Perceptually-Motivated Graphics

Katerina Mania
Technical University of Crete
k.mania@sussex.ac.uk
http://www.ece.tuc.gr/kmania

Erik Reinhard
University of Bristol, UK
reinhard@cs.bris.ac.uk
http://www.cs.bris.ac.uk/~reinhard

Abstract

In this half-day tutorial, we give an overview of the uses of knowledge about the human visual system, as applied to several aspects of computer graphics. In particular, we show how human visual perception applies to the optimization of rendering algorithms, display algorithms, as well as virtual environments. Examples are shown for applications such as real-time rendering, high quality rendering, material editing using images, and training and knowledge transfer in virtual environments. The aim is to show that the human visual perception literature harbours a rich source of knowledge that can be directly applied to improve a wide range of algorithms and technologies in computer graphics.

Keywords

Human visual perception, rendering, display algorithms, virtual environments

Outline of tutorial

Rendering algorithms may benefit from insights into human vision, for the purpose of cutting computational corners that remain below the visible threshold. For real-time rendering and simulation technology, psychophysical investigations can be carried out into the degree of similarity between the original and a synthetic simulation. The question then is whether we can interrogate human cognitive systems which are activated by interacting with a given real-world scene, to see if the same cognitive responses can be evoked under varied levels of scene fidelity. The scope of the proposed tutorial is two-fold and will be organized in two main themes:

- To explore how and which insights from human vision could be exploited towards optimizing rendering speed and quality of synthetic imagery. Such principles could be applied to non-real-time and to real-time rendering algorithms, image
- quality metrics as well as to display technologies such high dynamic range imaging.
- To explore how insights from human perception and cognition could be exploited towards behaviourally realistic Virtual Environments. Such principles could be applied to selective real-time rendering algorithms, positive transfer of training as well as to optimizations for latency degradations and predictive tracking.

Rendering

Rendering algorithms, which take as input a scene description, and produce images which can later be displayed, fall into two categories: those that produce images in (near-) real-time with the highest quality possible within a given frame-rate, and those that produce the highest possible quality at any computational cost.

In particular the former category of rendering algorithms may benefit from insights into human vision, for the purpose of cutting computational corners that remain below the visible threshold, while improving the overall frame-rate. Level-of-detail algorithms, which simplify geometry based on visibility, are among the most well-known examples. Here, less geometry to render means faster frame-rates.

In the case of non-real-time high quality rendering algorithms, there is also a role to play for knowledge of human vision. For instance, the representation of advanced materials such as bi-directional reflectance distribution functions can be optimised through perceptual guidance. Further, stopping criteria for high quality rendering algorithms can be perceptually informed.

Display algorithms

Displays form the interface between machine and human. Presenting an image to the observer induces a percept that needs to be controlled. A simple example of where perception creates an unintended effect is when images taken on a sunny day are displayed indoors under office lighting. Without specific correction, the image will appear too blue, and therefore does not match the memory of the photographer. The reason is that humans adapt to the prevailing lighting conditions, so that in the outdoors environment, the blue illumination from the sky is discounted by the human visual system. This does not happen when an image of the same environment is reproduced in a different viewing environment.

With display technology advancing at a rapid rate, improving both contrast and dynamic range, the mismatch between the environment in which the image was created, and the environment in which the image is viewed, can potentially become quite large, and involve a host of different effects. Hence, in addition to the usual gamma correction, further corrections may have to be pre-applied, anticipating the state of adaptation of the observer.

Simulation Technology

It is not computationally feasible to immerse a person into an interactive artificial environment which exactly mimics the panoply and complexity of sensory experiences associated with a “real” scene. For a start, it is technologically challenging to control all of the sensory modalities to render the exactly equivalent sensory array as that produced by real world interaction. When visual (or interaction) fidelity is increased, the system responsiveness decreases, resulting in reduced frame rates and added visual/tracking latency. It is argued that training in a VE with maximum fidelity would result in positive transfer equivalent to real-world training since the two environments would be impossible to differentiate. Robust metrics are essential in order to assess the fidelity of VE implementations as well as optimizing system design comprising of computer graphics imagery, display technologies and 3D interaction metaphors across a range of application fields.

A goal of Virtual Environment (VE) systems is to provide users with appropriate sensory stimulation so that they act and react in similar ways in the virtual world as they would in the natural world. The research community is challenged to investigate the factors that make virtual reality technologies effective for training purposes.

Psychophysics comprises a collection of methods used to conduct non-invasive experiments on humans, the purpose of which is to study mappings between events in an environment and levels of sensory responses to those events. Psychophysical validation can be subdivided into two nested levels, which we will refer to as Behavioral Realism and Reality Benchmarking. Behavioral Realism, simply put, claims that if a synthesized image or an interactive simulation (e.g., a driving simulator) can support natural behavior (e.g., high speed curve obstacle avoidance), then that technique has captured some of the behaviorally relevant portions of the real image or situation. Several experiments involving driving simulators and dynamic facial animation will be described to concretely demonstrate Behaviorial Realism techniques. The second level of psychophysical validation, Reality Benchmarking, is similar to Behavioral Realism, but attempts to provide a more quantitative measurement. This is accomplished by explicitly investigating how the rendered scene or simulation compares perceptually to its real counterpart. This type of validation will be illustrated by examples from research on the spatial awareness in Virtual Environments.

Physical and psychophysical fidelity issues in the assessment of virtual environments will be emphasised. Input from spatial cognition will be drawn towards optimizations of real-time selective rendering algorithms. Specifications for correct matching between the psychophysical characteristics of the displays and the human users' sensory and motor systems will be discussed as well as some examples of the consequences when systems fail to be physically well matched to their users. Measuring sensitivity of the human to Virtual Environments’s shortcomings such as latency will be explored as well as how such work could influence predictive tracking algorithms. A summary of current presence research funded by the EU Future Emerging Technologies program will be given.

Course syllabus

Introduction **(Mania, 10 mins.)**

Rendering **(Reinhard, 45 mins.)**

- Colourimetry
- Colour appearance modeling
- Dynamic range
- Tone reproduction

Display technologies **(Reinhard, 45 mins.)**

- Visible difference vs. visual equivalence
- Basic concepts of level-of-detail algorithms
- Stopping criteria
- Material perception
- Example: Image-based material editing

Simulation technologies **(Mania, 70 mins.)**

- Fidelity metrics for real-time Virtual Environment simulations
- Spatial awareness in complex, immersive, interactive VR systems.
- Functional Relationships between perception and physics?
- Perceptually-based real-time selective rendering algorithms
- Behavioural Realism: Is the simulation sufficient to support complex behaviors
- Presence research
- Human sensorimotor adaptation to feedback delay and tolerance and visuomotor adaptation to delayed feedback

Discussion **(Both, 10 mins.)**

Presenters

Dr Katerina Mania received her Ph.D in Computer Science from the University of Bristol in 2001 which was funded by Hewlett Packard Laboratories. Prior to that, she worked at HP Labs as a member of technical staff from 1996-1998. She was appointed as a Lecturer in Multimedia Systems at the University of Sussex, UK (Department of Informatics, 2001-2005) and currently serves as an Assistant Professor in the Department of Electronic and Computer Engineering of the Technical University of Crete, Greece since December 2005. In 2003 she worked on perceptual sensitivity to tracking latency at NASA Ames Research Centre, USA.

Exploiting fundamental memory research, she explores how “close” photorealistic simulations for training are compared with their real-world counterpart, from a cognitive rather than a physics point of view. She is currently working on perceptually-based real time selective rendering algorithms, on visualizing uncertainty in archaeological reconstructions based on possibility theory and Bayesian approaches, on real-time embodiment in virtual spaces based on non-invasive sensor data for mapping user states (emotional, physical) to the dimensions of form, colour, sound and movement, on ubiquitous computing and multimodal interfaces for interacting with digital heritage artifacts and on synthetic facial emotion cues.

Dr Katerina Mania currently serves as one of the Associate Editors for Presence, Teleoperators and Virtual Environments (MIT Press) as well as for ACM Transactions on Applied Perception.

Dr Erik Reinhard received his Ph.D. in Computer Science from the University of Bristol in 2000. Afterwards he was a post-doctoral researcher at the University of Utah (2000-2002) and assistant professor at the University of Central Florida (2002-2005). He started a lectureship at the University of Bristol in January 2006, and is currently senior lecturer (associate professor) at the same university.

His work focuses on rendering and display algorithms, with a particular interest in the application of perceptual knowledge in these areas of computer graphics. He has published several psychophysical investigations in premier venues such as SIGGRAPH, ACM Transactions on Applied Perception, and the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization (APGV). In addition, he was the lead author on the first book on high dynamic range imaging (with Greg Ward, Sumanta Pattanaik, and Paul Debevec), and is currently finalising a second book on color imaging. Both books directly relate perceptual issues to display algorithms, as well as other applications in computer graphics.

Further, he founded ACM Transactions on Applied Perception with Heinrich Buelthoff, and is editor-in-chief of this journal. He also acted as programme co-chair for APGV in 2006 with Bill Thompson. Finally, Erik is member of the CIE Technical Committee TC8-08, ``Testing of Spatial Colour Appearance Models``, dealing with the specification of a procedure to validate tone reproduction operators.

Display Algorithms and Rendering

Erik Reinhard
reinhard@cs.bris.ac.uk

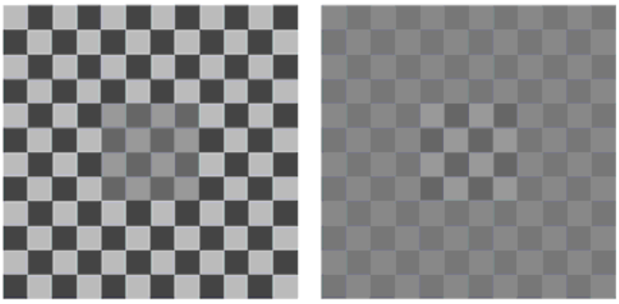
Human Vision

- Light enters the eye through the pupil
- Image formed on the retina
- Non-linear processing in eye and brain
- Human Vision is not a Simple Linear Light Meter

- Perception relates to low-level processing
- Cognition relates to high-level processing, and includes memory, thought, emotion,...

Size Constancy

Contrast Constancy

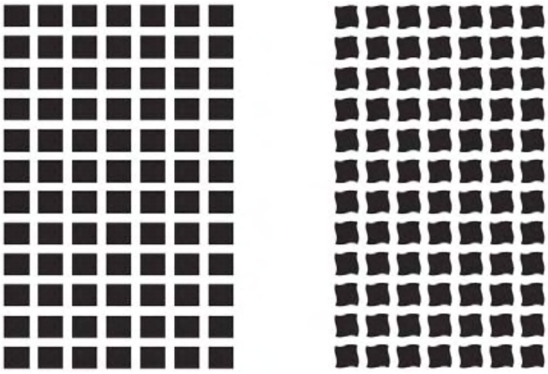


Lightness Constancy

Cafe Wall

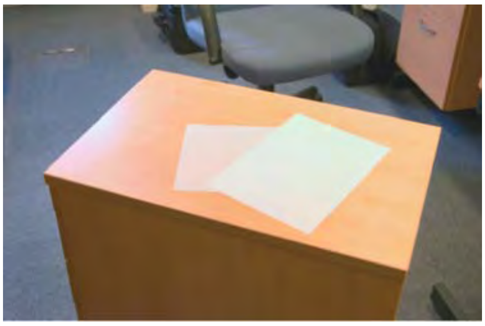
Cornsweet-Craik-O'Brien

Hermann Grid Illusion



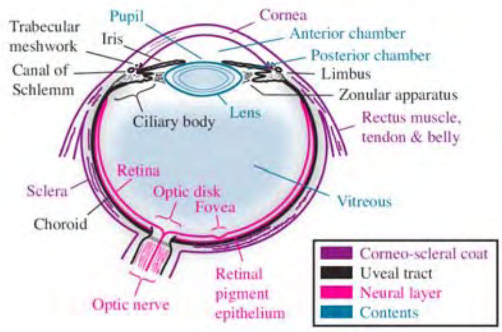
Gelb Illusion

Gelb Illusion

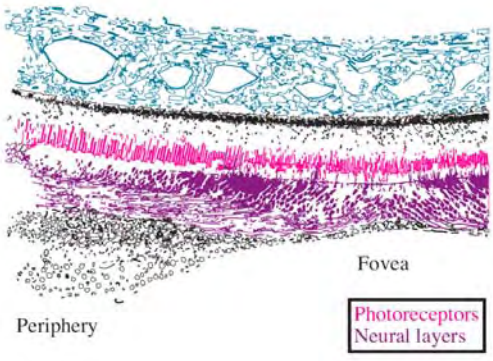


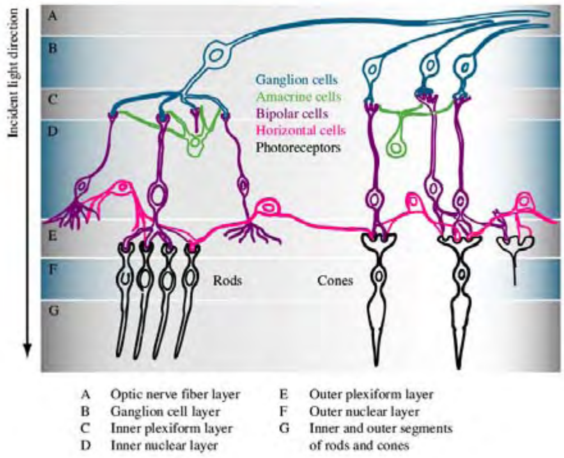
Gelb Illusion

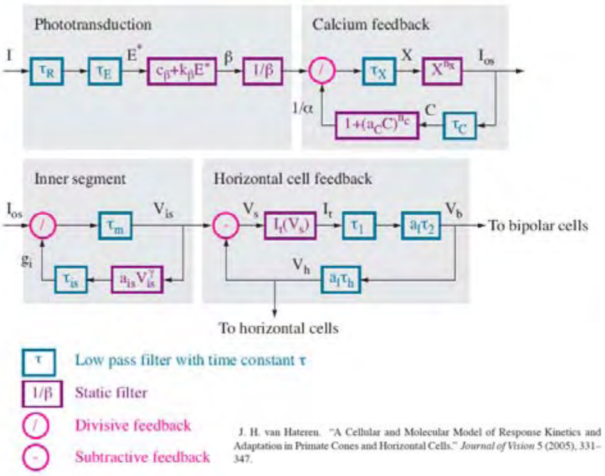
Eye



Retina

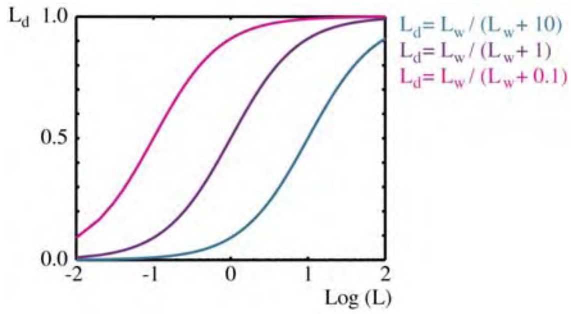






J. H. van Hateren. "A Cellular and Molecular Model of Response Kinetics and Adaptation in Primate Cones and Horizontal Cells." *Journal of Vision* 5 (2005), 331–347.

Photoreceptors – Simplified Response



Simplified Response with Lateral Connectivity

$$L_d(x,y) = \frac{L_w(x,y)^n}{L_w^{\mathrm{LPF}}(x,y)^n + L_w(x,y)^n}$$

Centre-Surround Processing

- Lateral connectivity:
  - Accumulate responses over a given neighbourhood of receptors
  - Subtract combined response from signal
  - ON-pathway
  - OFF-pathway
  - Common processing block in the HVS
- Can be modelled with difference of Gaussian process

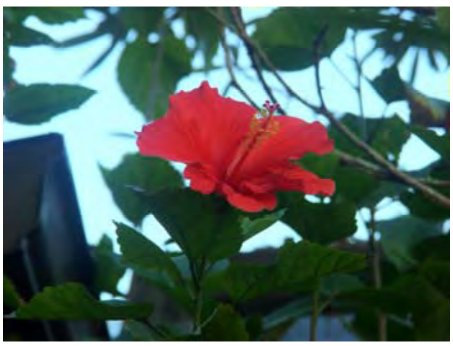
Difference of Gaussian

Difference of Gaussian

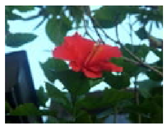
Scale-Space

Scale-Space

Opponent Processing



Opponent Processing



Opponent Processing

- Yields a decorrelated colour space
  - In practice close to independent axes
  - Helps overcome the communication bottleneck of the optic nerve
- Color opponency is multiplexed with spatial centre-surround organisation

Applications

- Retinal processing is relatively well understood
- Complicated computational model can be well approximated with simple sigmoids, which are used in many applications, including
  - Colour appearance modeling
  - Tone reproduction

(In red: discussed next)

Applications

- Spatial processing used in:
  - Visible difference predictors
  - Tone reproduction
- Opponent processing used in:
  - Colour difference metrics
  - Colour transfer between images
  - Transmission primaries

Colour Appearance Modeling

- Predict appearance of colour in the context of an environment
- Need tristimulus value of colour, plus a description of the environment
- Three components to algorithm:
  - Chromatic adaptation
  - Non-linear compression
  - Calculation of appearance correlates

CIECAM02 – Chromatic Adaptation

- von Kries hypothesis: cones are independent
  - Convert image to cone response space
  - $$R_c = \left[\frac{Y_w D}{R_w} + 1 - D\right] R,$$
    $$G_c = \left[\frac{Y_w D}{G_w} + 1 - D\right] G,$$
    $$B_c = \left[\frac{Y_w D}{B_w} + 1 - D\right] B.$$

CIECAM02 – Chromatic Adaptation

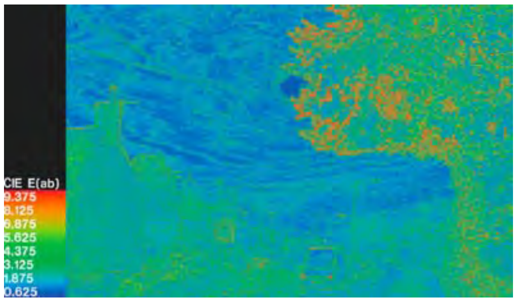
Colour Difference Metrics

- Convert to colour opponent space (CIE LAB)
- Compute Euclidian distance between pairs of corresponding pixels

Losslessly Encoded Image

JPEG Encoded Image

Colour Difference (CIE ΔE^*_{ab})



Tone Reproduction

- Reduce dynamic range of an image to fit the capabilities of a given display device
- Many solutions possible, often drawing from human visual perception
- Several operators in essence implement photoreceptor responses by using sigmoids

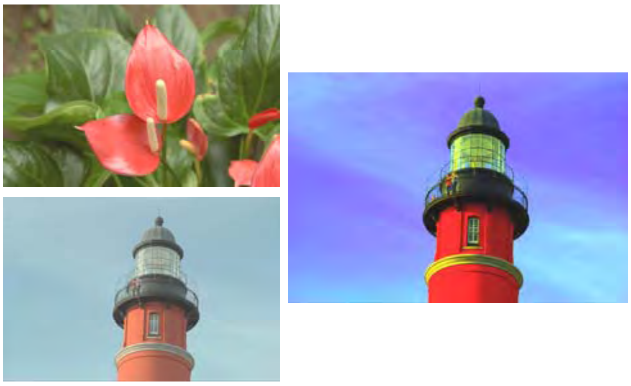
Sigmoidal Tone Reproduction

Comparison

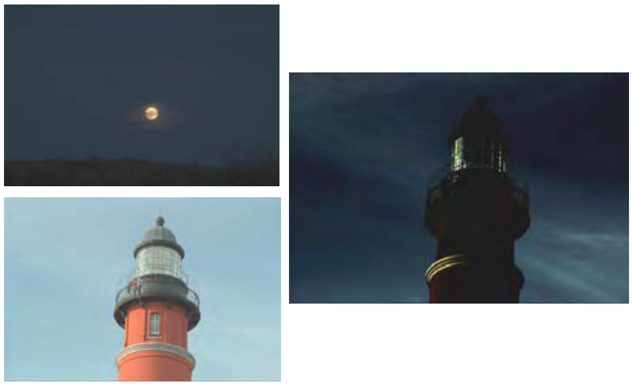
Colour Transfer

- Replace colour palette of an image what that of another
- Algorithm
  - Convert to decorrelated colour opponent space
  - Compute mean and standard deviation along each of the three colour axes
  - Shift and scale target image pixels to attain mean and standard deviation of source image

Colour Transfer



Colour Transfer



Transmission Primaries

- Conserve bandwidth by
  - Conversion to colour opponent space
  - Encode luminance at high spatial resolution
  - Encode chromatic channels at lower spatial resolution

32x Subsampling (All Channels)

32x Subsampling (Chromatic Channels Only)

Rendering

- Image synthesis involves the simulation of light
  - Electromagnetic Radiation (Maxwell's Equations)
  - Geometric Optics

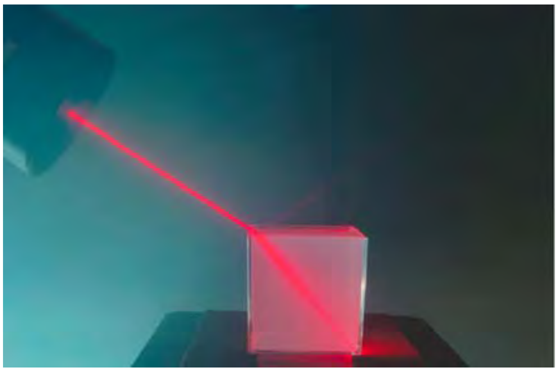
$$L_o(\mathbf{P}, \Theta_o) = \int_{\Omega_i} f_r(\mathbf{P}, \Theta_i, \Theta_o)\, L_i(\mathbf{P}', \Theta_i) \cos(\theta_i) d\omega_i,$$

  - Ray Tracing
  - Radiosity
  - Projective Algorithms

Example: Ray Tracing

$$\begin{aligned} L_o(\mathbf{P}, \Theta_o) = L_e(\mathbf{P}, \Theta_o) \\ &+ \sum_S \int_{\mathbf{P}_S \in S} v(\mathbf{P}, \mathbf{P}_S)\, f_r^{\text{diff}}(\mathbf{P})\, L_e(\mathbf{P}_S, \Theta_o')\, \cos(\theta_S)\, d\omega_S \\ &+ \int_{\Theta_R \in \Omega_R} f_r^{\text{spec}}(\mathbf{P}, \Theta_R, \Theta_o)\, L_o(\mathbf{P}_R, \Theta_R)\, \cos(\theta_R)\, d\omega_R \\ &+ \rho_d(\mathbf{P})\, L_a(\mathbf{P}). \end{aligned}$$

Ray Tracing



Ray Tracing

Radiosity

$$\begin{aligned} L(\mathbf{P}) = L_e(\mathbf{P}) \\ &+ \frac{\rho_d(\mathbf{P})}{\pi} \int_{\text{all } \mathbf{P}'} L(\mathbf{P}')\, \frac{\cos(\Theta_i)\cos(\Theta_o')}{\|\mathbf{P}' - \mathbf{P}\|^2}\, v(\mathbf{P}, \mathbf{P}')\, dA', \end{aligned}$$

Radiosity

Rendering: Other Phenomena

- Polarization
- Birefringence
- Interference
- Iridescence
- Diffraction
- Scattering (Rayleigh, Mie)

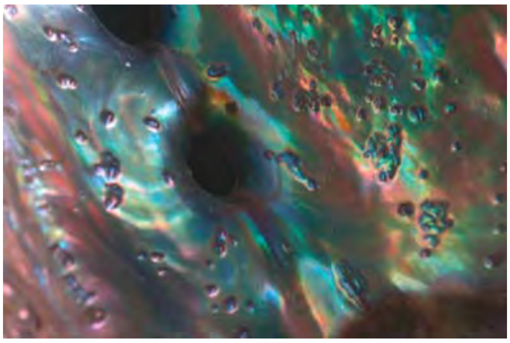
Polarization

Polarization

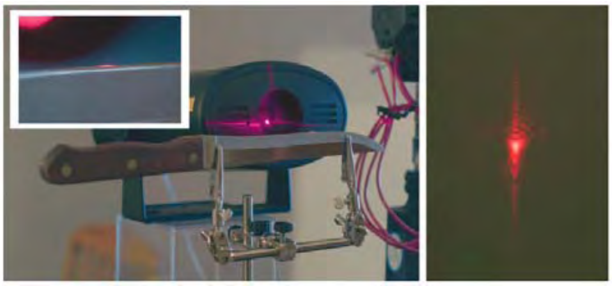
Birefringence

Interference

Iridescence

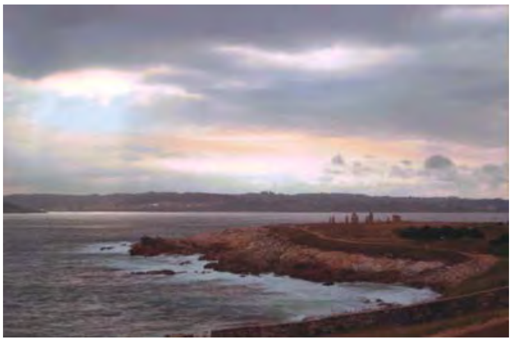


Diffraction



Rayleigh Scattering

Mie Scattering



Rendering / Image Processing

- Many visual phenomena ignored for computational efficiency
- Further computational advantages may be gained from considering human vision
- In addition, new applications in image processing are becoming available

Medium Level Processing

- To help understand its environment, the human visual system:
  - Solves an underconstrained problem
  - Makes assumptions on its environment
- Examples
  - Light comes from above
  - Dark is deep
  - Environment is "normal"

Medium Level Processing

- Consequences:
  - Some image degradations are easily spotted
  - Others tend to go unnoticed
- Challenge in graphics:
  - Avoid computations that lead to visual improvements that go unnoticed
  - Develop applications that exploit particularities of human visual processing

Examples

- Visible Difference Metrics
- Visual Equivalence
- Level of Detail Management
- BRDF Representation
- Stopping Criteria for Radiosity
- Image-based Material Editing

Visible Difference Predictor

Visual Equivalence

BRDF Representation

Stopping Criteria for Radiosity

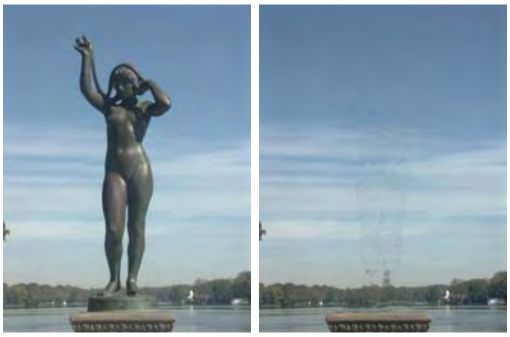
Image-Based Material Editing

- Change the pixels in a photograph such that an object appears to be made from a different material
- Various transforms possible, but here we focus on transparency

Image-Based Material Editing

- On the basis of an input image and an alpha matte delineating the object of interest:
  - Create lighting environment by removing the object using inpainting techniques
  - Recover depth map using dark-is-deep-paradigm
  - Texture map environment onto recovered geometry

Inpainting



Depth Recovery

Bilateral filtering used to help remove textures

General Texture Mapping

Transparency

Image-Based Material Editing

- Problem is under-constrained
  - Physically accurate simulation is impossible
  - Perceptually plausible results are obtainable
  - Visual equivalence achieved
- Humans can detect transparency easily, but cannot predict its visual appearance very well.

Conclusions

- Knowledge of the intricacies of the human visual system help us solve many computational problems
- Low level processing:
  - Colour appearance modelling
  - Tone reproduction
  - Colour transfer
  - Visible difference predictors
  - Edge Detection
  - ...

Conclusions

- Medium level processing:
  - Visual equivalence
  - Level of detail management
  - BRDF representation
- Gives guidance to where and when computational optimisations can be expected

Conclusions

- High level visual processing
  - Computational models of cognition are sparse
  - Not (yet) a well understood mechanism
  - Difficult to apply directly in graphics/vision/image processing

Acknowledgements

Many images taken from the forthcoming book:

E Reinhard, E A Khan, A O Akyuz, and G M Johnson,
“Color Imaging: Fundamentals and Applications”,
A K Peters, 2008

Acknowledgements

- Greg Ward, Erum Khan, Oguz Akyuz, Garrett Johnson, Roland Fleming, Heinrich Buelthoff, Peter Shirley, Michael Ashikhmin, Mike Stark, Mark Colbert, Jared Johnson, Ben Long, Timo Kunkel, Tania Pouli, Veronica Sundstedt, Diego Gutierrez, Jorge Lopez

Fidelity Metrics for Computer Graphics Simulations

Katerina Mania
k.mania@ced.tuc.gr
http://www.ece.tuc.gr/kmania

Visual Fidelity,

Visual fidelity **refers to the degree to which visual features in the VE conform to visual features in the real environment**
Waller et al. 1998

- *Physical* realism, in which the synthetic scene is an accurate point-by-point representation of the spectral radiance values of the real scene
- *Photorealism*, in which the synthetic scene produces the same visual response as the real scene even if the physical energy depicted from the image is different compared to the real scene
- *Functional realism*, in which the same information is transmitted in real and synthetic scenes while users perform visual tasks targeting transfer of training in the real world Ferwerda 2001

Interface or interaction fidelity **refers to the degree to which** the simulator technology (visual and motor) is perceived by a trainee to duplicate the operational equipment and the actual task situation**.**
Waller et al. 1998

Measures of Simulation Fidelity

- **System characteristics optimised (latency, FoV, resolution, etc.)**
  Adelstein et al. 2003, Ellis et al. 2002, Arthur 2000
- **Thresholds for human sensitivity to dynamic anomalies applied to physics-based animation**
  O Sullivan et al. 2003
- **Task performance efficiency in real-world task situation and 3D simulations**
  Kort et al. 2003, Mania et al. 2003, 2005
- **Presence as a metric to assess the effectiveness of a VE, or aspects of a VE according to its success in enhancing presence**
  Stanney et al. 1998

Measures of Simulation Fidelity

- **Effective VEs should maximise the efficiency of human task performance in VEs**
  Stanney et al. 1998
- **VEPAB: VE Performance Battery, a move towards benchmarking VE performance**
  Lampton et al. 1994

-**Vision:** Acuity-Colour Perception-Object recognition-Size/Distance estimation
-**Locomotion:** Walking, Navigational Tasks
-**Manipulation:** Grasping Objects
-**Tracking:** Use head movements to move cursors on targets
-**Reaction time:** Reporting time when seeing objects

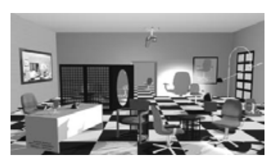
Memory for Places

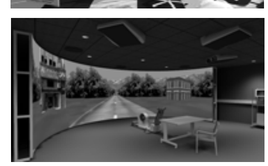
Research Philosophy

Perception and action in natural settings using

Computer Graphics to generate natural but well controlled stimuli of objects and scenes

Virtual Reality to study perception and action in a closed loop





Spatial memory tasks

- **Memory for places is a task often incorporated in fidelity benchmarking processes**
  Dihn et al. 1999, Bailey & Witmer 1994, Bliss et al. 1997, Waller et al. 1998, Ruddle 2006, Mania et al. 2005, Mourkoussis et al. 2007
- **The utility of VEs for any applications for which they are being proposed is predicated upon the accuracy of the spatial representation formed in the VE**

Spatial memory tasks

- **Effects of tactile, olfactory, audio and visual sensory cues on participants' memory recall of building spatial layout**
  Dihn et al. 1999
- **Form and development of spatial representation in VE training in relation to either real-world training or training with maps, photographs and blueprints**
  Bailey & Witmer 1994, Bliss et al. 1997, Waller et al. 1998, Ruddle 2002, 2006
- **Questionnaires, drawing recall and recall/recognition memory tasks, Sketches post exposure, actual navigation in real building, etc**

Way-finding metrics

- **Time taken to complete the task, distance travelled, or number of errors made**
- **Counting the number of times each part of an environment is visited or searched**
- **Frequency with which movements are made in different directions (e.g., forward vs. backward)**
- **Frequency with which collisions are made with the fabric of the environment**
- **Time Classification are: (a) whether or not a user is locomoting (stationary vs.traveling), and (b) whether or not a user is changing the direction in which they are looking within a given frame of reference (static vs. looking around)**
- **Errors during spatial 'decision points'**

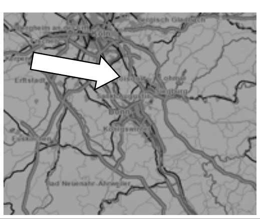
Memory Awareness Studies

Funded by Hewlett Packard Laboratories

- **Investigating the effect of different viewing conditions (direct perception on a real-world setting versus various CG representations) on participants object-location recall**
- **Main premise: Accuracy of performance per se does not reflect the cognitive activity that underlies performance in memory tasks**
  Mania, Troscianko, Hawkes & Chalmers 2003, Mourkoussis et al. 2006

Spatial updating and navigation in virtual environments

What information is necessary for humans in order to maintain a consistent spatial map of their surroundings - to keep us "on the map"?



Motivation

Riecke et al., APGV 2004

Disorientation in virtual environments and multimedia spaces
- Even for simple navigation tasks
- In spite of full cognitive control

However, in real world: Quick, automatic, and effortless spatial orientation, without thinking much (even ants can navigate well)

What is missing in many VR applications?

What is essential for quick and intuitive spatial orientation?
- **"Automatic spatial updating"**

Under what conditions does spatial updating work?

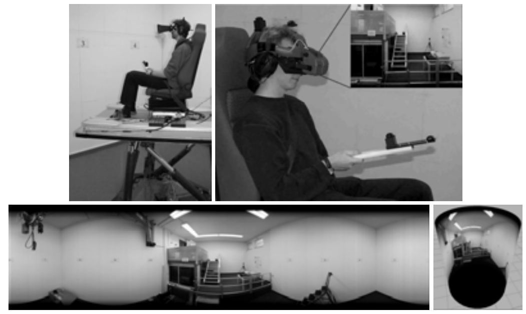
Approach: Experiments in a real room and VR replica

Visual-Vestibular Integration

Riecke et al., APGV 2004

- **Use motion platform to dissect spatial updating into visual and vestibular components**
- **Use rapid pointing to learned landmarks to study reflex-like processes and to avoid cognitive influence**
- **Compare performance in real and simulated environments**

Experimental Design: Setup *Riecke et al., APGV 2004*



Experiment Results *Riecke et al., APGV 2004*

- **Overall small response times & high pointing accuracy**
  - ➢ Good spatial orientation in VR *is* possible
  - ➢ Ease & intuitive usability of our rapid pointing paradigm
- **Whenever useful visual cues were available:**
- **Performance independent of turning angle**
  - ➔ automatic spatial updating
- **IGNORE harder than UPDATE**
  - ➔ obligatory spatial updating
- **Visual cues take priority over vestibular cues**
- *Performance in VR was almost as good as in the real world*

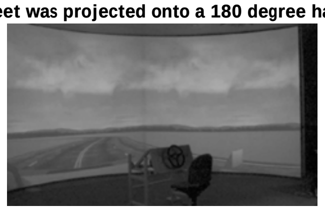
Spatial updating and navigation in virtual environments

- **Visual cues alone, without any concurrent vestibular cues, can be sufficient for „turning the world inside our head“, even against our own conscious will**
  - ➔ reflex-like, obligatory spatial updating
  - • This might **not** be sufficient for more complex environments and tasks (navigation)
- **Psychophysical experiments in VR (Virtual Tübingen) can help to** ***understand*** **and** ***quantify*** **how humans navigate**

Experimental Design

Cunningham et al., Journal of Vision, 2001

- **Testing temporal delays using a driving task in virtual reality**
- **Subjects were asked to drive to the end of the street without leaving the road.**
- **They were asked to remain in their assigned lane, if possible.**
- **A realistic street was projected onto a 180 degree half cylinder screen.**



Extension to VR

Mohler et al., APGV 2004

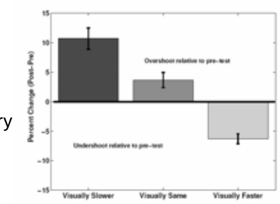
- **University of Utah**
- **Treadmill in wide-screen VR environment**

Adaptation in VR

Mohler et al., APGV 2004

Two experimental questions:
- Can we get adaptation in VR?
- Does adaptation in VR carry over to real world?

Task similar to previous study
- Training phase done in VR



Running versus walking

Mohler et al., APGV 2004

Human gait transitions:
- Walking->running 2.1m/s, running->walking 1.9m/s
- Explanations mainly based on energy expenditure

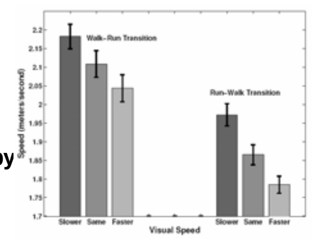
Is there a role of visual information in the gait transition from walking<->running?
- Manipulation of this easy to do in VR

Running versus walking

Mohler et al., APGV 2004

Significant effect of visual input on the transition

Hypothesis:
Visual input used by humans to *initiate* transition



Virtual Reality for studying human locomotion abilities

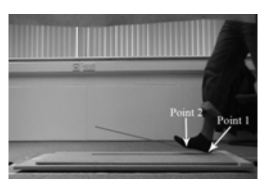
Studies have shown that human action systems are calibrated by the properties of visual input
- Optic flow of environment during locomotion for control of speed

VR (displayed via screens or HMDs, driven by locomotion interfaces) provides a flexible paradigm for studying the relation between visual input and motor output

Perception of slant

Both real & virtual environment studies have demonstrated that perceptual estimates of geographical slant are overestimated, whereas haptic estimates are more veridical when participants judge hills from a stationary point without actually walking on the hills
- Bhalla and Proffitt, 1999; Creem-Regehr et al. 2004

Action-based measure of slant



The measure is in effect the angle of the foot during walking. The foot angle can be defined as the angle formed by the horizontal plane and the imaginary line between the lowest two points of one's foot before it touches the ground.

This work explores the validity of the measure and if it is malleable to factors such as the materials of ramps that participants walk on, the degrees of inclination of the ramps, steps taken with the dominant or the non-dominant foot and the variability of measurements when the measuring task is repeated in both real and immersive virtual environments.

Mourkoussis, N., Rivera, F., Troscianko, T., Mania, K., Hawkes, R. (2007)

Comparative VR/Real Study: Design

Environment
- Real room & photorealistic VR equivalent

Conditions
- 6
- 2 levels of surface slipperiness x 3 levels of surface inclination

Levels of slipperiness
- Low: carpet
- High: ceramic tiles

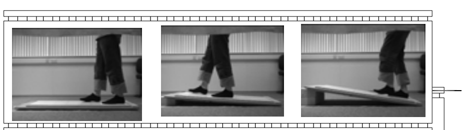
Levels of inclination
- 0, 5, 10 degrees

Participants
- 96 Sussex staff & students
- 48 per environment

Data extraction
- Angle of foot as measured between horizontal plane & foot's lowest two points from video recorded frames

Comparative VR/Real Study: Environments

Comparative VR/Real Study: Conditions in Real Study



3 degrees of inclination either with tiles or carpet

VR study

Conclusions

- **The results imply that the derived foot measure is modulated by motoric caution. If the "reality" of the VR environment is relatively high, results should reveal the same degree of caution as in the real world, but that is not the case here.**
- **It seems reasonable to assume that a larger derived foot angle arises from a longer stride, and thus a faster gait, implying that the walker is less cautious. On this interpretation of the results, people become more cautious when the ground plane is steep, slippery, or virtual.**

Selective Rendering

Previous work on real-time selective rendering

- Selective rendering in high quality the foveal region based on gaze information
  MCConcie, Loschky 1997; Watson et al. 1997
- Selective rendering in high quality the foveal area based on a priory knowledge of the users task focus
  Cater, Chalmers & Ward 2003
- Selective rendering based on saliency models
  Yee et al. 2001; Harber et al. 2001

Rendering in high quality the foveal region based on gaze information

- Gaze-dependent processing can economize in computational complexity by rendering in high resolution only those parts of image which are at the focus of attention
- However it encounters the problem of keeping up with updating the multi-resolution display after an eye movement without disturbing the visual processing
  - 5 ms of update time is allowed after fixation

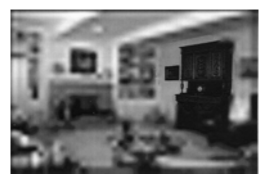
Rendering in high quality the foveal region based on gaze information

- Evaluated the effectiveness of high detail insets in HMDs
- Subjects performed a search task with different display types
- Each display type was a combination of two variables: Inset size and peripheral resolution
- Results not significantly distinguishable, but observers found search targets faster for the fine resolution no inset condition

Watson et al. 1997

Rendering in high quality the foveal region based on gaze information

- Measured image quality judgements
- Observers examining complex scenes with an eye linked multiple resolution display
- Images filtered with a radius of 4.1 degrees of foveal area statistically indistinguishable with a full high-resolution display



McConkie and Loschky 1997, 2000

Rendering in high quality the foveal region based on gaze information

- Gaze-dependent processing can economize in computational complexity by rendering in high resolution only those parts of image which are at the focus of attention
- However it encounters the problem of keeping up with updating the multi-resolution display after an eye movement without disturbing the visual processing
  - 5 ms of update time is allowed after fixation

Rendering in high quality the foveal area based on a priory knowledge of the users task focus

- Salient objects that would normally attract the viewer's attention are ignored if they are not relevant to the task at hand
- Viewers presented with two levels of quality in animations
- Results indicated that 2 degrees of foveal area rendered in high quality had the same results as the high quality animation

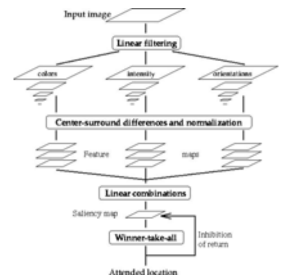
Cater, Chalmers & Ward 2003

Selective rendering based on saliency models

- **What is a saliency model?**
  - Based of the existence in the brain of a specific visual map encoding for local visual saliency
    Koch and Ullman 1985
  - Its purpose being to represent the saliency at every location in the visual field by a scalar quantity, and to guide the selection of attended locations based on spatial distribution of saliency
    Wooding 2002

Selective rendering based on saliency models

- **Modelled as a dynamic neural network.**
- **Attempts to predict given an input image**
- **Which regions of interest in the image will automatically and unconsciously draw one's attention**

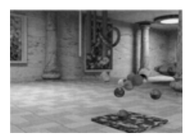


Itti et al. 1998

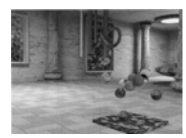
Selective rendering based on saliency models

- **A 4 to 10 times speedup over the time it would have taken to render the image in full in pre-rendered animation**
  Yee et al. 2001
- **However, such bottom-up visual attention models do not always predict attention regions in a reliable manner**
  Marmitt, Duchowski 2002
  - Studies indicated that the correlation between actual human and artificially presented scan-paths was much lower than predicted.
  - Probably because of the algorithm's lack of memory

Selective rendering based on saliency models



Original Image



Final rendering

Saliency map

Aleph Map

Yee et al. 2001

The Schema Theory Studies (preliminary)

Communication with Prof. Bill Brewer, Uni. Of Illinois, USA
Funded by EPSRC, UK in collaboration with Prof. T. Troscianko & Hewlett Packard Labs

- **Focusing on memory for places (rooms)**
- **Selective rendering reported in literature could be fovea-focused or task-focused.**
  Cater, Chalmers & Ward, 2003
- **Functional realism: The same information is transmitted in real and artificial scenes.**

The Schema Theory Studies (preliminary)

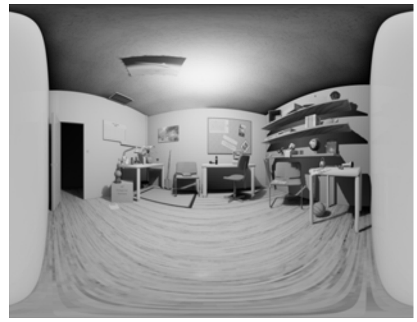
Communication with Prof. Bill Brewer, Uni. Of Illinois, USA
Funded by EPSRC, UK in collaboration with Prof. T. Troscianko & Hewlett Packard Labs

- **Schemata are knowledge structures or sets of expectations based on past experience**
- **An individual's prior experience will influence how he or she perceives, comprehends and remembers new information**
- **Information slots which have not been filled with perceptual information, are filled by default assignments based on stereotypic expectations from past experience**
  Brewer & Treyens, 1981

Selective rendering based on schemas

- Schema consistent spatial elements associated with one environment are expected to be found in this environment
- Information slots which have not been filled with perceptual information, are filled by default assignments based on stereotypic expectations from past experience
  Kuipers 1975

Fish-eye view



Schema Influence on Memory

- Memory performance is frequently influenced by schema-based expectations
  Brewer & Treyens, 1981
- Relevant research has shown that an *activated schema* can aid retrieval of information in a recall task
- Schemata are used to guide the search for information in memory; thus information which is not related to the schema being used in retrieval will be harder to recall than information which is schema related

Conclusions

- Schema objects seemed to require gross quality of the rendering (render less)
- Non-schema objects need to be noticed by detailed inspection (render more)
- The schema/non schema dichotomy taps into two separate processes each of which plays a role in our encoding/recall of a scene
  Mania, Robinson, Brandt 2005

Troscianko, T., Mourkoussis, N., Rivera, F., Mania, K., Dixon, T., Hawkes, R. (2007)

Presence Research

Presence in VEs

- *Virtual Environments:* The sense of being 'there'; the degree to which the users feel that they are somewhere other than they physically are while experiencing a computer generated simulation
  Barfield & Weghorst 1993, Slater & Usoh 1998
- An epiphenomenon for design?
  Ellis 1996
- GOAL: Equation of presence that allows to trade off factors against each other while still maintaining the same level of presence.

Presence – does it exist?

- *Virtual Environments:* **Responding equally to 'virtual' events as to real events at many levels – from physiological responses through to behavioural and cognitive responses, including what can be picked from EEG and fMRI**
  Slater 2004
- **There is no scientific evidence – but it has been demonstrated it exists – a powerful application is in psychotherapy**
- **Anxiety a surrogate for presence**
  - Fear of heights, Emmelkamp et al. 2002, Meehan et al. 2002
  - Fear of flying, Rothbaum et al. 1996
  - Arachnophobia, Carlin et al. 1997
  - Agoraphobia, Banos et al. 2003
  - Burns treatment, Hoffman et al. 2004

Presence – does it exist?

http://www.hitl.washington.edu/projects/exposure

Fear of public speaking

Pertraub, Slater & Barker 2002

Measuring Presence

- **Response to events with 'reflex' reactions**
  Held and Durlach 1987, Loomis 1992)
- **'Noise' added to real images, until it is impossible to be distinguished from the virtual image**
  Loomis 1992, Schloerb 1995
- **Hand-held slider; continuous rating of presence**
  Freeman et al 1999
- **Physiological Measures**
  Meehan 2000
- Breaks-in-Presence, BCIs, Neuro-correlates
  Slater 2006
- *Most common:* **Questionnaires**
  Witmer & Singer 1998, Barfield & Hendrix 1995, Slater et al 1998).

Presence is 'enhanced' with…

- **Wider Field of View**
- **Faster frame rate**
- **Lower latency**
- **Sound rather than no sound**
- **Haptics rather than no haptics**
- **Stereo rather than mono**
- **Head tracking rather than none**
- **Visual realism does not show up on the list..**
- **People do respond to virtual characters that are far from visually realistic**
  Pertraub, Slater & Barker, 2002
- **Sometimes people do not report greater presence between real world and simulation**
  Usoh et al. 2001, Mania et al. 2003

Presence and Task Performance

- **The relationship between presence and task performance is not clear although it is commonly though to be positive**
- **Causal or Correlational? It is argued that variables that increase presence, also increase task performance independently of their effect on presence**
  Welch 1999
- **Contradictory results**
  Witmer & Singer 1994, Singer et al. 1995

Presence and Aftereffects

- **Possible associations between presence and aftereffects may aid effective design of VEs**
- **Negative correlation has been reported - not always significant**
  Witmer & Singer 1994, Mania & Chalmers 2001
- **Positive correlation has been reported also**
  Wilson et al. 1997
- **Or, the greater the aftereffects, the greater the increase in presence over the period of exposure in the VE**
  Welch 1997

Measuring Perceptual Sensitivity to Head Tracking Latency

2003, research conducted at
NASA Ames Research Centre
Human Factors Research and Technology Division
Spatial Perception and Advanced Displays Laboratory

Ellis, S.R., Mania, K., Adelstein, B.D., Hill, M. (2004)

Mania, K., Adelstein, B., Ellis, S.R., Hill, M. (2004)

Latency: Definition

- **Latency is characterized as the time lag between a user's action and the system's response to this action.**
- **Human factors literature has established that these delays have a significant impact on user performance and user impressions of simulation fidelity of a training system**
  Ellis, Young, Ehrlich, & Adelstein, 1999, Jung, Adelstein & Ellis, 2000
- **Latency forces users to slow down to preserve manipulative stability**
- **Trade-off between latency and update rate**

Previous work at NASA Ames

- **Focused on precision, stability, efficiency and complexity of operation interaction with latency plagued systems**
- **First measures of human operator's discrimination of the consequences of latency during head or hand (tracked) movements**
  Ellis, Young, Ehrlich, & Adelstein 1999a,b, Adelstein, Lee & Ellis 2003
- **Predictive tracking, predictive compensation**
  Azuma & Bishop, 1994, Jung, Adelstein & Ellis 2000

Why Psychophysics for/in VE?

Quantify perceptual tolerances that are relevant to Virtual Environment (VE) system use

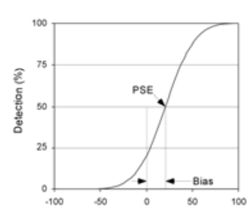
- Establish guidelines and specifications for the design, implementation, and effective deployment of VE systems and interfaces

Ultimately, to use appropriately implemented and well calibrated VE systems to rapidly prototype psychophysical (and other performance) studies

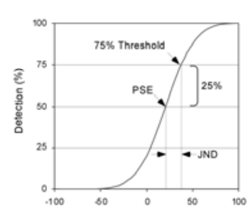
We want to measure human performance, not system artifact!

Psychometric Function

Features of the ogive



Point of Subjective Equality (PSE) and bias with respect to reference stimulus



Just Noticeable Difference (JND) for psychometric function *symmetric* about PSE

Psychometric Function

Features of the ogive

- Point of Subjective Equality/Equivalence (PSE)
  - Bias in observer's response
  - Criterion dependent
  - Question posed as a source of bias
- Just Noticeable Difference (JND)
  - Generally defined by ½ of stimulus difference between 1st and 3rd detection quartiles
  - For symmetric functions, the amount of additional stimulus difference to increase detection by 25% from PSE
  - JND is related to variance and is therefore a statistical measure of detectability

Summary Comments on Methods

Method choice should depend on objectives

Use Method of Constant Stimuli first, when have insufficient knowledge of detection capacity

- Measure *d*-prime and FA rates
- Time-consuming (inefficient)

Method of Limits w/ U-D Adapting Staircases

- Can select U-D ratio to concentrate data in region of interest
  - Efficient (fewer observations than Constant Stimuli)
- Does not measure FA rate
  - Has a prescribed *d´* for given *M*-alternative forced choice

Caveat: Pure perception experiments may be far removed from ecological experience—i.e., detached from realistic action and task performance

Project's goals

- **Measure perceptual thresholds to latency using psychophysics (interleaved staircases, method of limits)**
- **Determine generality of results in plain environments (with or without shear of motion of an object against a background) and in 'meaningful' spaces**

Psychophysics 1

- **Psychophysics is the scientific study of the relation between stimulus (φ) and sensation (ψ)**
- **A *difference threshold* (DL) is defined as the amount of change in a stimulus (Δφ) required to produce a *Just Noticeable Difference* (JND) in sensation**
- **Basic method: We present a series of stimuli to participants and we ask them to report whether they perceive the stimulus presented or not**

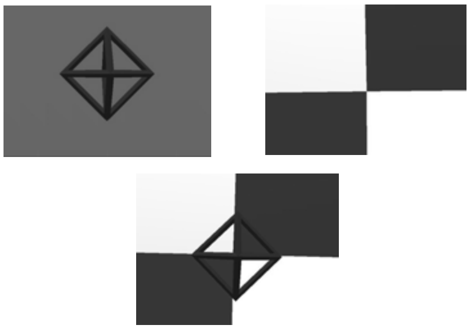
Psychophysics 2

- **Preliminary observations are made for locating the range of values always perceived and seldom perceived**
- **During the main experiment, a count of 'different' (stimuli detected) or 'same' (stimulus not detected) is kept**
- **For each stimulus value, the proportion of 'different' responses is computed**
- **Stimulus intensity is often plotted in the x-axis and the proportion of yes responses on the y-axis**
- **A curve is subsequently fitted to the plotted data points. If enough measurements are made, psychometric functions often follow a particular S shape called an *ogive* - Best fit Gaussian**

Method of Limits (Staircase method)

- **The experimenter begins by presenting a sequence of stimuli that progressively increase or decrease in value (ascending or descending series)**
- **When the observer's response changes, the stimulus value is recorded and the direction of the stimulus sequence is reversed from ascending to descending or ascending and vice versa**
- **This procedure is continued until a sufficient number of response transition points have been recorded**

Visual Conditions – Shear of motion



Visual Conditions - Radiosity

Set-up

Experimental method

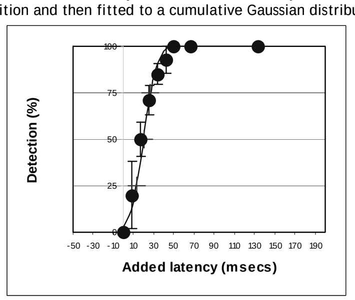
- Latency conditions were presented in sequential pairs, one being a reference (R) and the other the probe level (P) composed of the reference plus an added latency, in 8.5 msecs steps
- 6 staircases per visual condition (three ascending and three descending) with each set of two staircases being interleaved to prevent subjects' prediction (9 sections, 3 sections per visual condition)
- Descending staircases began with a pair comprising the stable R level and a 133 msecs probe. Ascending staircases started with a pair comprising the stable R level and a probe level of the minimum system latency (equal to the R level in this case)

Results 1

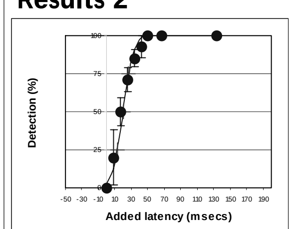
The accumulated data were compiled into detection rate versus added latency (the amount added above the system's minimum latency of 10 msecs) for each visual condition and then fitted to a cumulative Gaussian distribution

JND 6.8 ms

PSE 19.9

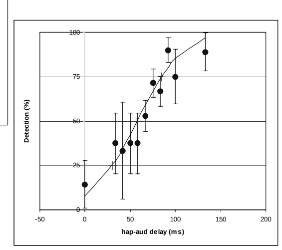


Results 2



JND 27.2

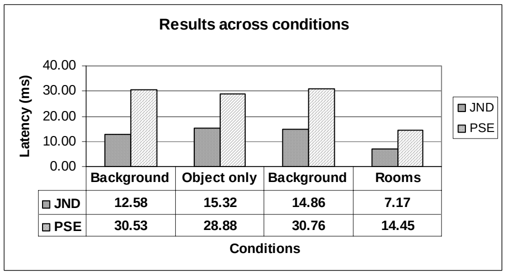
PSE 57.7



JND 6.8 ms

PSE 19.9

Results across visual conditions



	Background	Object only	Background	Rooms
JND	12.58	15.32	14.86	7.17
PSE	30.53	28.88	30.76	14.45

Conclusions

- There is no such thing as a cut-off threshold (JND), there are only conditional probabilities of detection, due to variation of human's responses
- A JND of ~15 ms is the detectable change (DL) after PSE (chance level), up to the 75% detection point
- Or if the intensity of latency is PSE, latency is increased to PSE+JND at the point of detection probability 75%
- Or JND is the amount of stimulus change that would be noticeable after chance level (PSE) at the convention defined 75% detection probability

Conclusions

- Previous studies revaled identical thresholds for different base latencies meaning that users of long latency VE systems will be as sensitive to changes in latency as those who use prompter systems

- Virtual Reality system designers should expect users to generally be able to notice changes in latency, when the change is around ~15 ms (but it could be even less dependent on scene context)

Acknowledgments

Nick Mourkoussis, Bernard Adelstein, Stephen R. Ellis, Douglas Cunningham, Bernard Riecke, Betty Mohler, Fiona Rivera