Measuring and Modeling the Feature Detection Threshold Functions of Colormaps

Presented by Jerry Yin
Which colourmap is the best at visualizing the data?
Paper contributions

- Paper type: evaluation
- Describes way to measure frequency-dependent discriminative power function of a colourmap
  - *Discriminative power*: ability to distinguish different colours
  - *Frequency-dependent*: more later
- Defines metric for “overall discriminative power” across entire range of a colourmap
Spatial frequency

- Discriminative power depends on spatial frequency
- Uniform colour spaces (UCS) intended to be *visually* uniform
  - Based on measurements between *large* patches of uniform colour
- Thus, uniform colour spaces may not actually appear uniform in high-frequency datavis contexts!

The bands visually disappear at different heights along the image.
Spatial frequency

- Discriminative power depends on spatial frequency
- Uniform colour spaces (UCS) intended to be *visually* uniform
  - Based on measurements between *large* patches of uniform colour
- Thus, uniform colour spaces may not actually appear uniform in high-frequency datavis contexts!

The bands visually disappear at different heights along the image.
Spatial frequency

- Discriminative power depends on spatial frequency
- Uniform colour spaces (UCS) intended to be *visually* uniform
  ○ Based on measurements between *large* patches of uniform colour
- Thus, uniform colour spaces may not actually appear uniform in high-frequency datavis contexts!

![Graph showing contrast sensitivity vs. spatial frequency]

- Uniform color spaces
- Many features in data visualization

- red-green
- yellow-blue
- luminance

Contrast sensitivity

Spatial frequency (cycles/deg)
Empirical study

- Paper devises empirical study for measuring discriminative power across multiple spatial frequencies
- Used 600×600px images
- For each column, participants click the area where the sinusoidal pattern disappears
- Tested nine colour sequences and three frequencies (10px, 15px, 45px)
  - For each sequence, tested 30 locations
### Tested colourmaps

<table>
<thead>
<tr>
<th>Colormap</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>Rainbow colormap. This version comes from Paraview software</td>
</tr>
<tr>
<td>CW</td>
<td>The Moreland cool-warm colormap</td>
</tr>
<tr>
<td>ECW</td>
<td>Extended cool-warm colormap from Samsel</td>
</tr>
<tr>
<td>BOD</td>
<td>Blue-orange-red divergent colormap from Samsel</td>
</tr>
<tr>
<td>GP</td>
<td>A uniform grey colormap</td>
</tr>
<tr>
<td>GR</td>
<td>Approximately equiluminous green-red colormap</td>
</tr>
<tr>
<td>BY</td>
<td>An approximately equiluminous yellow-blue colormap</td>
</tr>
<tr>
<td>VI</td>
<td>Viridis colormap prized for its uniformity</td>
</tr>
<tr>
<td>TH</td>
<td>A colormap sometimes used in thermal imaging</td>
</tr>
</tbody>
</table>
Results

VI viridis
GR green-red
GP greyscale
BY blue-ylw
RA rainbow
CW cool-warm
BOD blu-orang
ECW ext. cool-warm
TH thermal
Results

BY blue-yellow
GR green-red
CW cool-warm
VI viridis
GP greyscale
RA rainbow
BOD blue-orange
ECW extended cool-warm
TH thermal
Results

- Ran 2-way ANOVA
- Arcs indicate where differences not statistically significant
- Ran Tukey HSD test (another significance test), horizontal bars show cases where colourmaps were not significantly different
Which colourmap should I use?

- Despite having the highest discriminative power, the thermal colourmap is *confusing.*
Which colourmap should I use?

- Despite having the highest discriminative power, the thermal colourmap is confusing.
- Same also applies to divergent colourmaps, to some degree.
Reweighting CIELAB

- Discriminative power should correspond to distance traversed by colourmap in uniform colour space
- Paper describes simplistic way to reweight CIELAB space to take into account the measured values in the paper
  - Equal weight is given to the 10px, 15px, and 45px cases
(Own) critique

- Instead of reweighting CIELAB in a way that is good for all datasets, maybe it would be better to collect data for many frequencies and reweight based on data that is currently being plotted.
- Minimum discriminative power may be a better metric than mean discriminative power.
- Outliers were manually removed.
- Sample size a bit small: only 21 - 35 participants per colourmap.