An Exploration of Visual Indexing: Problems and Prospects

Corinne Jörgensen
Peter Jörgensen

Florida State University
School of Information Studies
Ascendancy of Images

- 1994
  - VCR
  - Video Laser Disc
- 2004
  - Digital Camera
  - DVD
Early Work

- Markey, 1984,
- Rorvig et al, 1987, Library Trends
- Beck & Jones, 1989, American Antiquity
- Jörgensen, C., Jörgensen, P., Hogan, M., 1993., ICHIM ’93
Rorvig (1987)

- *First, image description is a time-consuming process, often requiring high-level personnel for accurate results* (Rorvig 1987)

- *Second, individual differences in image perception give rise to extraordinary idiosyncrasy in the assignment of image terms.* (Markey 1984)

- Variation in the explicitness with which attributes are described and definitions are applied;
- Differences among different analyst's perceptions; and
- Changes in an analyst's perceptions over time.

Hogan et al. (1991)

- Visual Thesaurus: *collections of images in a hypermedia environment (vs. print)*
- Thesaurus organization reveals
  - *pre-defined structures*
  - *relationships among items*
Features (Assumptions?)

- skirt hierarchical term associations
- lack of reliance on textual indexing
- independence from a particular taxonomy
- end-user orientation
Words are not enough

- Spatial relationships of objects
  - Position
  - Relative Size
- Colors
- Textures
Problems w/ Visual Thesauri

- *Print visual dictionaries* - 1991 analysis
- Top level: facet or theme “transportation”
- Objects obtain meaning by *context*
- Difficulty grasping underlying structure of visual taxonomies
- Cultural biases
Relationships

- Part-Whole
- Similarity (limited)
  - form
  - function
Hogan, et al. (1991)

- the system must be able to identify the “relevant” semantic parts of an image
- At that time, methods for computing visual similarity in early stages of development
Current Implementations

- *Visual Thesaurus*
- As tool for displaying text relationships
- As interface for image browsing and retrieval
Visual Thesaurus for Text
Advantages of Text

- Easily processed
  - ASCII – Unicode
- Compact – efficient
  - high compression – to a point
- Independent of object
Advantages of Images

- Richness
- Language (not culture) independence
- Right brain - Parallel processing
- Object differentiation
- Orientation independence
- Color/Texture/Space
- Easier to remember
Non-issues?

- Bandwidth
- Storage
- File Format
Primarily Text - Browsing
Alexandria Digital Library UCSB

Text Metadata
Text Browsing
WebSEEK

A Content-Based Image and Video Search and Catalog Tool for the Web

( press here to Browse all subjects )

Animals
birds, dinosaurs, monkeys, fishes

Cats
leopards, lions, kittens, cheetahs

Horror
godzilla, aliens, skeletons, monsters

Nature
sunsets, flowers, weather, mountains

Architecture
bridges, lighting, domes, heating

Celebrities
bullocks, aniston, monroe, keanu

Humour
simpsons, beavis, dilbert, ren/stimpy

Sports
baseball, basketball, swimming, hockey, olympics, surfing

Art
painting, illistr, sketching, ceramics, monet, van-gogh

Dogs
bulldogs, puppies, coyotes, wolves

Movies
batman, starwars, jurassic, python, blade runner, actors

Transportation
cars, planes, titanic, motorcycles, porsches

Astronomy
nasa, planets, eclipses, space

Food
apples, beer, pizza, cakes, fruits, veges

Music
beadles, metal, rock, cure, zeppelin, guitars

Travel
asia, europe, newyork, paris, australia, mexico

http://persia.ee.columbia.edu:8008/
<table>
<thead>
<tr>
<th>Animals</th>
<th>Geography</th>
<th>Photography</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture</td>
<td>Governments</td>
<td>Plants</td>
</tr>
<tr>
<td>Art</td>
<td>Graphics</td>
<td>Products</td>
</tr>
<tr>
<td>Astronomy</td>
<td>History</td>
<td>Recreation</td>
</tr>
<tr>
<td>Books</td>
<td>Holidays</td>
<td>Religion</td>
</tr>
<tr>
<td>Buildings</td>
<td>Horror</td>
<td>Science</td>
</tr>
<tr>
<td>Clothing</td>
<td>Household</td>
<td>Sports</td>
</tr>
<tr>
<td>Comedy</td>
<td>Humour</td>
<td>Television</td>
</tr>
<tr>
<td>Computer</td>
<td>Insects</td>
<td>Terrorism</td>
</tr>
<tr>
<td>Computers</td>
<td>Movies</td>
<td>Toys</td>
</tr>
<tr>
<td>Dance</td>
<td>Music</td>
<td>Transportation</td>
</tr>
<tr>
<td>Education</td>
<td>Musicals</td>
<td>Travel</td>
</tr>
<tr>
<td>Entertainment</td>
<td>Mythology</td>
<td>Universities</td>
</tr>
<tr>
<td>Environment</td>
<td>Nature</td>
<td>Videos</td>
</tr>
<tr>
<td>Fashion</td>
<td>Objects</td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>People</td>
<td></td>
</tr>
<tr>
<td>Games</td>
<td>Photographs</td>
<td></td>
</tr>
</tbody>
</table>

Jørgensen & Jørgensen, ASIST 2004
User-Supplied Image (QBE)
Amore

Find images visually similar to the gif/jpg image whose URL is:

```
ccer.theo.uu.nl/hiero/G/G1.gif
```

Retrieve images

For visual similarity search Shape is: Very Important
For visual similarity search Color is: Not Important

Search
Shape Recognition
*MARS*

- *User identification of shape of interest*
Shape Recognition

*MARS*
Shape Recognition

Blobworld

- Somewhere between “stuff” and “things” of Tang (1999)

- Automatic segmentation of images

http://elib.cs.berkeley.edu/photos/blobworld/
**Step 2:**

Adjust the weights below if you’d like, then click "Submit."

<table>
<thead>
<tr>
<th>How important is the selected region?</th>
<th>Not</th>
<th>Somewhat</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How important are the features of this region?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
</tr>
<tr>
<td>Texture</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Shape/Size</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How important is the background (everything outside the region)?</th>
<th>Not</th>
<th>Somewhat</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Blobworld
Shape Recognition

eVe

The Query Image is a photograph of a leopard. The Object Map shows how eVe interprets the 'objects' in the image. The results from the search are shown below.

The top 4 search results from the non-object based engine include two "relevant" results and two obviously incorrect results.
eVe's top four search results are all "relevant" images of large cats.

Automatic Segmentation of image

http://www.evisionglobal.com/tech/overview.html

Jørgensen & Jørgensen, ASIST 2004
Clustering/Browsing

http://elib.cs.berkeley.edu/

Jørgensen & Jørgensen, ASIST 2004
Q/RBE

retrieved 46 images (results 1~15)
randomize

http://persia.ee.columbia.edu:8008/

Jörgensen & Jörgensen, ASIST 2004
QBE - Vector

CANDID

public.lanl.gov/kelly/CANDID/
QBE - ?

ImageRETRO

<table>
<thead>
<tr>
<th>Interestate links</th>
<th>The Mili Diamonds</th>
<th>Download-site</th>
</tr>
</thead>
<tbody>
<tr>
<td>KABELPON</td>
<td>Helpdesk</td>
<td></td>
</tr>
<tr>
<td>THEEUROPE!</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

April 7th 1998

40 images 4 steps All loaded Fixed reduction to 25%

Reset Back History

Jørgensen & Jørgensen, ASIST 2004
QBE

ImageRover

Select relevant images to guide search.

Images found

<table>
<thead>
<tr>
<th>Images selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>159074</td>
</tr>
<tr>
<td>129349</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Images found</th>
</tr>
</thead>
<tbody>
<tr>
<td>129516</td>
</tr>
<tr>
<td>59117</td>
</tr>
<tr>
<td>159069</td>
</tr>
<tr>
<td>159277</td>
</tr>
<tr>
<td>69745</td>
</tr>
<tr>
<td>129612</td>
</tr>
<tr>
<td>129528</td>
</tr>
<tr>
<td>17307</td>
</tr>
<tr>
<td>59143</td>
</tr>
<tr>
<td>592</td>
</tr>
</tbody>
</table>

- Dump statistics
- Feature subspace: LSH/visual
- Relevance feedback: Adaptive

Number of returned images: 10

Search  Reset form
QBE + metadata

LCPD
QBE – user supplied image
MetaSEEK
Search by Color

FOCUS - UMass
Search by Color

Histogram Tweak Tool – WebSEEK

Color Histogram Tweak Tool
Java Applet

Adjust the histogram and make a new search.

Clear  Reload  Redraw  Search  Sticky Colors
Color & Texture

Texture Palettes

[Image of a software interface with texture palettes and selection options]
Color & Texture

C-Bird

Content-Based Search
Search by Texture and Colour

Result: (Images 1 through 11 from 11)
Color & Texture

C-Bird
Color & Texture

JACOB online demo. You are the guest number

Set the color/texture ratio.

Color/Texture (0=texture only, 1=color only): 0.5

Returned images: 4

Set the color weights and the texture coarseness of direct query.

Texture coarseness (1=coarse, 0= fine): 0.6
Sketch

DrawSearch
Sketch

QBIC

http://www.hermitagemuseum.org/fcgi-bin/db2www/qbicLayout.mac/qbic?selLang=English

Jørgensen & Jørgensen, ASIST 2004
Sketch

QBIC
Assisted Query Construction

- Users can’t draw
  - especially using a mouse
- Users can’t name colors
- Users can’t describe textures
QBE

Spatial Iconographic Palettes

Visual Index of all Uniform Polyhedra

Click on the image to obtain a high-resolution image and some geometrical information on the chosen polyhedron. An index with individual icons is also available.
Search by Color

Color Palettes
Spatial Iconographic ImageScape

Sketching interface

Results
Multiple Access Methods

VisualSEEK

- Color
- Texture
- Spatial/Icono
- Vectors
Multiple - Iconographical

Percevia

- GIS
- Shape
- Color
- etc.
Multiple - Iconographical

Percevia
Spatial Iconographic
Scifinder

http://info.cas.org/SCIFINDER/SCHOLAR/options/sm.html
Direct Manipulation

QuickTime™ and a Graphics decompressor are needed to see this picture.
Visual Thesauri for Image Browsing and Retrieval

- Many examples
- Limited implementations of similarity
- Visual similarity takes precedence over other types of relationships
- BUT what is order of precedence of visual primitives?
Visual Thesauri for Image Browsing and Retrieval

- *Users cannot choose relationships*
- *Users cannot choose levels of similarity*
Relationships are multivalent

“Apple” - multiple syntactic, inferential relationships
color (red), shape (round), texture (smooth, shiny, hard)
basic object type: food > fruit > apple > Macintosh
Relationships are multivalent

- **uses**: juice, pie, applesauce
- **activities**: eating, cooking, bobbing for
- **parts**: skin, seeds, stem
- **meaning**: temptation, poison
Begin

Cast
Corinne – Corinne Jörgensen
Peter – Peter Jörgensen

Producers
Corinne & Peter Jörgensen
Florida State University
School of Information Studies
(soon to be College of Information)
Tallahassee, FL

Production took place on location in
Tallahassee, FL
Providence, RI

© 2004 Florida State University