Requirements (WHAT)

Annotations can be:
- **statistical**
  - word length
- **syntactic**
  - parts-of-speech
- **semantic**
  - sentiment tags
- **structural**
  - page margins
- **domain-specific**
  - proper names

- **categorical**
- **ordered**
- **quantitative**
- **boolean**

- **of any textual scope**
- **overlapping**

Pop-out is key

Characters/words are marks that are fairly densely packed and regularly spaced, and that already make use of some visual channels.

To make highlighting detectable, need to maximize pop-out.

Common highlighting techniques (HOW)

- Each technique can also encode boolean features (scope of paper limited to this consideration)
- 9 techniques used in user studies

3 User Studies

- Performed using Amazon Mechanical Turk
- Analysis techniques: ANOVA and Tukey HSD
- Unwanted variation
  - Individual difference: normalized each participant’s responses with respect to their performance range
  - Learning curve: discarded first trials in first study, added training trials in others
  - Fatigue effects: not observed

Study 1: Ranking Techniques

- **Goal**: rank techniques with respect to pop-out
- 673 words, 20 randomly highlighted
  - Find as many highlighted words as possible within a time limit
- 45 participants
- 3 trials per technique (27 trials total) per participant
  - trials ordered randomly

**Study 1 - results**

<table>
<thead>
<tr>
<th>Technique</th>
<th>Perf. Rank</th>
<th>Mean/Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>font size</td>
<td>A</td>
<td>0.80 (0.12)</td>
</tr>
<tr>
<td>border</td>
<td>B</td>
<td>0.84 (0.18)</td>
</tr>
<tr>
<td>background</td>
<td>A</td>
<td>0.79 (0.14)</td>
</tr>
<tr>
<td>red</td>
<td>A</td>
<td>0.70 (0.10)</td>
</tr>
<tr>
<td>bold</td>
<td>C</td>
<td>0.74 (0.15)</td>
</tr>
<tr>
<td>shadow</td>
<td>B</td>
<td>0.65 (0.19)</td>
</tr>
<tr>
<td>underlined</td>
<td>D</td>
<td>0.50 (0.23)</td>
</tr>
<tr>
<td>spacing</td>
<td>E</td>
<td>0.15 (0.14)</td>
</tr>
<tr>
<td>italic</td>
<td>F</td>
<td>0.15 (0.14)</td>
</tr>
</tbody>
</table>

**Study 1 - discussion**

Possible explanations of strong results:

- Increased font size: sticks out from cap line, fill white space
- Border: makes the target appear larger
- Colour: strong pop-out effect
  - background may outperform text colour because coloured area is larger

Possible explanations of weak results:

- Letter spacing: already a normal feature of text
- Italics: slanted character features already found in text

Study 2: Search with Distractor

- **Goal**: determine how different techniques (A,B) interfere when used in the same text
  - Is relative strength of techniques a factor?
- 20 highlighted words for each of A, B, A+B
  - must choose words highlighted only with A
- 30 participants
- All pairs of techniques tried (72 trials total) per participant

**Study 2 - results**

<table>
<thead>
<tr>
<th>distractor technique</th>
<th>Perf. Rank</th>
<th>Mean/Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>font size</td>
<td>A</td>
<td>0.79 (0.13)</td>
</tr>
<tr>
<td>border</td>
<td>B</td>
<td>0.70 (0.10)</td>
</tr>
<tr>
<td>background</td>
<td>A</td>
<td>0.67 (0.11)</td>
</tr>
<tr>
<td>red</td>
<td>A</td>
<td>0.65 (0.13)</td>
</tr>
<tr>
<td>bold</td>
<td>C</td>
<td>0.61 (0.12)</td>
</tr>
<tr>
<td>shadow</td>
<td>B</td>
<td>0.50 (0.19)</td>
</tr>
<tr>
<td>underlined</td>
<td>D</td>
<td>0.50 (0.23)</td>
</tr>
<tr>
<td>spacing</td>
<td>E</td>
<td>0.30 (0.15)</td>
</tr>
<tr>
<td>italic</td>
<td>F</td>
<td>0.30 (0.15)</td>
</tr>
</tbody>
</table>

Study 2 results: weaker techniques

did not expect improvements

Study 2 results: stronger techniques

Fig. 5: Effect of distractor thickness on performance. The dashed line shows the baseline performance without any distractor. The filled circles represent the mean performance with a thickness of 1.5 mm, and the open circles represent the performance with 1.0 mm thickness. The error bars denote the 95% confidence intervals. (From Strobelt et al., 2016, figure 4.)
Study 3: Visual Conjunctive Search

- **Goal:** How strong is a combination of techniques (A, B) compared to each alone?
- 20 highlighted words for each of A, B, A+B — must choose only A+B
- 24 participants
- All pairs of techniques tried (36 trials total) per participant

**Results**

- Only underlined + spacing showed improvement over both individually

Only one feature

Choose a technique with strong pop-out

**Examples:**
- Font size
- Borders
- Yellow background

Same visibility; conjunction unimportant

Choose techniques with strong pop-out that do not significantly interfere with each other

**Examples:**
- Bold + yellow background
- Border + red
- Font size + yellow background
- Font size + border

Conjunction of features more important than each individually

Choose techniques that scored high in visual conjunction test

**Examples:**
- Border + red
- Font size + red
- Font size + yellow background

One feature significantly more important than the other

Choose techniques such that one has significantly higher pop-out

**Examples:**
- Yellow background + spacing
- Font size + underlined
- Border + italics

Guidelines

**Scenarios:**
- Only one feature should be highlighted
- Both features should have the same visibility; conjunctive visual search is not important
- Conjunction of features is more important than each individually
- One feature is significantly more important than the other
- Both features should have the same visibility; their conjunction should be easy to see

**Comments/Critiques**

- The guidelines for some scenarios are **very** similar, and multiple examples cover multiple scenarios
  - 3 studies for 5 scenarios
  - Some scenario refactoring would not be amiss
- I would have liked to see a larger scope
  - The authors don’t misrepresent the scope
  - A larger scope would be a lot more work
  - **BUT** a larger set of matrices might reveal more clusters to fit the scenarios better

Discussion/Future Work

Increase **scope**

- Combinations of more than two techniques
- Include more techniques (e.g., different colour combinations
- Include categorical/ordered/quantitative data
- Include tasks that require context/analysis
- Consider overlay visualizations

Comments/Critiques

- I would have liked to see a statement of expected results, based on existing understanding of marks and channels

Questions?