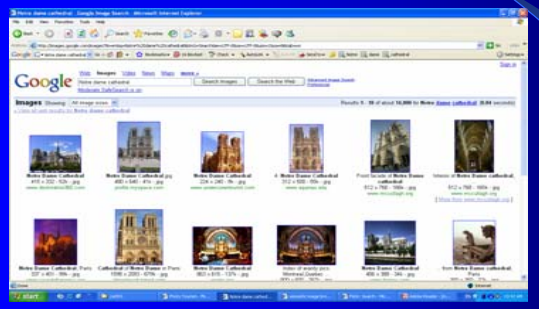
Image and Video Collections

Aili Li
CSE 591:Visual Analytics
Prof. Klaus Mueller

Background

- Digital photograph + internet sharing!

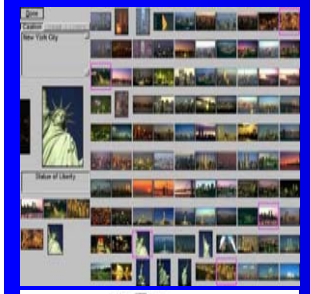


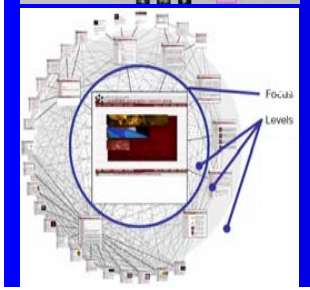
- Lacked methods for interactive browsing and searching!

Challenges

- How to organize unstructured collection of photographs?
  - By annotated with caption and keywords, the photos can be indexed and searched as textual documents
  - By measuring the similarities between images
  - Most annotation have been done by manually, automatic annotation is not used in image browsing
  - Lack methods to monitor and evaluate the automatic image analysis algorithm
- Image browsing is different from image querying

Structure of This Talk





1 2
3 4

Image Browsing by Similarity

Semantic Image Browser

MorieGraph

Photo Tourism

Image Browsing by Similarity

[Rodden et al, CHI01]

- CHI 2001

❖Does Organization by Similarity Assist Image Browsing?

Overview

[Rodden et al, CHI01]

- Construct arrangement of images based on:
  - Caption similarity
  - Visual similarity
- Performed two experiments
  - if arrangement of images benefits users
  - Users preferences

Interface of the Experiment Software

[Rodden et al, CHI01]

- 10x10 caption-based arrangement of 100 images of NY

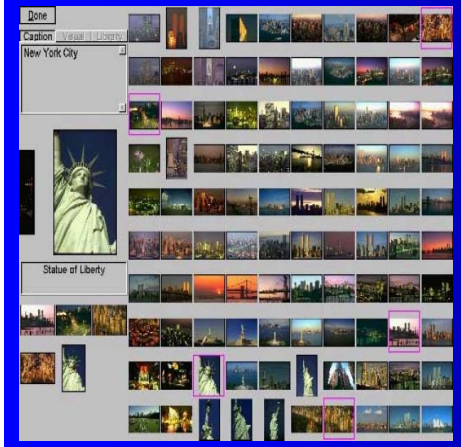


Image Arrangement

[Rodden et al, CHI01]

- Multi-Dimensional Scaling (MDS)
  - treats inter-object dissimilarities as distances in some high dimensional space, and then attempts to approximate them in a low dimensional output.

1. Create a similarity matrix
2. Find 2D configuration of points using MDS
3. Place thumbnail images at these points

Remove Overlaps

[Rodden et al, CHI01]

- Arrangement of 100 images of Kenya, based on visual similarity.
  - Left: MDS arrangement, Right: 12x12 grid (remove overlap)

Change Grid Size

[Rodden et al, CHI01]

- Arrangement of 100 images of Kenya, based on visual similarity.
  - Left: MDS arrangement, Right: 10x10 grid (maximize thumbnail size)

The Task

- The task was the same for both experiments
- Participants were given the following written description of it:

You have been asked to choose photographs to illustrate a set of "destination guide" articles for a new "independent travel" World Wide Web site. Each article will be an overview of a different location, and is to appear on a separate page. The articles have not yet been written, so all you have are short summaries to indicate the general impression that each will convey. You also have 100 photographs of each location, and your task is to choose 3 of the photos (to be used together) for each article. It is entirely up to you to decide on the criteria you use to make your selections—there are no "right" answers, and you are not bound by the given summaries.

Experiment I

[Rodden et al, CHI01]

- Goal:
  - If users find either of the similarity-based arrangement useful
  - If it was helpful to have both arrangements available
- Participants
  - 18 participants were all attendees of "infodesign 99"
- Apparatus
  - For each of the four places: New York, Paris, Kenya and Alaska
  - Created two 12x12 grid arrangements of 100 images, (visual similarity and caption similarity).

Results

Statement	*Agreement* 0	1	2	3	4	*Median score*
"the arrangement of photos by **caption** similarity was useful"	0	3	3	9	3	3
"the arrangement of photos by **image** similarity was useful"	3	4	3	6	2	2
"it was useful to have two different views of the same set of photos"	1	4	5	4	4	2

	Visual	Caption	no
Average spent time	37%	63%	
Heavily favored in	3 (ct<22%)	7 (ct > 85%)	8
First chose to search	14	40	

Experiment II

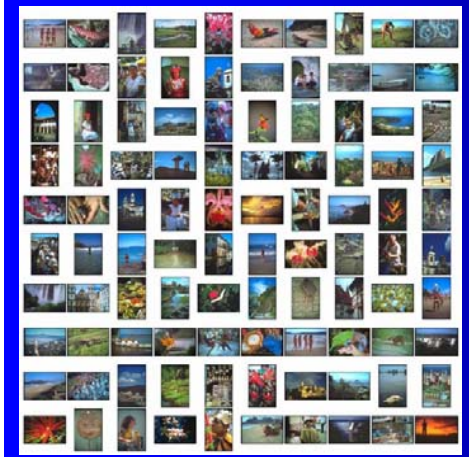
[Rodden et al, CHI01]

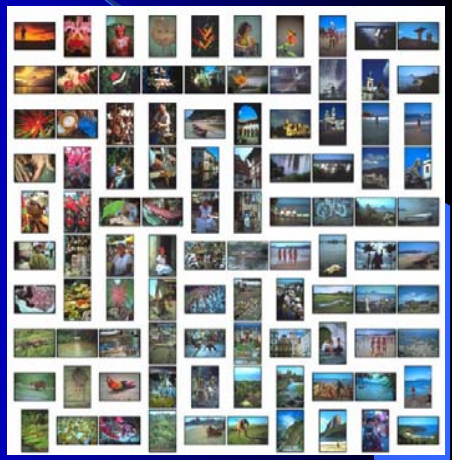
- Goal:
  - If users would prefer a similarity-base arrangement to a random arrangement
  - If a similarity-based arrangement would help users to carry out the given task more quickly
- Participants
  - 10 students of graphic design from Anglia Polytechnic Univ.
- Apparatus
  - Created two 10x10 grid arrangements of 100 images, (visual similarity and randomly) for each of 9 places.

Arrangements

[Rodden et al, CHI01]

- Two 10x10 grids of 100 images of Brazil





Left: randomly arranged

Right:visual similarity

Procedure

[Rodden et al, CHI01]

- Part one:
  - Group X {Denmark, Jamaica and Nepal}
  - Group Y {Death Valley, Ireland and Kenya}

1st	2nd	3rd	4th	5th	..	9th	10th
X v	X r	Y v	Y r	X r	..	Y v	Y r
Y r	Y v	X r	X v	Y v	..	X r	Y v

- Part two:
  - Group Z {Brazil, Canada and Yellowstone National Park}
  - The students had to choose one of them to use initially, and could then switch between the two views as they wished

Results

[Rodden et al, CHI01]

Statement	*Median score visual*	*random*	*p*
"it was enjoyable to use"	4	2.5	0.019
"it made it easy for me to find the photos I wanted"	5	3	0.007
"it made it easy to find photos that complemented each other"	4.5	3	0.133

	Visual	Random	both
Average spent time	66%	34%	
Heavily favored in	5 (vt>73%)	2 (vt<22%)	3
First chose to search	21	9	
Preference	4	1	5

Discussion

[Rodden et al, CHI01]

- Arranging by similarity does seem to be useful
- A caption-based arrangement helps to break down the set
  - Its usefulness is affected by the level of details of the captions
- A visual-based arrangement helps to divide the set into simple genres
  - Cause adjacent images to appear to "merge"
- Preferences may simply is due to individual differences

Semantic Image Browser

[Yang et al, VAST06]

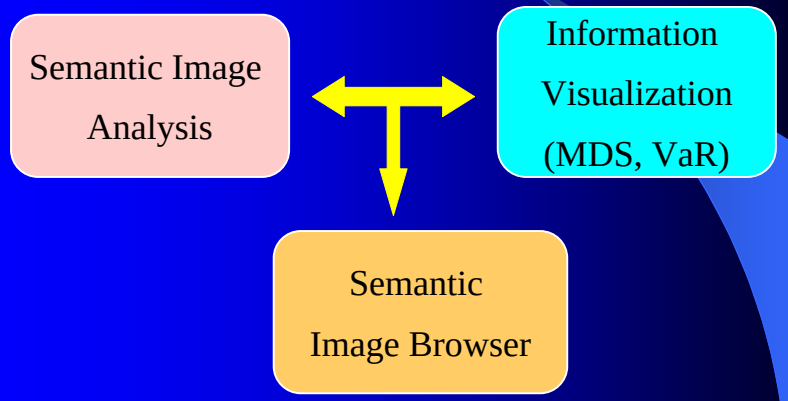
- VAST 2006

❖Semantic Image Browser:

Bridging Information Visualization with Automated Intelligent Image Analysis

Overview

[Yang et al, VAST06]



Automatic Annotation Engine

[Yang et al, VAST06]

- Concept-sensitive image content analysis technique
- Pre-defined salient object detect functions
  - Low-level automatic image segmentation
  - Classification by using Support Vector Machine

Figure:The semantic image classification results for the concept “sea world” with the salient objects, such as “sand field”

Image Overview

[Yang et al, VAST06]

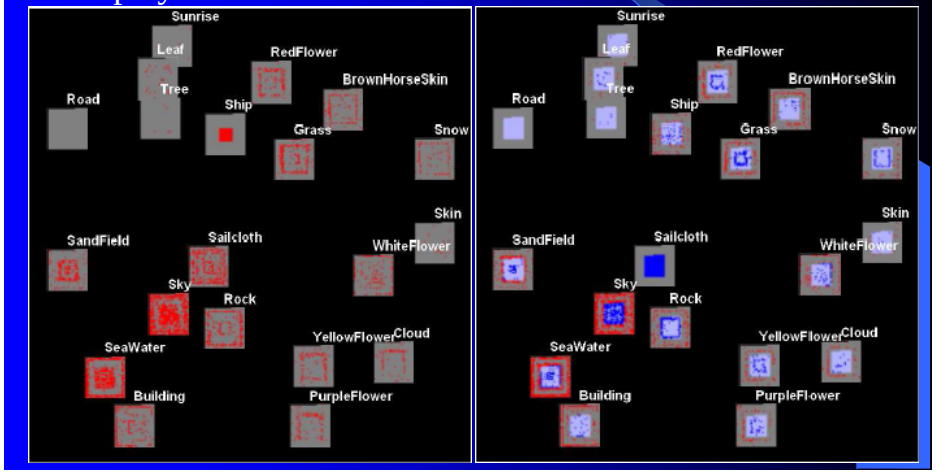
- MDS layout
- Interactions
  - Reordering
  - Dynamic scaling
  - Relocation
  - Distortion
  - Showing original image
  - Zooming and panning

Figure:An MDS image overview of Corel collection (1100 images)

Content Overview

[Yang et al, VAST06]

- Generated by visualizing the content dataset in VaR display



Interactions In Content Overview

[Yang et al, VAST06]

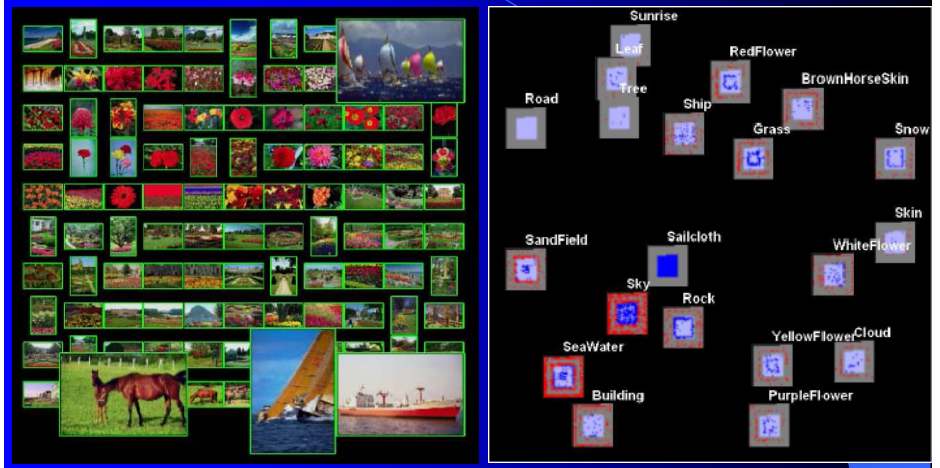
- VaR display provides a rich set of interaction tools
  - Clutter reduction
  - Reordering
  - Detection of correlations
- Combine search for images with (without) a certain content
  - Reduce a selected subset by requiring that the search results must /not contain a certain content
  - Increase a selected subset by adding images with /not a content

Image/Content Overview

- Each image collection has both image overview and content overview
- Selected image are highlighted in both overviews
- Use different views in different exploration stages
  - Select images by sample image via image overview
  - Select images by content via content overview
- VaR content overview is more stable than image MDS image overview, better for large image collections

Annotation Evaluation

[Yang et al, VAST06]



Left: Good annotation for "red flower", Right: Bad annotation for sailcloth

Monitor Annotation Process

[Yang et al, VAST06]

- There are some features more closely related to the annotation than others
- Many features used for classification are nearly identical
  - could be removed from the annotation process

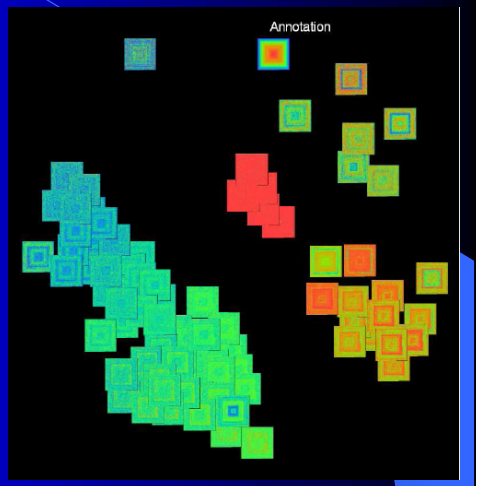


Figure: The feature-content dataset of the Corel collection (89 dimensions, 10,471 items) in the VaR display

Contributions

[Yang et al, VAST06]

- A novel semantic image browser
- An MDS image layout based on semantic similarities
- An VaR content display
- A rich set of interaction tools
  - Combinable searching by sample and searching by contents
- Visualizations and interactions
  - Allow image analysts to visually monitor, evaluate and improve their annotation process.

Photo Tourism Demo

[Snavely et al, SIGGRAPH06]

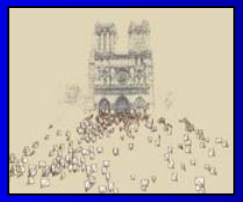
Photo Tourism Overview

[Snavely et al, SIGGRAPH06]

Input photographs

Scene reconstruction

Photo Explorer



Relative camera positions and orientations

Point cloud

Sparse correspondence

Scene Reconstruction

[Snavely et al, SIGGRAPH06]

- Automatically estimate
  - position, orientation, and focal length of cameras
  - 3D positions of feature points

Feature detection

Pairwise feature matching

Correspondence estimation

Incremental structure from motion

Reconstruction Performance

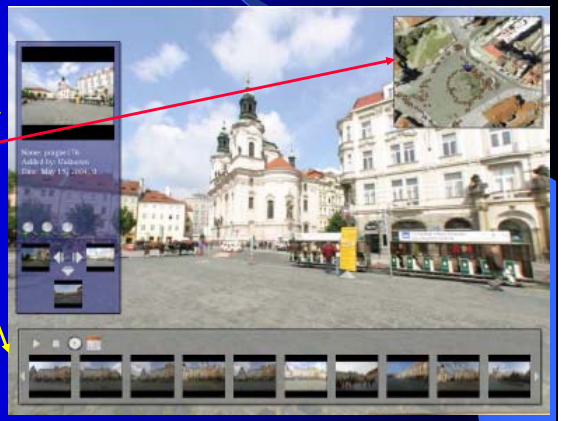
[Snavely et al, SIGGRAPH06]

- For photo sets from the Internet, 20% to 75% of the photos were registered
- Most unregistered photos belonged to different connected components
- Running time:
  - < 1 hour for 80 photos
  - > 1 week for 2600 photo

User Interface Layout

[Snavely et al, SIGGRAPH06]

- Main view
  - Information and search pane
  - Thumbnail pane
  - Map pane



Rendering The Scene

[Snavely et al, SIGGRAPH06]

Rendering Transitions

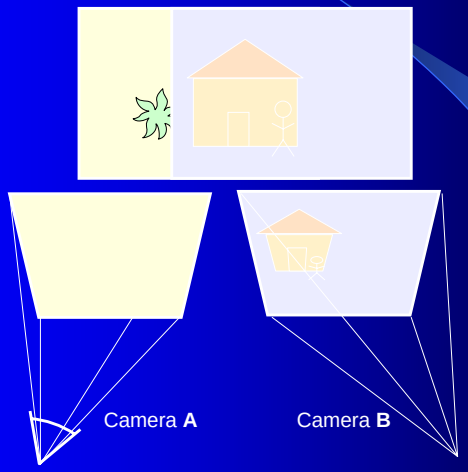
[Snavely et al, SIGGRAPH06]

Rendering Transitions

[Snavely et al, SIGGRAPH06]

Rendering Transitions

[Snavely et al, SIGGRAPH06]



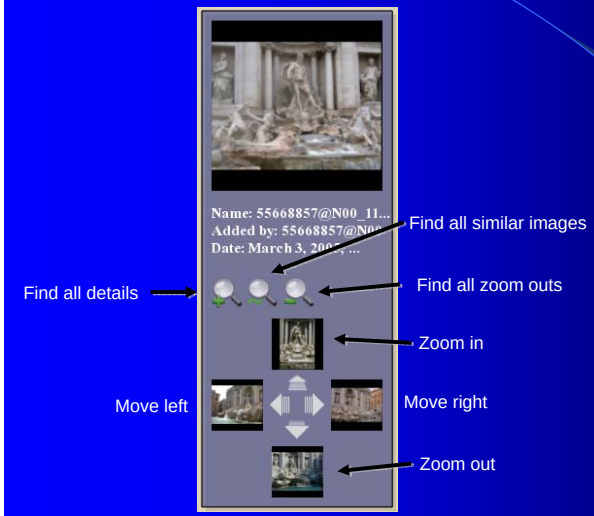
Navigation Controls

[Snavely et al, SIGGRAPH06]

- Free-flight navigation
- Object-based browsing
- Relation-based browsing
- Overhead map

Moving Between Related Views

[Snavely et al, SIGGRAPH06]



Object-based Browsing

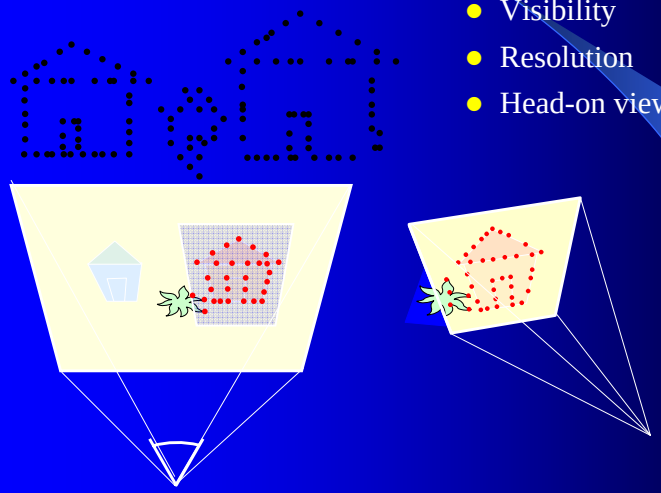
[Snavely et al, SIGGRAPH06]

The user drags a rectangle around Neptune in one photo, and the camera moves to a new, high-resolution photograph of the statue

Object-based Browsing

[Snavely et al, SIGGRAPH06]

- Visibility
- Resolution
- Head-on view



Annotation Transfer

[Snavely et al, SIGGRAPH06]

Three regions were annotated in the photograph on the left; the annotations were automatically transferred to the other photographs, a few were show on the right

Contributions

[Snavely et al, SIGGRAPH06]

- Automated system for registering photo collections in 3D for interactive exploration
- Structure from motion algorithm demonstrated on hundreds of photos from the Internet
- Photo exploration system combining new image-based rendering and photo navigation techniques

MorieGraphs

[Kelly and Ma, InfoVis 03]

- InfoVis 2003
  - **MorieGraphs:**

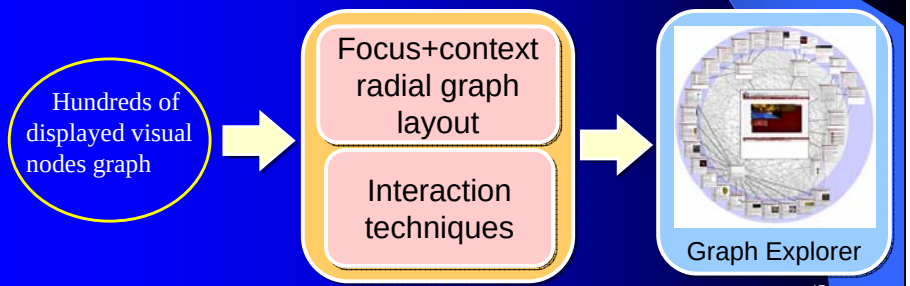
    Radial Focus+Context Visualization and Interaction for Graphs with Visual Nodes


(b) Level 1 Focused

Overview

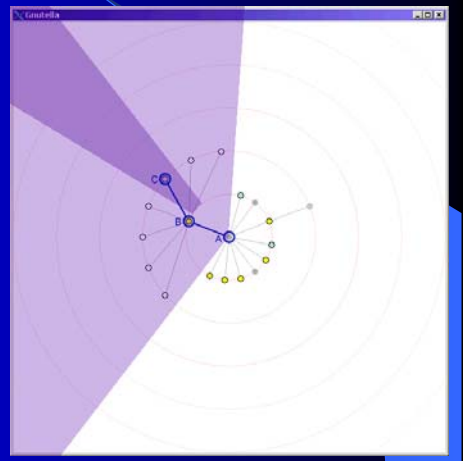
[Kelly and Ma, InfoVis 03]

- Create MorieGraph:
  - Choose focus, which will be the root of a spanning tree created from the visual node graph
  - Position this tree using the new focus+context radial layout
  - Render the tree and its visual node elements



Radial Focus+Context Graph Layout

- Radial layout
- Radial focus+context graph layout
  - Both radial spacing and node size decrease as the distance from the focus node increases

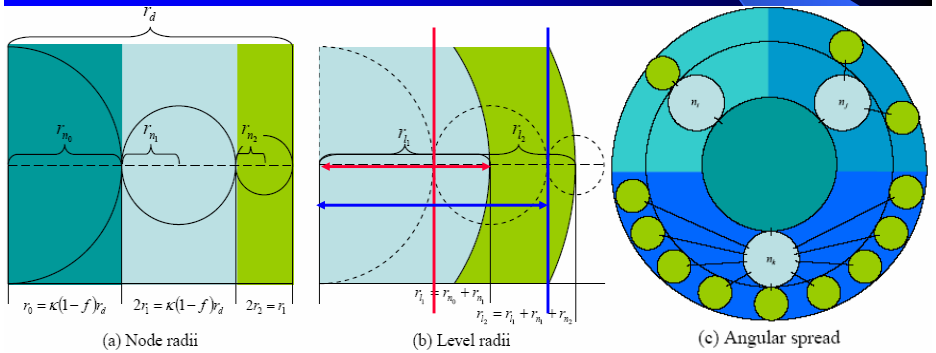


[courtesy of Ka-Ping Yee]

Layout Calculations

[Kelly and Ma, InfoVis 03]

- Focus strength f, display's radius r_d
- Node radii
- Level radii
- Angular spread



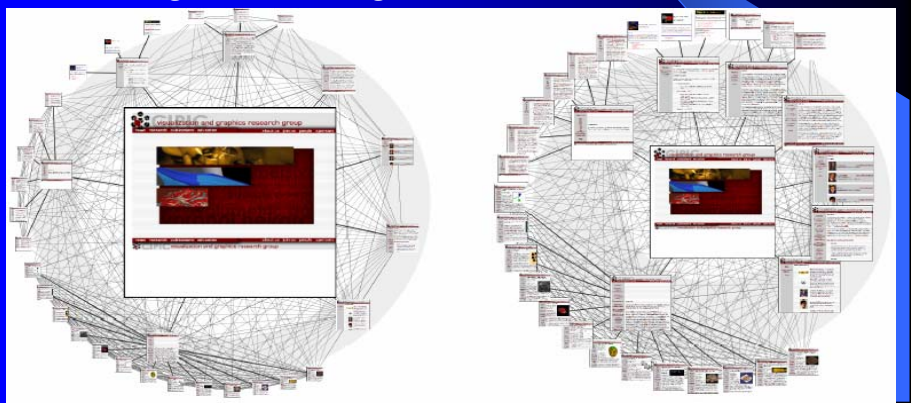
(a) Node radii (b) Level radii (c) Angular spread

Changing Focus Strength

[Kelly and Ma, InfoVis 03]

Left: The rest of graph is pushed to the periphery by increasing the focus strength

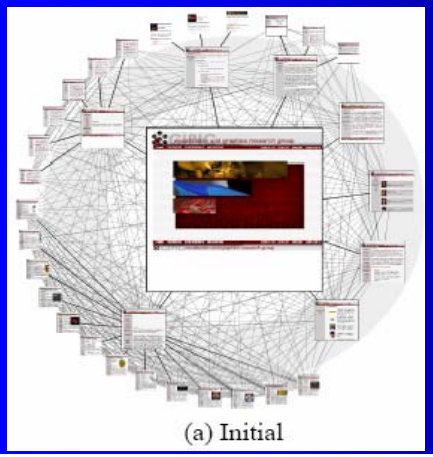
Right: Conversely, more room is allocated to the focus' children by decreasing the focus strength



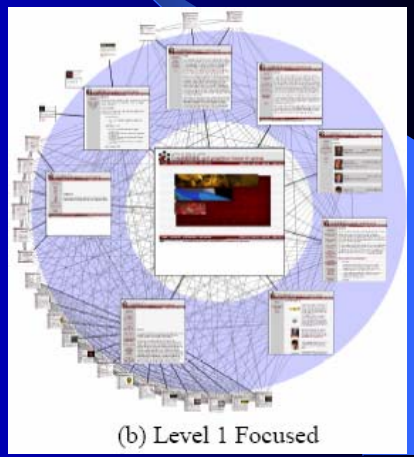
Level Highlighting

[Kelly and Ma, InfoVis 03]

- Highlight level 1



(a) Initial

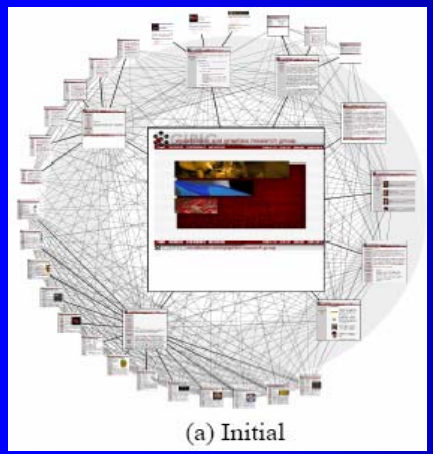


(b) Level 1 Focused

Level Highlighting

[Kelly and Ma, InfoVis 03]

- Highlight level 2

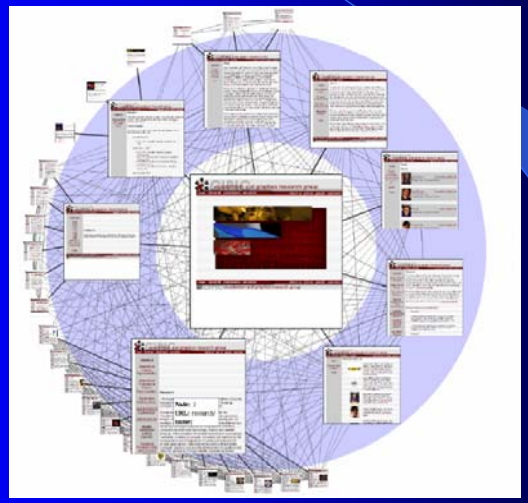


(a) Initial

(c) Level 2 Focused

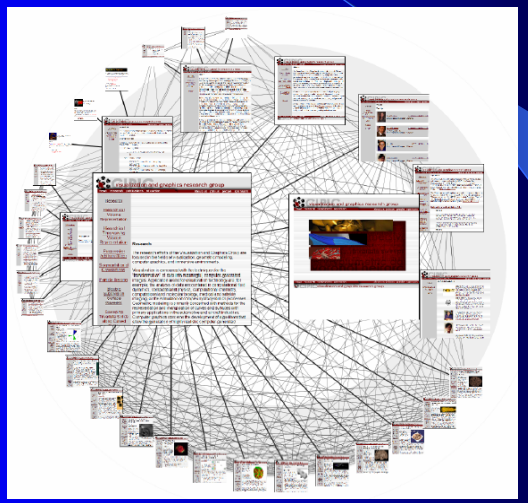
Animated Navigation

[Kelly and Ma, InfoVis 03]



Animated Navigation

[Kelly and Ma, InfoVis 03]



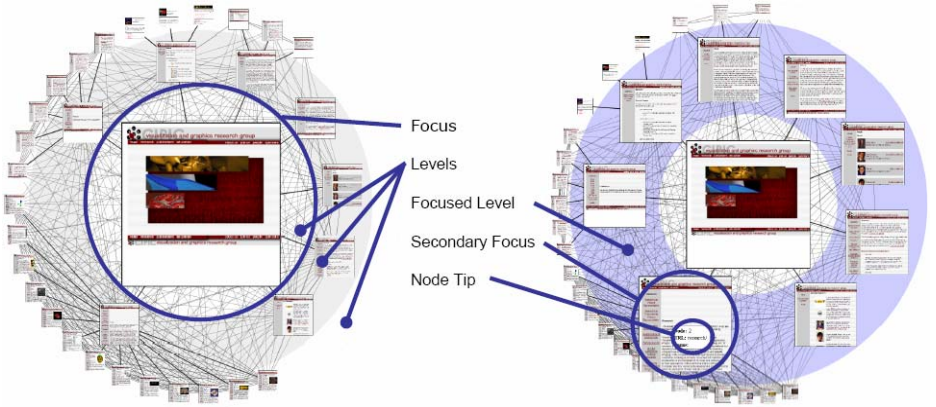
Animated Navigation

[Kelly and Ma, InfoVis 03]

Applications

[Kelly and Ma, InfoVis 03]

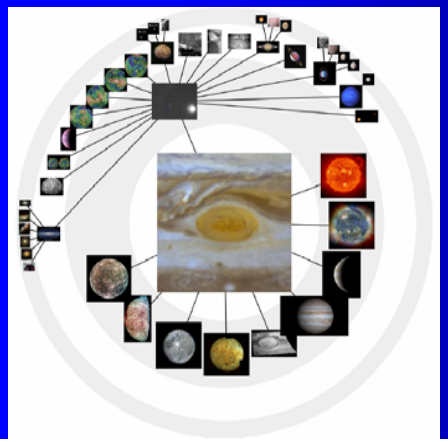
- Web-pages linking graph

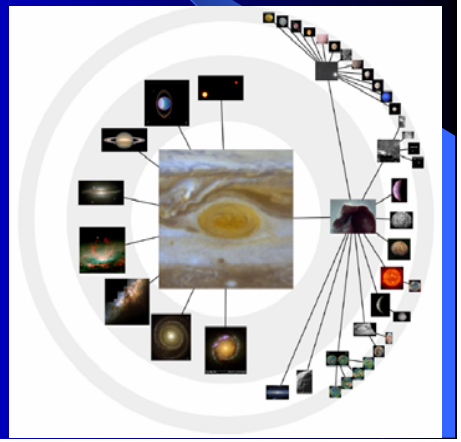


Applications

[Kelly and Ma, InfoVis 03]

- Image database grouped by stellar hierarchy
- Image database grouped by source mission

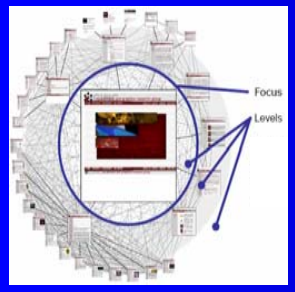




Talk Summary

- Annotation is important for image searching and browsing
- Arrangement of images helps image browsing
- Image browsing can benefit from combining with
  - Semantic image analysis
  - Information visualization techniques (MDS, VaR ..)
  - Imaged-based modeling
  - Imaged-based rendering
  - …
- Annotation can be evaluated, monitored and improved by some image browser

The End



Thank you !