Evaluating photo aesthetics using machine learning

Domen Pogačnik, Robert Ravnik, Narvika Bovcon, Franc Solina

Ljubljana, 2012
Motivation

- Low cost & ease of manipulation
- Social networks

[http://www.quora.com/How-many-photos-are-uploaded-to-Facebook-each-day](http://www.quora.com/How-many-photos-are-uploaded-to-Facebook-each-day)
Simplicity

- Clear center of interest (subject)
- No or little clutter
Composition

- Compositional rules (rule of thirds, golden section, rabatment, ...)

![Composition example](image-url)
Color selection

- Unusual lighting condition
- As a design technique (color schemes)
Our approach

- Characteristics of HQ photos → calculable features
- Features → machine learning
- Distinguish between low and high quality photos
- Subject identification
- 74 features (color selection, composition, simplicity, visual complexity)
Color selection

- Hues (unique hues count, peaks)
- Avg saturation, avg brightness
- Color schemes

Hue histogram calculated on 30 bins
Composition

- Rule of thirds, golden section, rabatment
- Different models for evaluation
Simplicity & visual complexity

- Simplicity
  - Subject & background comparison

- Visual complexity:
  - Image edges
  - Image compression
Flickr dataset

- Recently uploaded, picks of the day
- Evaluated by experienced photographers (color scheme, composition, overall)
- Subject identified manually
- 258 photos (each with at least 3 evaluations)
Flickr experiment

- Overall score (3 class attribute: low, average, high)
- Excluded photos rated as average 258 → 114 photos
- Used best 28 features (reliefF metric)
- 95% CA with SVM classifier (10-fold cross validation)

<table>
<thead>
<tr>
<th>Feat. used</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>28</th>
<th>73</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA [%]</td>
<td>80.2</td>
<td>91.3</td>
<td>94.7</td>
<td>95.3</td>
<td>95.3</td>
</tr>
</tbody>
</table>
DPChallenge dataset (portraits)

- Manual subject identification = time consuming
- Face detector
- Photos part of photographic contests
DPChallenge experiment

- Ratings range: 1..10 (5.55 avg)
- < 4.5 (low quality), > 6.5 (high quality)
- 1048 photos (at least 100 evaluations)
- Used best 41 features (reliefF)
- 75% CA with SVM classifier (10-fold cross validation)

<table>
<thead>
<tr>
<th>Feat. used</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>41</th>
<th>71</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA [%]</td>
<td>69.8</td>
<td>71.1</td>
<td>72.1</td>
<td>72.4</td>
<td>74.8</td>
<td>73.4</td>
</tr>
</tbody>
</table>
Feature interpretation

- **Unique hues**
  - Up to 4 preferred (DP)
  - 15 (max), 3 (median)

- **Aspect ratio**
  - 0.8 .. 1.2 preferred
  - 0.47 (min),
  - 2.83 (max),
  - 1.0 (median)
Feature interpretation (II)

- Composition
  - Rule of thirds, golden section
  - Subject center distance to nearest vertical line
  - Distances up to 11% preferred (rule of thirds)
  - Up to 15% for golden section
Feature interpretation (III)

- Subject distance from image borders
  - Not desirable to be near image border
  - > 38% from left border, 23%-86% bottom (DP)

- Detected image edges
  - Higher number more appealing
  - More distinguishable subject

- Spatial distribution of detected edges
  - Distance from image center & borders
Feature interpretation (IV)

- Visual complexity
  - Detected image edges
  - Compressed image file size
    - > 101kB file size preferred (340k pixels)
    - 11kB (min), 388 (max), 101kB (median) - DPChallenge
Conclusion

- Characteristics of HQ photos → calculable features
- Features → machine learning
- Distinguish between low and high quality photos
- 95% CA (Flickr), 75% CA (DPChallenge)
- Different types of photos (landscape, portrait, ...)
- Automatic subject recognition