Washington Post, 2018]
Area Marks and Color Hue & Saturation

Map by Matthew T. Campbell
Spatial Graphics and Analysis Lab
Department of Cartography and Geography
East Central University (Oklahoma)
Map Template courtesy of www.mymaps.com

Respondents through March 1, 2003
Size Encoding

[M. Ericson, New York Times]
Dasymetric Dot Density
2016 ELECTION MAP
Each figure represents 250,000 votes

Votes are distributed by state as accurately as possible while keeping national totals correct.
Location within each state is approximate.
Cartograms

US Presidential Election 2016
Results mapped at county level showing the candidate with the largest vote share in each area

Overall result:
Trump
62,979,636 votes (46.1%)
306 electoral votes (won 306 in the Electoral College)

Clinton
65,844,610 votes (48.2%)
232 electoral votes (received 232 in the Electoral College)

Other candidates
7,604,213 votes (5.7%)

Vote share of candidate with most votes

[Map of U.S. showing cartogram with vote shares for Trump and Clinton, and areas not voted for either candidate.]

Map by Benjamin Hennig
www.viewsfromtheworld.net

[Text on image: Not voted for either Trump or Clinton.]

Reference map
Cartograms

• Data: geographic geometry data & two quantitative attributes per region (one part-of-whole)

• Derived data: new geometry derived from the part-of-whole attribute

• Tasks: trends, patterns, comparisons, part-of-whole

• How: area marks from derived geometry, color hue/saturation/luminance

• Scalability: thousands of regions

• Design choices:
  - Colormap
  - Geometric deformation
Rectangular Cartogram

1860 Presidential Election

<table>
<thead>
<tr>
<th>CANDIDATE</th>
<th>PARTY</th>
<th>ELECTORAL VOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abraham Lincoln</td>
<td>Republican</td>
<td>180</td>
</tr>
<tr>
<td>John C. Breckinridge</td>
<td>Southern Democratic</td>
<td>72</td>
</tr>
<tr>
<td>John Bell</td>
<td>Constitutional Union</td>
<td>39</td>
</tr>
<tr>
<td>Stephen A. Douglas</td>
<td>Northern Democratic</td>
<td>12</td>
</tr>
</tbody>
</table>

Each state is sized by the number of electoral votes it had in 1860.

[New York Times]
Hexagonal Cartogram

District totals by category

189 17 6 14 21 44 144

MAJORITY

= one district

[FiveThirtyEight, 2018]
Non-Contiguous Cartogram

[M. Bostock, 2012]
World Cartograms
World Population

[M. Newman, 2009]
World Energy Consumption

[M. Newman, 2009]
House Races: Map?

House Race Ratings by the Cook Political Report

- 183 solid Democratic seats
- 145 solid Republican seats
- 7 vacancies
- 218 seats needed for House majority

[New York Times, 2018]
House Races: Cartogram?

[D. Koop, DSC 530, Spring 2019]
Maps Aren't Always Best: Close House Races

12 Lean Democratic
- AZ-02 Open (McSally)
- CA-49 Open (Issa)
- CO-06 Coffman
- IA-01 Blum
- KS-03 Yoder
- MI-11 Open (Trott)
- MN-02 Lewis
- MN-03 Paulsen
- NV-03 Open (Rosen)
- NJ-11 Open (Frelinghuysen)
- PA-07 Vacant (formerly Dent)
- VA-10 Comstock

31 Tossups
- CA-10 Denham
- CA-25 Knight
- CA-39 Open (Royce)
- CA-45 Walters
- CA-48 Rohrabacher
- FL-26 Curbelo
- FL-27 Open (Ros-Lehtinen)
- IL-06 Roskam
- IL-12 Bost
- IA-03 Young
- KS-02 Open (Jenkins)
- KY-06 Barr
- ME-02 Poliquin
- MI-08 Bishop
- MN-01 Open (Walz)
- MN-08 Open (Nolan)
- NJ-03 MacArthur
- NJ-07 Lance
- NM-02 Open (Pearce)
- NY-19 Faso
- NY-22 Tenney
- NC-09 Open (Pittenger)
- NC-13 Budd
- OH-01 Chabot
- PA-01 Fitzpatrick
- TX-07 Culberson
- TX-32 Sessions
- UT-04 Love
- VA-02 Taylor
- VA-07 Brat
- WA-08 Open (Reichert)

25 Lean Republican
- AR-02 Hill
- CA-50 Hunter
- FL-15 Open (Ross)
- FL-16 Buchanan
- GA-06 Handel
- GA-07 Woodall
- IL-13 Davis
- IL-14 Hultgren
- MO-02 Wagner
- MT-AL Gianforte
- NE-02 Bacon
- NY-24 Katko
- NY-27 Collins
- NC-02 Holding
- OH-12 Balderson
- PA-10 Perry
- PA-16 Kelly
- SC-01 Open (Sanford)
- TX-23 Hurd
- TX-31 Carter
- VA-05 Open (Garrett)
- WA-03 Herrera Beutler
- WA-05 McMorris Rodgers
- WV-03 Vacant (formerly Jenkins)
- WI-01 Open (Ryan)

[New York Times, 2018]
Maps Aren't Always Best: Obama Targets

If President Obama were to win all of the states above this line, he would need an additional 17 electoral votes from states below it in order to win in 2012.

Circles are sized according to the number of electoral votes in 2012.

Percentage with bachelor’s degrees or higher

Margins of victory in 2008

D. Koop, DSC 530, Spring 2019
D3 Map Example