Visual Saliency

Derek Juba and Amitabh Varshney

What Controls Visual Attention?
Visual Attention in VR

- Attention Management
  - Color
  - Texture
  - Geometry
  - Lighting
- Influence visual attention in HMDs
  - Validate through user studies

The Last Inch Problem

Image courtesy of Howard Hughes Medical Institute
Visual Attention

A way of seeing is also a way of not seeing – a focus on object A involves a neglect of object B

Kenneth Burke

• Where we look has significant implications for:
  – what we perceive
  – how we interpret
  – how we act
• Visual attention is the primary filter by which we can cope with our immense sensory bandwidth
  – Retinal information is too vast and most of it has no survival value
• Eye-tracking can quantify overt visual attention

Visual Abstractions

from Visual Intelligence by Ann Marie Barry
Visual Abstractions

- Abstraction involves estimation of visual importance – what do people see?

from [Yarbus 1967]

What Controls Visual Attention

Bottom-up Image Properties (task independent)

Top-down Semantics and Task-driven Properties

[Yarbus 1967]
Controlling Visual Attention

Visual saliency is driven by local contrast:

- Luminance
- Chrominance
- Texture Detail
- Orientation
- Curvature
- Motion
- ...

Visual Attention

The Feast of Belshazzar, Rembrandt, 1635
Applications

- Image saliency maps
  - Tsotsos et al. 95, Milanese et al. 94, Itti et al. 98, Rosenholtz 99

- Applications: compression and cropping
  - Privitera and Stark 99, Chen et al. 03, Suh et al. 03

Visual Scalability

- Camera
- Display

![Graph showing visual scalability over years](image)
Gigapan Robot

http://www.gigapan.org

Digital Camera Technology

1 Gigapixel to 50 Gigapixel
to
Brady et al. Nature, June 20
Personal Drones

Quad, Hexa, and Octocopters

Gigapixel Landscape Image
Panning is Impractical

Can’t miss the salient ...
What is *Interesting*?

Multiscale Aggregation

- Works well on small images
- If we have many more scales …
  - Large regions dominate small regions
  - Wait… we don’t want to miss the small regions

- Traditional multiscale saliency is insufficient
Our Sliding-Window Aggregation

- We see different things at different zoom levels
- One saliency map per level
- Only aggregate up to 4x
- Use a sliding-window across scales
- Why 4x?
  Eye resolution difference ~5x

Mount Whitney (5.0 GPix)
Saliency has a 3D Component

3D features pop out quickly

2D features not pre-attentive

Based on Enns and Rensink 90
Distinctive 3D structure pops out pre-attentively

What is Salient?

- High curvature is important

- but not always…
Mesh Saliency

Saliency should find regions that are different from their surrounding context

Armadillo Leg  Curvature  Saliency

Lee, Varshney, Jacobs, SIGGRAPH 2005

Mesh Saliency

Saliency should capture interesting features at all meaningful scales

Saliency at a large object scale  Saliency at a small object scale
Saliency Computation Overview

- **Curvature**
- **Center-Surround**
- **Saliency Maps at multiple scales**
- **Nonlinear Normalization**
- **Mesh Saliency**

Center-Surround Operator

Gaussian-weighted average is:

\[
G(\bar{\mathcal{C}}(v), \sigma) = \frac{\sum_{x \in N(v, 2 \sigma)} \bar{\mathcal{C}}(x) \exp\left[-\|x - v\|^2 / (2\sigma^2)\right]}{\sum_{x \in N(v, 2 \sigma)} \exp\left[-\|x - v\|^2 / (2\sigma^2)\right]}
\]

\(\bar{\mathcal{C}}(x)\): Mean curvature at vertex \(v\)

Gaussian Weights
Center-Surround Operator

Saliency map at each scale $i$ is:

$$\mathcal{S}(v) = |G(\otimes(v), \sigma_i) - G(\otimes(v), 2\sigma_i)|$$

$\sigma_i \in \{2\varepsilon, 3\varepsilon, 4\varepsilon, 5\varepsilon, 6\varepsilon\}$, $\varepsilon = 0.3\%$ of the diagonal of the object
Mesh Saliency Results

Stanford Armadillo

Cyberware Isis

Saliency Applications

Simplification: Scale the quadric error by the saliency to preserve more triangles for salient regions

Cyberware Male

Mesh Saliency

Viewpoint Selection: Find the viewpoint that maximizes the sum of the visible saliency

• Gradient-descent-based optimization
Validation of Mesh Saliency

Kim, Varshney, Jacobs, Guimbretiere
ACM Transactions on Applied Perception
2010

Original Model
Computed Saliency
Fixations of 6 users over first 3 seconds

Visual Persuasion

- We can draw viewer attention in several ways
- Obtrusive methods like arrows or flashing pixels
  - Distracts the viewer from exploring other regions
- Principles of visual perception used by artists and illustrators
  - Gently guide to regions that they wished to emphasize

Kim and Varshney, IEEE Visualization 2006, TVCG 2006
Visual Persuasion (Art)

The Feast of Belshazzar, Rembrandt, 1635

Joseph’s Bloody Coat Brought to Jacob, Velasquez 1630
Visual Persuasion (Photography)

1. What’s in focus? What’s out?
All the catchers are blurry — with two exceptions, Manager and President. With the oval lens they used in the original photo, they would have been blurry as well. Pratt said this is the most obvious that the photo has been altered.

2. The lighting
Pratt said point to the lighting. The fans’ heads are creating well-defined shadows on their chests, and third baseman and Little League basemen. Bush has a streak of hair shining on his head, right side of his face, while all the other fans have shadows on their faces.

3. Comparing brightness
Pratt used measurement tools in Adobe Photoshop to check the colors of the skin tones. Bush’s skin was about 50 percent brighter than the crowd members around him.

4. Using adjustments
Pratt used another technique to adjust the image’s brightness. “The less adjustments highlight the differences in the lighting,” Pratt said.

Saliency-Guided Visual Persuasion

- Saliency Field
- Enhancement Operators
- Emphasis Field
- Saliency Enhancement
- Saliency-enhanced Volume Rendering
- Validation by eye-tracking based user study
Visual Persuasion - Brightness

User Study – Experimental Design

Extracting fixations from raw points
- Raw points: all points from the eye-tracker
- Saccade Removal
  - Velocity > 15°/sec
- Fixation combining
  - Filter out the points which stay less than 100ms within 15 pixels
  - Average eye locations within 15 pixels and 100ms
User Study – Result I

Traditional Volume Rendering
  With Fixation Points

Gaussian-based Enhancement
  With Fixation Points

Saliency Field

With Fixation Points

Saliency-guided Enhancement
  With Fixation Points

User Study – Result II

Saliency-guided Enhancement
  With Fixation Points

Gaussian-based Enhancement
  With Fixation Points
Data Analysis I

The percentage of fixations on the ROI for the original, Gaussian-enhanced, and Saliency-enhanced visualizations

Geometry-driven Visual Attention

The process involves:

1. Region Selection
2. Mesh Filtering
3. Model Displacement
4. Curvature Change Map
5. Stylized Rendering
6. Validation
Geometry-driven Visual Attention

(a)

(c)

Geometry-driven Visual Attention

(a)

(b)

(c)
Concluding Remarks

• Steady growth vs Disruptive advances
  – Important improvements within established frameworks
  – Developing new frameworks

• Disruption usually involves an enabling technology

• Eye tracking may be that technology
  – Approximately quantifies visual attention
  – Helps close the visual synthesis – visual perception loop
  – Allows us to reverse-engineer mental models of space

Concluding Remarks

• Every communication genre employs its distinct rhetorics
  – Movies vs documentaries
  – Posters vs newspapers
  – Photography vs art

• Possible elements of Visualization Rhetorics
  – Data: nouns
  – Visual Salience: adjectives
  – Visual Persuasion: verbs

• Could be the beginnings of a new and rich genre of visual communication

Or it could go the way of personal air cars …