An Introduction to Eyetracking-driven Applications in Computer Graphics

Eakta Jain
Assistant Professor
CISE, University of Florida
ejain@cise.ufl.edu
jainlab.cise.ufl.edu
Goals

• Applications that use eye tracking data (rather than saliency models)

• An introduction, rather than an exhaustive collection

• Organization:
  • Eyetracking-driven gaze behavior for animated characters
  • Gaze as a source of user priorities
Gaze Behavior for Animated Characters

• “Eyes are the window to the soul.”

• Very important to create lifelike eyes for virtual characters

A still from Polar Express
Critics’ comments included “lifeless eyes”
Challenges and Approaches

- Challenge 1: Modeling and realistic rendering (spheres, texture mapping, iris patterns, etc.)

- Collecting data on shape, appearance and movement

Pamplona et al. (2009), Photorealistic models for pupil light reflex and iridal pattern deformation, ACM TOG

Berard et al. (2014), High-Quality Capture of Eyes, ACM TOG
Challenges and Approaches

- Challenge 2: Animating gaze behavior
- Data driven models
- Playback of recorded animation

Lee, Badler and Badler (2002), *Eyes Alive*, ACM TOG

McDonnell et al. (2012), *Render me real?*, ACM TOG
Applications to Real-time Avatars

- Conversational agents
- Human-robot interaction

Niewiadomski et al. (2013), *Computational models of expressive behaviors for a virtual agent*, Social Emotions in Nature and Artifact

Moon et al. (2013), *Meet Me Where I’m Gazing*, HRI
Summary

• Modeling the eye (appearance, shape, movement) is crucial for creating compelling virtual characters

• For a good overview, see the state-of-the-art report

Gaze as a Source of User Priorities

• Explicit indicator of what the user wants (eye as cursor)

• Implicit indicator of what the user wants (eye reveals what is hard to articulate explicitly)
Gaze as Explicit Indicator
(Eye as Cursor)

• Can be faster than the mouse

• Especially important for users with hand mobility impairments

• Need to worry about the “Midas Effect”

Zhai et al. (1999), Manual and gaze input cascaded (MAGIC) pointing, CHI

Sibert and Jacob. (2000), Evaluation of eye gaze interaction, CHI
Gaze as Implicit Indicator

- Usability analysis
  - Can reveal bottlenecks in a user interface
  - Illustrate differences between systems via scan path analysis techniques


Jacob and Karn (2003), *Eye tracking in human-computer interaction and usability research: Ready to deliver the promises*, Mind
Gaze as Implicit Indicator

• Gaze contingent applications
  • Knowing where people look can lead to efficiency in rendering, modeling, compression, etc.
  • Covered by Sumanta Pattanaik in his session
Gaze as Implicit Indicator

- Knowing where people look provides insight into how the brain is processing the visual world
- Can be used to mimic high level operations such as painterly abstraction

DeCarlo and Santella (2002), *Stylization and abstraction of photographs*, ACM TOG
Gaze as Implicit Indicator

- Knowing where people look provides insight into how the brain is processing the visual world
- Can be used to mimic sophisticated high level operations

Gaze as input to image operations
  e.g., Cropping, Segmentation

Gaze as input to video operations
  e.g., Segmentation, Editing
Automatic methods using saliency models can fail in often simple cases, e.g., where the yellow light is visually salient, but not relevant to the context.

Santella et al. (2006), *Gaze-based Interaction for Semi-Automatic Photo Cropping*, ACM SIGCHI
Image Operation: Moves on stills

Comic panel

Moves-on-stills

Original images copyright MARVEL

Jain et al. (2012), *Inferring Artistic Intention in Comic Art from Viewer Gaze*, ACM SAP

Jain et al. (2016, to appear), *Predicting Moves-on-Stills for Comic Art using Viewer Gaze*, IEEE CG&A
Moves-on-Stills

- **Advantages:**
  - Engage the audience
  - Engage a strength of digital displays (material can be animated)
  - Keep unique characteristic of comic art (each panel is a moment frozen in time)

- **Challenges:**
  - Needs a semantic understanding of the image
  - Need to convert image understanding into a camera move
“...the artist must...secure control of the reader’s attention and dictate the sequence in which the reader will follow the narrative...”
Intensity, Color, Orientations

Emotional content

Hollywood trailers versus natural movies

Task: Judge the age or comment on clothes


Stimuli

Comic art

Photoessay

Amateur snapshots

Robot pictures,
Kang et al. (2009)

Jain et al. (2012), *Inferring Artistic Intention in Comic Art from Viewer Gaze*, ACM SAP
Experimental Setup

• Nine participants
• Calibration done to <1.5 degree error (30-40 pixels)
• Stimuli randomized across the four categories
• Comprehension questions at random points
• Self-paced with a minimum amount of time (4 seconds)
ROC Curves / Area Under Curve

Mean ROC area for each category. Error bars are standard error of the mean. (Gaze data on word bubbles discarded.)

p<0.05

Mean ROC area after leave-one-out.

Stimuli category

Robot
Amateur
Photoessay
Ironman
Watchmen

Watchmen-withtext
Ironman-withtext
Aligned Vector Distance

Sakoe and Chiba (1990)

Mean RMSD for each category. Error bars are standard error of the mean.
Finding

Increased consistency in gaze data of viewers for comic art

Artists are successful in designing a visual route and directing viewers to follow it

- artistic intention can be inferred from recorded gaze data
Image Operation: Moves on stills

Comic panel → Eyetracking device → Gaze data

Points of interest
- Pan
- Track

Framing window parameters (x,y,size)
Rendering the move on still
Image Operation: Moves on stills

Comic panel

Moves-on-stills

Original images copyright MARVEL

Jain et al. (2016, to appear), *Predicting Moves-on-Stills for Comic Art using Viewer Gaze*, IEEE CG&A
Results: World War Hulk

Jain et al. (2016, to appear), Predicting Moves-on-Stills for Comic Art using Viewer Gaze, IEEE CG&A
Gaze as Implicit Indicator

- Knowing where people look provides insight into how the brain is processing the visual world
- Can be used to mimic sophisticated high level operations

Gaze as input to image operations
e.g., Cropping, Segmentation

Gaze as input to video operations
e.g., Segmentation, Editing
Gaze as Input to Image Operations: Segmentation

Challenge: What is an appropriate segmentation?

Gaze as Input to Image Operations: Segmentation

Cluster data to determine regions of interest. Use gaze clusters to assemble super-pixels into segments.
Effects

Figure 1: Leveraging Gaze Data for Segmentation and Effects on Comics

Defocus    Recolor    Stereo
Effects

Problem: How to create effects on legacy comics?

Gaze as Input to Image Operations: Segmentation

Mishra, Aloimonos and Fah. (2009) *Active Segmentation with Fixation*, ICCV
Gaze as Input to Video Operations: Segmentation

Karthikeyan et al. (2015) Eyetracking assisted extraction of attentionally important objects from videos, CVPR

Spampinato et al. (2015) Using the Eyes to “See” Objects, ACM Multimedia
Gaze as Implicit Indicator

- Knowing where people look provides insight into how the brain is processing the visual world
- Can be used to mimic sophisticated high level operations

- Gaze as input to image operations
  - e.g., Cropping, Segmentation
- Gaze as input to video operations
  - e.g., Segmentation, Editing
Video Re-editing

Problem: How would we best present the narrative content in this scene?

Problem: Widescreen Video at a Reduced Aspect Ratio

Several intelligent retargeting operators:
For images, simply cropping is voted to be visually more pleasing.
(A Comparative Study of Image Retargeting, Rubinstein et al., ACM Transactions on Graphics, 2010)
Challenge: Narrative-Important Regions

Predicting is hard, but we can measure…

Original Widescreen Video (1.75:1)

Bottom-up factors, top-down influences, context, motion, audio…

Baluch and Itti (2011)  Rudoy et al. (2013)  Katti et al. (2014) …
Solution: Recording Viewer Gaze

Eyetracking data from six subjects
Place cropping window to best capture viewers’ gaze

Piecewise nonuniform cubic B-spline

Score each trajectory by the number of gaze points enclosed

Cinematic constraints:
Ease-in-ease-out
Knot distance constrains pan velocity
Switch between two trajectories based on a shift in viewer attention, subject to minimum distance (avoid a ‘jump’ cut)
Results

Herbie Rides Again 1974 (1.75:1)

Our Result (1:1)
no zoom

Our Result (1:1)
zoom parameter=1

Zoom achieved by changing the size of the cropping window based on the spread of gaze data in the scene
Validation Via Eyetracking

Blue markers: Result videos
Red markers: Original widescreen videos.

Knowing where people look provides insight into how the brain is processing the visual world.

Can be used to mimic sophisticated high level operations.

- Gaze as input to image operations:
  - e.g., Cropping, Segmentation

- Gaze as input to video operations:
  - e.g., Segmentation, Editing
Gaze as Input to Video Operations: Summarization

- Video operations such as summarization, recommendation, search, categorization, etc. are hugely relevant today.

- Challenge: Large amount of data to process. So, how can we obtain a prioritization?

- Gaze?
  - x-y locations tell us spatial prioritization
  - Temporal prioritization?
Gaze as Input to Video Operations: Summarization

Pupillary dilation: Autonomic nervous system response to emotional arousal

Katti et al. (2011) Affective Video Summarization and Story Board Generation using Pupillary Dilation and Eye Gaze, IEEE International Symposium on Multimedia (ISM)
Challenge

Change in pupillary diameter could be a result of changing screen brightness, as well as, viewer’s emotional response

Can we decouple the emotional response and light response?
Calibration to Changing Brightness

Measured light intensity versus grayscale image intensity of calibration slides

Light Intensity (lumens)

Grayscale Intensity (0-255)

Session 1
Session 2
Session 3
Session 4
Session 5
Linear Model of Pupillary Light Reflex

\[ d(t) = d_0 + k \times T(t - \Delta) \]
Subtracting out Pupillary Light Reflex

Example frame (high arousal)

- Average Intensity = 0.4502
- Measured diameter = 4.1985 mm
- Residual (our model) = 3.1962 mm

Example frame (lower arousal)

- Average Intensity = 0.0759
- Measured diameter = 4.1582 mm
- Residual (our model) = 3.6261 mm
Decay, a zombie video

**Scene 1** (moderately arousing)
- “He’s alive!”
- Camera pans to friend

**Scene 2** (low arousing)
- Furious woman shouting

**Scene 3** (highly arousing)
- Surprised zombie attack

**Scene 4**
- Money and Agius 2008
- Lee et al. 2012
- Lu and Grauman 2012
- Rasheed and Shah 2002
- Rasheed et al. 2005
- Picard et al. 2001
- Graham and Clifton 1966
- Lang and Cuthbert 1997

**Result**
Decoupling Light Reflex from Pupillary Dilation to Measure Emotional Arousal in Videos, ACM SAP

Gaze as Implicit Indicator

- Knowing where people look provides insight into how the brain is processing the visual world
- Can be used to mimic sophisticated high level operations

Gaze as input to image operations
e.g., Cropping, Segmentation

Gaze as input to video operations
e.g., Segmentation, Editing
Summary

• Applications that use eye tracking data (rather than saliency models)

• An introduction, rather than an exhaustive collection

• Organization:
  • Eyetracking-driven gaze behavior for animated characters
  • Gaze as a source of user priorities
13th ACM Symposium on Applied Perception (SAP)

Co-located with SIGGRAPH 2016 (Anaheim, USA)

July 22-23, 2016

Goals and Scope: To advance and promote research that crosses the boundaries between perception and disciplines such as graphics, visualization, vision, haptics and acoustics

Website: http://sap.acm.org

Long and short papers, posters

The strongest long papers have the option to be fast-tracked to a journal publication in the ACM Transactions on Applied Perception (TAP)