Image search through browsing using $\text{NN}^k$ networks

Daniel Heesch, Marcus Pickering, Stefan Rüger, Alexei Yavlinsky

TRECVID 2003
Overview

• Image and Collection Preprocessing
• Search and Relevance Feedback
• Temporal Browsing and $\text{NN}^k$ Browsing
• TVID Results
Preprocessing

- Use only common keyframes + LIMSI transcript
- Removal of bottom 51 lines
11 Primitive Features

• 4 Colour
  – global HSV, centre HSV, marginal RGB colour moments, colour structure descriptor

• 2 Structure
  – convolution map features on grey image

• 3 Texture
  – simple features on image tiles

• 1 Annotation
  – Bag of stemmed-words (tf-idf)

• 1 Localisation
  – Thumbnail of grey image
44x27 Thumbnail: Ad detection

- average pixel difference between two thumbnails
Distance of topic Q to image i given feature f

- \( \text{dist}_f: \) Manhatten
- KNN distance
  - positive examples (set Q)
  - negative examples (set N, random)

\[
d_f(Q, i) = \frac{\sum_{n \in N} (\text{dist}_f(n, i) + \varepsilon)^{-1}}{\sum_{q \in Q} (\text{dist}_f(q, i) + \varepsilon)^{-1} + \varepsilon}
\]
Fusion of features

- Convex combination

\[ D^w(Q, i) = \sum_f w_f d_f(Q, i) \]

- \( w \) is the “plasticity” of our retrieval system
Relevance Feedback

LAST NIGHT THERE DOZENS OF SMALL PLACES MERGED INTO ONE AND THE TINDER DRY BRUSH
Relevance Feedback

- Minimize

\[ \sum_{i} (D_{\text{user}}(Q, i) - D^{w}(Q, i))^2 \]

with respect to \( w \) and convexity constraint.
Browsing

- Hierarchical (not yet)
- In ranked list (not shown)
- Temporal
- Lateral
Temporal Browsing

- Movement along a sequence of shots
Temporal Browsing

- Movement along a sequence of shots
Temporal Browsing

- Movement along a sequence of shots
Temporal Browsing

- Movement along a sequence of shots

- Q: Add to query panel
- A: Add to assembly panel
Assembly panel
Pruning Panel
Lateral Browsing

• Images as vertices in a directed graph
• Instantiate arc \((i,j)\) iff there is a feature combination \(w\) such that \(j\) is closest to \(i\)
• \(\text{NN}^k\) network
**NN^k Network construction**

- For each image
  - for each \( w \) determine nearest neighbour and compute corresponding proportion of weight space (= edge weight)
  - store adjacent images and edge weights
Sampling the weight space
Rationale

- exposure of semantic richness
- user decides which image meaning is the correct one
- network precomputed -> interactive
- supports search without query formulation
Properties

• small average distance between any two vertices (three nodes for 32,000 images)
• high clustering coefficient: an image’s neighbours are likely to be neighbours themselves
• vertex degrees follow power-law distribution

-> scale-free small-world graph
Browsing interface

- Initial display:
  - query-by-example retrieval result
  - OR
  - high connectivity nodes (hubs)

- Clicking on an image moves it into the center and displays all adjacent nodes in the network
Both birds were graded from IN DANGER to THREATENED status of a few years ago later this year they could be removed from the
BOTH BIRDS WERE GRADED FROM IN DANGER TO THREATENED STATUS OF A FEW YEARS AGO LATER THIS YEAR THEY COULD BE REMOVED FROM THE
CLINIC PLANS TO TEST THE HEART SAVER ON HUMANS NEXT YEAR IF SUCCESSFUL FIRST OF ALL IT'S
BEING VERY Frustrated in years past that they never of mass produce the concept cars we so
SERENE SETTING SHELTER FROM THE WORLD
PAKISTAN TESTED its own medium range missiles both sides
MOST MAJOR INDIAN CITIES WITHIN RANGE OF THAT MISSILE AND
Observations

• Browsing can help to explore visual similarity
• Some task are impossible with browsing alone: find video shots with Senator Mark Sounder
• Browsing can be a fun activity
## Interactive runs

<table>
<thead>
<tr>
<th>Runs</th>
<th>Search</th>
<th>Relevance Feedback</th>
<th>Browsing</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>II</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>✔</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experimental design

- 4 subjects, 4 runs -> square lattice design

<table>
<thead>
<tr>
<th></th>
<th>T1-6</th>
<th>T7-12</th>
<th>T13-18</th>
<th>T19-25</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>I</td>
<td>II</td>
<td>III</td>
<td>IV</td>
</tr>
<tr>
<td>S2</td>
<td>IV</td>
<td>I</td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>S3</td>
<td>III</td>
<td>IV</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>S4</td>
<td>II</td>
<td>III</td>
<td>IV</td>
<td>I</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>MAP</th>
<th>RANK (out of 36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>S + RF + B</td>
<td>0.26</td>
<td>5</td>
</tr>
<tr>
<td>S + RF</td>
<td>0.26</td>
<td>4</td>
</tr>
<tr>
<td>S + B</td>
<td>0.23</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>0.13</td>
<td>27</td>
</tr>
</tbody>
</table>
Conclusions

• Competitive system: Three out of four runs among the top 8 (of 36)
• “Search by browsing“ a viable alternative to traditional search by example for visual topics