Multi-Facet Exploration of Image Collections with an Adaptive Multi-Focus Zoomable Interface

Sebastian Stober,
Christian Hentschel & Andreas Nürnberg
Outline

- Motivation
- Multi-Focus Visualization
  - The Galaxy Metaphor
  - The Problem: “Projection Errors”
  - The Adaptive SpringLens Approach
  - Demo
  - Evaluation (not in the paper!)
- Multi-Facet Similarity
  - Adaptable Music Similarity
  - Behind the Scenes
  - Automatic Similarity Adaptation (not in the paper!)
- Conclusions
- Appendix: Further Publications
How can we support users that cannot explicitly state what they are looking for?

Example:
Looking for a banner photo for a conference in Barcelona in a photo collection.
Motivation

- exploratory search problem
  - i.e. no query!!!
  - I just know, it is the right thing when I found it ;-) 

How can my computer help me?

- strategy:
  - give a general overview of what is available
  - let the user decide what region to explore deeper
  - provide guidance during exploration
The Galaxy Metaphor

- get general overview of what is available:
  - photos as stars
  - brightness = importance (manual rating)
  - some photos as thumbnails for orientation
    - 4 filter modes:
The Galaxy Metaphor

- get general overview
  - photos as stars
  - brightness = importance (manual rating)
  - some photos as thumbnails for orientation
  - 4 filter modes:
The Galaxy Metaphor

- get general overview
  - photos as stars
  - brightness = importance (manual rating)
  - some photos as thumbnails for orientation

4 filter modes:
The Galaxy Metaphor

- get general overview
  - photos as stars
  - brightness = importance (manual rating)
  - some photos as thumbnails for orientation

4 filter modes:

- sparse
The Galaxy Metaphor

- get general overview
  - photos as stars
  - brightness = importance (manual rating)
  - some photos as thumbnails for orientation

- 4 filter modes:
The Galaxy Metaphor

- get **general overview** of what is available:
  - photos as stars
  - brightness = importance (manual rating)
  - some photos as thumbnails for orientation
    - 4 filter modes:

- neighbors are similar
  - ... well, mostly (except for “projection errors”)
“Projection Errors”

- cause:
  - dimensionality reduction during projection:

  high-dimensional feature space

2D display

- similar
- dissimilar
General Idea

temporarily fix / highlight
the neighborhood in focus
Multi-Focus Adaptive SpringLens Distortion

- **primary fisheye lens**
  - enlarges region of interest
  - more space for details
  - preserves context

- **secondary fisheye lenses:**
  - extends focus to nearest neighbors
  - wormholes become evident
  - neighbors come closer together
Focus-Adaptive Visualization

strategy implementation:

- give a general overview of what is available
  - projection & filtering (galaxy metaphor)

- let the user decide what region to explore deeper
  - (primary) focus

- provide guidance during exploration
  - fisheye (preserves context)
  - secondary focus (wormhole metaphor)
Evaluation: Setting (not in paper!)

- retrieval task: finding photos for given topics
  - photos unknown to users (no bias)
- 4 collections (each 350 images and 5 non-overlapping topics)
- different exploration interface for each collection
  1. introduction & training
  2/3 only panning & zooming
  3/2 only Adaptive SpringLens (fisheye)
  4. combined interface

- 30 participants
  - 40% female
  - 19-32 years old
  - 30-60’ each
- recorded gaze (eye-tracker), webcam, audio comment

(submitted for NordiCHI)
Evaluation: PhotoGalaxy - Demo
Evaluation: Results (not in paper!)

usability: comparison with traditional panning & zooming:

- Helpfulness
- Simplicity
- Intuitivity

usefulness: percentage of annotated images by focus region

<table>
<thead>
<tr>
<th>focus region</th>
<th>primary</th>
<th>ext. primary</th>
<th>secondary</th>
<th>none</th>
</tr>
</thead>
<tbody>
<tr>
<td>same topic</td>
<td>37.75</td>
<td>4.27</td>
<td>30.74</td>
<td>4.38</td>
</tr>
<tr>
<td>other topic</td>
<td></td>
<td>4.49</td>
<td>13.24</td>
<td>2.08</td>
</tr>
<tr>
<td>no focus</td>
<td></td>
<td></td>
<td>30.74</td>
<td>2.08</td>
</tr>
<tr>
<td>total</td>
<td>37.75</td>
<td>8.75</td>
<td>43.98</td>
<td>9.52</td>
</tr>
</tbody>
</table>

(submitted for NordiCHI)
Motivation

How can we support users with different views on image similarity?
Adaptable Similarity

- support multiple facets of image similarity, e.g.:
  - EXIF time
  - GPS position
  - MPEG-7 ColorLayout, ScalableColor & EdgeHistogram

- each facet defined by a specific distance / similarity measure on a set of features

- let users interactively adapt aggregation of facet similarities

- real-time visualization update to make impact of changes immediately visible

- requires special data-structures
  (projection and neighbors cannot be pre-computed)
Behind the Scenes: Offline Preprocessing

- extract features
- pre-compute per-facet distances
- select representative landmark sample
- store only facet distances to landmarks
Behind the Scenes: Online Processing

projection facet distances

projection facet distance aggregator

projection

index

neighborhood facet distance aggregator

NN query

focus change

filtering & display

distortion

distortion

pipeline

pipeline
Similarity Adaptation (not in paper!)

- **motivation:**
  - try to have more relevant photos in secondary focus

- **idea:**
  - increase similarity of photos labeled (by user) as same topic

- **approach:**
  - derive similarity constraints from topic labels
  - apply metric learning approaches to learn facet weights

- **metric learning:**
  - interpretation as a binary classification problem
    [Cheng & Hüllermeier ‘06]
    - facet weights define the separating hyper plane
  - solved with (linear) support vector machine optimization
Change of the performance over the course of the simulation (iterations) for each topic of the three evaluation datasets (rows). Values are averaged over all possible seed images of the topic. Each plot shows the precision at rank 5 (left y-axis) for the initial weighting (baseline), the topic-specific adaptation and the query-specific adaptation. The solid line shows the average number of new relevant images in secondary focus (right y-axis) for the query-specific adaptation.
Conclusions

PhotoGalaxy

- uses galaxy metaphor
- alleviates “projection errors” with multi-focus distortion
- allows to adapt the underlying similarity measure
- fast enough for (near) real-time interaction
- usability and usefulness for exploration shown in study
- applicable in other domains – e.g. music, texts, …
- automatic similarity adaptation possible

future work

- better integration of panning and zooming with SpringLens
- extend GUI to allow object re-arrangement
August 17-18, 2010 – Johannes Kepler University Linz, Austria

Registration now open!

http://amr.dke-research.de/

Thank you for your attention!

demo videos: http://www.dke-research.de/aucoma/
contact: stober@ovgu.de
Appendix: Further Publications


Appendix: Facet Distance Definition

Definition. Given a set of features $F$, let $S$ be the space determined by the feature values for a set of objects $O$. A facet $f$ is defined by a facet distance measure $\delta_f$ on a subspace $S_f \subseteq S$ of the feature space, where $\delta_f$ satisfies the following conditions for any $x, y \in O$:

- $\delta(x, y) \geq 0$ and $\delta(x, y) = 0$ if and only if $x = y$
- $\delta(x, y) = \delta(y, x)$ (symmetry)

Optionally, $\delta$ is a distance metric if it additionally obeys the triangle inequality for any $x, y, z \in I$:

- $\delta(x, z) \leq \delta(x, y) + \delta(y, z)$ (triangle inequality)

Definition. The distance between objects $x, y \in O$ w.r.t. to the facets $f_1, \ldots, f_l$ and the individual facet distances $\delta_{f_1}(x, y), \ldots, \delta_{f_l}(x, y)$ is computed as weighted sum:

$$d(x, y) = \sum_{i=1}^{l} w_i \delta_{f_i}(x, y)$$