Siggarph ASIA 2011 Course:
Multidimensional Image Retargeting

Alessandro Artusi
CaSToRC Cyprus Institute

Tunç O. Aydın
Disney Research Zurich, Switzerland

Francesco Banterle*
Visual Computing Lab ISTI-CNR, Italy

Piotr Didyk
MPI Informatik, Germany

Elmar Eisemann
Télécom ParisTech / CNRS-LTCI, France

Diego Gutierrez
Universidad de Zaragoza, Spain

Rafał Mantiuk
University of Bangor, UK

Karol Myszkowski
MPI Informatik, Germany

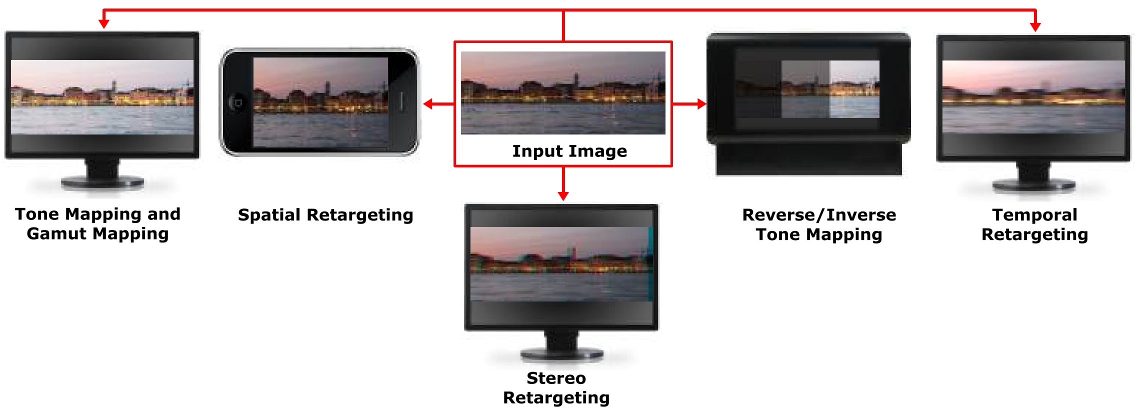


Figure 1: *The retargeting process; an input image/video is adapted to different displaying devices.*

Abstract

Retargeting is a process through which an image/a video is adapted from the display device for which it was meant (target display) to another one (retarget display). The retarget display has different features from the target one such as: dynamic range, discretization levels, color gamut, multi-view (3D), refresh rate, spatial resolution... This is a very relevant and hot topic in graphics, given the increasing number of display devices, from large, high-contrast screens to small cell phones with limited dynamic range; a lot of techniques are being published in different venues, and it's simply very hard to keep up.

For instance, one of the few cases for which retargeting can be potentially straightforward is when adapting images from a larger display (in term of resolution) to a smaller one with the same aspect ratio: a low-pass filter followed by downsampling can then achieve good quality results. However, for most cases retargeting can be an ill-posed problem, such as when displaying Low Dynamic Range (LDR) or 8-bit content on High Dynamic Range (HDR) displays. Such a problem requires the retargeting algorithm to generate new content which is missing in the input image/frame.

In this course, we will present the latest solutions and techniques for retargeting images along various dimensions such as dynamic range, colors, temporal and spatial resolutions, and offer for the first time a much-needed holistic view of the field. Moreover, we are going to show how to measure and analyze the changes applied to an image/video in terms of quality using both (subjective) psychophysical experiments and (objective) computational metrics.

The course should be of interest to anyone involved in graphics in its broader sense, given the almost unavoidable need to retarget results to different devices: from developer that are interested to implement retargeting techniques, to users that just need an overall perspective, for researchers fully engaged in developing multi-dimensional retargeting techniques, for whom this course will serve as a solid background for future algorithms.

*e-mail:frabante@gmail.com

1 Course General Information

1.1 Course Organizer

Francesco Banterle
e-mail: frabante@gmail.com or francesco.banterle@isti.cnr.it
Website: http//www.banterle.com/francesco
Address: Visual Computing Lab, ISTI-CNR, Via G. Moruzzin n.1, CAP 56124, Pisa, Italy
Telephone: +39 050 315 2921
Mobile phone: +39 338 1594 385
Fax: +39 050 315 2604

1.2 Lecturers

- Alessandro Artusi, CaSToRC Cyprus Institute, Cyprus
- Tunç O. Aydın, Disney Research Zurich, Switzerland
- Francesco Banterle, Visual Computing Lab ISTI-CNR, Italy
- Piotr Didyk, MPI Informatik, Germany
- Elmar Eisemann, Télécom ParisTech / CNRS-LTCI, France
- Diego Gutierrez, Universidad de Zaragoza, Spain
- Rafał Mantiuk, University of Bangor, UK
- Karol Myszkowski, MPI Informatik, Germany

1.3 Course Syllabus

Dynamic Range and Color Retargeting (Tone Mapping) (80 minutes)

1. The ingredients of tone mapping: (15 minutes, presenter): Rafał Mantiuk
   - Intent of tone mapping
   - LDR and HDR pixel values
   - Display models
   - The logarithmic domain and sensitivity to light
   - Algebra of tone mapping
2. Major Approaches to Tone Mapping (30 minutes, presenter): Rafał Mantiuk
   - Illumination and reflectance separation
   - Forward visual model
   - Forward and inverse visual model
   - Constraint mapping problem
3. Visual Illusions for Tone-Mapping (10 minutes, presenter): Rafał Mantiuk
   - Glare
   - Cornsweet illusion / countershading
4. Color Issues in Tone Mapping (25 minutes, presenter): Alessandro Artusi
   - Introduction to Color
   - Color is an Issue
   - Gamut Mapping
   - Color Correction

Reverse/Inverse Tone Mapping (60 minutes), Lecturer: Francesco Banterle

1. Increasing the Dynamic Range Problem: the problem to adapt legacy LDR content to HDR and high contrast displays (5 minutes)

2. Global Methods: dynamic range is increased applying a per pixel function (20 minutes)
   - Linear Models
   - Non-Linear Models
3. Local Methods: methods based on neighbors' information for expanding the dynamic range (20 minutes)
   - Frequency editing
   - Expand Maps
   - Selective methods
4. Evaluation and Conclusions (15 minutes)
   - Subjective Studies
   - Objective Studies

Image Spatial Resolution Retargeting (65 minutes) Lecturer: Diego Gutierrez

1. The Problem: definition of the context of main problem of image resizing (5 minutes)
2. Main Approaches:
   - Discrete algorithms: working on a representation of an image as a set of pixels or a graph (15 minutes)
   - Continuous algorithms: working on a representation of an image as a continuous signal (15 minutes)
   - A note on video resizing: (5 minutes)
3. Comparison of existing techniques: analysis of 8 of the most popular and recent resizing algorithms:
   - Subjective study: description and analysis of the results of a perceptual study (10 minutes)
   - Objective study: description and analysis of the results using 6 different existing metrics (10 minutes)
4. Conclusions (5 minutes)

Temporal Image Retargeting (70 minutes) Lecturers: Karol Myszkowski and Elmar Eisemann

1. Motivation (5 minutes)
   - Improvement of Perceived Quality: Reduction of Flickering, Hold-type Blur, and Judder Effect, as well as Enhancement of Motion Continuity, Gamut, and Spatial Resolution
   - Fixed Frame-rate Requirements
   - High-refresh Rate Displays
2. Human Visual System (HVS) Background (15 minutes)
   - Critical Flicker Frequency
   - Spatio-temporal Contrast Sensitivity
   - Temporal Integration in the Eye
   - Eye Movement Characteristics
3. Hold-type Blur Reduction (15 minutes)
   - Modern TV-sets: Backlight Flashing, Black Data Insertion, Blurred Frame Insertion, Frame Rate Doubling, Motion Compensated Inverse Filtering, In-between Frame Derivation based on Optical Flow
   - Rendering: Frame Warping and In-between Frame Insertion, Warping Artifacts Reduction, Interleaving Sharpened and Blurred Frames (exploiting temporal integration)
   - Discussion: Display Requirements (Temporal Response, Backlighting, Sync), Ghosting, Luminance Reduction, Flickering
4. Image Warping Techniques (15 minutes)
   - Per-pixel Methods
   - Mesh-based Methods
   - Edge-preserving Methods
   - Warping Artifact Reduction

5. Exploiting Temporal Integration for Image Enhancement (15 minutes)
    - Gamut Extension (Frame Rate Control)
    - Subimage Optimization
    - Flickering Reduction
6. Conclusions (5 minutes)

Image and Video Quality Assessment (70 minutes) Lecturer: Tunç O. Aydın

1. Introduction to quality assessment (10 minutes)
    - Subjective quality assessment through psychophysical experimentation
    - Objective quality assessment metrics
    - Comparison of objective and subjective methods
2. Image and video quality assessment metrics
    - Full-reference vs. no-reference quality assessment (5 minutes)
    - Various approaches to defining quality (10 minutes)
        - Visible Differences
        - Structural Similarity
        - Visual Equivalence
    - Image quality assessment metrics (20 minutes)
        - Grouping w.r.t. sophistication
        - Grouping w.r.t. task performed.
    - Video quality assessment metrics: what additional mechanisms are needed to handle video (10 minutes)
3. Calibration and validation experiments (10 minutes)
4. Conclusions (5 minutes)

Stereo Content Retargeting (60 minutes) Lecturers: Piotr Didyk

1. Motivation (5 minutes)
    - Stereo and Realism
    - 3D Display Devices
    - Stereo Retargeting
2. Depth Perception Background (10 minutes)
    - Depth Cues
    - Stereopsis
    - Visual Comfort
3. Stereo Content Adjustment (15 minutes)
    - Scene Adjustment
    - Disparity Mapping
    - Misperception
4. Perception-based Stereo Retargeting (20 minutes)
    - Disparity Model
    - Global Disparity Operators
    - Backward-compatible Stereo
    - Personalized Stereo
5. Disparity Metric (5 minutes)
6. Conclusions (5 minutes)

1.4 Intended Audience

This course is aimed at an audience interested in using and developing image/video re-targeting techniques for modern displays such as high dynamic range, high refresh rate, high resolution, mobile screens. Specifically display and mobile developers are the main target of this course. Nevertheless, many of the topics, such as adapting media from phones to tablets are becoming extremely important in many multimedia productions such as games, e-books, etc. Hence, we expect to draw the attention of people from many fields and believe that graduate students and researchers are likely to be interested in this course as well.

1.5 Prerequisites

For best results participants should have a basic understanding of the most commonly used video/image processing techniques, such as filtering (e.g. spatial and temporal), basic video/image operations (e.g. histograms manipulation, bilinear/nearest neighbors downsampling/upsampling), and very basic notions of high dynamic range imaging.

1.6 Level of Difficulty

Intermediate

2 Lecturers' Biographies

Alessandro Artusi
CaSToRC Cyprus Institute, Cyprus
email: artusialessandro4@googlemail.com
Dr. Alessandro Artusi is a Researcher at CaSToRC Cyprus Institute. He is working on High Dynamic Range Imaging, Image Processing applied on Computer Graphics, Colour Science and Visual Perception. He received a MSc in Computer Science from the University of Milan, Italy, in 1997 and a PhD in Computer Science (Computer Graphics) with distinction, from the Vienna University of Technology (VUT), Austria, in 2004. Dr. Artusi has covered several positions in several different academic institutions, and he also won an ERCIM European fellowship in the 2006. Recently he won a Ramon Cajal fellowship. He is author and co-author of one granted patent and five patents applications, and co-author of the book "Advanced High Dynamic Range Imaging Theory and Practice" edited by AK Peters (CRC Press) 2011. In 2009 he co-founded goHDR Ltd a spin-off company of the University of Warwick. In 2010, Dr. Artusi received the 1st prize Award at the International Entrepreneurship Competition held in Cyprus for the best business plan submitted on a research idea on High Dynamic Range Imaging. He is a member of the management committee of the European COST Action IC1005 "HDRi: The digital capture, storage, transmission and display of real-world lighting". He has served as Program Chair at VAST'09.

Tunç O. Aydın
Disney Research Zurich, Switzerland
email: tunc@disneyresearch.com
Tunç O. Aydın recently joined Disney Research Zurich as a Post-Doctoral Researcher. His main research interest lie in modelling various aspects of the human visual system, and applying these models to computer graphics and vision problems. He holds a PhD (summa cum laude) degree from the Computer Graphics Department of Max-Planck-Institut fr Informatik (2010), an M.S. degree from the College of Computing of Georgia Institute of Technology (2005), and a B.S degree from the Civil Engineering Department of Istanbul Technical University (2003). He also had a brief industry experience as a C++ developer.

Francesco Banterle
Visual Computing Lab ISTI-CNR, Italy
email: francesco.banterle@isti.cnr.it
Francesco Banterle is a post-doc researcher at the Visual Computing Laboratory at ISTI-CNR Italy . He received a Ph.D in Engineering from the International Digital Laboratory, WMG, University of Warwick in 2009. During his PhD he developed a new branch of High Dynamic Range (HDR) Imaging called Inverse Tone Mapping which bridges the gap between Low Dynamic Range Imaging and HDR Imaging. He holds a BSc (Magna cum Laude, 2004) and a MSc (Magna cum Laude, 2006) in Computer Science about Rendering from Verona University, Italy. During his doctorate he patented two patents as first author on the field of HDR Imaging. In 2009, he co-founded goHDR, a start-up company, where he developed the core technology. Before joining the Visual Computing he worked as intern at Arup Ltd , and for the University of Warwick where he developed new algorithms for HDR video cameras. He is first co-author of the book "Advanced High Dynamic Range" published by AK Peters in 2011 (CRC press). His main research fields are High Dynamic Range Imaging, Image Processing, Rendering, and Parallel Processing (GPUs and shared memory systems).

Piotr Didyk
MPI Informatik, Germany

email: pdidyk@mpi-inf.mpg.de

Piotr Didyk is a PhD student at MPI Informatik, Saarbrücken, Germany. Before joining the MPI, he received his M.Sc degree in Computer Science from University of Wrocław in Poland. In 2007, he was awarded with a fellowship award from "Polish Talents" organization supported by the Polish Academy of Science. In 2011, he worked as a visiting student at MIT. His work focuses on image quality enhancement as well as retargeting techniques in the context of new display technologies. He has developed techniques, which by exploiting properties of the human visual system, allow exceeding physical limitations of off-the-shelf displays. He has also contributed into the field of stereo-vision by developing a perceptual model for disparity.

Elmar Eisemann
Télécom ParisTech / CNRS-LTCI, France
email: elmar.eisemann@telecom-paristech.fr

Before being an associate professor at Telecom ParisTech, Elmar Eisemann was a senior scientist in the Cluster of Excellence (MMCI), Saarland University / MPI Informatik, Germany and head of the research group ECLEXIS until December 2009. He studied Mathematics at the University of Cologne and Computer Science at the Ecole Normale Superieure Paris (2001). He obtained Master (2004) and PhD. (2008) in Mathematics / Computer Science from Grenoble Universities. He worked abroad at MIT (2003), UIUC (2006), Adobe / Seattle (2007), and Adobe / Boston (2008). His interests include real-time rendering, shadow algorithms, global illumination, and GPU acceleration techniques. Together with Karol Myszkowski, he was the local organizer of EGSR 2010.

Diego Gutierrez
Universidad de Zaragoza, Spain
email: diegog@unizar.es

Diego Gutierrez is an Associate Professor at the Universidad de Zaragoza, in Spain, where he received his PhD in Computer Science. Hes published his research on physically based global illumination, perception and image processing techniques in top journals and conferences (including SIGGRAPH and Eurographics). He's currently Papers Chair for Applied Perception in Graphics and Visualization (APGV 2011), and has held other relevant positions such as Program Chair of SIGGRAPH Asia Sketches & Posters (2008), Papers Chair for ACM Graphite (2006), or Conference Chair for APGV 2010. Hes served on many other Program Committees, including SIGGRAPH Asia (2009) and Eurographics (2007, 2010, 2011). He's also an Associate Editor of three journals (IEEE Computer Graphics & Applications, ACM Transactions on Applied Perception and Computers & Graphics)

Rafałłc Mantiuk
University of Bangor, UK
email: mantiuk@bangor.ac.uk

Rafal Mantiuk is a lecturer (assistant professor) at Bangor University (UK) and a member of a Reasearch Institute of Visual Computing. Before comming to Bangor he received his PhD from the Max-Planck-Institute for Computer Science (2006, Germany) and was a postdoctoral researcher at the University of British Columbia (Canada). Rafal has published over 15 journal papers, including ACM SIGGRAPH/ACM Trans. on Graphics, Eurographics & EGSR/Computer Graphics Forum and IEEE Trans. on Image Processing, applied for several patents and was recognized by the Heinz Billing Award (2006). He is co-chair of the High Dynamic Range Area program at Eurographics 2011. Rafal Mantiuk investigates how the knowledge of the human visual system and perception can be incorporated within computer graphics and imaging algorithms. His recent interests focus on designing imaging algorithms that adapt to human visual performance and viewing conditions in order to deliver the best images given limited resources, such as computation time or display contrast.

Karol Myszkowski
MPI Informatik, Germany
email: karol@mpi-inf.mpg.de

Karol Myszkowski is a tenured senior researcher at the MPI Informatik, Saarbrucken, Germany. From 1993 to 2000 he served as an Associate Professor in the Department of Computer Software at the University of Aizu, Japan. In the period 19861992 he worked for Integra, Inc. a Japanbased company, developing rendering software for customers such as Toshiba Lighting, Shiseido, Matsushita Electric, Kandenko, and others. He received his PhD. (1991) and habilitation (2001) degrees in computer science from Warsaw University of Technology (Poland). His research interests include perception issues in graphics, high-dynamic range imaging, global illumination and rendering. Karol published and lectured on these topics widely including ACM Siggraph Courses in 2000, 2001, 2003, and 2006. He also co-chaired the Eurographics Rendering Symposium in 2001, the ACM Symposium on Applied Perception in Graphics and Visualization in 2008, and the Spring Conference on Computer Graphics 2008.

References

ADAMS, A., GELFAND, N., DOLSON, J., AND LEVOY, M. 2009. Gaussian kd-trees for fast high-dimensional filtering. *ACM Transactions on Graphics (TOG) 28*, 3, 1–12.

Adams, A., Baek, J., and Davis, M. 2010. Fast High-Dimensional Filtering Using the Permutohedral Lattice. *Computer Graphics Forum 29*, 2, 753–762.

Akyüz, A. O., Fleming, R., Riecke, B. E., Reinhard, E., and Bülthoff, H. H. 2007. Do hdr displays support ldr content?: a psychophysical evaluation. *ACM Trans. Graph. 26*, 3, 38.

Anstis, S. M., and Howard, I. P. 1978. A Craik-O'Brien-Cornsweet illusion for visual depth. *Vision Res.*, 18, 213–217.

Artamonov, O., 2004. X-bit's guide: Contemporary lcd monitor parameters and characteristics. http://www.xbitlabs.com/articles/monitors/display/lcd-guide_11.html, October.

Avidan, S., and Shamir, A. 2007. Seam carving for content-aware image resizing. *ACM Trans. Graph. 26* (July).

Aydin, T. O., Mantiuk, R., Myszkowski, K., and Seidel, H.-P. 2008. Dynamic range independent image quality assessment. *ACM Trans. Graph. 27*, 3, 1–10.

Aydin, T. O., Čadík, M., Myszkowski, K., and Seidel, H.-P. 2010. Video quality assessment for computer graphics applications. *ACM Trans. Graph. 29* (December), 161:1–161:12.

Banterle, F., Ledda, P., Debattista, K., and Chalmers, A. 2006. Inverse tone mapping. In *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, ACM, New York, NY, USA, 349–356.

Banterle, F., Ledda, P., Debattista, K., Chalmers, A., and Bloj, M. 2007. A framework for inverse tone mapping. *The Visual Computer 23*, 7, 467–478.

Banterle, F., Ledda, P., Debattista, K., and Chalmers, A. 2008. Expanding low dynamic range videos for high dynamic range applications. In *SCCG '08: Proceedings of the 4th Spring Conference on Computer Graphics*, ACM, New York, NY, USA, 349–356.

Banterle, F., Ledda, P., Debattista, K., Artusi, A., Bloj, M., and Chalmers, A. 2009. A psychophysical evaluation of inverse tone mapping techniques. *Computer Graphics Forum 28*, 1 (March), 13–25.

Barten, P. G. 1999. *Contrast sensitivity of the human eye and its effects on image quality*. SPIE – The International Society for Optical Engineering.

Berns, R. S., Motta, R. J., and Gorzynski, M. E. 1993. Crt colorimetry. parti:theory and parctice. *Color Research & Applications Journal 18*, 299–314.

1999. Gamut Mapping for Pictorial Images, IS&T - The Society for Imaging Science and Technology.

1999. General-Purpose Gamut-Mapping Algorithms: Evaluation of Contrast-Preserving Rescaling Functions for Color Gamut Mapping, IS&T - The Society for Imaging Science and Technology.

Bradshaw, M. F., and Rogers, B. J. 1999. Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Res. 39*, 18, 3049–56.

Brookes, A., and Stevens, K. 1989. The analogy between stereo depth and brightness. *Perception 18*, 5, 601–614.

Burr, D. 1987. Implications of the Craik-O'Brien illusion for brightness perception. *Vision Resarch 27*, 11, 1903–1913.

Calabria, A. J., and Fairchild, M. D. 2003. Perceived image contrast and observer preference: I. the effects of lightness, chroma, and sharpness manipulations on contrast perception. *The Journal of imaging science and technology 47*, 6, 479–493.

Chen, H., Kim, S.-S., Lee, S.-H., Kwon, O.-J., and Sung, J.-H. 2005. Nonlinearity compensated smooth frame insertion for motion-blur reduction in LCD. In *Proc. Multimedia Signal Processing, 2005 IEEE 7th Workshop on*, 1–4.

Chen, J., Paris, S., and Durand, F. 2007. Real-time edge-aware image processing with the bilateral grid. *ACM Transactions on Graphics (TOG) 26*, 3, 103.

Chiu, K., Herf, M., Shirley, P., Swamy, S., Wang, C., and Zimmerman, K. 1993. Spatially nonuniform scaling functions for high contrast images. In *Graphics Interface*, Citeseer, 245–253.

Cutting, J., and Vishton, P. 1995. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of Space and Motion (Handbook Of Perception And Cognition)*, Academic Press, W. Epstein and S. Rogers, Eds., 69–117.

Daly, S. 1993. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision*, 179.

Daly, S. 1993. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision*, 179–206.

DEELEY, R., DRASDO, N., AND CHARMAN, W. 1991. A simple parametric model of the human ocular modulation transfer function. *Ophthalmic and Physiological Optics 11*, 1, 91–93.

DIDYK, P., MANTIUK, R., HEIN, M., AND SEIDEL, H.-P. 2008. Enhancement of bright video features for HDR displays. In *Proceeding of Eurographics Symposium on Rendering 2008*, Eurographics, Blackwell Ltd, Computer Graphics Forum.

DIDYK, P., EISEMANN, E., RITSCHEL, T., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. Apparent display resolution enhancement for moving images. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2010, Los Angeles) 29*, 3.

DIDYK, P., EISEMANN, E., RITSCHEL, T., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. Apparent display resolution enhancement for moving images. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2010, Los Angeles) 29*, 4.

DIDYK, P., EISEMANN, E., RITSCHEL, T., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. Perceptually-motivated real-time temporal upsampling of 3D content for high-refresh-rate displays. *Computer Graphics Forum (Proceedings Eurographics 2010, Norrköpping, Sweden) 29*, 2, 713–722.

DIDYK, P., RITSCHEL, T., EISEMAN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. Adaptive image-based stereo view synthesis. In *Proc. VMV*.

DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. Adaptive image-space stereo view synthesis. In *Vision, Modeling and Visualization Workshop*, 299–306.

DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. A perceptual model for disparity. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2011, Vancouver) 30*, 4.

DOOLEY, R. P., AND GREENFIELD, M. I. 1977. Measurements of edge-induced visual contrast and a spatial-frequency interaction of the cornsweet illusion. *Journal of the Optical Society of America 67*.

DURAND, F., AND DORSEY, J. 2002. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics 21*, 3 (July).

DZMURA, M., AND LENNIE, P. 1986. Mechanisms of color constancy. *Journal of the Optical Society of America A 3*, 10, 1662–1672.

FEDOROVSKAYA, E., DERIDDER, H., AND BLOMMAERT, F. 1997. Chroma variations and perceived quality of color images of natural scenes. *Color Research and Application 22*, 2, 96–110.

FENG, X.-F., PAN, H., AND DALY, S. 2008. Comparisons of motion-blur assessment strategies for newly emergent LCD and backlight driving technologies. *Journal of the Society for Information Display 16*, 981–988.

FENG, X.-F. 2006. LCD motion blur analysis, perception, and reduction using synchronized backlight flashing. In *Human Vision and Electronic Imaging XI*, SPIE, vol. 6057, M1–14.

GOODMAN, J. 2005. *Introduction To Fourier Optics*. Roberts & Co.

GOREA, A., AND TYLER, C. W. 1986. New look at Bloch's law for contrast. *Journal of the Optical Society of America A 3*, 1, 52–61.

GREEN, P. 2000. Defining Colour Gamut Boundaries with a Test Target.

HERZOG, R., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. Spatio-temporal upsampling on the GPU. In *Proceedings of ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, 91–98.

HOFFMAN, D., GIRSHICK, A., AKELEY, K., AND BANKS, M. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision 8*, 3, 1–30.

HOWARD, I. P., AND ROGERS, B. J. 2002. *Seeing in Depth*, vol. 2: Depth Perception. I. Porteous, Toronto.

HULLIN, M. B., EISEMANN, E., SEIDEL, H.-P., AND LEE, S. 2011. Physically-based real-time lens flare rendering. *ACM Trans. Graph. (Proc. SIGGRAPH 2011) 30*, 4, 108:1–108:9.

JANSSEN, R. 2001. *Computational Image Quality*. SPIE Press.

KAKIMOTO, M., MATSUOKA, K., NISHITA, T., NAEMURA, T., AND HARASHIMA, H. 2005. Glare generation based on wave optics. *Computer Graphics Forum 24*, 2, 185–193.

KAKIMOTO, M., MATSUOKA, K., NISHITA, T., NAEMURA, T., AND HARASHIMA, H. 2005. Glare simulation and its application to evaluation of bright lights with spectral power distribution. In *ACM SIGGRAPH 2005 Posters*, ACM, New York, NY, USA, ACM, 42.

KALLONIATIS, M., AND LUU, C., 2009. Temporal resolution.

KARNI, Z., FREEDMAN, D., AND GOTSMAN, C. 2009. Energy-based image deformation. In *Proceedings of the Symposium on Geometry Processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SGP '09, 1257–1268.

KAWASE, M. 2005. Practical implementation of high dynamic range rendering. In *Game Developers Conference*.

KEIL, M. S. 2007. Gradient representations and the perception of luminosity. *Vision research 47*, 27, 3360–3372.

KINGDOM, F., AND MOULDEN, B. 1988. Border effects on brightness: a review of findings, models and issues. *Spatial Vision 3*, 4, 225–262.

KINGDOM, F. A. A., AND WHITTLE, P. 1996. Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing. *Vision Research 36*, 6, 817–829.

KLOMPENHOUWER, M. A., AND VELTHOVEN, L. J. 2004. Motion blur reduction for liquid crystal displays: Motion-compensated inverse filtering. In *Proceedings of SPIE*, vol. 5308, 690.

KOVALESKI, R. P., AND OLIVEIRA, M. M. 2009. High-quality brightness enhancement functions for real-time reverse tone mapping. *Vis. Comput. 25* (April), 539–547.

KRÄHENBÜHL, P., LANG, M., HORNUNG, A., AND GROSS, M. 2009. A system for retargeting of streaming video. In *ACM SIGGRAPH Asia 2009 papers*, ACM, New York, NY, USA, SIGGRAPH Asia '09, 126:1–126:10.

KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2007. Contrast restoration by adaptive countershading. *Computer Graphics Forum (Proc. Eurographics 2007) 26*, 3, 581–590.

KURITA, T. 2001. Moving picture quality improvement for hold-type AM-LCDs. In *Society for Information Display (SID) '01*, 986–989.

KŘIVÁNEK, J., FERWERDA, J. A., AND BALA, K. 2010. Effects of global illumination approximations on material appearance. In *ACM SIGGRAPH 2010 papers*, ACM, New York, NY, USA, SIGGRAPH '10, 112:1–112:10.

LANDIS, H. 2002. Production-ready global illumination. In *Siggraph Course Notes 16*.

LANG, M., HORNUNG, A., WANG, O., POULAKOS, S., SMOLIC, A., AND GROSS, M. 2010. Nonlinear disparity mapping for stereoscopic 3D. *ACM Trans. Graph. (Proc. SIGGRAPH) 29*, 4, 751–760.

LIN, W., GAI, Y., AND KASSIM, A. 2006. Perceptual impact of edge sharpness in images. *Vision, Image and Signal Processing, IEE Proceedings 152*, 2 (April), 215–223.

LIU, F., GLEICHER, M., JIN, H., AND AGARWALA, A. 2009. Content-preserving warps for 3D video stabilization. *ACM Transaction on Graphics (Proceedings SIGGRAPH) 28*.

LIVINGSTONE, M. 2002. *Vision and Art: The Biology of Seeing*. Harry N. Abrams.

LUBIN, J. 1995. A visual discrimination model for imaging system design and development. In *Vision models for target detection and recognition*, World Scientific, P. E., Ed., 245–283.

LUEBKE, D., WATSON, B., COHEN, J. D., REDDY, M., AND VARSHNEY, A. 2002. *Level of Detail for 3D Graphics*. Elsevier Science Inc., New York, NY, USA.

LUFT, T., COLDITZ, C., AND DEUSSEN, O. 2006. Image enhancement by unsharp masking the depth buffer. *ACM Transactions on Graphics 25*, 3, 1206–1213.

LUNN, P., AND MORGAN, M. 1995. The analogy between stereo depth and brightness: a reexamination. *Perception 24*, 8, 901–4.

1993. Gamut Mapping in Perceptual Colour Space, IS&T - The Society for Imaging Science and Technology.

MACKAY, D. M. 1973. Lateral interaction between neural channels sensitive to texture density? *Nature 245*, 5421, 159–161.

MAHAJAN, D., HUANG, F.-C., MATUSIK, W., RAMAMOORTHI, R., AND BELHUMEUR, P. 2009. Moving gradients: a path-based method for plausible image interpolation. *ACM Trans. Graph. 28* (July), 42:1–42:11.

MAHAJAN, D., HUANG, F.-C., MATUSIK, W., RAMAMOORTHI, R., AND BELHUMEUR, P. 2009. Moving gradients: A path-based method for plausible image interpolation. *ACM Transaction on Graphics (Proceedings SIGGRAPH '09) 28*, 3.

MÄKELÄ, P., ROVAMO, J., AND WHITAKER, D. 1994. Effects of luminance and external temporal noise on flicker sensitivity as a function of stimulus size at various eccentricities. *Vision Research 34*, 15, 1981–91.

MANTIUK, R., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2004. Perception-motivated high dynamic range video encoding. *ACM Transactions on Graphics (Proc. of SIGGRAPH) 23*, 3 (Aug.), 733.

MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2004. Visible difference predicator for high dynamic range images. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, 2763–2769.

MANTIUK, R., DALY, S., MYSZKOWSKI, K., AND SEIDEL, H. 2005. Predicting visible differences in high dynamic range images: model and its calibration. *Proc. SPIE*.

MANTIUK, R., DALY, S., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2005. Predicting visible differences in high dynamic range images - model and its calibration. In *Human Vision and Electronic Imaging X, IST SPIE's 17th Annual Symposium on Electronic Imaging*, B. E. Rogowitz, T. N. Pappas, and S. J. Daly, Eds., vol. 5666, 204–214.

MANTIUK, R., MANTIUK, R., TOMASZWESKA, A., AND HEIDRICH, W. 2009. Color correction for tone mapping. *Prooceedings of Eurographics 2009 28*, 2.

MARK, W. R., MCMILLAN, L., AND BISHOP, G. 1997. Post-rendering 3D warping. In *Proceedings of the 1997 symposium on Interactive 3D graphics*, 7–16.

MASIA, B., AGUSTIN, S., FLEMING, R. W., SORKINE, O., AND GUTIERREZ, D. 2009. Evaluation of reverse tone mapping through varying exposure conditions. *ACM Trans. Graph. 28*, 5, 1–8.

MCCANN, J. J. 2005. Do humans discount the illuminant? In *Proceedings of SPIE*, SPIE, vol. 5666, 5666–9.

MEYLAN, L., DALY, S., AND SÜSSTRUNK, S. 2006. The Reproduction of Specular Highlights on High Dynamic Range Displays. In *IST/SID 14th Color Imaging Conference*.

MEYLAN, L., DALY, S., AND SÜSSTRUNK, S. 2007. Tone Mapping For High Dynamic Range Displays. In *Electronic Imaging*, vol. 6492.

MOON, P., AND SPENCER, D. E. 1944. On the Stiles-Crawford Effect. *Journal of the Optical Society of America 34*, 6 (June), 319–329.

MOROVIC, J., AND R, L. M. 2000. Calculating medium and image gamut boundaries for gamut mapping. *Color Research & Applications Journal 25*, 394–401.

MOROVIC, J. 1998. *To Develop a Universal Gamut Mapping Algorithm*.

NAKAMAE, E., KANEDA, K., OKAMOTO, T., AND NISHITA, T. 1990. A lighting model aiming at drive simulators. In *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, vol. 24, 395–404.

NEHAB, D. F., SANDER, P. V., LAWRENCE, J., TATARCHUK, N., AND ISIDORO, J. 2007. Accelerating real-time shading with reverse reprojection caching. In *Proceedings of the 22nd ACM SIGGRAPH/EUROGRAPHICS symposium on Graphics hardware*, Eurographics Association, 25–35.

2004. Interactive Perception Based Model for Characterization of Display device, IS&T - The Society for Imaging Science and Technology.

NEUMANN, A., ARTUSI, A., ZOTTI, G., NEUMANN, L., AND PURGATHOFER, W. 2007. Accurate display gamma functions based on human observation. *Color Research & Applications Journal 32*, 310–319.

NICOLAS BONNIER, FRANCIS SCHMITT, M. H., AND LEYNADIER, C. 2007. Spatial and color adaptive gamut mapping: A mathematical framework and two new algorithms. *Proc. of the 15th Color Imaging Conference*, 267–272.

OPPENHEIM, A., SCHAFER, R., AND STOCKHAM, T. 1968. Nonlinear filtering of multiplied and convolved signals. *Proceedings of the IEEE 56*, 8, 1264–1291.

PAJAK, D., HERZOG, R., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2011. Scalable remote rendering with depth and motion-flow augmented streaming. *Computer Graphics Forum 30*, 2, 415–424.

PALMER, S. E. 1999. *Vision Science: Photons to Phenomenology*. The MIT Press.

PAN, H., FENG, X.-F., AND DALY, S. 2005. LCD motion blur modeling and analysis. In *Proc. ICIP*, 21–24.

PARIS, S., AND DURAND, F. 2009. A fast approximation of the bilateral filter using a signal processing approach. *International journal of computer vision 81*, 1, 24–52.

PÉREZ, P., GANGNET, M., AND BLAKE, A. 2003. Poisson image editing. *ACM Transaction on Graphics 22*, 3, 313–318.

POYNTON, C. 1993. ”gamma” and its disguises: The nonlinear mappings of intensity in perception, crts, film and video. *SMPTE Journal 102*, 1099–1108.

PRATT, W. K. 1991. *Digital image processing (2nd ed.)*. John Wiley & Sons, Inc., New York, USA.

PRITCH, Y., KAV-VENAKI, E., AND PELEG, S. 2009. Shift-map image editing. In *ICCV'09*, 151–158.

PURVES, D., SHIMPI, A., AND LOTTO, R. 1999. An empirical explanation of the Cornsweet effect. *The Journal of Neuroscience 19*, 19, 8542–8551.

RAMANARAYANAN, G., FERWERDA, J., WALTER, B., AND BALA, K. 2007. Visual Equivalence: Towards a new standard for Image Fidelity. *ACM Transactions on Graphics (Proc. of SIGGRAPH) 26*, 3. Article 76.

RATLIFF, F. 1971. Contour and contrast. *Proceedings of the American Philosophical Society 115*, 2, 150–163.

REMPEL, A. G., TRENTACOSTE, M., SEETZEN, H., YOUNG, H. D., HEIDRICH, W., WHITEHEAD, L., AND WARD, G. 2007. Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs. *ACM Trans. Graph. 26*, 3, 39.

RITSCHEL, T., SMITH, K., IHRKE, M., GROSCH, T., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2008. 3D unsharp masking for scene coherent enhancement. *ACM Transactions on Graphics (Proc. SIGGRAPH) 27* (August), 90:1–90:8.

RITSCHEL, T., IHRKE, M., FRISVAD, J. R., COPPENS, J., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2009. Temporal glare: Real-time dynamic simulation of the scattering in the human eye. *Computer Graphics Forum (Proc. EUROGRAPHICS 2009) 28*, 3 (March), 183–192.

ROGERS, B., AND GRAHAM, M. 1983. Anisotropies in the perception of three-dimensional surfaces. *Science 221*, 4618, 1409–11.

ROKITA, P. 1993. A model for rendering high intensity lights. *Computers & graphics 17*, 4, 431–437.

RUBINSTEIN, M., SHAMIR, A., AND AVIDAN, S. 2008. Improved seam carving for video retargeting. *ACM Trans. Graph. 27* (August), 16:1–16:9.

RUBINSTEIN, M., SHAMIR, A., AND AVIDAN, S. 2009. Multi-operator media retargeting. *ACM Trans. Graph. 28* (July), 23:1–23:11.

RUBINSTEIN, M., GUTIERREZ, D., SORKINE, O., AND SHAMIR, A. 2010. A comparative study of image retargeting. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia) 29*, 5.

SCHERZER, D., YANG, L., MATTAUSCH, O., NEHAB, D., SANDER, P. V., WIMMER, M., AND EISEMANN, E. 2011. A survey on temporal coherence methods in real-time rendering. In *In State of the Art Reports Eurographics. May 2010.*,

SESHADRINATHAN, K., AND BOVIK, A. 2007. A structural similarity metric for video based on motion models. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 1, I–869–I–872.

SHI, L., WANG, J., DUAN, L., AND LU, H. 2009. Consumer video retargeting: context assisted spatial-temporal grid optimization. In *Proceedings of the seventeen ACM international conference on Multimedia*, ACM, New York, NY, USA, MM '09, 301–310.

SIMPSON, G. 1953. Ocular haloes and coronas. *British Journal of Ophthalmology 37*, 8, 450–486.

SITTHI-AMORN, P., LAWRENCE, J., YANG, L., SANDER, P. V., NEHAB, D., AND XI, J. 2008. Automated reprojection-based pixel shader optimization.

SMITH, K., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Beyond tone mapping: Enhanced depiction of tone mapped HDR images. In *EUROGRAPHICS 2006 (EG'06)*, Blackwell, Vienna, Austria, E. Gröller and L. Szirmay-Kalos, Eds., vol. 25 of *Computer Graphics Forum*, Eurographics, 427–438.

SMITH, K., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Beyond tone mapping: Enhanced depiction of tone mapped HDR images. *Computer Graphics Forum (Proc. of EUROGRAPHICS) 25*, 3, 427–438.

SMITH, K., LANDES, P.-E., THOLLOT, J., AND MYSZKOWSKI, K. 2008. Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. *Computer Graphics Forum (Proc. EUROGRAPHICS) 27*, 2, 193–200.

SMITH, K. 2008. *Contours and Contrast*. PhD thesis, MPI Informatik, Saarbruecken, Germany.

SPENCER, G., SHIRLEY, P., ZIMMERMAN, K., AND GREENBERG, D. P. 1995. Physically-based glare effects for digital images. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 325–334.

STEVENS, J., AND STEVENS, S. 1963. Brightness function: Effects of adaptation. *JOSA*.

STICH, T., LINZ, C., WALLRAVEN, C., CUNNINGHAM, D., AND MAGNOR, M. 2011. Perception-motivated interpolation of image sequences. *ACM Transactions on Applied Perception (TAP) 8*, 2, 11:1–11:25.

TEMPLIN, K., DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2011. Apparent resolution enhancement for animations. In *27th Spring Conference on Computer Graphics*, 85–92.

TOMASI, C., AND MANDUCHI, R. 1998. Bilateral filtering for gray and color images. In *International Conference on Computer Vision*, Narosa Publishing House, 839–846.

TYLER, C. W. 1975. Spatial organization of binocular disparity sensitivity. *Vision Res. 15*, 5, 583 – 590.

van den Berg, T. J. T. P., Hagenouw, M. P. J., and Coppens, J. E. 2005. The ciliary corona: Physical model and simulation of the fine needles radiating from point light sources. *Investigative Ophthalmology and Visual Science 46*, 2627–2632.

Vos, J. J., and Boogaard, J. 1963. Contribution of the Cornea to Entoptic Scatter. *Journal of the Optical Society of America 53*, 7 (July), 869–873.

Walter, B., Drettakis, G., and Parker, S. 1999. Interactive rendering using the render cache. In *Proceedings of the 10th Eurographics Workshop on Rendering*, vol. 10, Citeseer, 235–246.

Wang, Z., and Bovik, A. C. 2002. A universal image quality index. *IEEE Signal Processing Letters 9*, 3 (March), 81–84.

Wang, Z., and Bovik, A. C. 2006. *Modern Image Quality Assessment*. Morgan & Claypool Publishers.

Wang, L., Wei, L.-Y., Zhou, K., Guo, B., and Shum, H.-Y. 2007. High dynamic range image hallucination. In *Proceedings of Eurographics Symposium on Rendering*.

Wang, Y.-S., Tai, C.-L., Sorkine, O., and Lee, T.-Y. 2008. Optimized scale-and-stretch for image resizing. In *ACM SIGGRAPH Asia 2008 papers*, ACM, New York, NY, USA, SIGGRAPH Asia '08, 118:1–118:8.

Wang, Y.-S., Fu, H., Sorkine, O., Lee, T.-Y., and Seidel, H.-P. 2009. Motion-aware temporal coherence for video resizing. *ACM Trans. Graph. 28* (December), 127:1–127:10.

Wang, Y.-S., Lin, H.-C., Sorkine, O., and Lee, T.-Y. 2010. Motion-based video retargeting with optimized crop-and-warp. *ACM Trans. Graph. 29* (July), 90:1–90:9.

Westheimer, G. 1986. The eye as an optical instrument. In *Handbook of Perception and Human Performance: 1. Sensory Processes and Perception*, K. Boff, L. Kaufman, and J. Thomas, Eds. Wiley, New York, 4.1–4.20.

Whittle, P. 1986. Increments and decrements: Luminance discrimination. *Vision Research 26*, 10, 1677–1691.

Wolberg, G. 1998. Image morphing: A survey. *The Visual Computer 14*, 8.

Wolf, L., Guttmann, M., and Cohen-Or, D. 2007. Non-homogeneous content-driven video-retargeting. In *Proceedings of the Eleventh IEEE International Conference on Computer Vision (ICCV-07)*.

Wüller, D., and Gabele, H. 2007. The usage of digital cameras as luminance meters. In *Proceedings of SPIE*, SPIE, vol. 6502, 65020U–65020U–11.

Yoshida, A., Mantiuk, R., Myszkowski, K., and Seidel, H.-P. 2006. Analysis of reproducing real-world appearance on displays of varying dynamic range. *Computer Graphics Forum (Proc. EUROGRAPHICS 2006) 25*, 3 (March), 415–426.

Yoshida, A., Ihrke, M., Mantiuk, R., and Seidel, H.-P. 2008. Brightness of the glare illusion. In *Proceedings of the ACM Symposium on Applied Perception in Graphics and Visualization*, ACM, 83–90.

Zavagno, D., and Caputo, G. 2001. The glare effect and the perception of luminosity. *Perception 30*, 2, 209–222.

Zavagno, D. 1999. Some new luminance-gradient effects. *Perception 28*, 835–838.

Multidimensional Image Retargeting

Alessandro Artusi
CaSToRC Cyprus Institute

Tunç O. Aydın
Disney Research Zurich, Switzerland

Francesco Banterle
Visual Computing Lab ISTI-CNR, Italy

Piotr Didyk
MPI Informatik, Germany

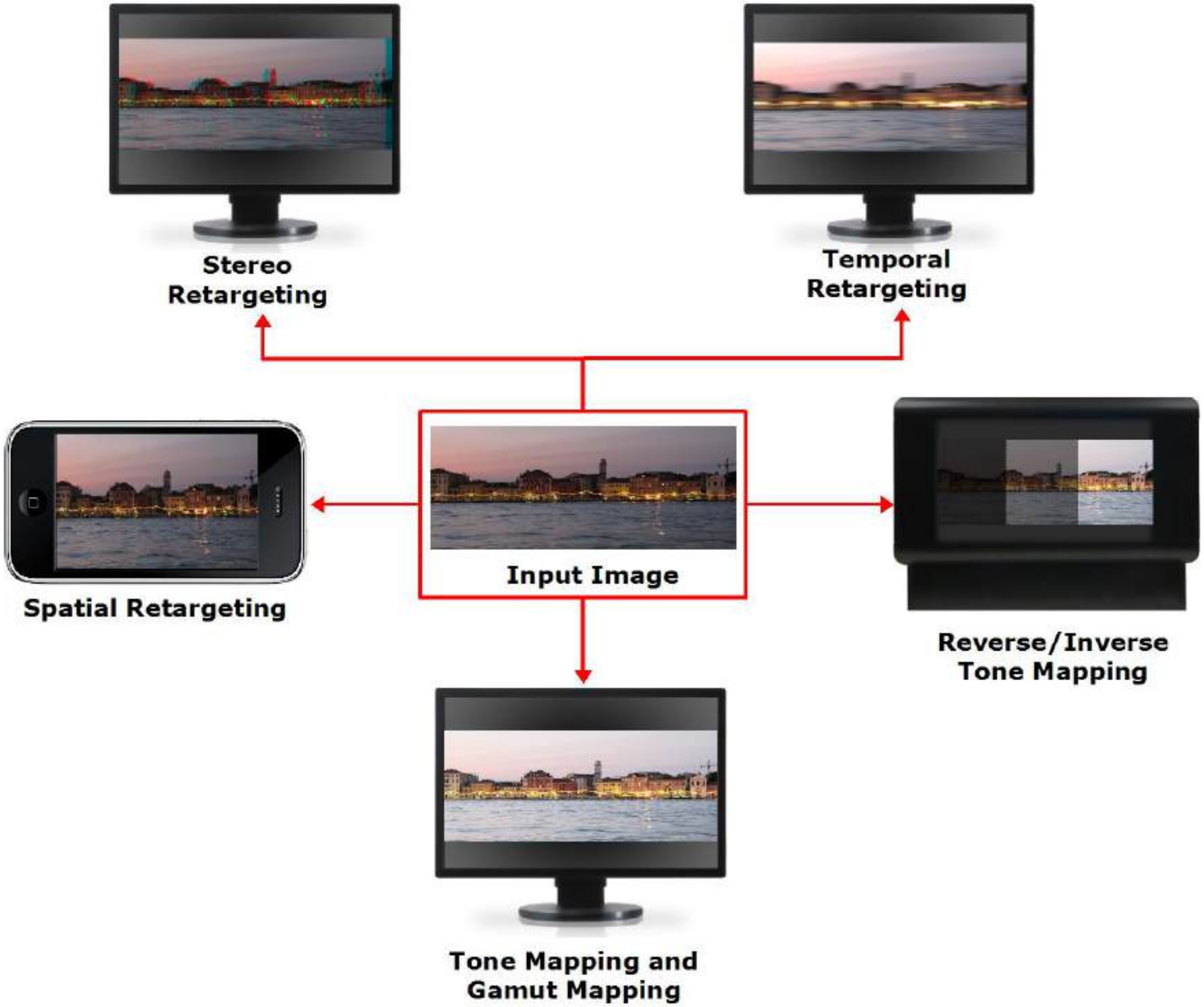
Elmar Eisemann
Telecom ParisTech / CNRS-LTCI, France

Diego Gutierrez
Universidad de Zaragoza, Spain

Rafał Mantiuk
University of Bangor, UK

Karol Myszkowski
MPI Informatik, Germany

Introduction: The Problem



Introduction: Outline

- **Dynamic Range and Color Retargeting (~80 mins):**
  - Rafał Mantiuk, Karol Myszkowski, and Alessandro Artusi
- **Reverse/Inverse Tone Mapping (~60 mins) :**
  - Francesco Banterle
- **Image Spatial Resolution Retargeting (~65 mins) :**
  - Diego Gutierrez
- **Temporal Image Retargeting (~70 mins) :**
  - Karol Myszkowski and Elmar Eisemann
- **Image and Video Quality Assessment(~70 mins) :**
  - Tunç O. Aydın
- **Stereo Content Retargeting (~60 mins):**
  - Piotr Didyk

Multidimensional retargeting: Tone Mapping

Dr. Rafal Mantiuk

• PRIFYSGOL CYMRU • UNIVERSITY OF WALES BANGOR 1884

Bangor University, UK

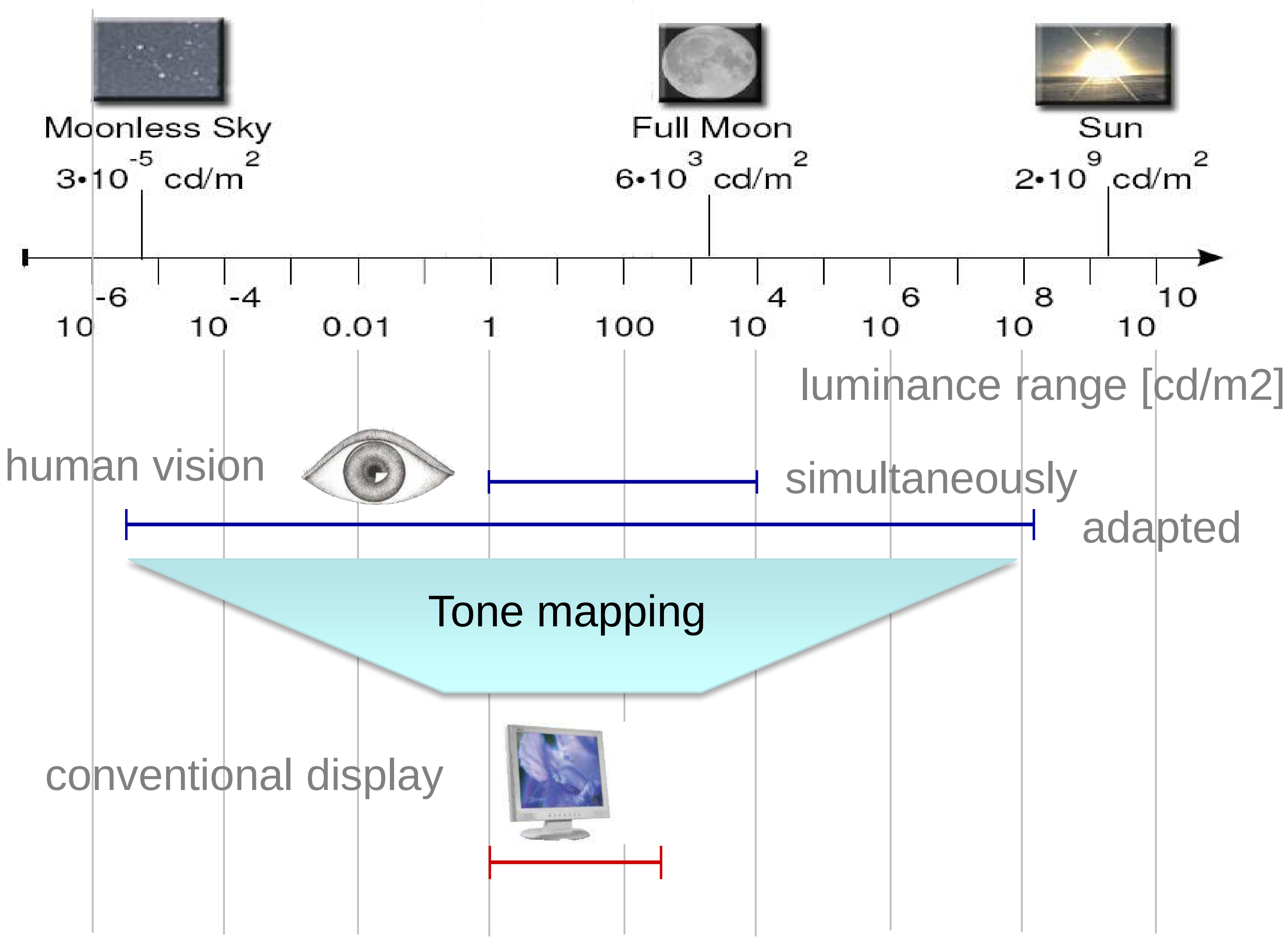
Research Institute of Visual Computing

http://www.bangor.ac.uk/mantiuk/

Check the latest version of the slides

- http://www.bangor.ac.uk/mantiuk/

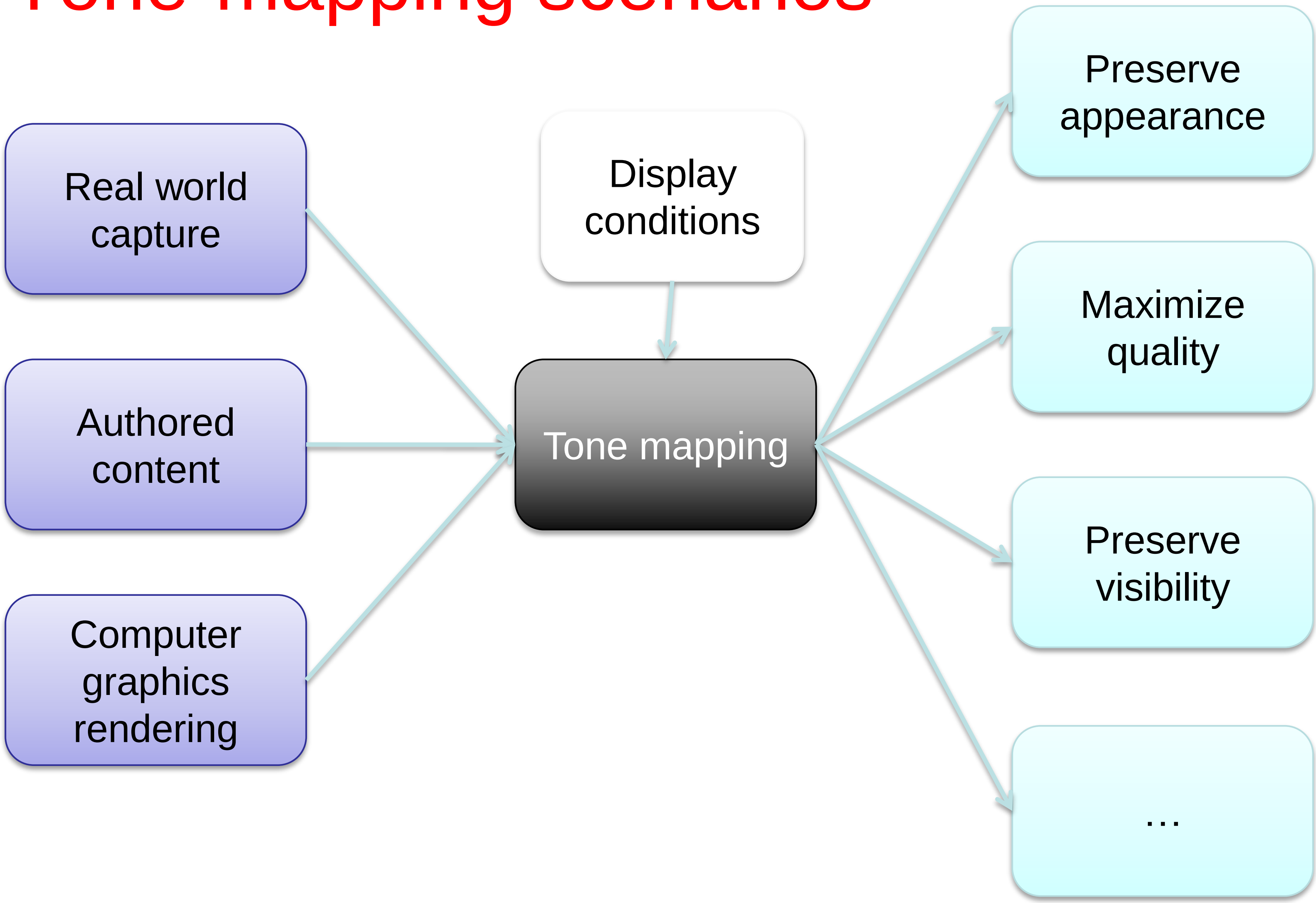
Tone-mapping problem



Tone Mapping?

- HDR ?
- Or something else ?

Tone-mapping scenarios



Input and output

- HDR
- (approximate) physical units
- luminance
- linear RGB
- scene-referred

- LDR (SDR)
- pixel values
- luma
- gamma corrected R’G’B’
- display referred

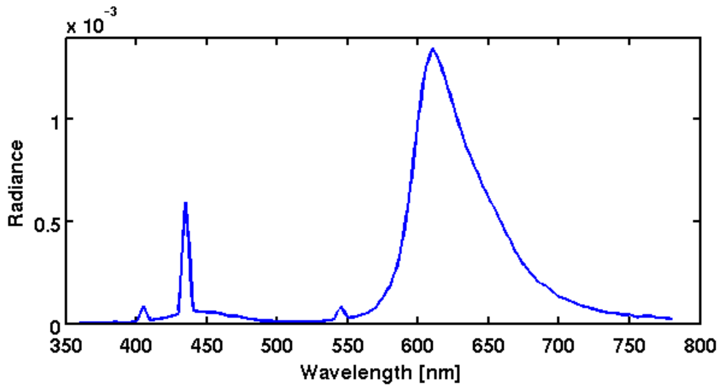
Luminance

- Luminance – perceived brightness of light, adjusted for the sensitivity of the visual system to wavelengths

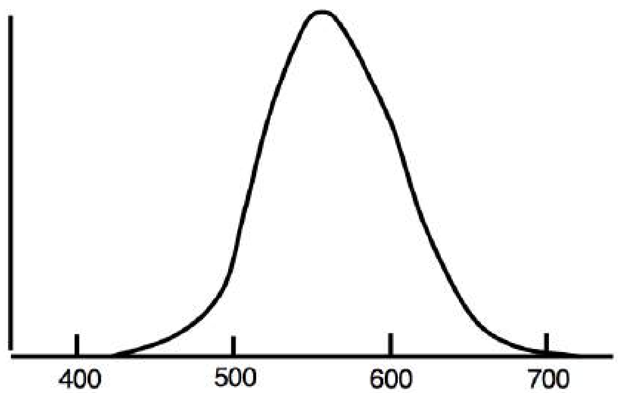
Luminance

$$L_V = \int_0^\infty L(\lambda) \cdot V(\lambda) d\lambda$$

Light spectrum (radiance)



Luminous efficiency function (weighting)



Do HDR images contain luminance values?

- Not exactly, because:
  - a) the combination of camera red, green and blue spectral sensitivity curves will not match the luminous efficiency function
  - b) the multi-exposure techniques do not capture absolute luminance values, only relative (luminance factor)
- But they contain a good-enough approximation for most applications
  - For multi-exposure camera capture the error in luminance measurements is 10-15%

Sensitivity to luminance

- Weber-law – the just-noticeable difference is proportional to the magnitude of a stimulus

Ernst Heinrich Weber
[From wikipedia]

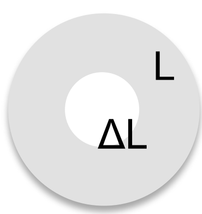
The smallest detectable luminance difference

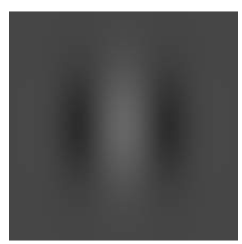
Background (adapting) luminance

$$\frac{\Delta L}{L} = k$$

Constant

Typical stimuli:





Consequence of the Weber-law

- Smallest detectable difference in luminance

$$\frac{\Delta L}{L} = k,$$

L	ΔL
100 cd/m^2	1 cd/m^2
1 cd/m^2	0.01 cd/m^2

- Adding or subtracting luminance will have different visual impact depending on the background luminance
- Unlike LDR luma values, HDR luminance values are not perceptually uniform!

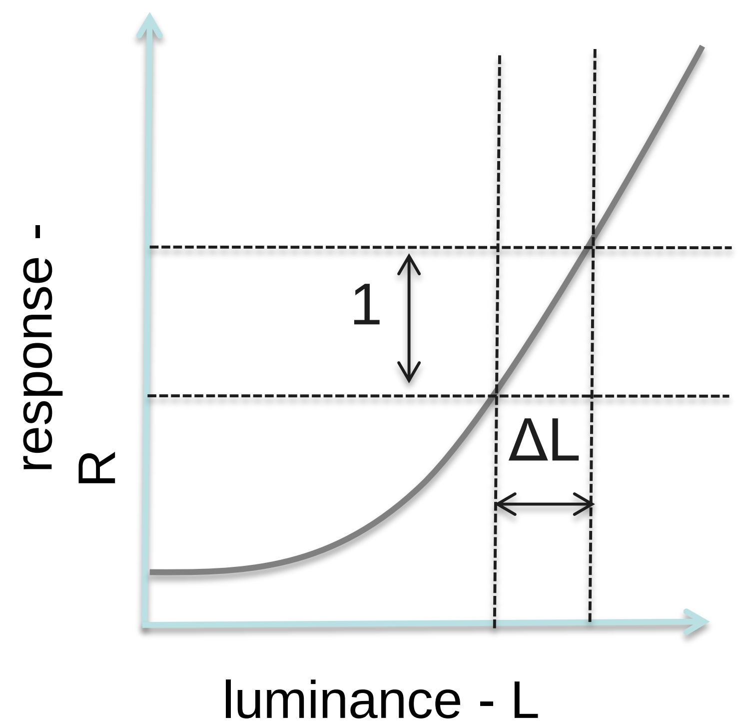
How to make luminance (more) perceptually uniform?

- Using Fechnerian integration

$$dR(L) = \frac{1}{\Delta L(L)}$$

Derivative of response

Detection threshold



Luminance transducer:

$$R(L) = \int_0^L \frac{1}{\Delta L(l)} dl$$

Assuming the Weber law

$$\frac{\Delta L}{L} = k,$$

- and given the luminance transducer

$$R(L) = \int_0^L \frac{1}{\Delta L(l)} dl$$

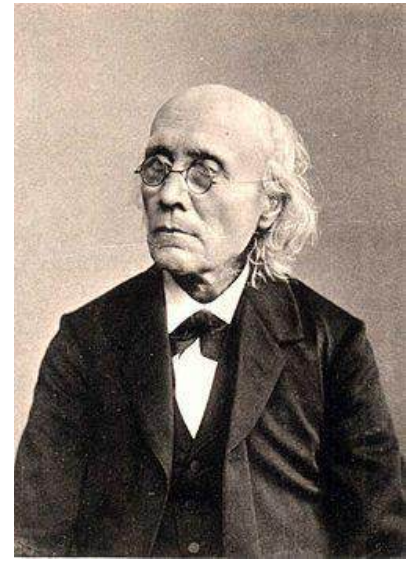
- the response of the visual system to light is:

$$R(L) = \int \frac{1}{kL} dL = \frac{1}{k} \ln(L) + k_1$$

Fechner law

$$R(L) = a \ln(L)$$

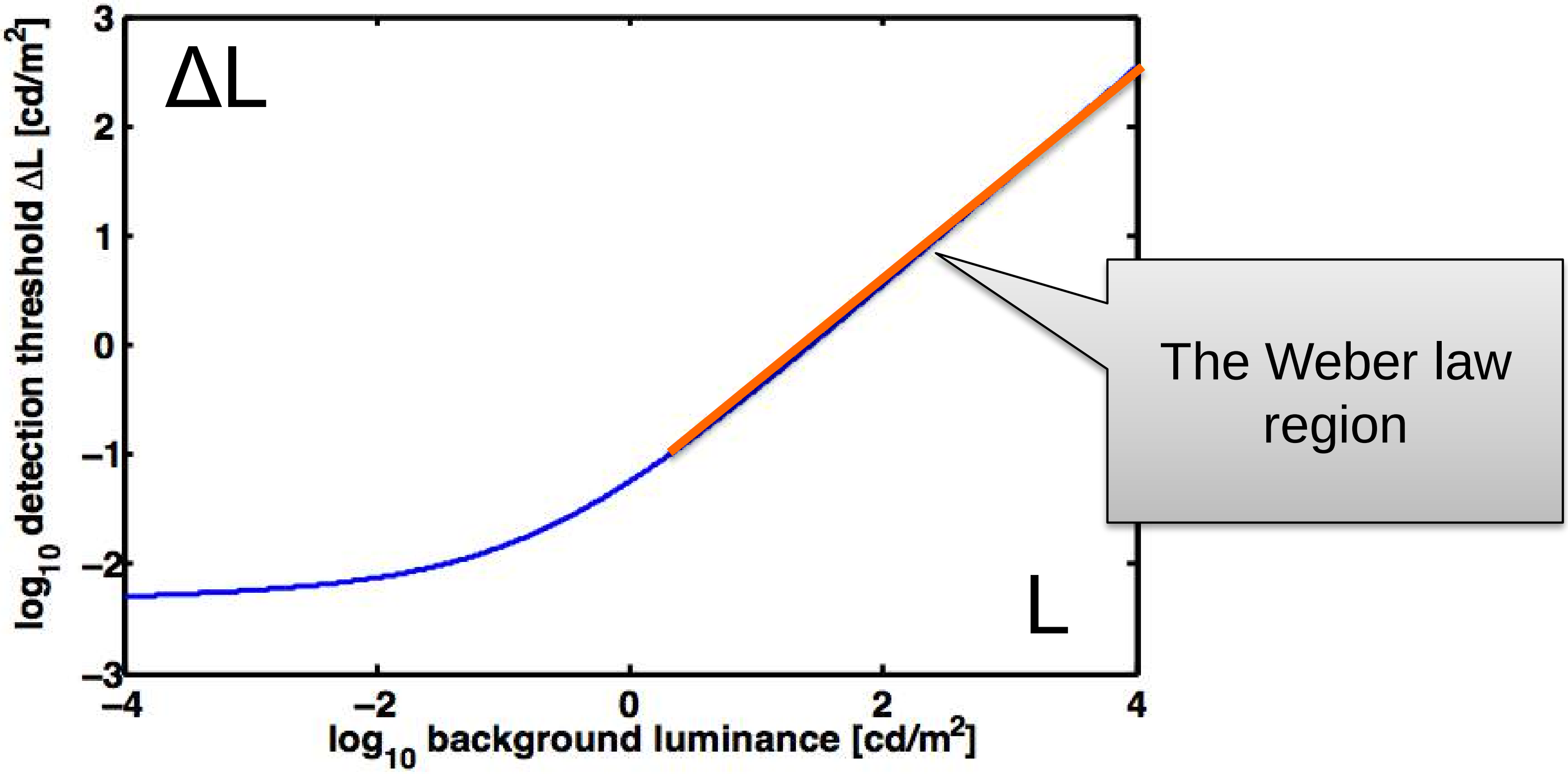
- Practical insight from the Fechner law:
  - The easiest way to adopt image processing algorithms to HDR images is to convert luminance (radiance) values to the logarithmic domain



Gustav Fechner
[From Wikipedia]

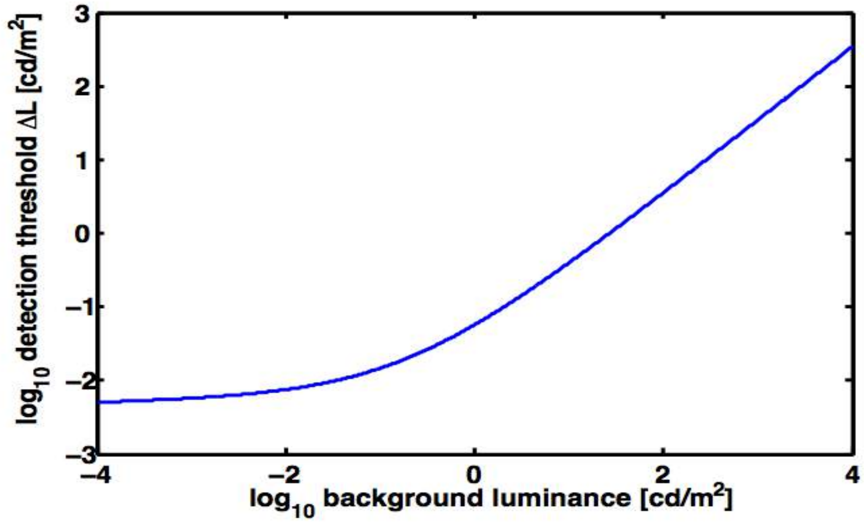
But…the Fechner law does not hold for the full luminance range

- Because the Weber law does not hold either
- Threshold vs. intensity function:



Weber-law revisited

- If we allow detection threshold to vary with luminance according to the t.v.i. function:

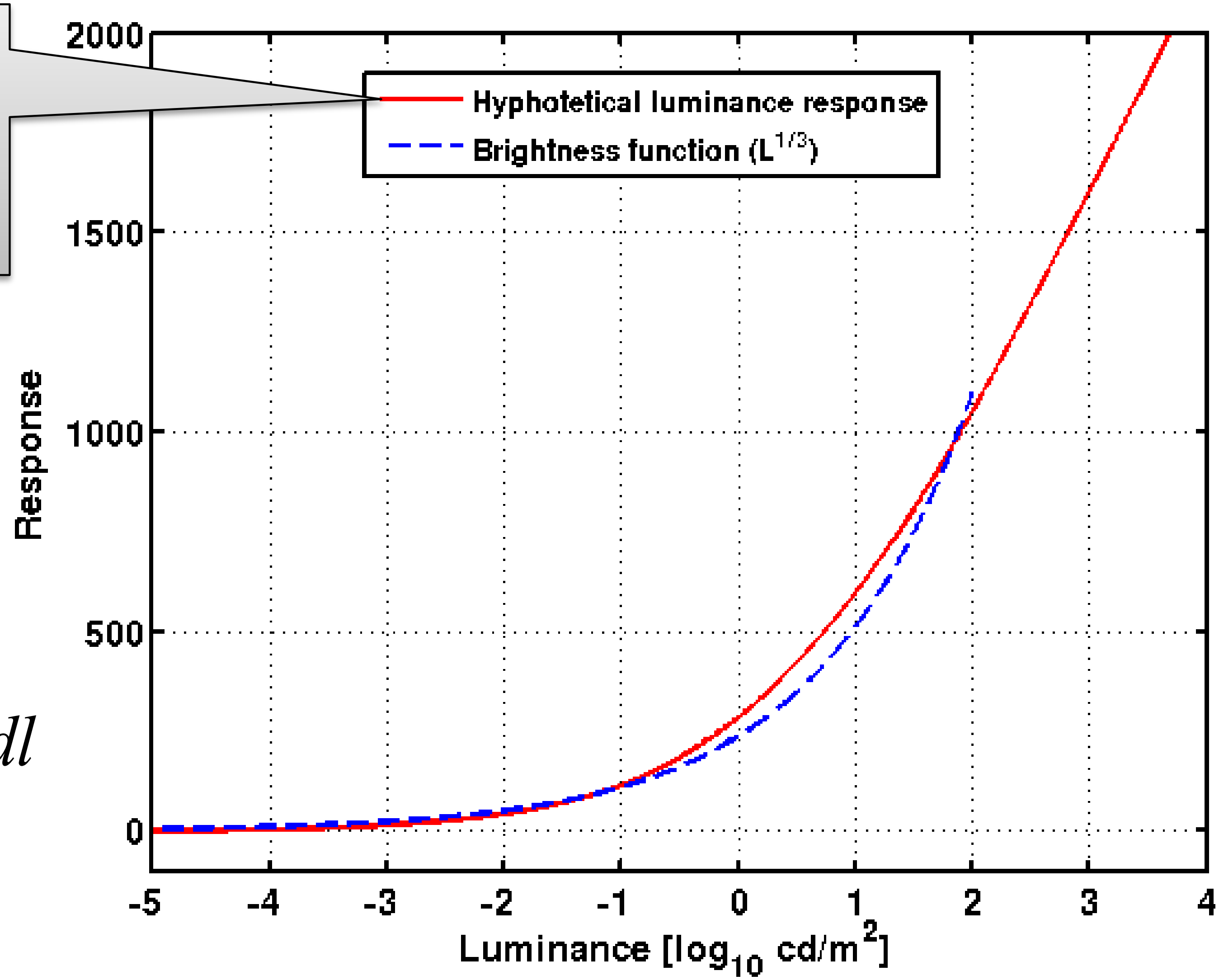


- we can get more accurate estimate of the "response":

$$R(L) = \int_0^L \frac{1}{\Delta L(l)} dl$$

Fechnerian integration and Steven's law

Function derived from the t.v.i. function



$$R(L) = \int_0^L \frac{1}{\Delta L(l)}\, dl$$

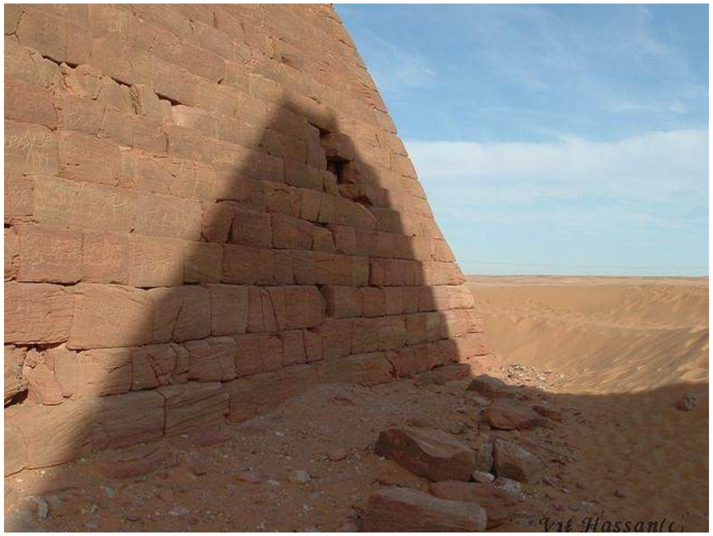
Major approaches to tone-mapping

- Illumination & reflectance separation
- Forward visual model
- Forward & backward visual model
- Constraint mapping problem

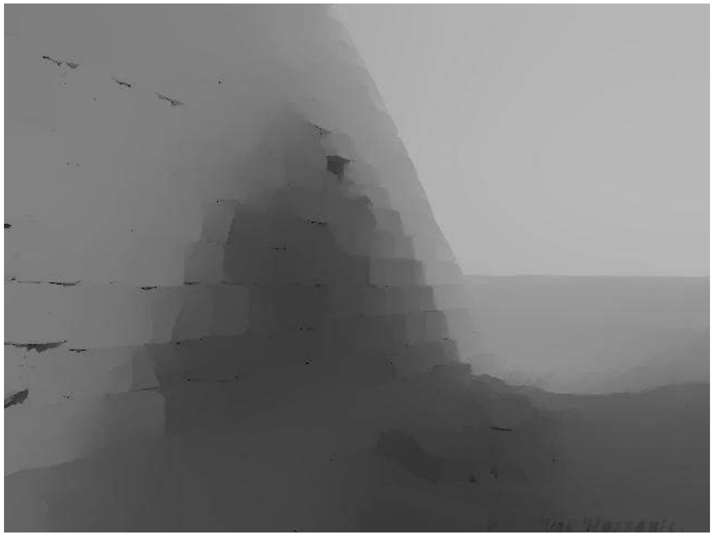
Major approaches to tone-mapping

- Illumination & reflectance separation
- Forward visual model
- Forward & backward visual model
- Constraint mapping problem

Illumination & reflectance separation



Input



Incoming illumination

Reflectance

- Different to intrinsic images in CV where *outgoing* illumination is relevant.

Illumination and reflectance

Illumination

- Sun ≈ 10^9
- Lowest perceivable luminance ≈ 10^{-6}
- Dynamic range can easily exceed 3-4 $\log_{10}$ units in a scene
- Visual system partially discounts illumination

Reflectance

- White ≈ 90%
- Black ≈ 3%
- Dynamic range < 100:1
- Reflectance critical for object & shape detection

Reflectance & Illumination TMO

- Distortions in reflectance are more apparent than the distortions in illumination.
- Tone mapping could preserve reflectance but compress illumination

Illumination

Tone-mapped image

$$I_d = R \cdot T(I)$$

Reflectance

Tone-mapping

- for example: $I_d = R \cdot L^{1/\gamma}$

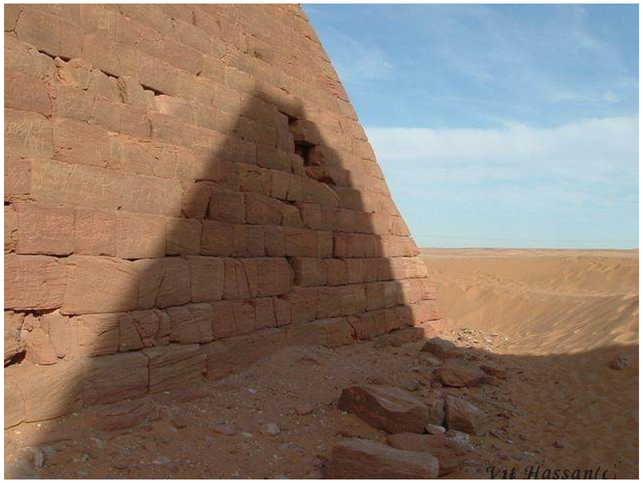
How to separate the two?

- (Incoming) illumination – slowly changing
  - except very abrupt transitions on shadow boundaries
- Reflectance – low contrast and
  high frequency variations

Gaussian filter

$$f(x) = \frac{1}{2\pi\sigma_s} e^{\frac{-x^2}{2\sigma_s^2}}$$

- First order approximation



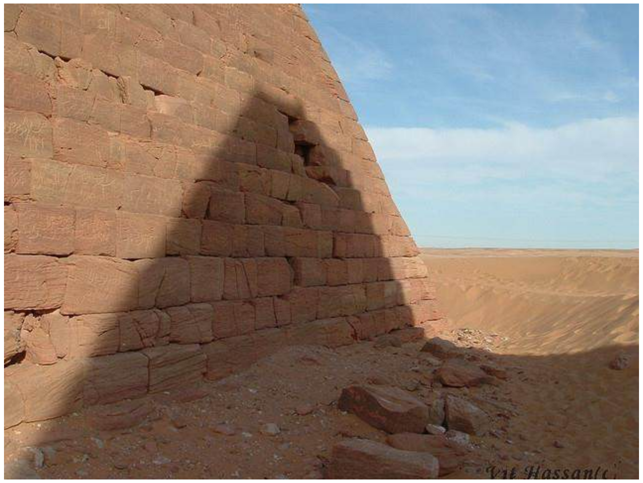
- Blurs sharp boundaries
- Causes halos

Tone mapping result

Bilateral filter

$$I_p \approx \frac{1}{k_s} \sum_{t \in \Omega} f(p-t)\, g(L_p - L_t)\, L_p$$

- Better preserves sharp edges



- Still some blurring on the edges
- Reflectance is not perfectly separated from illumination near edges

Tone mapping result

WLS filter

- Weighted-least-squares optimization

Make reconstructed image *u* possibly close to input *g*

Smooth out the image by making partial derivatives close to 0

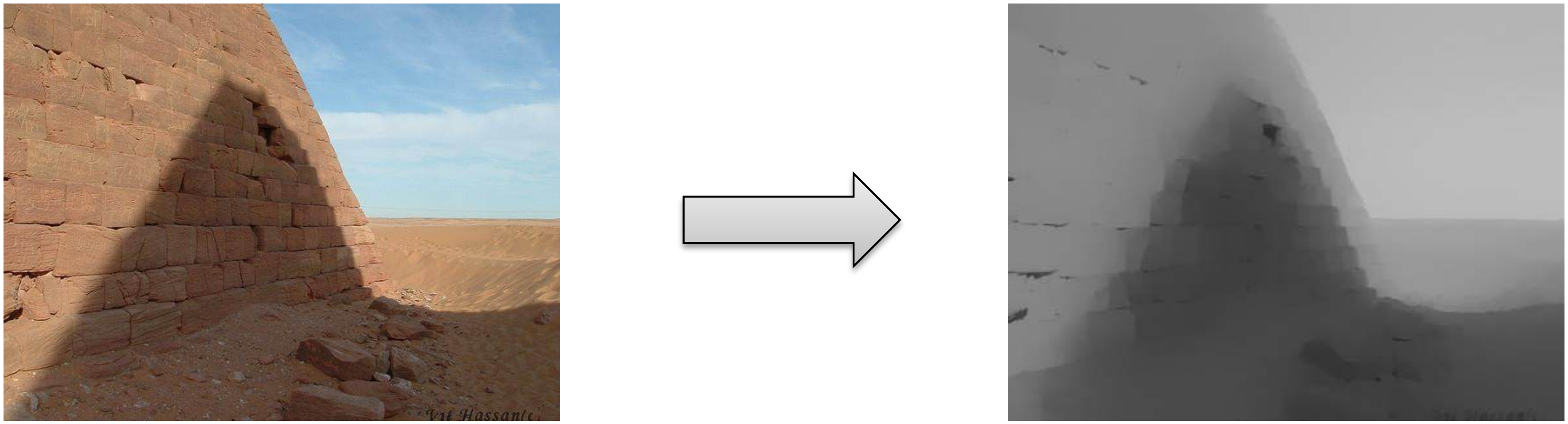
$$\sum_p \left(\left(u_p - g_p\right)^2 + \lambda \left(a_{x,p}(g) \left(\frac{\partial u}{\partial x} \right)_p^2 + a_{y,p}(g) \left(\frac{\partial u}{\partial y} \right)_p^2 \right) \right) \quad \text{-> min}$$

Spatially varying smoothing – less smoothing near the edges

- [Farbman et al., SIGGRAPH 2008]

WLS filter

- Stronger smoothing and still distinct edges



Tone mapping result

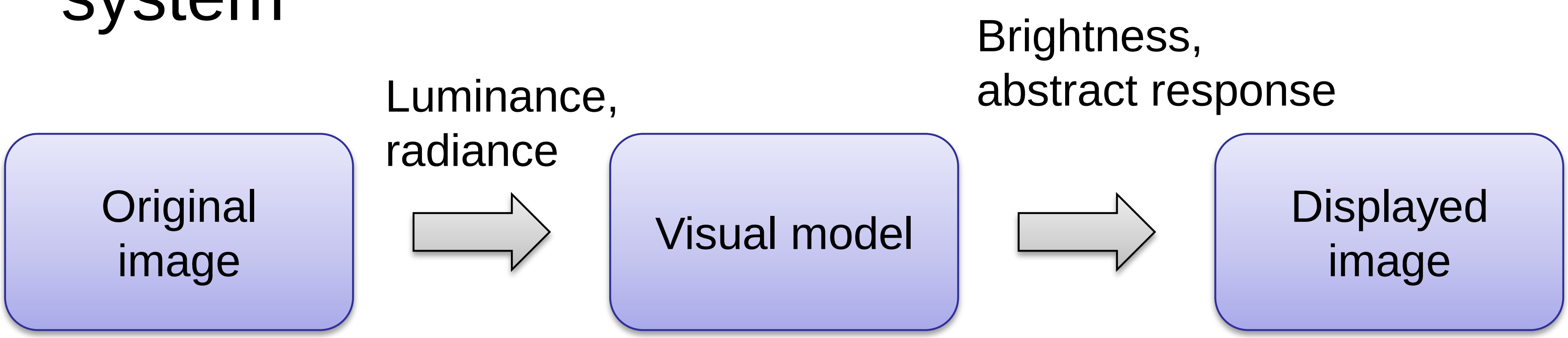
- Can produce stronger effects with less artifacts

Major approaches to tone-mapping

- Illumination & reflectance separation
- Forward visual model
- Forward & backward visual model
- Constraint mapping problem

Forward visual model

- Mimic the processing in the human visual system



- Assumption: what is displayed is brightness or abstract response of the visual system

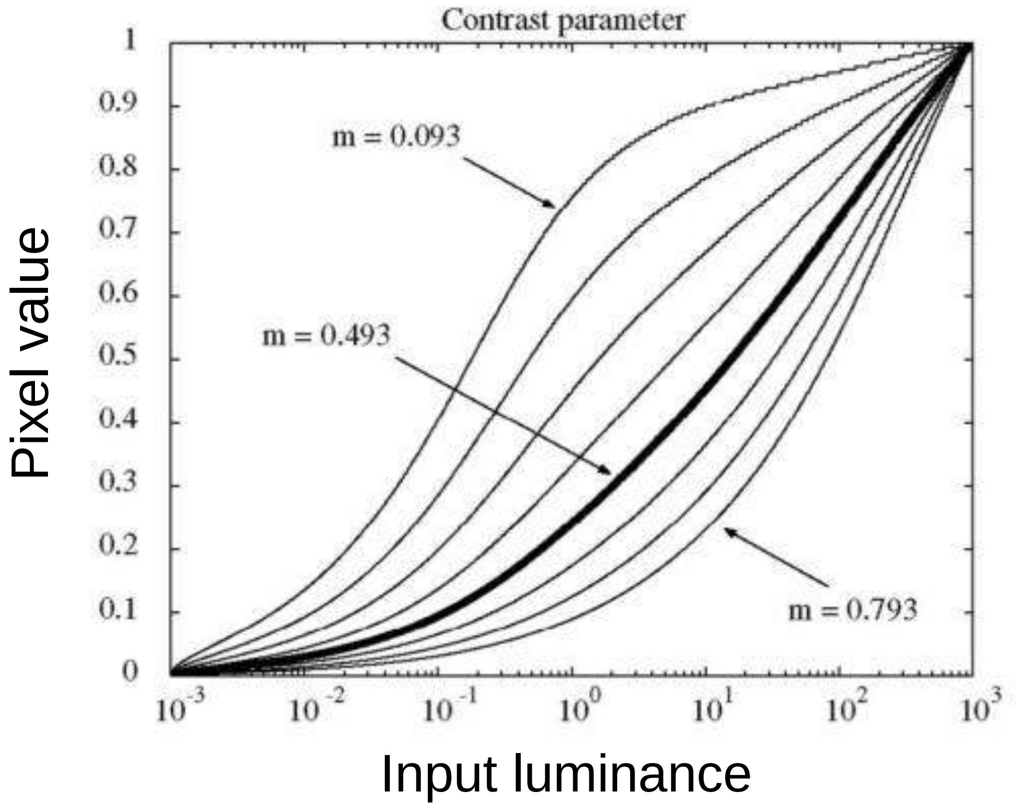
Photoreceptor response

- Dynamic range reduction inspired by photoreceptor physiology
  - [Reinhard & Devlin '05]

$$V = \frac{I}{I + \sigma(I_a)} V_{max}$$

$$\sigma(I_a) = (f I_a)^m.$$

- From gamma to sigmoidal response:



Results: photoreceptor TMO

Our operator

Bilateral filtering

Trilateral filtering

Histogram adjustment

Photographic tonemapping (global)

Photographic tonemapping (local)

Logarithmic mapping

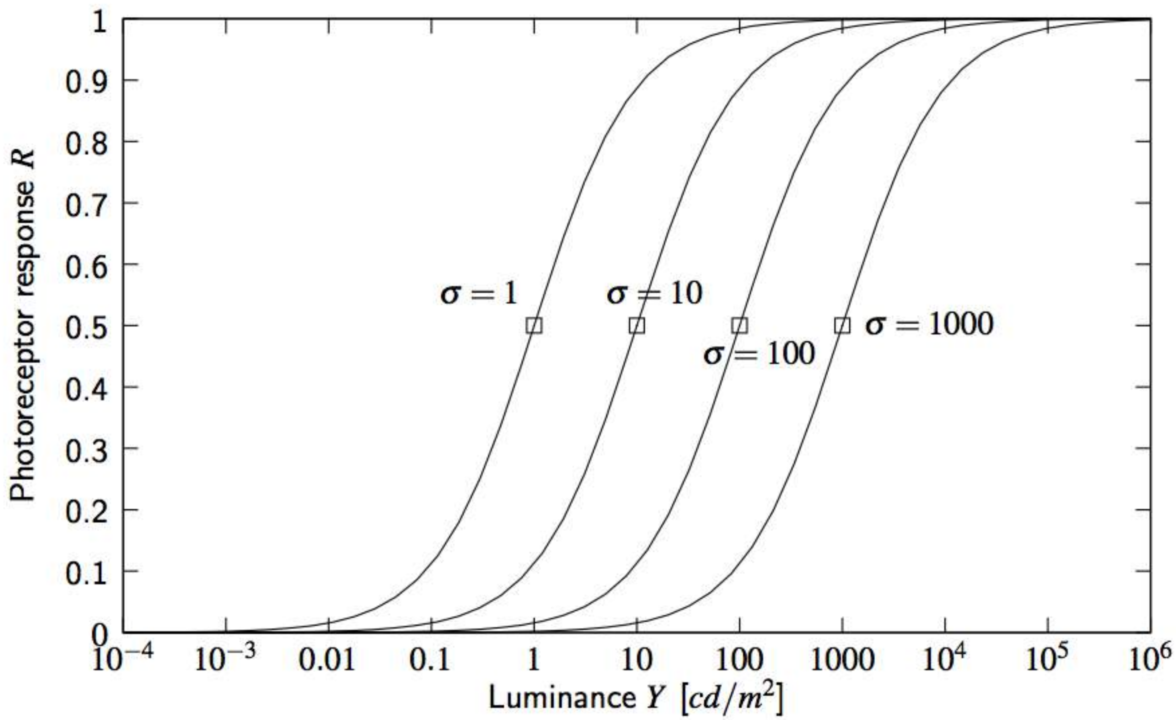
Adaptive logarithmic mapping

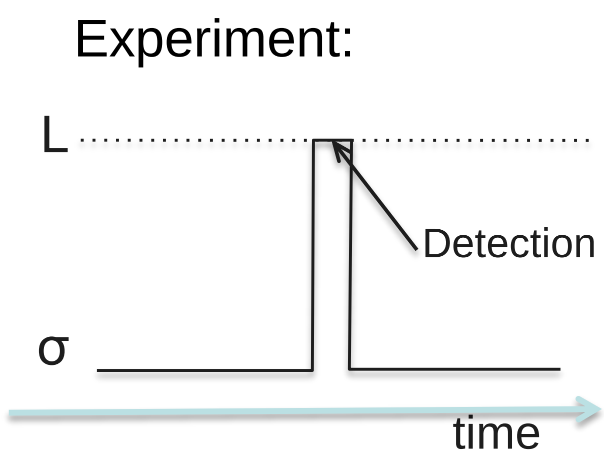
Ashikhmin's operator

Photoreceptor models

- Naka-Rushton equation:

$$\frac{R}{R_{max}} = \frac{Y^n}{Y^n + \sigma^n}$$

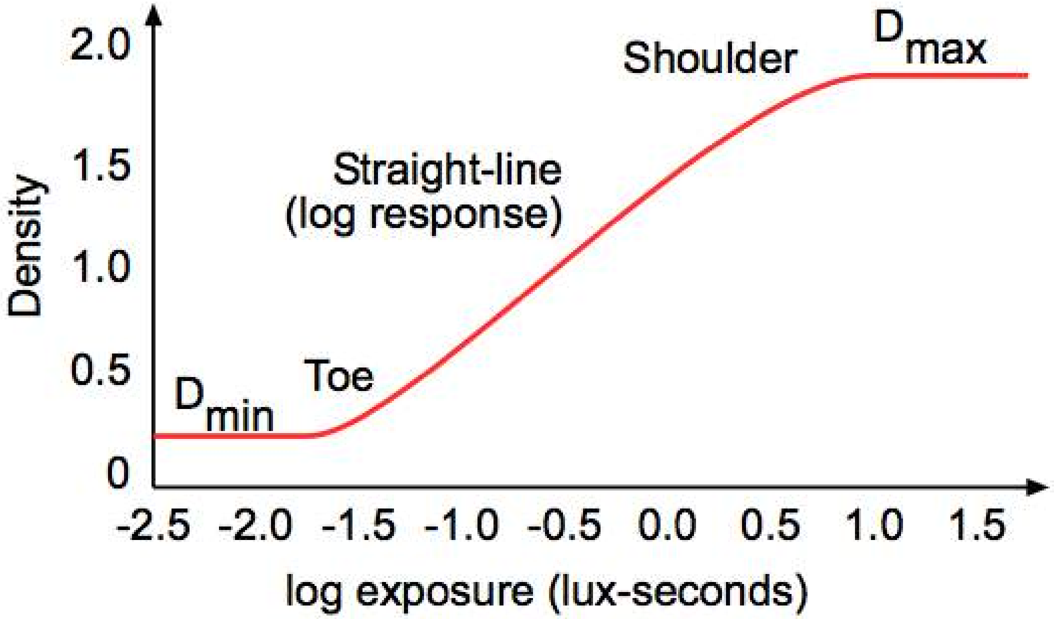




- Response of the photoreceptor to a short flicker of light - less applicable to viewing static images

Sigmoidal tone-curves

- Very common in digital cameras
  - Mimic the response of analog film
  - Analog film has been engineered for many years to produce optimum tone-reproduction (given that he tone curve must not change)



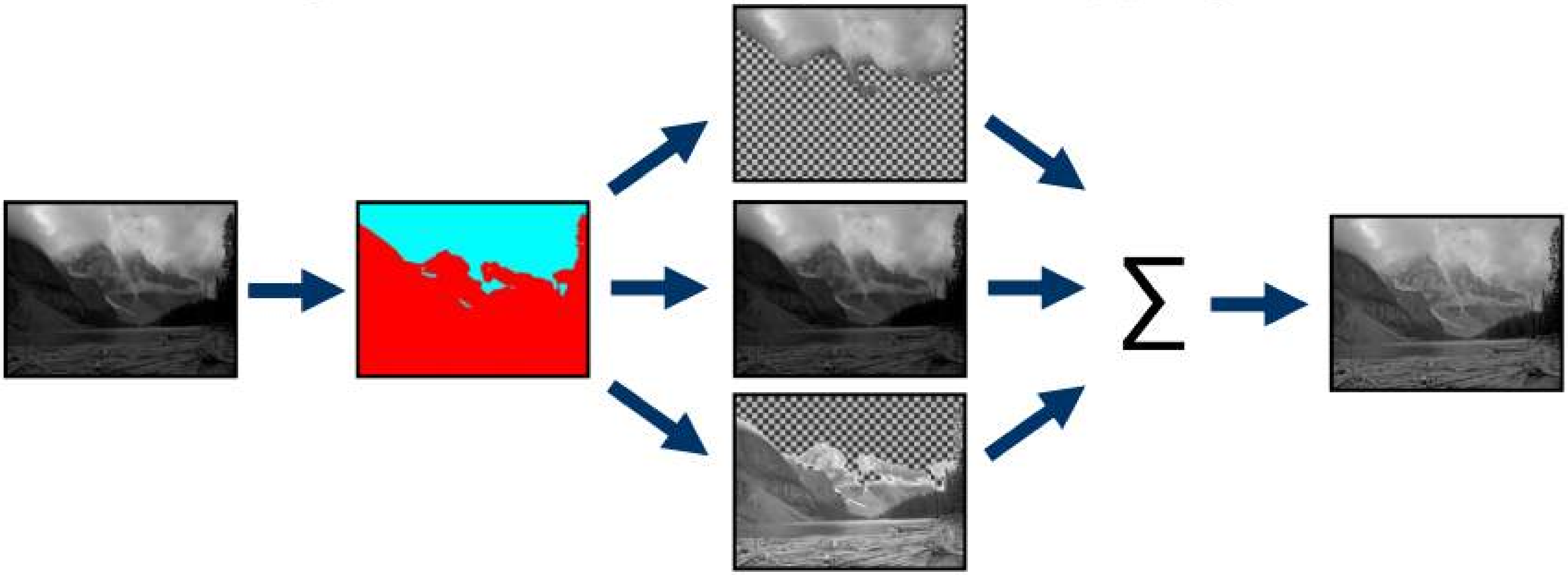
- Effectively the most commonly used tone-mapping!

Why sigmoidal tone-curves work

- Because they mimic photoreceptor response
  - Unlikely, because photoreceptor response to steady light is not sigmoidal
- Because they preserve contrast in mid-tones, which usually contains skin color
  - We are very sensitive to variation in skin color
- Because an image on average has Gaussian distribution of log-luminance
  - S-shape function is the result of histogram equalization of an image with a Gaussian-shape histogram

Lightness perception

- Lightness perception in tone-reproduction for high dynamic range images [Krawczyk et al. '05]
- Based on Gilchrist lightness perception theory

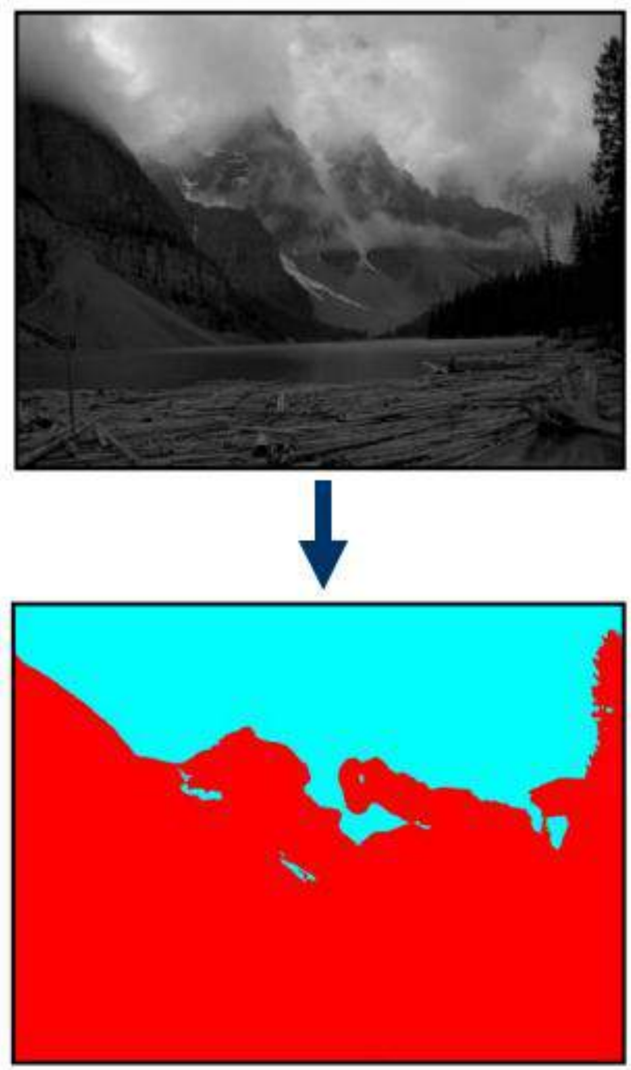


- Perceived lightness is anchored to several frameworks

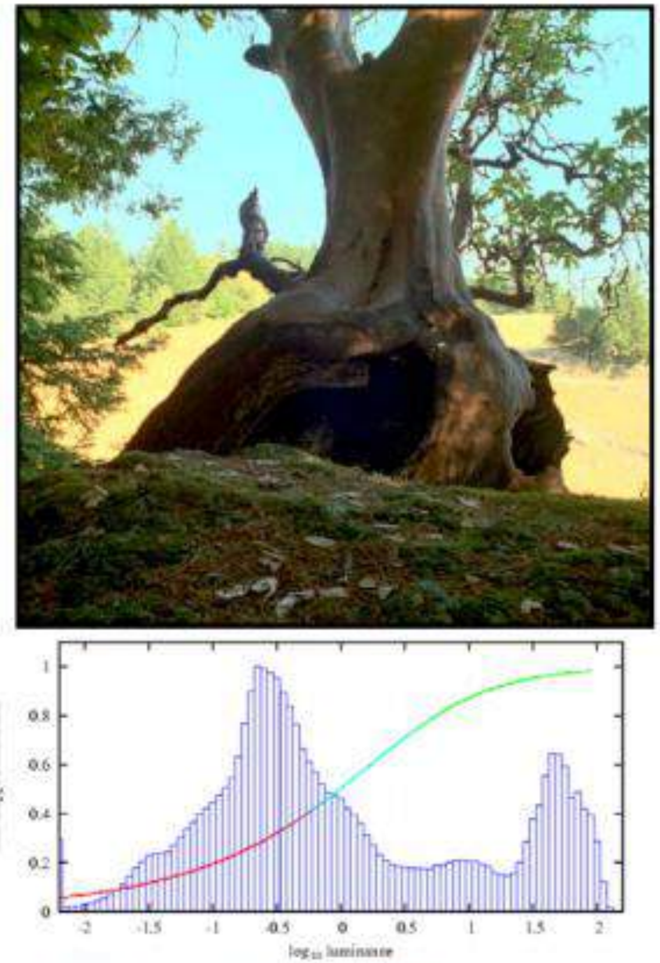
Gilchrist lightness perception theory

- Frameworks – areas of common illumination
- Anchoring – the tendency of
  - highest luminance
  - largest area

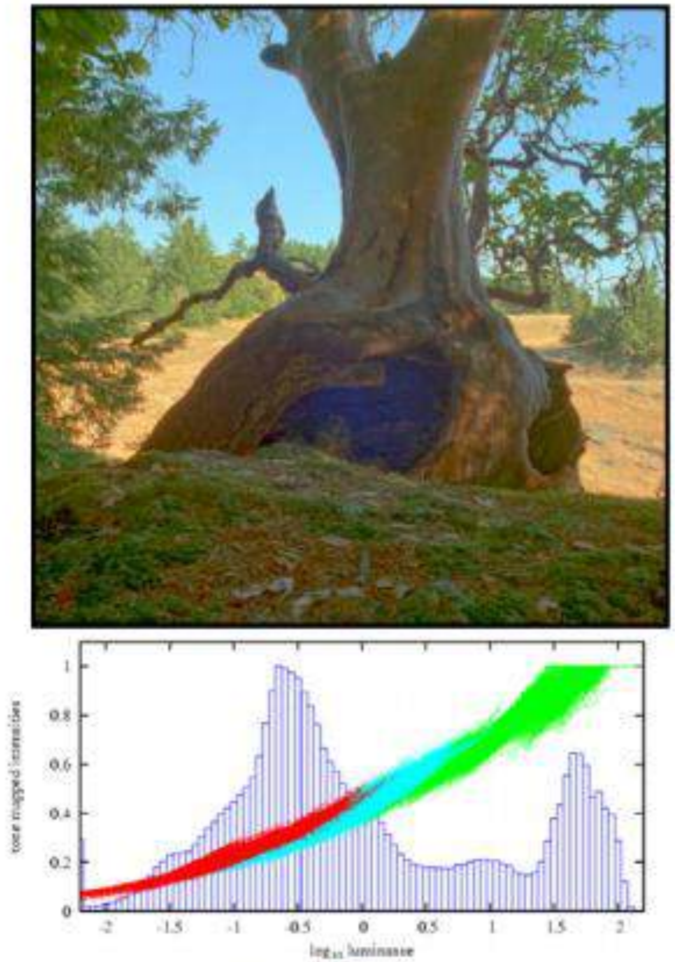
  to appear white
- Tone-mapping
  - Rescale luminance in each framework to its anchor



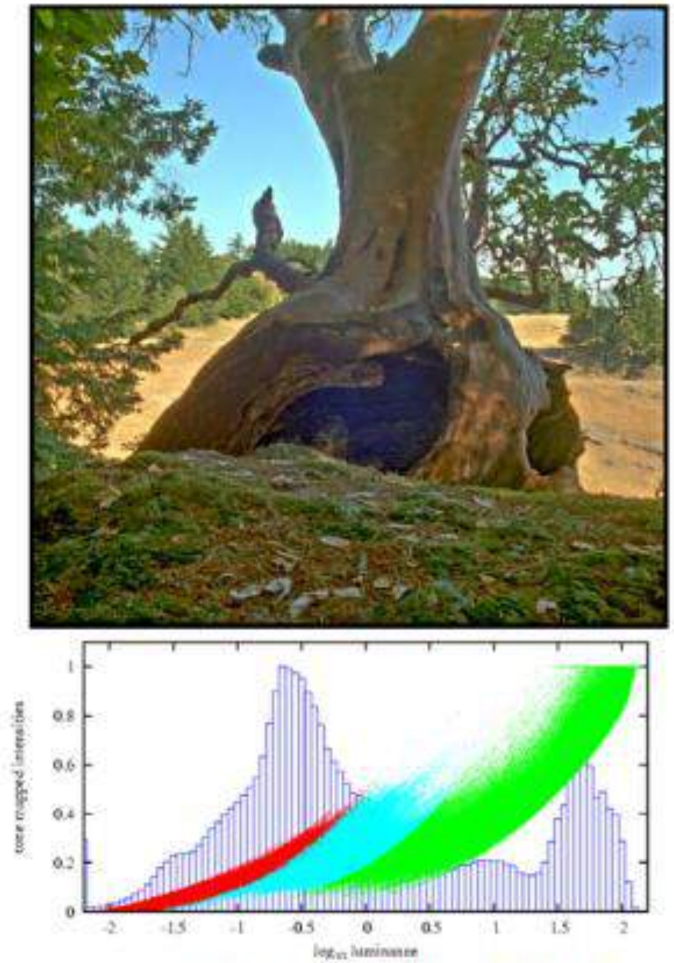
Results – lightness perception TMO



Photographic Tone Reproduction



Bilateral Filtering

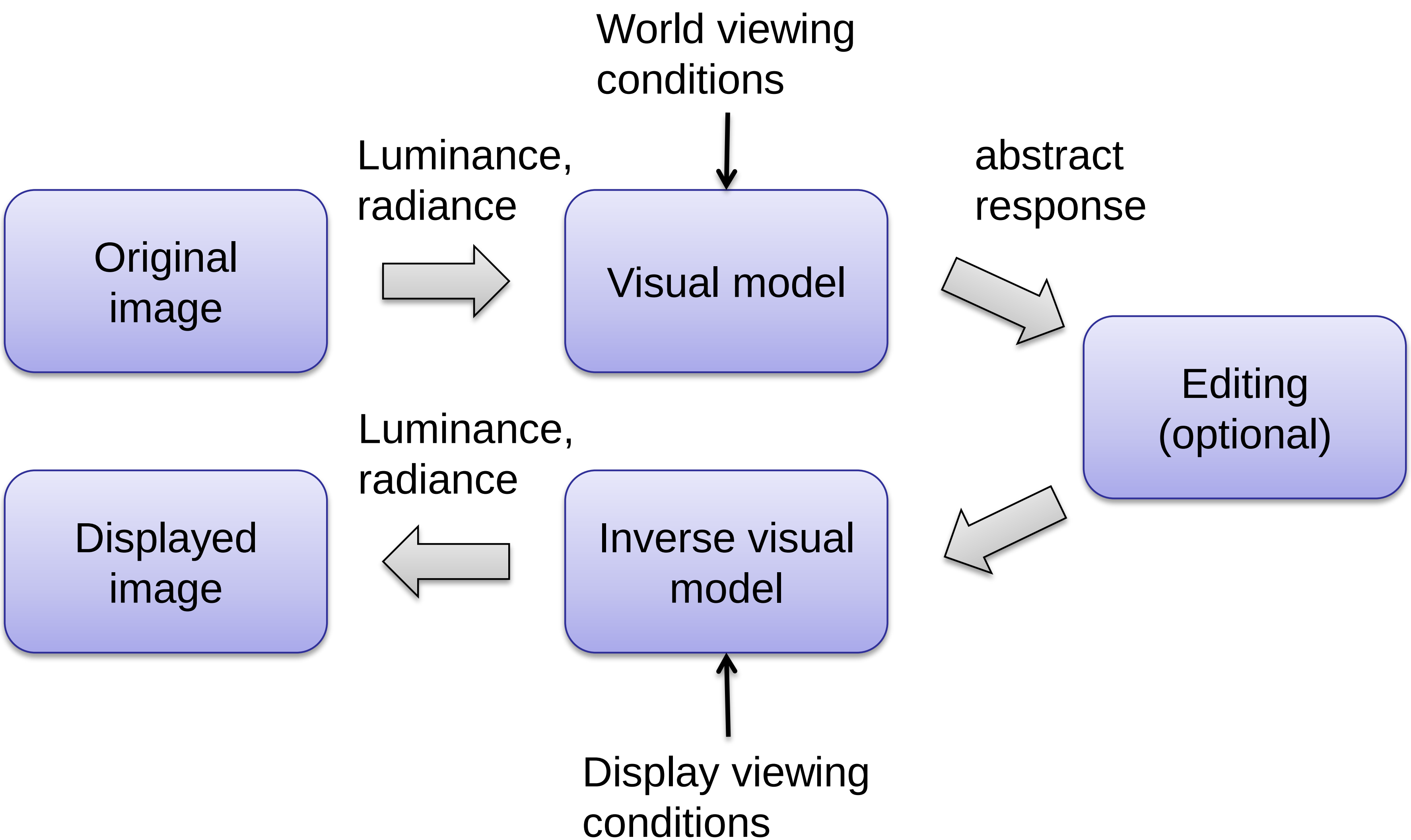


Presented Computational Model

Major approaches to tone-mapping

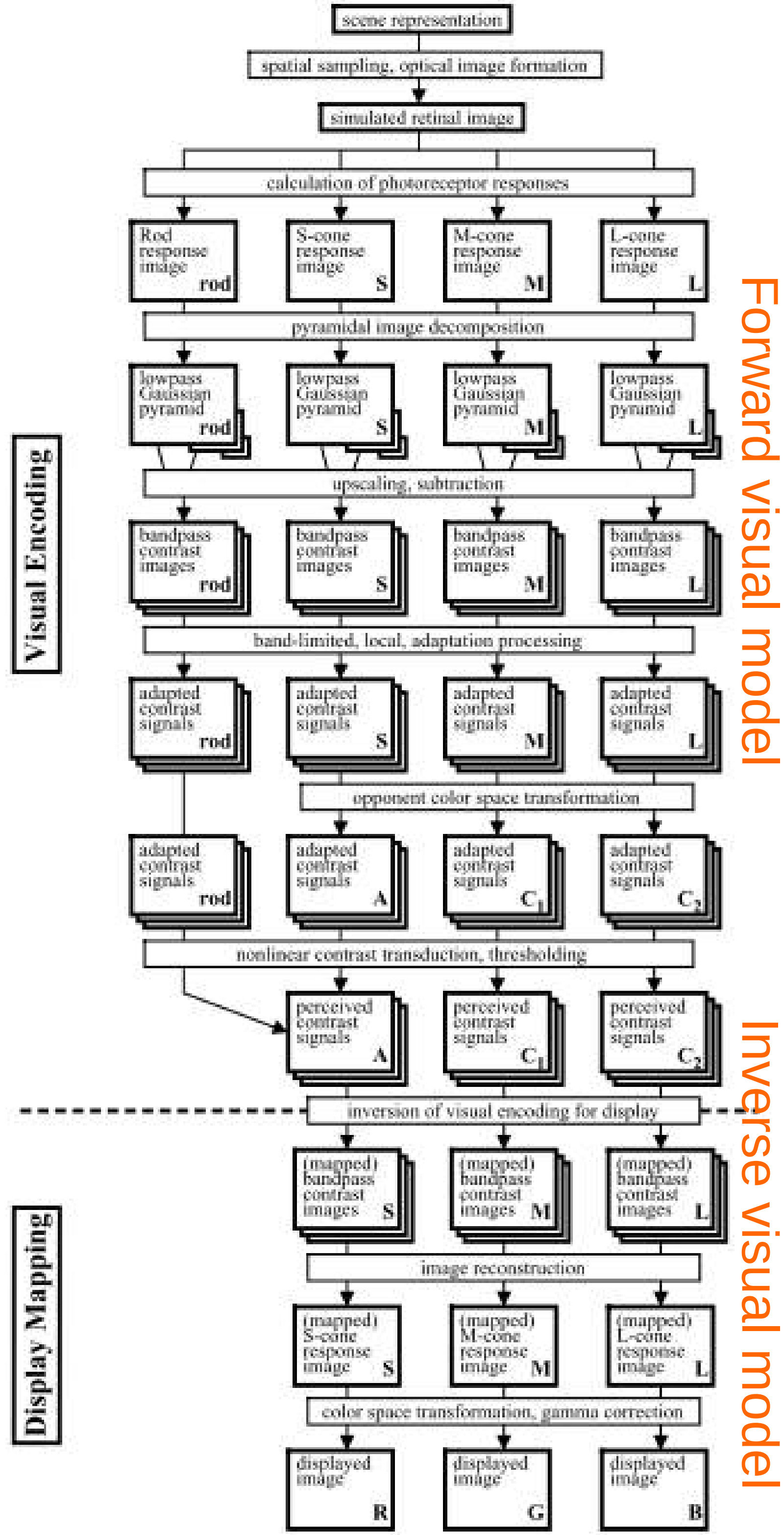
- Illumination & reflectance separation
- Forward visual model
- Forward & backward visual model
- Constraint mapping problem

Forward and inverse visual model

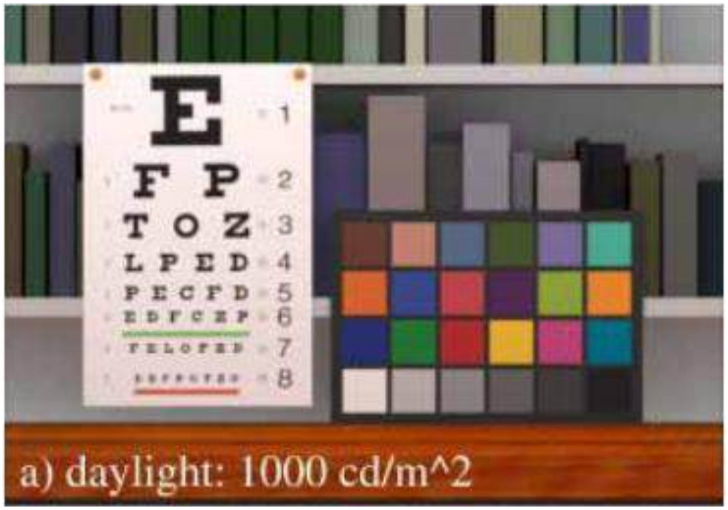


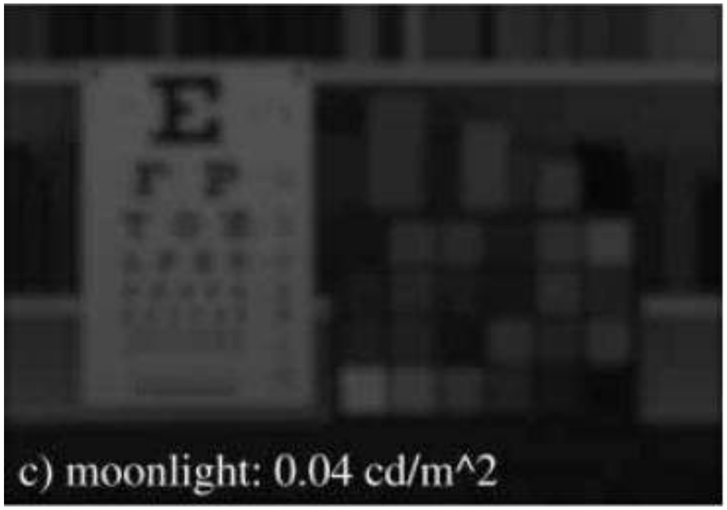
Multi-scale model

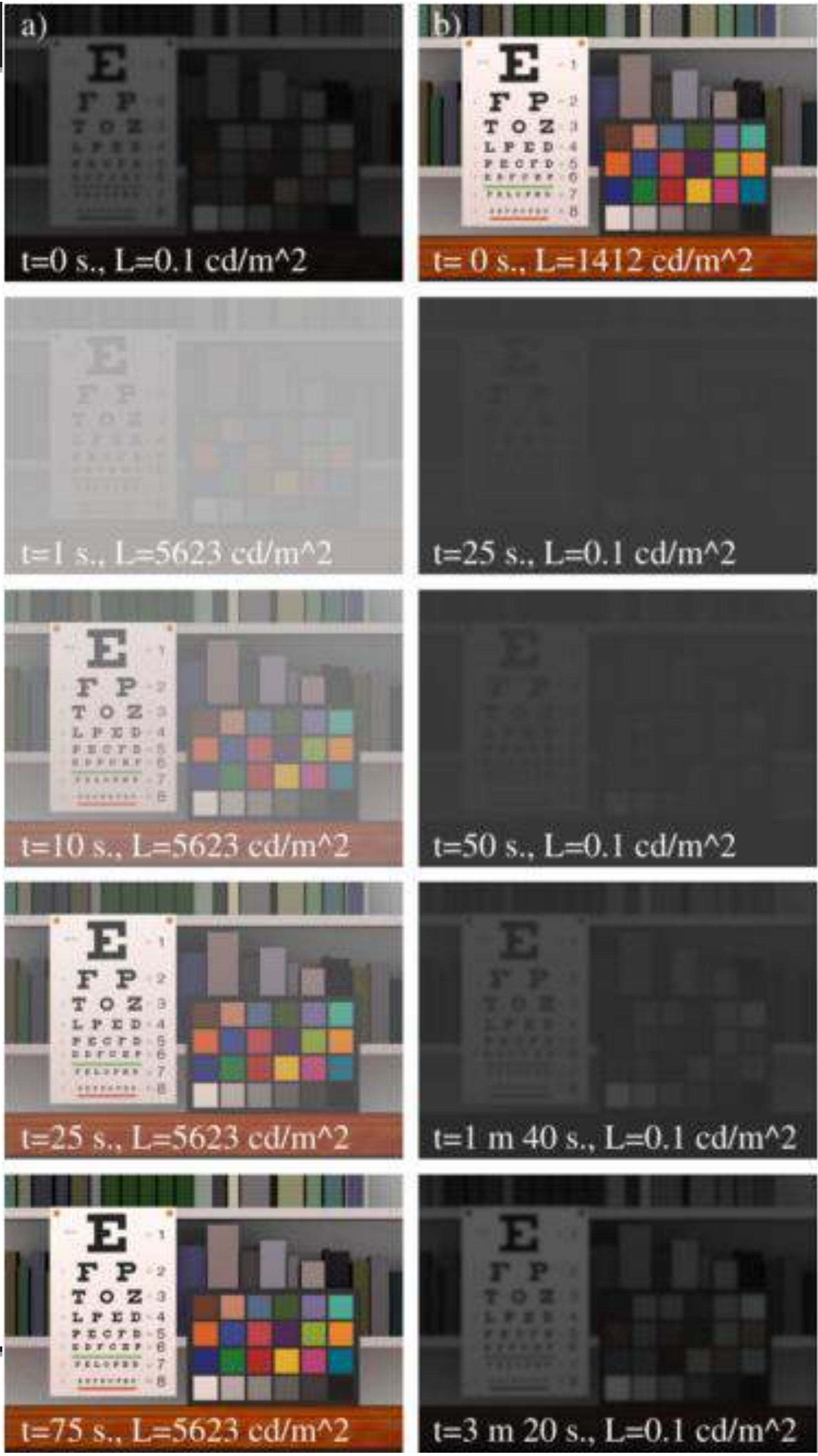
- Multi-scale model of adaptation and spatial vision and color appearance
  - [Pattanaik et al. '98]
- Combines
  - psychophysical threshold and superthreshold visual models
  - light & dark adaptation models
  - Hunt's color appearance model
- One of the most sophisticated visual models



Results – multiscale model …







Forward and inverse visual model

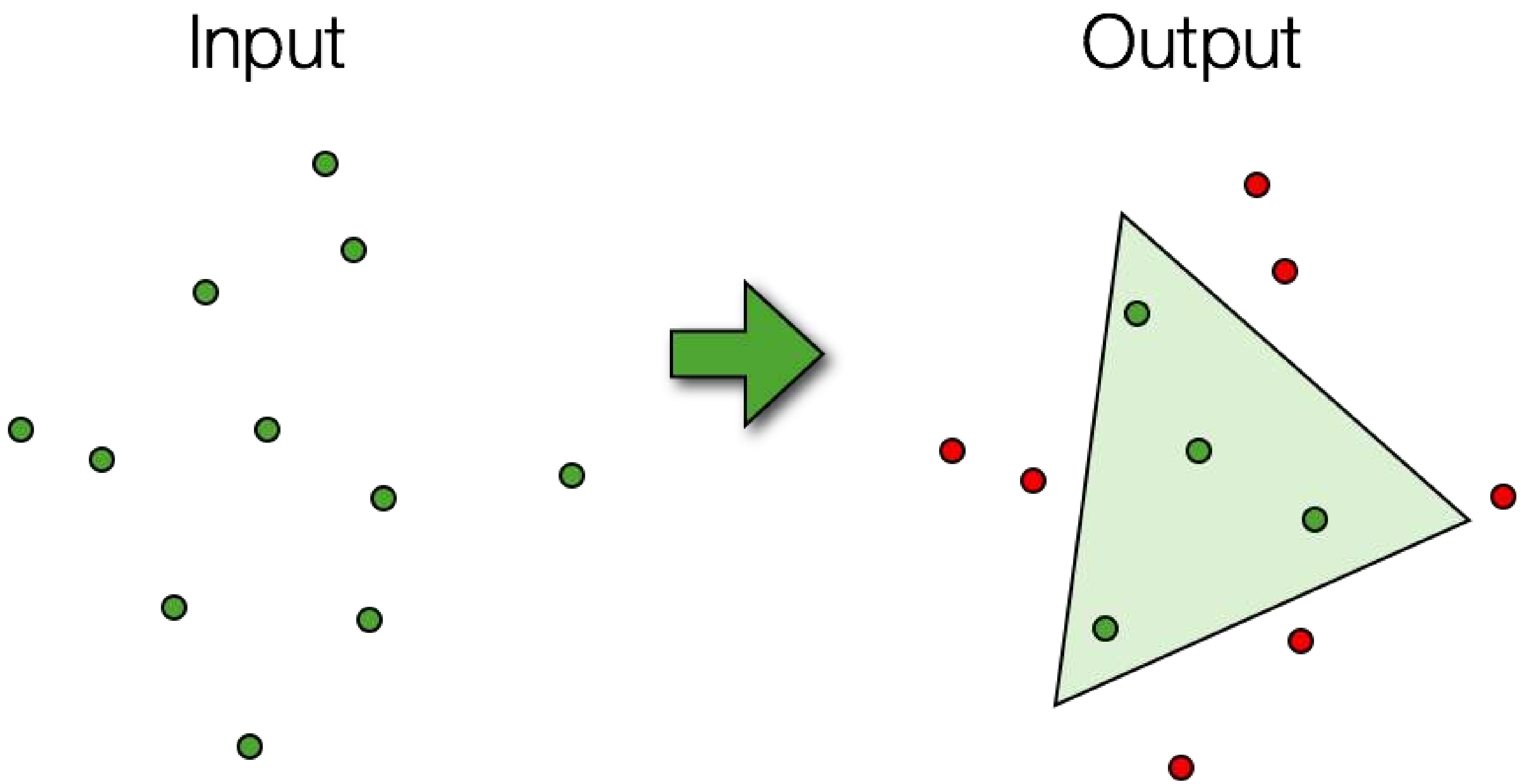
- Advantages of F&I visual models
  - Can render images for different viewing conditions
    - Different state of chromatic or luminance adaptation
  - Physically plausible
    - output in the units of luminance or radiance
- Shortcomings F&I visual models
  - Assume that a standard display can reproduce the impression of viewing much brighter or darker scenes
  - Cannot ensure that the resulting image is within the dynamic range of the display
    - Not necessary meant to reduce the dynamic range
  - Visual models are difficult to invert

Major approaches to tone-mapping

- Illumination & reflectance separation
- Forward visual model
- Forward & backward visual model
- Constraint mapping problem

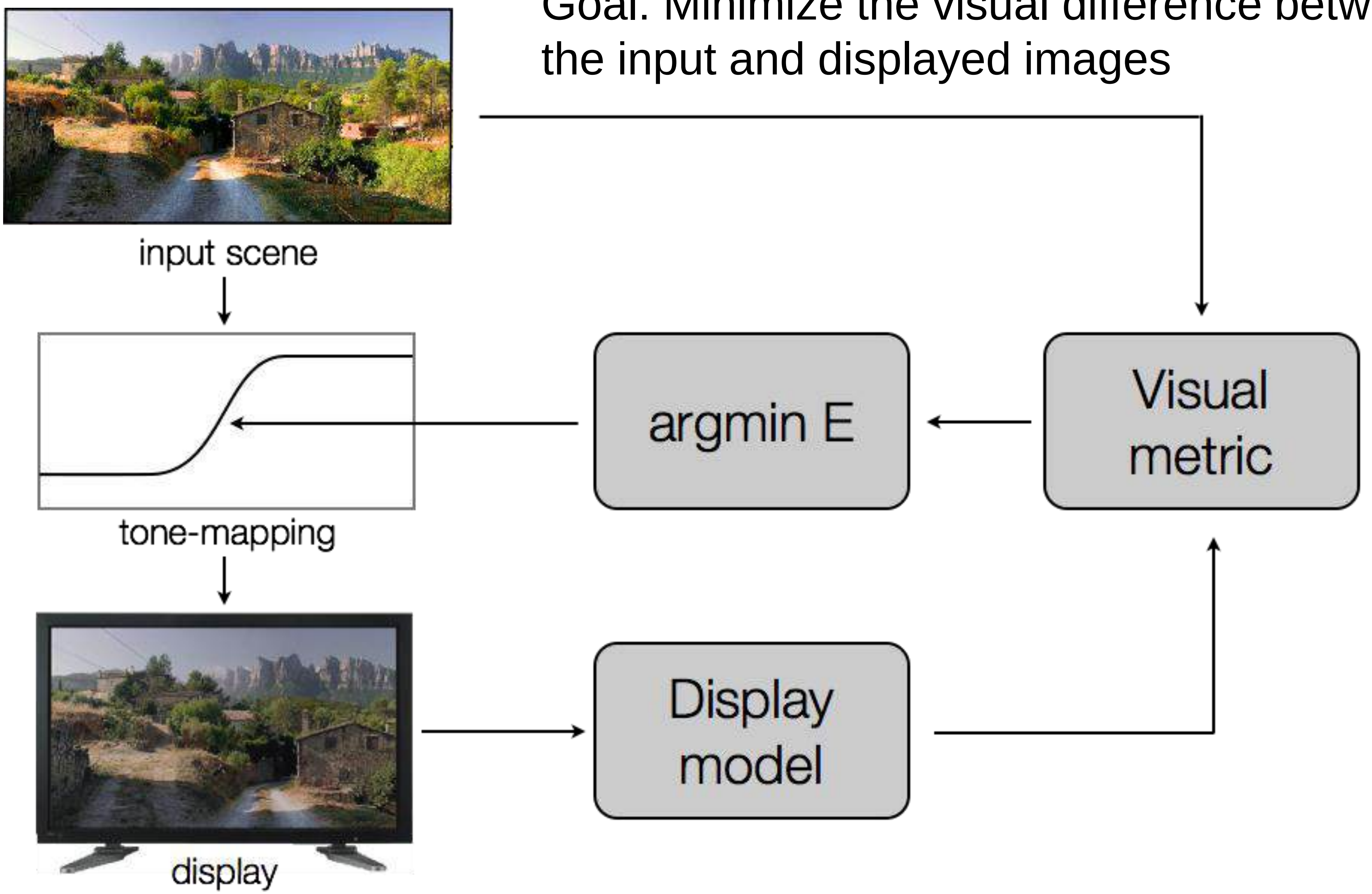
Constraint mapping problem

- Goal: to restrict the range of values while reducing inflicted damage



Display adaptive tone-mapping

Goal: Minimize the visual difference between the input and displayed images

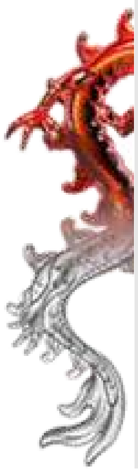


Thank you

Apparent Contrast and Brightness Enhancement

Karol Myszkowski

MPI Informatik

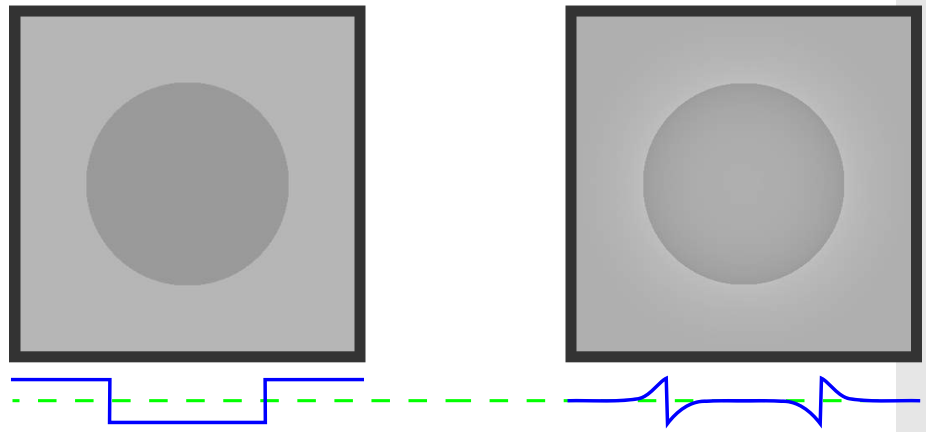


Motivation

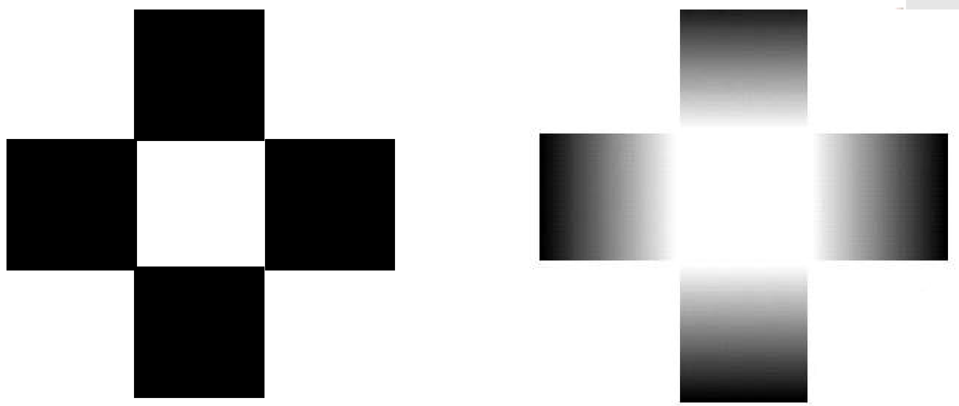
- Image display
  - Limited dynamic range of existing display technology
  - Cannot match to physical contrast and brightness of real world scenes
  - Physical match not really required for good reproduction of image appearance
- Modern tone mapping operators good at optimizing the physical contrast and luminance use
- Human preference
  - Enhanced contrast and brightness improve image appearance
- Can we still boost the contrast and brightness impression?

Human perception

- Spatial vision
  - Image appearance can be strongly affected by skillful introduction of intensity gradients between neighboring pixels
- Cornsweet illusion
  - Apparent contrast boost



- Glare illusion
  - Apparent brightness boost



Contrast Enhancement: Motivation

HDR image (reference)

tone mapping result

- Usual contrast enhancement techniques
  - either enhance everything
  - or require manual intervention
  - change image appearance
- Tone mapping often gives numerically optimal solution
  - no dynamic range left for enhancement

Krawczyk et al. EG2007

Overview

Reference HDR Image

Tone Mapped Image

Measure Lost Contrast at Several Feature Scales

Enhance Lost Contrast in Tone Mapped Image

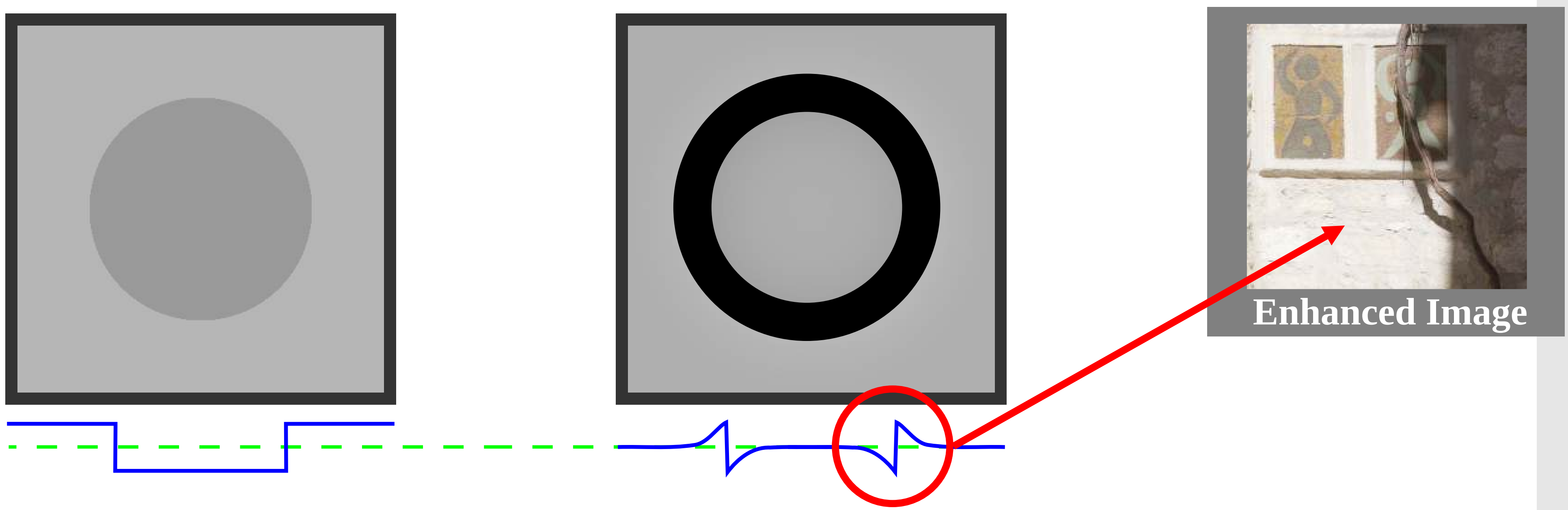
Enhanced TM Image

communicate lost image contents

maintain image appearance

Krawczyk et al. EG2007

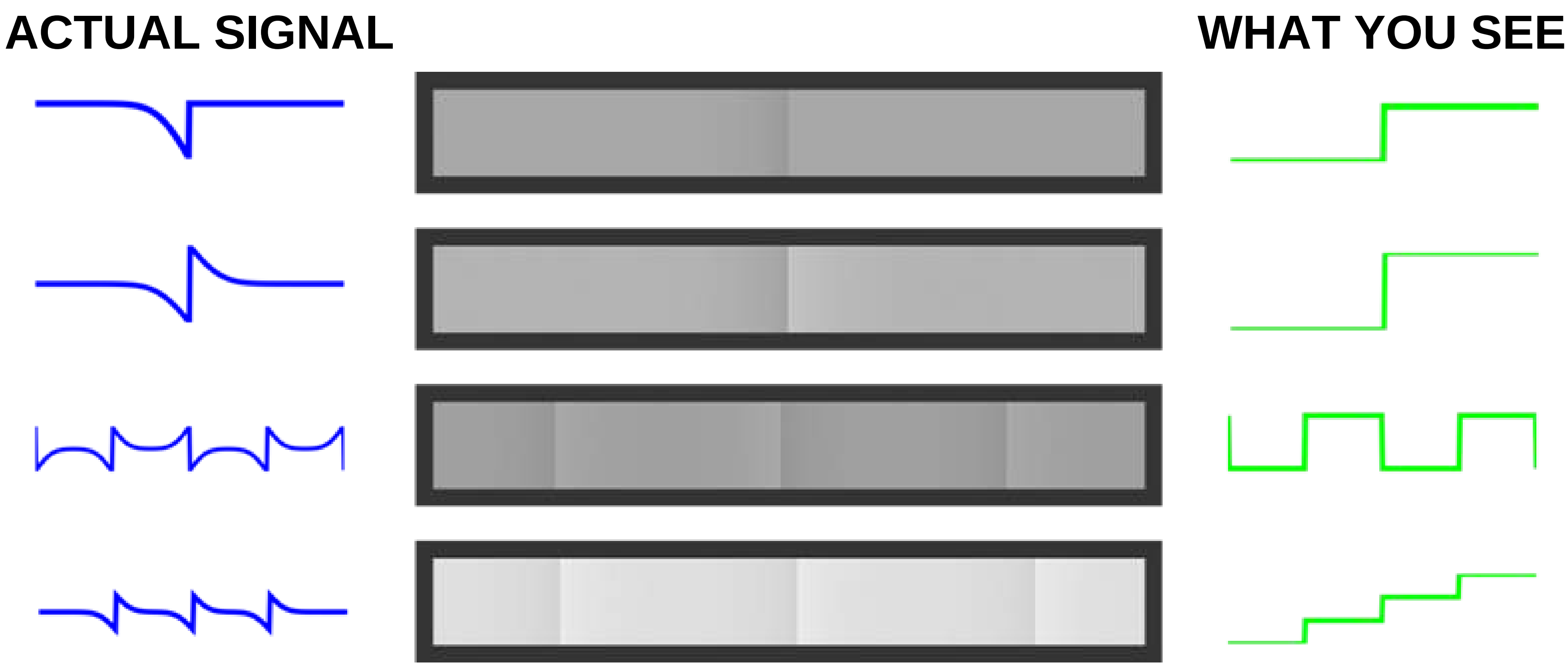
Cornsweet Illusion



- Create apparent contrast based on Cornsweet illusion
- **Countershading**
  - gradual darkening / brightening towards a contrasting edge
  - contrast appears with 'economic' use of dynamic range

Krawczyk et al. EG2007

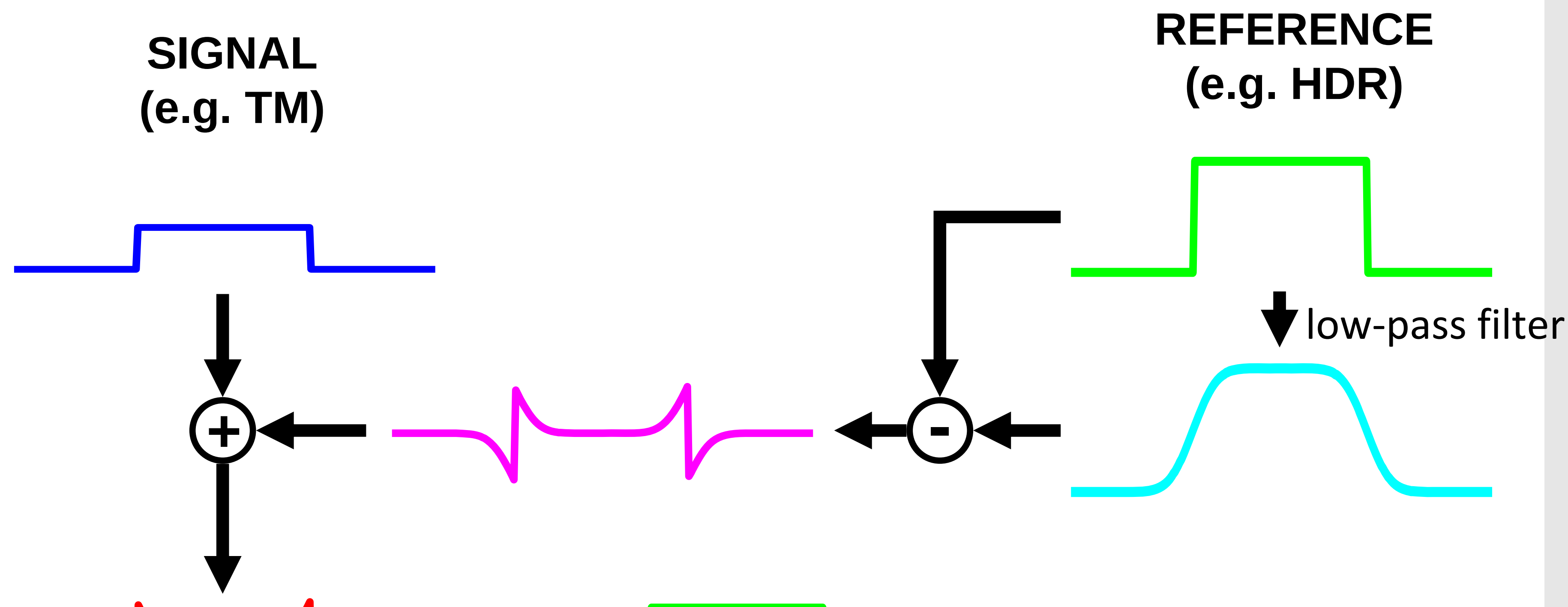
Details of Contrast Illusion

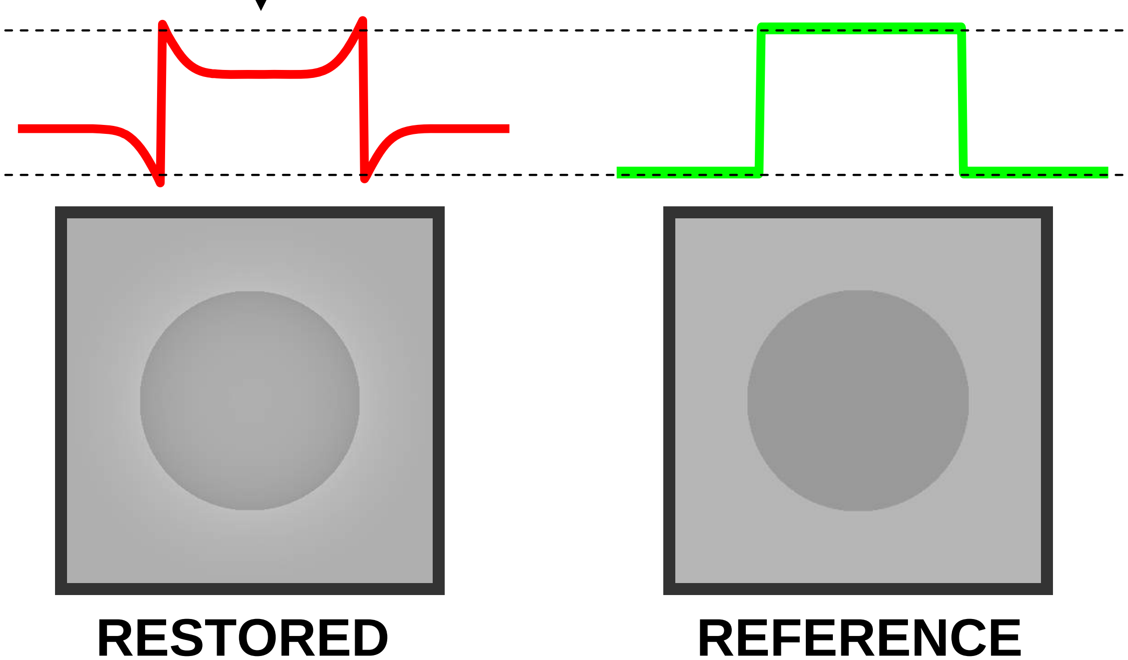


1. Contrast between areas caused by luminance profiles
2. Properties:
   - shape of the profile matches the shape of the enhanced feature
   - amplitude of the profile defines the perceived contrast
   - noise (texture) does not cancel the illusion
   - profiles should not be discernible

Krawczyk et al. EG2007

Construction of Simple Profile (1/2)

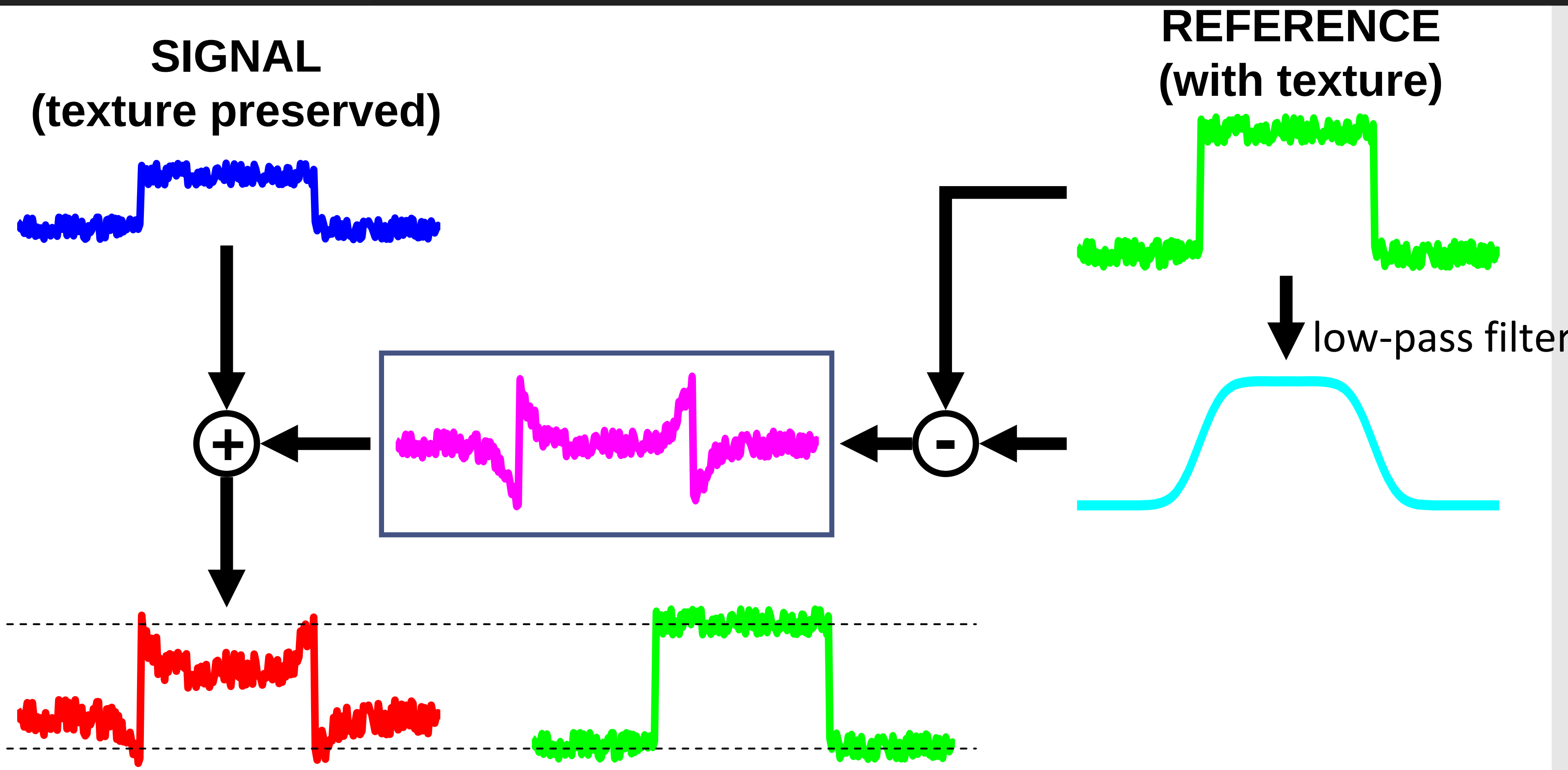




- Profile from low-pass filtered reference
- Size and amplitude adjusted manually
- This is **unsharp masking**

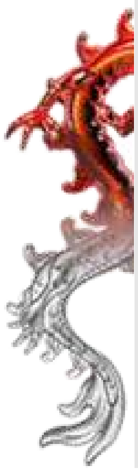
Krawczyk et al. EG2007

Construction of Simple Profile (2/2)

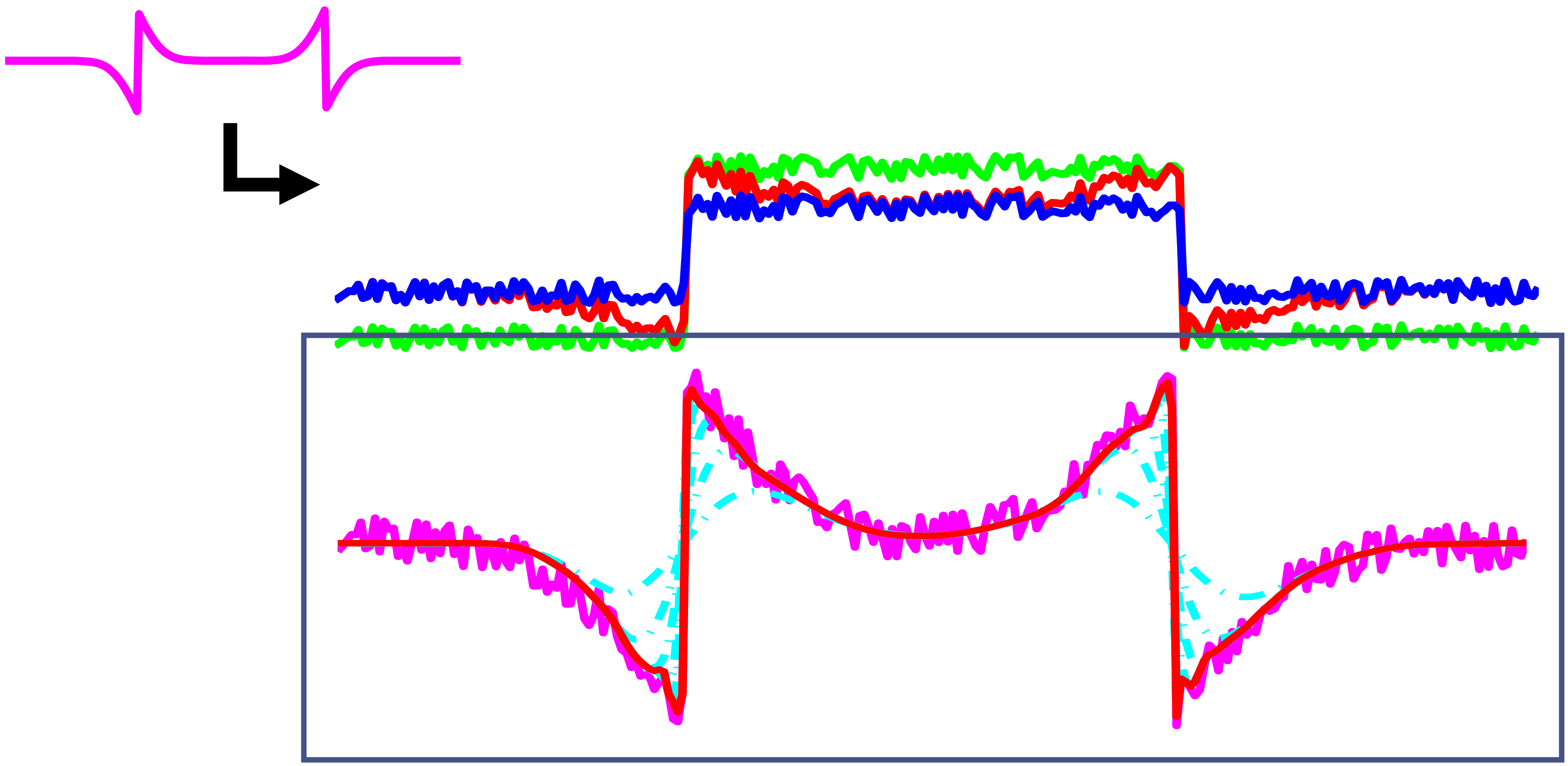


Well preserved signal is exaggerated by **unsharp masking**

Krawczyk et al. EG2007



Correct Profile for Textured Area



- Profile constructed directly from the reference image contains high frequency features which exaggerate texture
- **Sub-band components allow to select features**
  - **high frequency component present only at high contrast edge**

Krawczyk et al. EG2007

Multi-resolution Contrast Metric

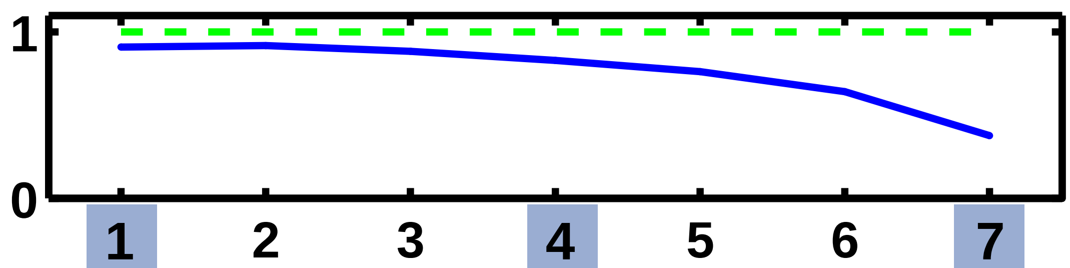
Reference HDR Image

Measure Lost Contrast at Several Feature Scales

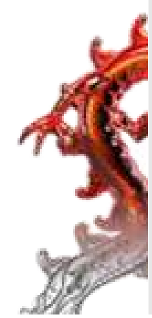
$$C_l = \frac{|Y - Y_{mean}|}{Y_{mean}}$$

Tone Mapped Image

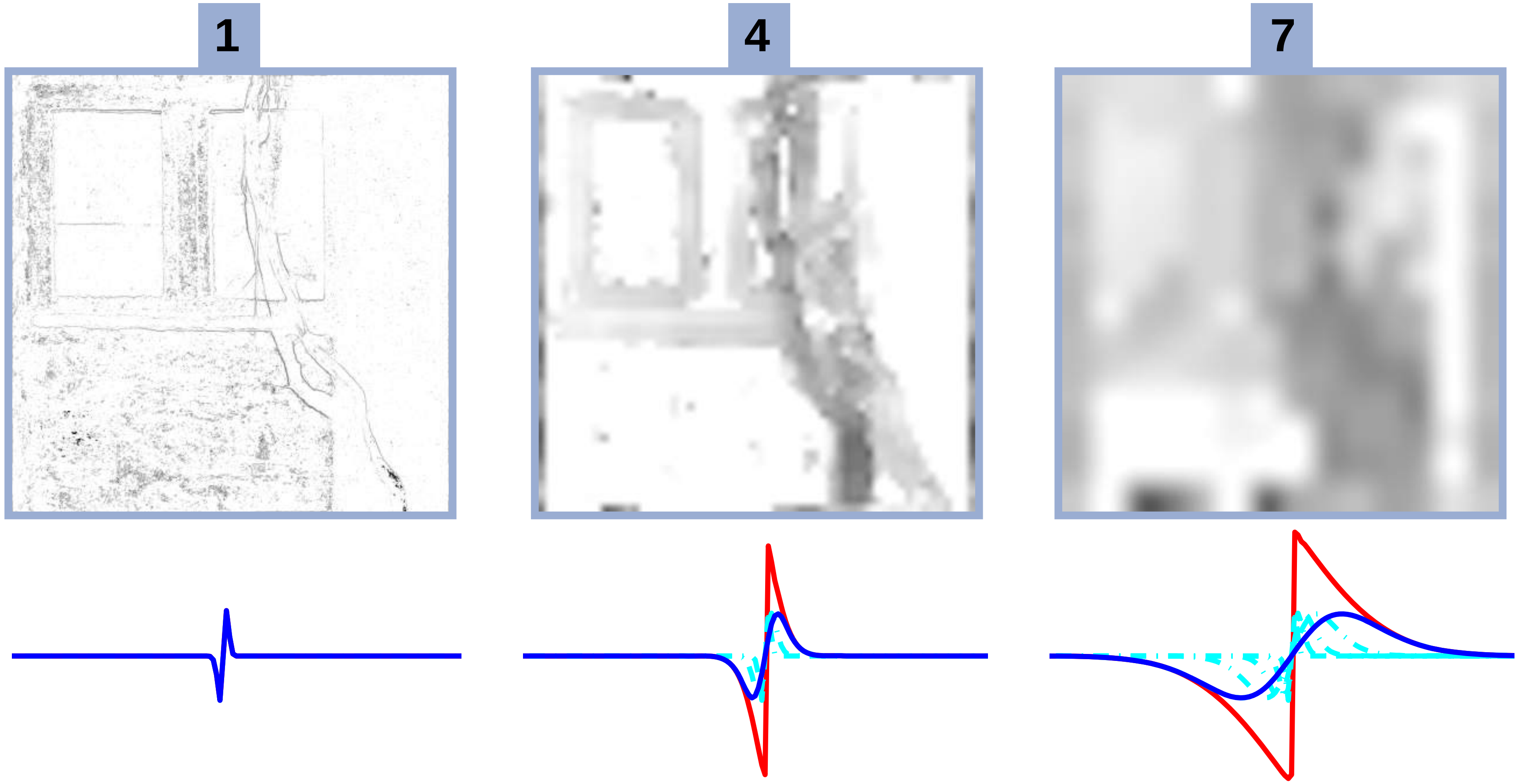
$$R_l = \frac{C_l^{inp}}{C_l^{ref}}$$



Contrast ratios at several scales

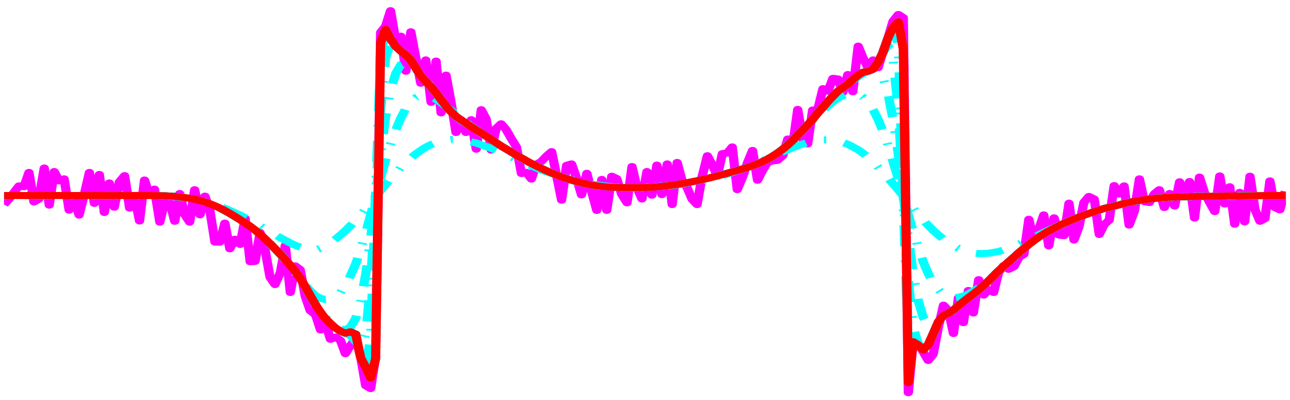


Link: Contrast Metric & Profiles



1. Contrast ratio at each scale defines the sub-band amplitude (blue)
2. Contrast for larger scales appears also on smaller scales
   - the full profile is always reconstructed (red)
3. Scale of contrast measure defines the profile size

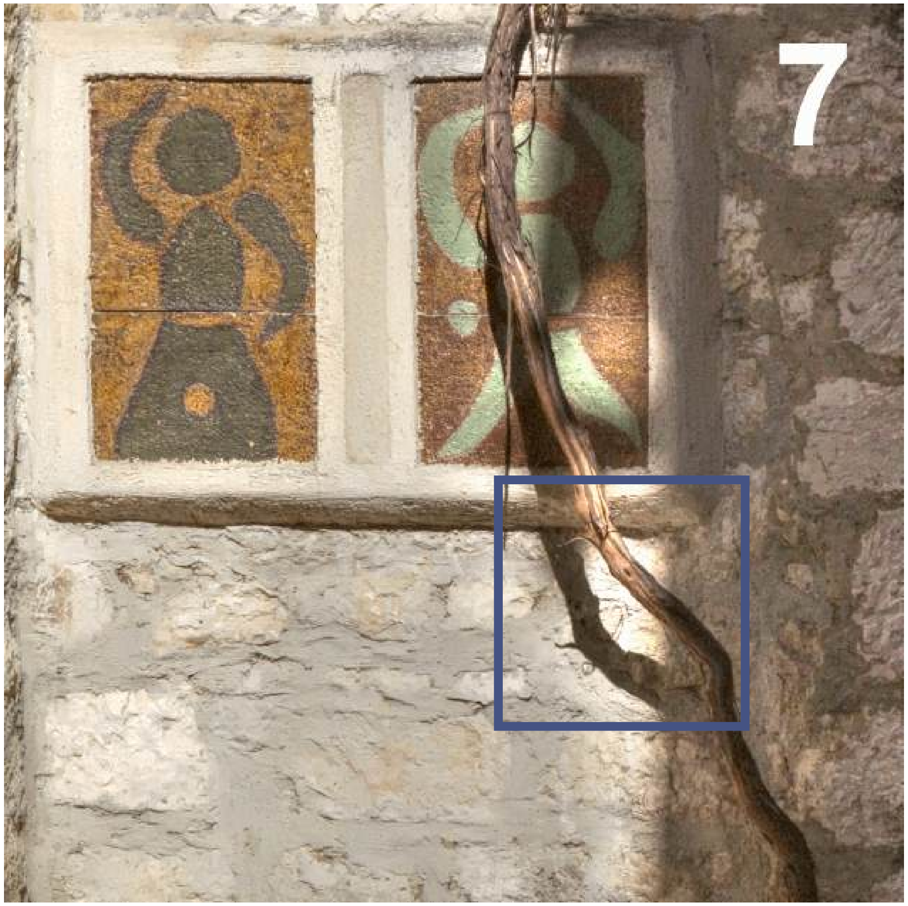
Formula: Countershading Profile



$$P = \sum_{l=1}^{N} \underbrace{(1 - \uparrow R_l)}_{\text{amplitude of profile}} \times \underbrace{(\log Y^{ref}_{\sigma(l-1)} - \log Y^{ref}_{\sigma(l)})}_{\text{sub-band component of profile}}$$

1. Contrast ratio $\boldsymbol{R_l}$ on scale $\boldsymbol{l}$ drives the amplitude of sub-band component of profile at size $\boldsymbol{l}$
2. Sum of $\boldsymbol{N}$ sub-band components gives the countershading profiles $\boldsymbol{P}$ that match the contrasts in the reference image

Adaptive Countershading

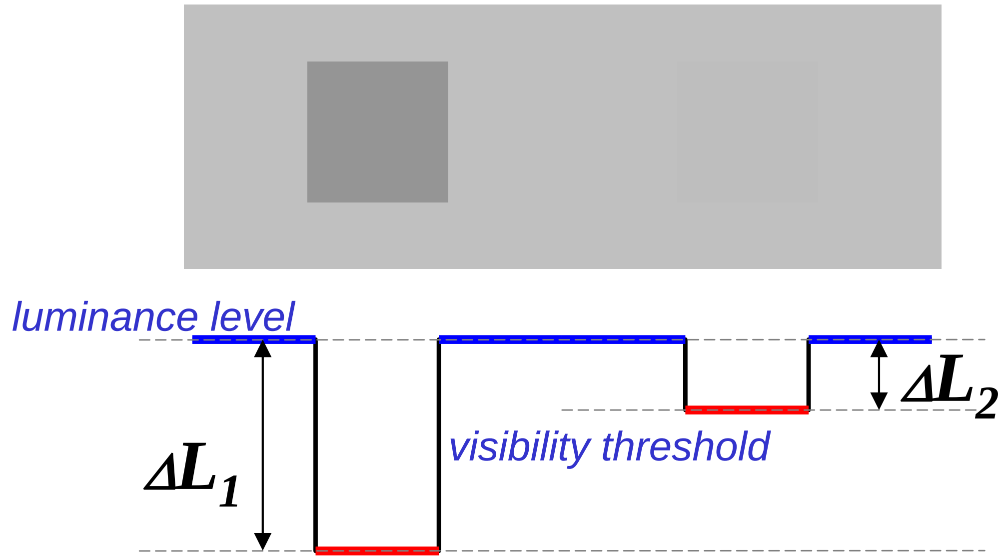


final contrast restoration

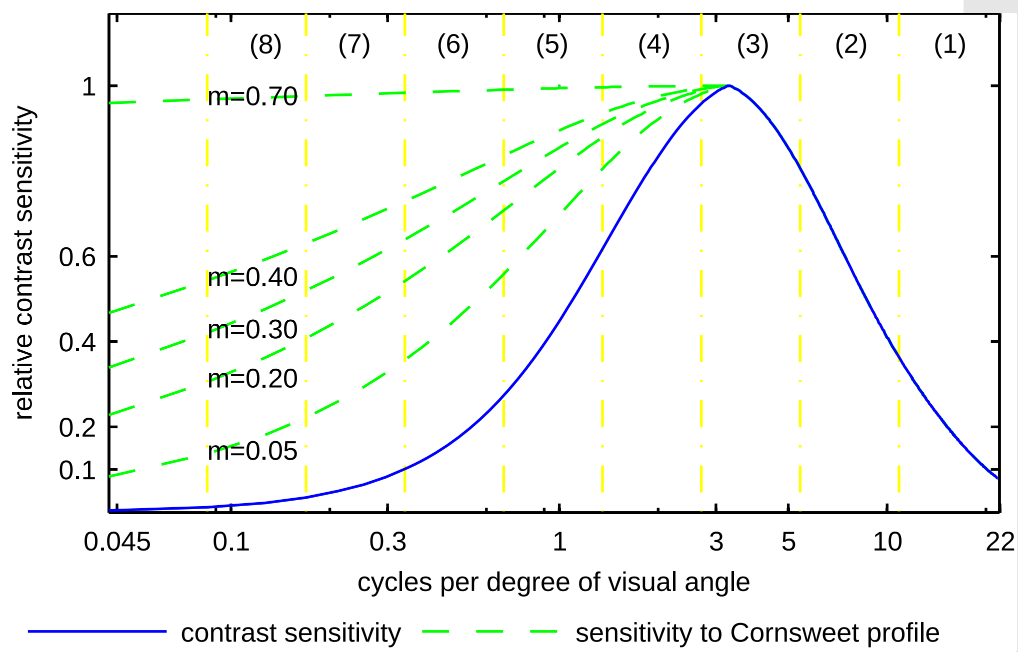
progress of restoration

- **Objectionable visibility of countershading profiles**

Visual Detection Model

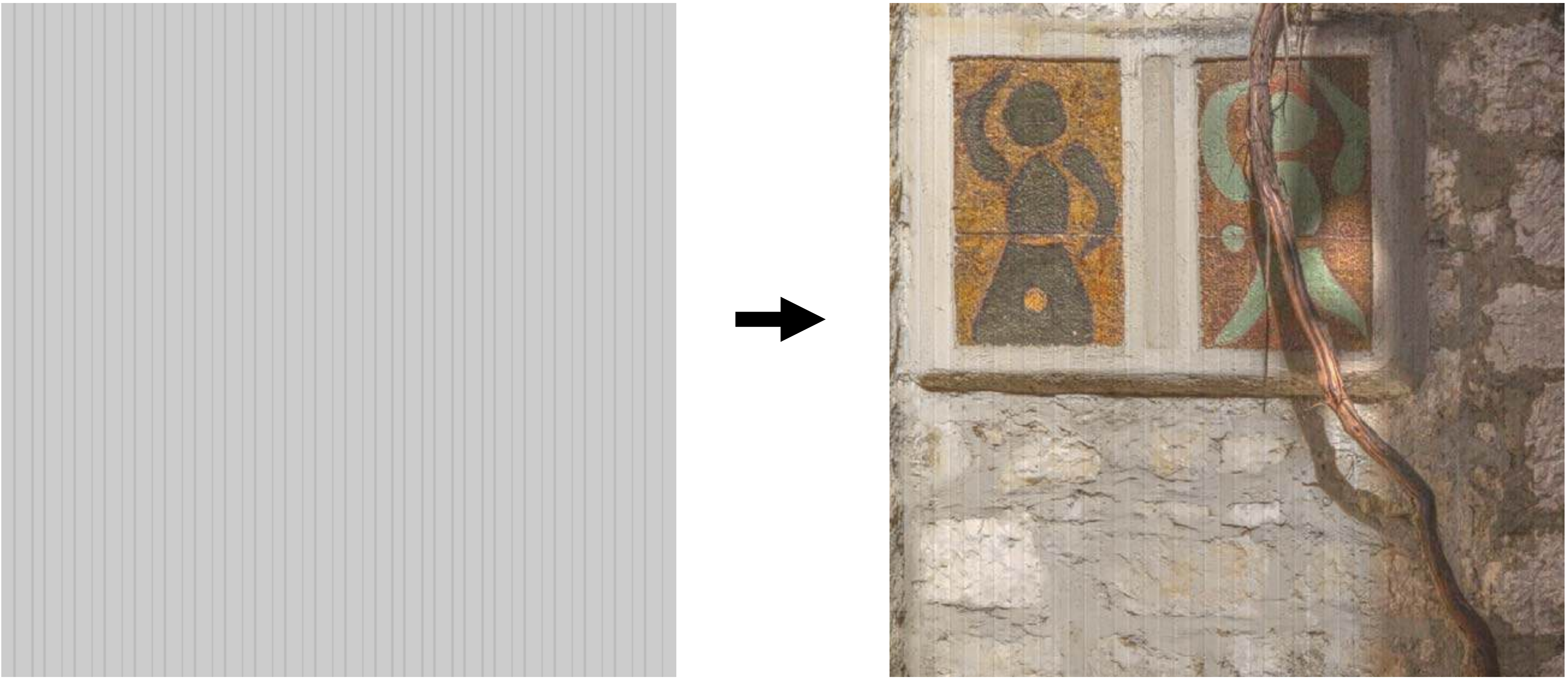


[Dooley and Greenfield, 1977]



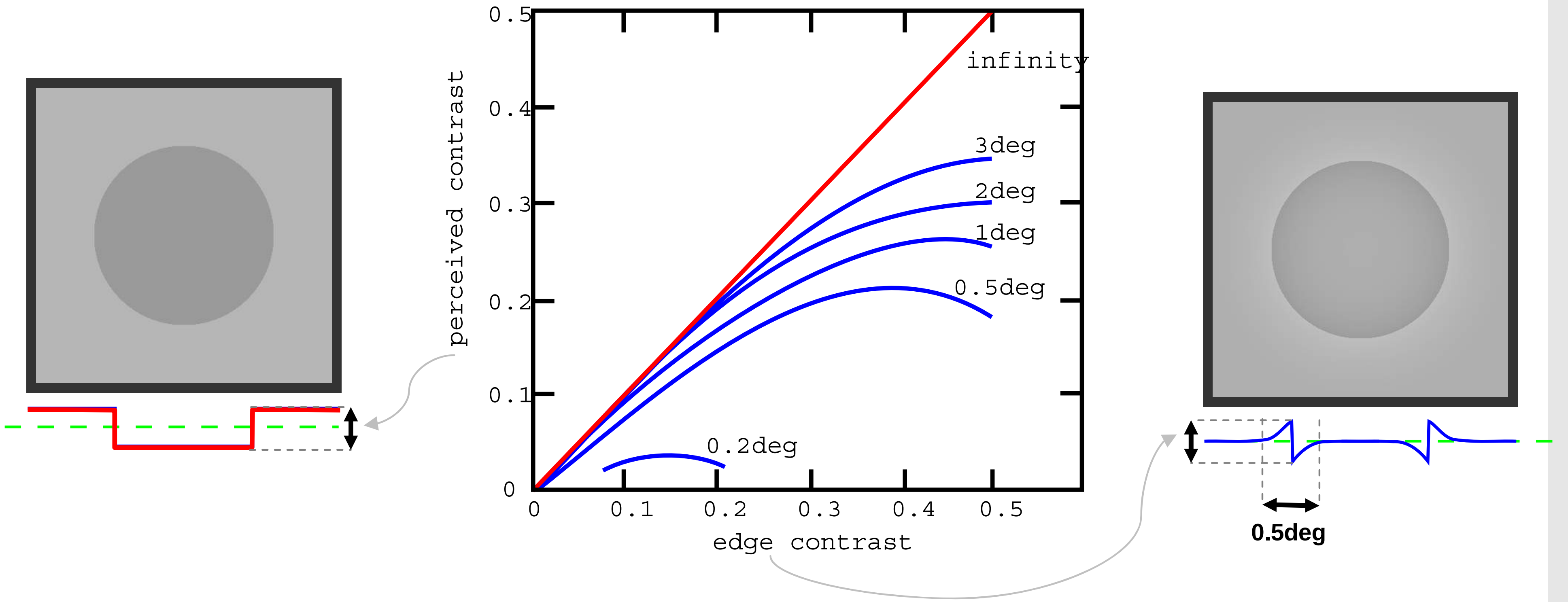
- **Luminance masking**
  - absolute luminance level L defines minimum perceivable luminance difference ΔL
  - defined by t.v.i. functions
- **Spatial contrast sensitivity**
  - reduced sensitivity to low frequencies
  - defined by CSF functions
  - improved by supra-threshold measurements of Cornsweet profile

Hiding Countershading Profiles



- Contrast masking
  - existing contrast masks new signals of similar orientation and frequency
  - defined by a power function of contrast present in an area
- <u>Essential improvement</u>
  - previous models allow for rather small amplitudes of profiles

Limits of Countershading Profiles



- Measurements plot for the Cornsweet effect
  - contrast at the profile edge (x) and the matching contrast at the step edge (y)
- Masking allows for stronger enhancement
- Maximum correction depends on profile size
  - natural images unlikely require correction of a large contrast with a small profile

Adaptive Countershading

without visual model

with visual model

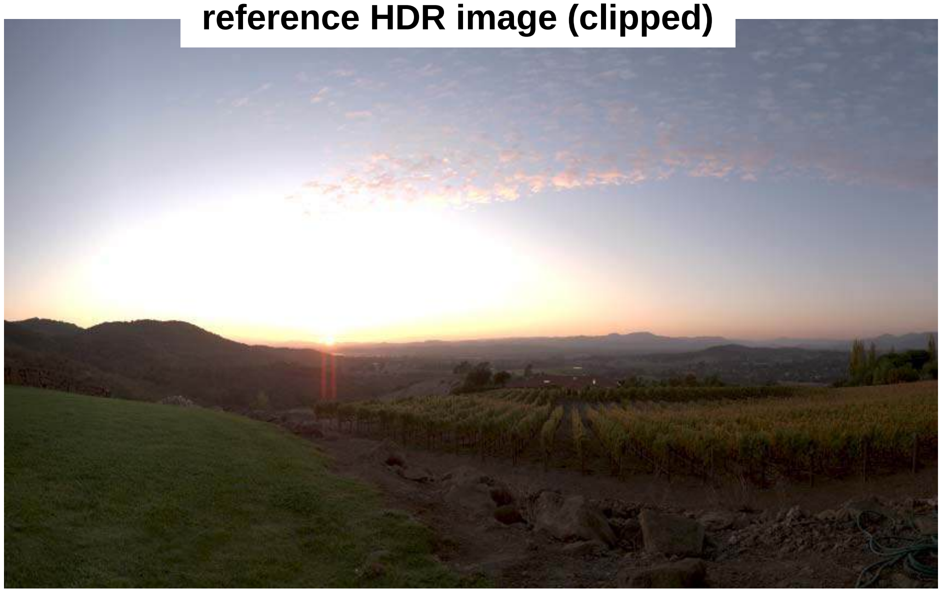
Restoration of TM Images (1/3)

(a) global tone mapping

(b) contrast equalization tone mapping

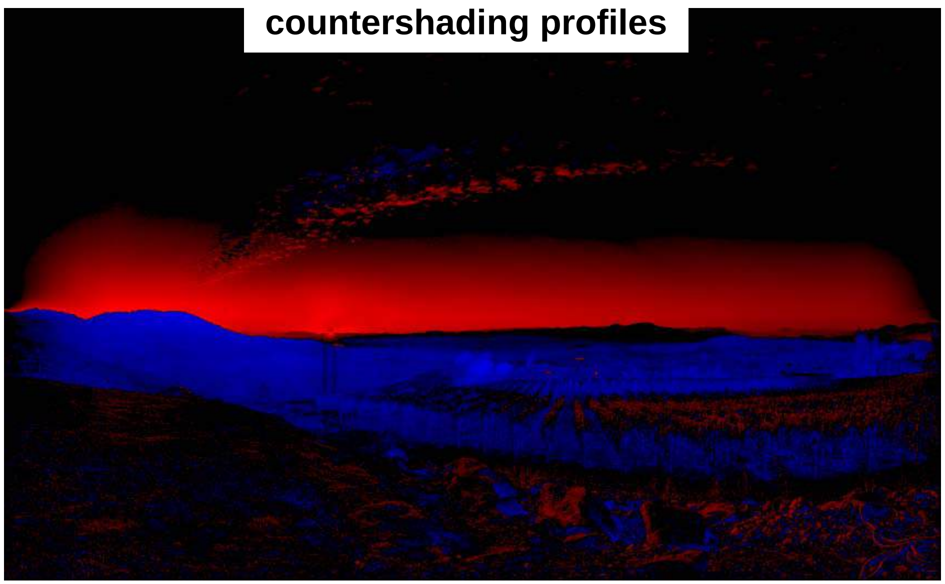
Restoration of TM Images (2/3)

reference HDR image (clipped)



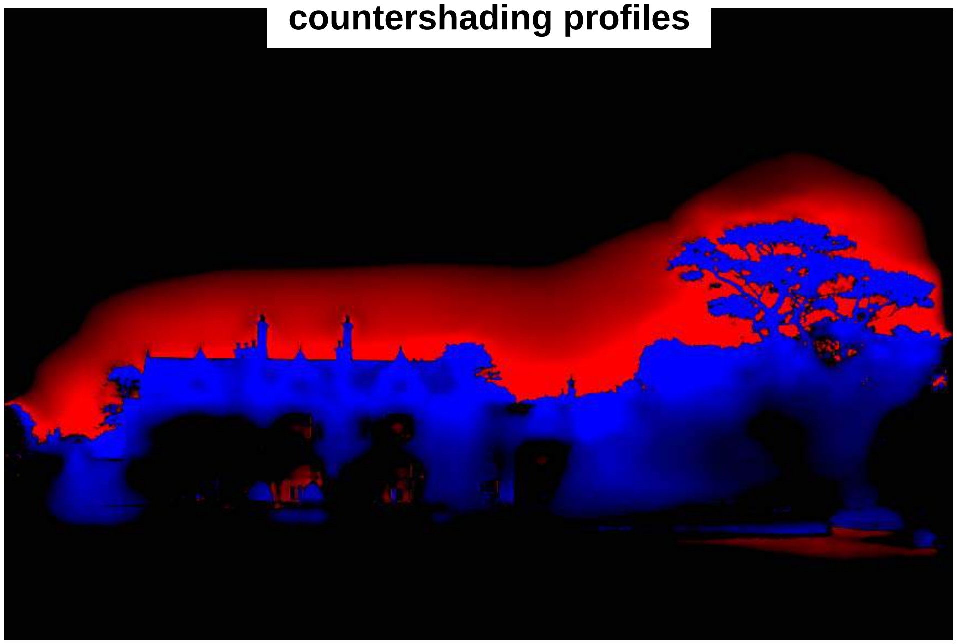
countershading of tone mapping

countershading profiles



tone mapping

Restoration of TM Images (3/3)



C-shading vs. Unsharp Mask

adaptive countershading

unsharp masking

tone mapping

Countershading Variants

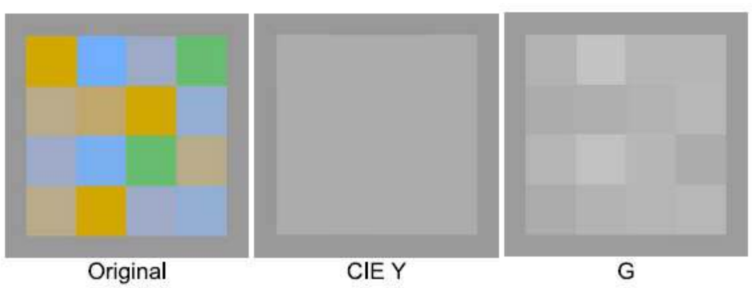
- Traditional countershading
  - performed in the achromatic channel to enhance perceived luminance contrast

Cross-modal approach

  - Use depth signal to derive countershading profile
  - Countershading over chromatic channels enhances the overall image contrast

Color2Grey:

  - dimensionality reduction 3->1: may lead to information loss
  - countershading in the achromatic channel used to reproduce lost chromatic contrast



Purpose: Contrast Restoration

Reference Depth Map

Measure Missing Contrast at Several Feature Scales

Enhance Missing Contrast in The Input Image

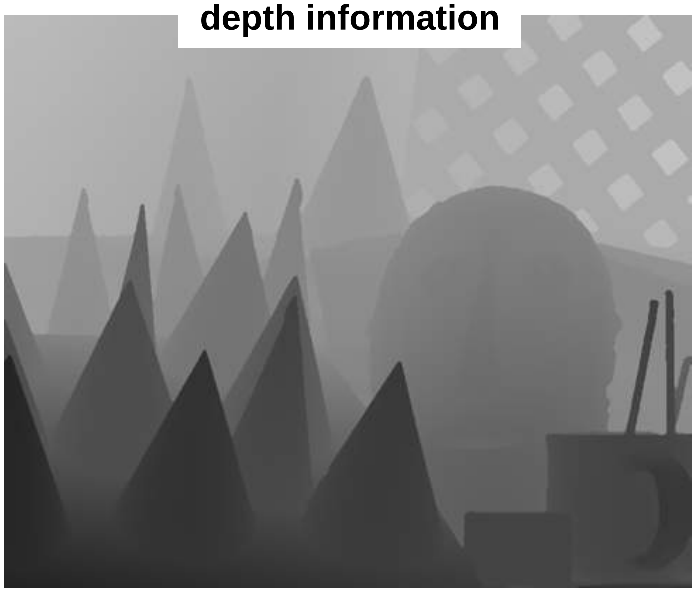
Input Image

Enhanced Image

Luft et al. SIG2008

Depth Map as Contrast Reference

depth information



original image

adaptive countershading

depth darkening [Luft2006]

Luft et al. SIG2008

Colourfulness Countershading

- "Strasbourg": Gradient method tone mapping, strong global contrast loss so strong restoration effect.
- Colourfulness contrast at border between sky and buildings
  - promotes FG/BG separation
  - creates impression of greater dynamic range
  - increases impression of depth

Smith et al. EG2006

Countershading Results (original)

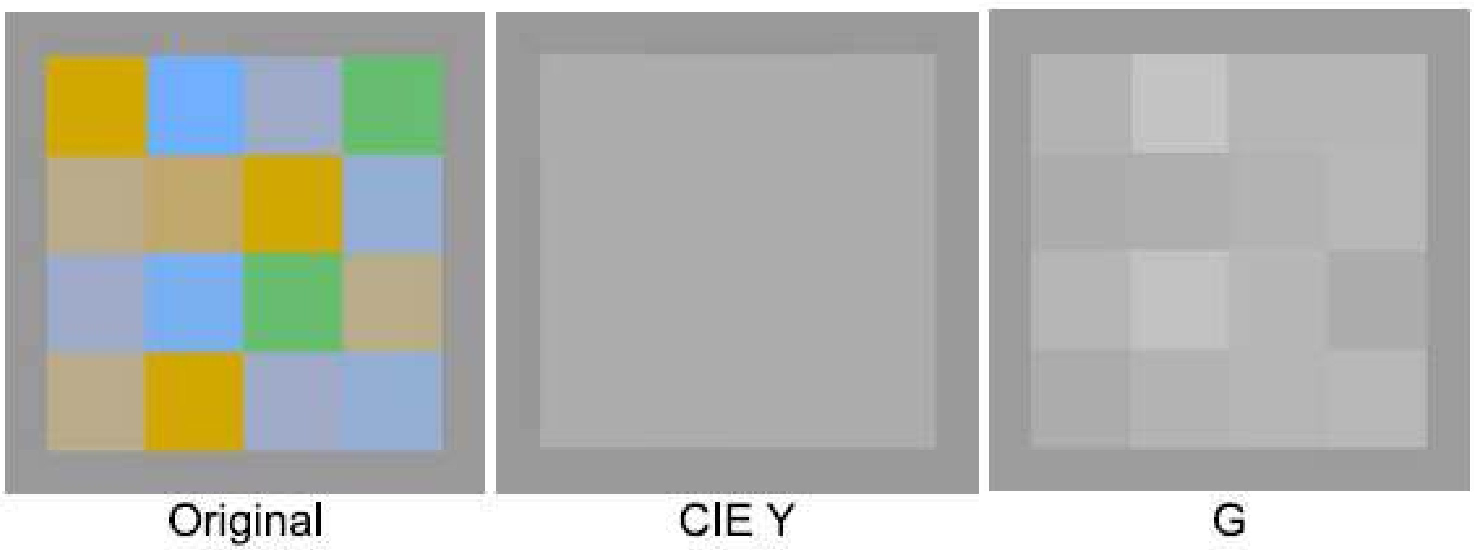
Smith et al. EG2006

Countershading Results (chroma enhancement)

Smith et al. EG2006

Color2Grey Application

- Isoluminant color pattern transformed to grey G using Helmholz-Kohlraush effect, which takes into account the contribution of chromatic component into brightness



Smith et al. EG2008

Color2Grey Application

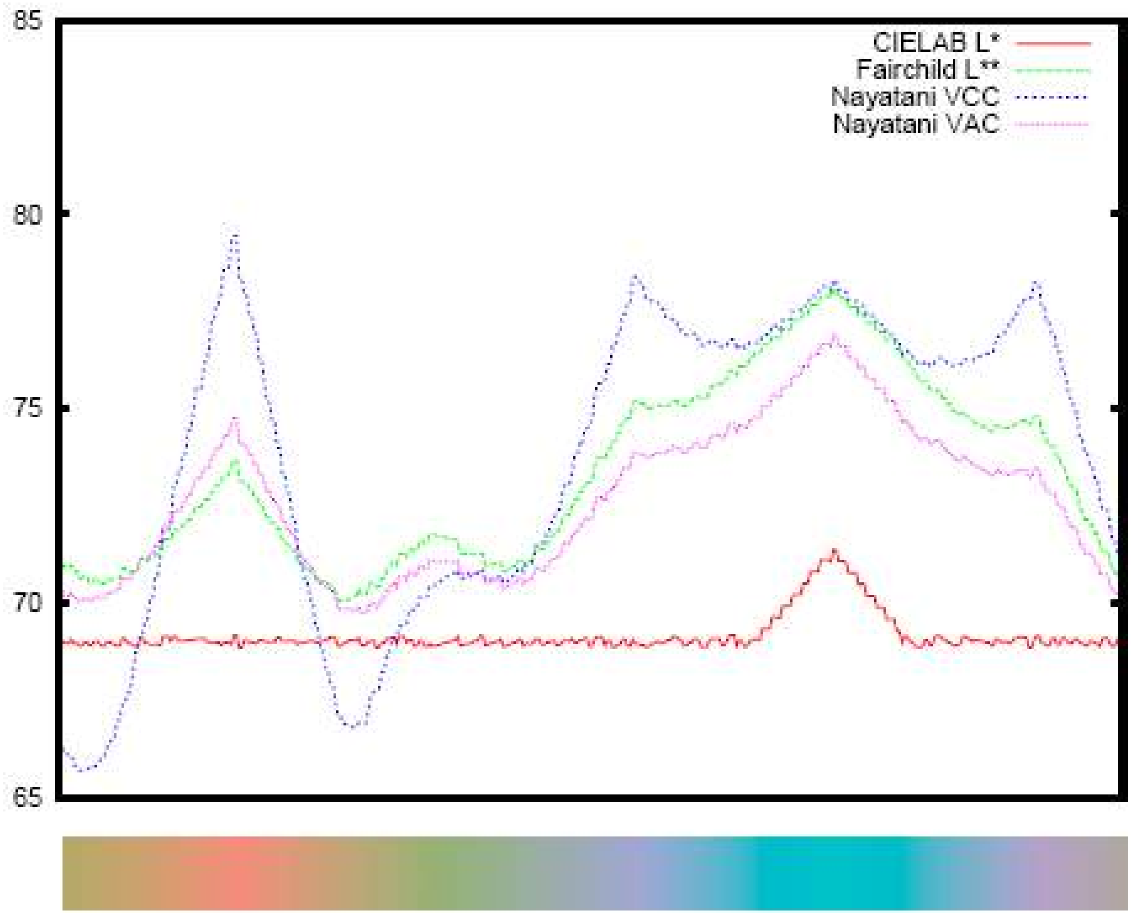


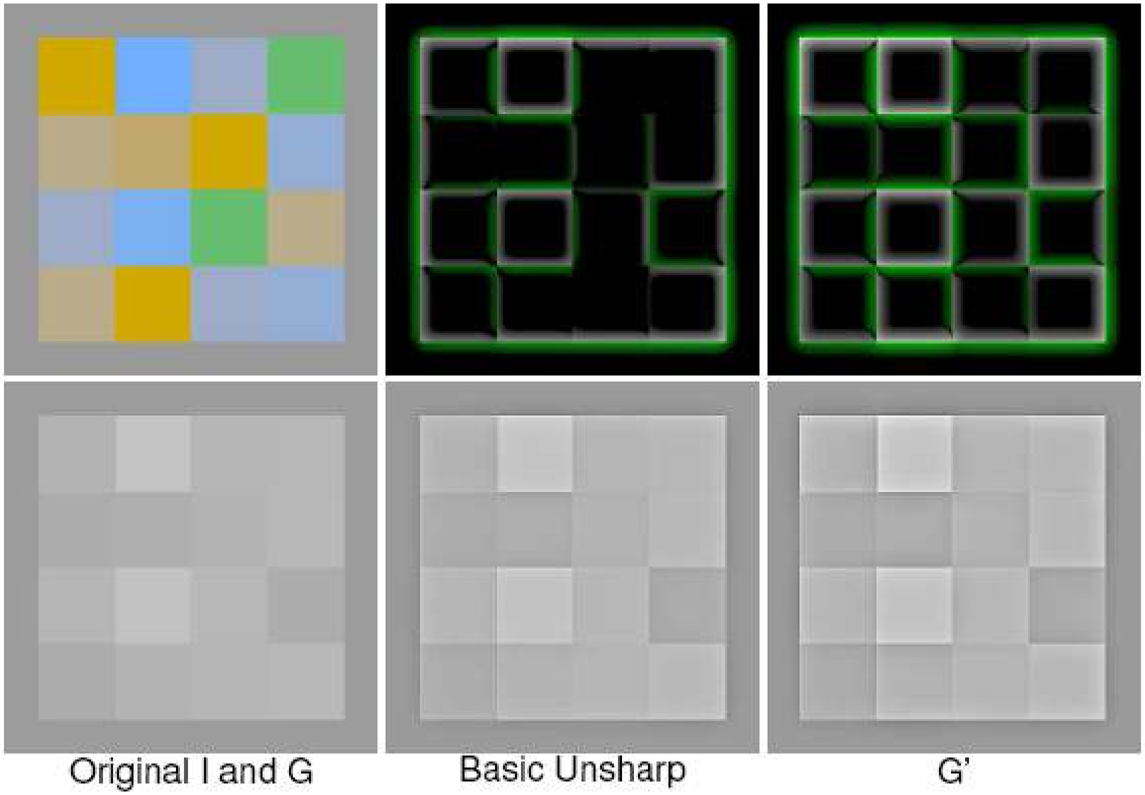
Figure 1: *Lightness values from various H-K effect predictors applied to a spectrum of isoluminant colours, compared to CIE L**.

Smith et al. EG2008

Color2Grey Application

- G'_{L*} : The effect of adding multi-resolution countershading correction $h_i(G_{L*})$ (*upper-left*) to the greyscale image G_{L*} (*lower-left*)

$$G'_{L^*} = G_{L^*} + \sum_{i=0}^{n-1} k_i \lambda_i h_i(G_{L^*})$$

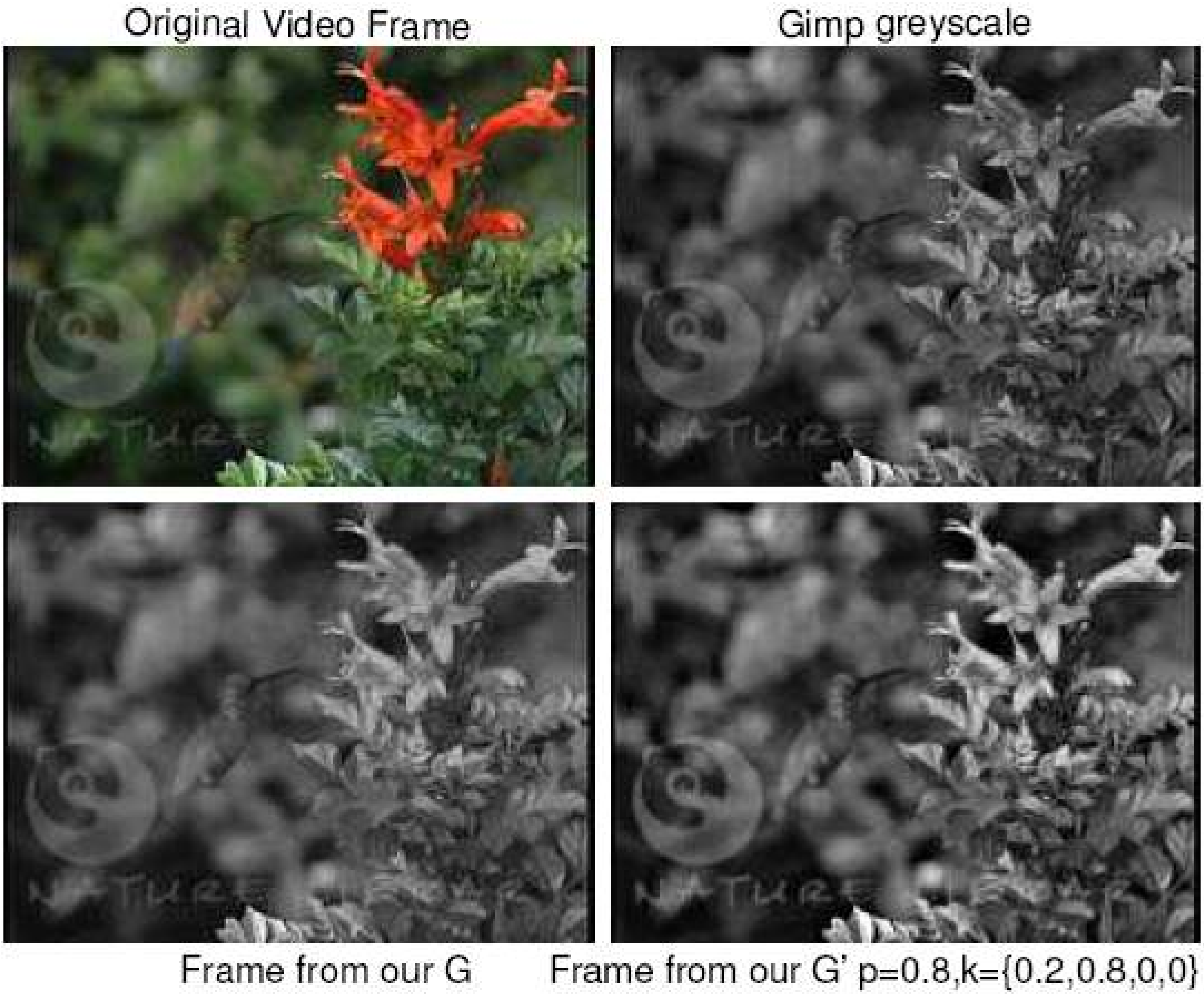


The correction is driven by contrast in chroma channels of the original image I (*upper-left*)

$$\lambda_i = \left(\frac{\Delta E(h_i(I))}{|h_i(G_{L^*})|} \right)^p$$

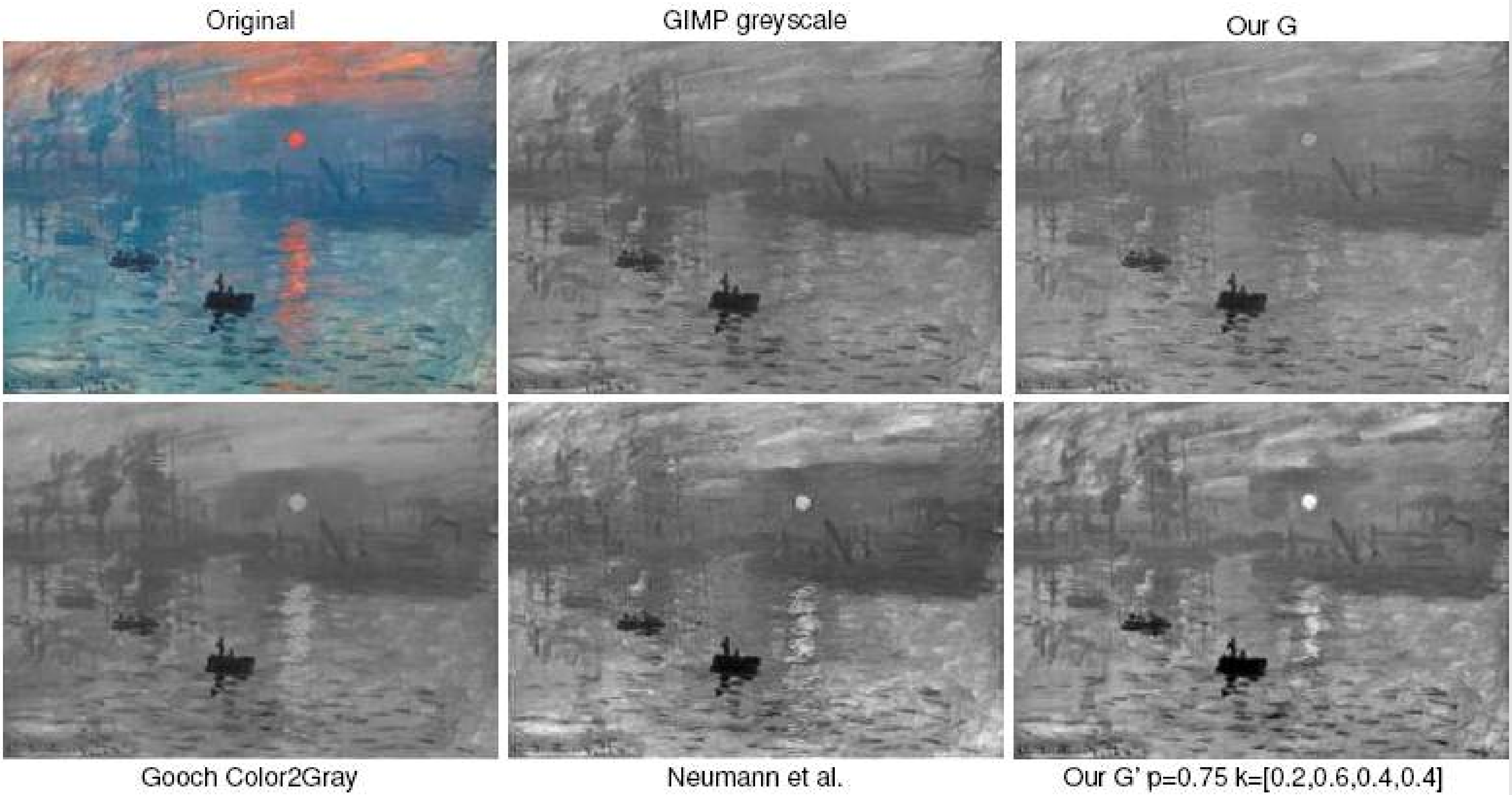
Smith et al. EG2008

Color2Grey Application



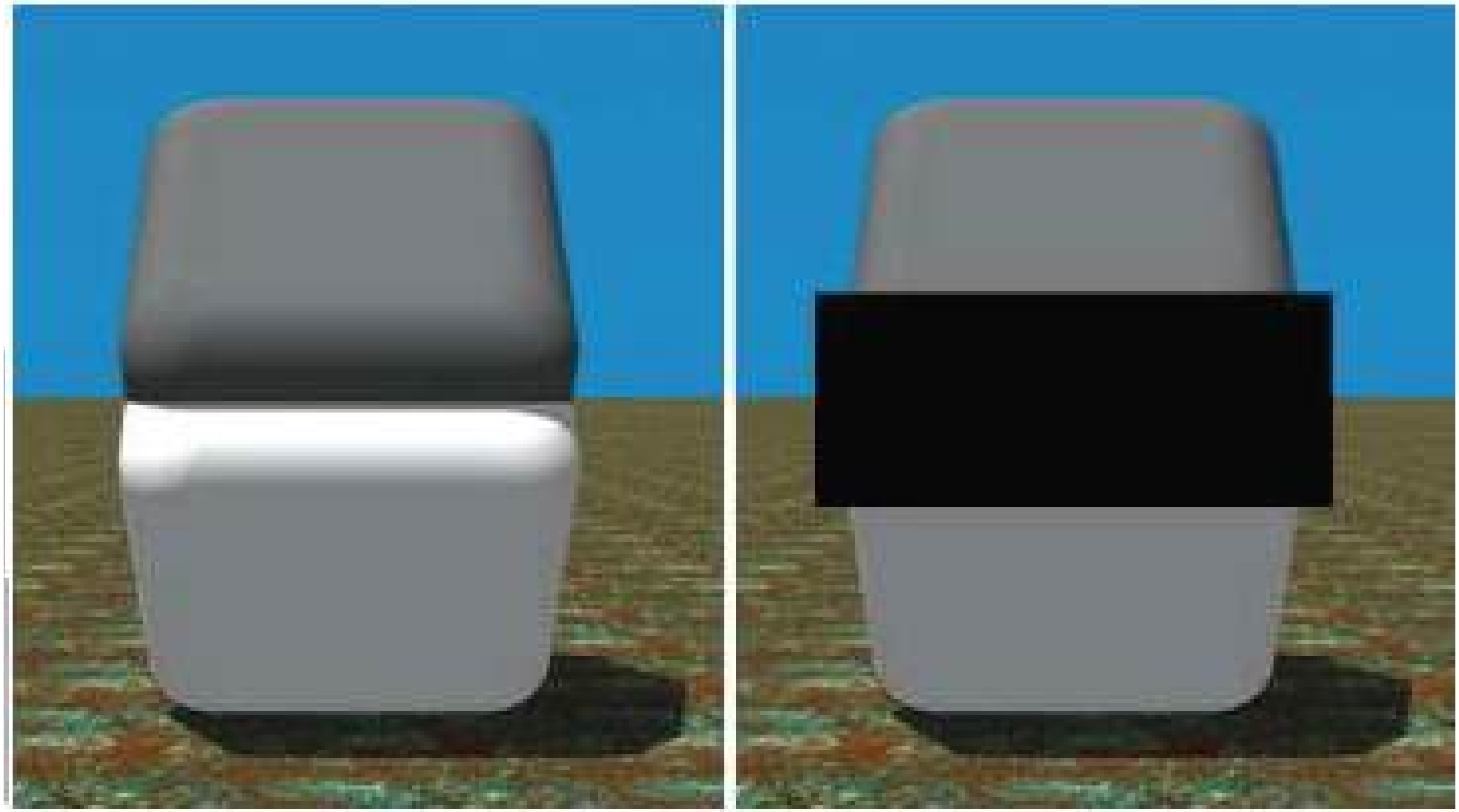
Smith et al. EG2008

Color2Grey Application



Smith et al. EG2008

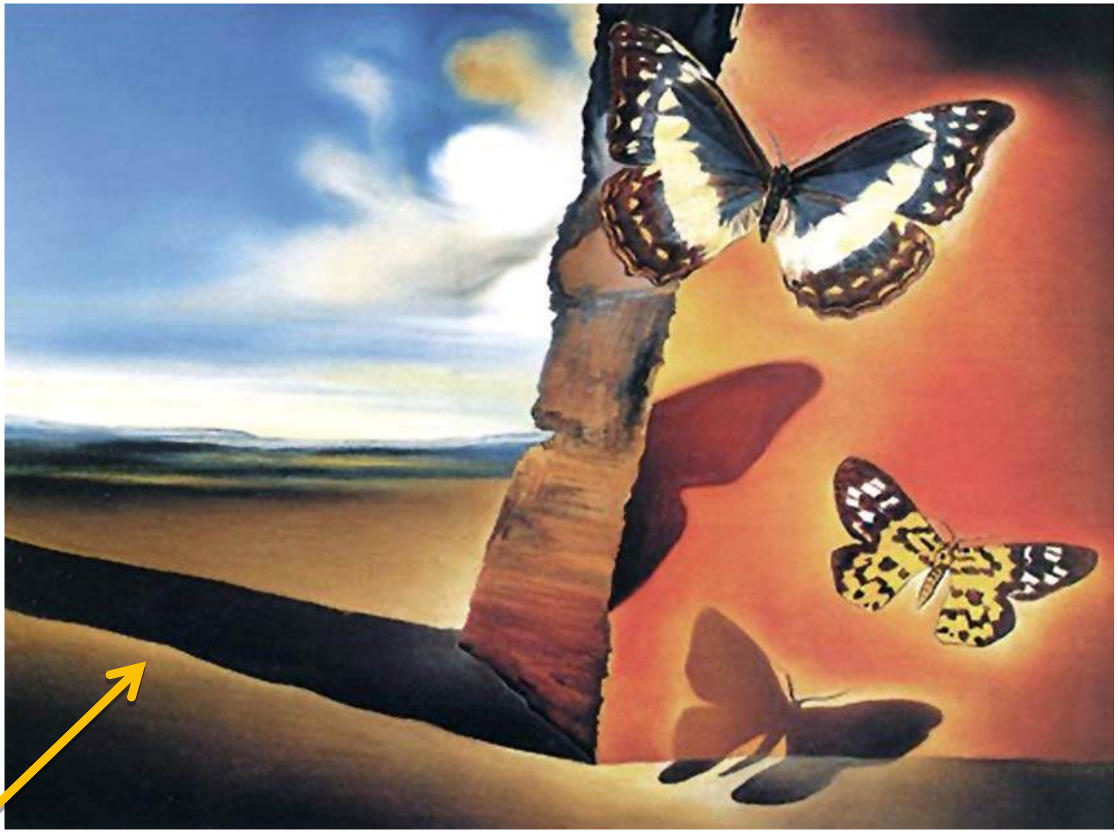
Countershading in 3D?



3D Cornsweet Illusion

Purves-Lotto illusion: much stronger effect in 3D

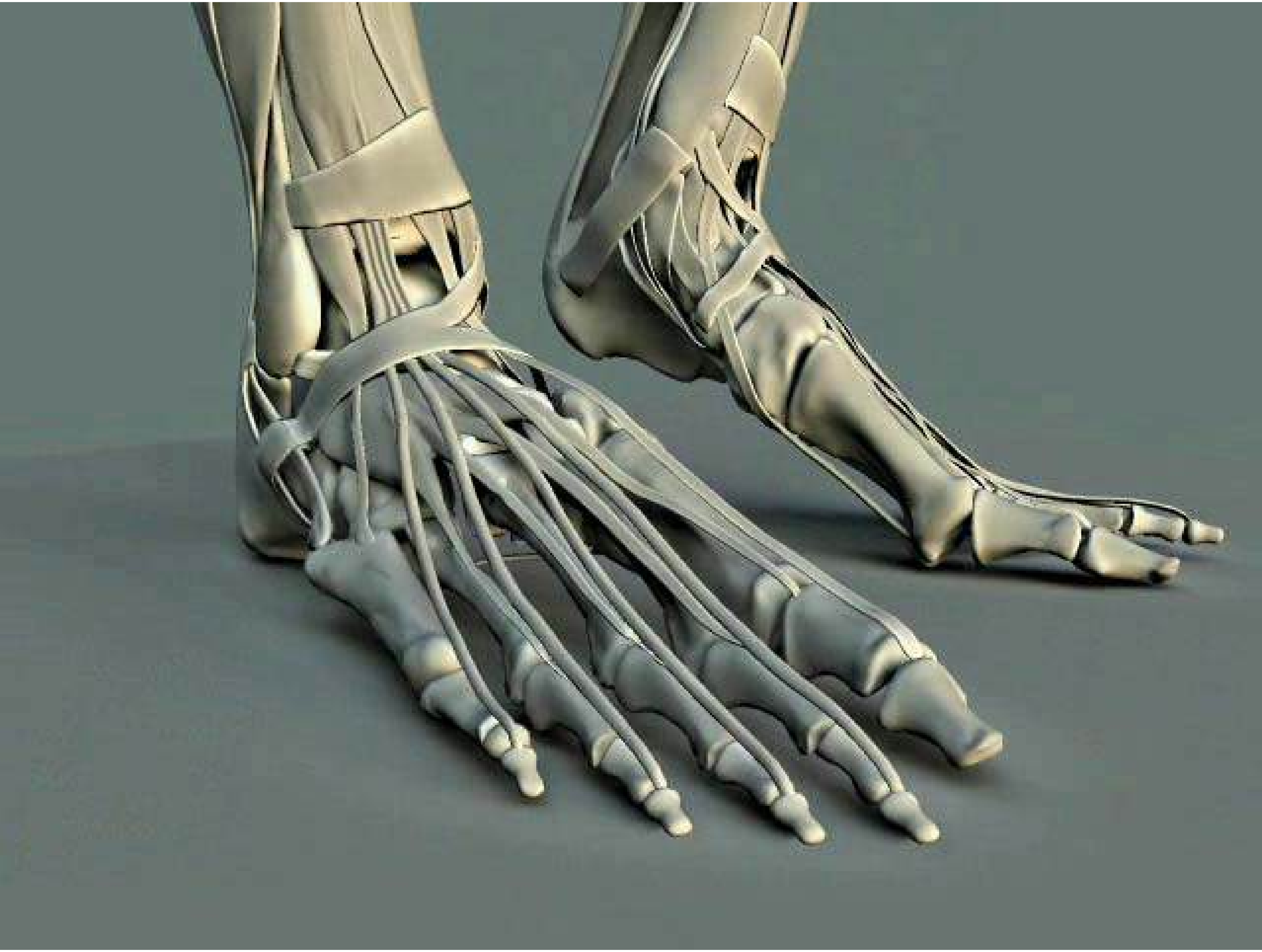
Scene-aligned Countershading



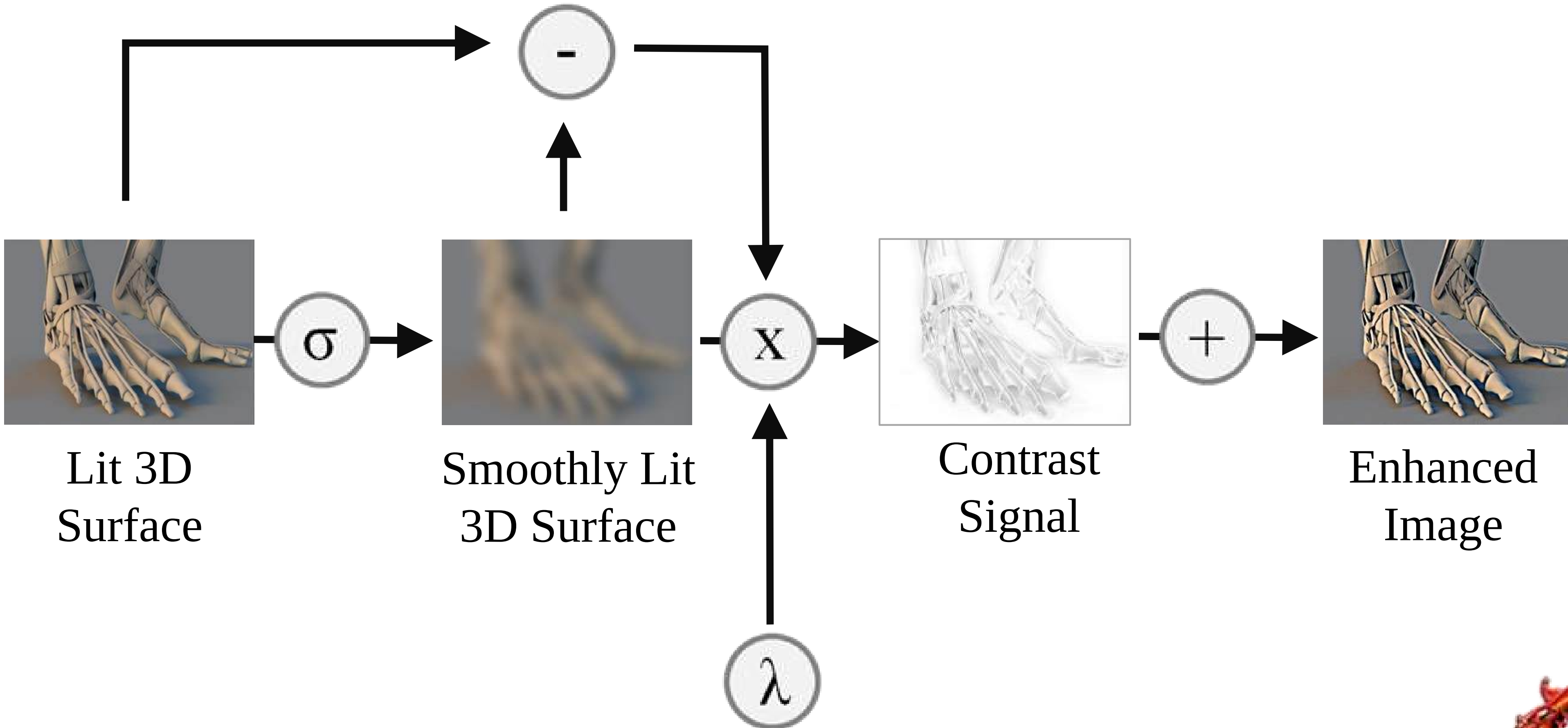
S. Dalí, *Landscape with butterflies*

Scene-aligned Countershading

G. Seurat, *Bathers at Asnieres*



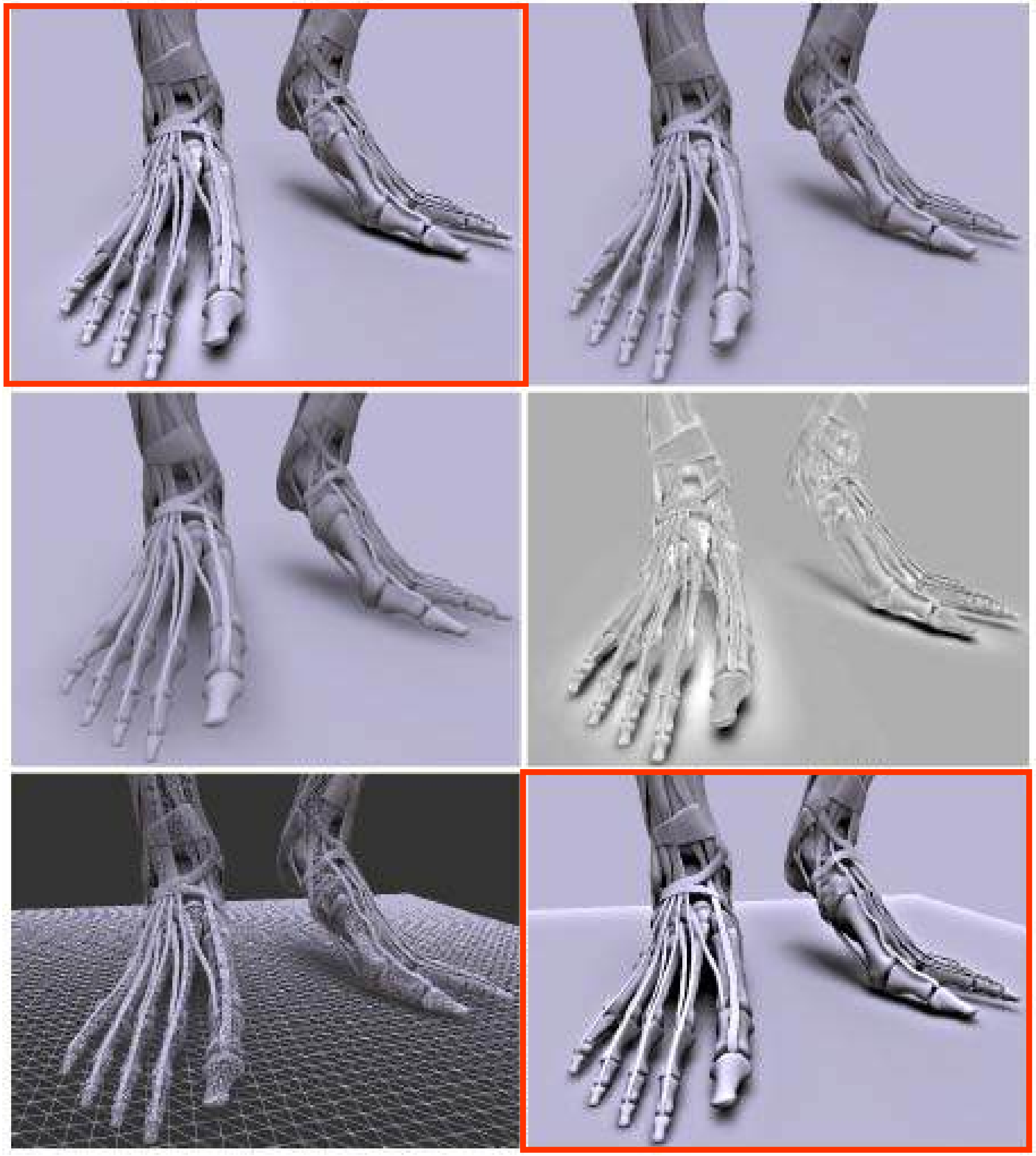
3D Unsharp Masking



$$U(S) = S + \lambda(S - S_\sigma)$$

Ritschel et al. SIG2008

3D Unsharp Masking



3D unsharp masking

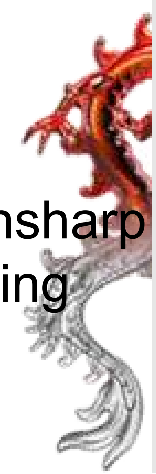
Original image

3D blurred signal

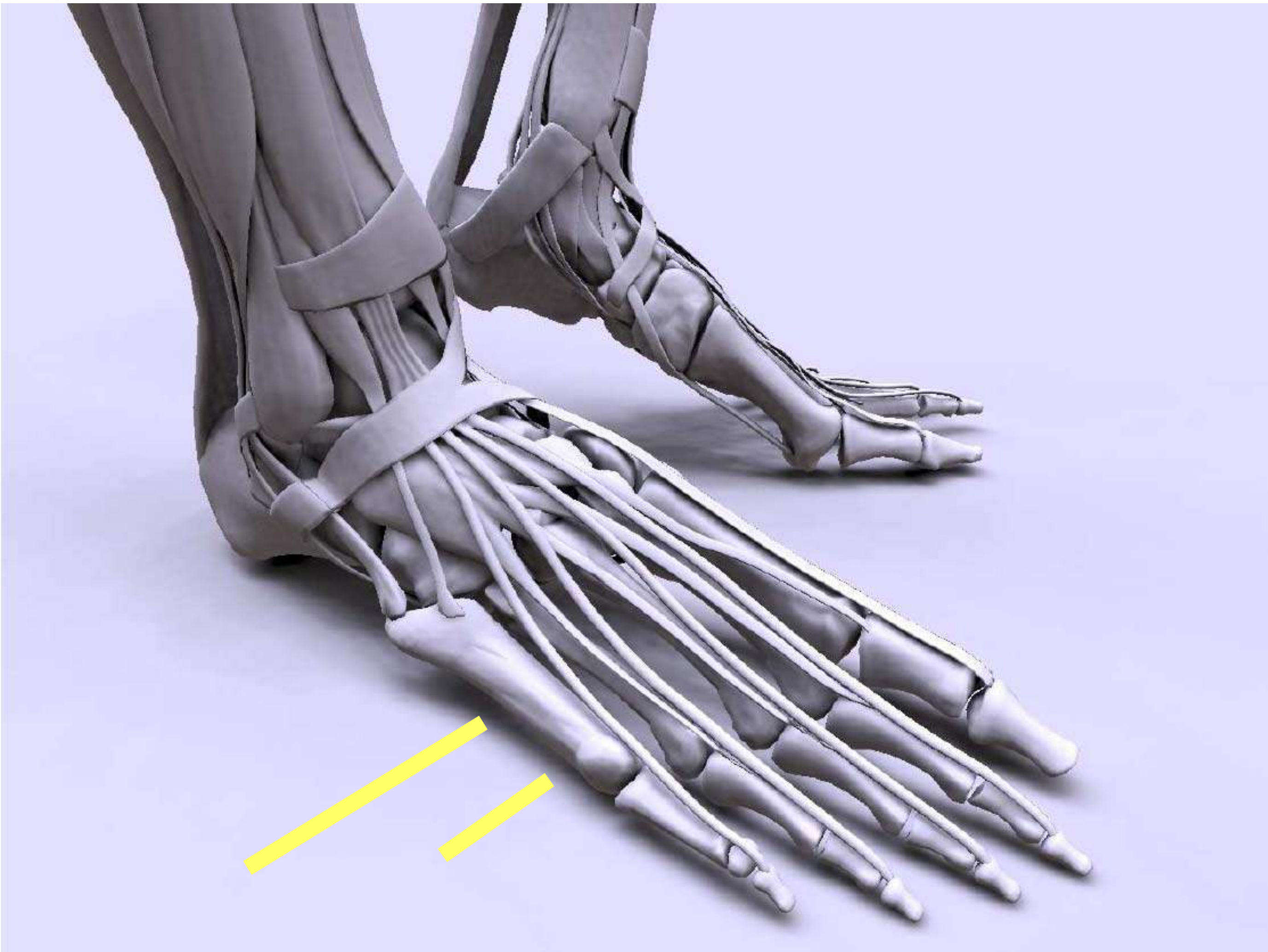
Enhancement signal

Mesh

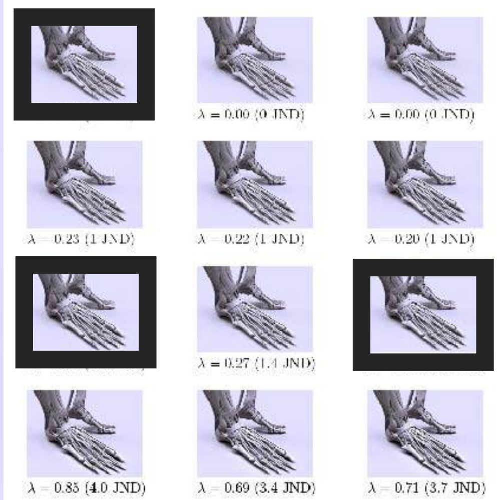
2D unsharp masking



Adjustable Effect



Width σ



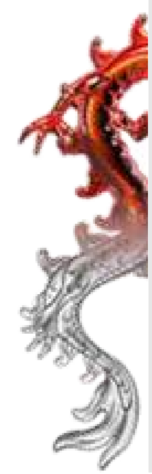
Strength λ

$$U(S) = S + \lambda(S - S_\sigma)$$

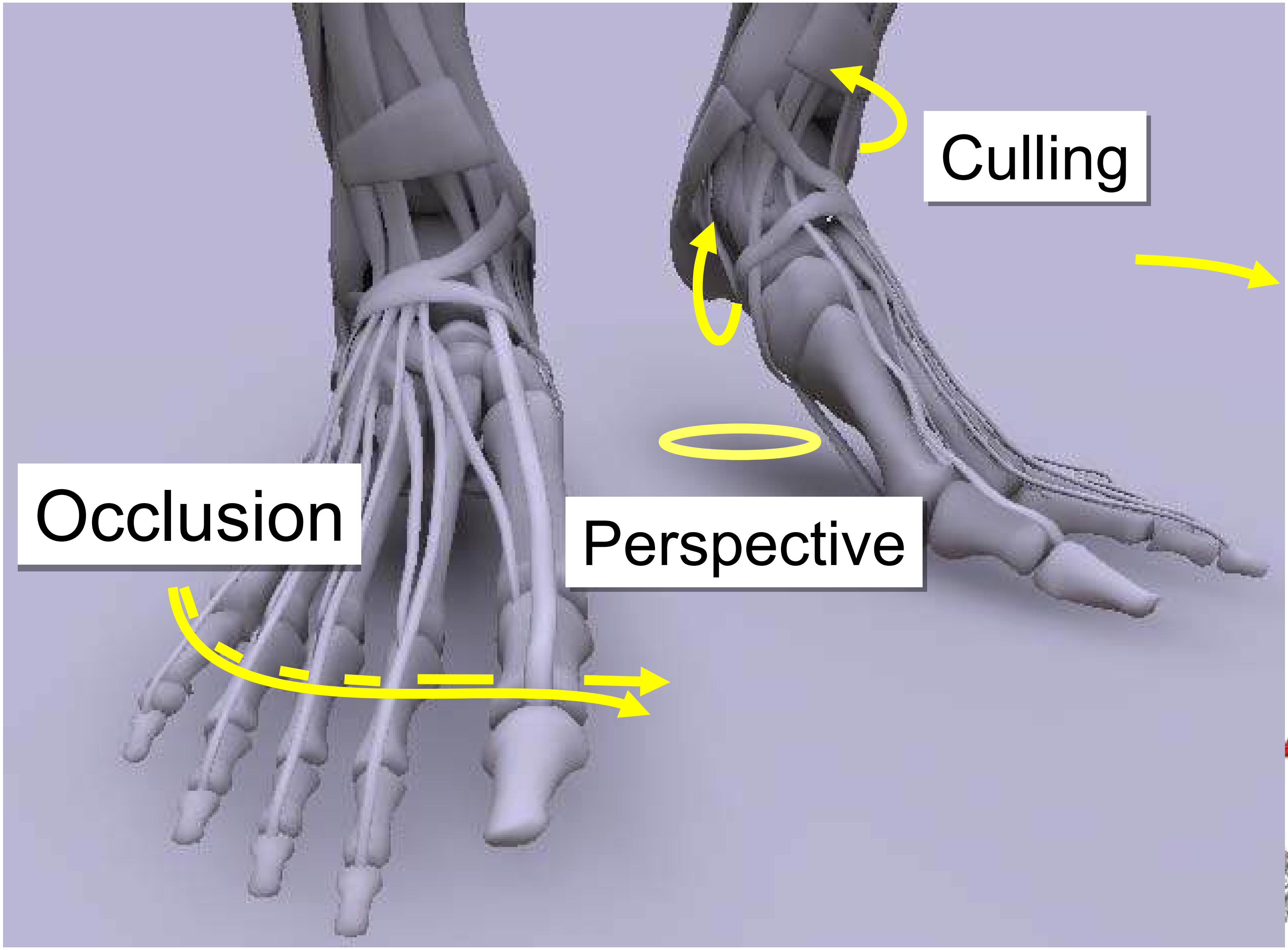
2D vs. 3D Unsharp Masking Comparison

SIGGRAPHASIA2011 HONG KONG

	2D	3D
Signal	Image	Lit Surface
Smoothing	(Gaussian) Image Blur	Laplacian Surface Blur
Representation	Pixels	Lit vertices and pixels
Smoothness σ	Image distance	Geodesic world distance
Strength λ	Factor	Factor



3D Unsharp Masking: Scene Coherence



Ritschel et al. SIG2008

Complex Mesh

3D unsharp masked rendering

Original rendering

Ritschel et al. SIG2008

Enhanced Text Contrast in the Shadow

3D unsharp masking

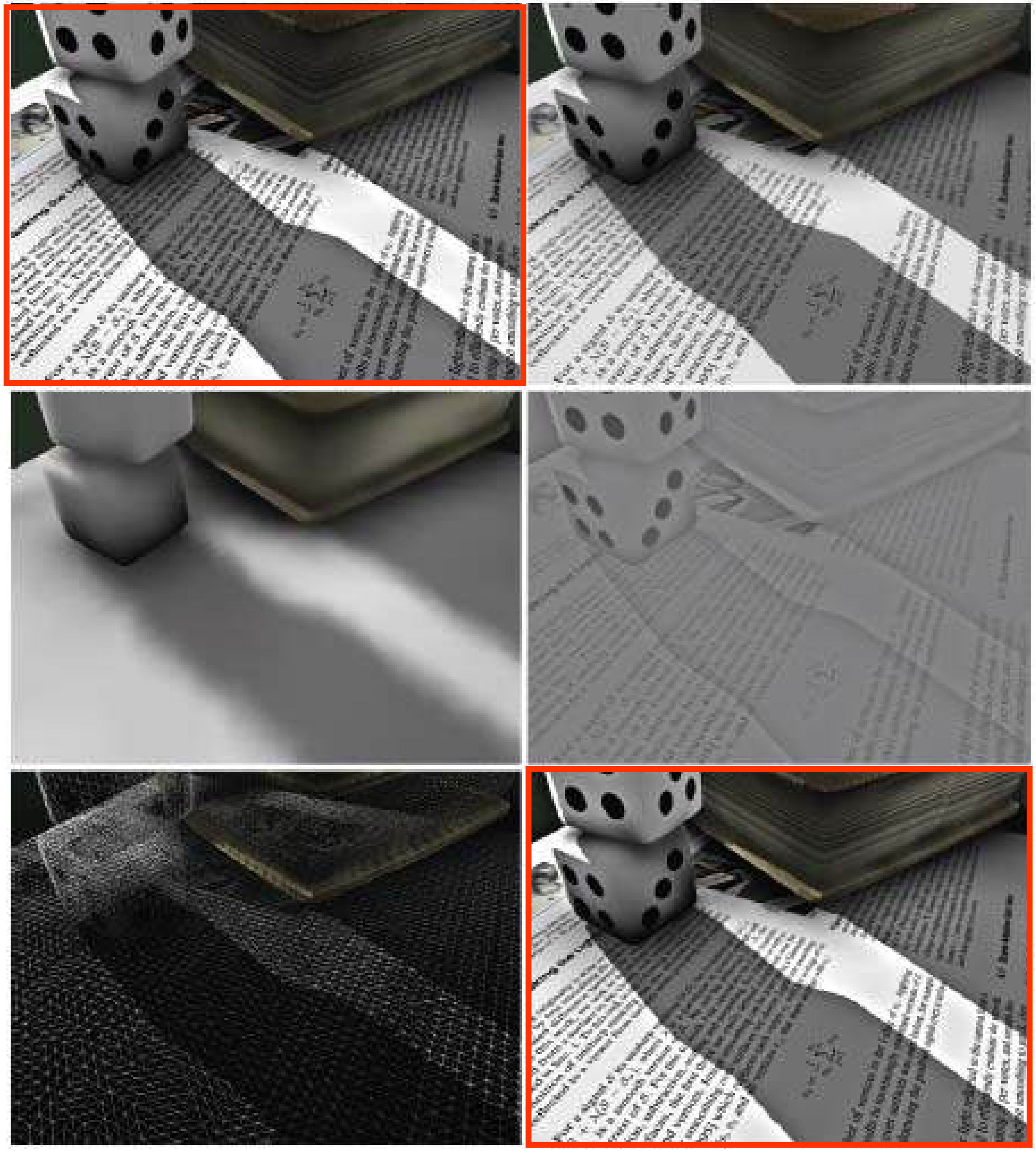
Original image

3D blurred signal

Enhancement signal

Mesh

2D unsharp masking

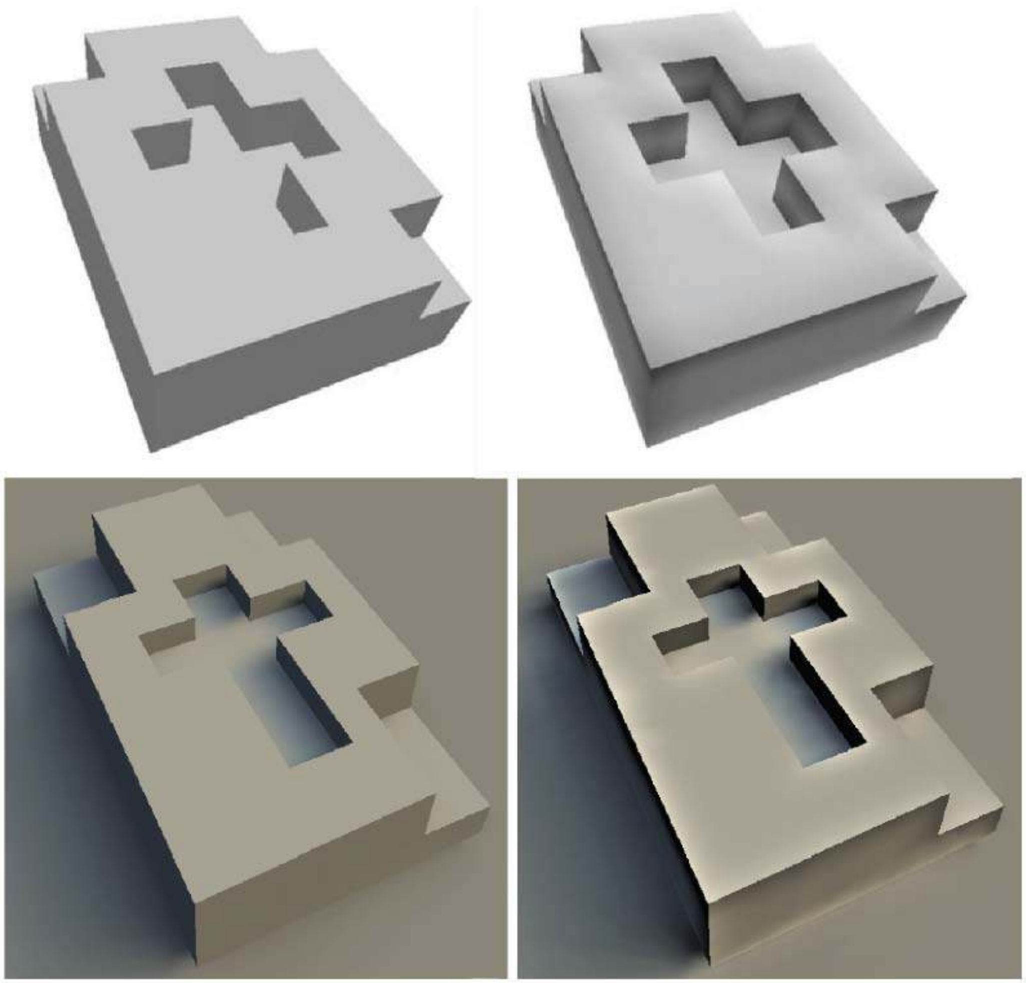


Results – Legibility

Normal Enhancement

- Only **geometric** term
  - Shadows ?
  - Hightlights ?
  - Reflectance ?
- Vertex resolution
- 3D unsharp masking: Pixel resolution

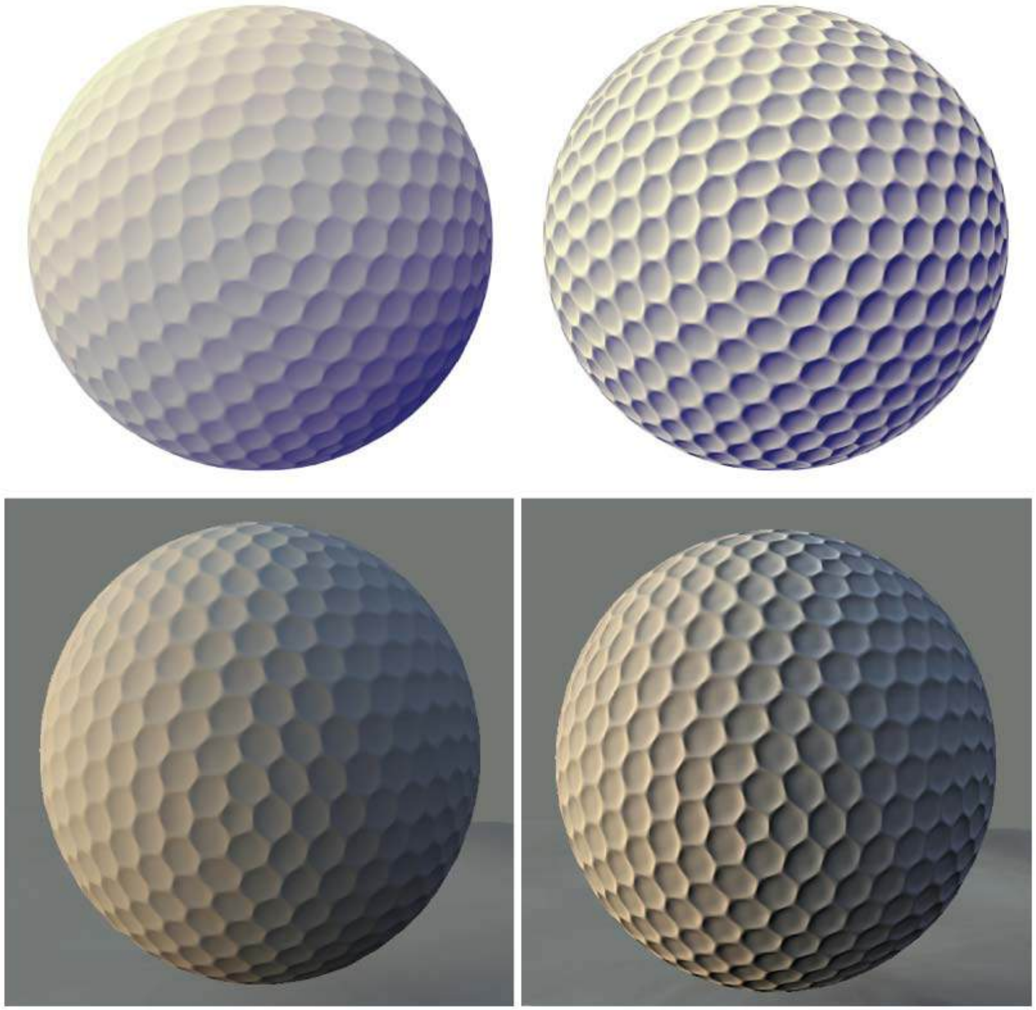
Cignoni et al. ´05, C & G Vol. 29



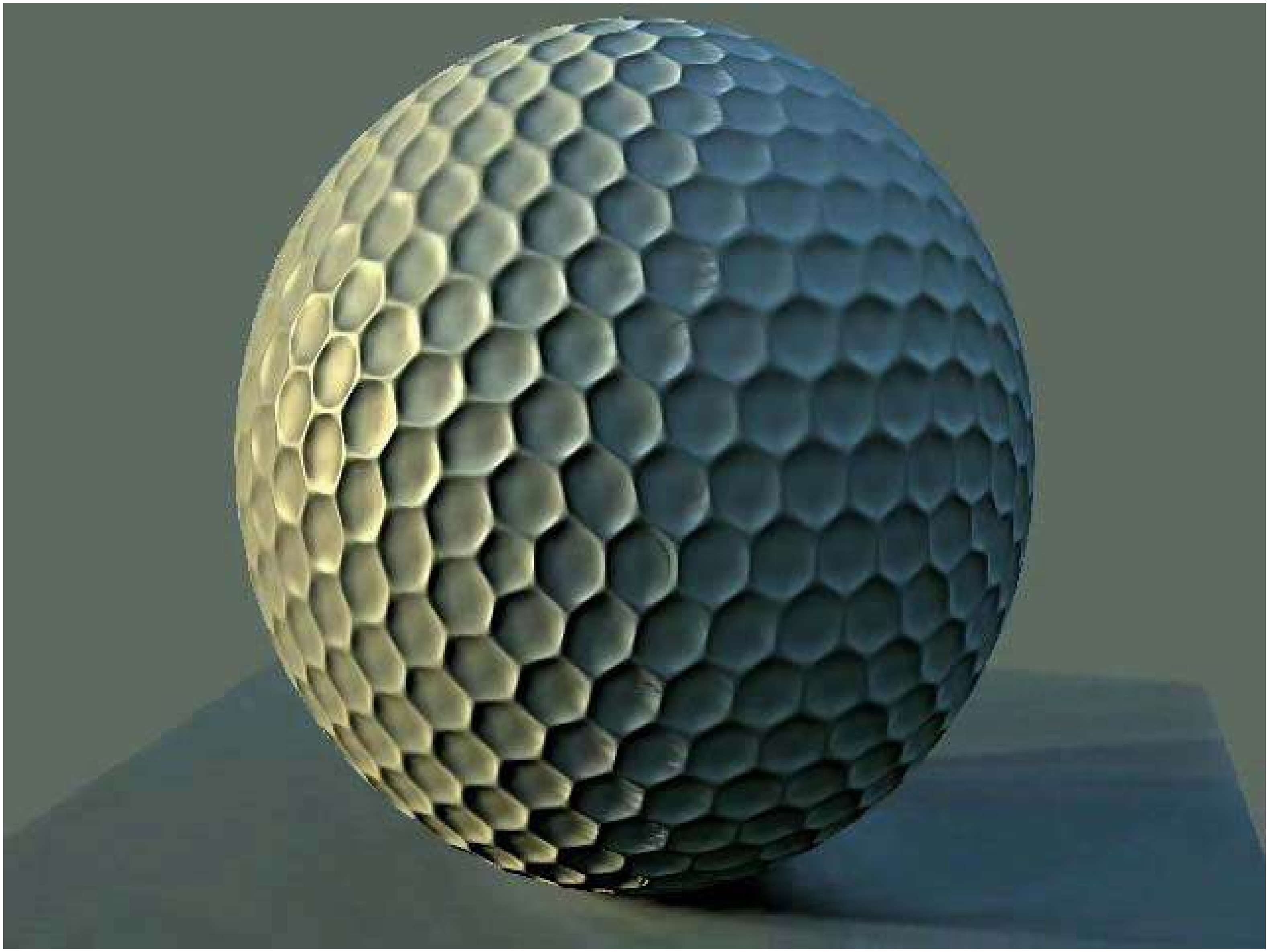
Exaggerated Shading

- Object enhancement
  - Illuminate each vertex at grazing angle
  - Improves geometry understanding
  - Highlights?
  - Shadows?
- Scene enhancement
  - Change everything
- Both have applications

Rusinkiewicz et al., SIGGRAPH'06



Specular Shading

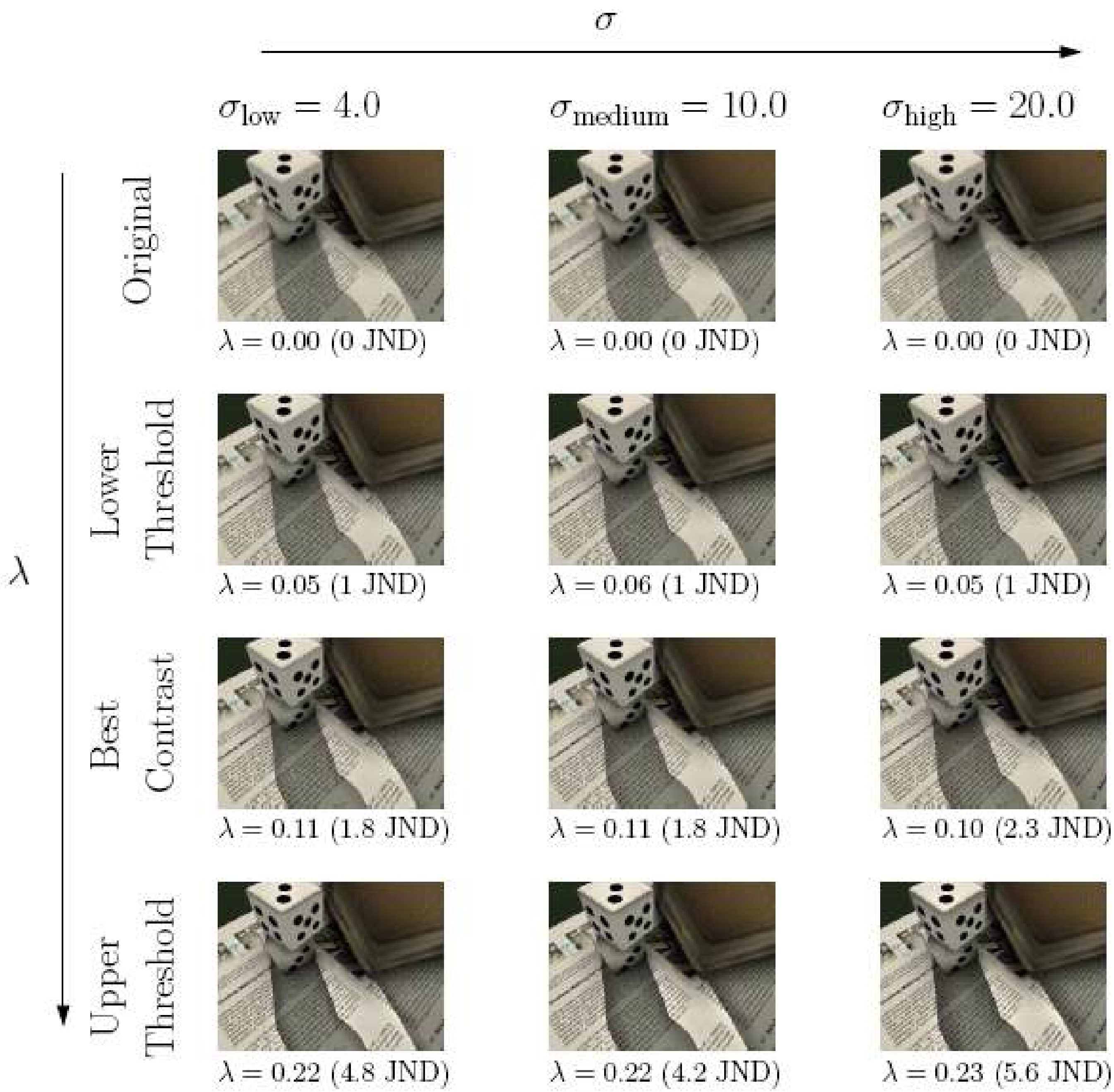


Study

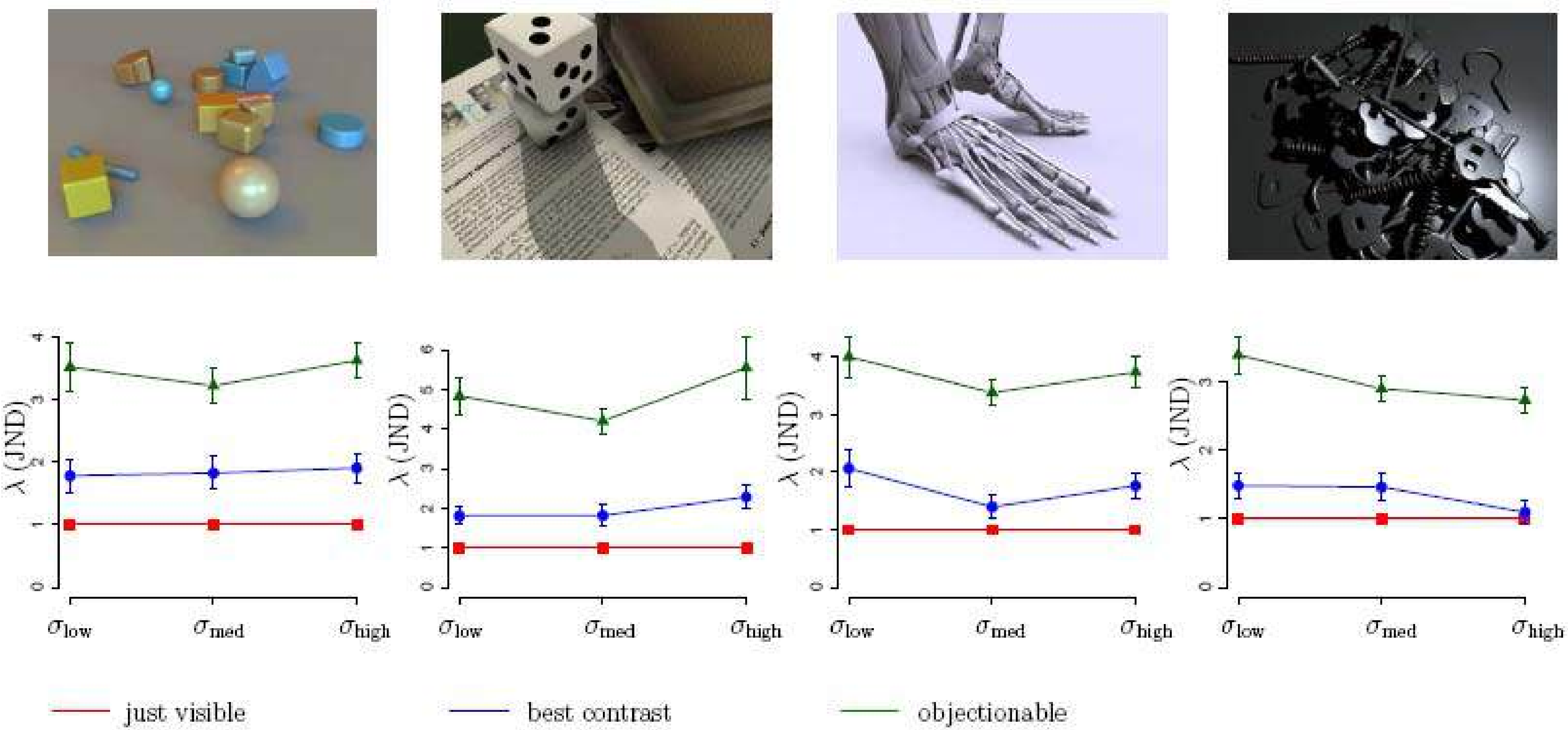
- Goals
  - Find suitable settings
  - See limitations
  - Rank preference
- Method of adjustments
  - Strength λ: adjustable
  - Fixed width $\boldsymbol{\sigma}$: low, medium, high
  - 4 scenes,15 participants
  - Task: Find such λ that:
    1. Added enhancement is ***just noticeable***
    2. Added enhancement becomes ***objectionable***
    3. Image appearance is ***preferred***

Ihrke et al. SPIE2009

Results


σ
$\sigma_{\text{low}} = 4.0$
$\sigma_{\text{medium}} = 10.0$
$\sigma_{\text{high}} = 20.0$
λ
Original
Lower Threshold
Best Contrast
Upper Threshold
$\lambda = 0.00$ (0 JND)
$\lambda = 0.00$ (0 JND)
$\lambda = 0.00$ (0 JND)
$\lambda = 0.05$ (1 JND)
$\lambda = 0.06$ (1 JND)
$\lambda = 0.05$ (1 JND)
$\lambda = 0.11$ (1.8 JND)
$\lambda = 0.11$ (1.8 JND)
$\lambda = 0.10$ (2.3 JND)
$\lambda = 0.22$ (4.8 JND)
$\lambda = 0.22$ (4.2 JND)
$\lambda = 0.23$ (5.6 JND)

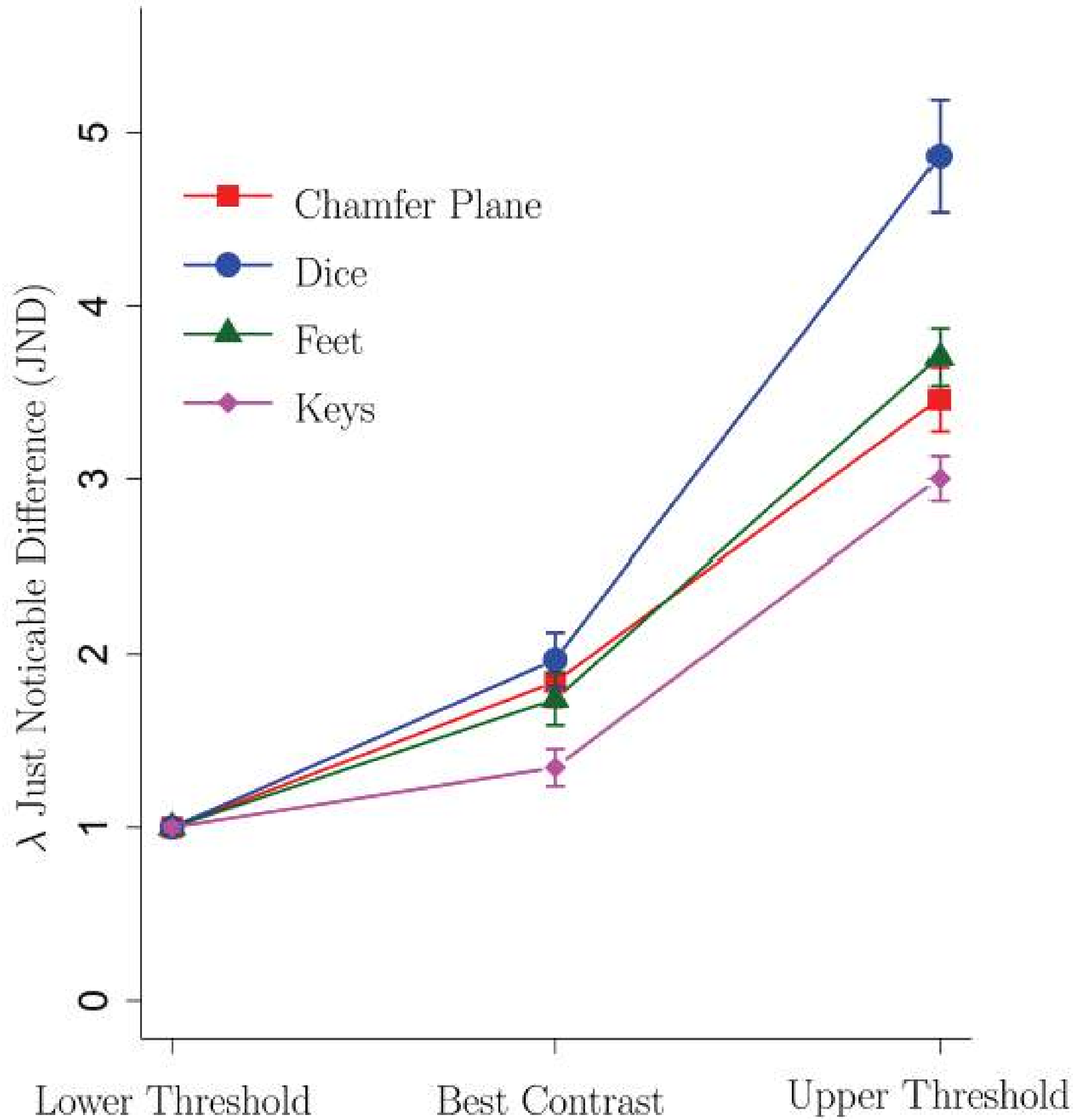
Results



Ihrke et al. SPIE2009

Results

- 2 JND
  – preferred
- 4 JND
  – objectionable



Ihrke et al. SPIE2009

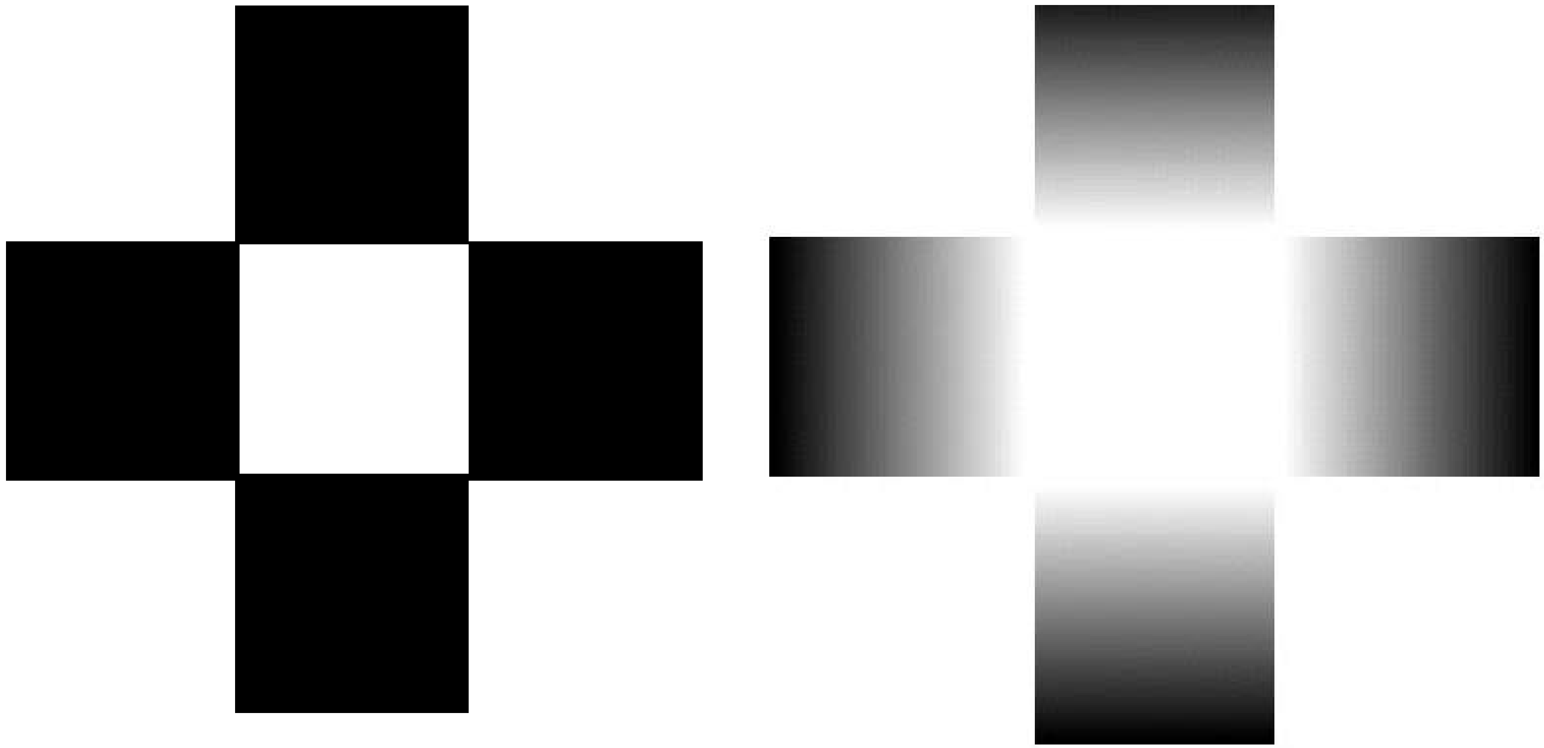
Summary

Better communicate image contents with a minimal change to image appearance

- Application of Cornsweet illusion to image enhancement
  - Generalization of unsharp masking
  - Automatic enhancement given the reference data:
    - HDR image
    - depth information
    - shading in 3D scene
  - Scene consistent 3D unsharp masking leads to even stronger effects

Glare Illusion

- Glowing effect [Zavagno and Caputo 2001]



SIGGRAPHASIA2011 HONG KONG

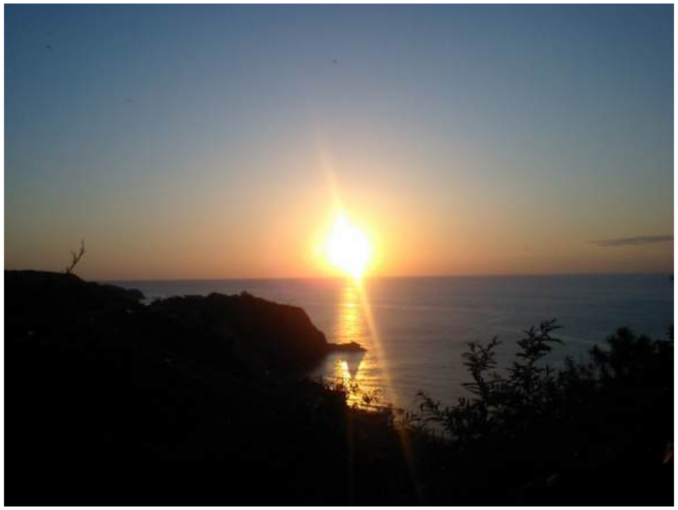
Glare Illusion

"Alan Wake" © Remedy Entertainment

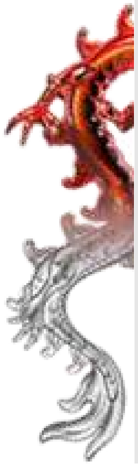
Glare Illusion in Different Media

Arts

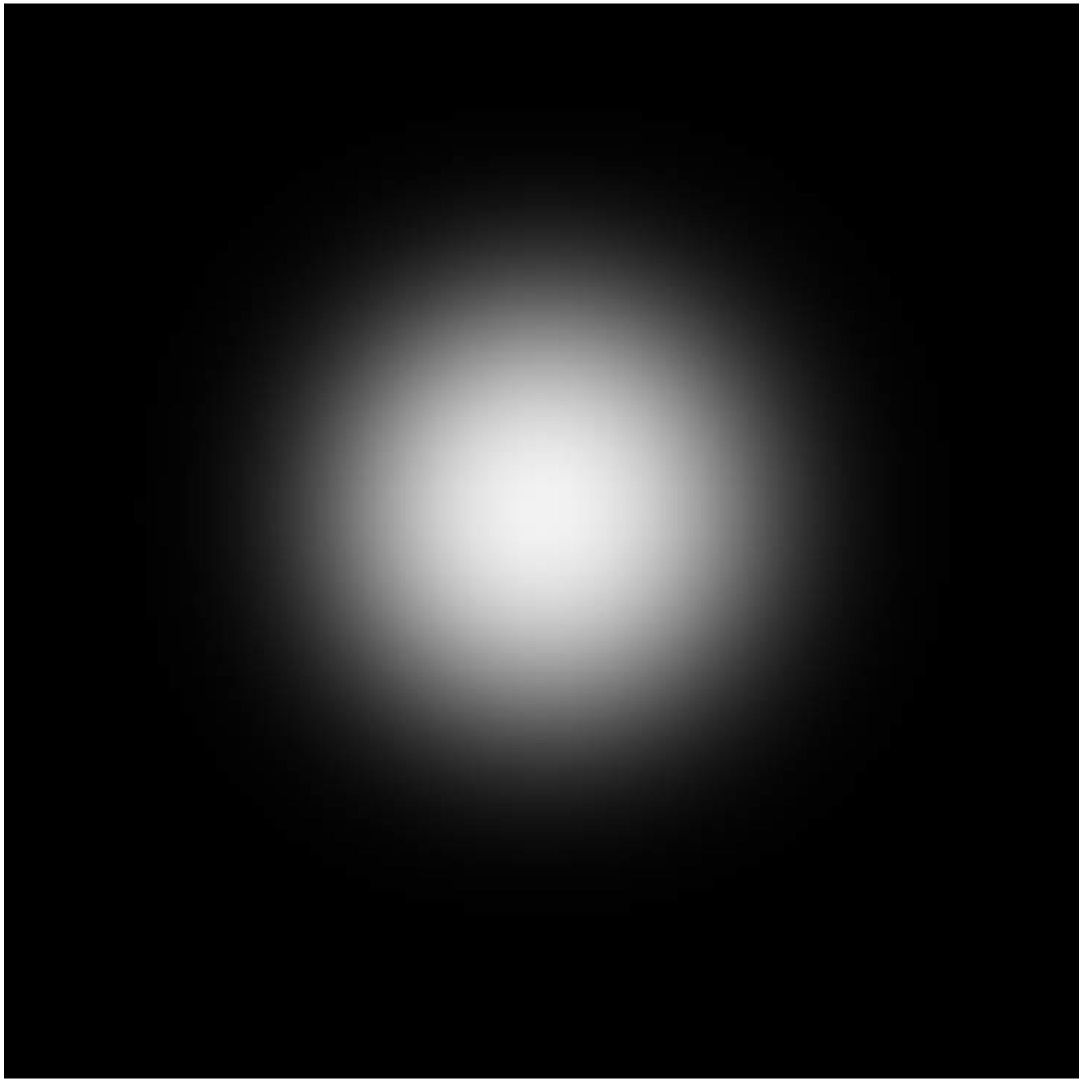
Computer Games



Photography



In Games

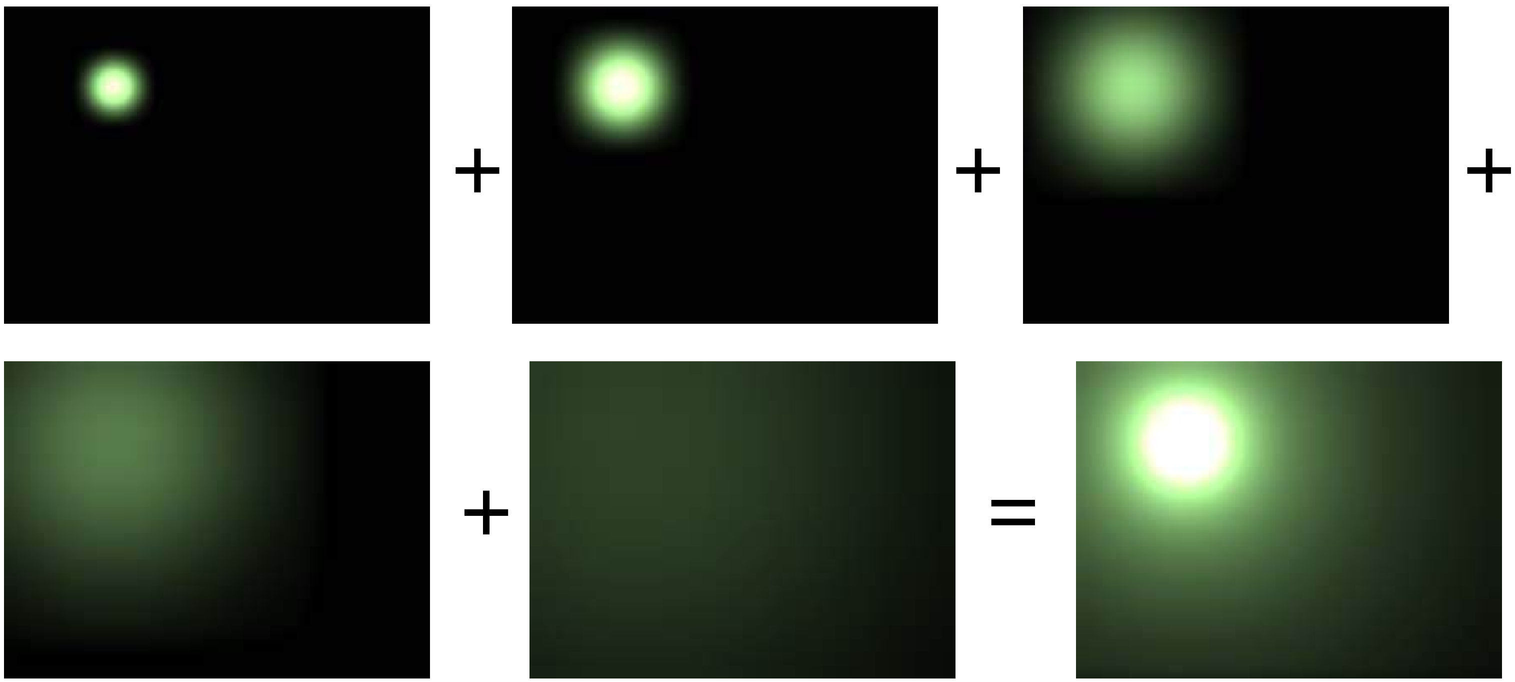


- Simple approximation: convolution with Gaussian
- Already does a good job in conveying brightness

Yoshida *et al*. (2008)

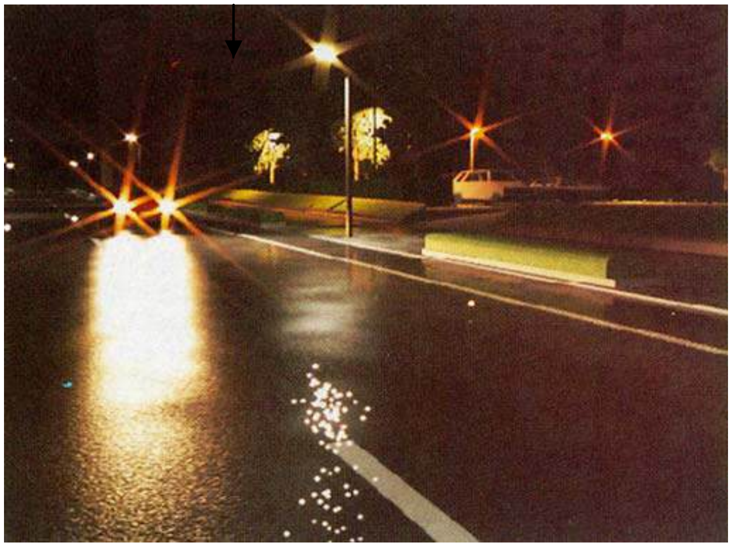
In Games

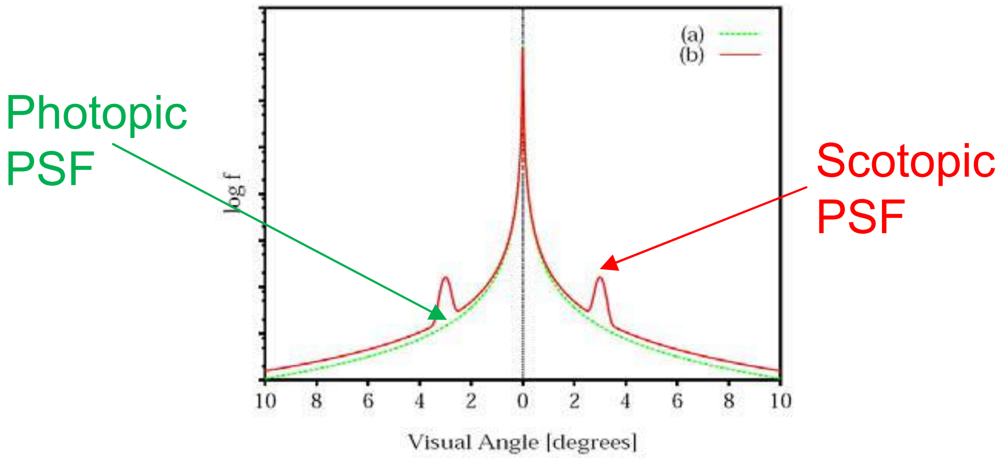
- Kawase, Practical Implementation of High Dynamic Range Rendering, Game Developer's Conference 2004



Glare in Realistic Rendering

- Optics-based models for rendering glare illusion
  - [Nakamae et al. 1990]
  - [Rokita 1993]
  - [Ward Larson et al. 1997]
  - [Kakimoto et al. 2004, 2005]
  - [Van den Berg et al. 2005]
  - [Spencer et al. 1995]





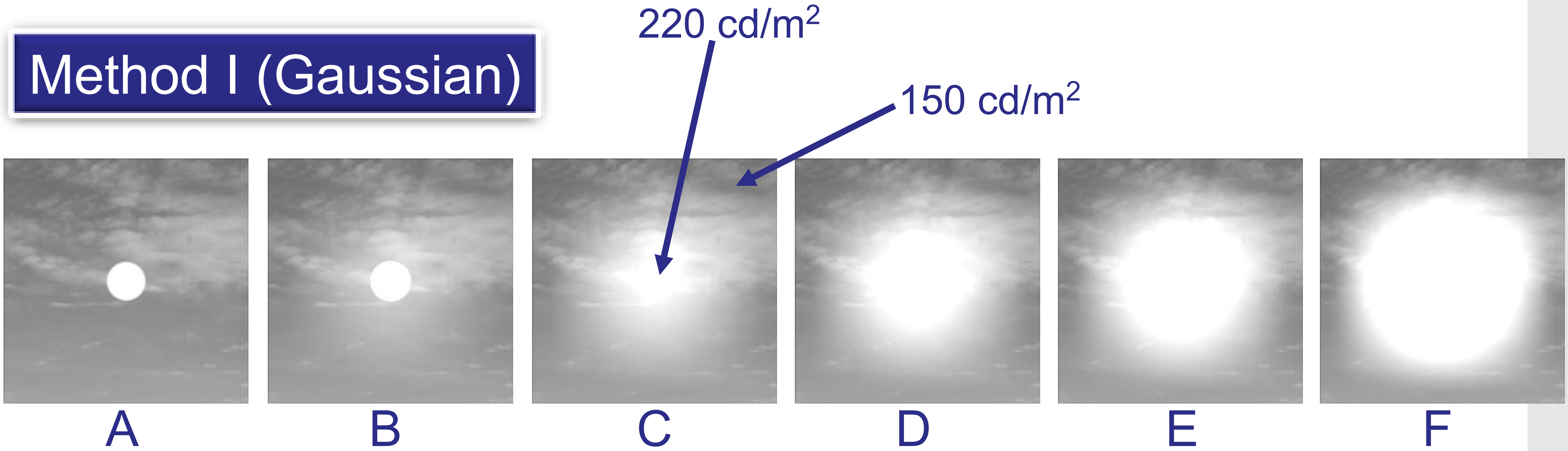
Psychophysical Experiment

- Goal
  - Measuring the brightness boosts caused by glare illusion
- 2 methods, 6 patterns for each
  - Gaussian: blurring kernel
    - Cheap approximation
  - Spencer et al.: human eye's PSF (disability glare)
    - Optical correctness
- 10 subjects
  - 20 minutes per person
- Barco Coronis Color 3MP Diagnostic Luminance Display (max. 430 cd/m^2)
- Dimly illuminated room (60 lux)

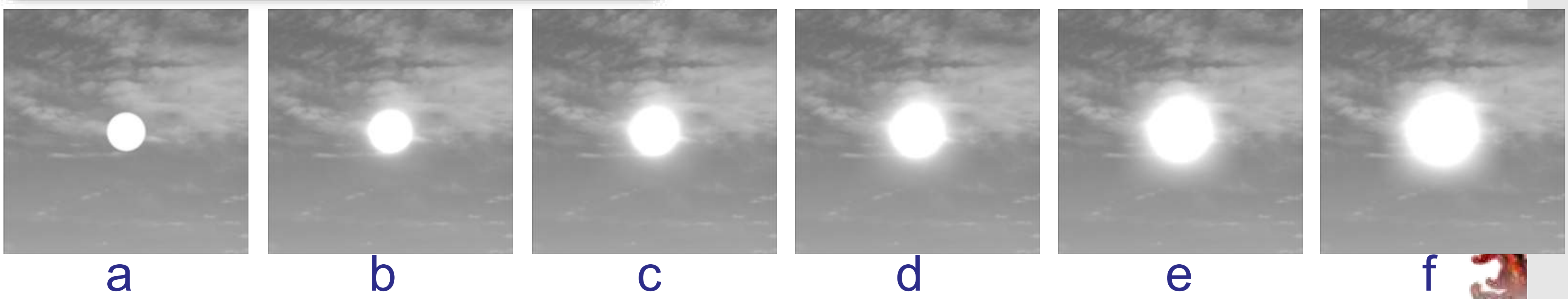
Yoshida et al. APGV2008

Stimuli

Method I (Gaussian)

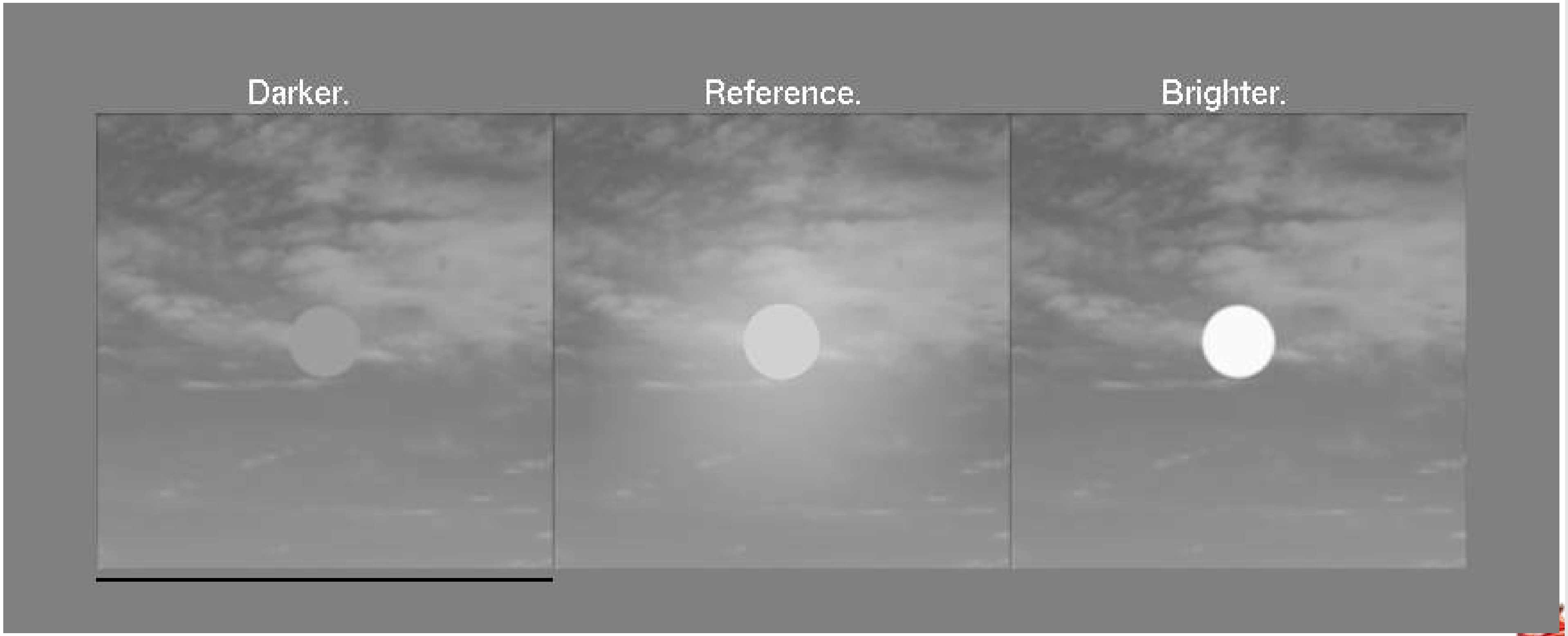


Method II (Spencer et al.)



Yoshida et al. APGV2008

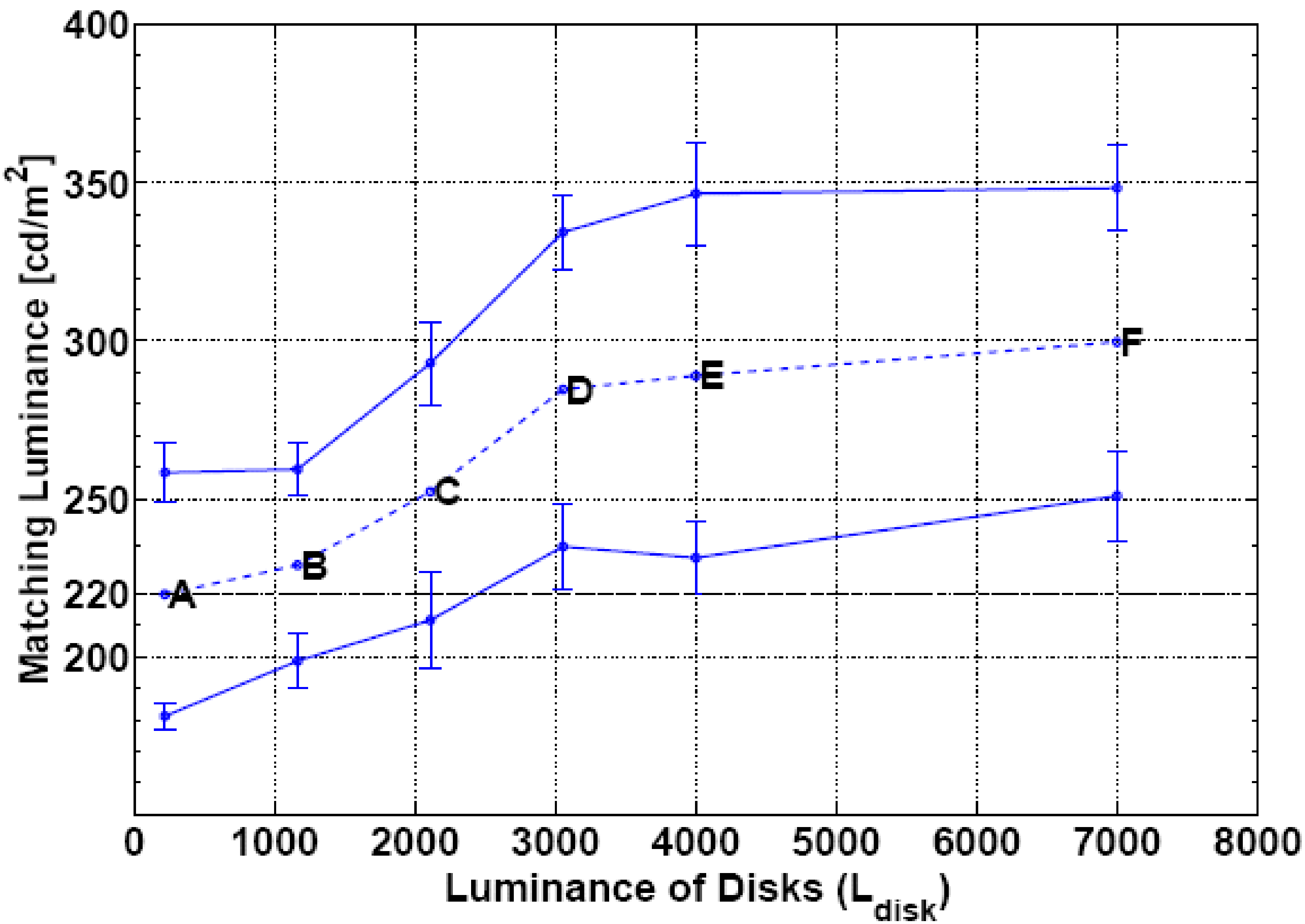
Perceptual Experiment

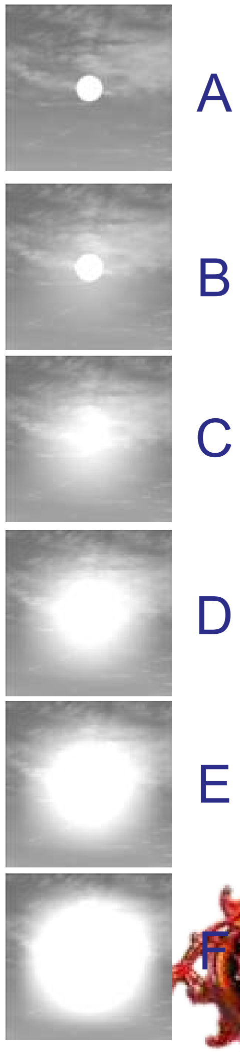


Task: Adjust the target disk luminance as close as possible to that of the reference, but slightly yet visibly darker/brighter.

Yoshida et al. APGV2008

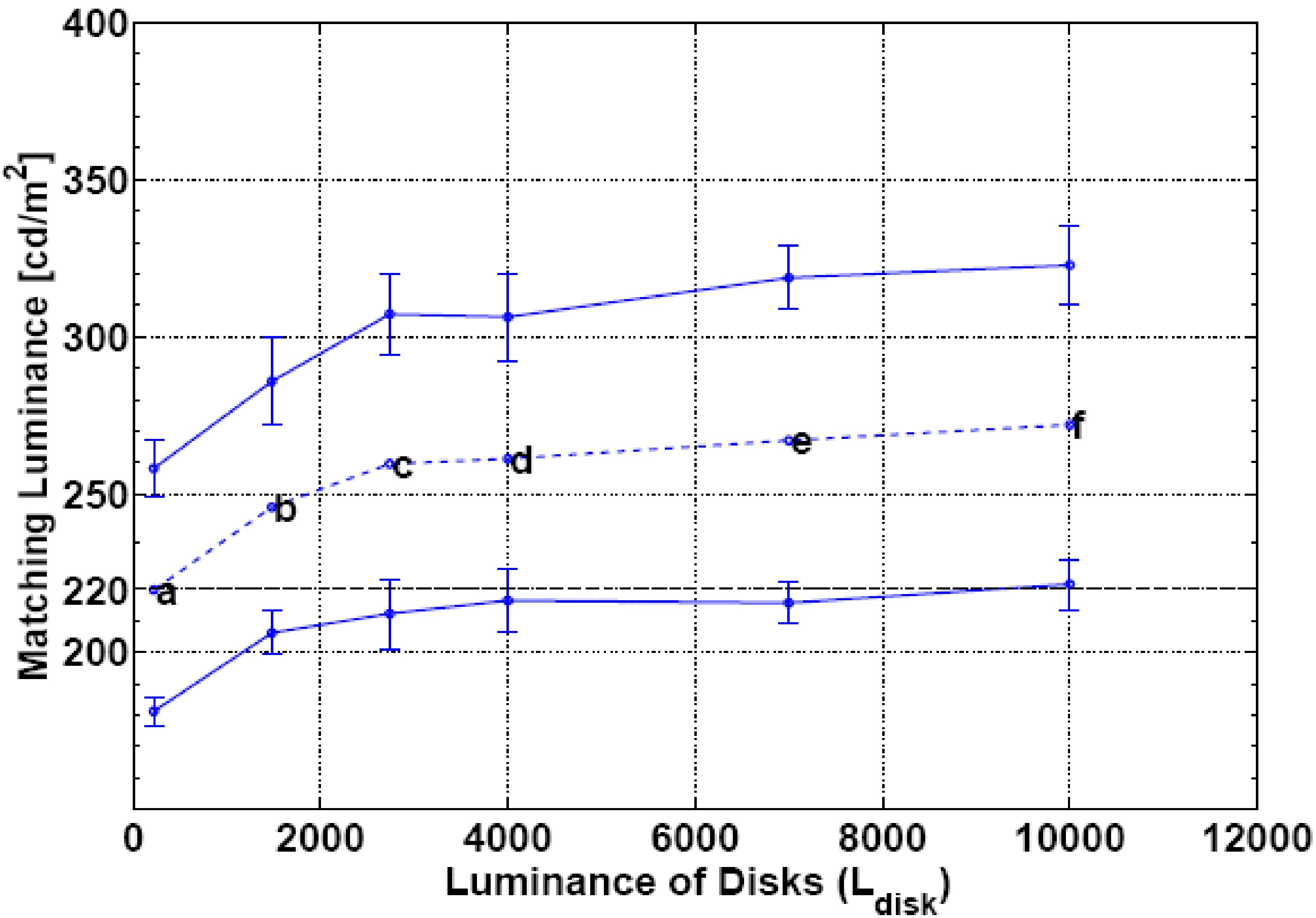
Method I (Gaussian)

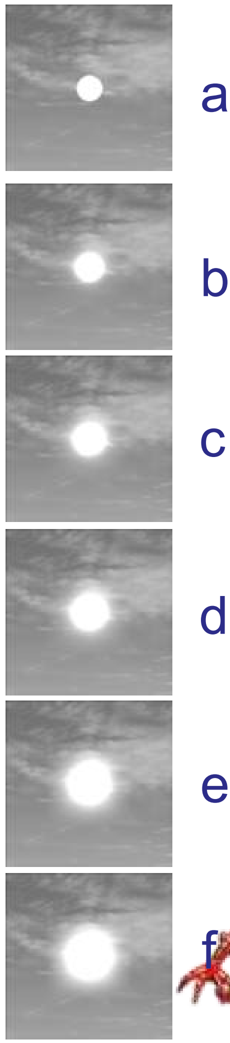




Yoshida et al. APGV2008

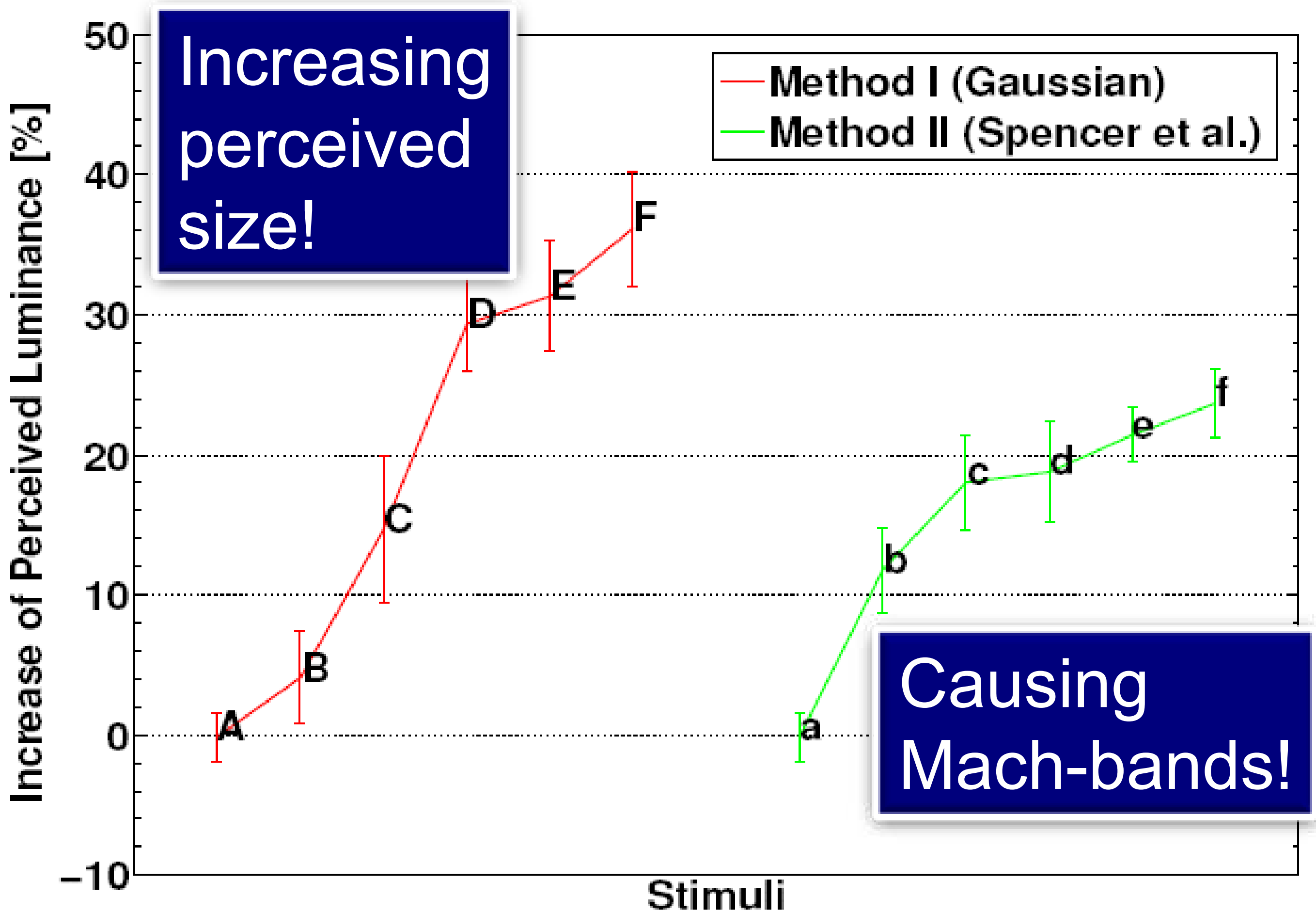
Method II (Spencer et al.)

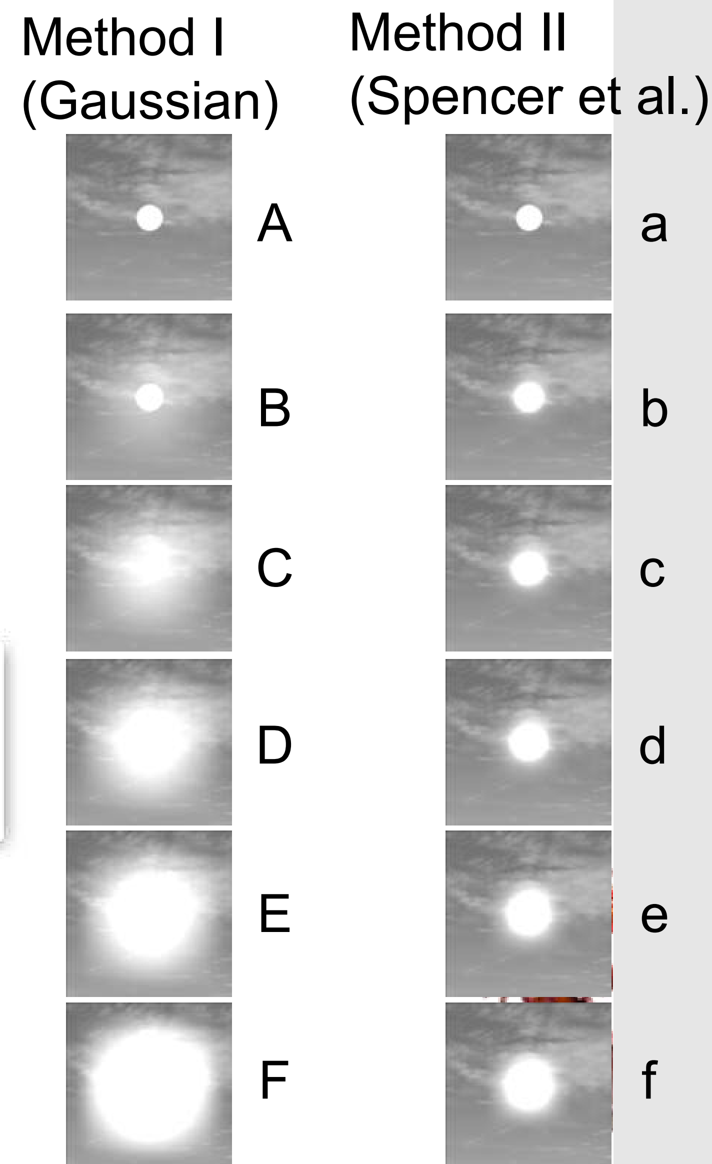




Yoshida et al. APGV2008

Trade-offs





- Measuring brightness boost of the glare illusion
  - Increasing the perceived luminance by 20 – 35 %
  - Gaussian blurring is equally effective
- Trade-offs for both Gaussian and human eye's PSF

Yoshida et al. APGV2008

Dynamic Glare

- **Realism**
  Colorful haloes around bright lights by camera or eyes
- **Temporal glare**
  Changes over time (in eyes)
- **Motivation**
  Model of dynamic human eye to simulate temporal glare
- **Study**
  Can temporal glare boost even further boost brightness?

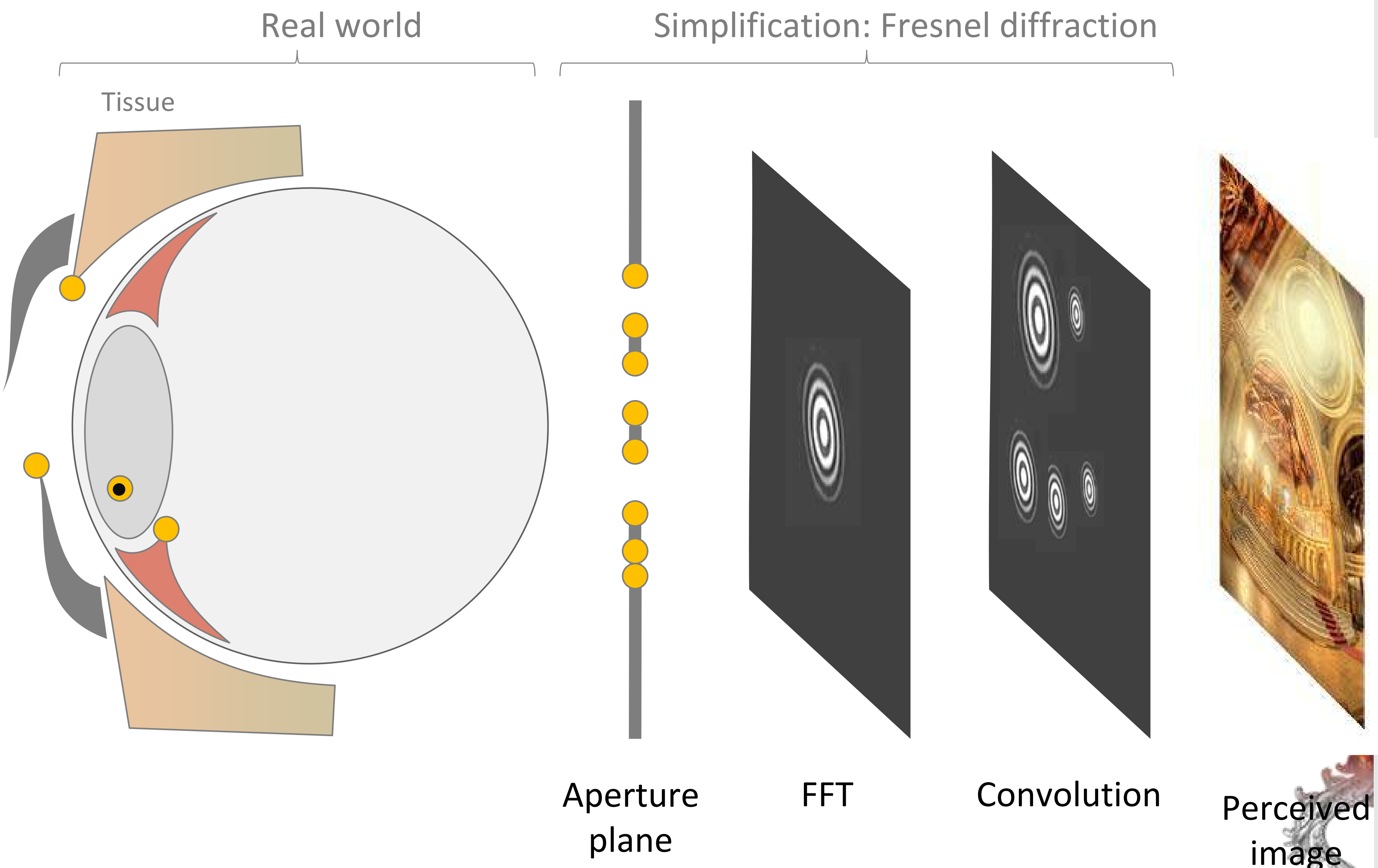
Ritschel et al. EG2008

Point spread function (*PSF*)

- **P**oint
  Spread
  Function
- Key to glare modeling
- Describes, how
  a **pixel** maps to
  a **pattern** under
  an **aperture**

Ritschel et al. EG2008

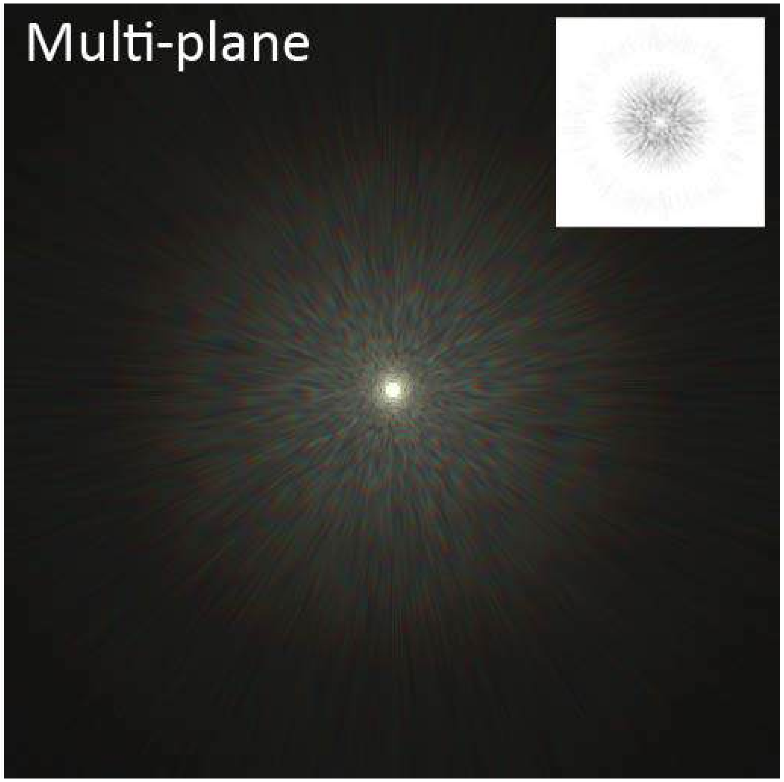
Our Simplified Model



Ritschel et al. EG2008

Diffraction: Single vs. Multi Aperture Planes

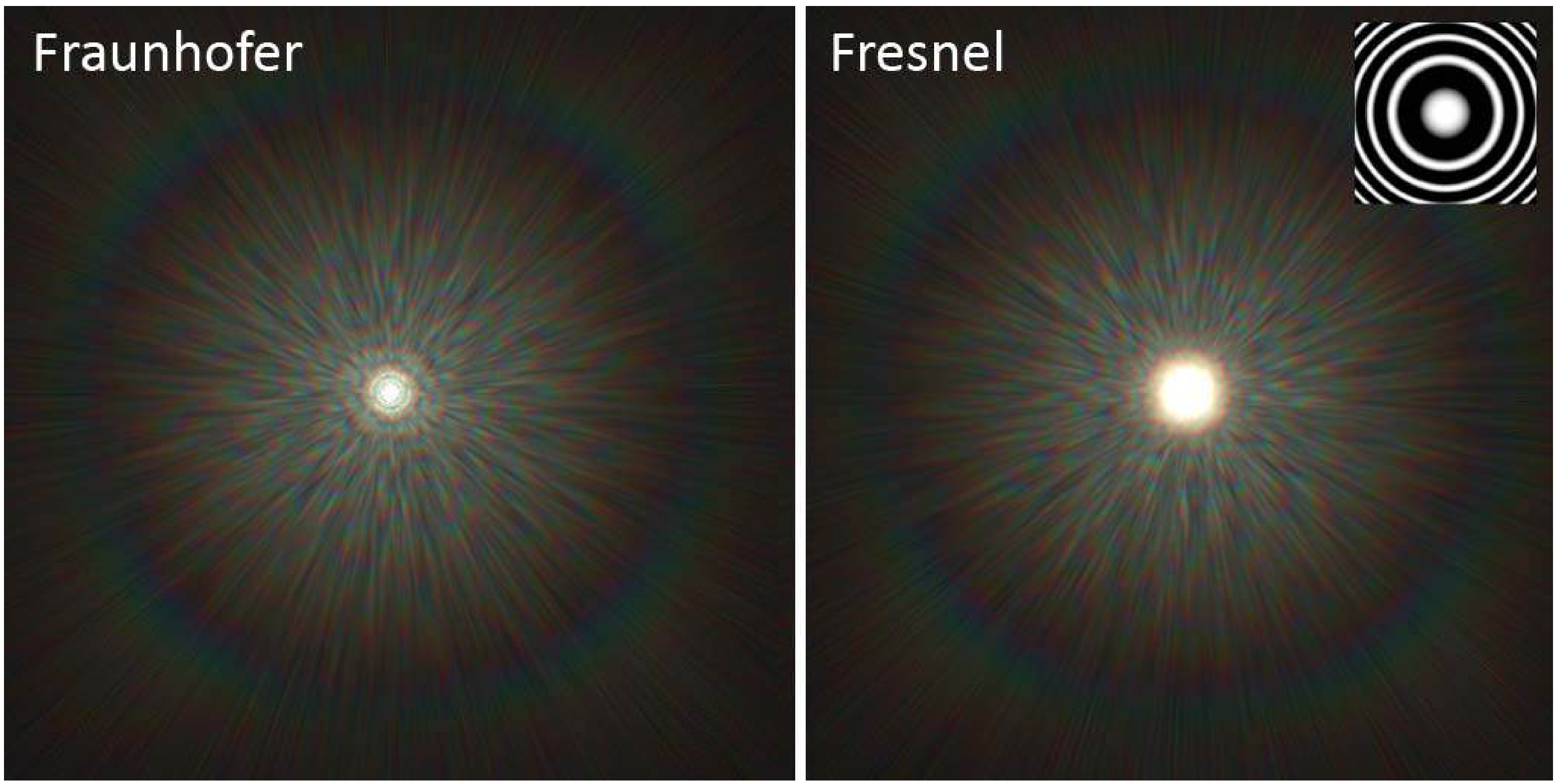
$$
\begin{aligned}
L_{\mathrm{i}}(x_{\mathrm{i}}, y_{\mathrm{i}}) &= K \left| \mathcal{F}\left\{ P(x_{\mathrm{p}}, y_{\mathrm{p}}) E(x_{\mathrm{p}}, y_{\mathrm{p}}) \right\}_{p=\frac{x_i}{\lambda d}, q=\frac{y_i}{\lambda d}} \right|^2 \\
K &= 1/(\lambda d)^2 \\
E(x_{\mathrm{p}}, y_{\mathrm{p}}) &= e^{i \frac{\pi}{\lambda d}(x_{\mathrm{p}}^2 + y_{\mathrm{p}}^2)}
\end{aligned}
$$



Ritschel et al. EG2008

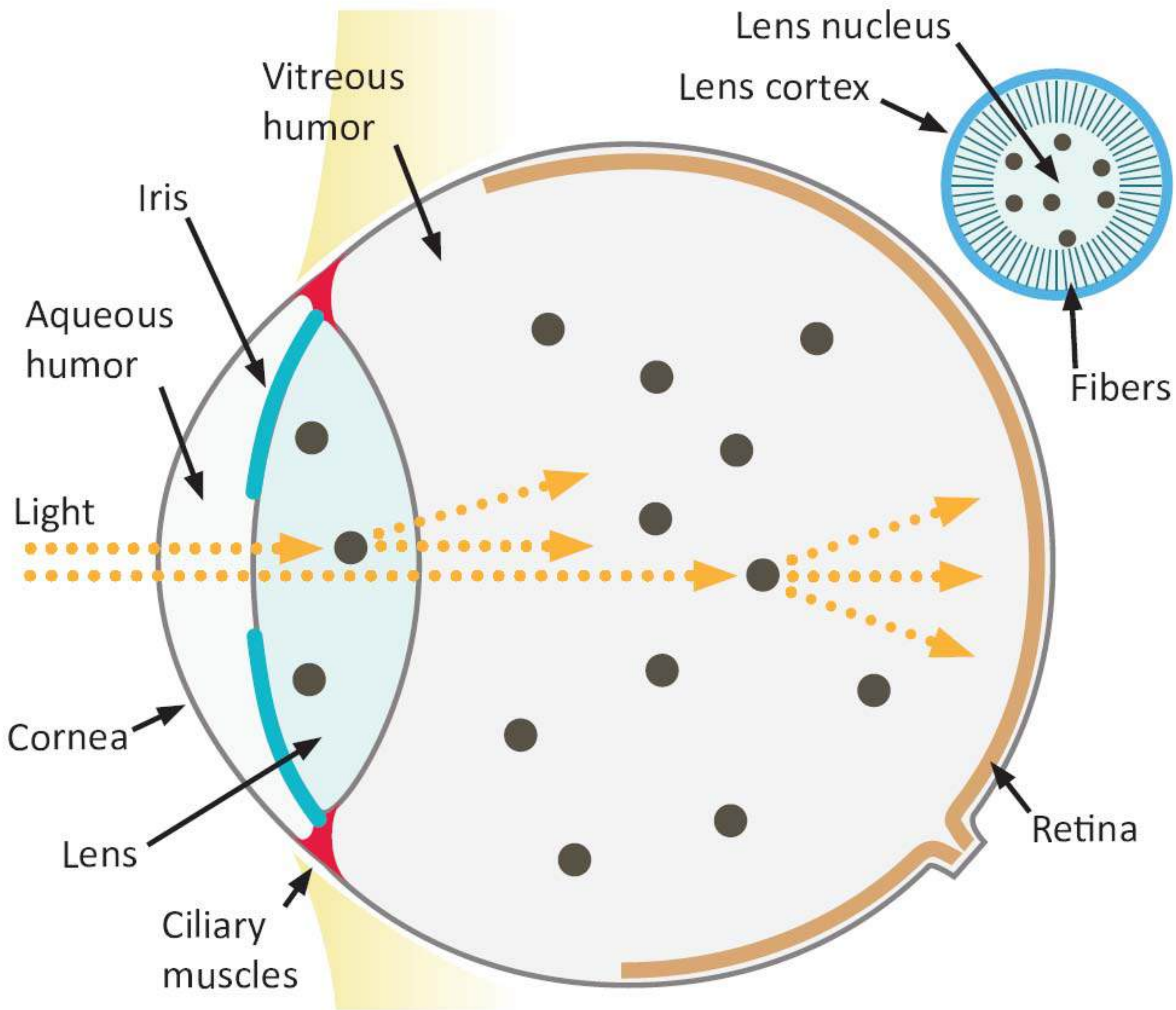
Diffraction: Fraunhofer vs. Fresnel

$$
\begin{aligned}
L_{\mathrm{i}}(x_{\mathrm{i}}, y_{\mathrm{i}}) &= K \left| \mathcal{F}\{P(x_{\mathrm{p}}, y_{\mathrm{p}}) E(x_{\mathrm{p}}, y_{\mathrm{p}})\}_{p=\frac{x_i}{\lambda d}, q=\frac{y_i}{\lambda d}} \right|^2 \\
K &= 1/(\lambda d)^2 \\
E(x_{\mathrm{p}}, y_{\mathrm{p}}) &= e^{i \frac{\pi}{\lambda d}(x_{\mathrm{p}}^2 + y_{\mathrm{p}}^2)}
\end{aligned}
$$



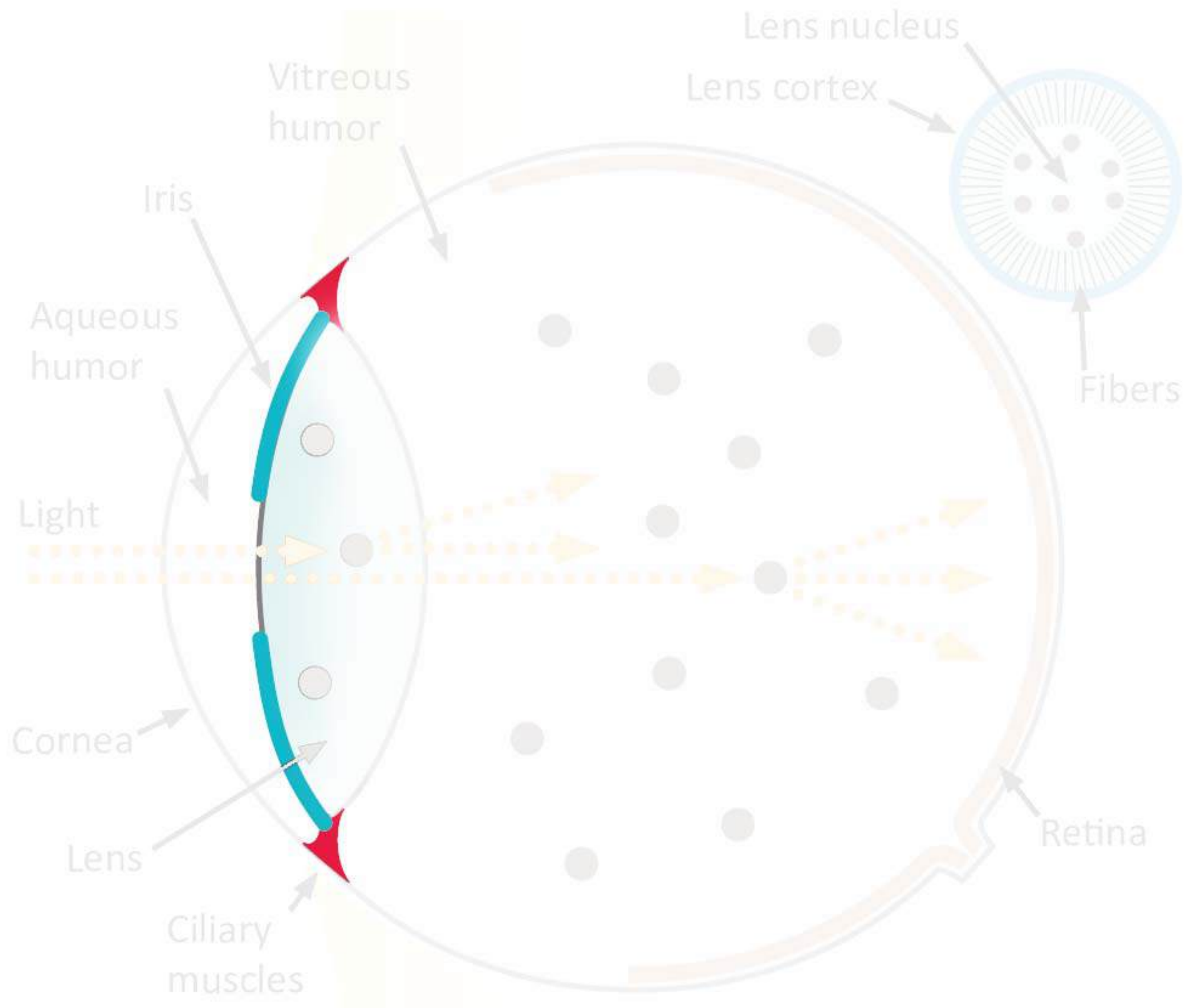
Ritschel et al. EG2008

Temporal Glare Pipeline



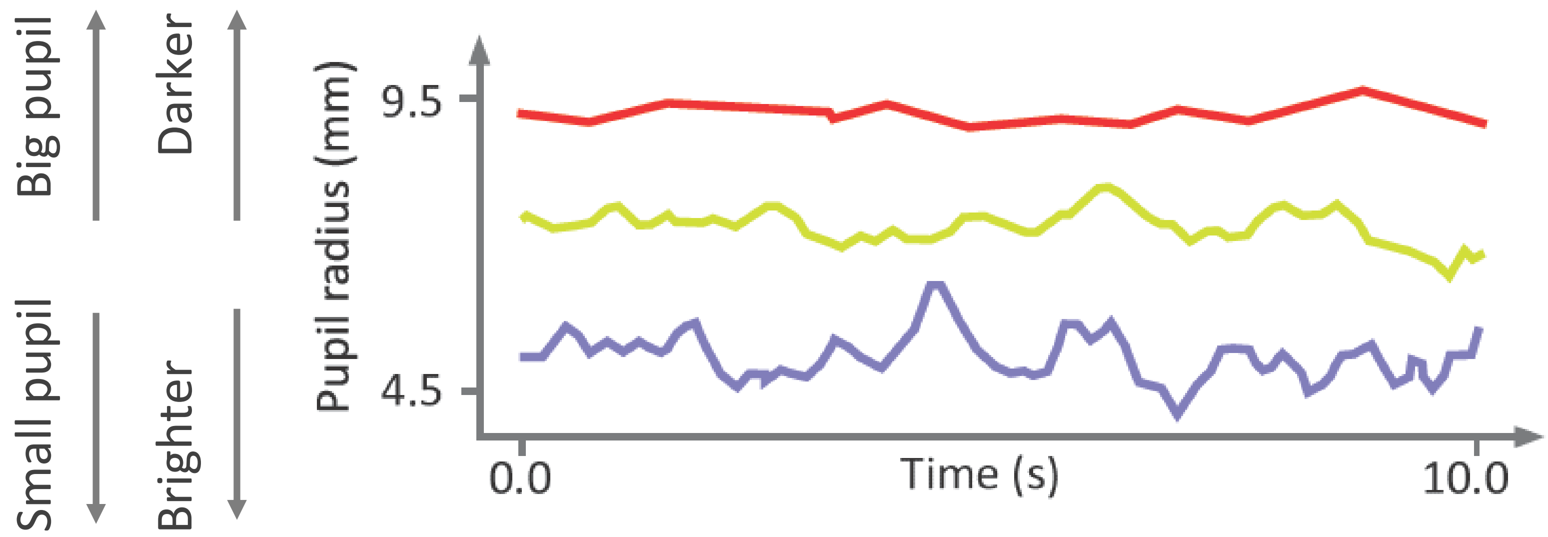
Ritschel et al. EG2008

Aperture: Pupil

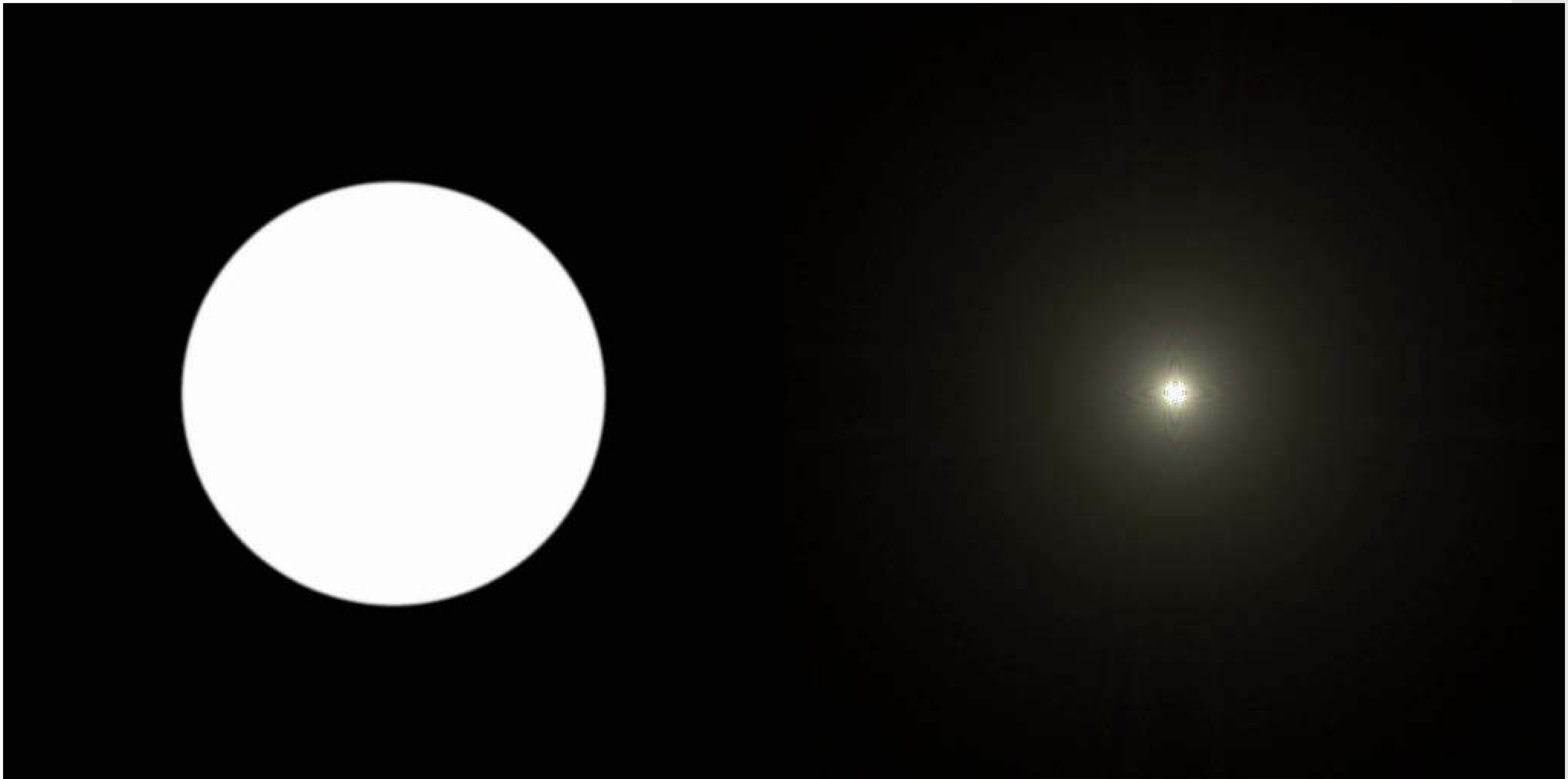


Aperture: Pupil

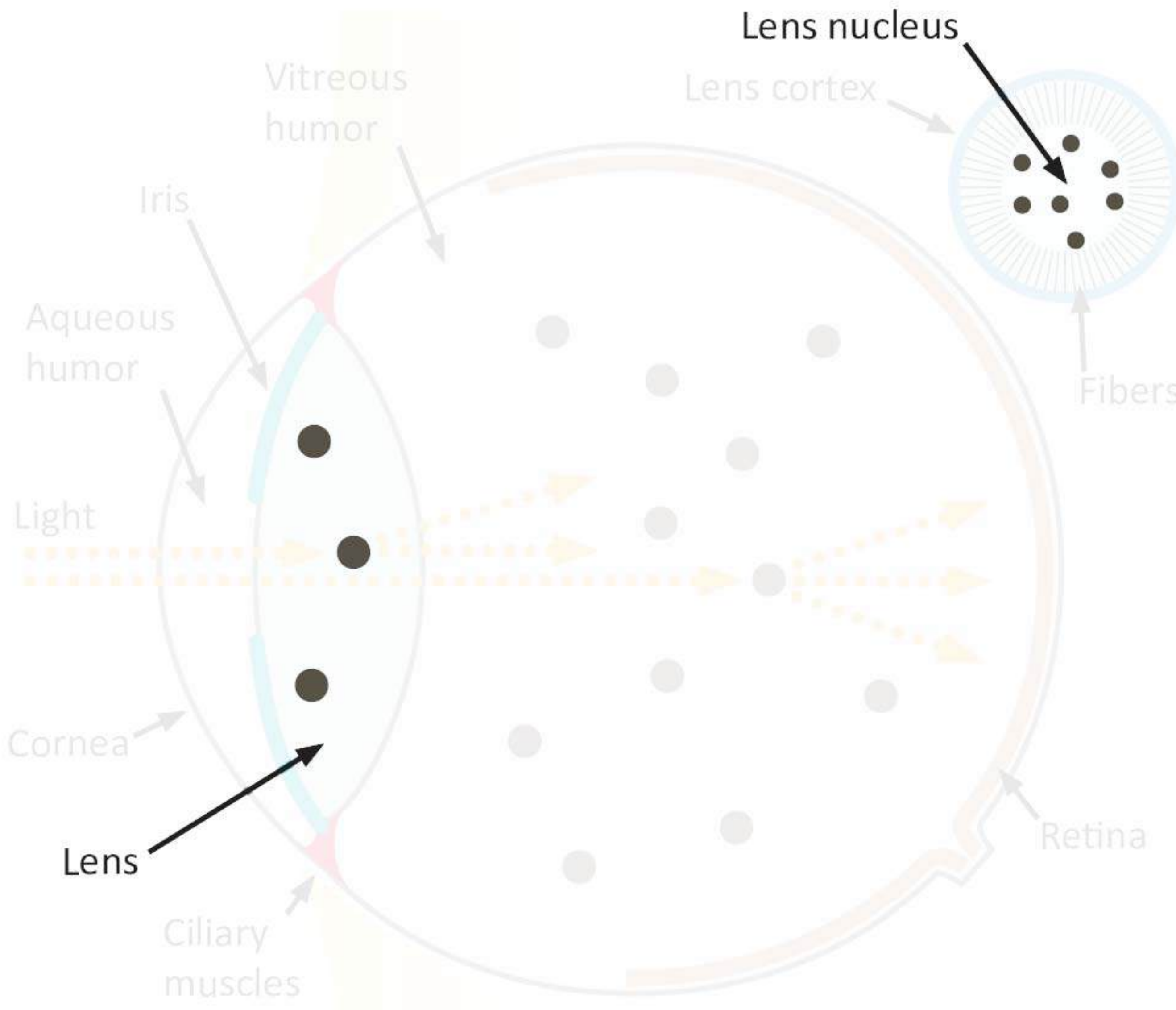
- **Adaptation**
- Can convert HDR image into pupil size
- Pupillary hippus:
  Strong contrast between glare source and background
- Stronger for smaller pupils, i.e. bright conditions



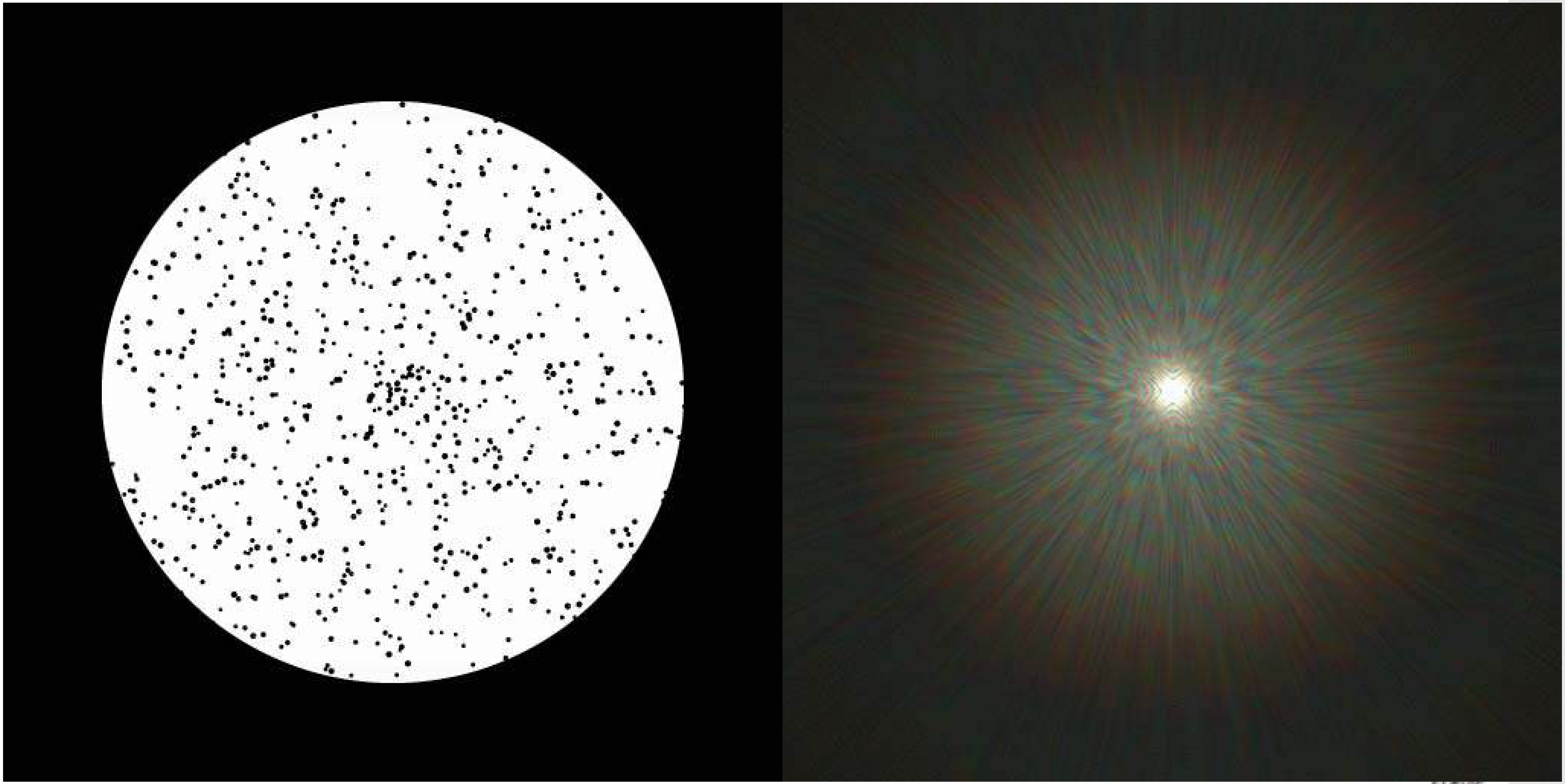
Aperture: Pupil



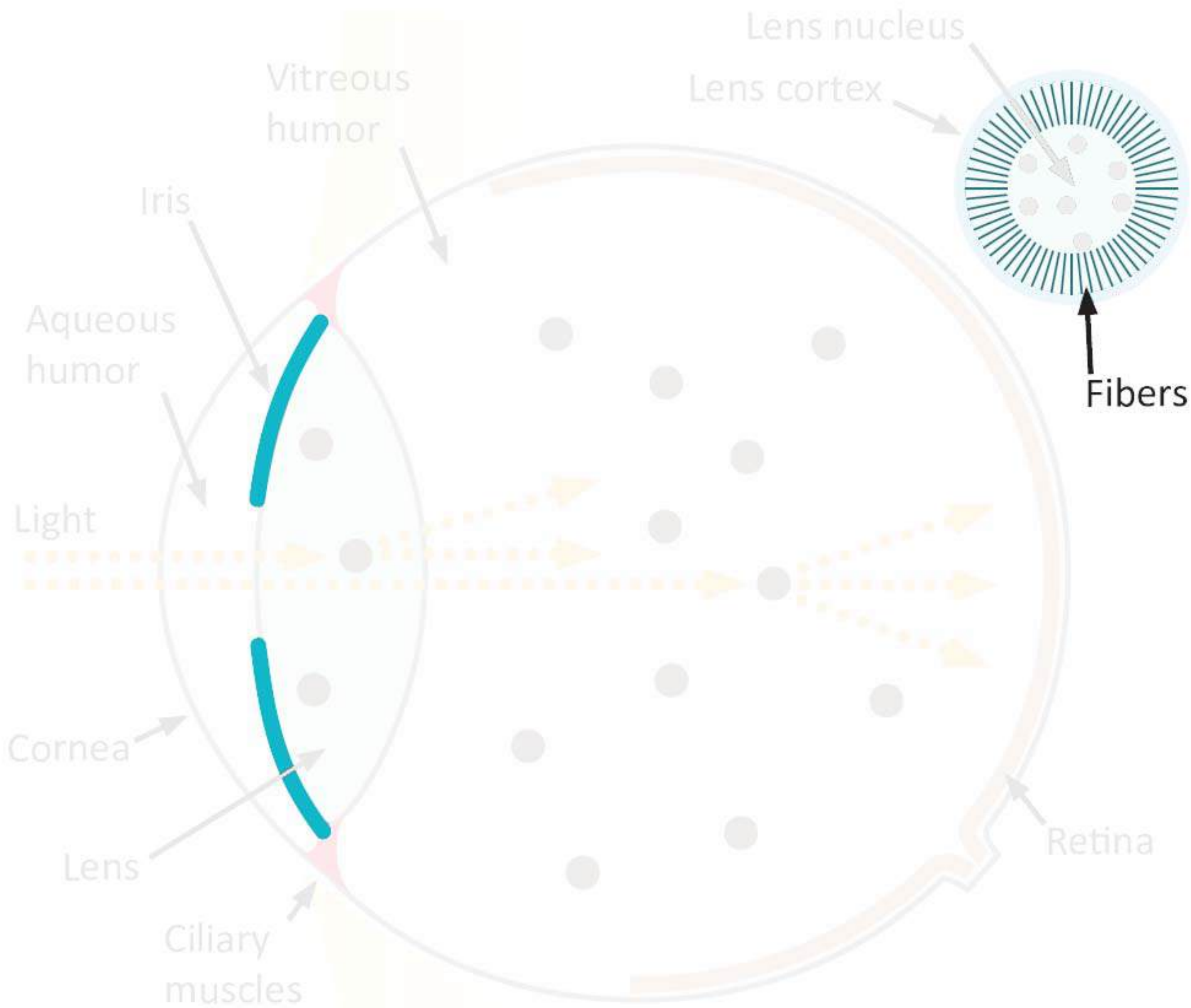
Aperture: Lens



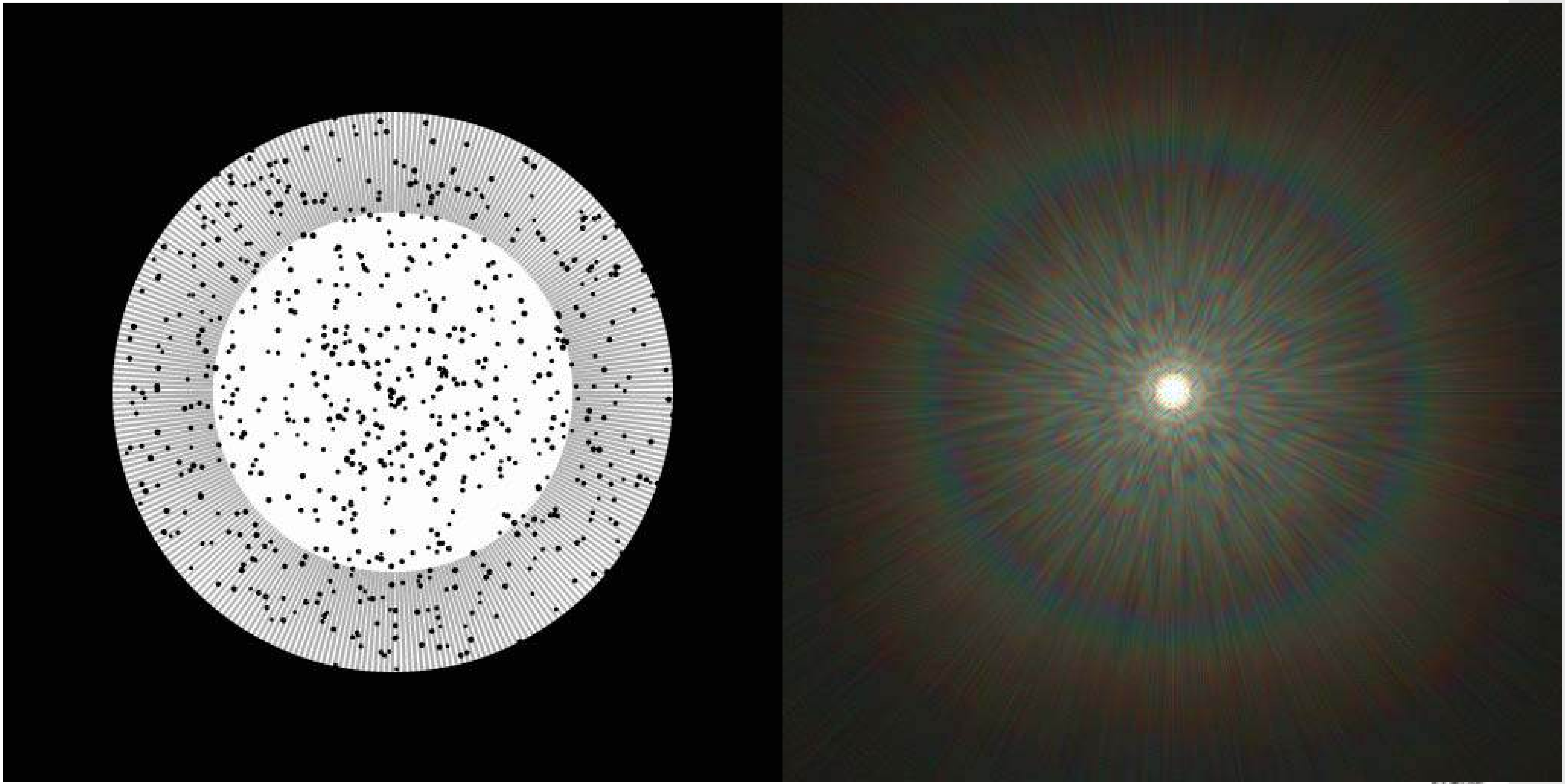
Aperture: Lens



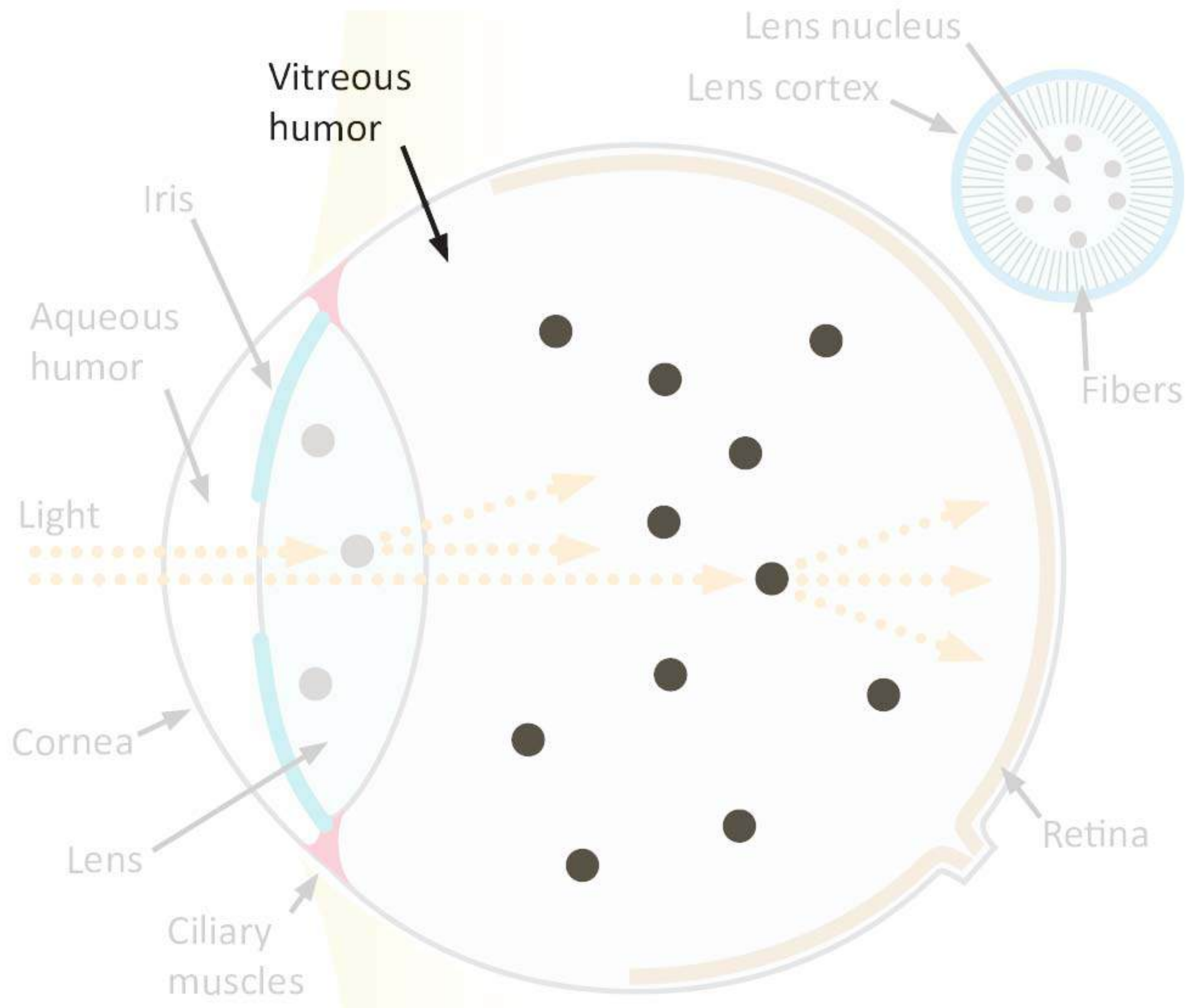
Aperture: Gratings / Lens fibers



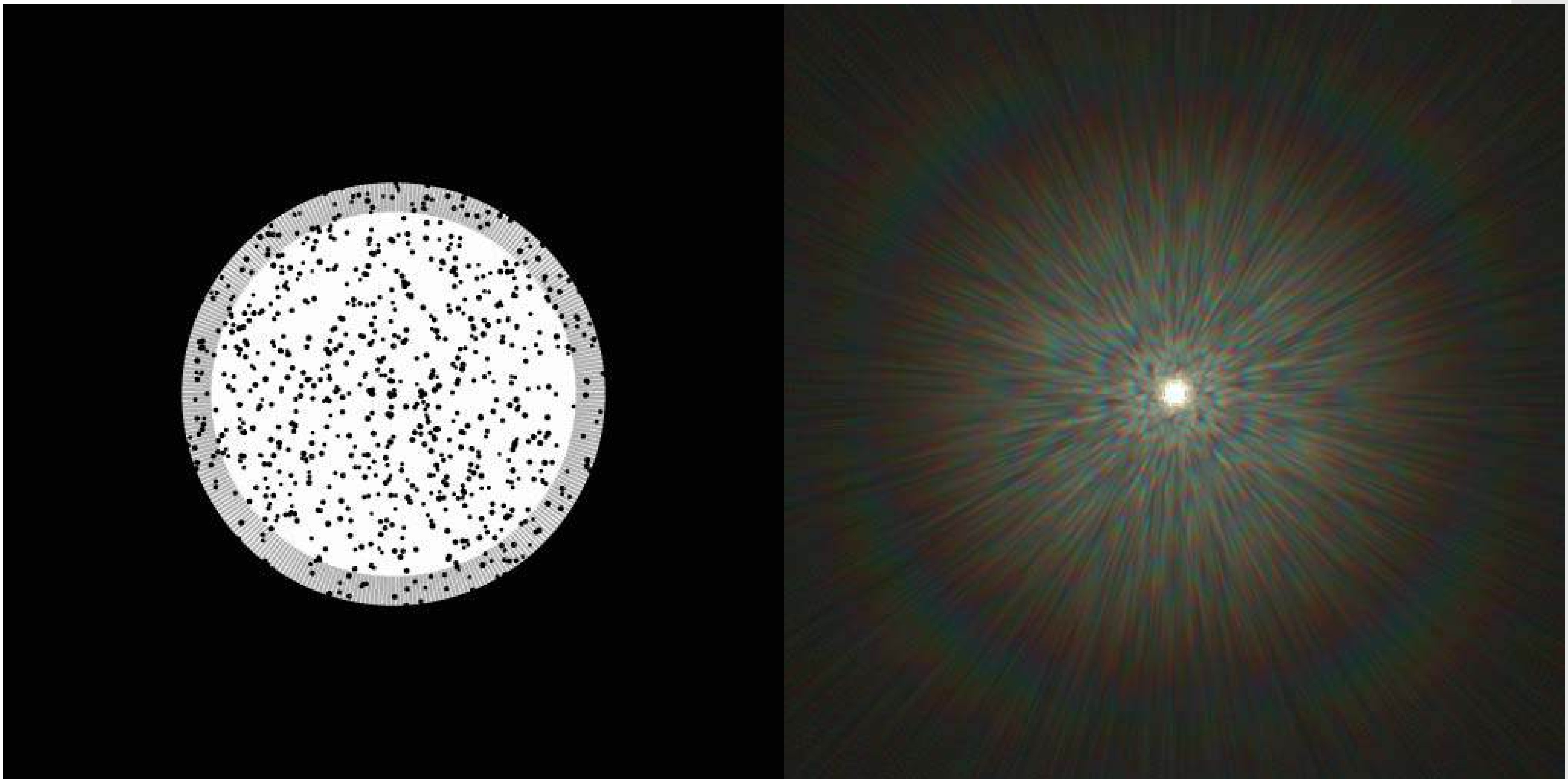
Aperture: Gratings / Lens fibers



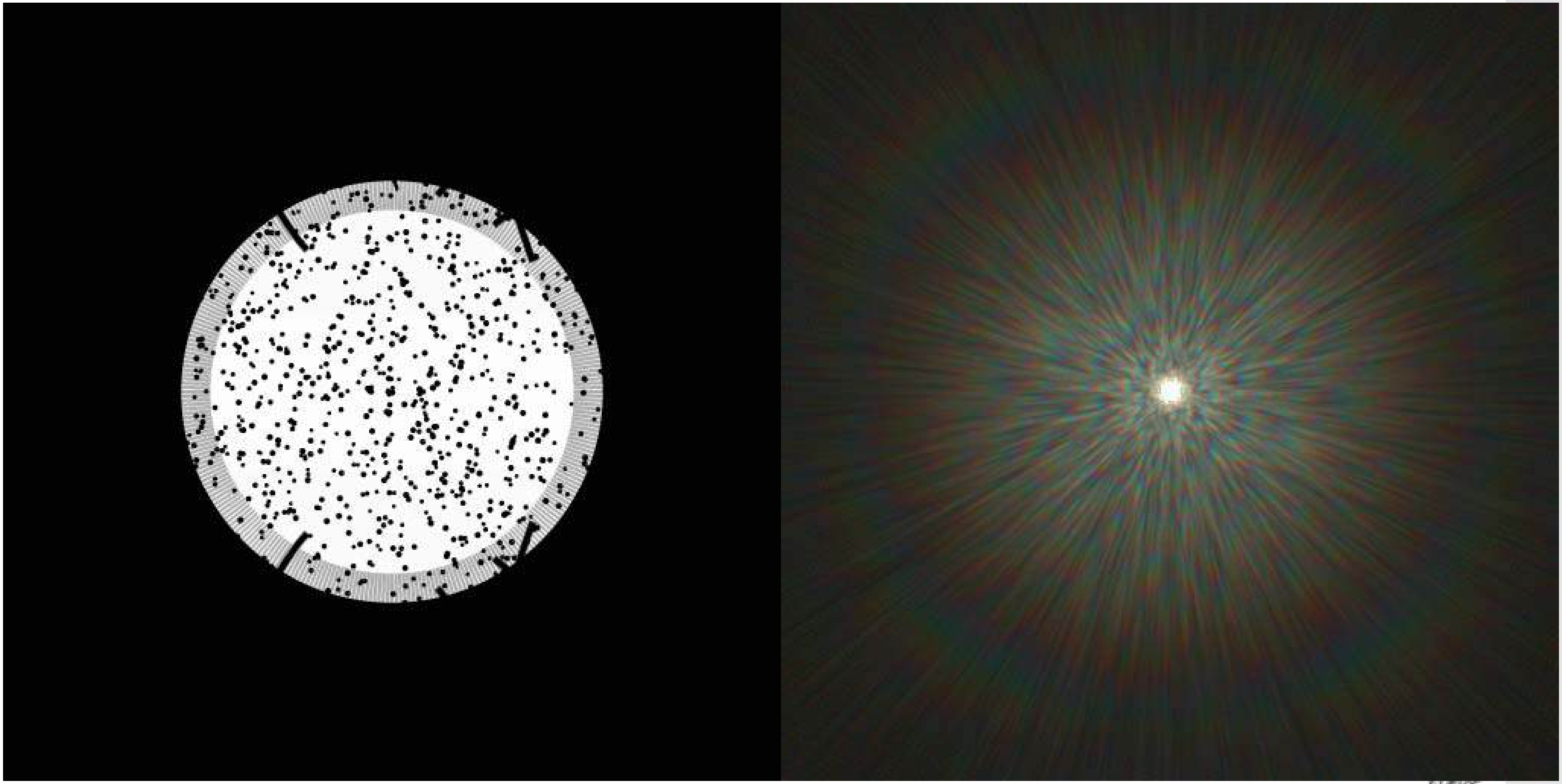
Aperture: Vitreous Humor



Aperture: Vitreous Humor

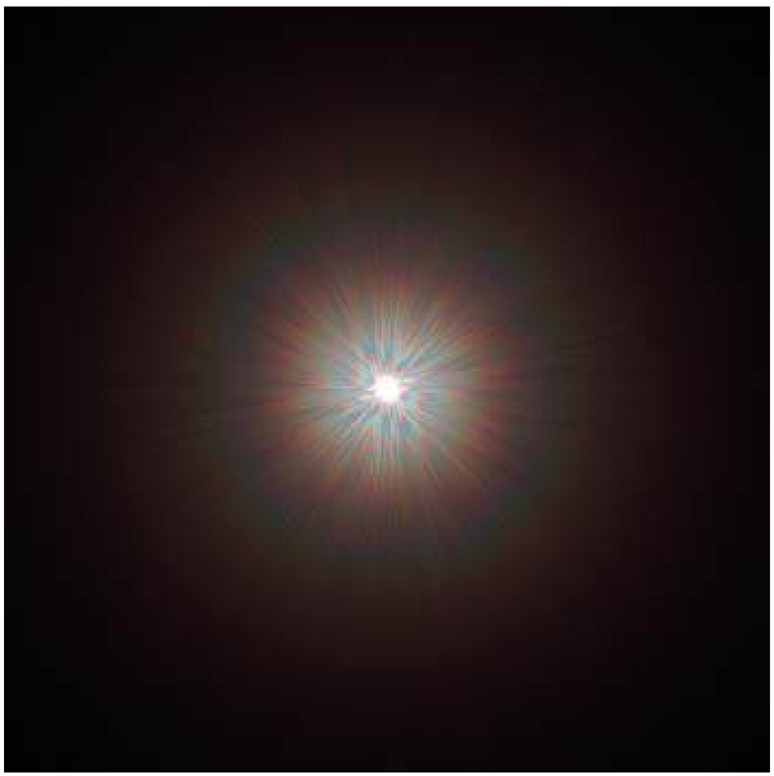


Aperture: Eyelashes (*optional*)

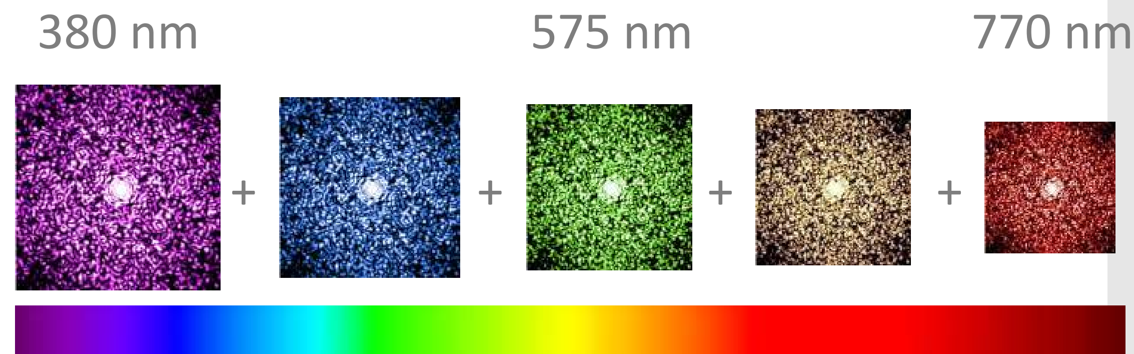


Chromatic Blur

- Compute one wavelength - Get others for free!
- They are scaled copy of base wavelength, i.e. 575 nm (approximation)



=



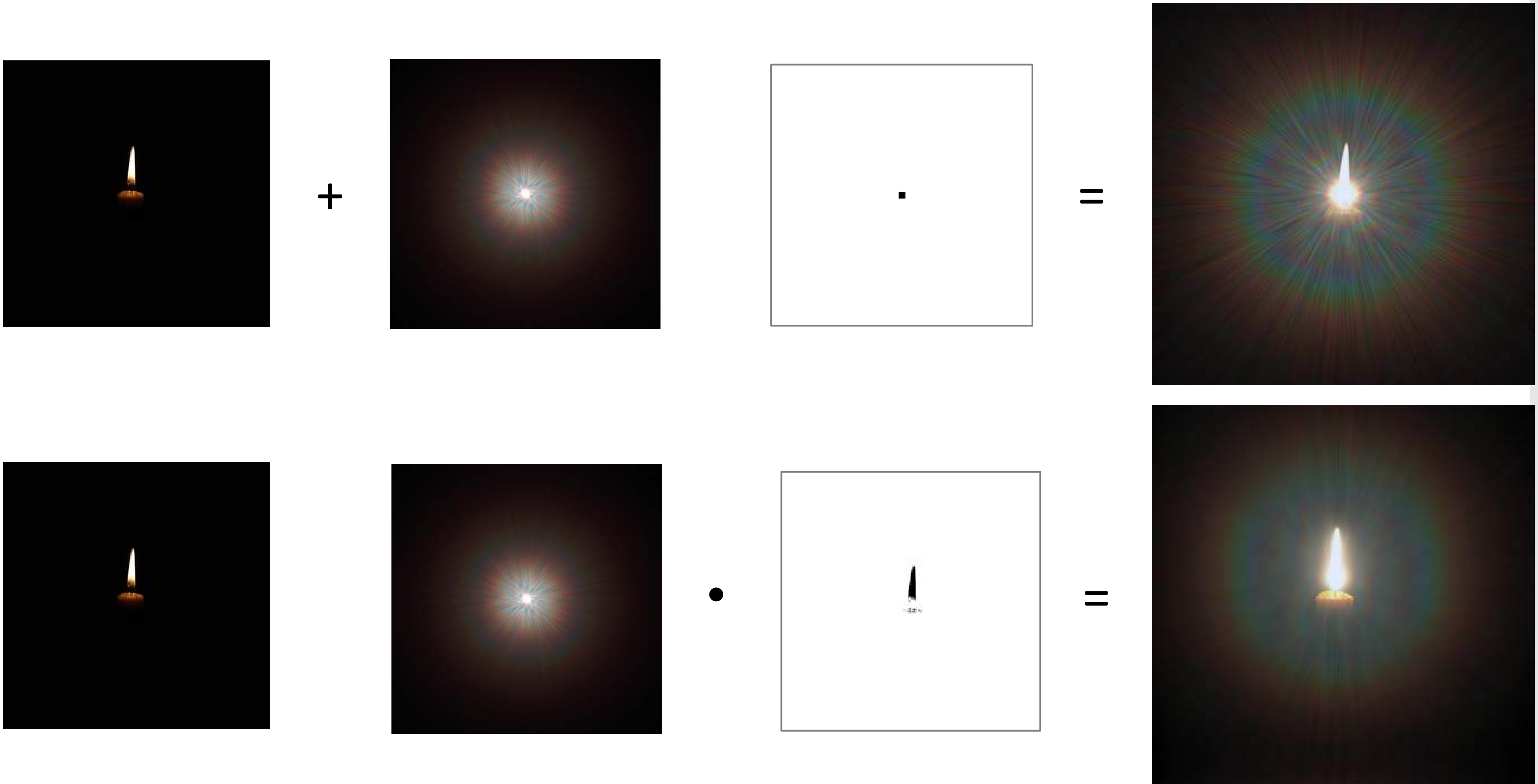
$$F_s(\mathbf{x}) = \sum_{i=0}^{n-1} s(\lambda_i) F_{575\,\mathrm{nm}}(\mathbf{x}_i)$$

$$\lambda_i = 380\,\mathrm{nm} + i\,\frac{770\,\mathrm{nm} - 380\,\mathrm{nm}}{n}$$

$$\mathbf{x}_i = \mathbf{x}\,\frac{575\,\mathrm{nm}}{\lambda_i}\ .$$

Convolution

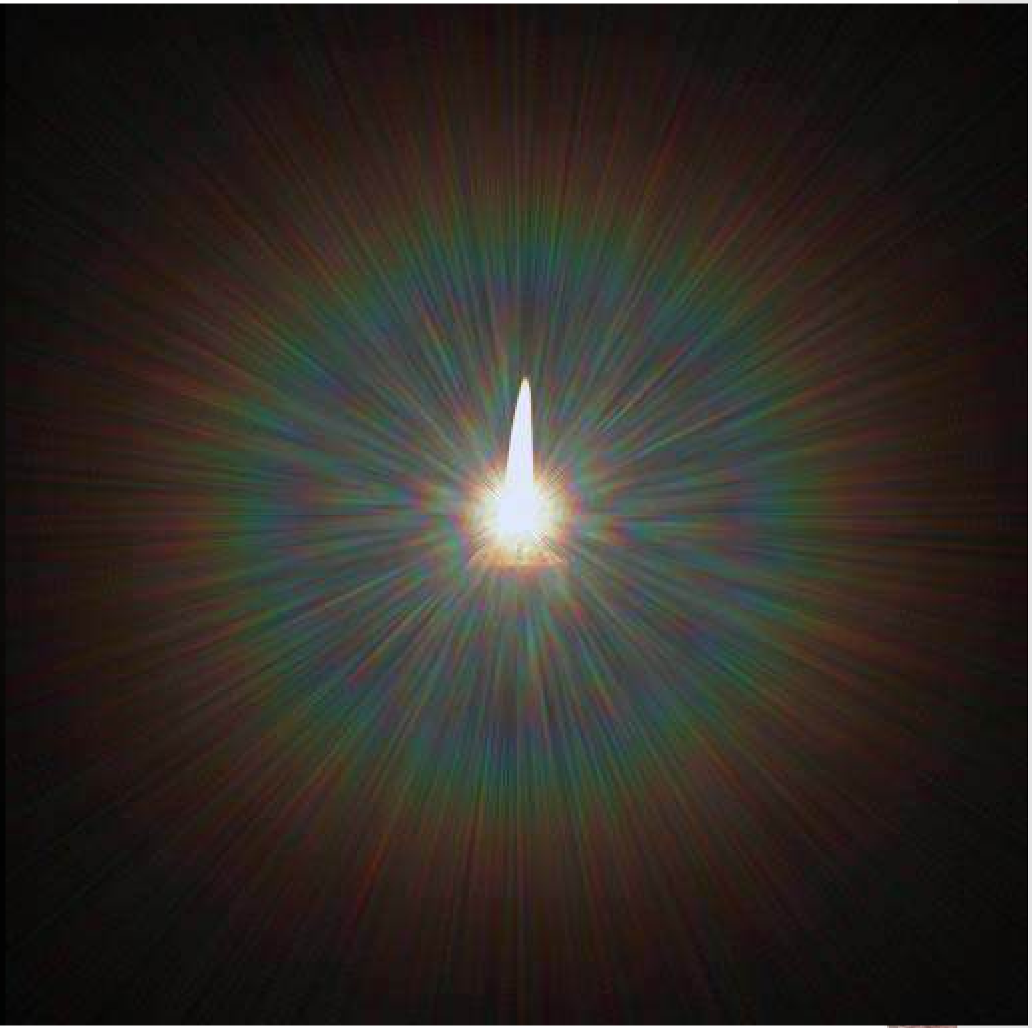
HDR image PSF Bright pixels Billboard



Convolution

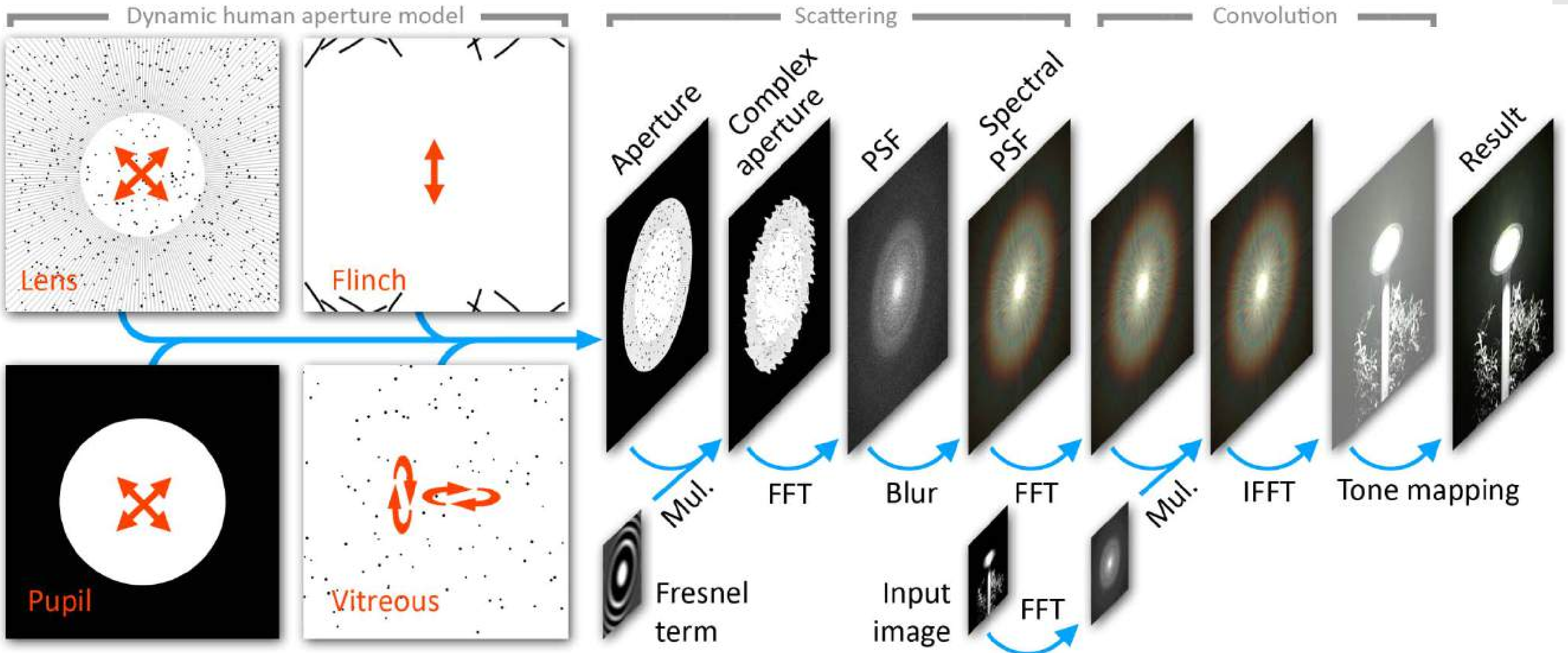
Convolution

Convolution



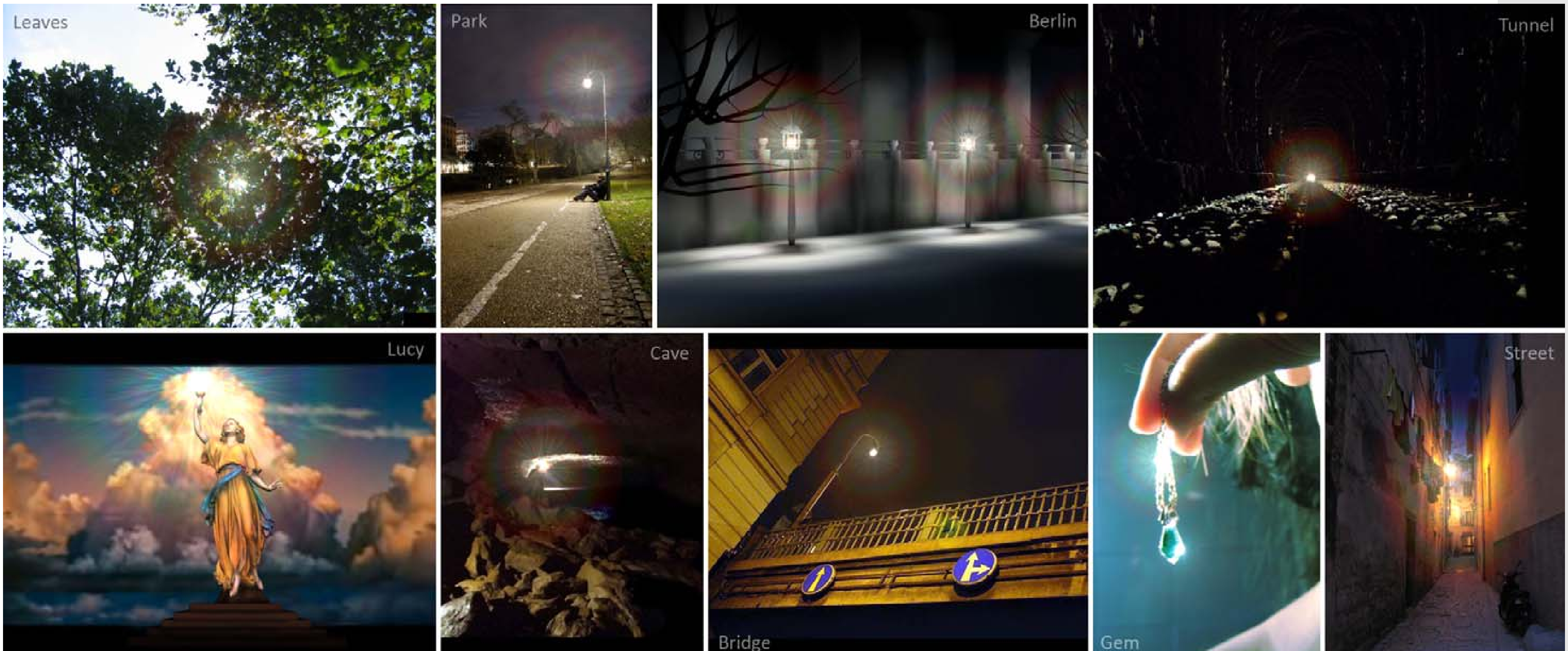
Billboard

Temporal Glare Pipeline



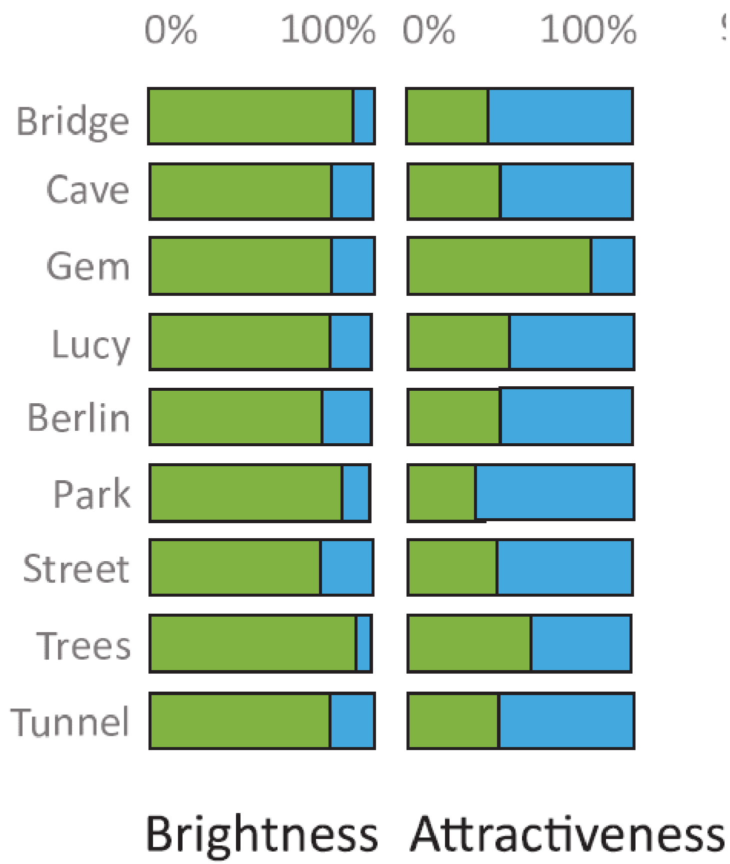
Results: Study

1. Two-alternative-forced-choice (bright, attractive, real)
   10 subjects
2. Method of adjustment
   4 subjects



Results: Study

1. Two-alternative-forced-choice (bright, attractive)
   10 subjects
2. Method of adjustment: dynamic glare ~5% brighter
   4 subjects



Summary/Limitations

- Glare illusion might boost apparent brightness up to 30%
- Comprehensible model of light scattering in the eye taking into account dynamic eye elements
- Real-time rendering
- Model might miss important parts
- Model might contain unimportant parts
  - No differential study
- Other temporal low-level eye physics like
  - Floaters
  - Local adaptation (“After images”)

http://www.mpi-inf.mpg.de/resources/hdr/TemporalGlare/

Acknowledgements

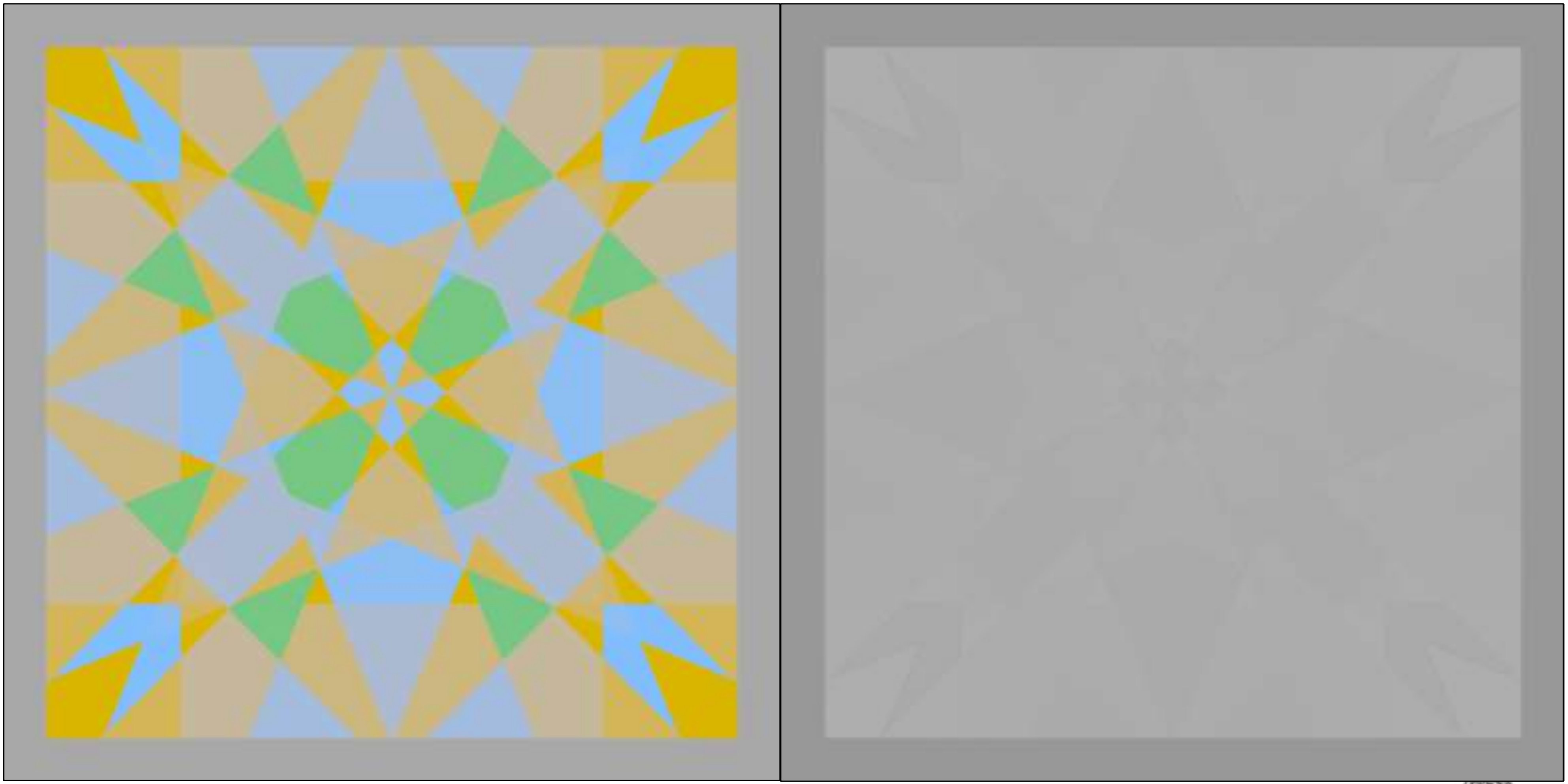
- I would like to thank Grzegorz Krawczyk, Tobias Ritschel, Kaleigh Smith, Akiko Yoshida, and Matthias Ihrke for help in preparing slides.

Retargeting Color Content: Color Issues in Tone Mapping

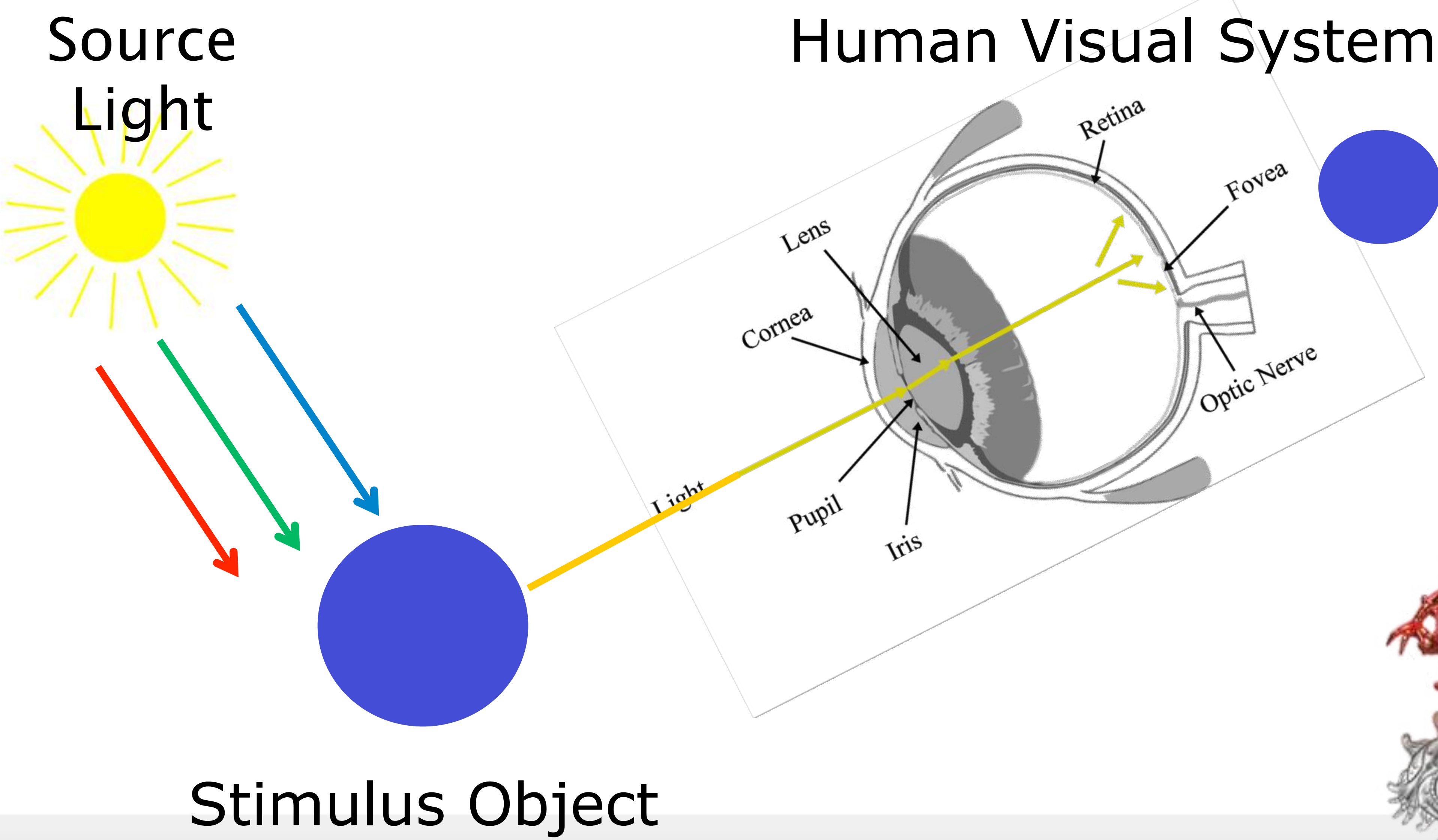
Alessandro Artusi

Cyprus Institute, CaSToRC, Cyprus

Introduction to Color



What is Color?



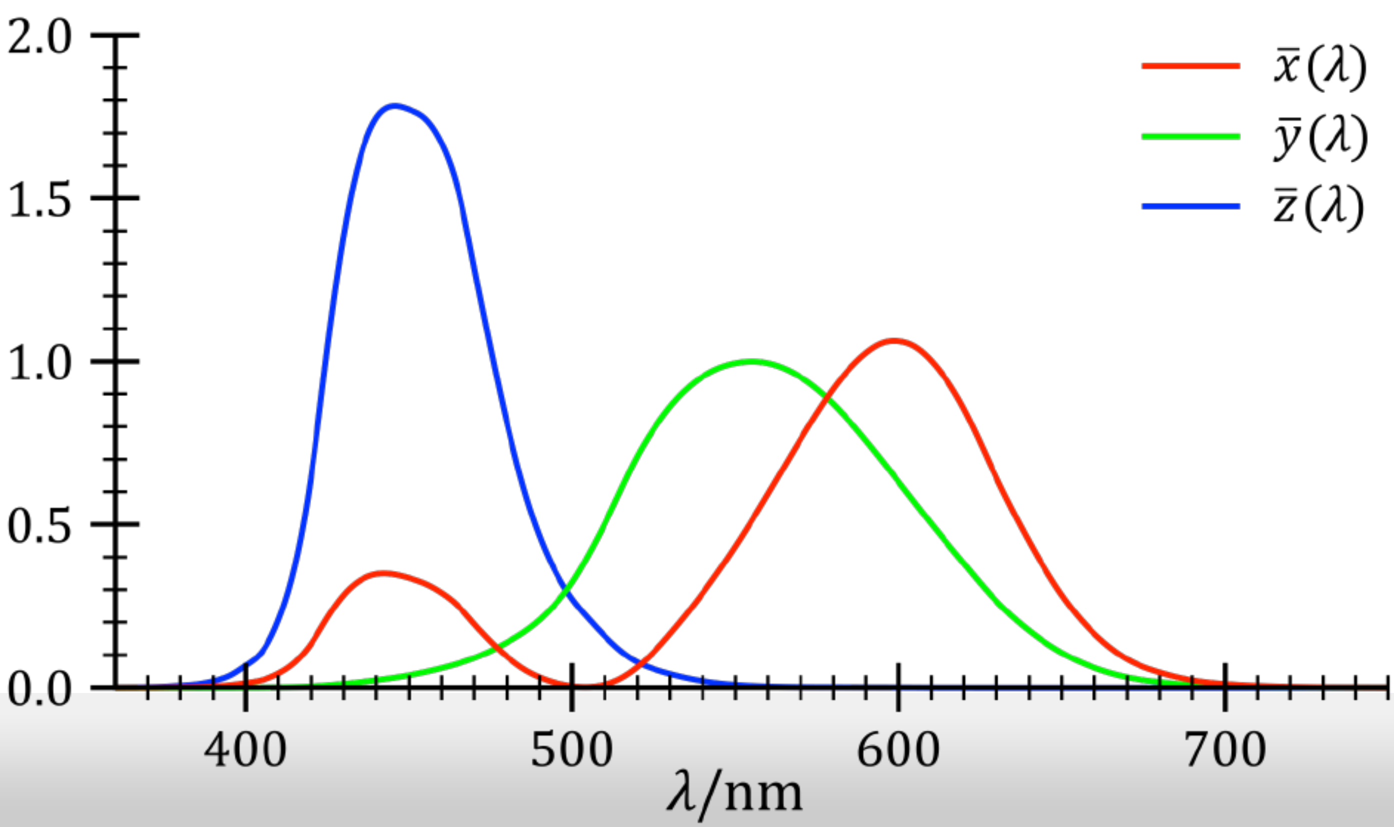
Quantifying Color

$I(\lambda)$	SPD of the light
$\rho(\lambda)$	Reflectance of the object
$\bar{x},\bar{y},\bar{z}(\lambda)$	CIE color matching functions

$$X = \int_0^\infty I(\lambda)\rho(\lambda)\bar{x}(\lambda)d\lambda$$

$$Y = \int_0^\infty I(\lambda)\rho(\lambda)\bar{y}(\lambda)d\lambda$$

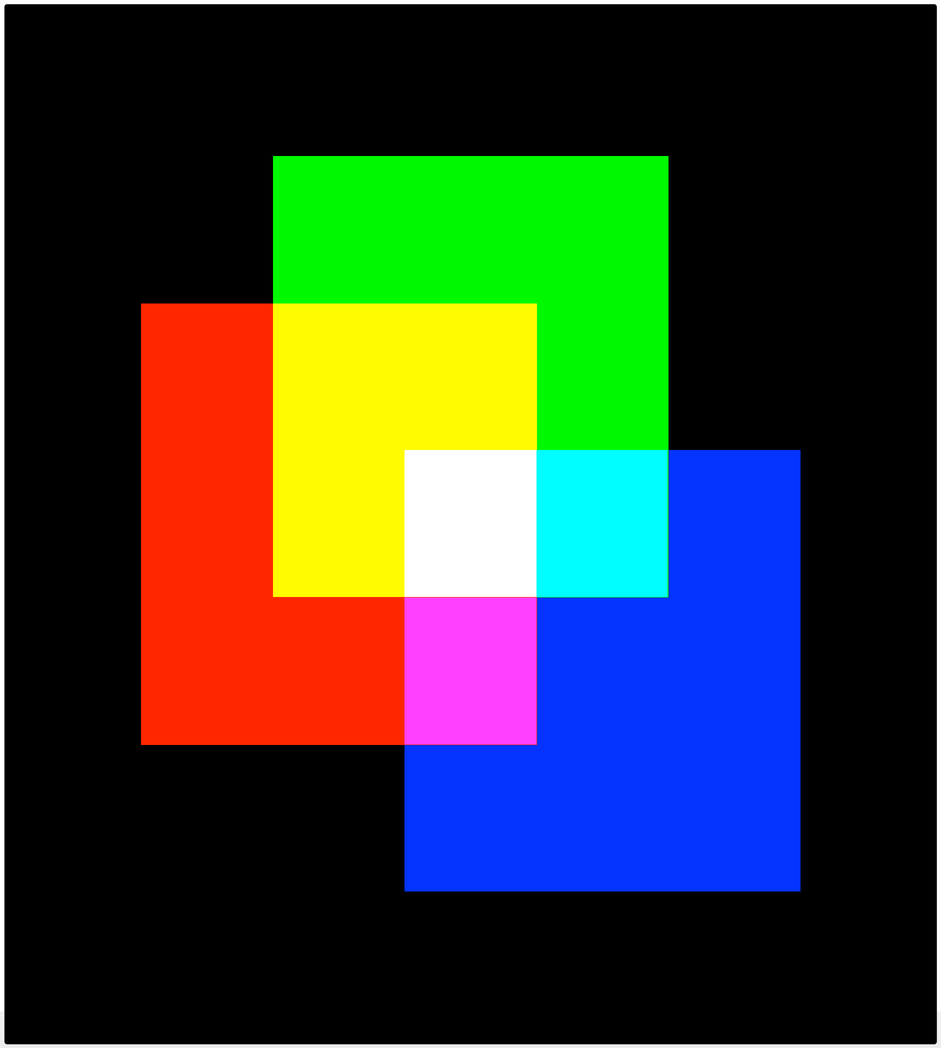
$$Z = \int_0^\infty I(\lambda)\rho(\lambda)\bar{z}(\lambda)d\lambda$$



How Color is Produced?

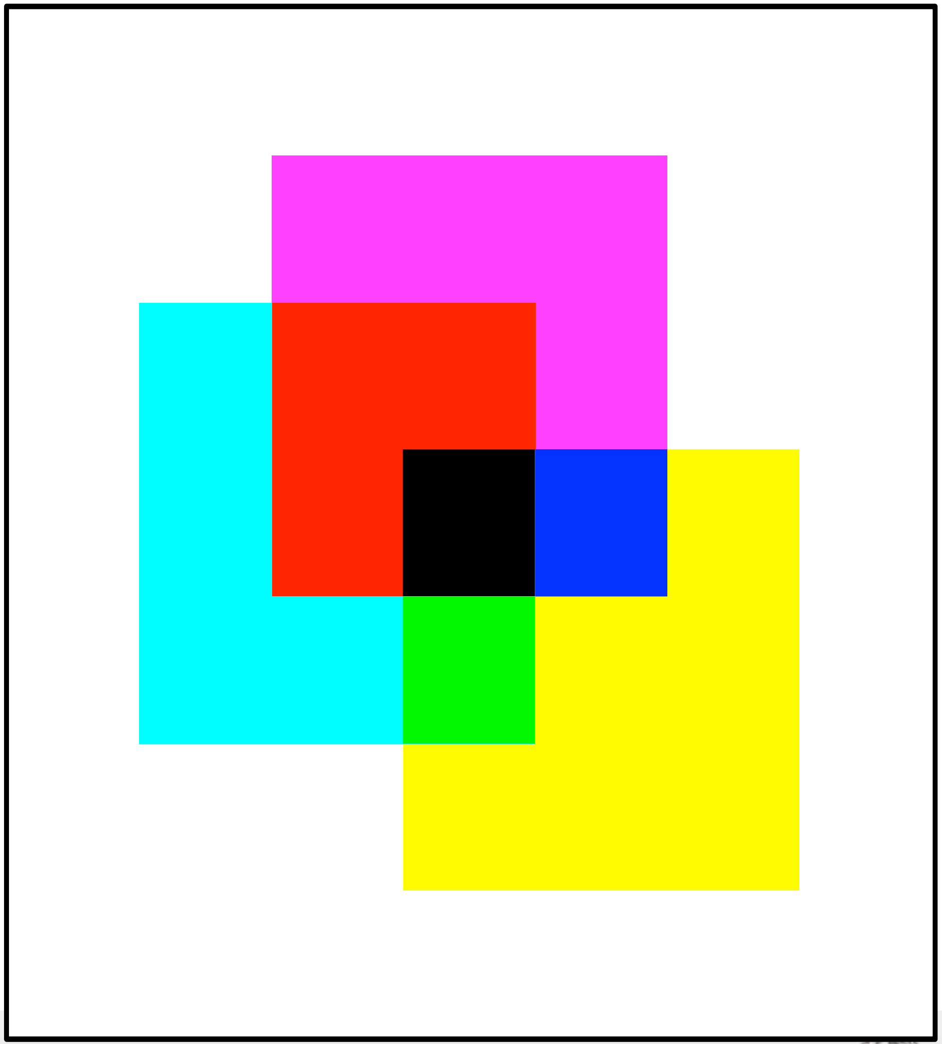
SIGGRAPHASIA2011 HONG KONG

Additive



(a)

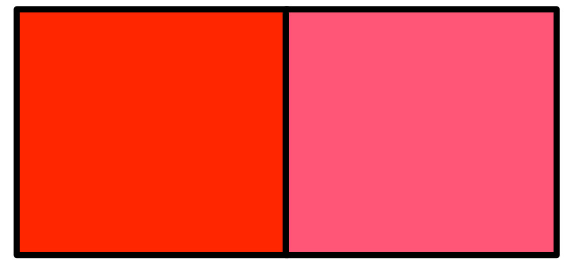
Subtractive



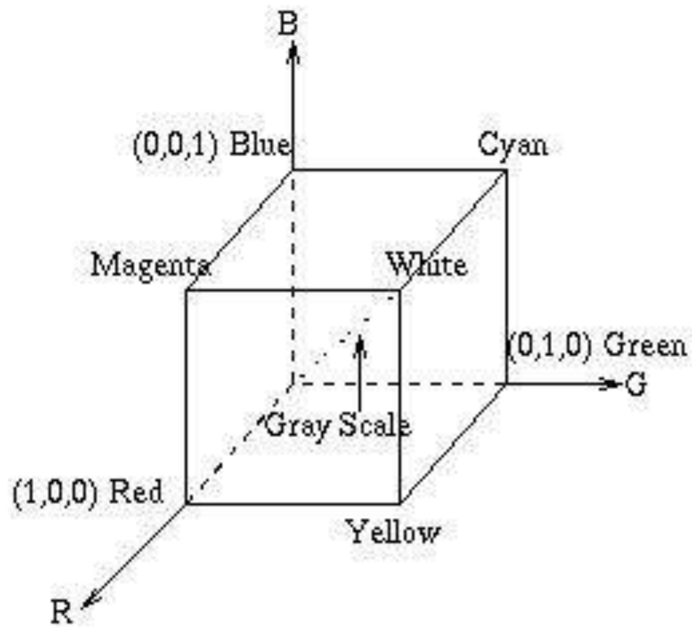
(b)

Color Space

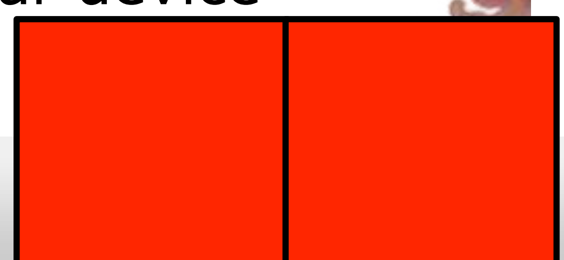
- Device dependent: the description of color information is related to the characteristics of a particular device



  - Set of primaries
  - Technology



- Device independent: the description of color information is not dependent from the characteristics of a particular device
  - CIEXYZ, CIELab, CIELuv etc...



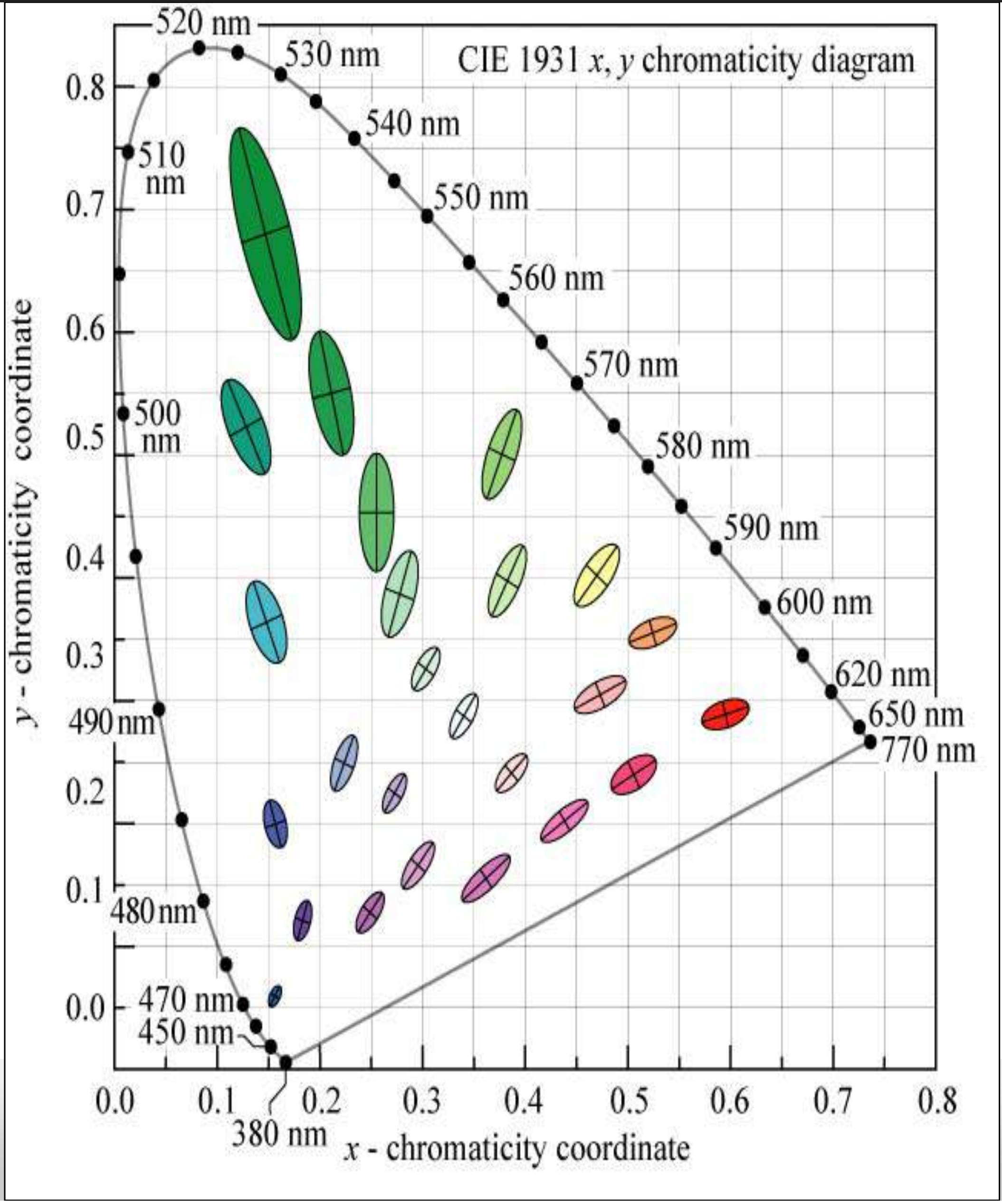
Chromaticity Diagram and MacAdam's Ellipses

SIGGRAPHASIA2011 HONG KONG

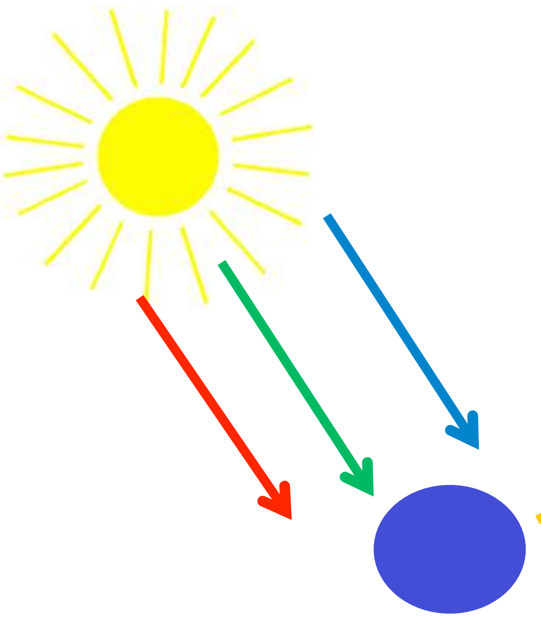
- **MacAdam's Ellipses**
  - contains all colors which are indistinguishable to an human observer from the color at the center of the ellipse
  - the contour of the ellipse represents the just noticeable differences of chromaticity

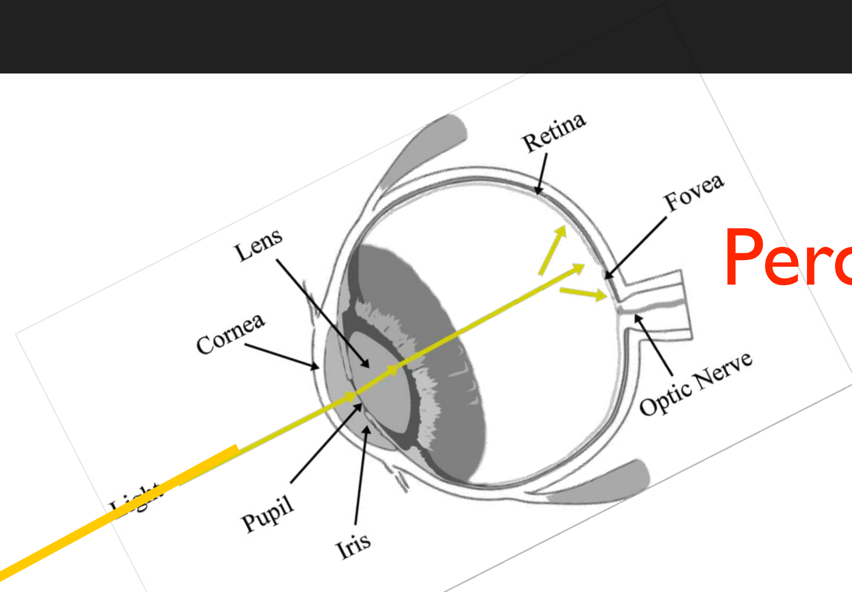
$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$



Color Attributes by the CIE





Perception

- **Hue**
- **Saturation**
- **Lightness**

- **Hue** The degree to which a stimulus can be described as similar to or different from stimuli that are described as red, green, blue, and yellow.
- **Saturation** is the colorfulness of an area judged in proportion to its brightness.
- **Lightness** Human vision has a nonlinear perceptual response to luminance: The perceptual response to luminance is called lightness.

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 \qquad 0.008856 < \frac{Y}{Y_n}$$

Color in High Dynamic Range

SIGGRAPHASIA2011 HONG KONG

- Color Ratio (Schlick 1994)

$$RGB_{out} = \frac{RGB_{in}}{L_{in}} L_{out}$$

RGB_{in}	Color Input
RGB_{out}	Color Output
L_{in}	Luminance Input
L_{out}	Luminance Output

$\longrightarrow$

Mantiuk et al.. "Color Correction for Tone Mapping", Proceedings Eurographics 2009.

Color in High Dynamic Range

- Saturation Control (Thumblin and Turk 1999)

$$RGB_{out} = \left(\frac{RGB_{in}}{L_{in}}\right)^{s} L_{out}$$

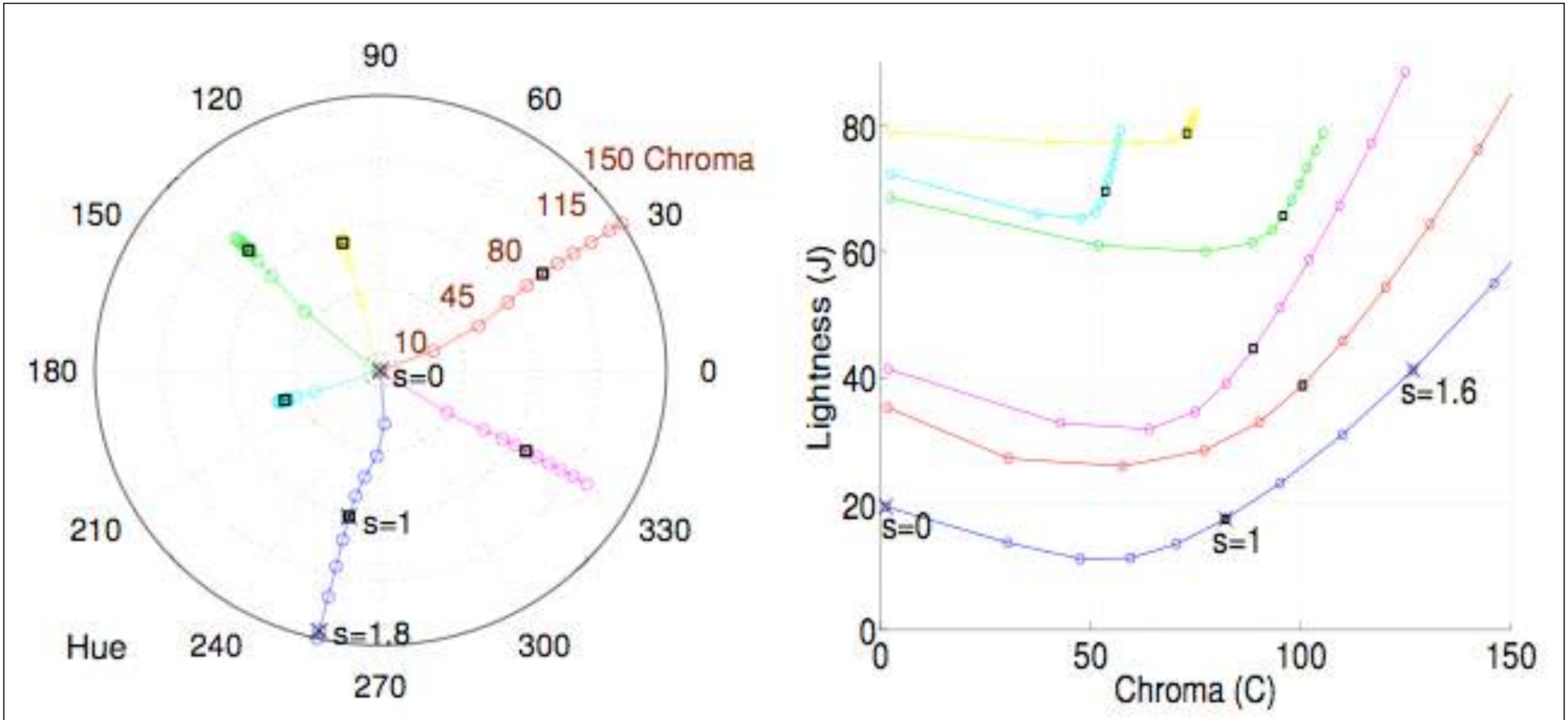
s Saturation Parameter

c Contrast Compression

Under-saturated colors for S=C.

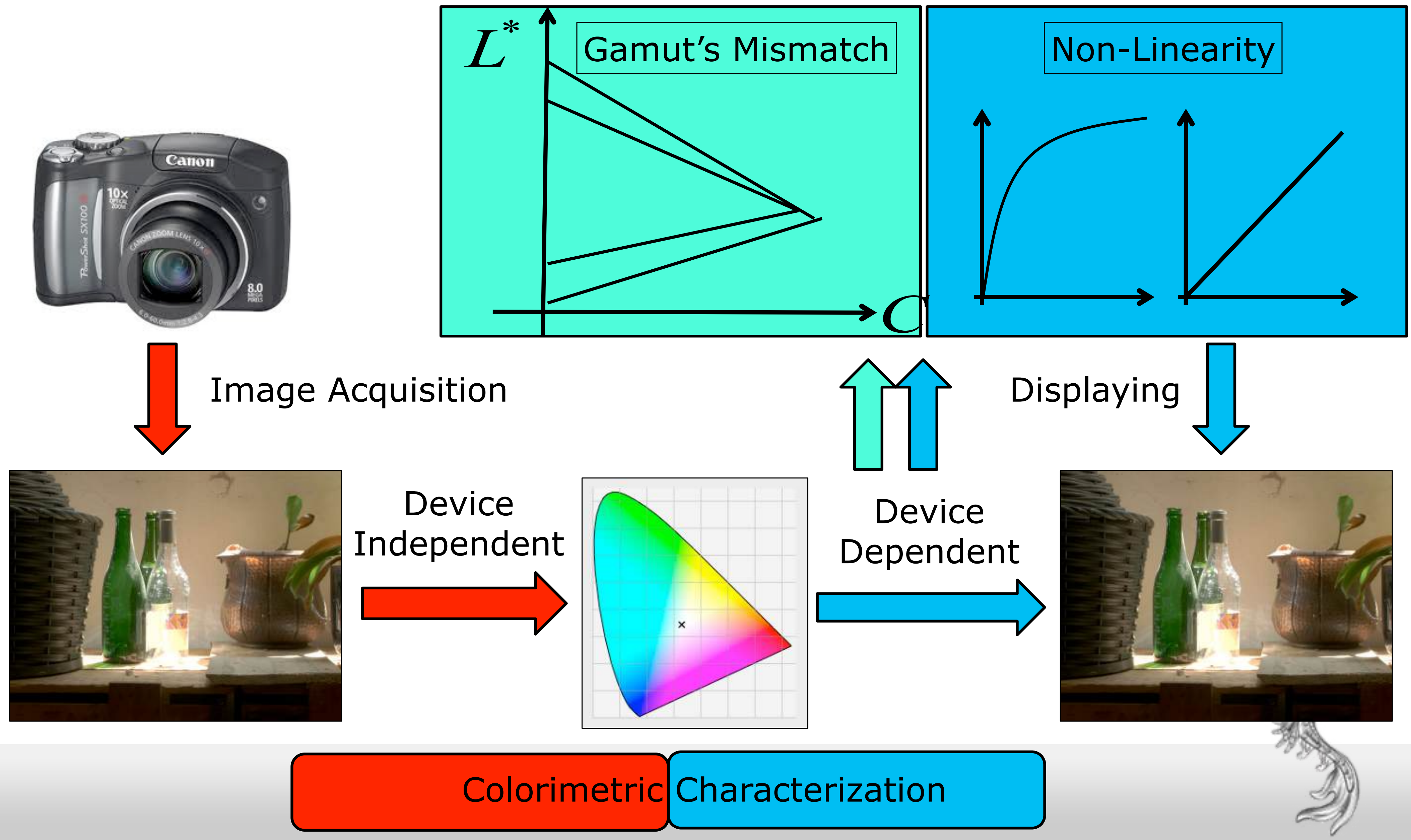
Mantiuk et al..“Color Correction for Tone Mapping”, Proceedings Eurographics 2009.

Color in High Dynamic Range



Mantiuk et al..“Color Correction for Tone Mapping”, Proceedings Eurographics 2009.

Color Rendering Pipeline (8 Bit)



Colorimetric Characterisation of a Device

Device

Minimisation Process

Spectrophotometer

RAW Data are Linear

R G B

3x3 Matrix

X Y Z

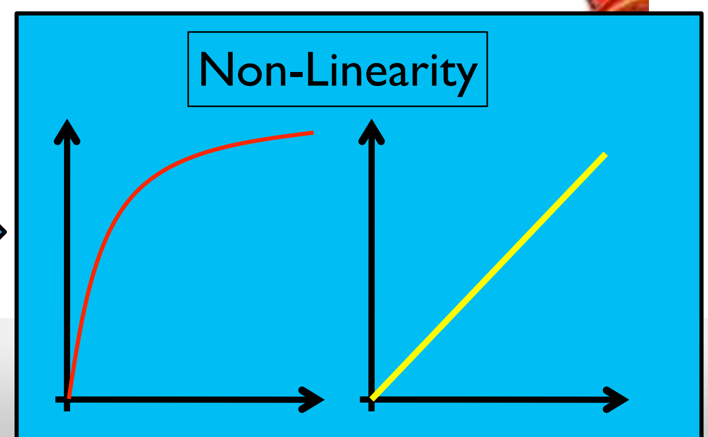
$$redTRC^{-1}[linR] = [deviceR]$$
$$greenTRC^{-1}[linG] = [deviceG]$$
$$blueTRC^{-1}[linB] = [deviceB]$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \begin{bmatrix} red_X green_X blue_X \\ red_Y green_Y blue_Y \\ red_Z green_Z blue_Z \end{bmatrix}^{-1} = \begin{bmatrix} linR \\ linG \\ linB \end{bmatrix}$$

X Y Z

R G B

3x3 Primaries Matrix

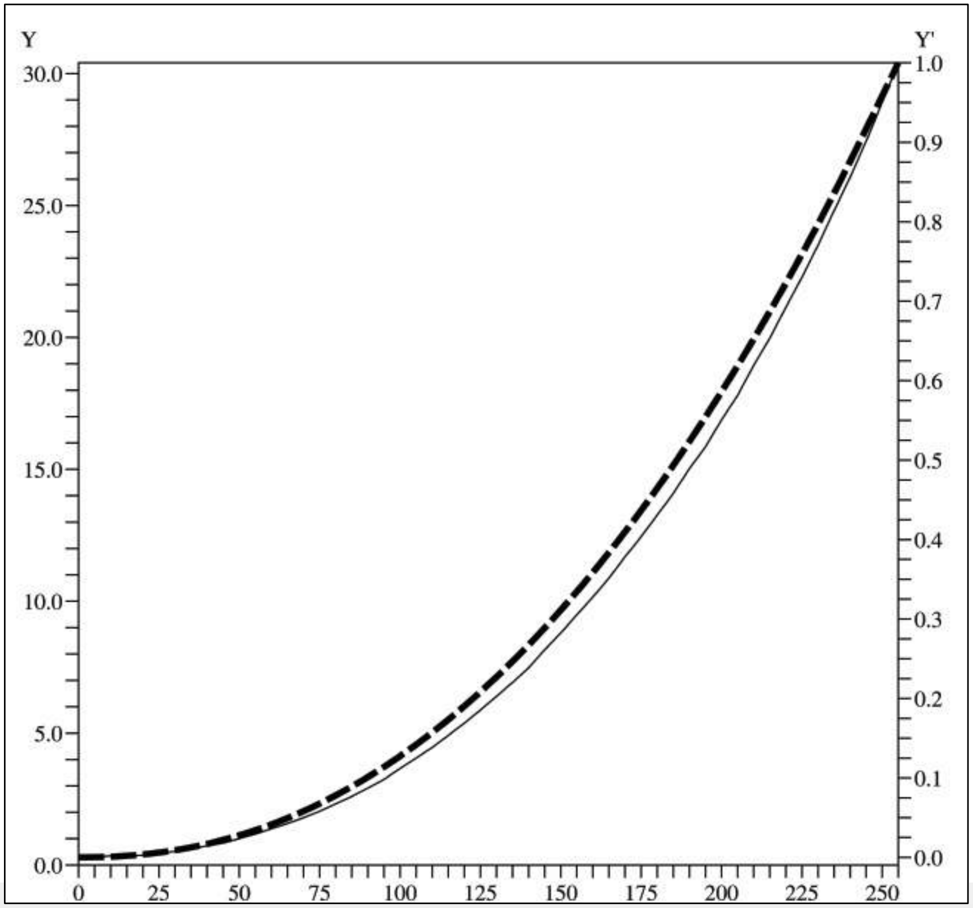


Gamma – Curve

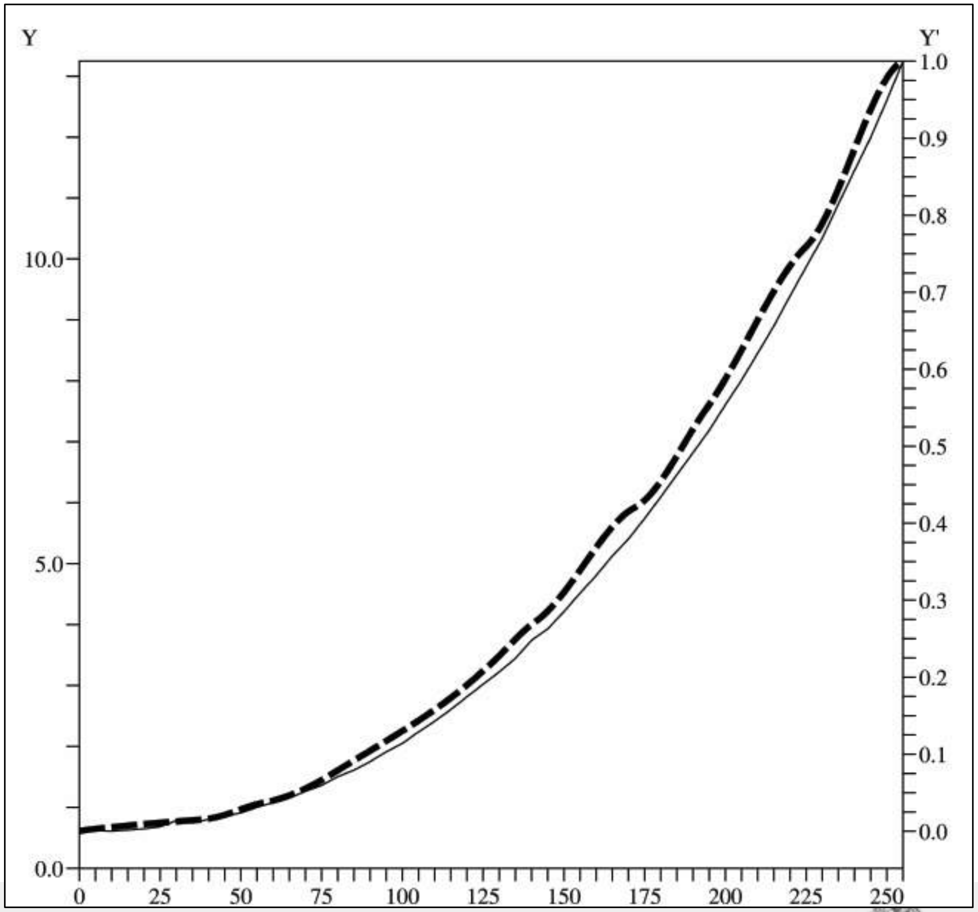
....

$$R(d_i) = ((1-b)d_i + b)^{\gamma}$$

Gamma Response for RED



Gamma Response for BLUE



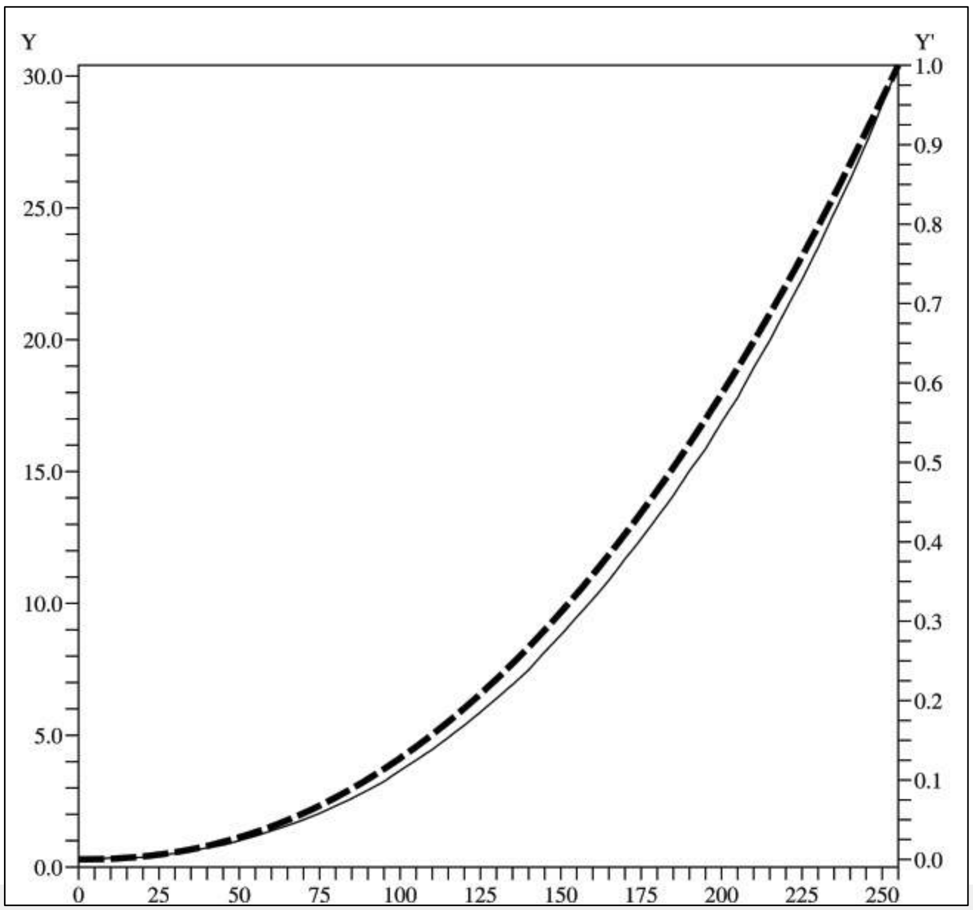
A. Neumann, A. Artusi, L. Neumann, G. Zotti and W. Purgathofer "Accurate Display Gamma Function based on Human Observation".

Gamma – Curve

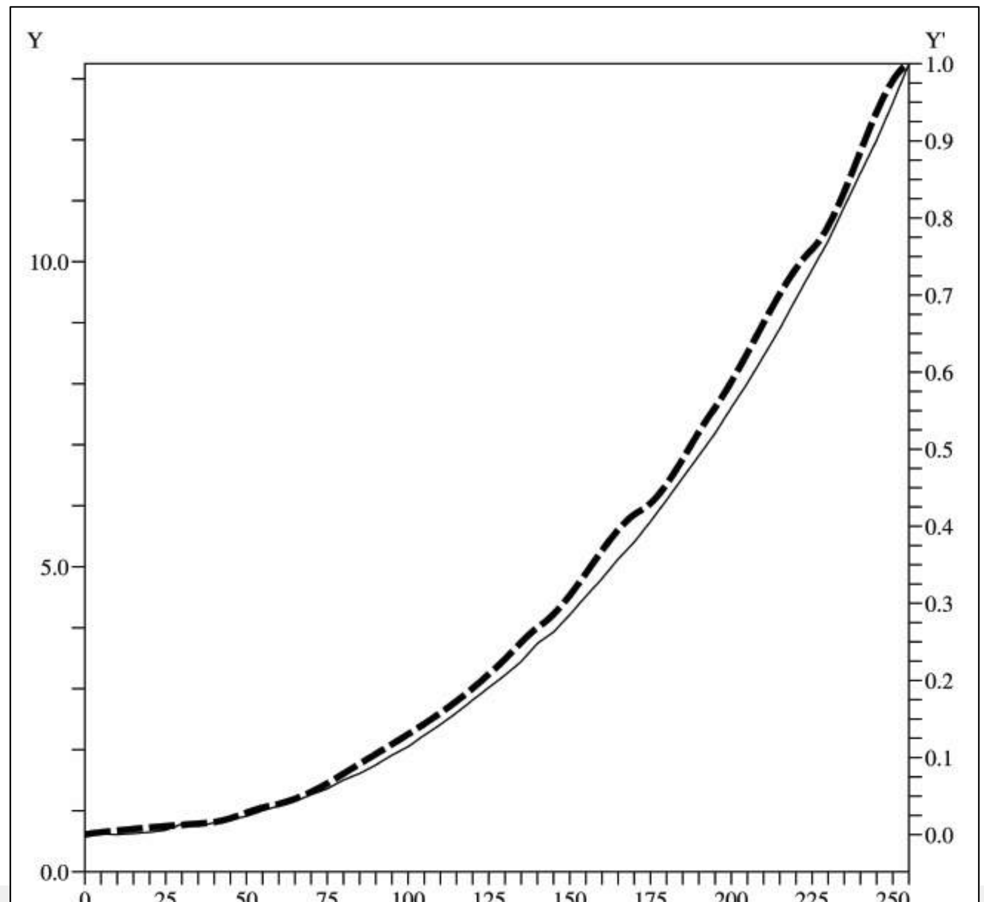
SIGGRAPHASIA2011 HONG KONG

$$R(d_i) = \left((1-b)d_i + b\right)^{\gamma}$$

Gamma Response for RED



Gamma Response for BLUE



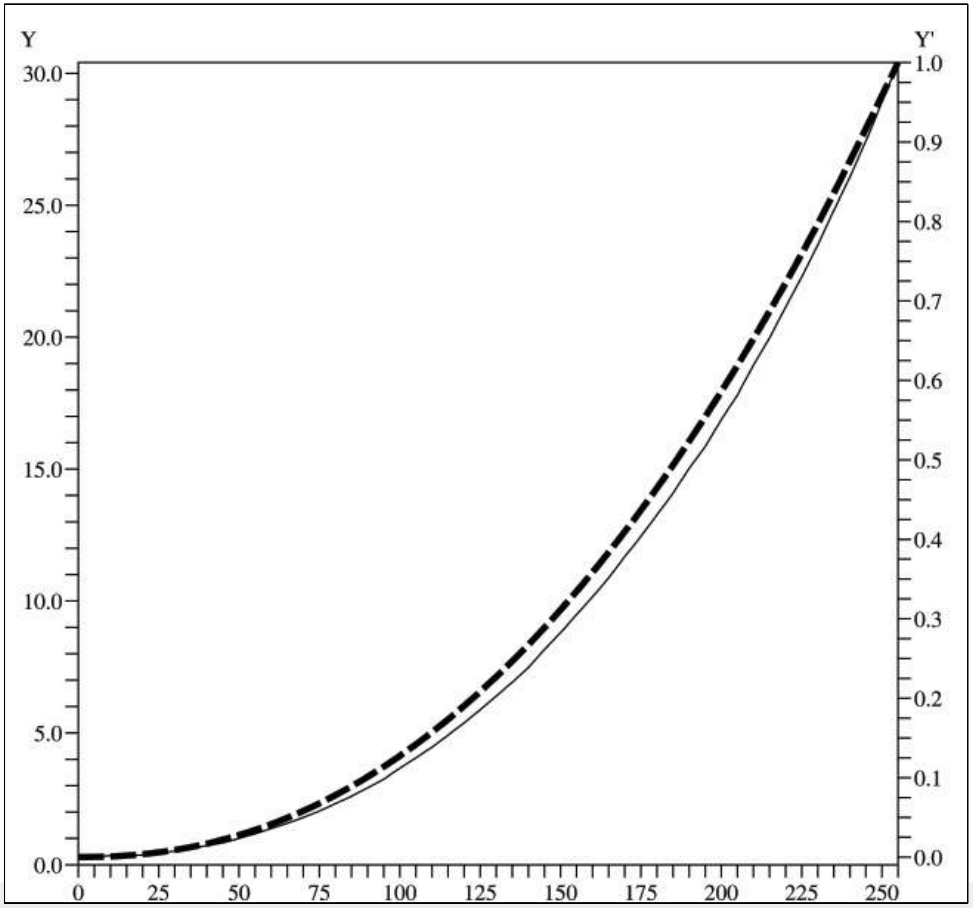
A. Neumann, A. Artusi, L. Neumann, G. Zotti and W. Purgathofer “Accurate Display Gamma Function based on Human Observation“.

Gamma – Curve

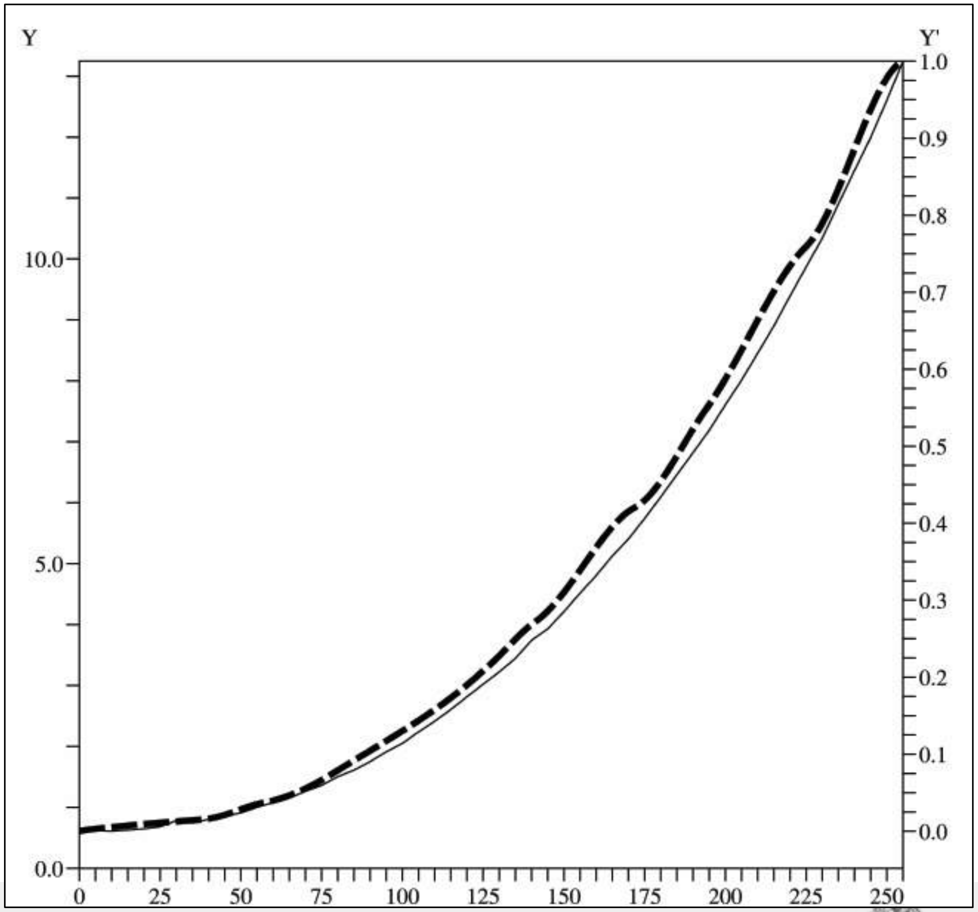
SIGGRAPHASIA2011 HONG KONG

$$R(d_i) = ((1-b)d_i + b)^{\gamma}$$

Gamma Response for RED



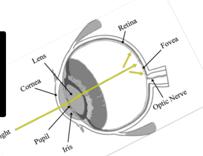
Gamma Response for BLUE



A. Neumann, A. Artusi, L. Neumann, G. Zotti and W. Purgathofer “Accurate Display Gamma Function based on Human Observation“.

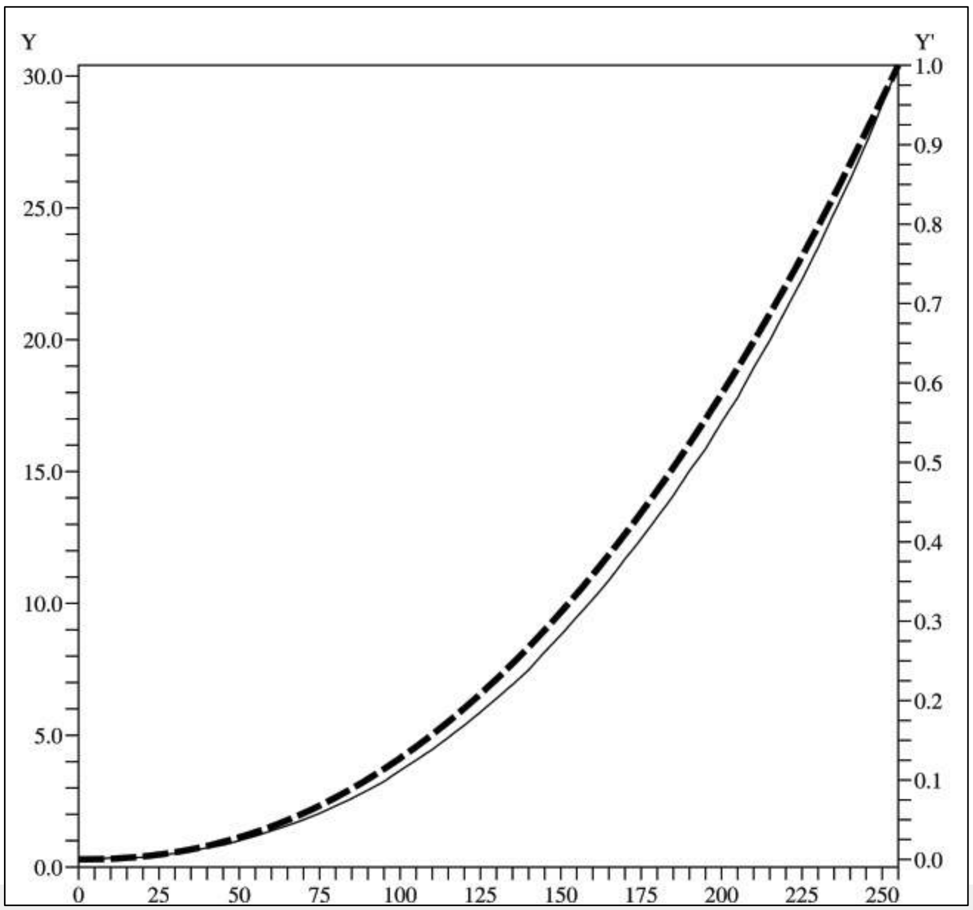
Gamma – Curve

SIGGRAPHASIA2011 HONG KONG

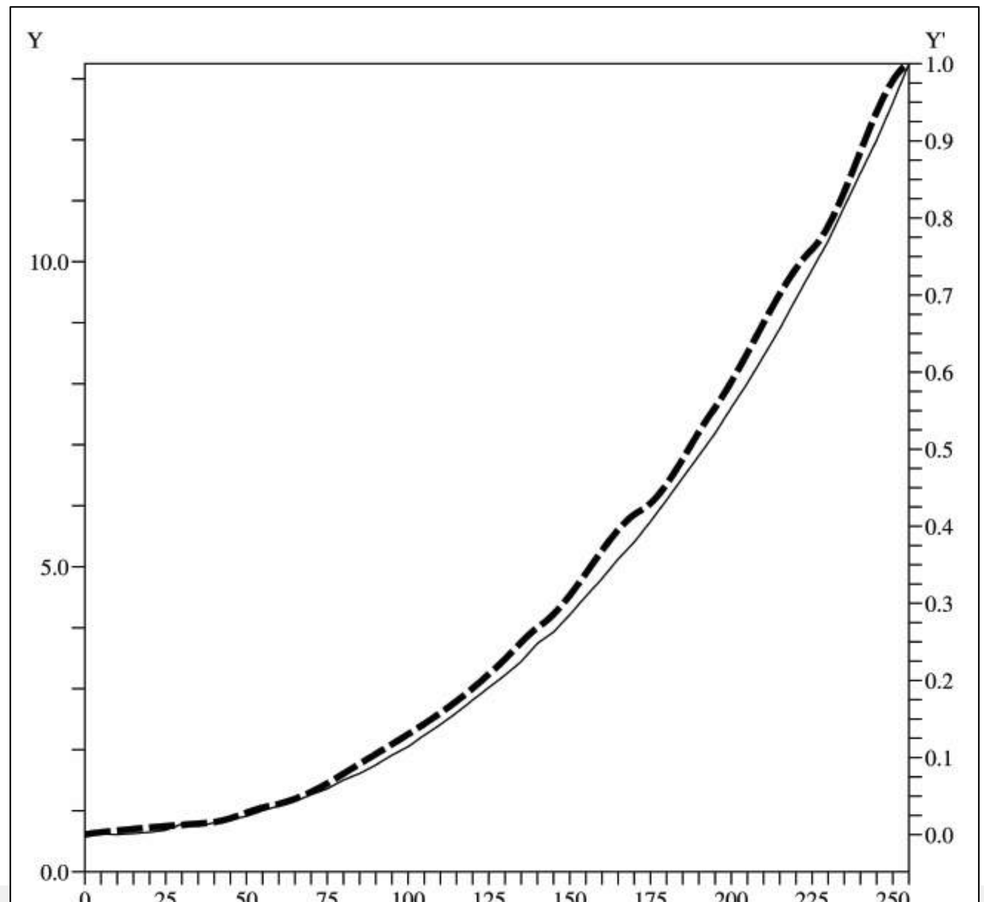


$$R(d_i) = \left((1-b)d_i + b\right)^{\gamma}$$

Gamma Response for RED

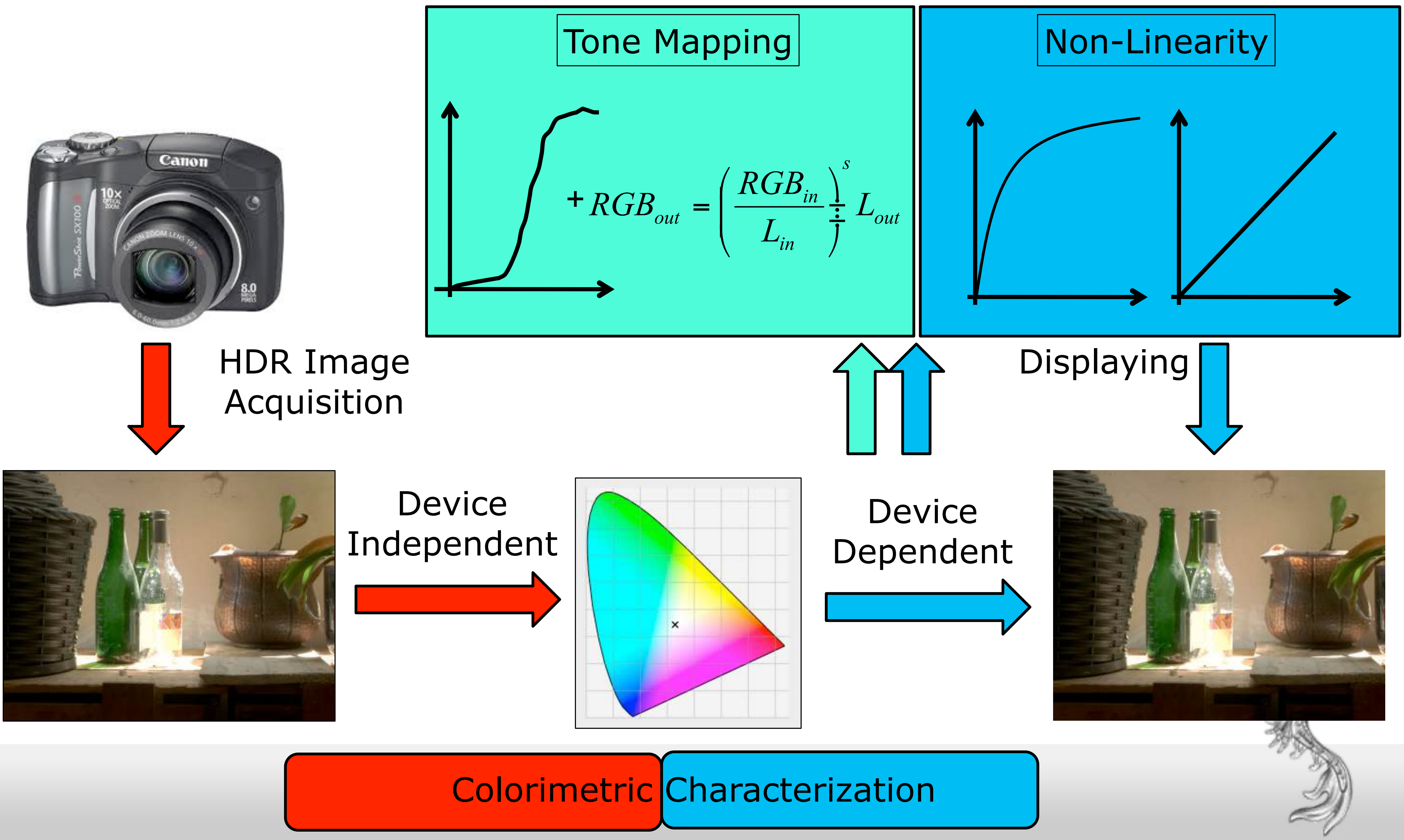


Gamma Response for BLUE



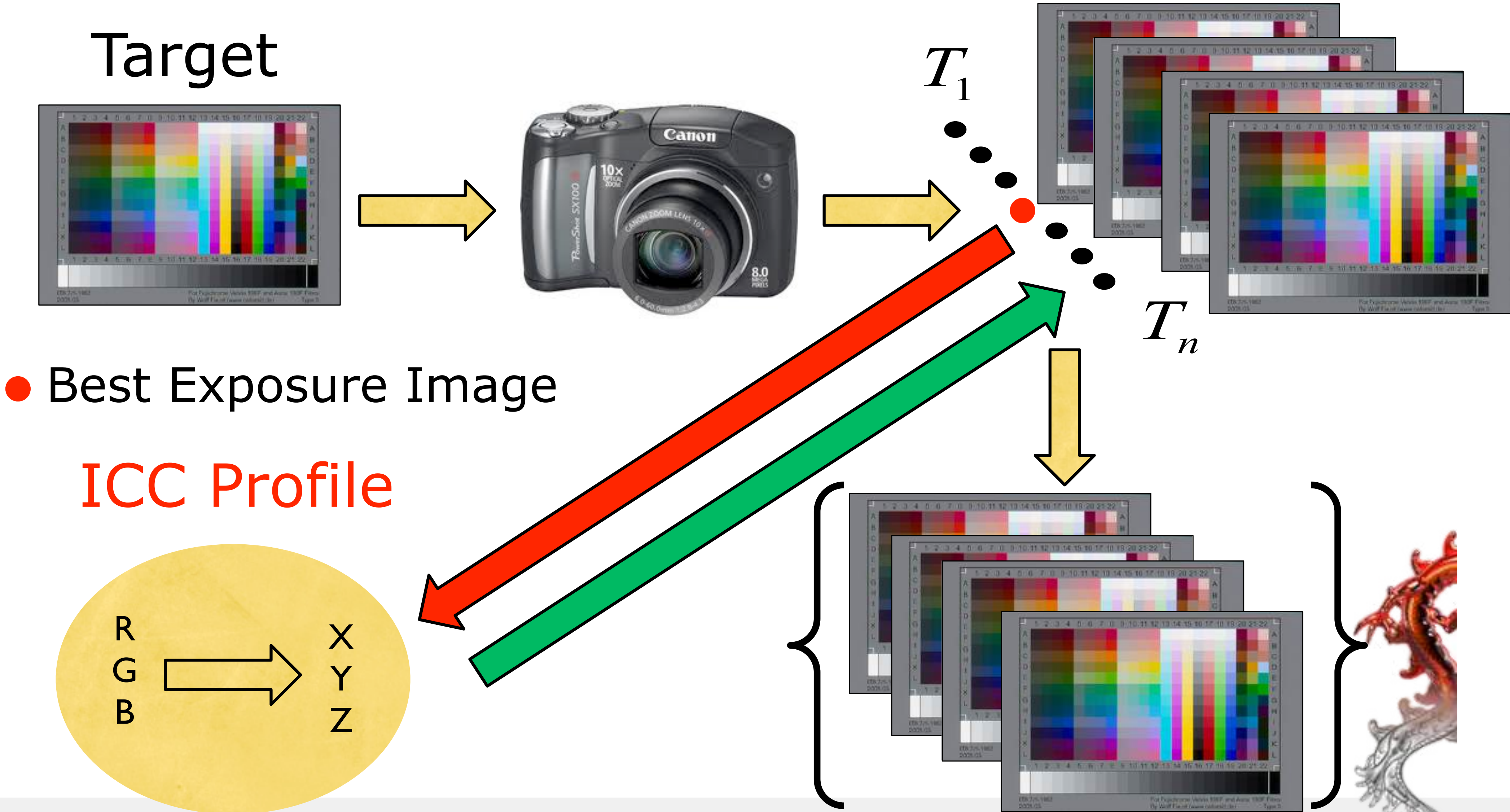
A. Neumann, A. Artusi, L. Neumann, G. Zotti and W. Purgathofer “Accurate Display Gamma Function based on Human Observation“.

Color Rendering Pipeline in HDR



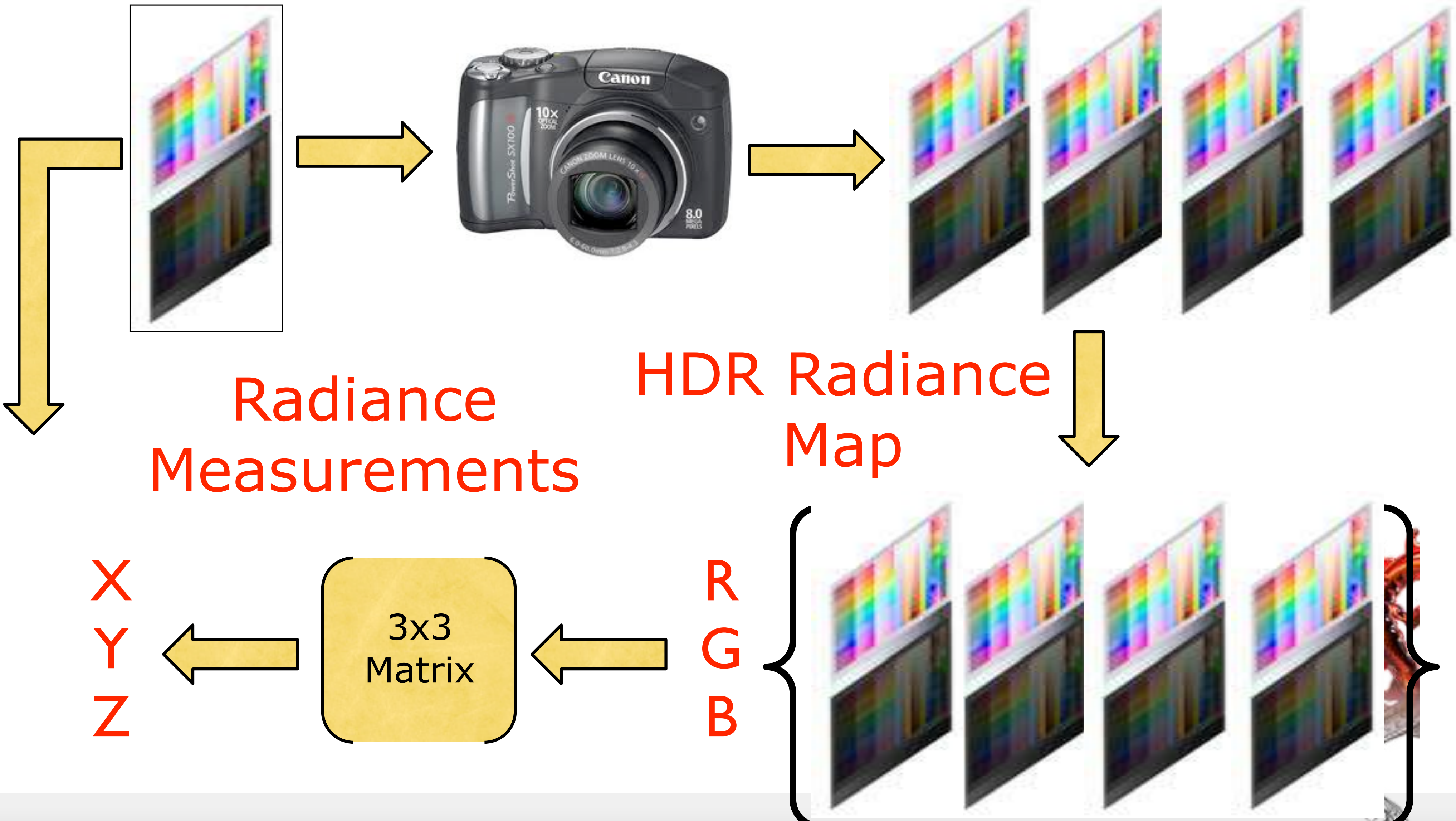
HDR ICC Profile

SIGGRAPHASIA2011 HONG KONG



Goesele et al. "Color Calibrated High Dynamic Range Imaging with ICC Profiles."

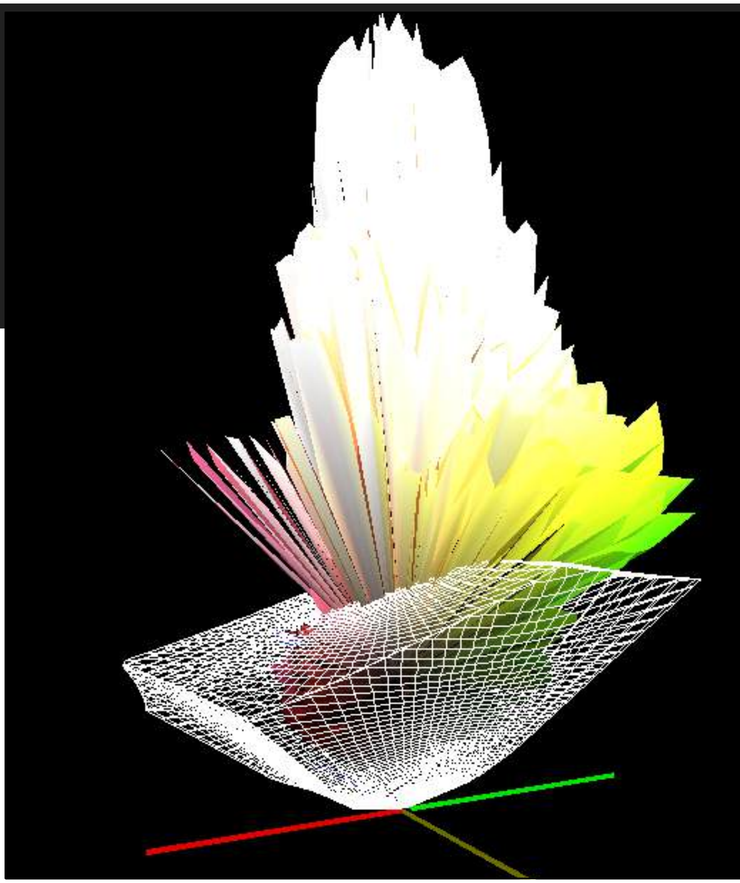
HDR Colorimetric Camera Characterization



Min H. Kim et al. “Characterization of High Dynamic Range Imaging.”

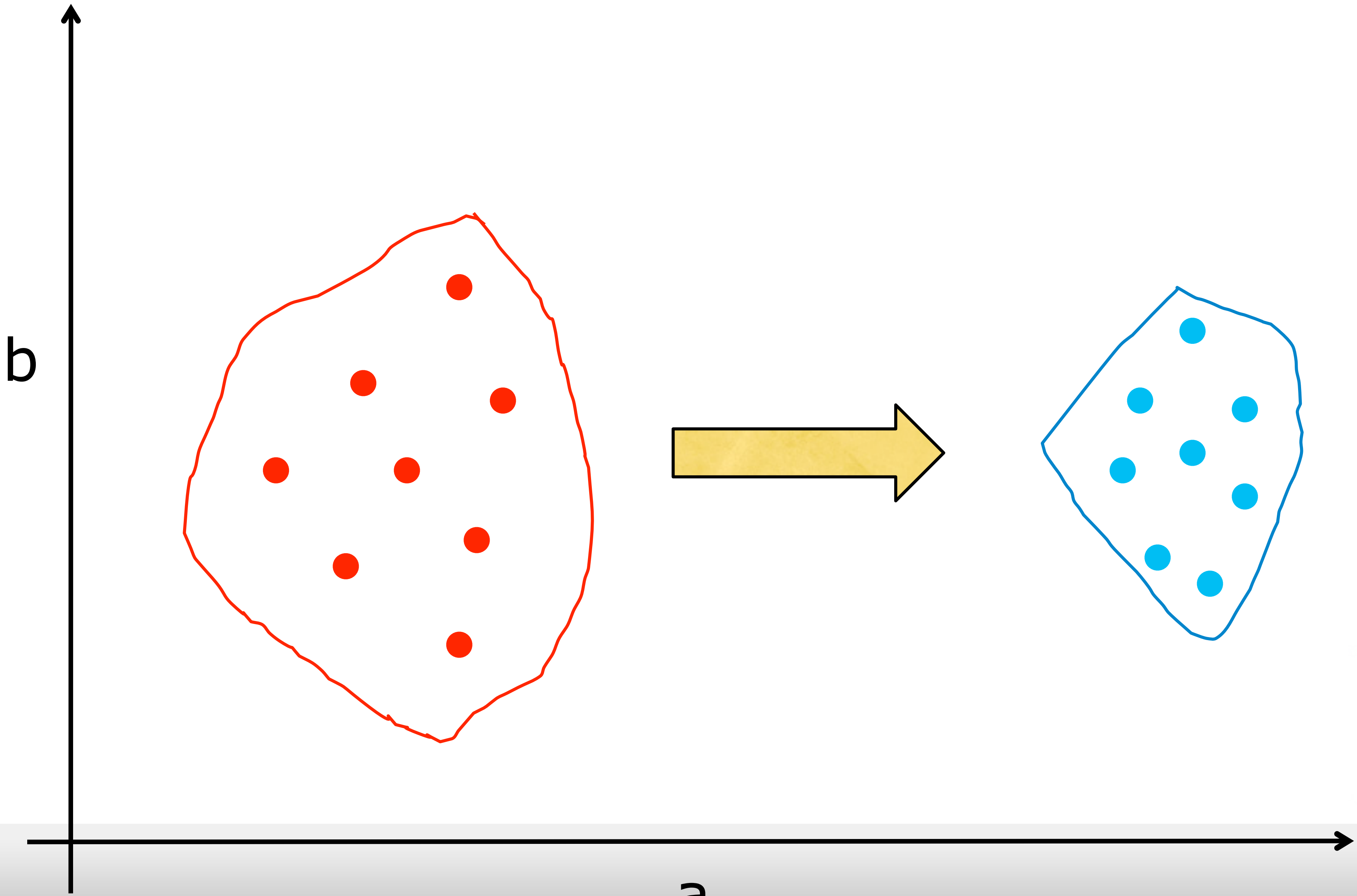
Color Gamut

SIGGRAPHASIA2011 HONG KONG

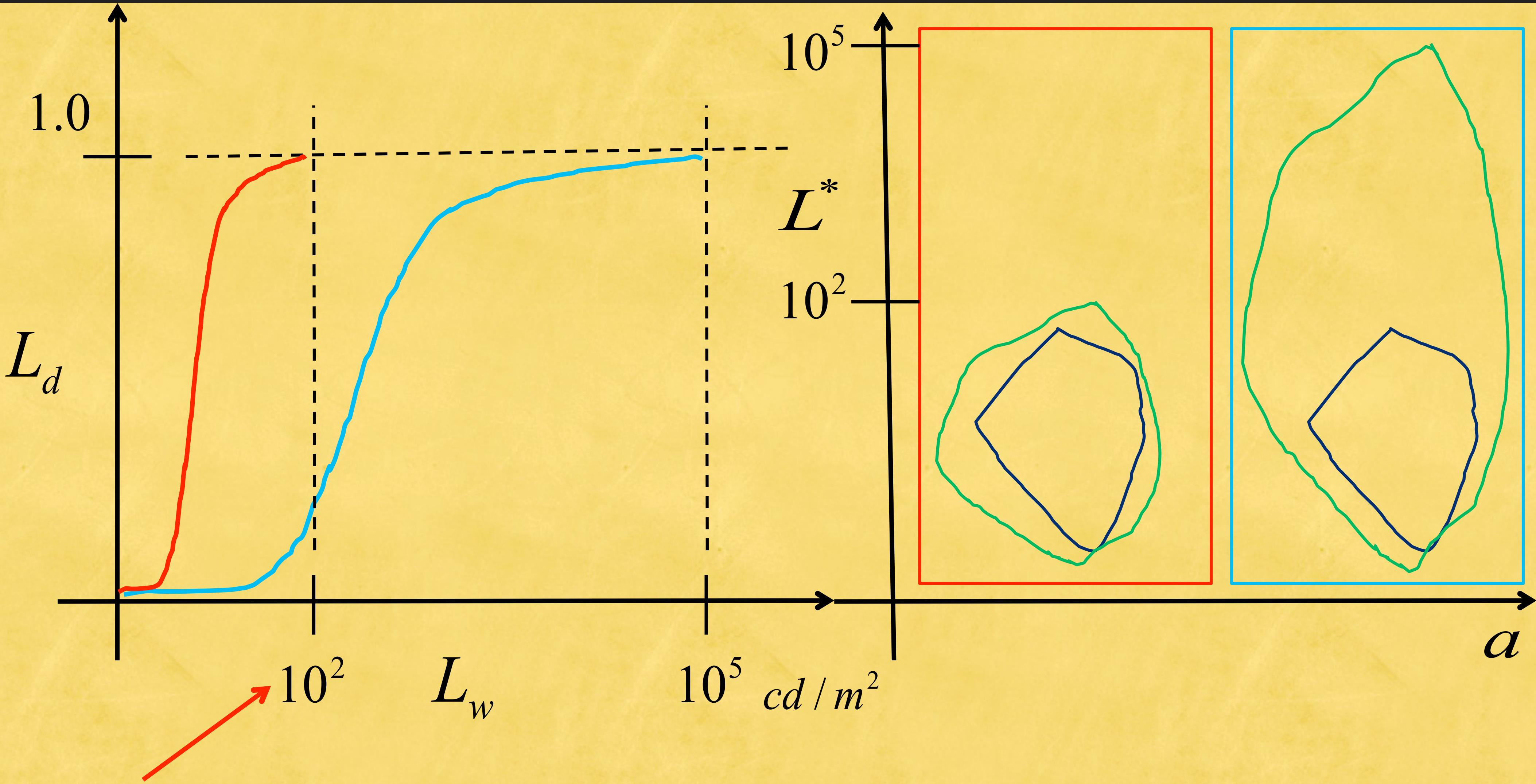


- **Device**
  - Set of colors reproducible by the device
- **Image**
  - Set of colors that compose the image

Gamut Mapping



Gamut vs. Tone Mapping

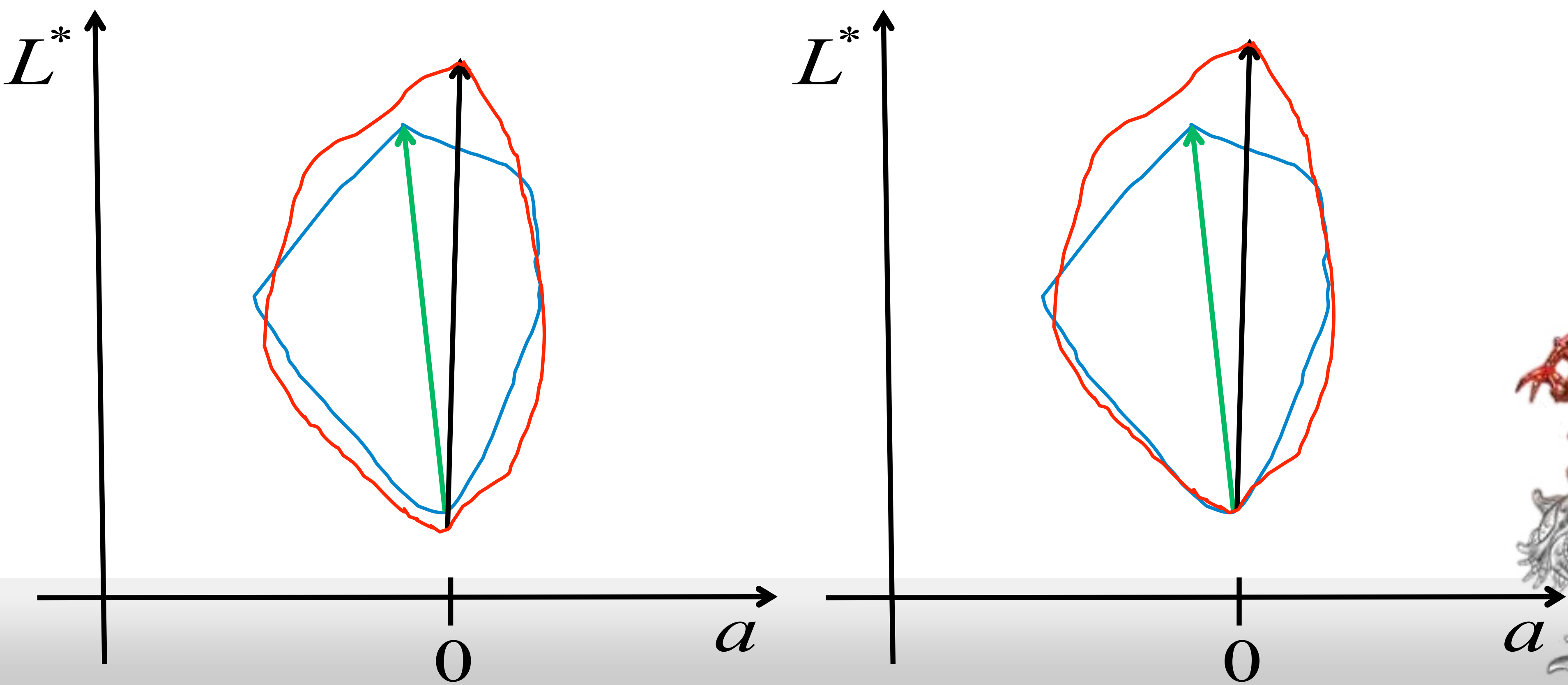


Not HDR Content

Gamut Mapping Aims (CS)

SIGGRAPHASIA2011 HONG KONG

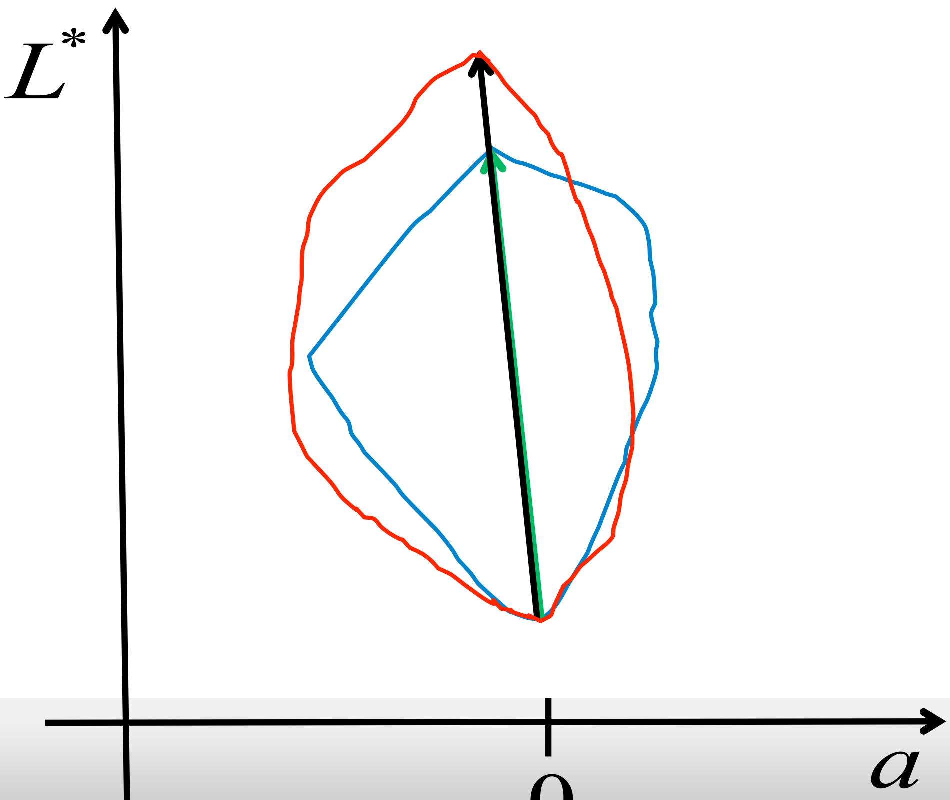
- Gray axes alignment, mapping white to white and black to black

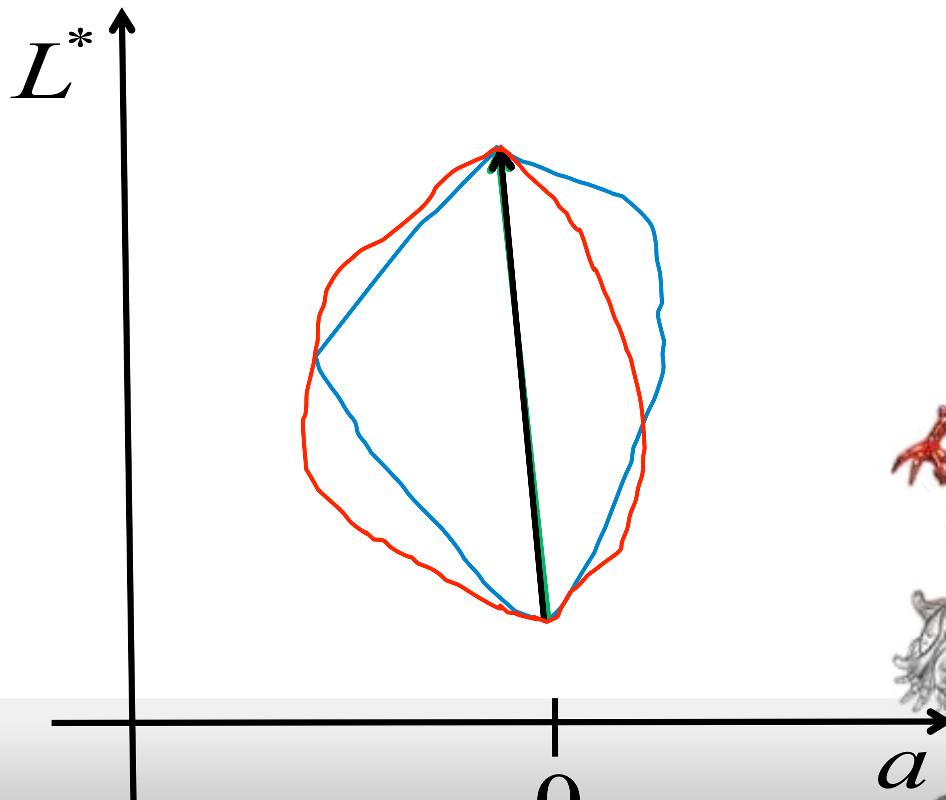


Gamut Mapping Aims (CS)

SIGGRAPHASIA2011 HONG KONG

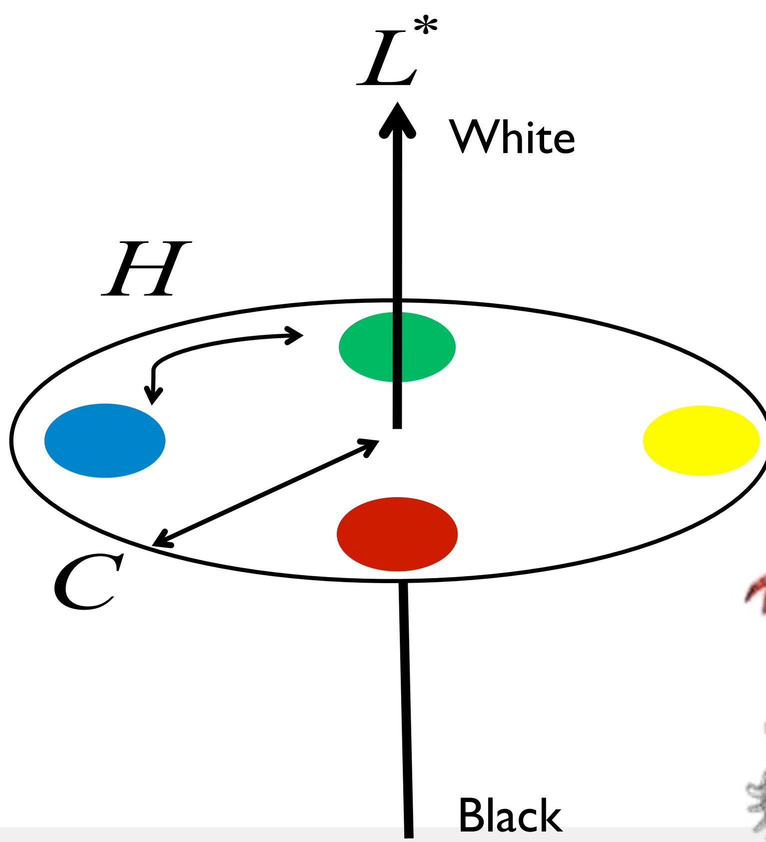
- Gray axes alignment, mapping white to white and black to black





Gamut Mapping Aims (CS)

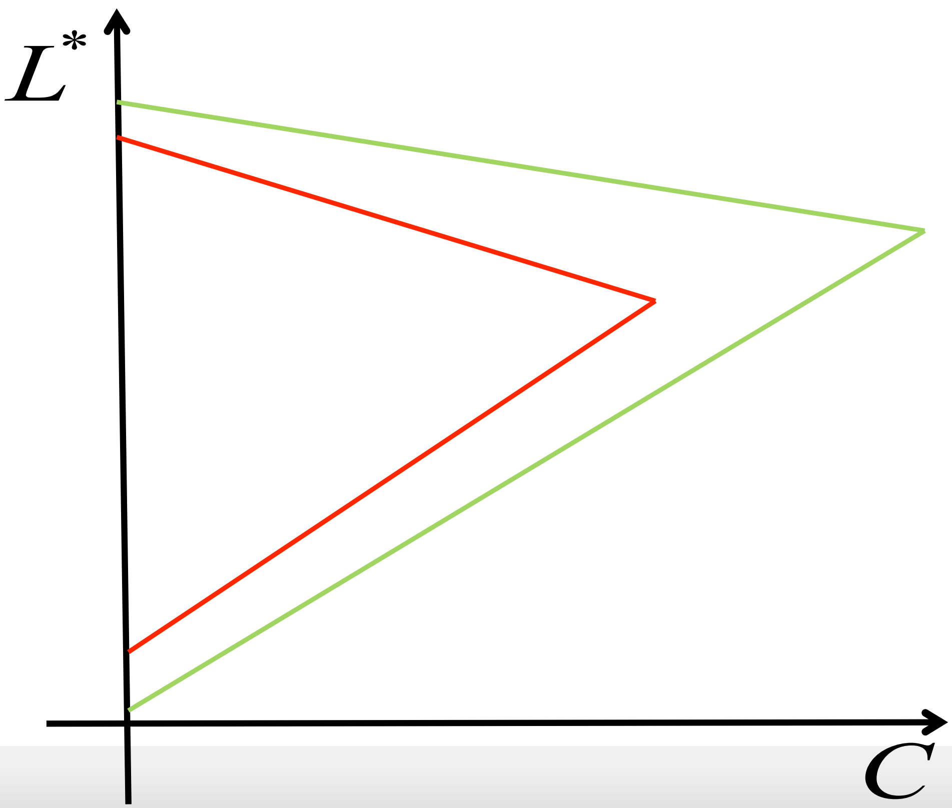
- Unchanged the Hue shift, will keep the overall image appearance

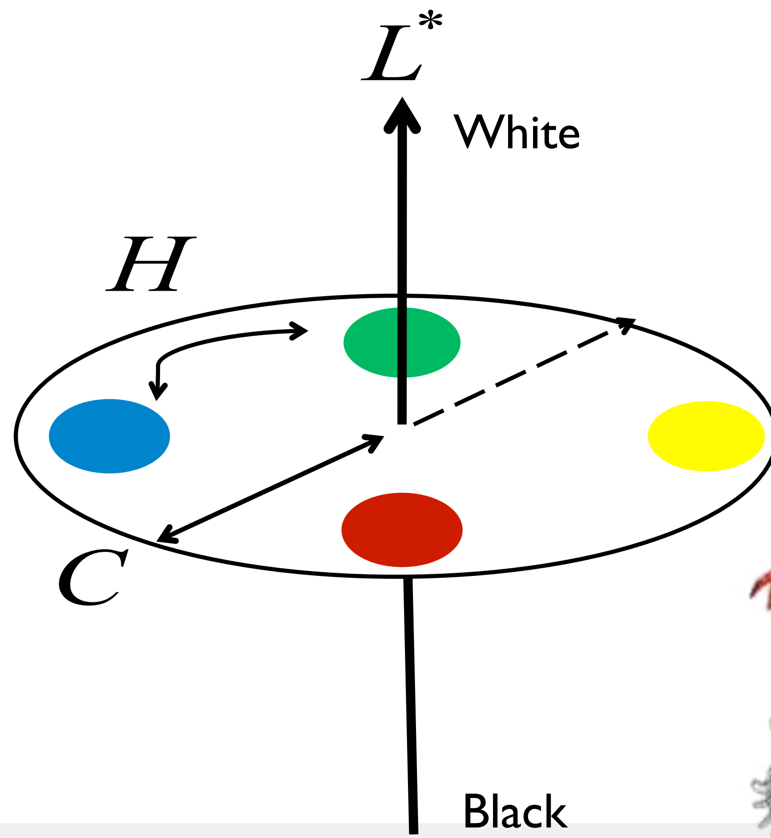


Gamut Mapping Aims (CS)

SIGGRAPHASIA2011 HONG KONG

- Unchanged the Hue shift, will keep the overall image appearance

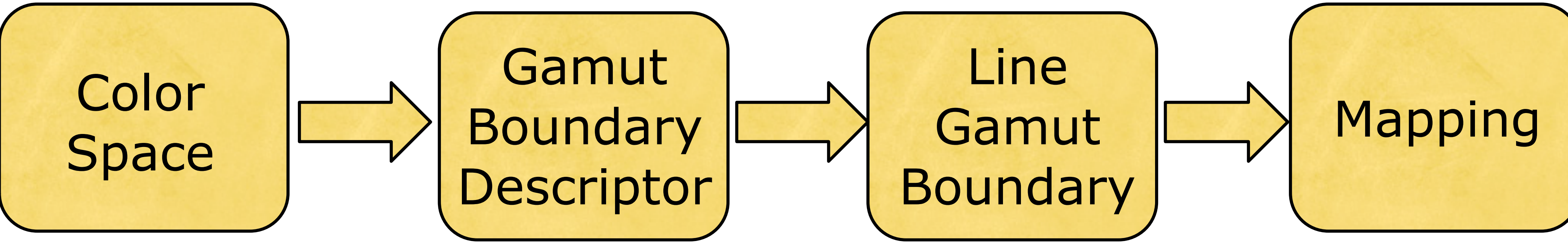




Gamut Mapping Aims (CS)

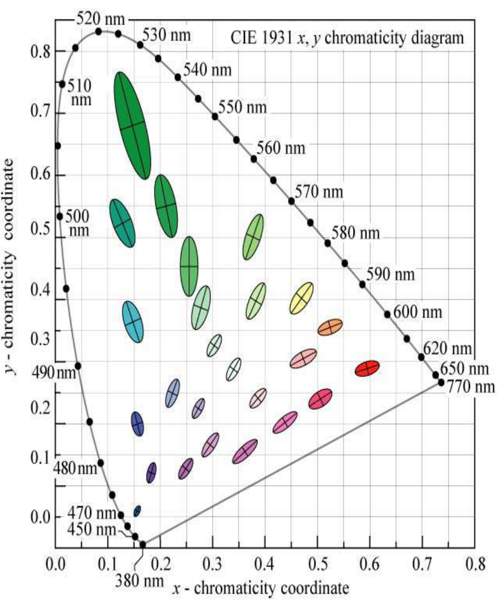
- Limiting out of gamut colours
  - Soft clipping can be afterwards adopted to eliminate these extremes
- Increase Image saturation
  - Destination gamut has reduced saturation
  - Helps maintaining the original chroma differences of the input Image

Gamut Mapping Pipeline



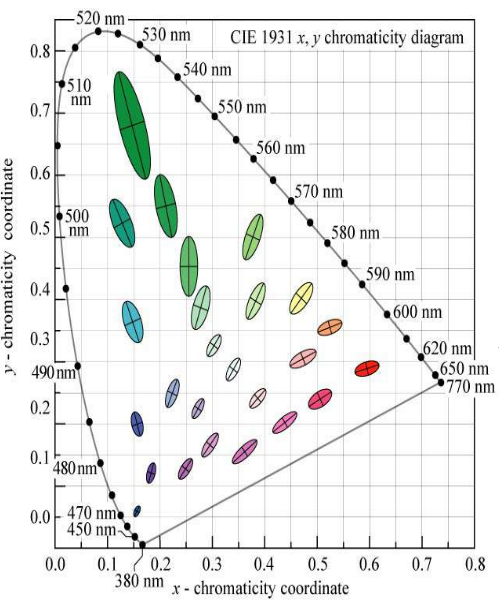
Gamut Mapping Pipeline

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

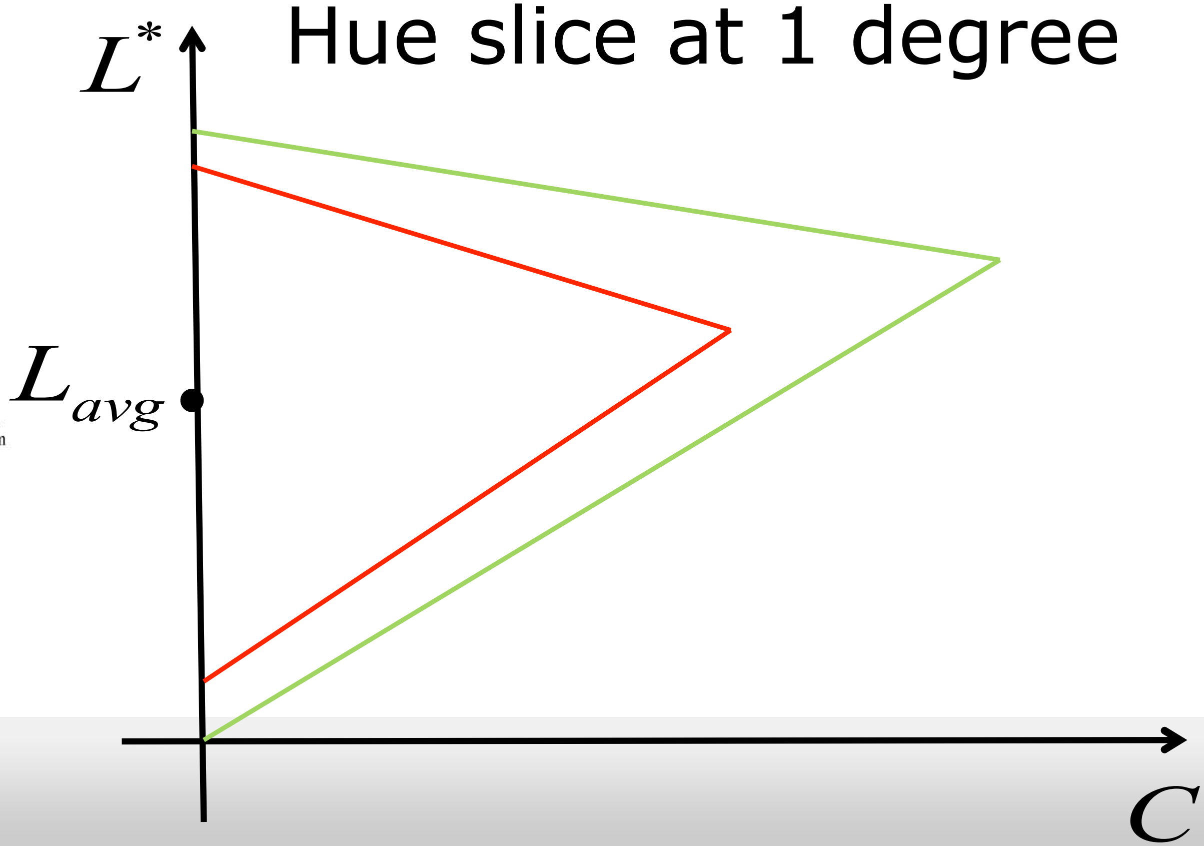


Gamut Mapping Pipeline

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

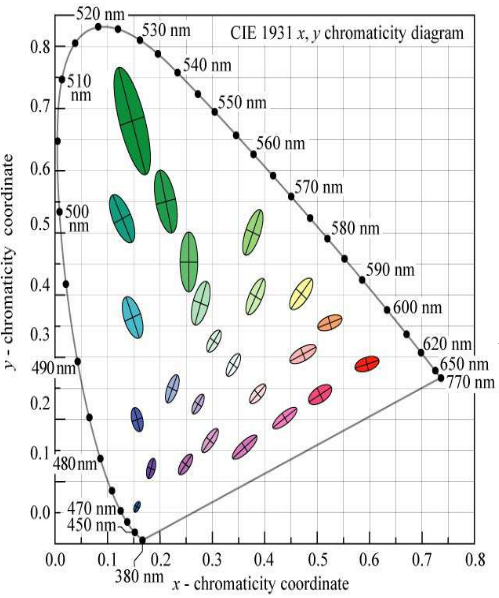


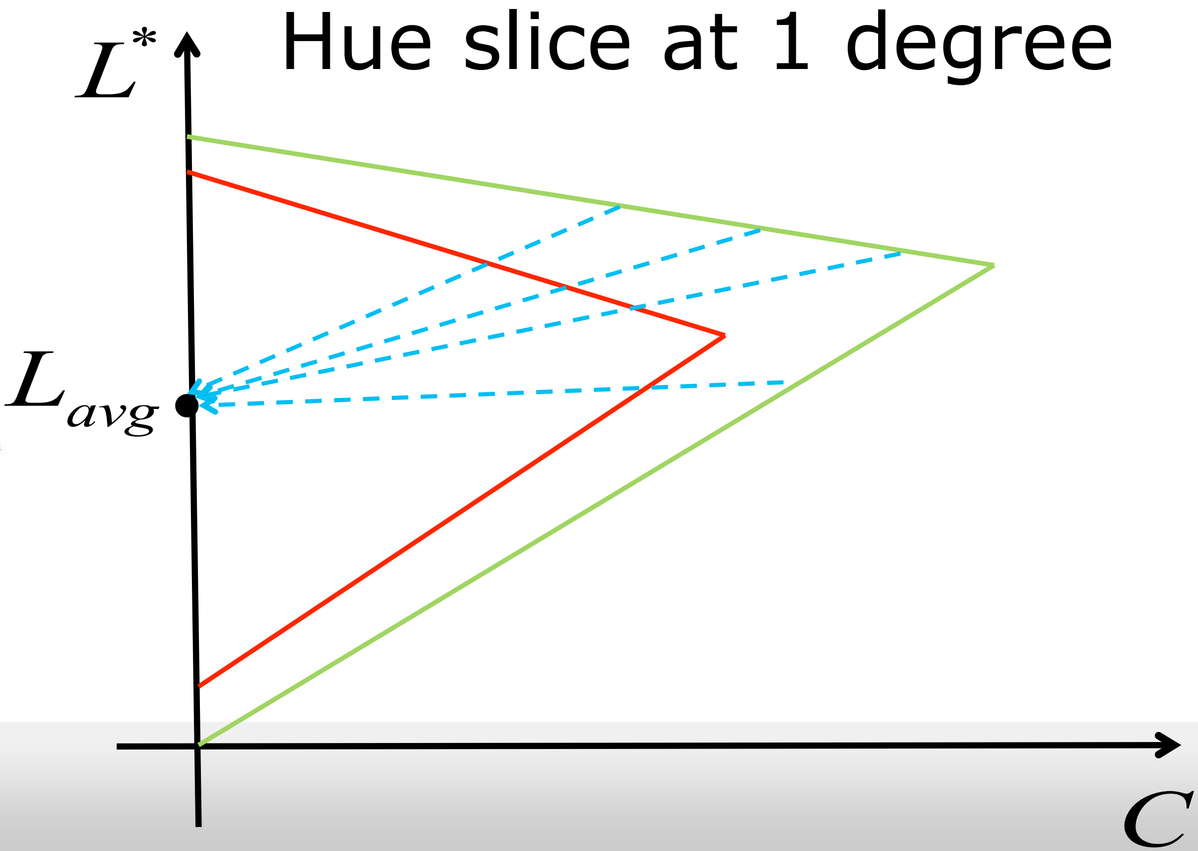
Hue slice at 1 degree



Gamut Mapping Pipeline

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

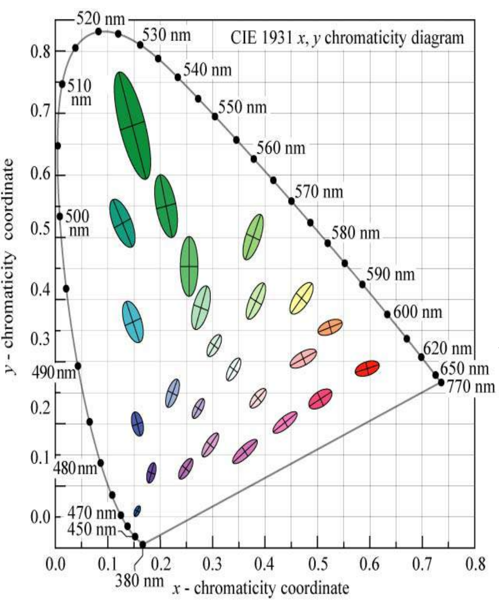


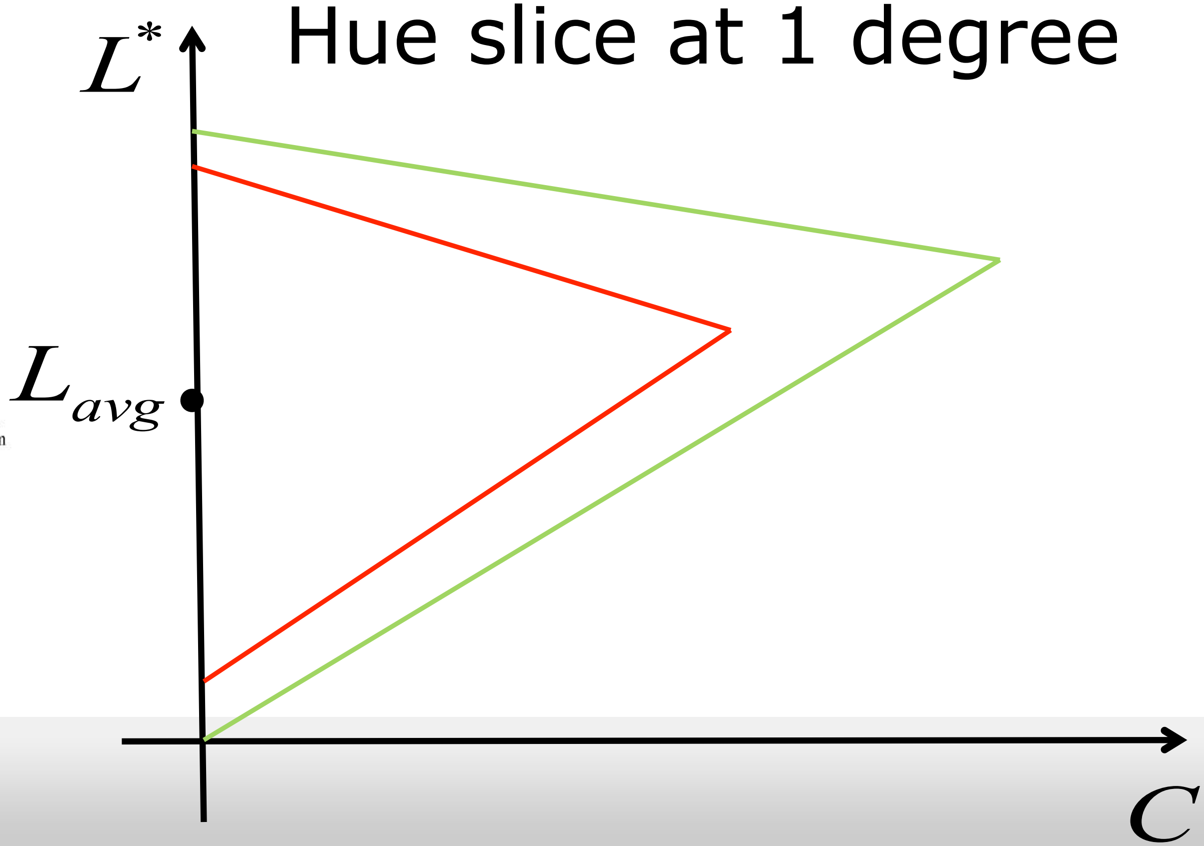


Gamut Mapping Pipeline

SIGGRAPHASIA2011 HONG KONG

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

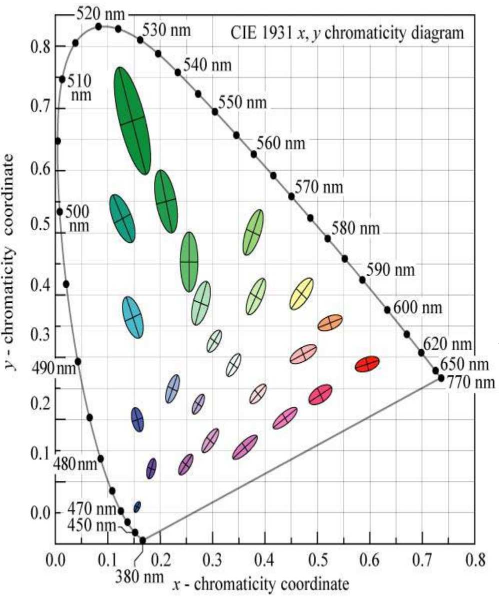


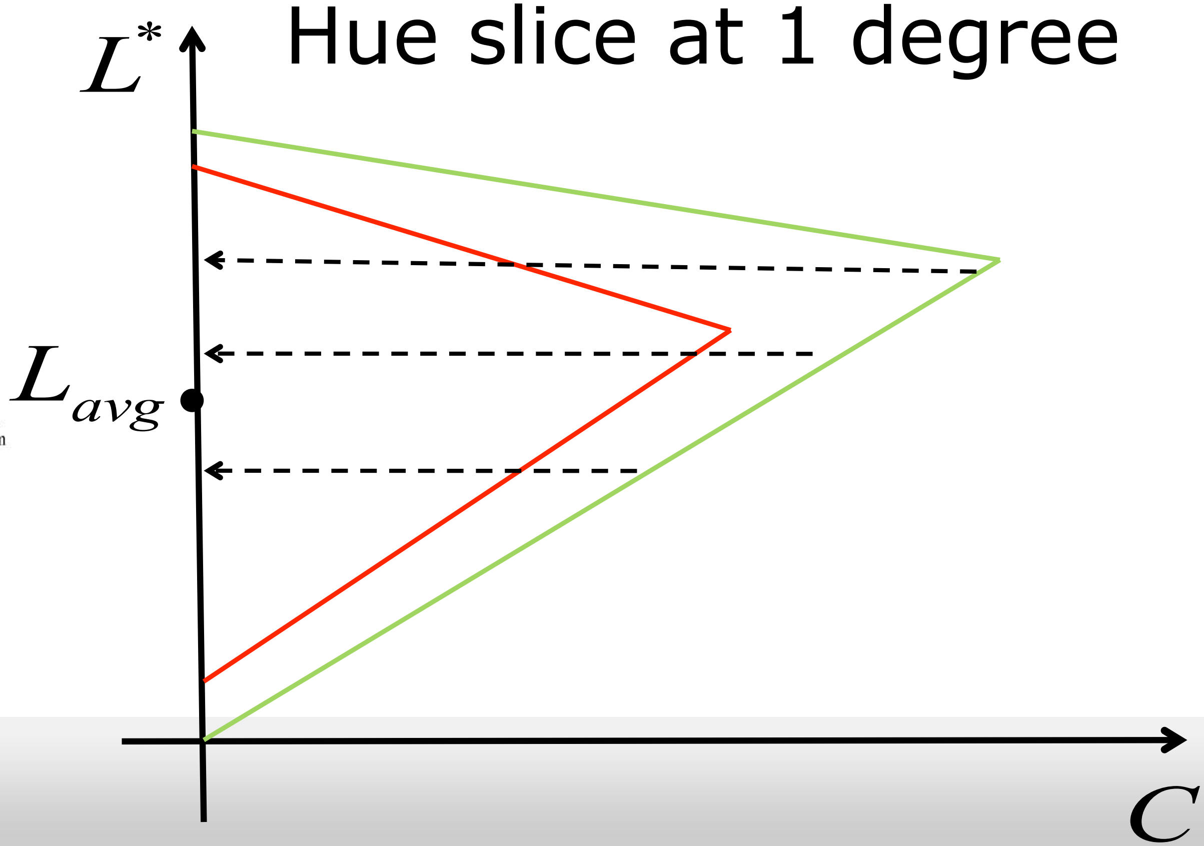


Gamut Mapping Pipeline

SIGGRAPHASIA2011 HONG KONG

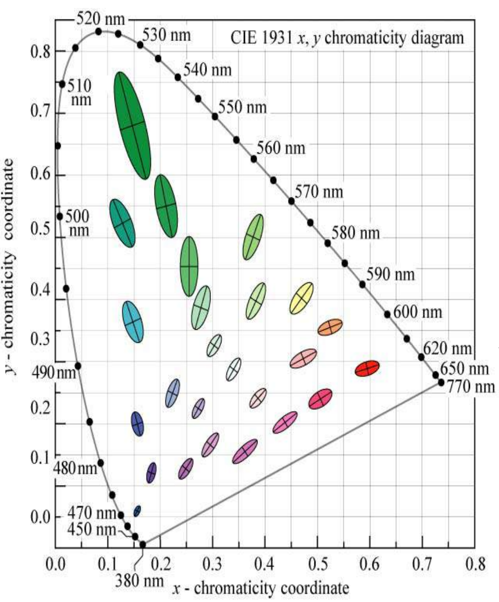
Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping



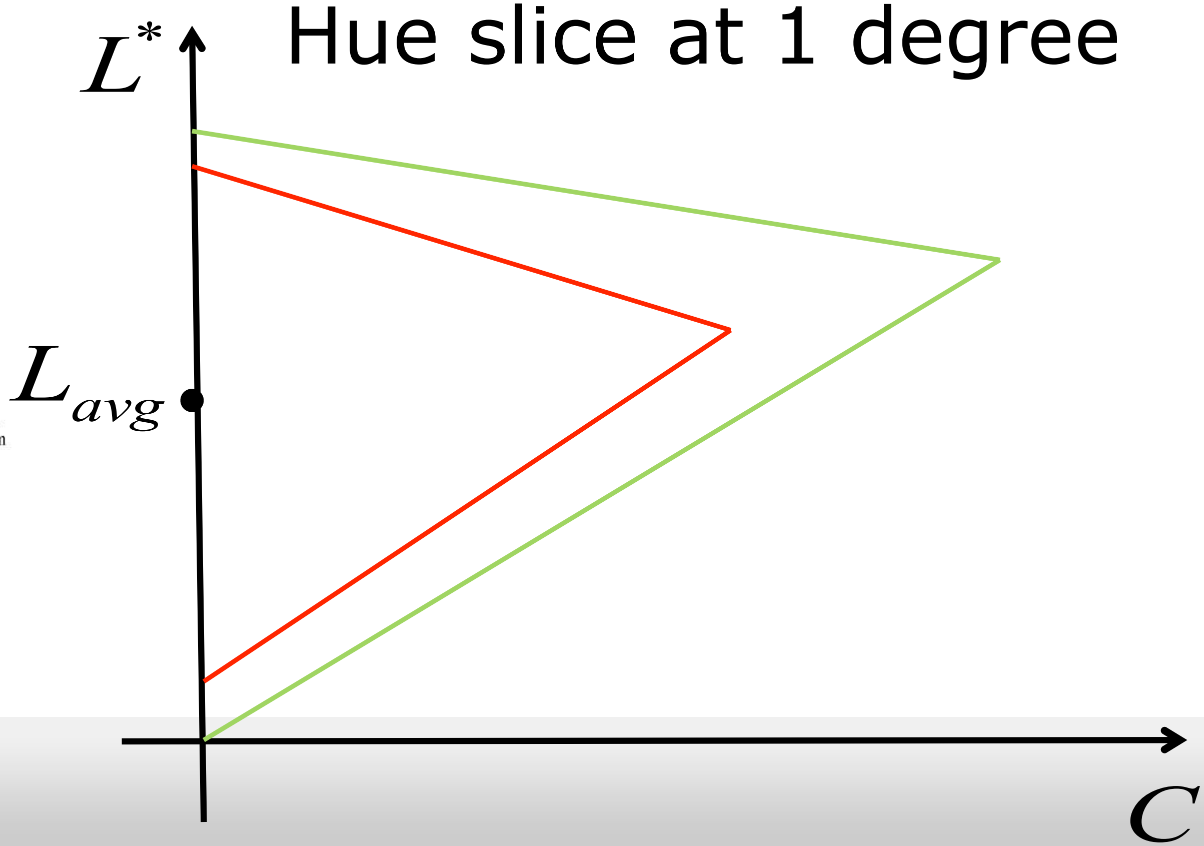


Gamut Mapping Pipeline

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping



Hue slice at 1 degree

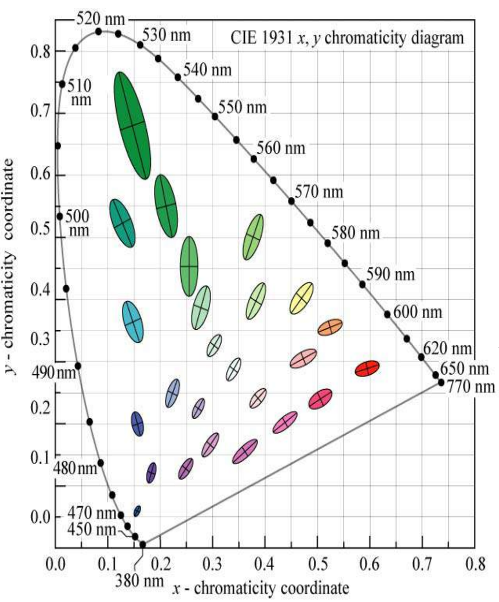


Gamut Mapping Pipeline

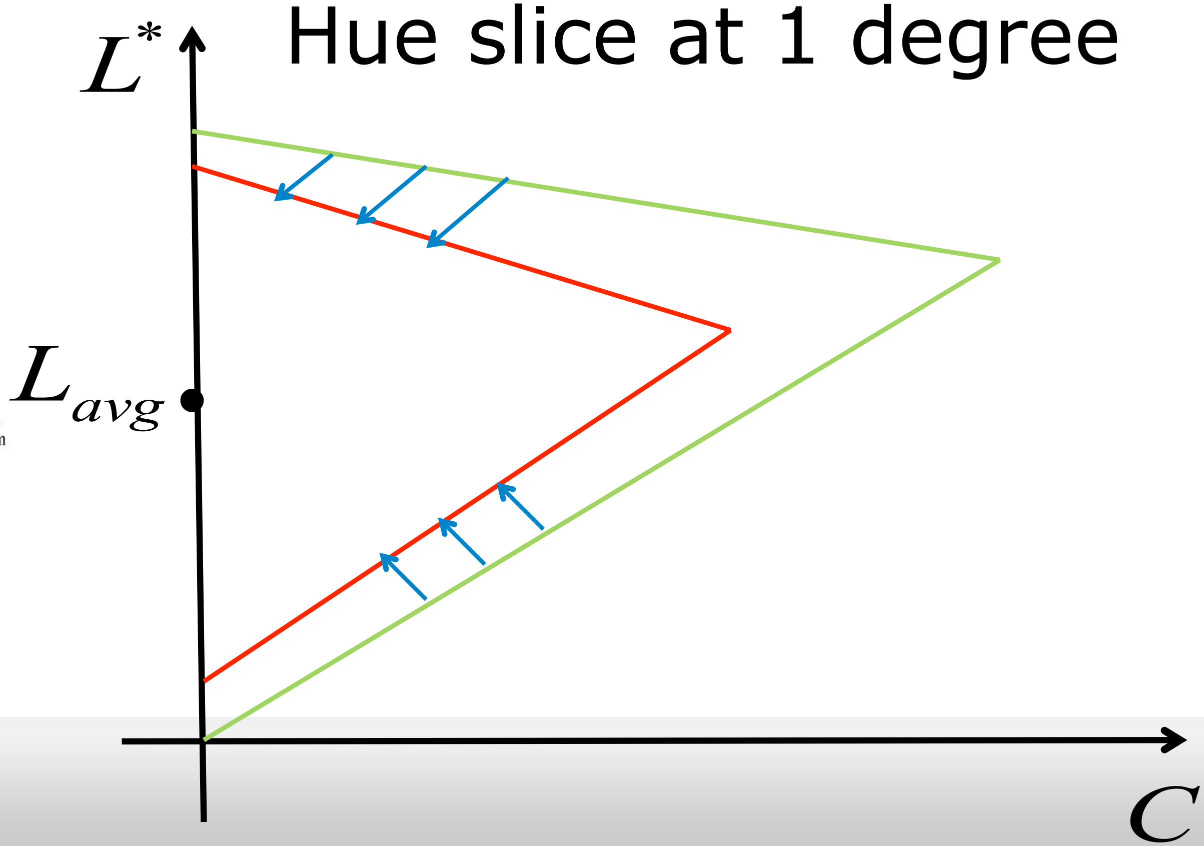
SIGGRAPHASIA2011 HONG KONG

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

Clipping



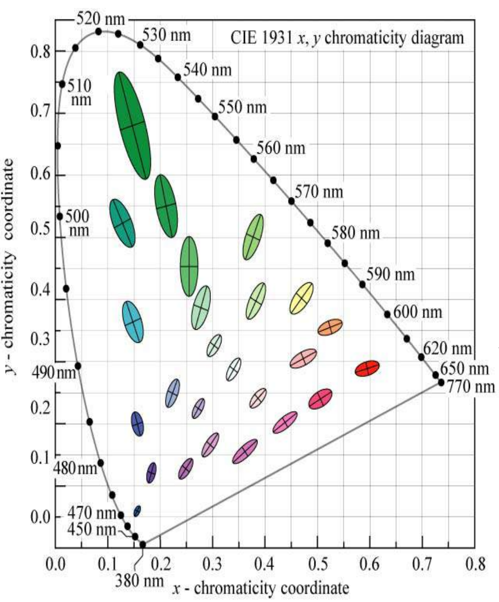
Hue slice at 1 degree



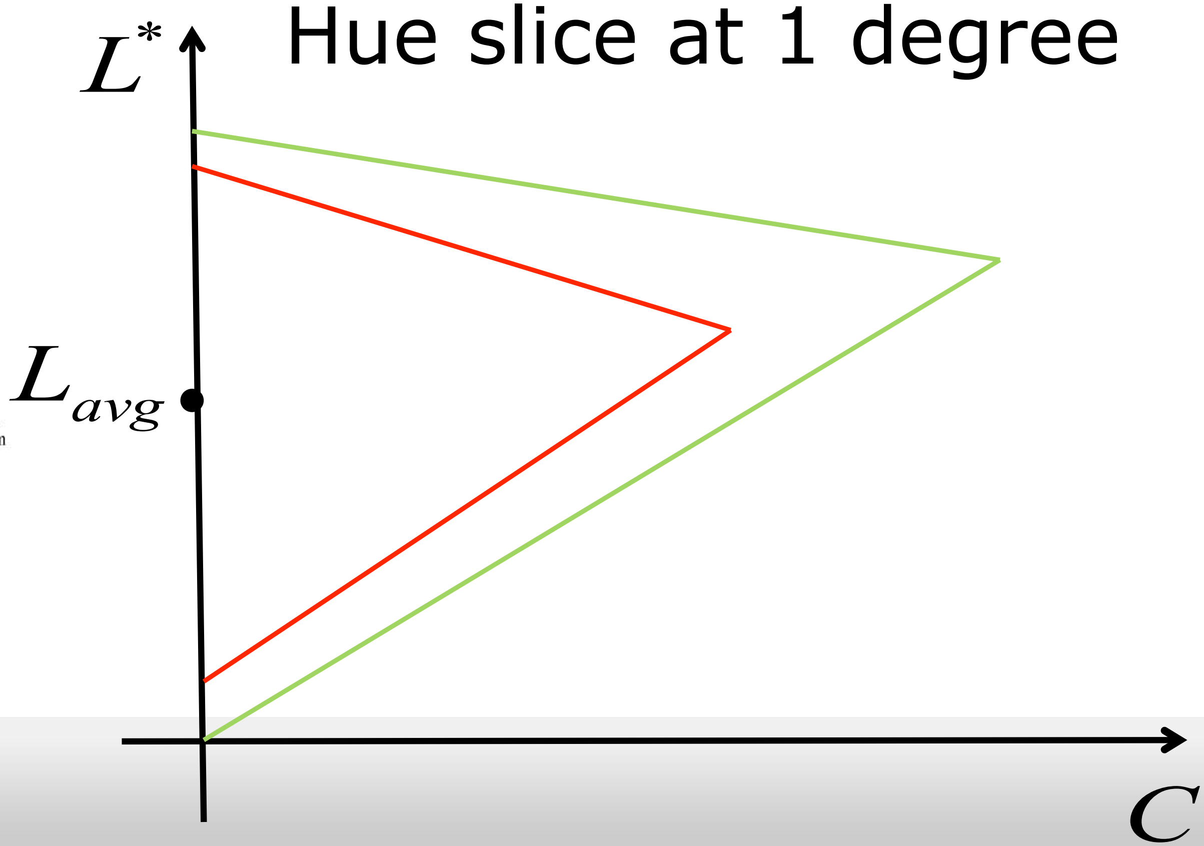
Gamut Mapping Pipeline

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

Clipping



Hue slice at 1 degree

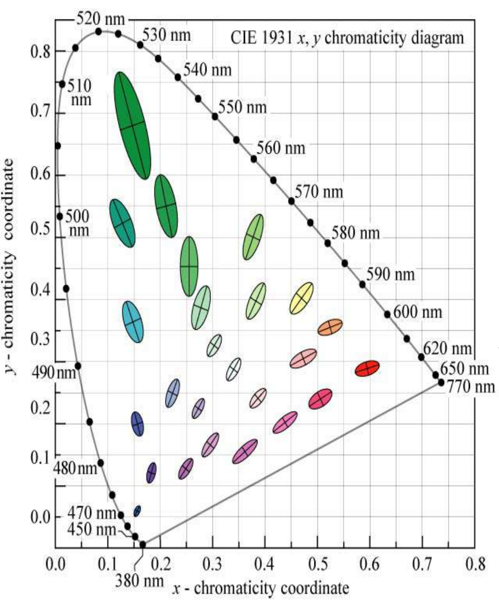


Gamut Mapping Pipeline

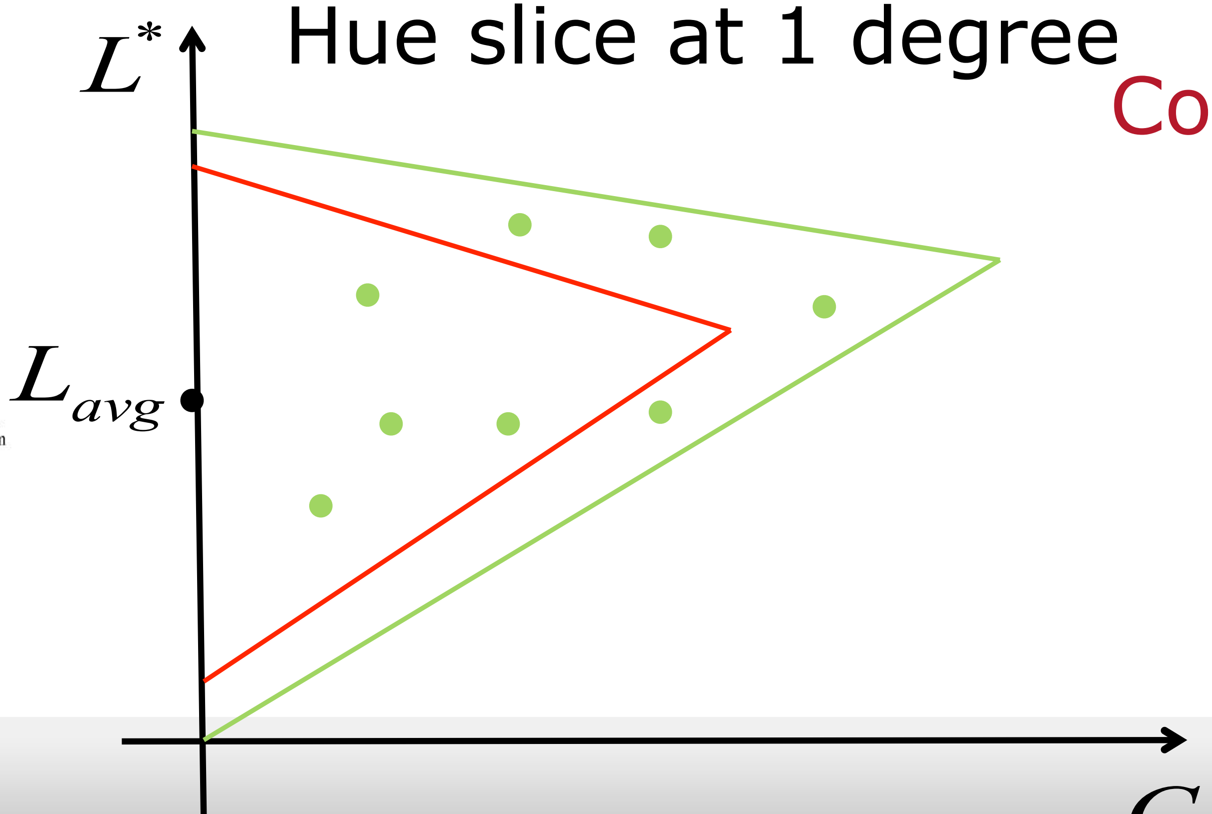
Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

Clipping

Compression



Hue slice at 1 degree

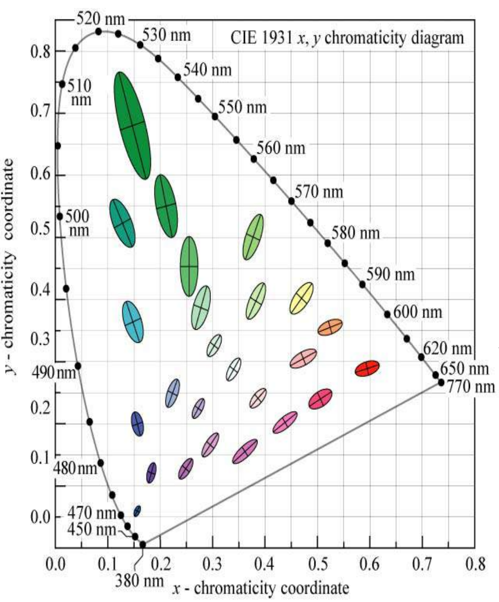


Gamut Mapping Pipeline

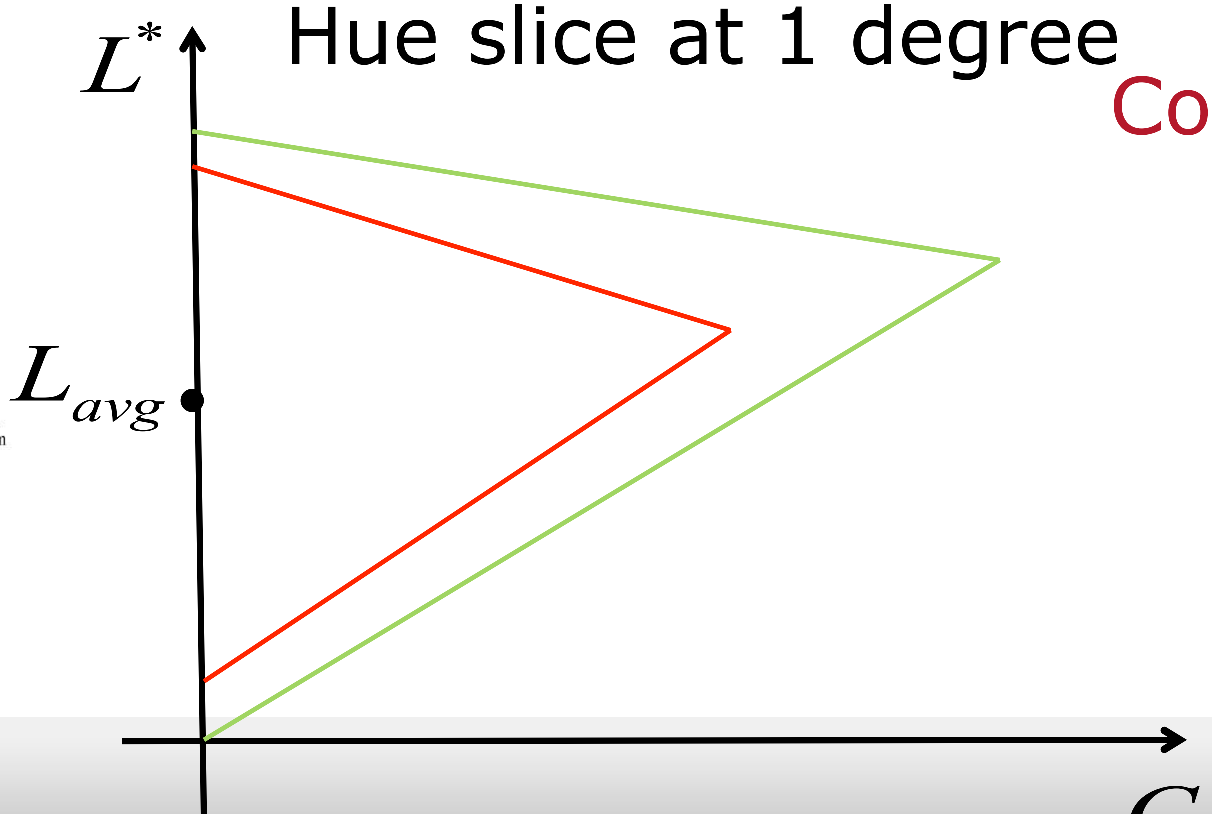
Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

Clipping

Compression



Hue slice at 1 degree

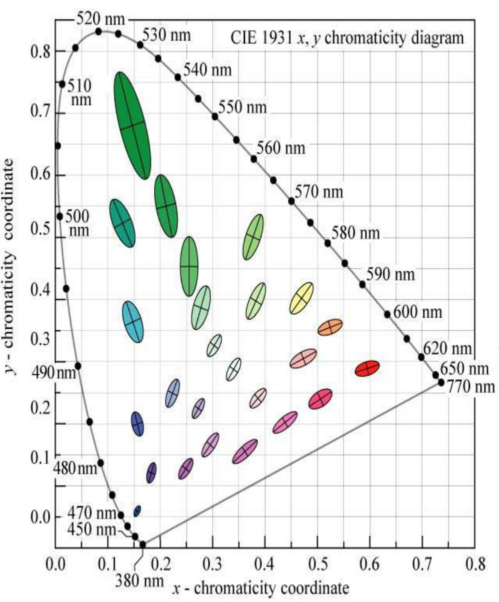


Gamut Mapping Pipeline

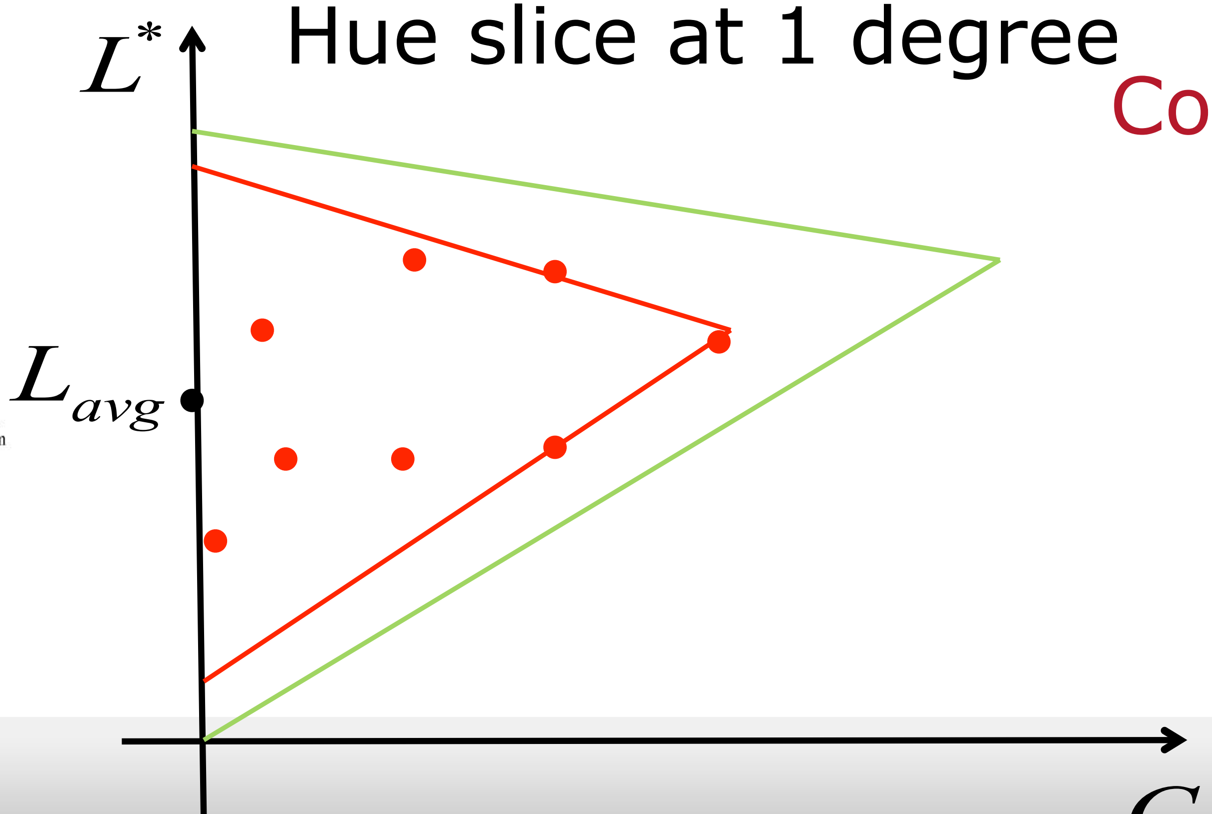
Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping

Clipping

Compression

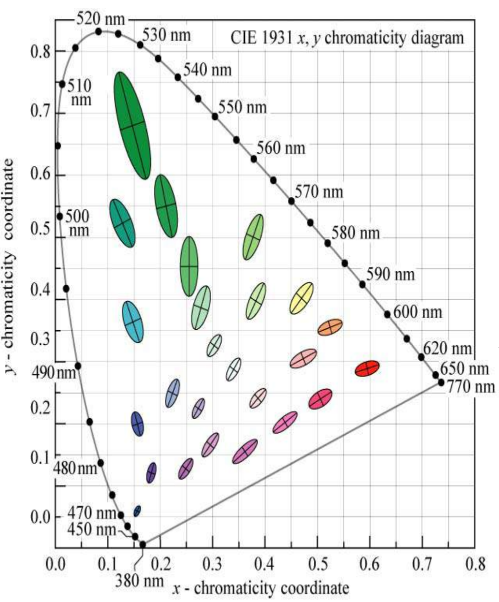


Hue slice at 1 degree

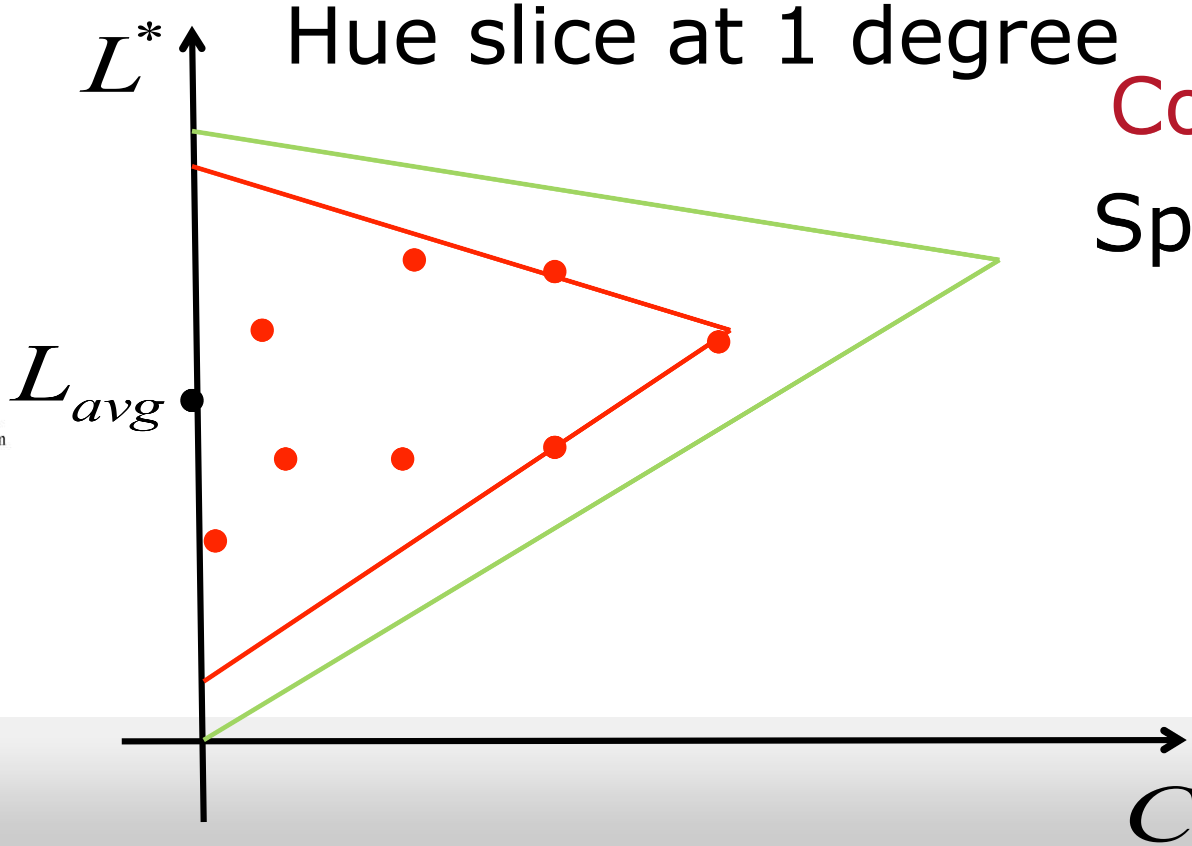


Gamut Mapping Pipeline

Color Space → Gamut Boundary Descriptor → Line Gamut Boundary → Mapping



Hue slice at 1 degree



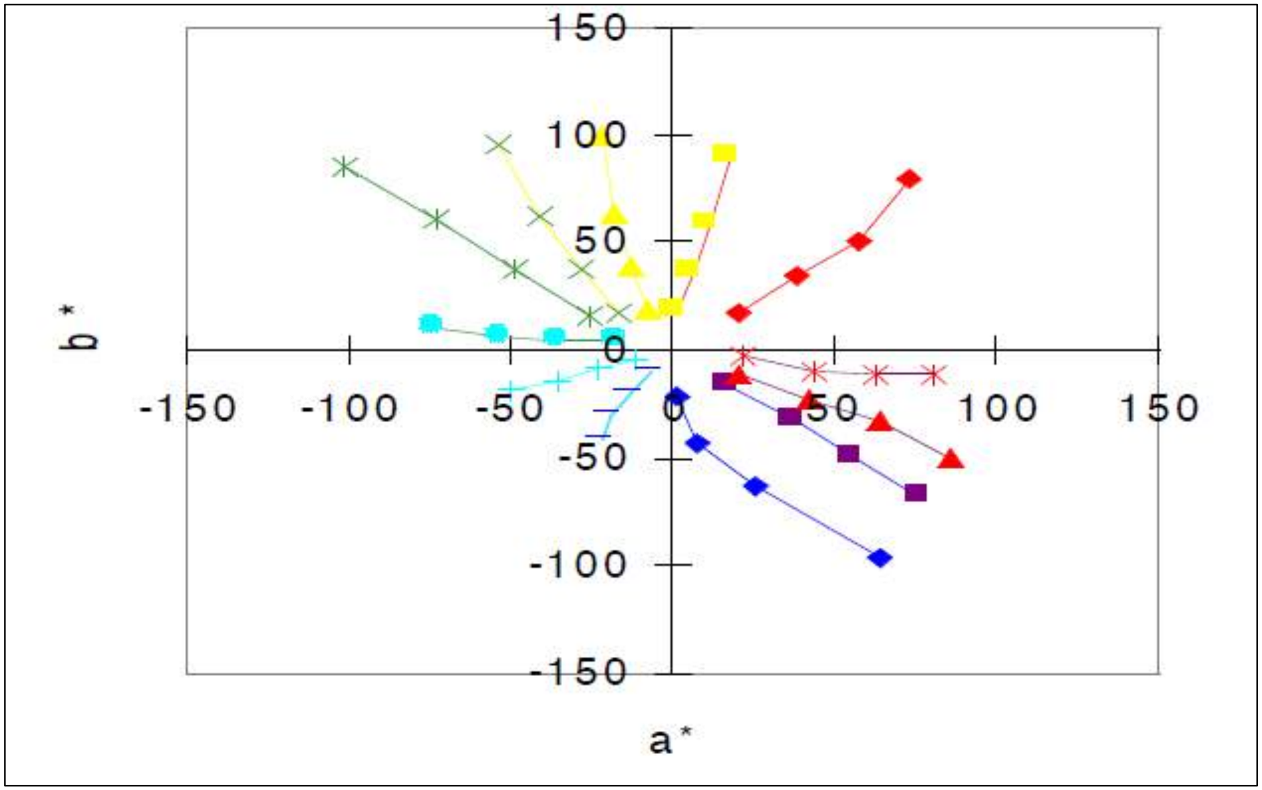
Clipping

Compression

Spatial GMAs

Color Space Issue

SIGGRAPHASIA2011 HONG KONG



- Gamut Mapping that preserves metric hue angle
  - No Hue shift after compression or clipping
- CIELab is suffering of non linearity in blue regions, but also in red regions

Braun and Fairchild. “Color Gamut Mapping in Hue-Linearized CIELab Color Space”

Point-wise Gamut Mapping Techniques

- **Clipping**

It changes colours which are outside of the destination gamut, mapping them on the boundaries of the destination gamut

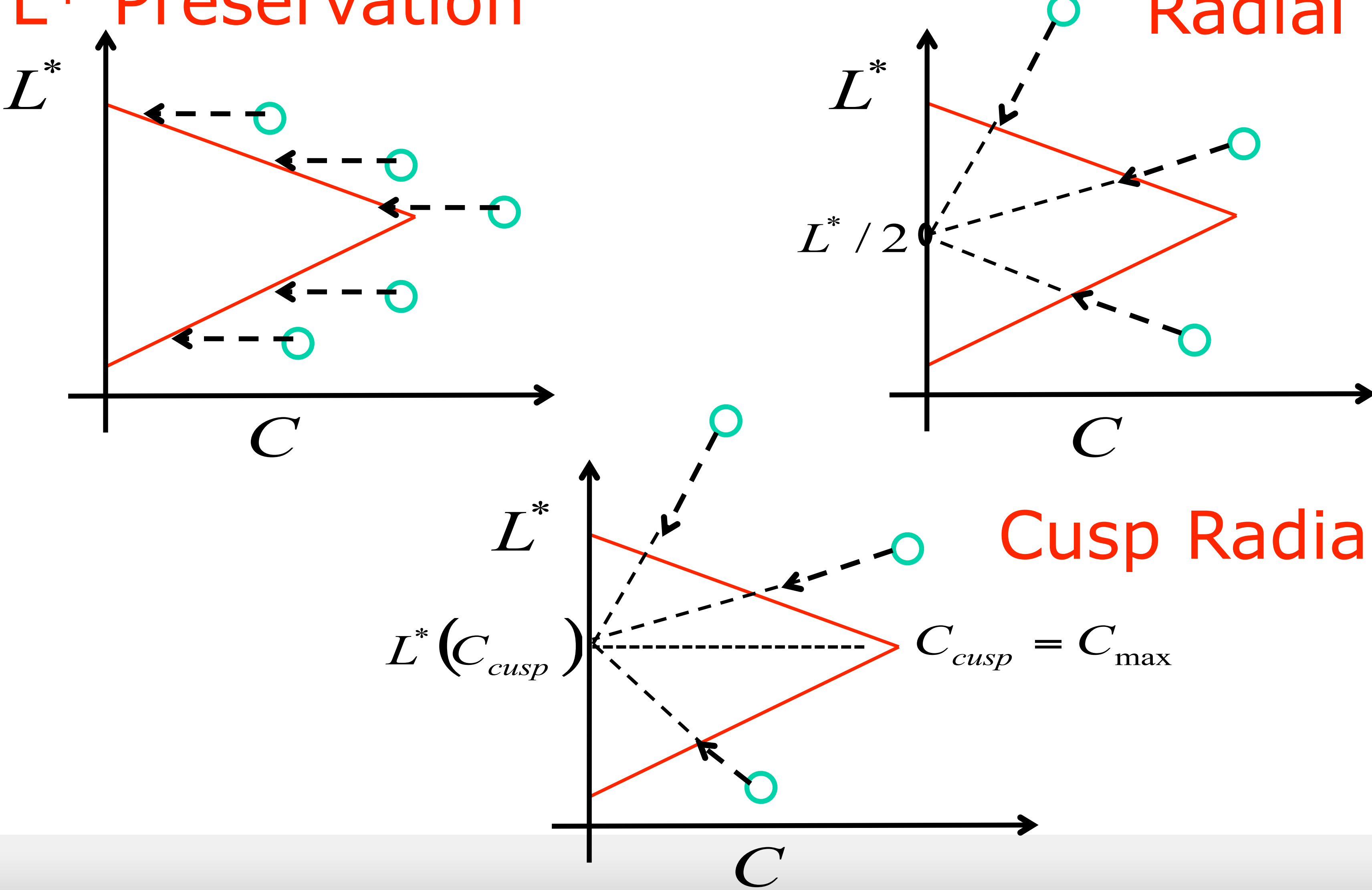
- **Horizontal** (lines of constant lightness)
- **Radial to a centre of Gravity**
  - Centre of lightness axis (Constant)
  - Lightness corresponding to the Chroma Cusp (variable)
- **Distance in CIELab**
  - To the colour boundary of the destination gamut that has the smallest distance (**HPMin**ΔE **Clipping**)

Clipping

L* Preservation

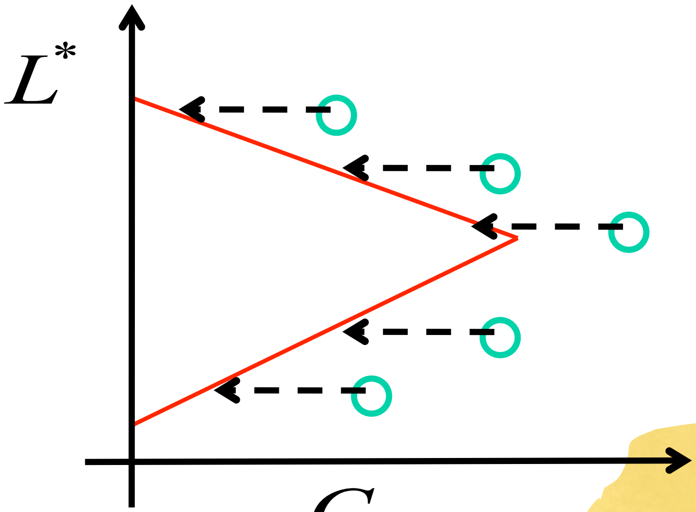
Radial to L*/2

Cusp Radial

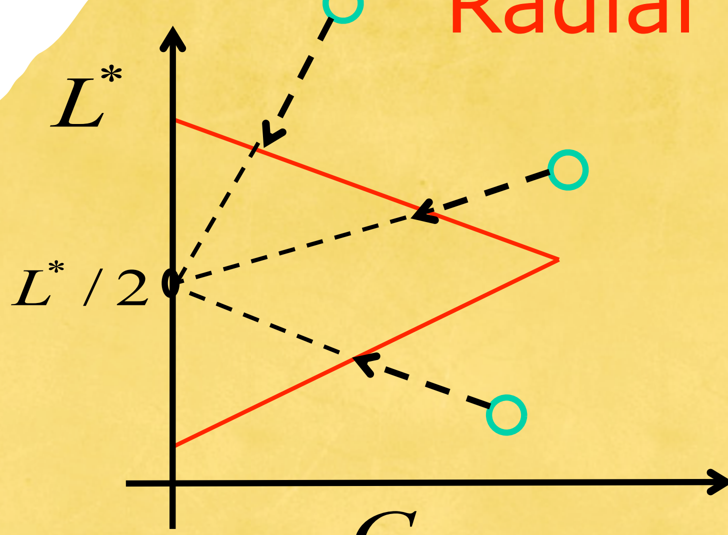


Clipping

L* Preservation

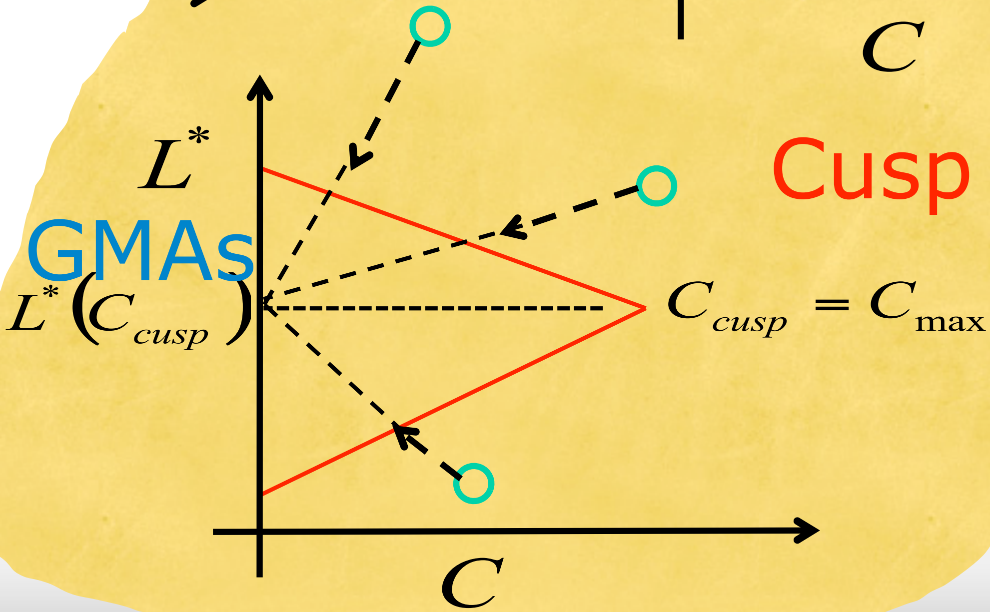


Radial to L*/2



Cusp Radial

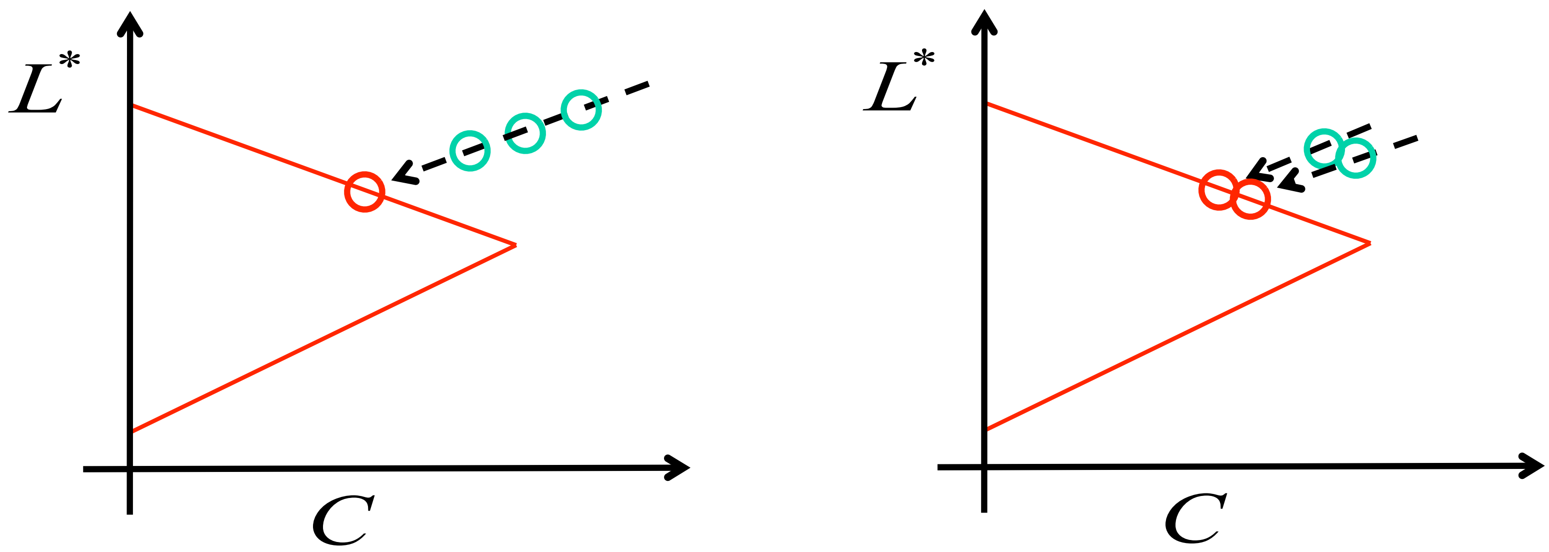
Simultaneous GMAs



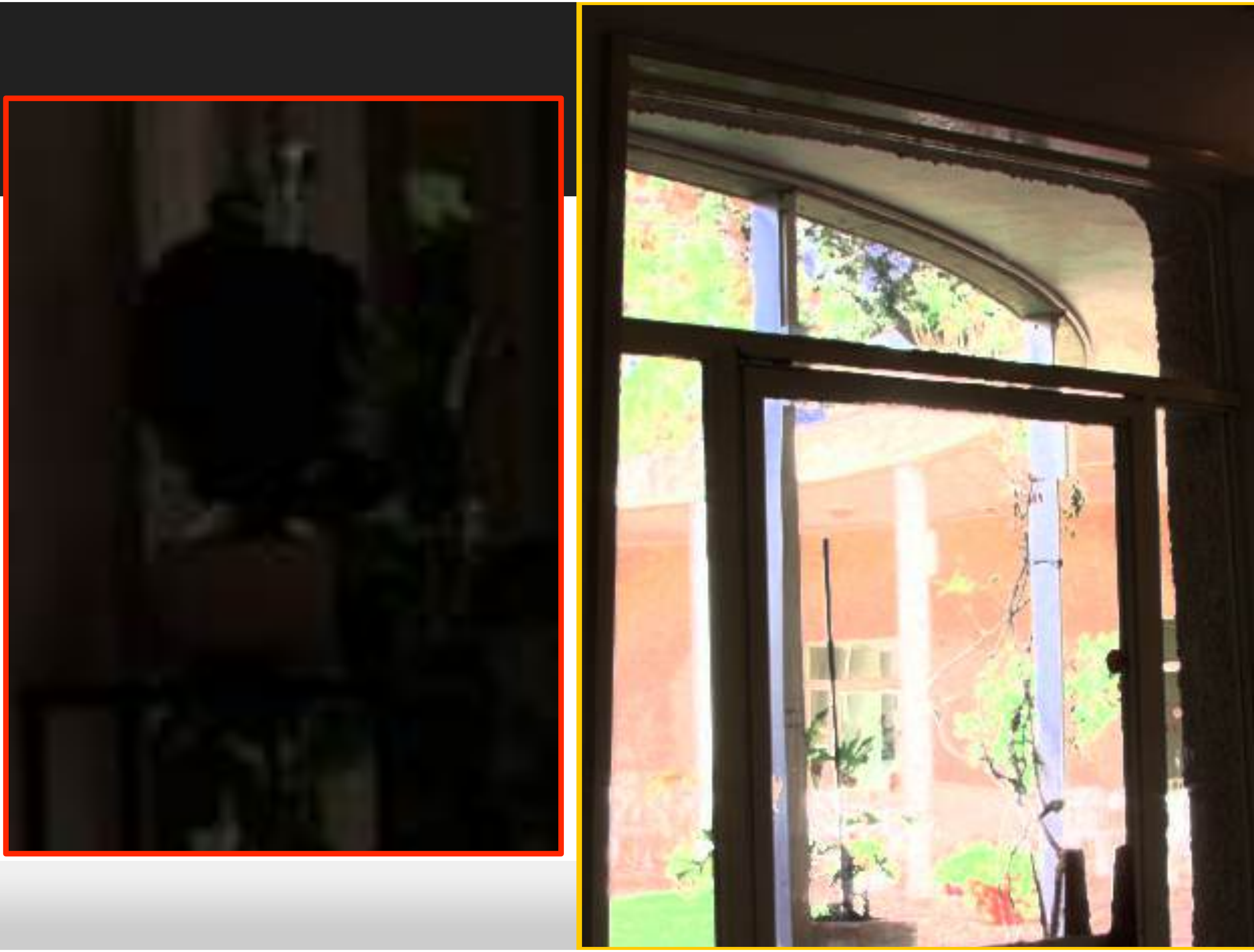
Clipping – Major Drawbacks

SIGGRAPHASIA2011 HONG KONG

Erase Local Image variation (Details)



Preserve Saturation



Point-wise Gamut Mapping Techniques

- **Compression**

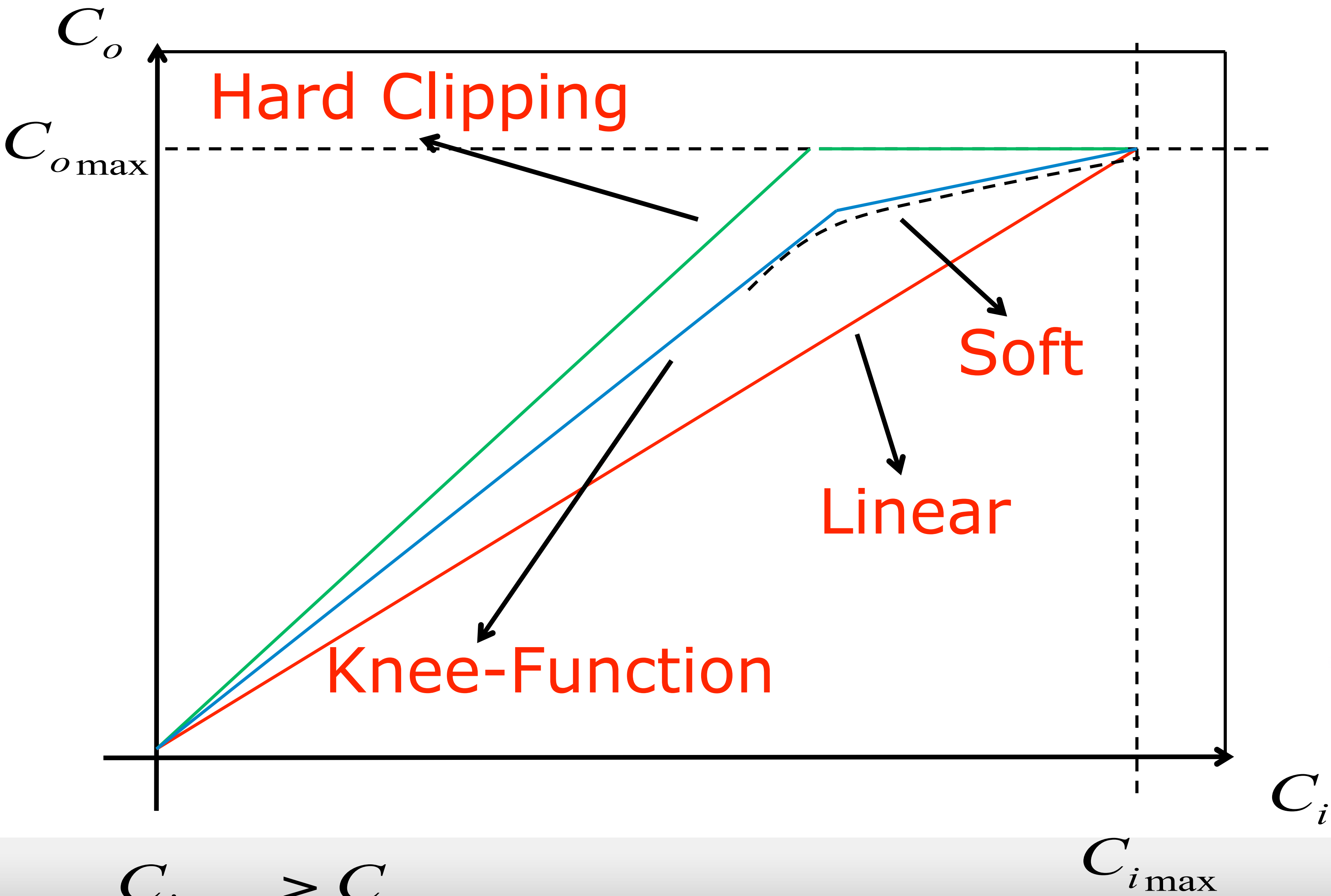
It makes changes to all the colors of the source gamut to be accommodated into the destination gamut .

  - **Linear**
  - **Sigmoid**
  - **Knee-function**

- **Parametric**

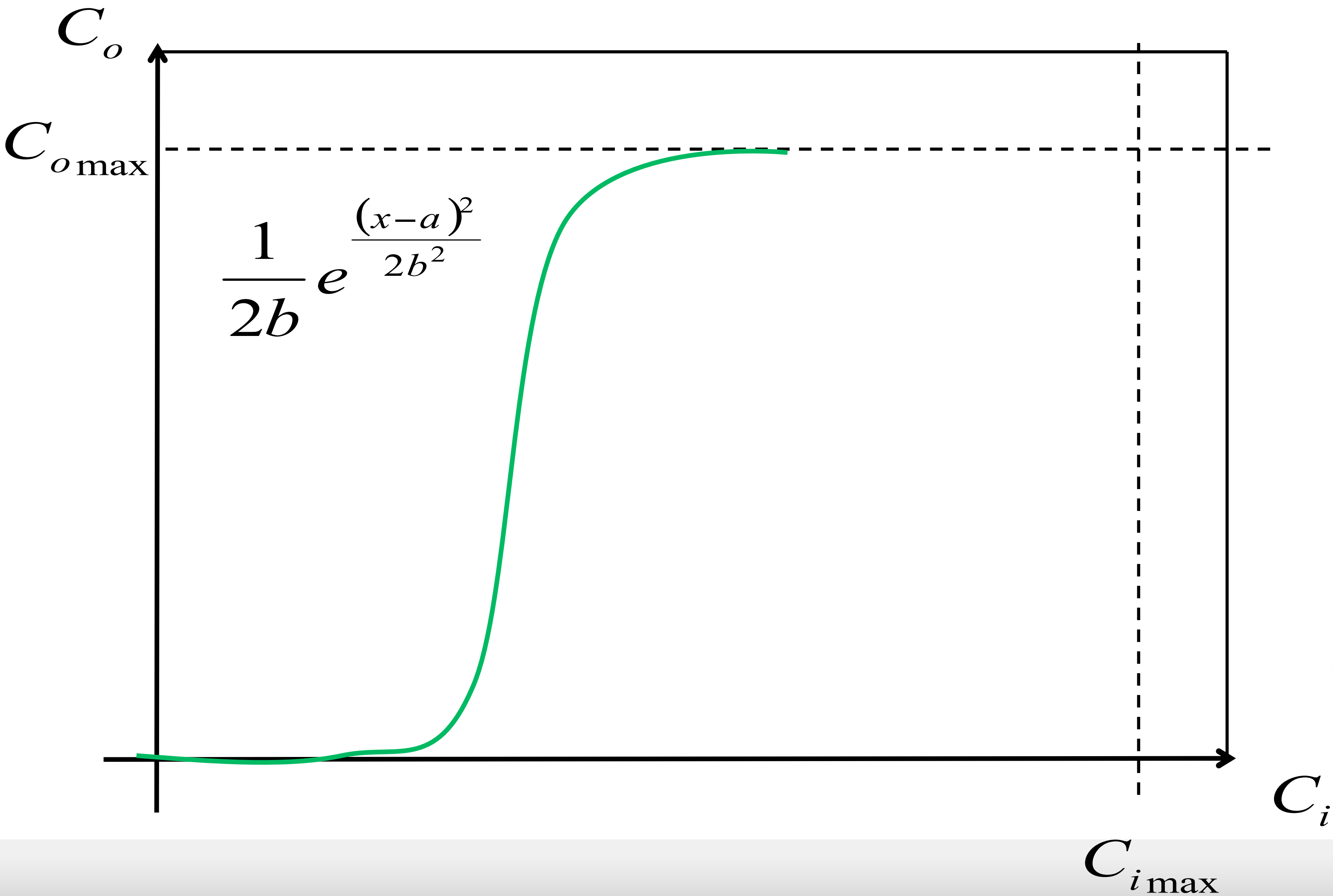
The behaviour change based on the shapes of the two gamut's (source and destination) at the hue angle, or it depends from user parameters. (Clipping and Compression)

Compression

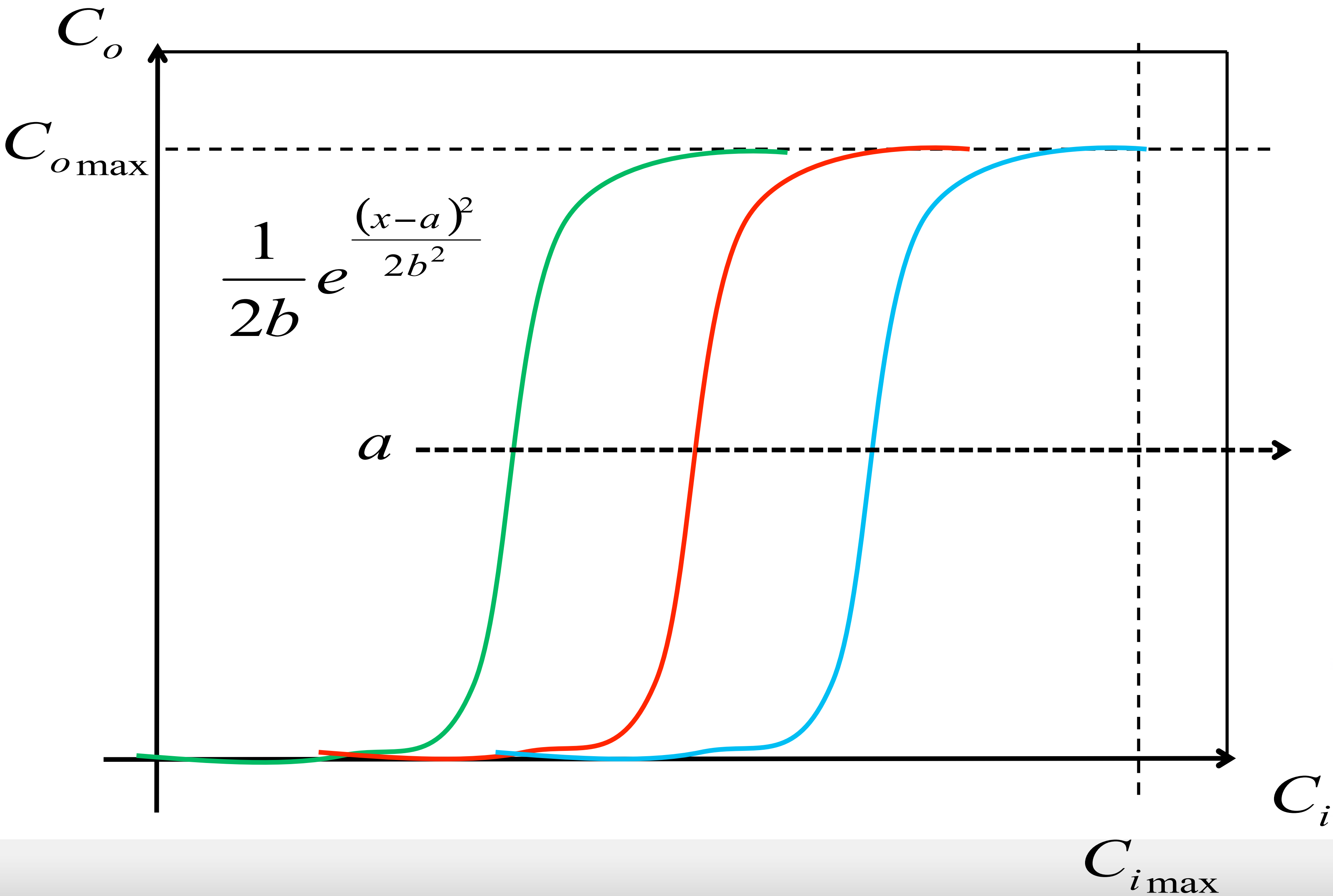


$C_{i\max} > C_{o\max}$

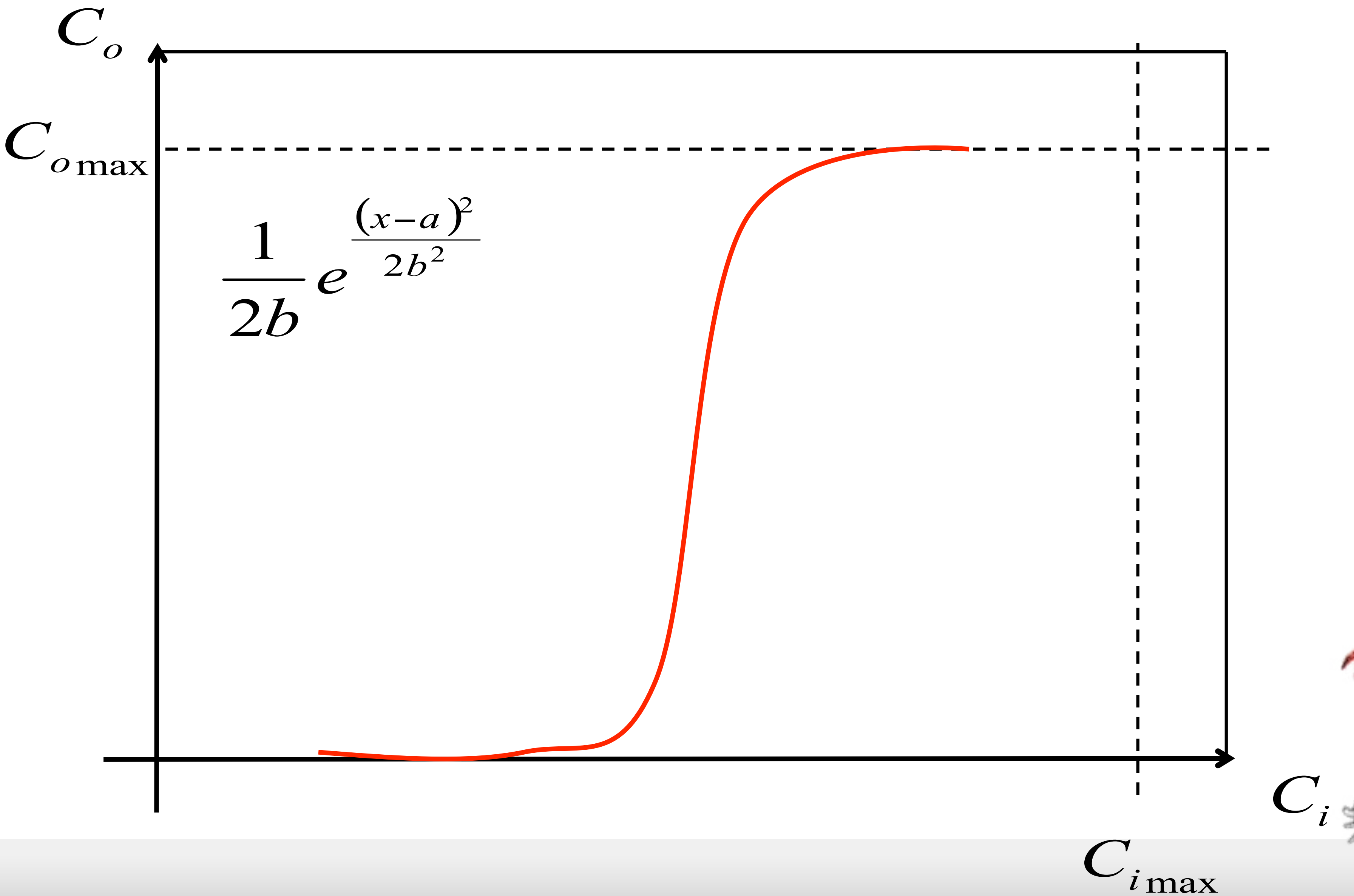
Compression



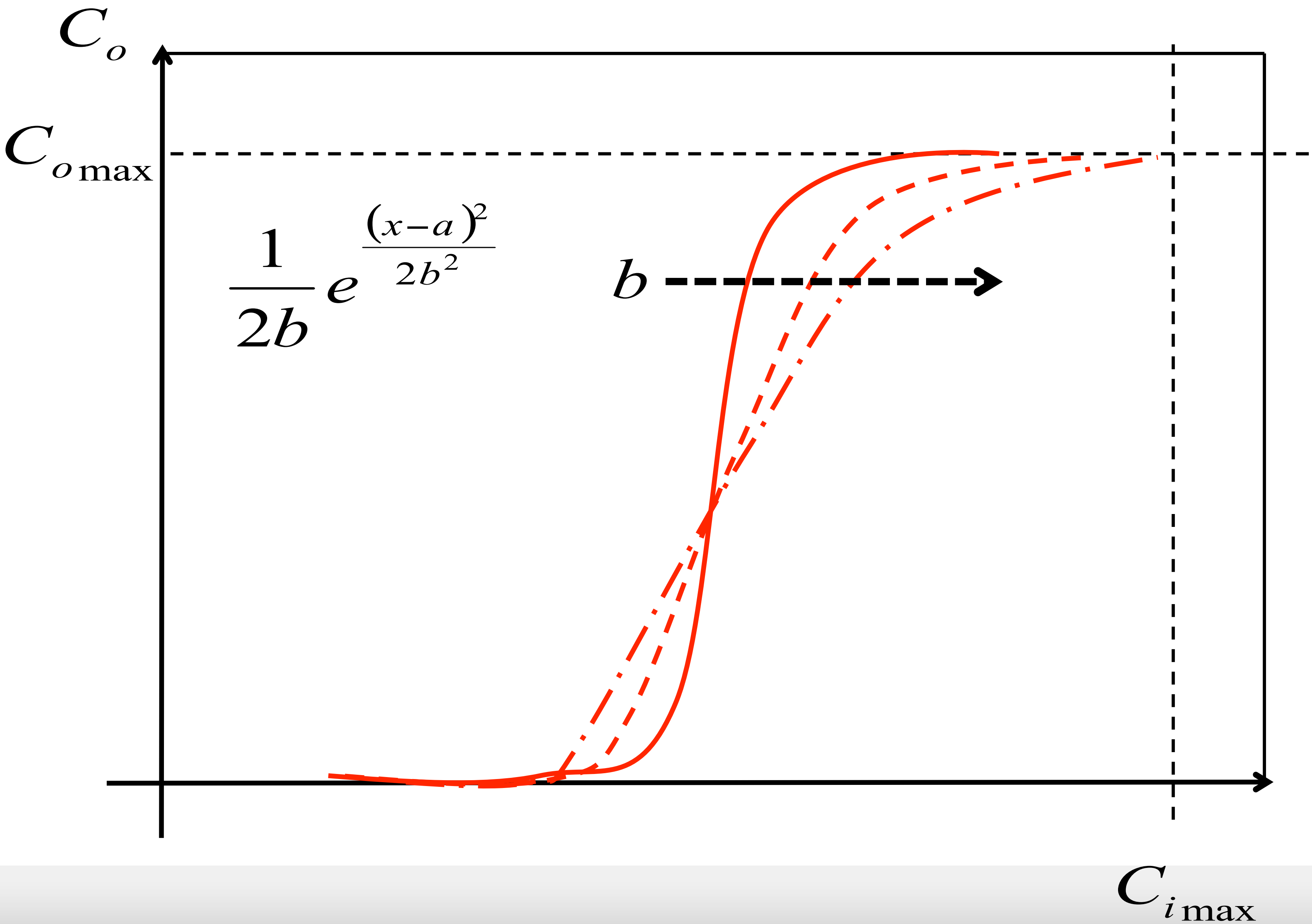
Compression



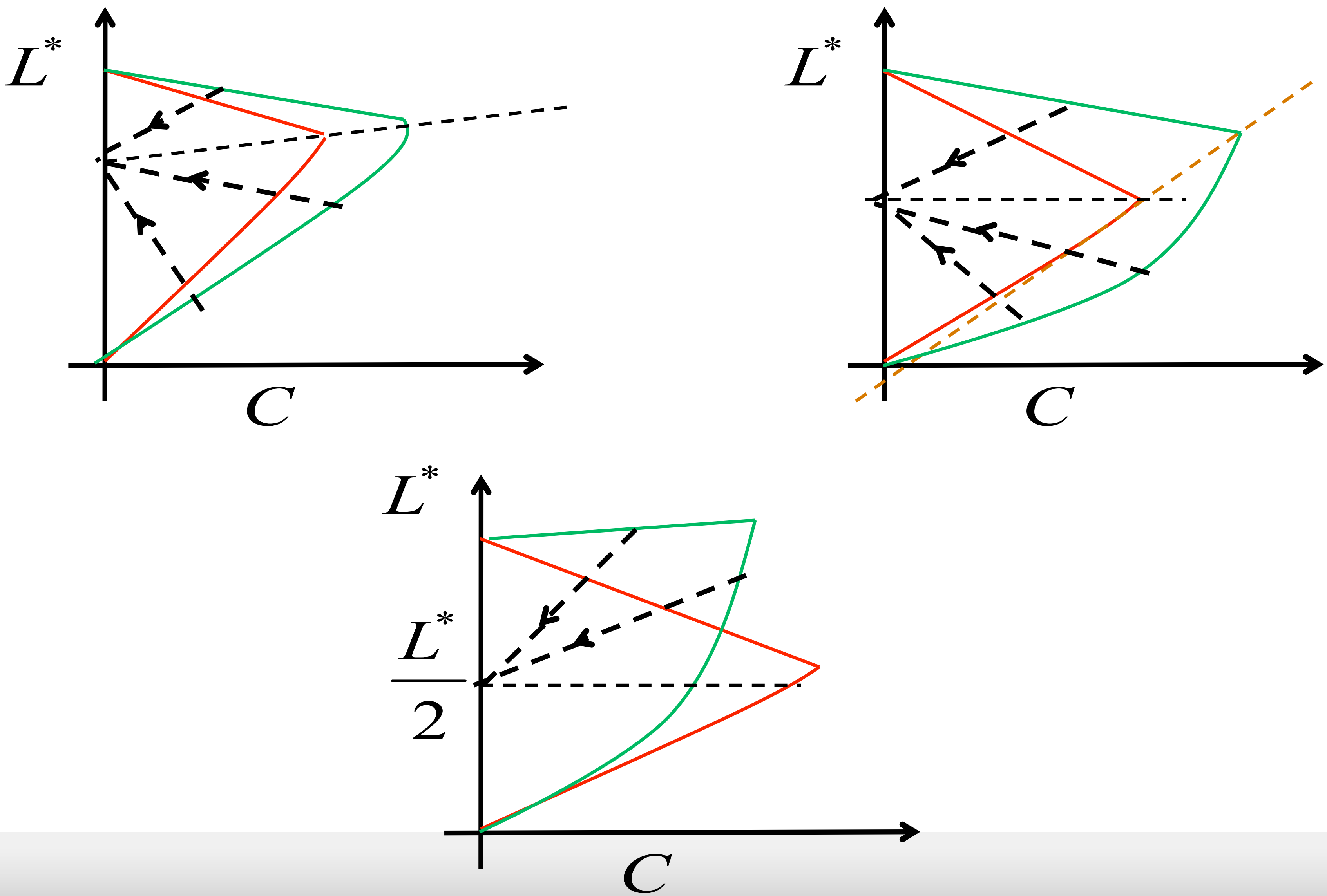
Compression



Compression



Parametric



Preservation of Spatial Details

- **Optimization**

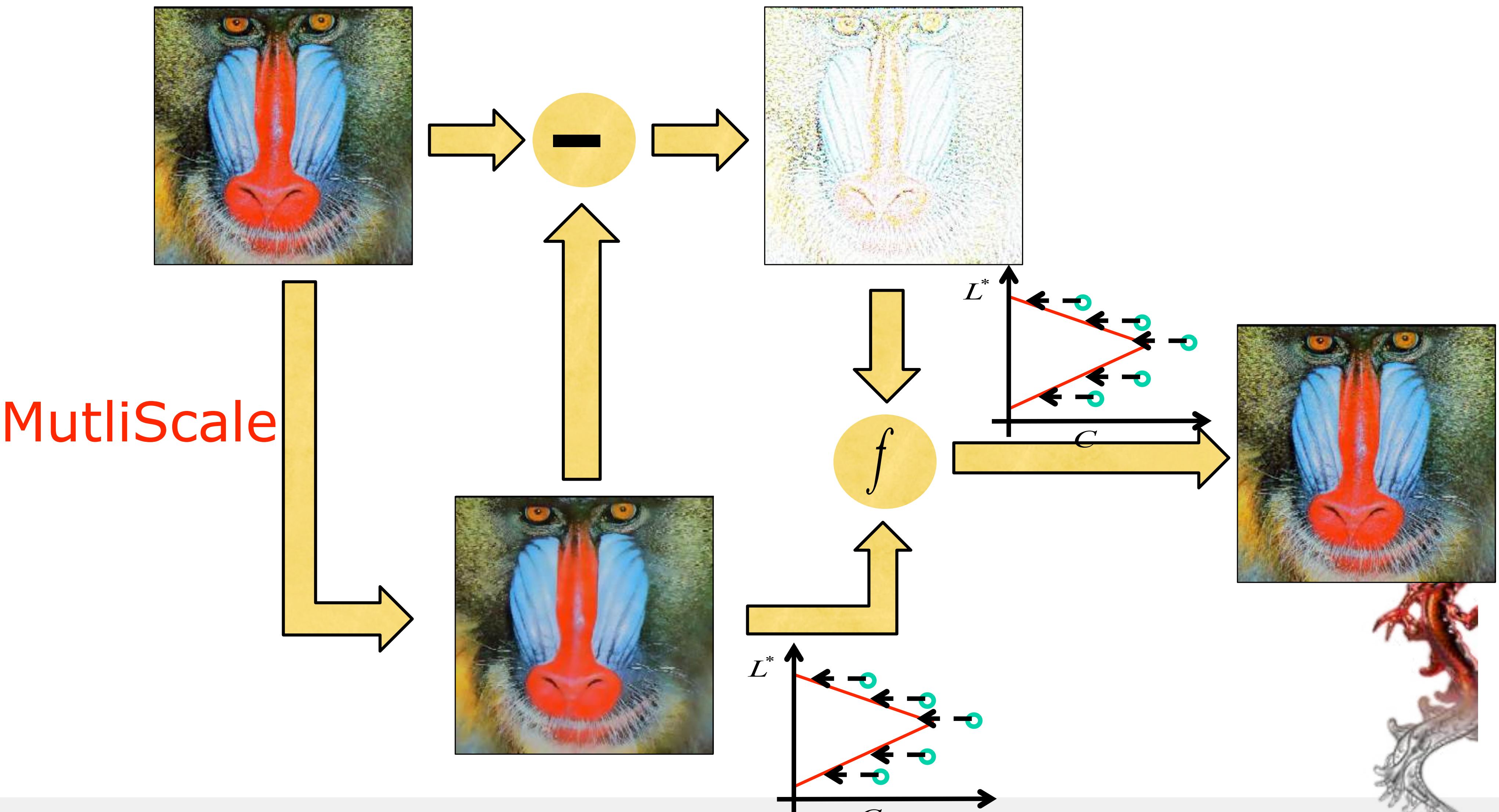
Making use of Human Visual System Models minimize the perceived differences between the input and output image.

- **Multiscale**

Re-inserts high-frequency information content in the gamut mapped image (clipped).

  - Clipping – loss of details
  - General framework has been proposed that includes the different cases

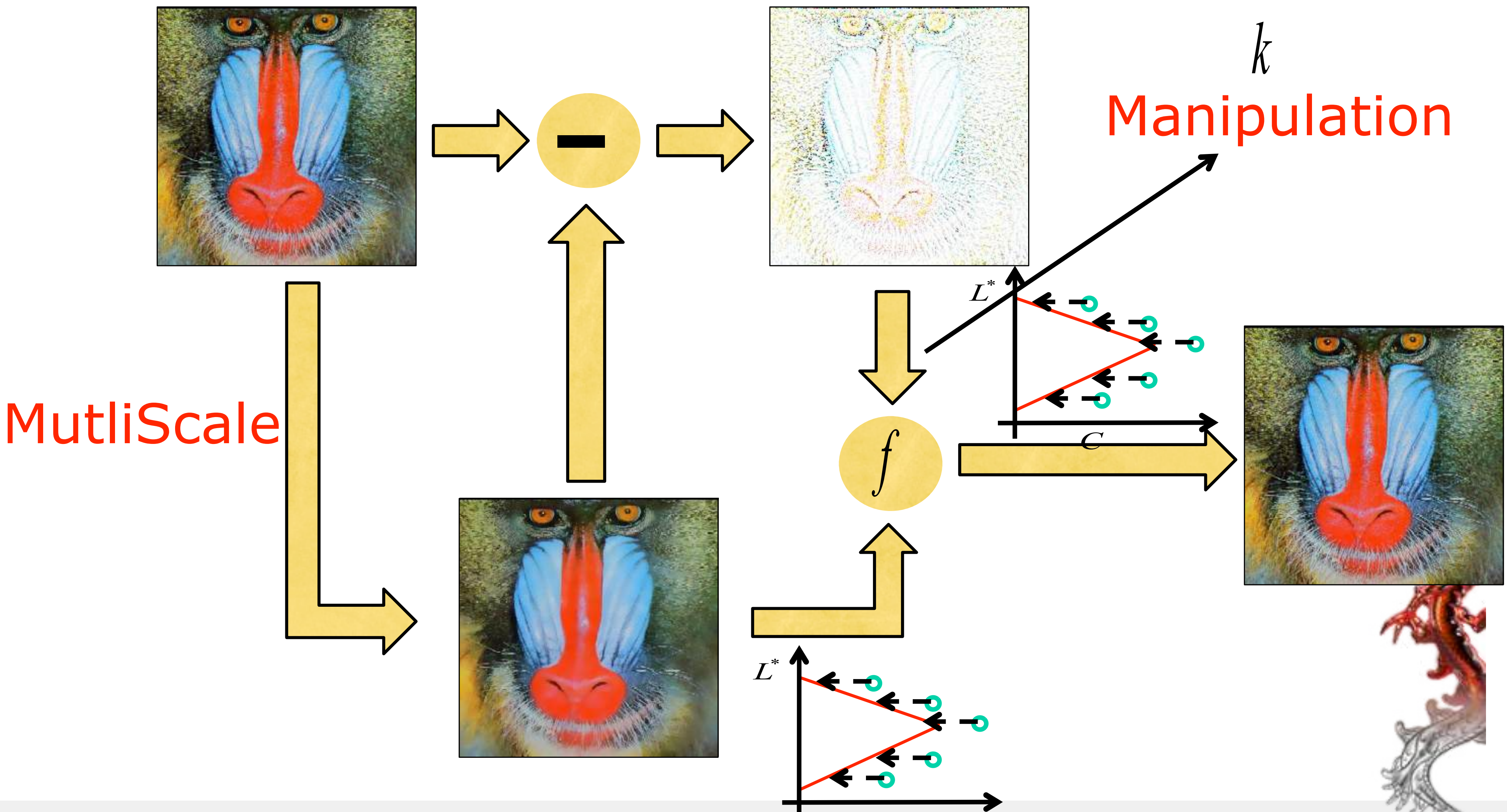
Preservation of Spatial Details



Bonnier et al. "Spatial and Color Adaptive Gamut Mapping: A Mathematical Framework and Two New Algorithms."

Preservation of Spatial Details

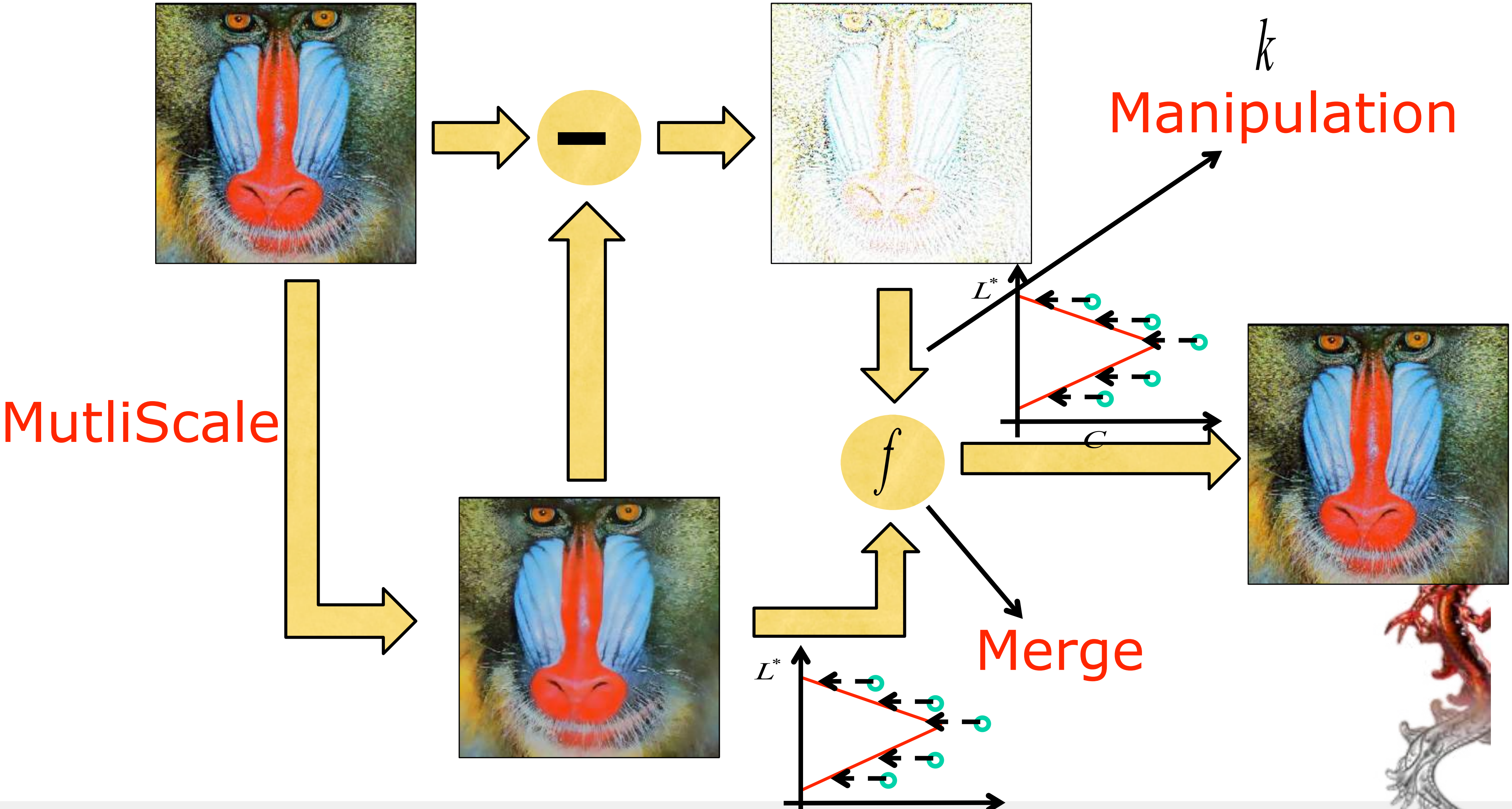
SIGGRAPHASIA2011 HONG KONG



Bonnier et al. “Spatial and Color Adaptive Gamut Mapping: A Mathematical Framework and Two New Algorithms.”

Preservation of Spatial Details

SIGGRAPHASIA2011 HONG KONG



Bonnier et al. “Spatial and Color Adaptive Gamut Mapping: A Mathematical Framework and Two New Algorithms.”

Mantiuk et al. “Color Correction for Tone Mapping”

Automatic estimation of desaturation (s) factor in function of contrast compression (c) (<u>non-linear color correction</u>).

$$C_{out} = \left(\frac{C_{in}}{L_{in}}\right)^{s} L_{out}$$

$$s(c) = \frac{(1+k_1)\, c^{k_2}}{1+k_1\, c^{k_2}}$$

k1=2.3892, k2=0.8552

s = f(c) determined based on results of perceptual experiment

Mantiuk et al. “Color Correction for Tone Mapping”

luminance(C_{in}) = luminance(C_{out})

$$C_{out} = \left(\left(\frac{C_{in}}{L_{in}} - 1 \right) s + 1 \right) L_{out}$$

k1=2.3892, k2=0.8552

$$s(c) = \frac{(1+k_1)\, c^{k_2}}{1+k_1\, c^{k_2}}$$

Unchanged luminance value after color correction
(luminance preserving solution)

Conclusions

- Works on high dynamic range imaging have mostly operated on luminance (lightness) information
  - some works start to appear proposing solution for color saturation, acquisition of colorimetric correct high dynamic range color values, and color appereance
- In Color Science a lot of works have been presented in the context of colorimetric characterisation, color appearance and gamut mapping on low dynamic range [0, 100]
  - Some of these works have been extended or re-used for high dynamic range applications
  - Tone mapping can bee seen as an extension or a particular case of gamut mapping (if we consider only the luminance information)
  - Many gamut mapping works started to analyse the details preservation on color information

Conclusions

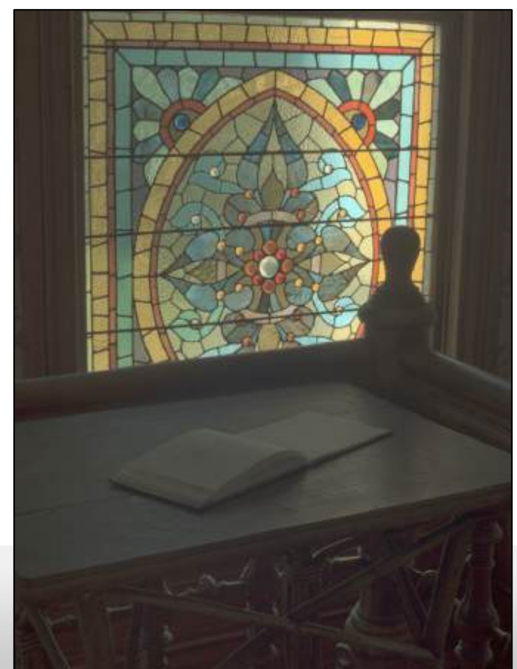
- Works on high dynamic range imaging have mostly operated on luminance (lightness) information
  - some works start to appear proposing solution for color saturation, acquisition of colorimetric correct high dynamic range color values, and color appereance
- In Color Science a lot of works have been presented in the context of colorimetric characterisation, color appearance and gamut mapping on low dynamic range [0, 100]
  - Some of these works have been extended or re-used for high dynamic range applications
  - Tone mapping can bee seen as an extension or a particular case of gamut mapping (if we consider only the luminance information)
  - Many gamut mapping works started to analyse the details preservation on color information

Low Dynamic Range [0,100]

Acknowledgments

- Image IM2-Color (slide 2) Courtesy of Laszlo Neumann
- Material from the paper "Color Correction for Tone Mapping" Courtesy of Rafal Mantiuk
- Image Bottles (slides 12 and 15) Courtesy of Francesco Banterle
- Images (slides 30 and 41) Courtesy of Ela Sikudova





Retargeting From LDR to HDR: Reverse/Inverse Tone Mapping

Dr. Francesco Banterle

Visual Computing Lab, ISTI-CNR, Italy

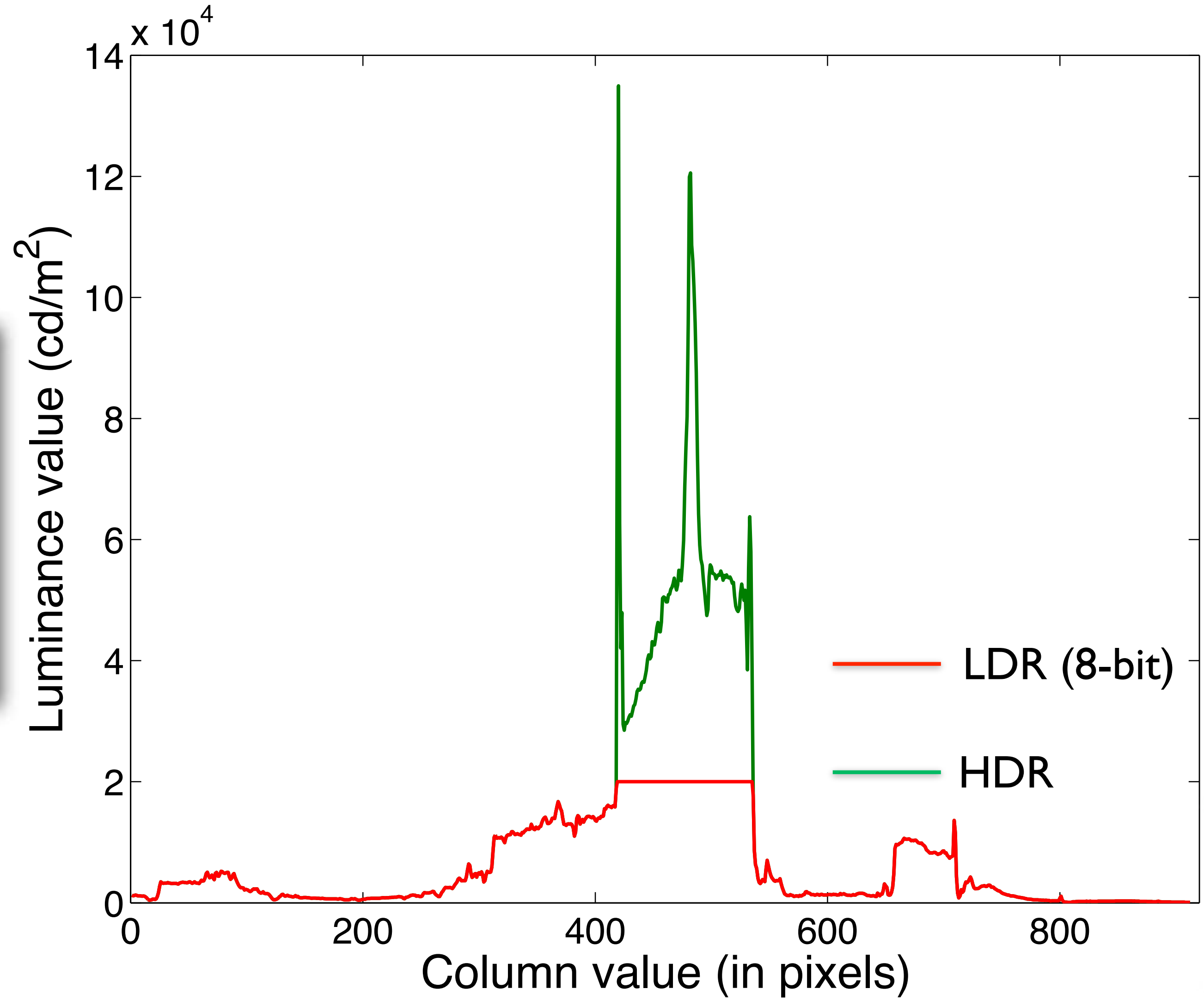
francesco.banterle@isti.cnr.it

frabante@gmail.com

Outline of the Talk

- An Overview on Reverse/Inverse Tone Mapping
- Expansion Methods:
  - Global Methods
  - Expand Map Methods
  - Classification Methods
  - User Based Methods
- Evaluation:
  - Psychophysical Experiments
  - Computational Metrics
- Conclusions

Overview on Reverse/Inverse Tone Mapping



Overview on RTM/ITM: Why?

SIGGRAPHASIA2011 HONG KONG

- Why do we need RTM/ITM?
  - **We want to retarget LDR content into HDR monitors, applications (i.e. Image Based Lighting), and editing!**

- The general operator is typically defined as:

$$g(I) = \mathbb{D}_{\mathrm{i}}^{w \times h \times c} \rightarrow \mathbb{D}_{\mathrm{o}}^{w \times h \times c}$$

- Common steps of these operators:
  - Linearization of the LDR image
  - Noise and quantization reduction
  - Luminance Expansion

Overview on RTM/ITM: Why?

- Why do we need RTM/ITM?
  - **We want to retarget LDR content into HDR monitors, applications (i.e. Image Based Lighting), and editing!**

- The general operator is typically defined as:

$$g(I) = \underbrace{\boxed{\mathbb{D}_{\mathrm{i}}^{w \times h \times c}}}_{\text{LDR}} \rightarrow \mathbb{D}_{\mathrm{o}}^{w \times h \times c}$$

- Common steps of these operators:
  - Linearization of the LDR image
  - Noise and quantization reduction
  - Luminance Expansion

Overview on RTM/ITM: Why?

- Why do we need RTM/ITM?
  - **We want to retarget LDR content into HDR monitors, applications (i.e. Image Based Lighting), and editing!**

- The general operator is typically defined as:

$$g(I) = \boxed{\mathbb{D}_{\mathrm{i}}^{w \times h \times c}} \rightarrow \boxed{\mathbb{D}_{\mathrm{o}}^{w \times h \times c}}$$

LDR HDR

- Common steps of these operators:
  - Linearization of the LDR image
  - Noise and quantization reduction
  - Luminance Expansion

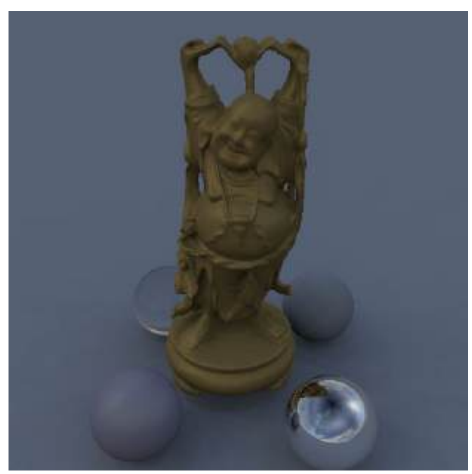
Global Methods (I)

- Landis [Landis02] proposed a simple function for generating HDR images for VFX:

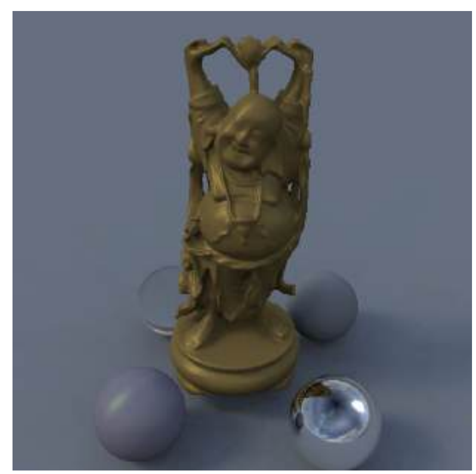
$$L_{\mathrm{w}}(\mathbf{x}) = \begin{cases} (1-k)L_{\mathrm{d}}(\mathbf{x}) + kL_{\mathrm{w,\,max}}L_{\mathrm{d}}(\mathbf{x}) & \text{if } L_{\mathrm{d}}(\mathbf{x}) \geq R; \\ L_{\mathrm{d}}(\mathbf{x}) & \text{otherwise,} \end{cases}$$

$$k = \left(\frac{L_{\mathrm{d}}(\mathbf{x}) - R}{1-R}\right)^{\alpha},$$

Original LDR EM



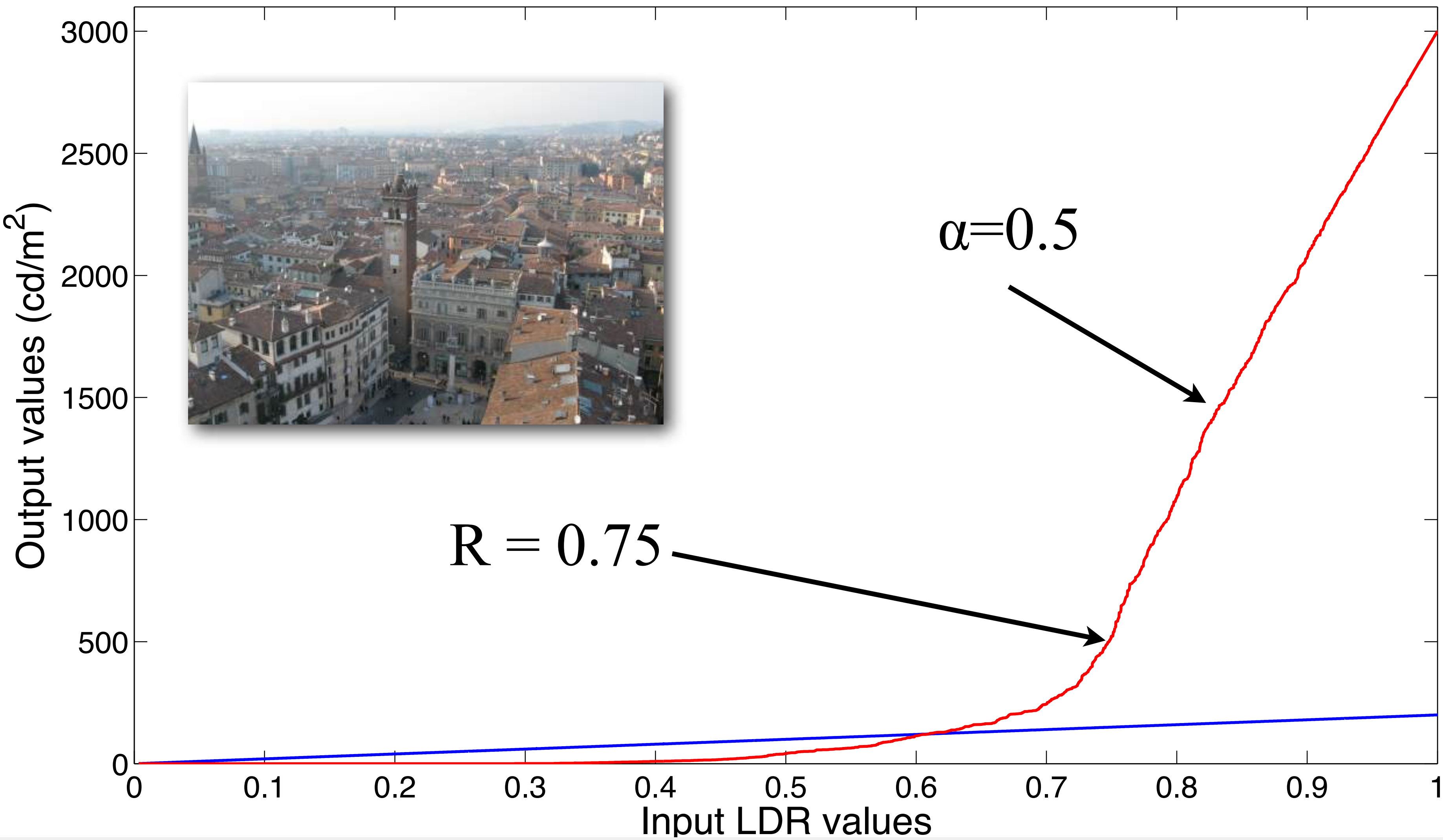
Rendered with LDR EM



Rendered ITMO EM

LDR Environment map is courtesy of H. Landis [Landis 02]

Global Methods (II)



Global Methods (III)

- Akyüz et al. [AFR*07] shown that "***a simple linear scale can provide an HDR experience***" based on psychophysically experiments:

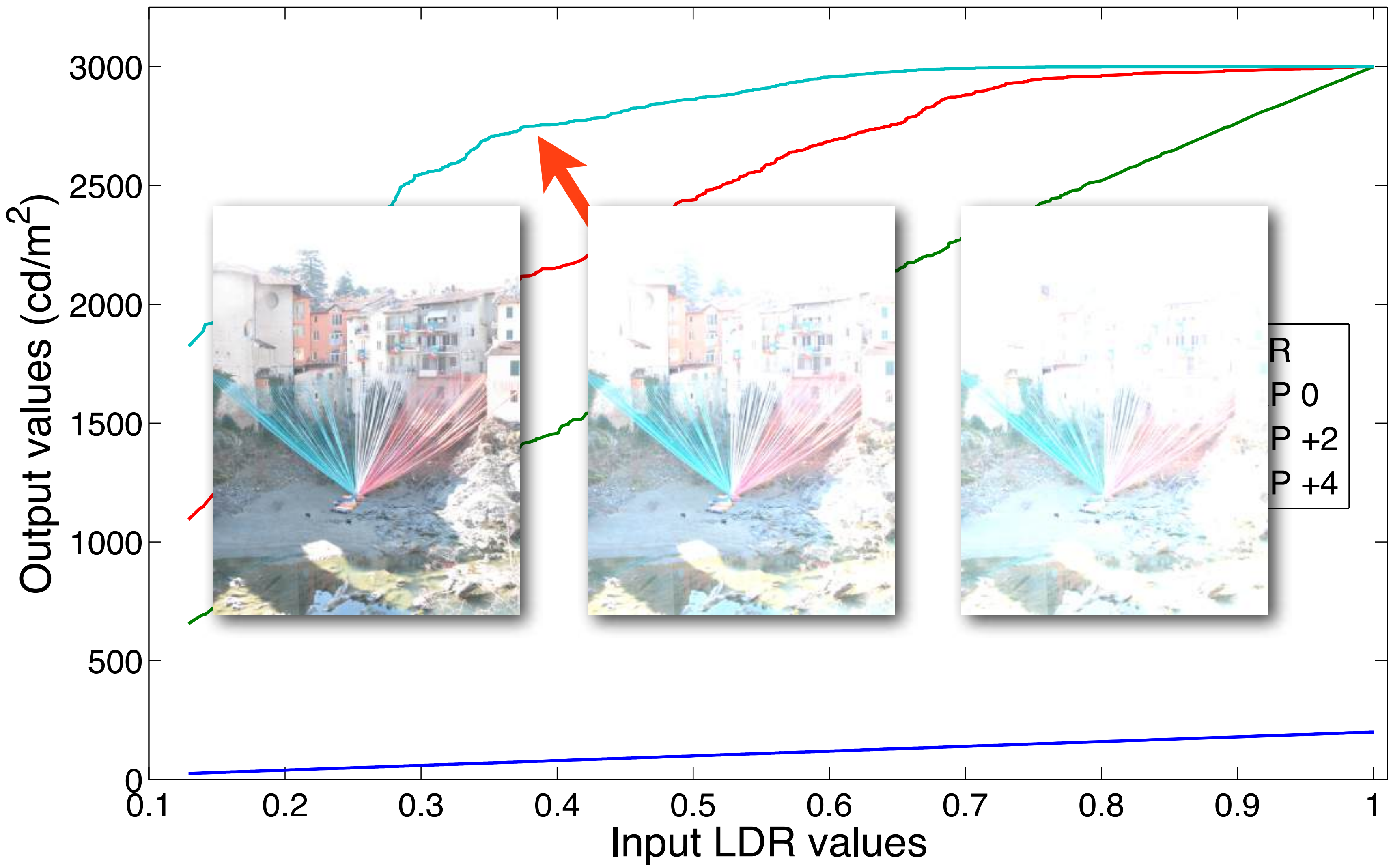
$$L_{\mathrm{w}}(\mathbf{x}) = k\left(\frac{L_{\mathrm{d}}(\mathbf{x}) - L_{\mathrm{d,\ min}}}{L_{\mathrm{d,\ max}} - L_{\mathrm{d,\ min}}}\right)^{\gamma}$$

- Masia et al. [MAF*09] shown that for over-exposed images a non-linear function (gamma) needs to be applied. This non-linearity depends on exposedness of the image:

$$L_{\mathrm{w}}(\mathbf{x}) = L_{\mathrm{d}}(\mathbf{x})^{\gamma} \qquad \gamma = 10.44k - 6.282$$

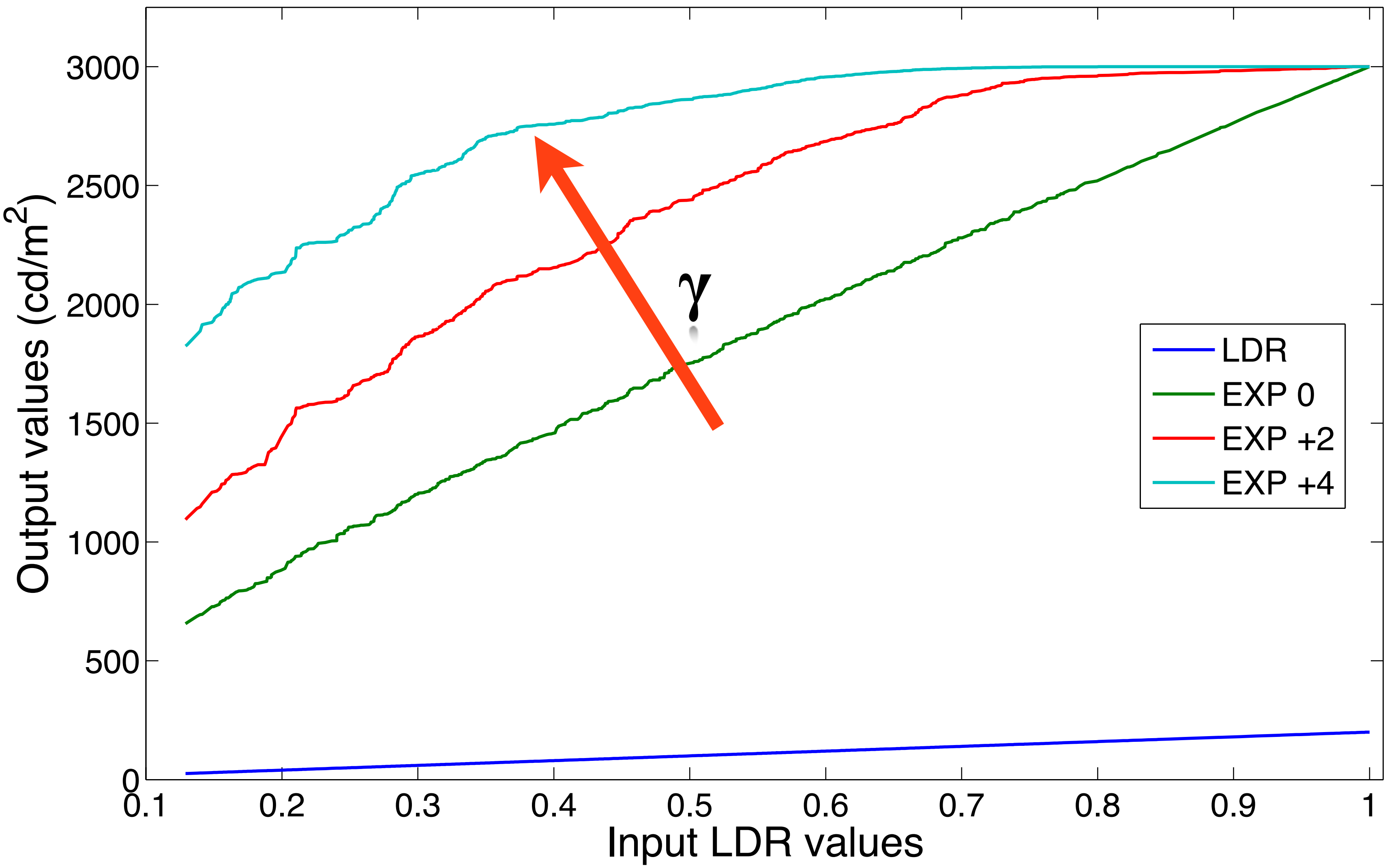
$$k = \frac{\log L_{\mathrm{d,\ avg}} - \log L_{\mathrm{d,\ Min}}}{\log L_{\mathrm{d,\ Max}} - \log L_{\mathrm{d,\ Min}}} \quad k > 0.65$$

Global Methods (IV)

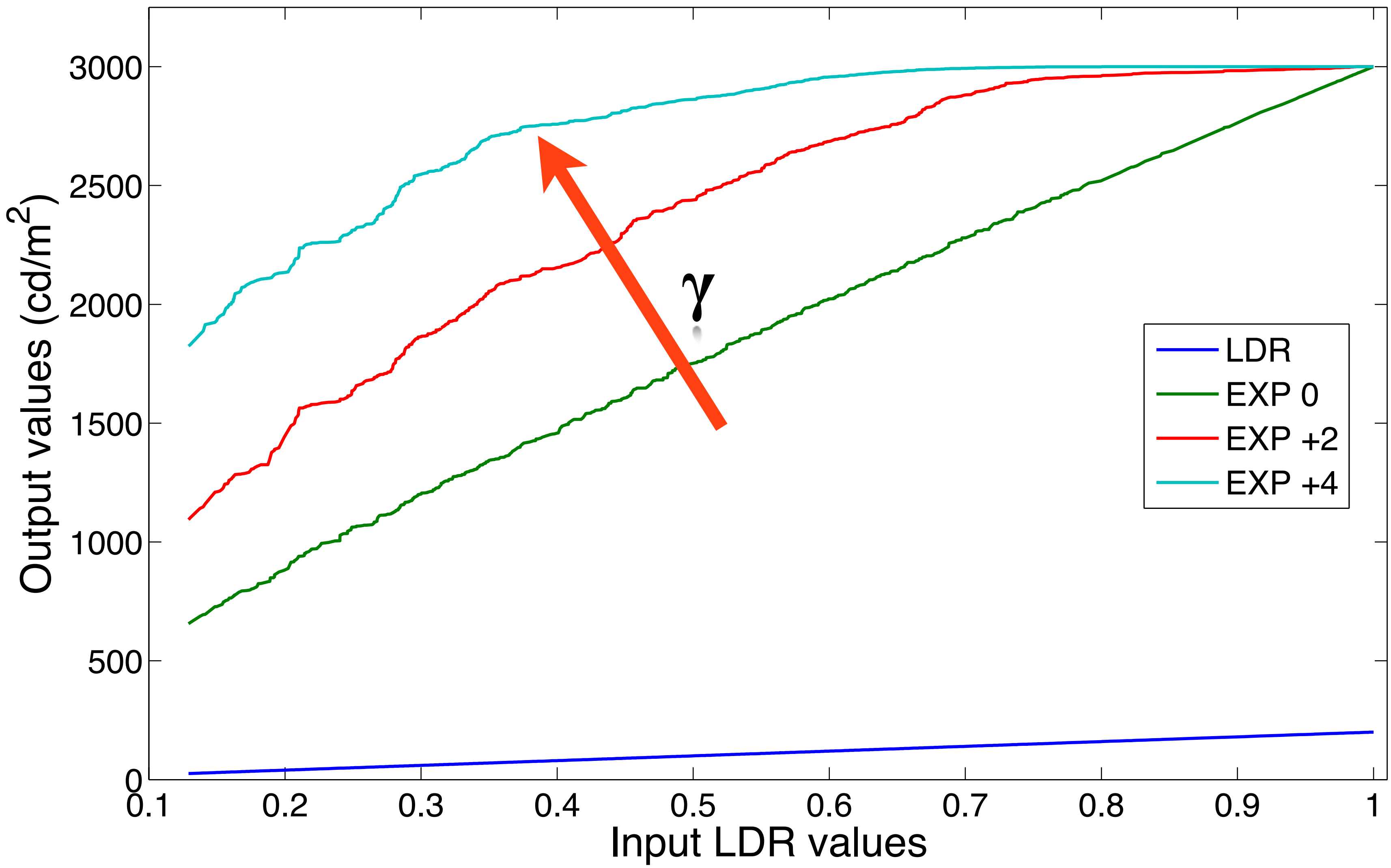


Global Methods (IV)

SIGGRAPHASIA2011 HONG KONG



Global Methods (IV)



Classification Methods: Highlights Reproduction on HDR Monitors (I)

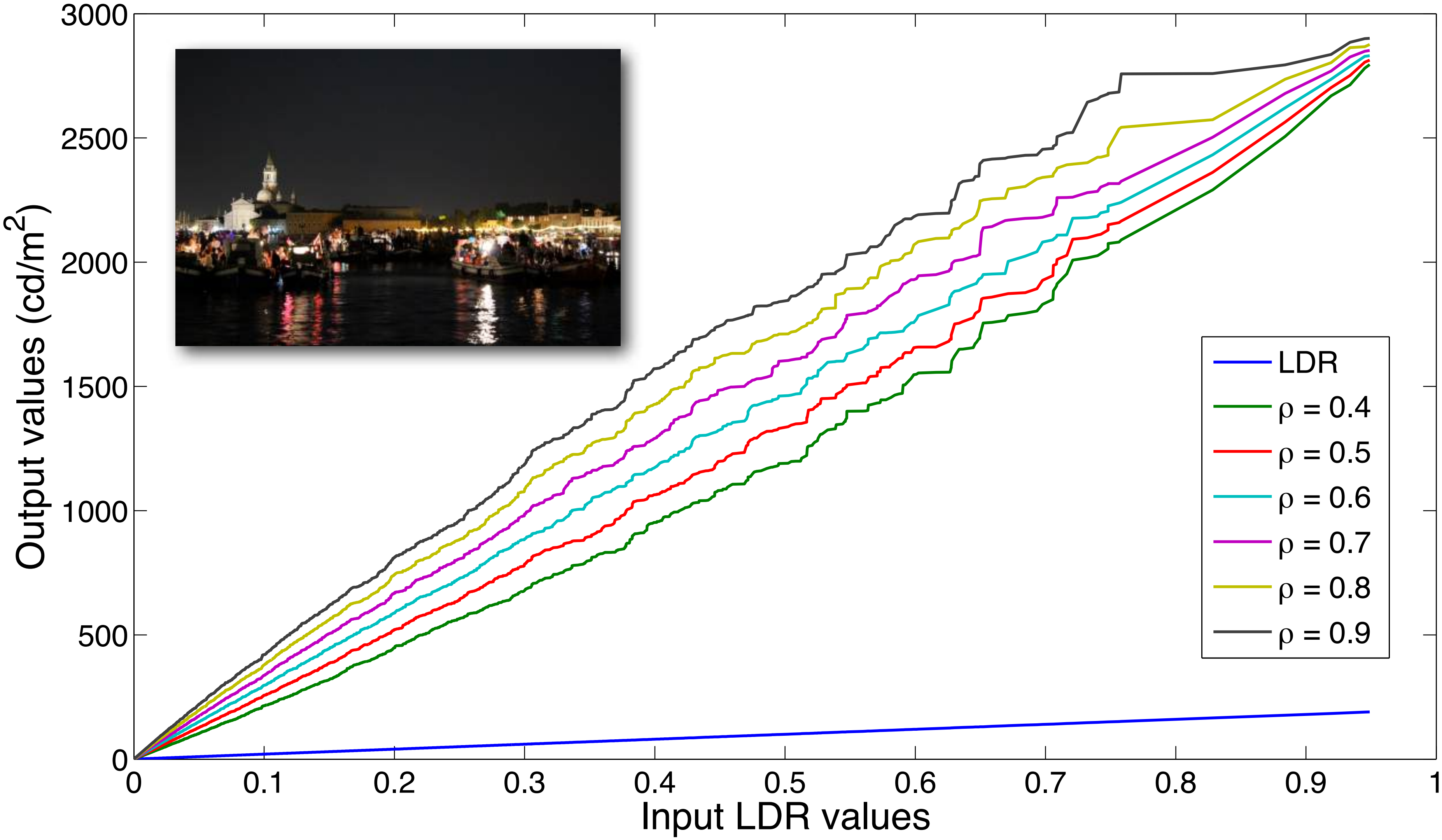
- Meylan et al. [MDDS06, MDS07] present a classification approach:
  - Expand highlights and specular surfaces ($\omega > 0$)
  - ω is computed using robust thresholding
  - Expansion using a two-scale model:

$$L_{\mathrm{w}}(\mathbf{x}) = f(L_{\mathrm{d}}(\mathbf{x})) = \begin{cases} s_1 L_{\mathrm{d}}(\mathbf{x}) & \text{if } L_{\mathrm{d}}(\mathbf{x}) \leq \omega \\ s_1\omega + s_2(L_{\mathrm{d}}(\mathbf{x}) - \omega) & \text{otherwise} \end{cases}$$

$$s_1 = \frac{\rho}{\omega} \quad s_2 = \frac{1-\rho}{L_{\mathrm{d,\ Max}} - \omega}$$

  - To avoid contouring low-pass filtering on expanded regions

Classification Methods: Highlights Reproduction on HDR Monitors (II)

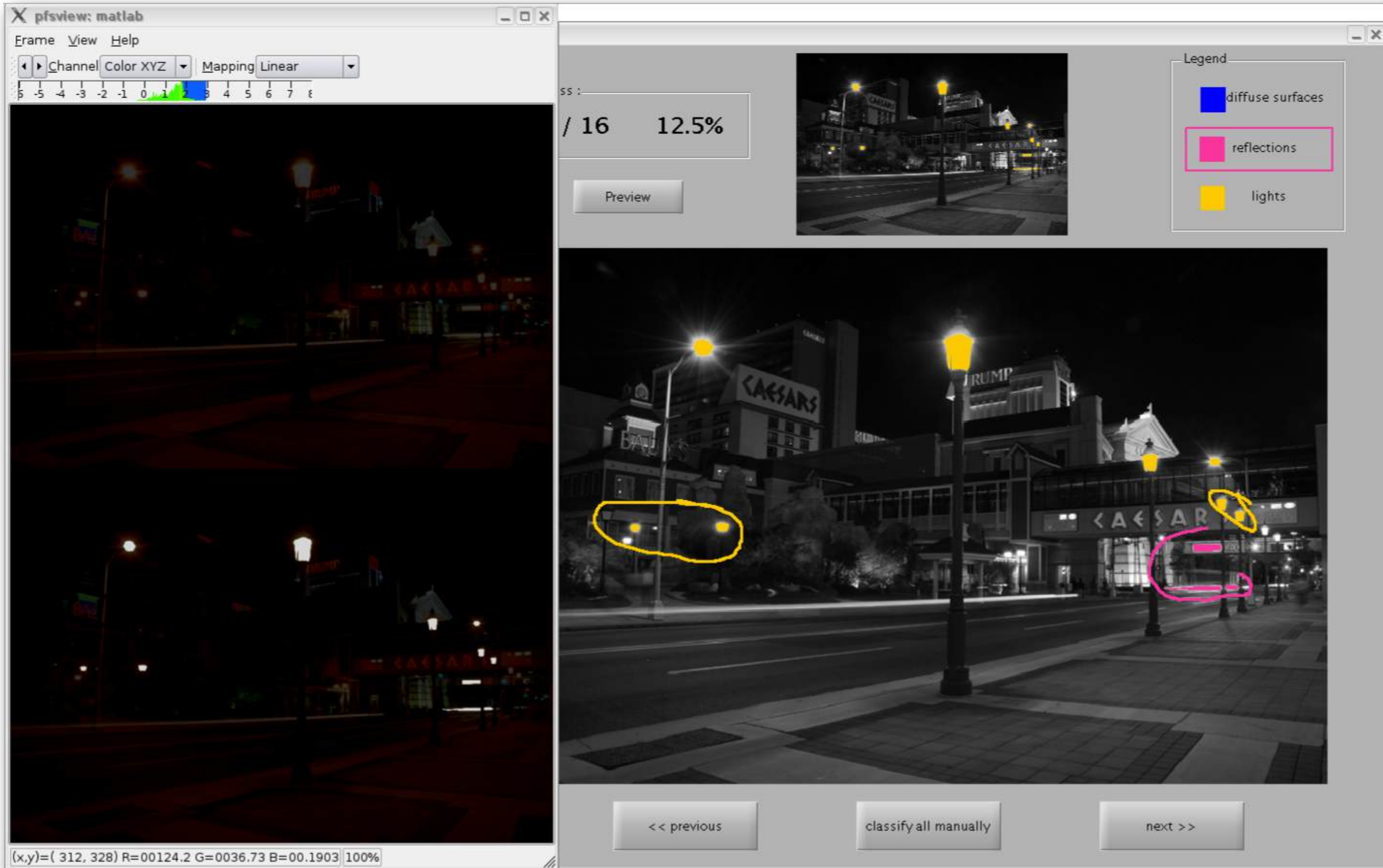


Classification Methods: Enhancement of Bright Videos (I)

- Didyk et al. [DMHS08] extended Meylan et al.'s method:
  - Three classification areas: diffuse, reflections, and lights
  - Automatic Classifier (AC):
    - SVM + Nearest Neighbor + Tracking $\Rightarrow$ 3% error
  - User interface for adjusting the AC errors
  - Non-linear adaptive tone curve for expanding the range based on the histogram of the region:
    - Bilateral filtering layers separation (high and low frequencies) for avoiding contouring

SIGGRAPHASIA2011 HONG KONG

Classification Methods: Enhancement of Bright Videos (II)



Classification Methods: Selective Reverse Tone Mapping (I)

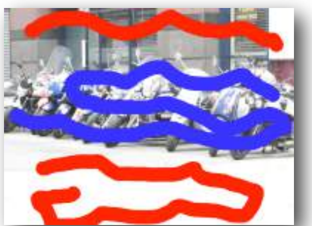
- Masia et al. [MFSG10] proposed a novel approach based on saliency:
  - **Range Expansion (RE)**: pice-wise linear expansion using the zonal system by Adams (9 zones):

$$p = \left(\frac{\exp(v \sin(\pi \frac{z-1}{16})) - 1}{\exp(v) - 1} \right)^{\frac{1.0}{2.2}} \qquad v = 5.25 \quad z \in [0, 9]$$

  - **Labeling**:
    - salient objects and background discrimination using different techniques:
      - learning-based saliency detection (Liu et al. [LSZ*07])
      - *saliency cuts* (Fu et al.[FCLL08])
    - Different Labels ⇒ Different RE functions

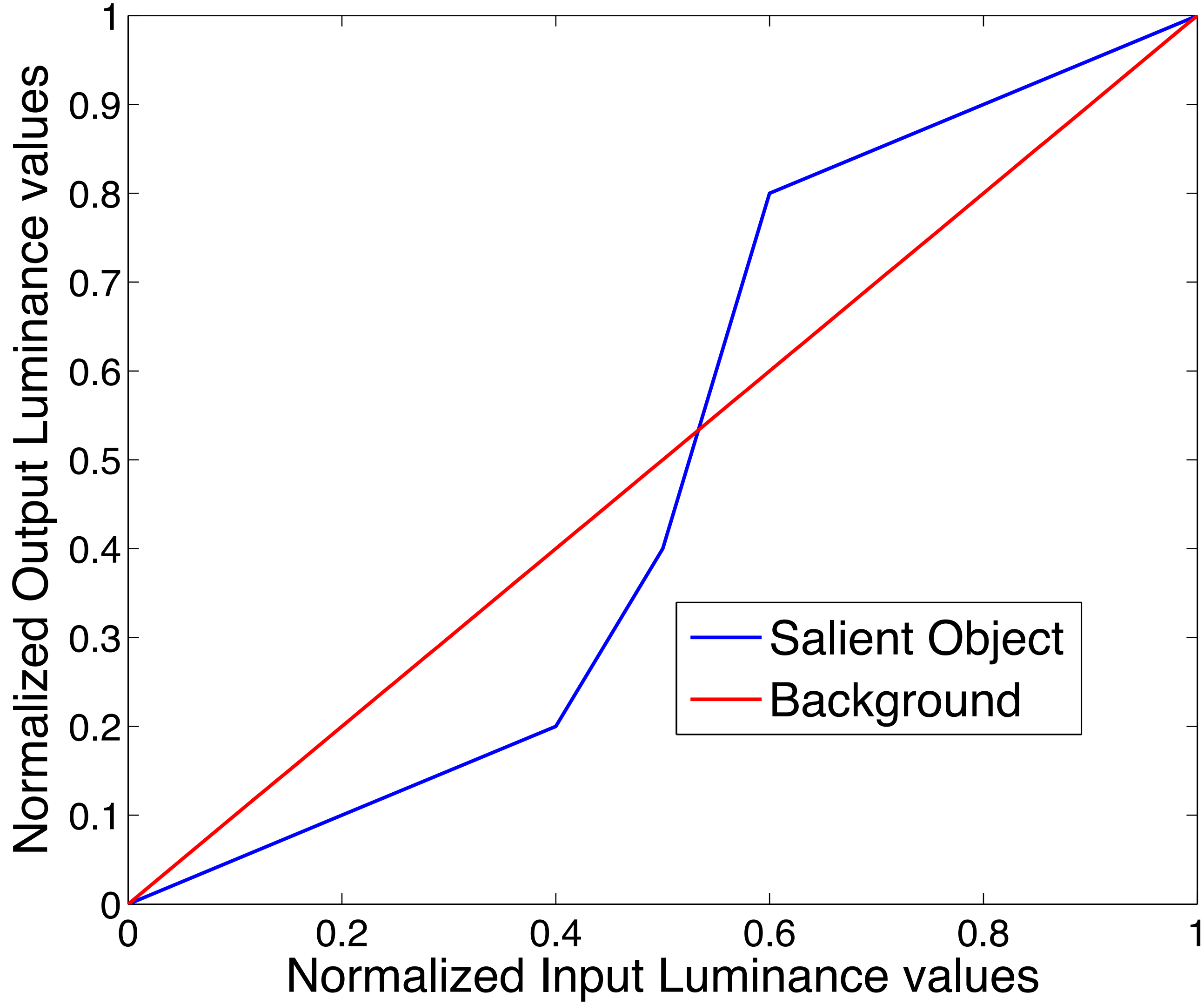
Classification Methods: Selective Reverse Tone Mapping (II)

Input



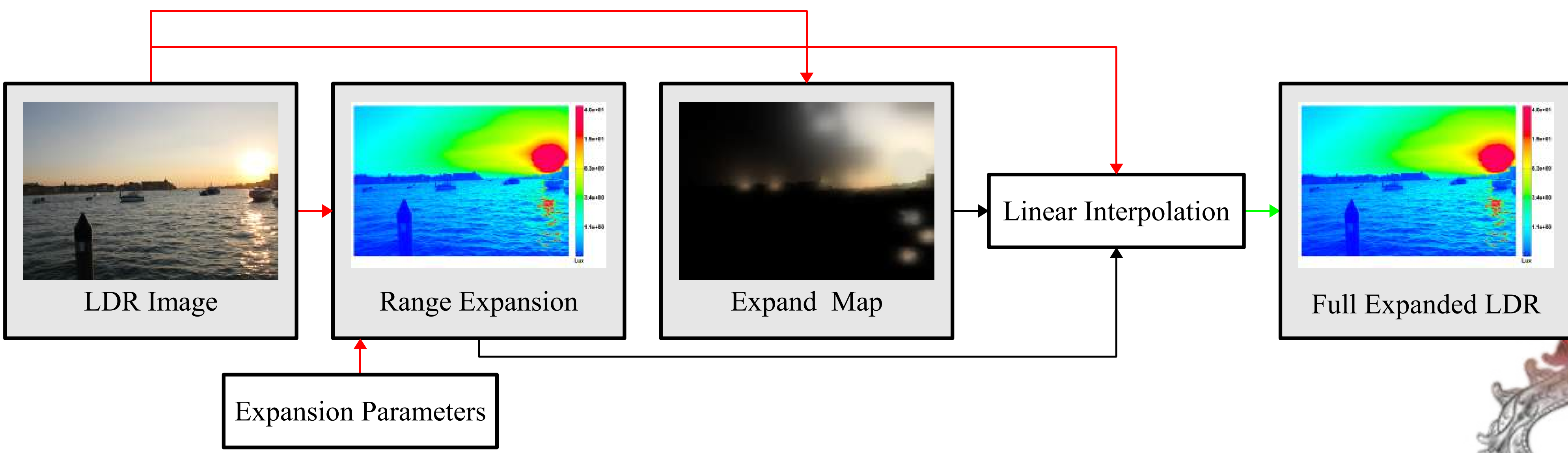
Auto-Labeling

Binary Mask



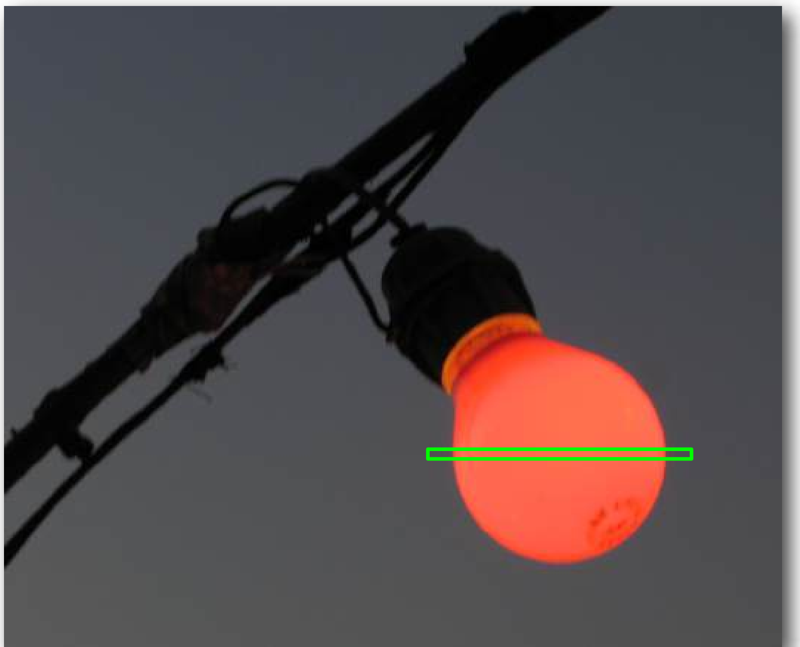
Expand Maps Methods: Non-Linear Expansion using Expand Maps (I)

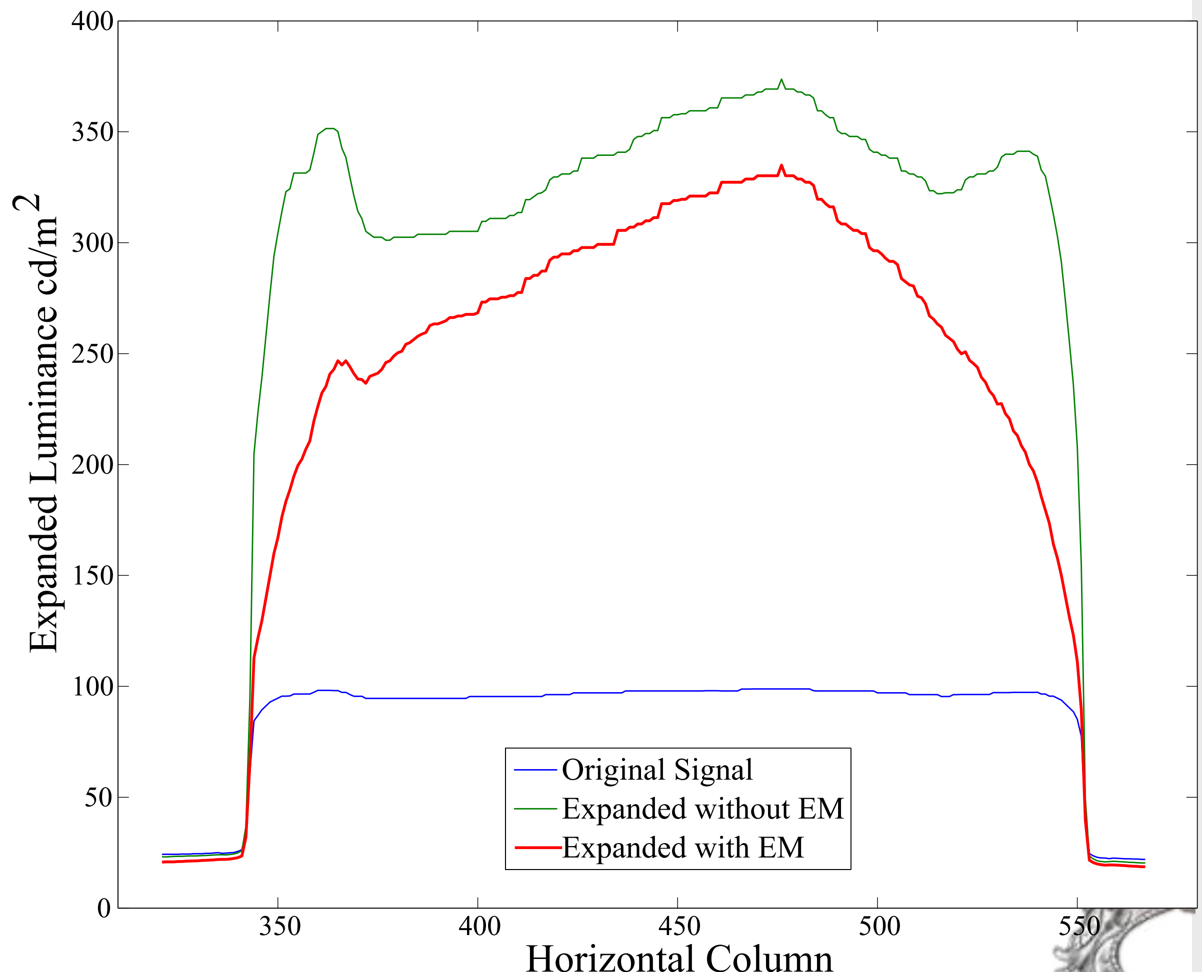
- Banterle et al. [BLDC06,BLDBC07,BLDC08,B09] presented a general and real-time framework:
  - **Range Expansion**: non-linear (inverting an TMO; other functions)
  - **Expand Map**: sampling+density estimation+cross bilateral (avoiding contouring and compression artifacts)



Expand Maps Methods: Non-Linear Expansion using Expand Maps (II)

SIGGRAPHASIA2011 HONG KONG





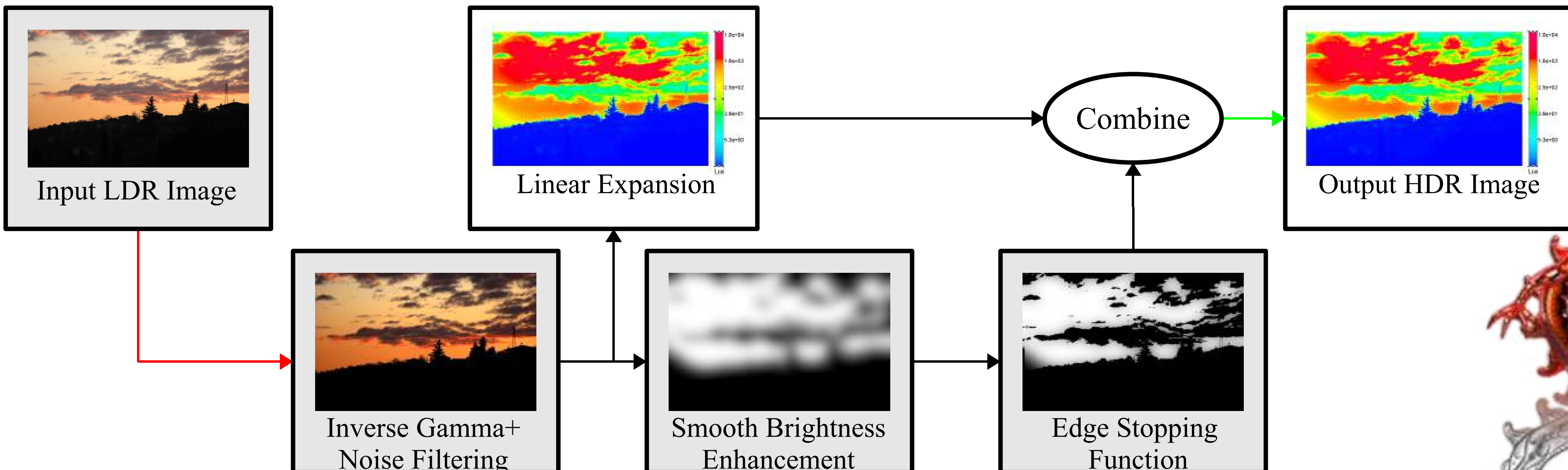
Expand Maps Methods: Non-Linear Expansion using Expand Maps (II)

IBL using original HDR

IBL using expanded LDR

Expand Maps Methods: LDR2HDR (I)

- Rempel et al. [RTS*07] presented a similar work of Banterle et al.:
  - **Range Expansion**: linear
  - **Expand Map**: thresholding+filtering+edge stopping



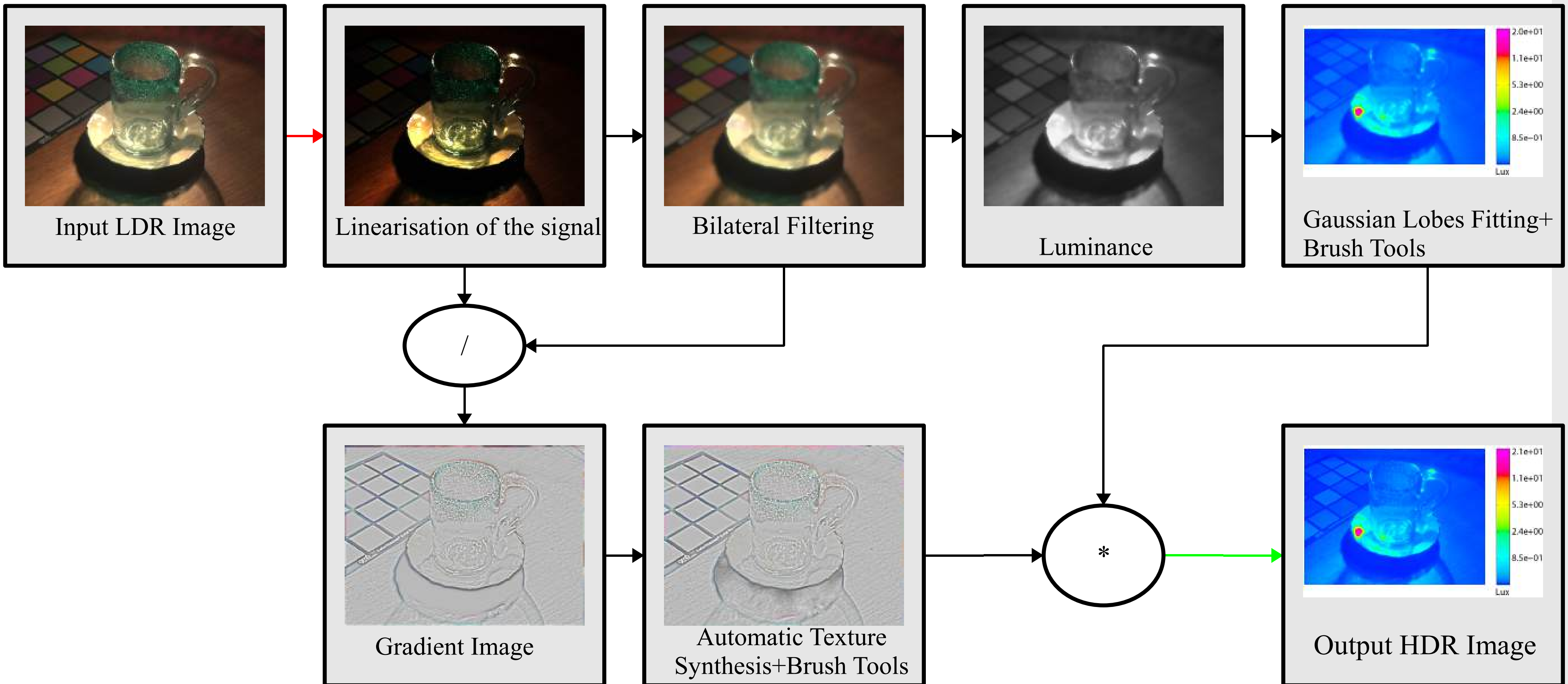
Expand Maps Methods: LDR2HDR (II)

- A variant of the algorithm was presented by Kovaleski and Oliveria [KO09] using the bilateral grid to improve the quality of the Expand Map.

User Based Methods: Hallucination (I)

- Wang et al. [WWZ*07] proposed the first user based method with reconstruction of details:
  - **HDR frequencies using the bilateral filter**: base (low) and detail (high) layers
  - **Automatic Expansion (base layer)**: saturated regions are fitted using 2D Gaussian lobes (elliptical)
  - **Reconstruction (detail layer)**:
    - Automatic: texture synthesis
    - User-based: Stamp tool (similar to the Healing tool of Photoshop 7)
    - NOTE: other images can be used as source for the missing details

User Based Methods: Hallucination (II)



Mexican Mug's image is courtesy of Ahmet Oguz Akyuz

User Based Methods: Hallucination (III), Copying Fine Details in the Detail Layer

SIGGRAPHASIA2011 HONG KONG

User Based Methods: Hallucination (III), Copying Fine Details in the Detail Layer

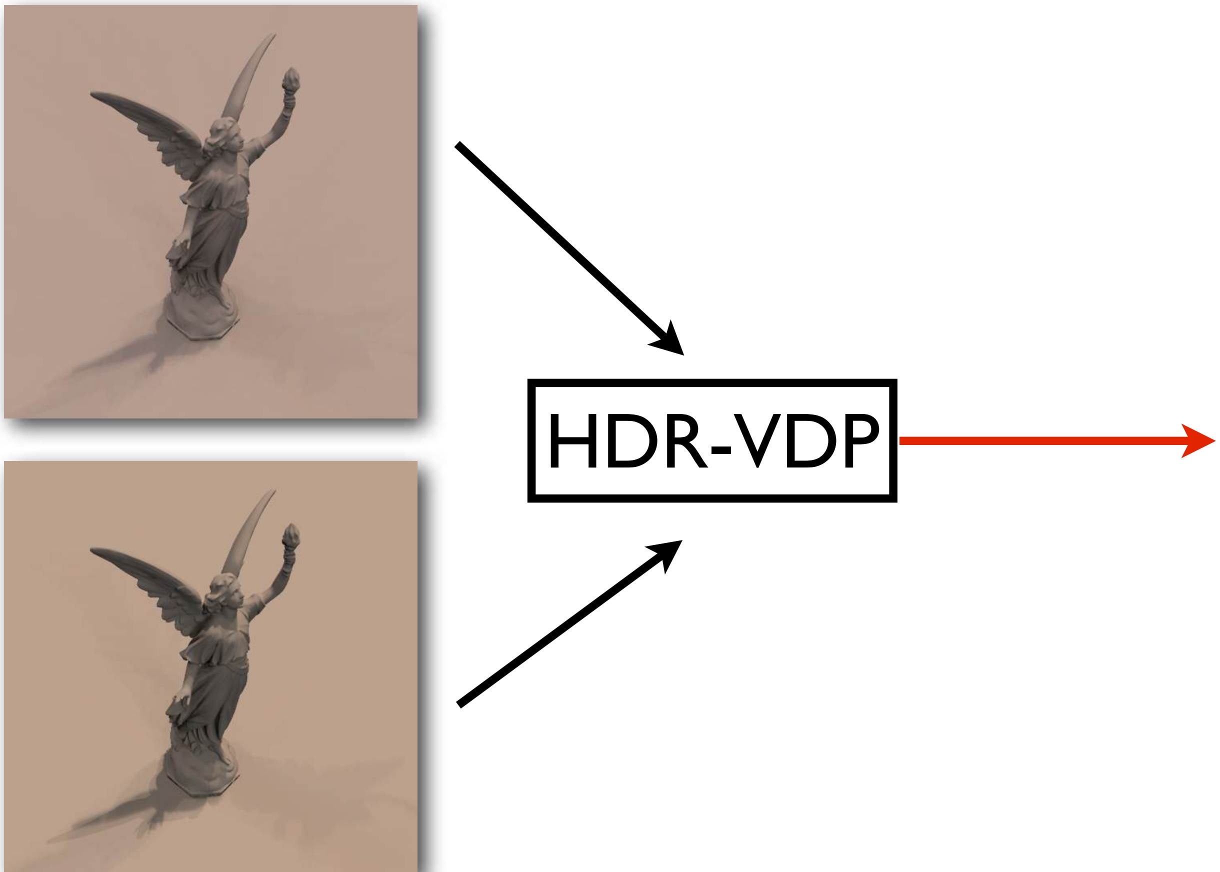
Evaluation: Why validation

- Need to evaluate different expansion methods against a "*ground truth*". Why?
  - To understand weak features or drawbacks
  - To understand important features
- rTMO/iTMO techniques do not generate exact luminance values
- Evaluation:
  - Perceptual Image Metrics: not exact comparison as in the PSNR, RMSE, etc.
  - Psychophysical Experiments

Evaluation: Perceptual Image Metrics

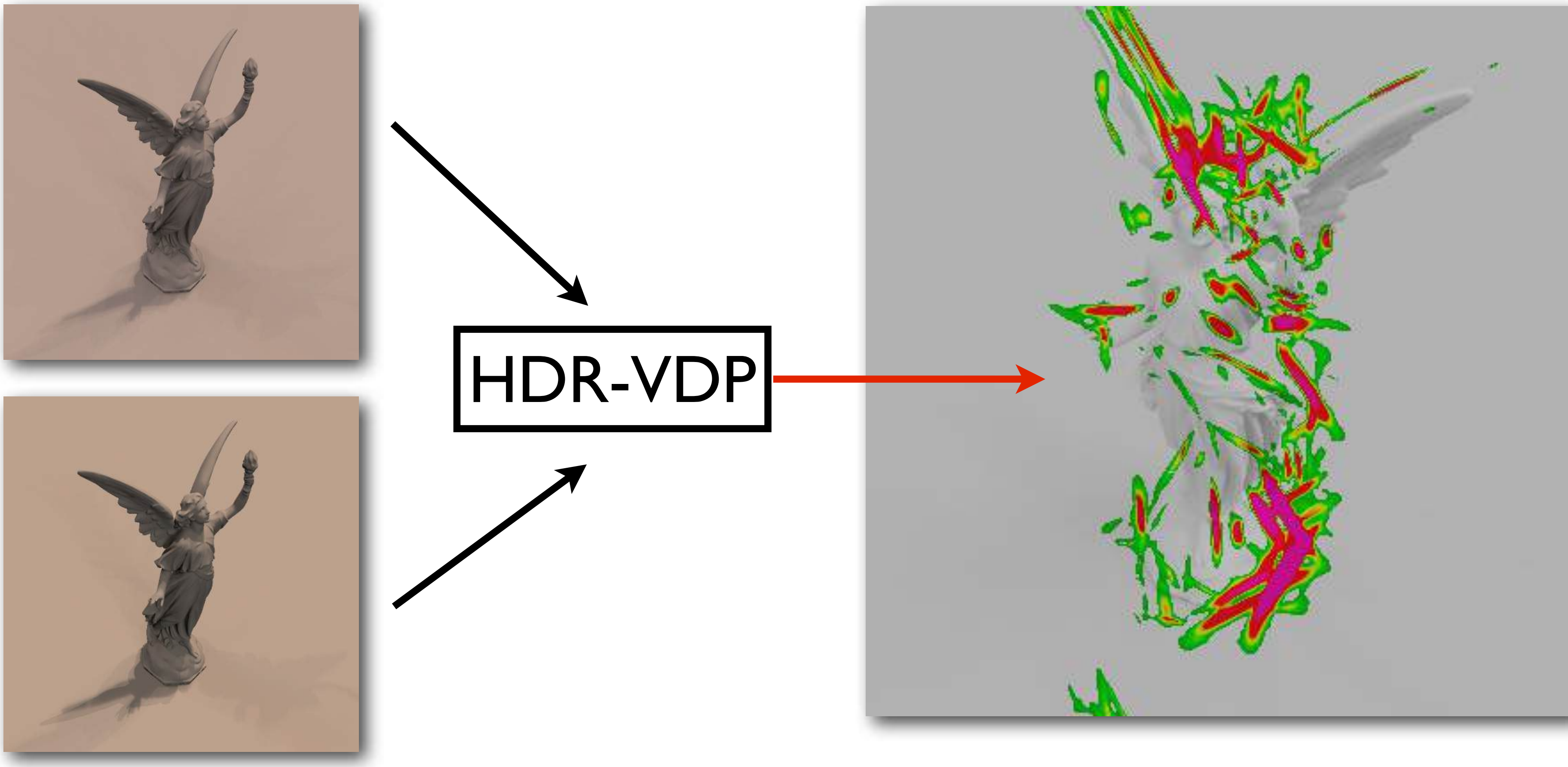
- **HDR-VDP** (current version 2.1) [MDMS04,MKRH11]: determines the probability for each pixel of being different:
  - Banterle et al. [BLDC06,BLDCB07,BLDC08,B09] used it to validate that their models were performing better than a simple non-linear expansion, validate against other methods, etc.
- **DI-IQA** [AMMS08]: detects changes in details visibility, quantization artifacts. Validation of the quality in general:
  - Masia et al. [MAF*09] and Kovaleski and Oliveria [KO09] used it to prove that their methods introduce less distortions during LDR expansion

Evaluation: Perceptual Image Metrics (II)



Lucy model is courtesy of the Stanford 3D Scanning Repository

Evaluation: Perceptual Image Metrics (II)



Lucy model is courtesy of the Stanford 3D Scanning Repository

Evaluation: Psychophysical Experiments

- Pairwise comparisons of HDR videos [DMHS08]:
  - validation of the method against LDR, and LDR2HDR
- Pairwise comparisons of HDR images [BLD*09]: comparisons for image visualization and IBL:
  - quantization artifacts need to be handle for better quality.
  - IBL needs non-linear expansion.
- Rating of HDR images and tone mapped expanded images [MAF*09]:
  - understanding preferences in very over-exposed area
  - understanding artifacts in expanded images.

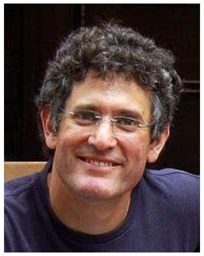
Conclusions:

- LDR Expansion for HDR applications:
  - LDR expansion methods are needed to be used in HDR applications (HDR visualization, Image Based Lighting, etc.)
  - The size of over/under-exposed areas is a limitation when recreating the content
- What's important?
  - To have non-linearity or controllable expansion functions
  - Avoid artifacts' boosting: quantization and JPEG-like compression
  - Take care of over-exposed areas differently

Spatial Retargeting

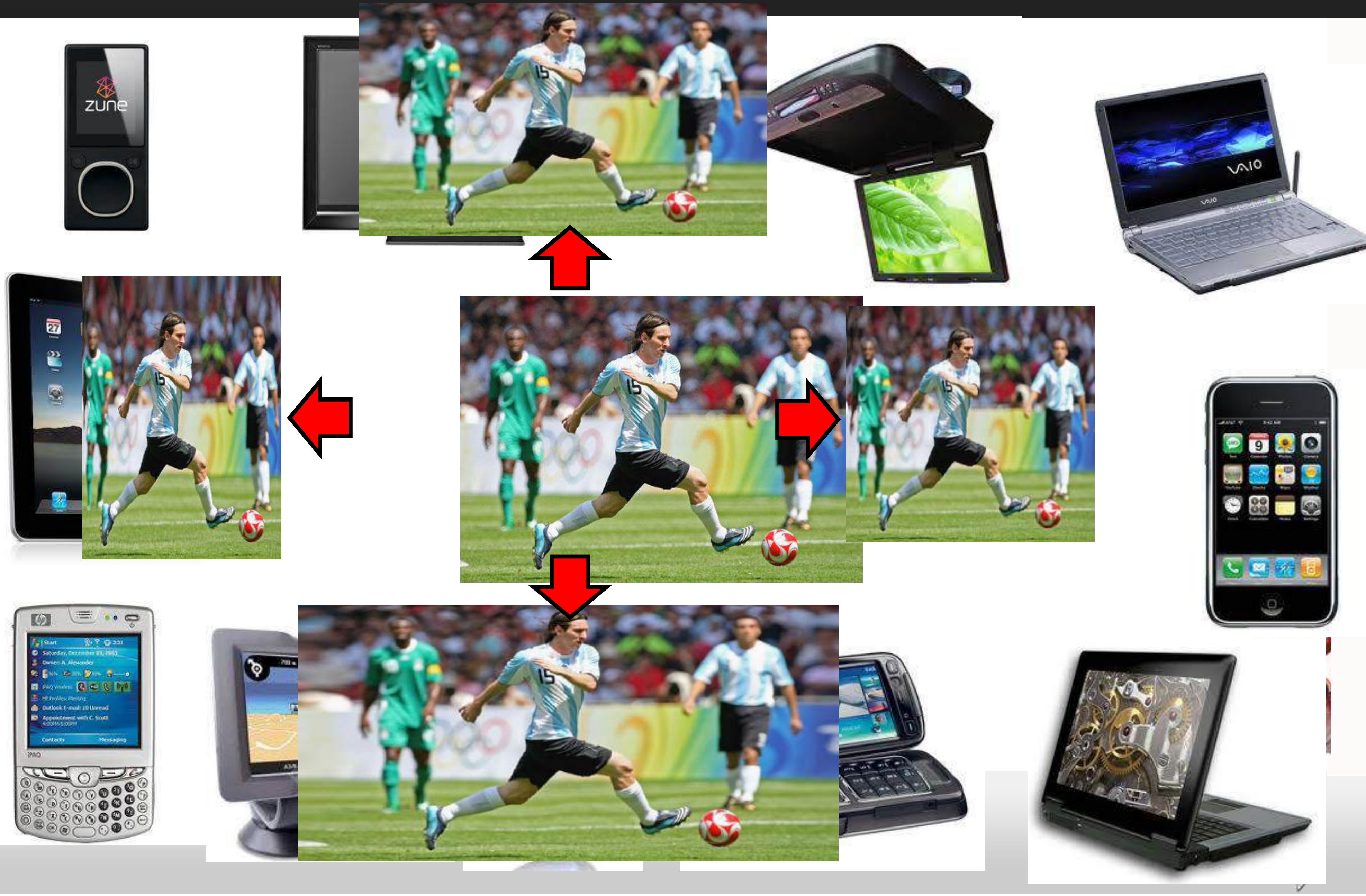
Diego Gutierrez
Universidad de Zaragoza

(slides material also from Miki Rubinstein, Olga Sorkine, Arik Shamir and Susana Castillo)

The Retargeting Problem

SIGGRAPHASIA2011 HONG KONG



Common solutions

- Homogeneous squeezing/stretching
- Cropping [DeCarlo and Santella 2002; Viola and Jones 2004…]
- Hybrid solution [modern TV sets]


original


squeeze


crop


hybrid

Visual Media Retargeting: Siggraph Asia Course 2009

Ariel Shamir

The Interdisciplinary Center, Herzliya

Olga Sorkine

New York Univeristy

Visual Media Retargeting: An Example

[Avidar & Shamir 2007]

Visual Media Retargeting: Scaling

SIGGRAPHASIA2011 HONG KONG

Scaling

[Avidar & Shamir 2007]

Visual Media Retargeting: Seams

Insert & remove seams

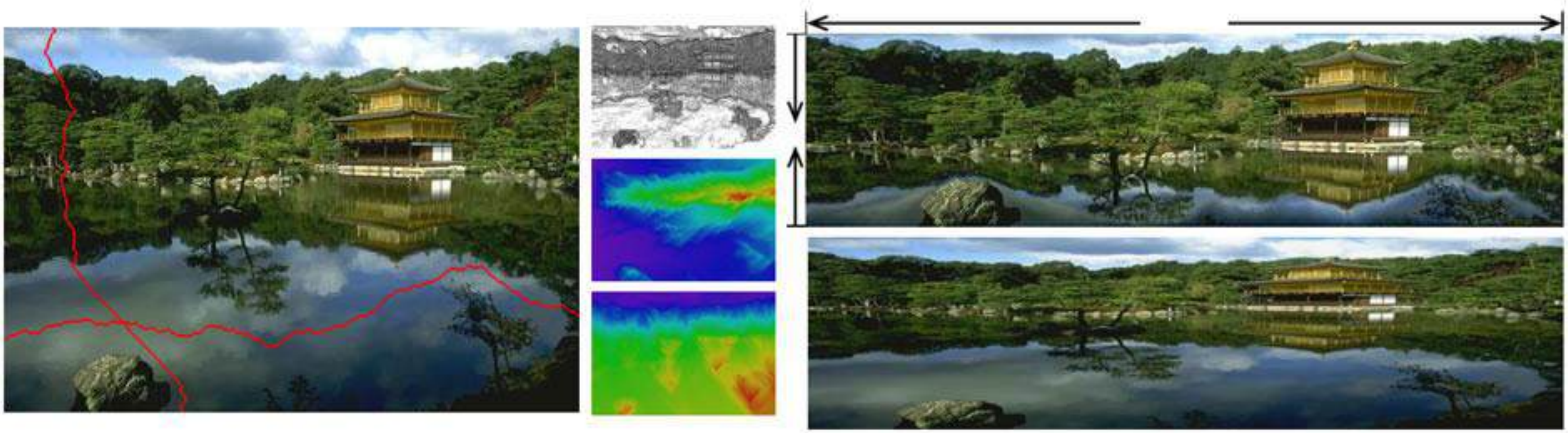
Scaling

[Avidar & Shamir 2007]

Visual Media Retargeting: Energy Concept

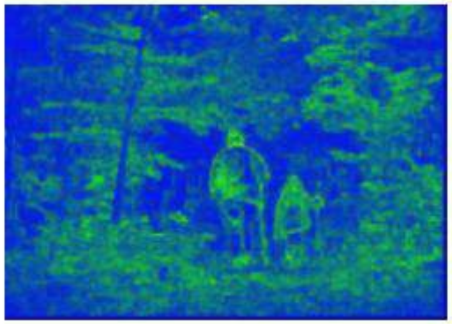
1. Define an energy function **E**(**I**) (interest, importance, saliency…)

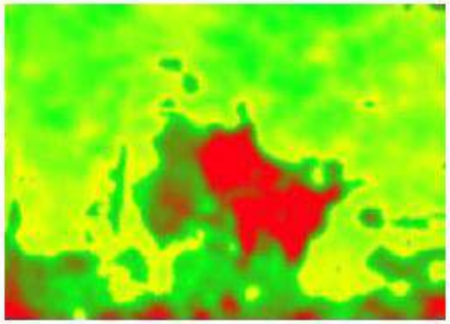
2. Use some operator(s) to change the image **I**



[Avidar & Shamir 2007]

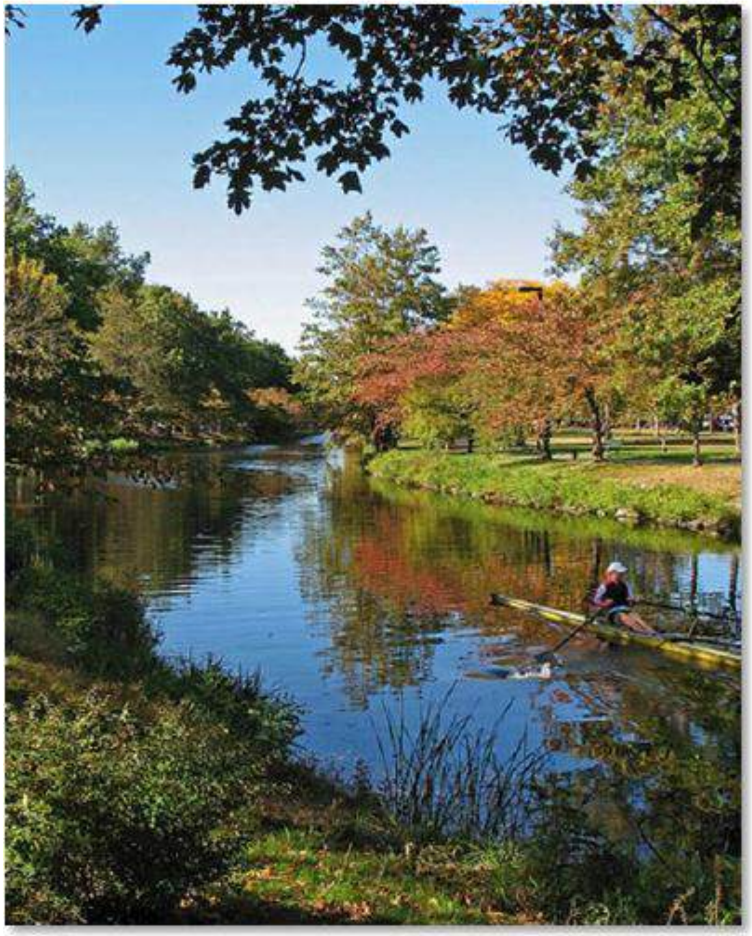
- Magnitude of gradients (simple)
- Saliency (e.g. Itty’s method) - multires

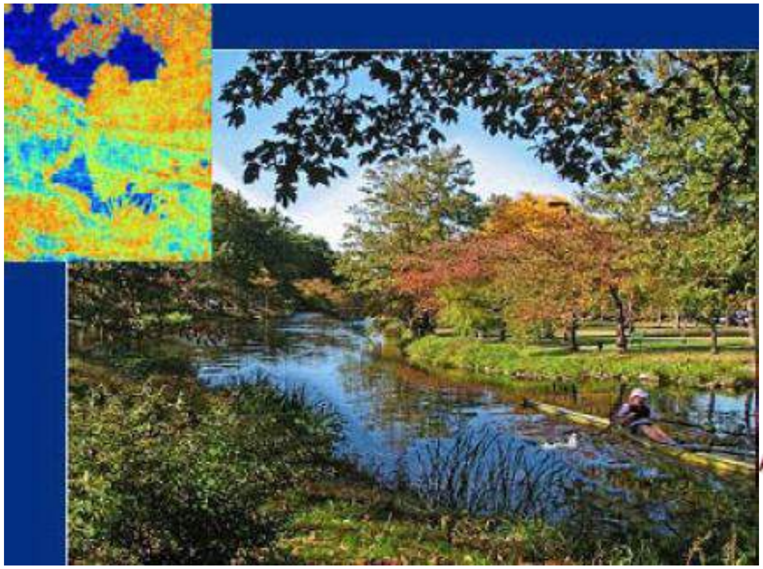




[Shamir and Sorkine 2009]

Different energy functions

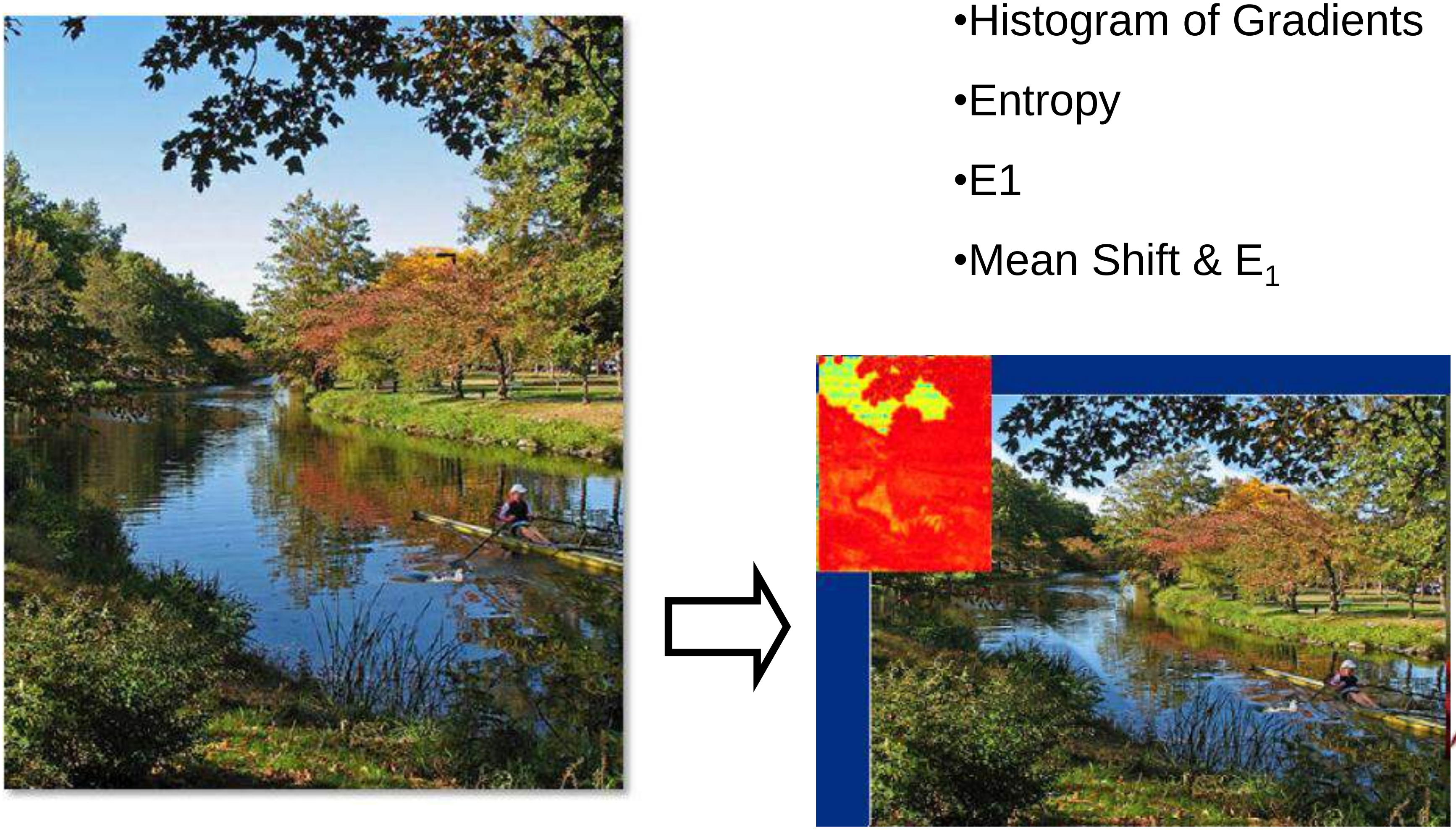




- Histogram of Gradients
- Entropy
- E1
- Mean Shift & E_1

[Shamir and Sorkine 2009]

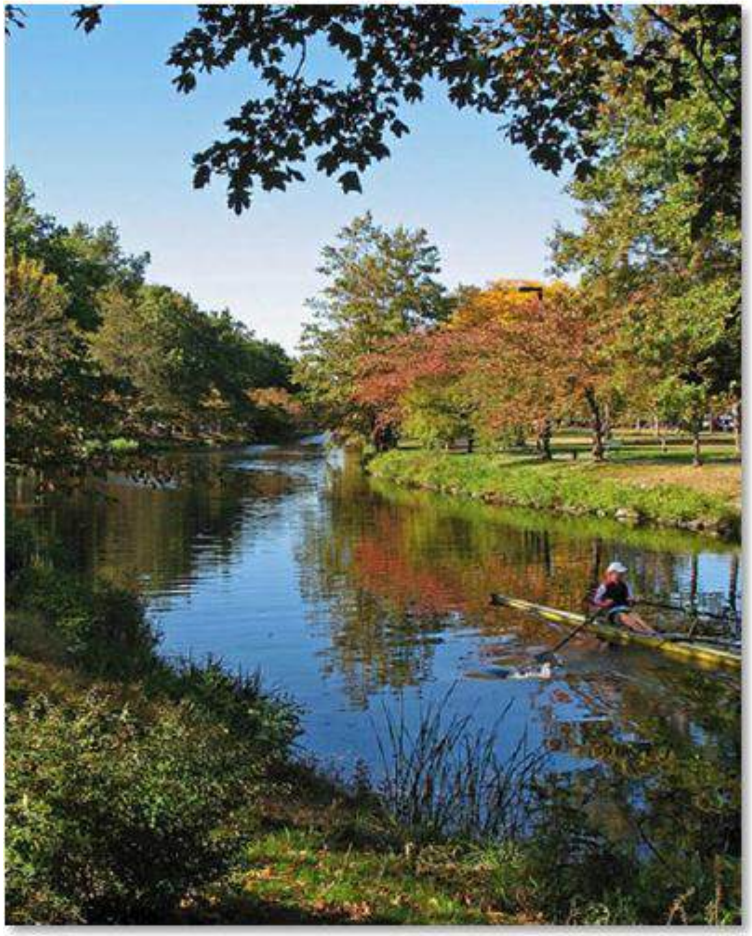
Different energy functions



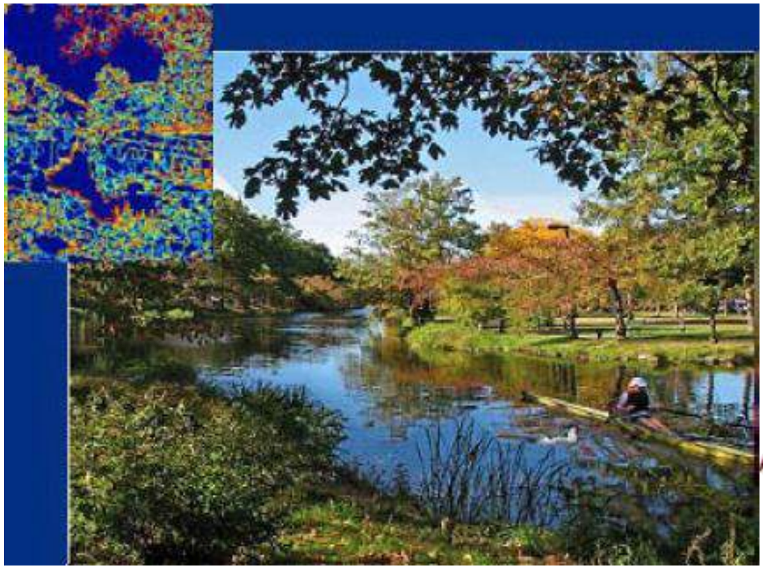
- Histogram of Gradients
- Entropy
- E1
- Mean Shift & E_1

[Shamir and Sorkine 2009]

Different energy functions



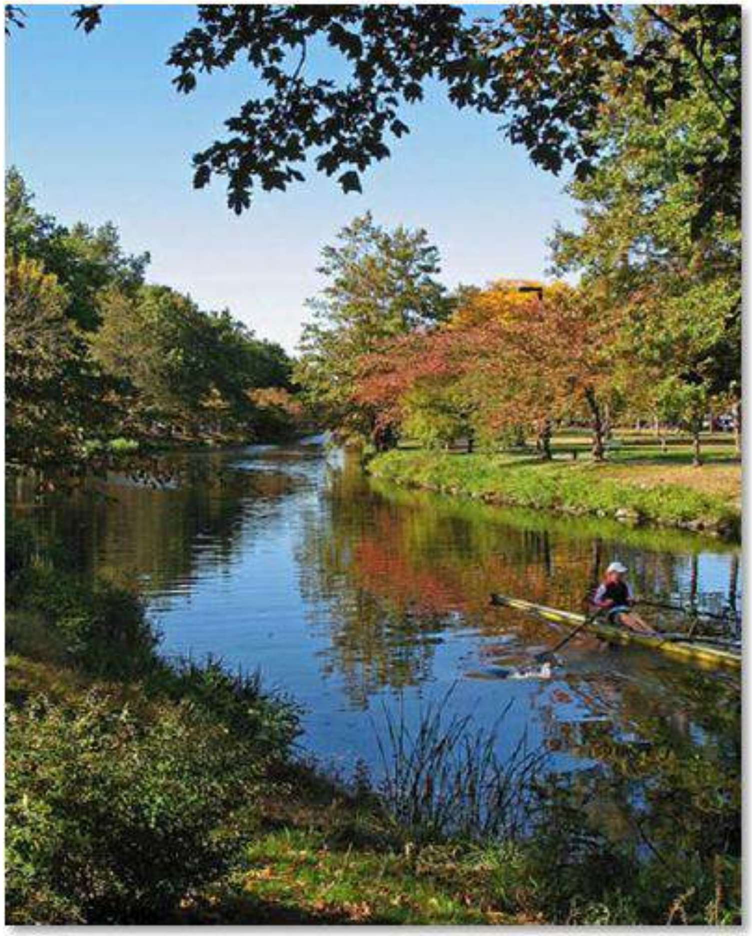
- Histogram of Gradients
- Entropy
- E1
- Mean Shift & E_1



[Shamir and Sorkine 2009]

Different energy functions

- Histogram of Gradients
- Entropy
- E1
- Mean Shift & E_1



[Shamir and Sorkine 2009]

Simple operators: cropping

- Crop s.t. important (salient) parts remain
- Use domain-specific tools, such as face detector, gaze estimation… [DeCarlo and Santella 2002; Viola and Jones 2004]

original crop

[Shamir and Sorkine 2009]

Simple operators: scaling

- Cam combine with cropping techniques (done on modern TV sets – center remains, peripheral data is scaled)
- Distorts content but is perfectly temporally coherent (video)

original squeeze hybrid

[Shamir and Sorkine 2009]

Discrete vs continuous

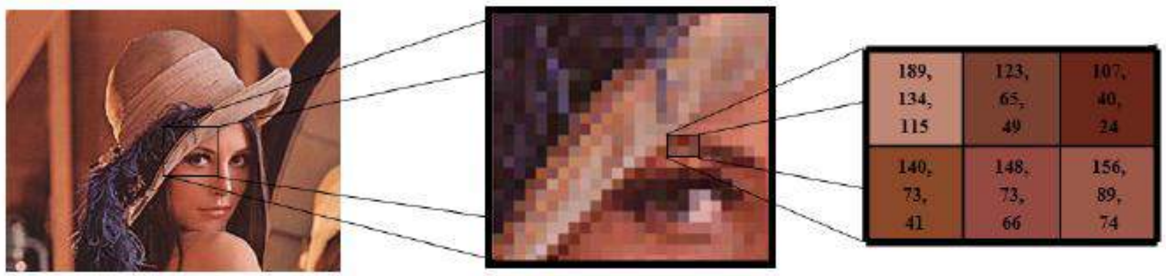


Figure 2: *A digital image as a 2D discrete grid of pixels. In this case the cells contain 3 values of RGB color.*

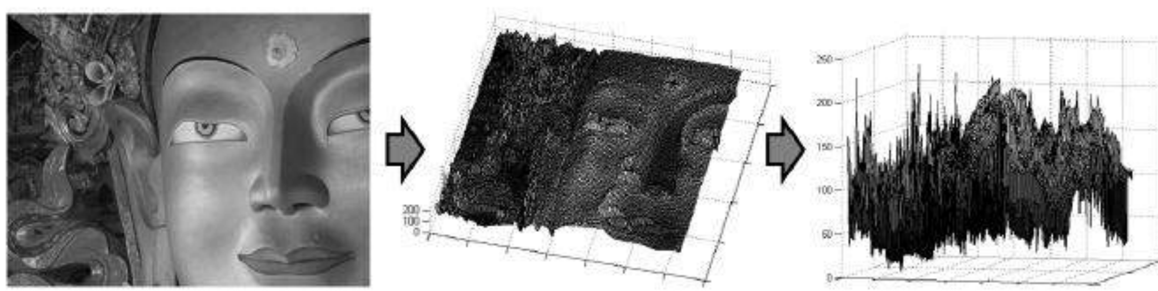


Figure 3: *A digital image as a sampling of a continuous function.*

[Shamir and Sorkine 2009]

Problem statement

- Given an image **I** of size (*n x m)*, we want to produce an image **I*** of size (*n* x m*)* which is a good representative of image **I**
- **But what is a "good representative"?** No definitions exist
- Goals of a retargeting algorithm:
  - 1. The important *content* of **I** should be preserved in **I***.
  - 2. The important *structure* of **I** should be preserved in **I***.
  - 3. **I*** should be *artifact*-free

Discrete approaches

- **Seam carving for content aware image resizing**
  SIGGRAPH 2007
  S. Avidan and A. Shamir
- **Improved seam carving for video retargeting**
  SIGGRAPH 2008
  M. Rubinstein, A. Shamir and S. Avidan
- **Seam carving for Media Retargeting**
  Trans. Of the ACM
  S. Avidan and A. Shamir
- **Multi-Operator Media Retargeting**
  SIGGRAPH 2009
  M. Rubinstein, A. Shamir and S. Avidan

Continuous approaches

- **Feature-aware textureing**

  EGSR 2006

  R. Gal, O. Sorkine and D. Cohen-Or
- **Non-homogeneous content-drive video retargeting**

  ICCV 2007

  L. Wolf, M Guttmann and D. Cohen-Or
- **Optimized scale-and-stretch for image resizing**

  SIGGRAPH ASIA 2008

  Y. Wang, C. Tai, O. Sorkine and T. Lee
- **Shrinkability maps for content-aware video resizing**

  Pacific Graphics 2008

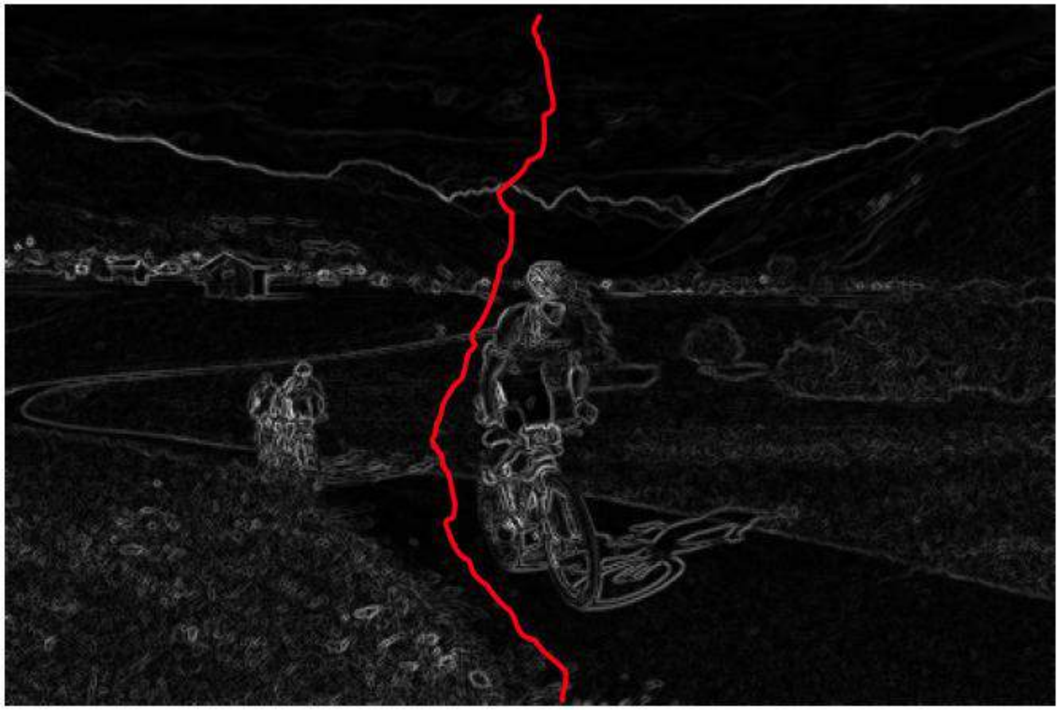
  Y. Zhang, S. Hu and R. Martin

[Shamir and Sorkine 2009]

Discrete example: Seam carving

[Rubinstein, Avidan and Shamir 2007]

Seam carving



[Rubinstein, Avidan and Shamir 2007]

Seam carving

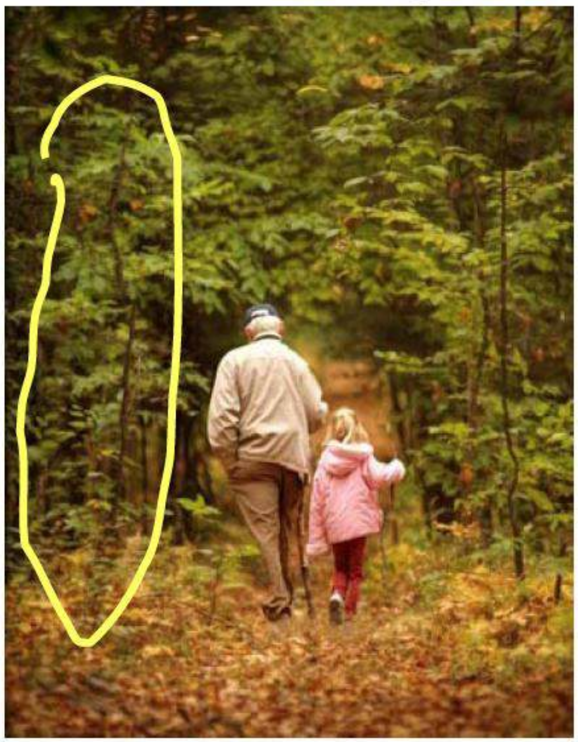
[Rubinstein, Avidan and Shamir 2007]

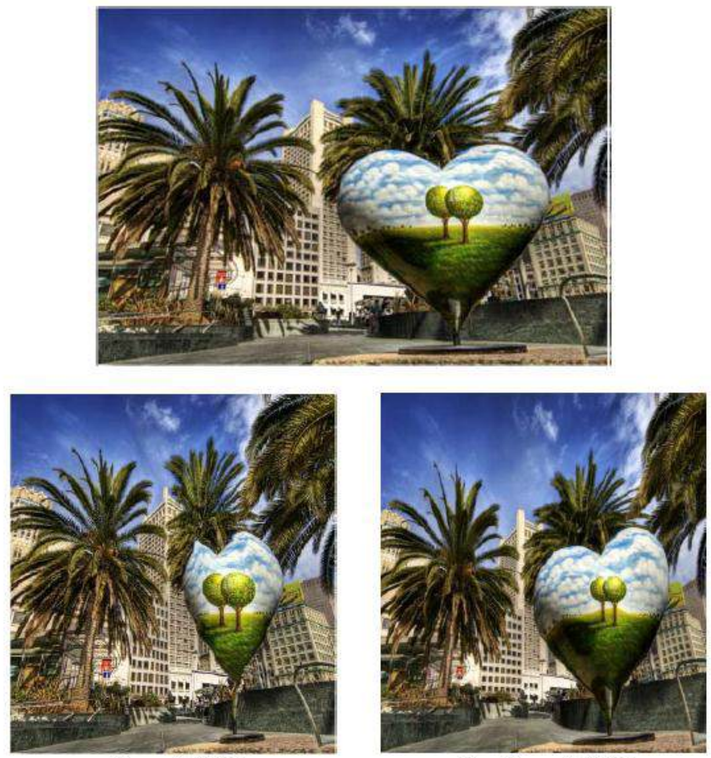
Seam carving

[Rubinstein, Avidan and Shamir 2007]

Seam carving: problems

- Discrete and greedy – may break structures
- Can run out of good seams in one direction



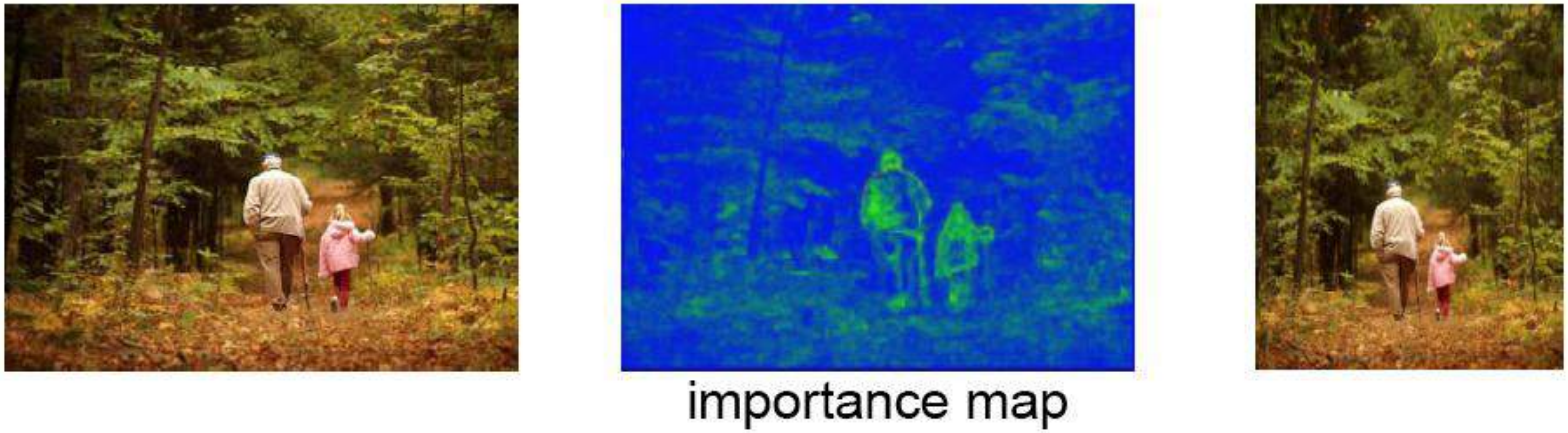


direct SC indirect SC

[Rubinstein, Avidan and Shamir 2007]

Continuous example: Warping

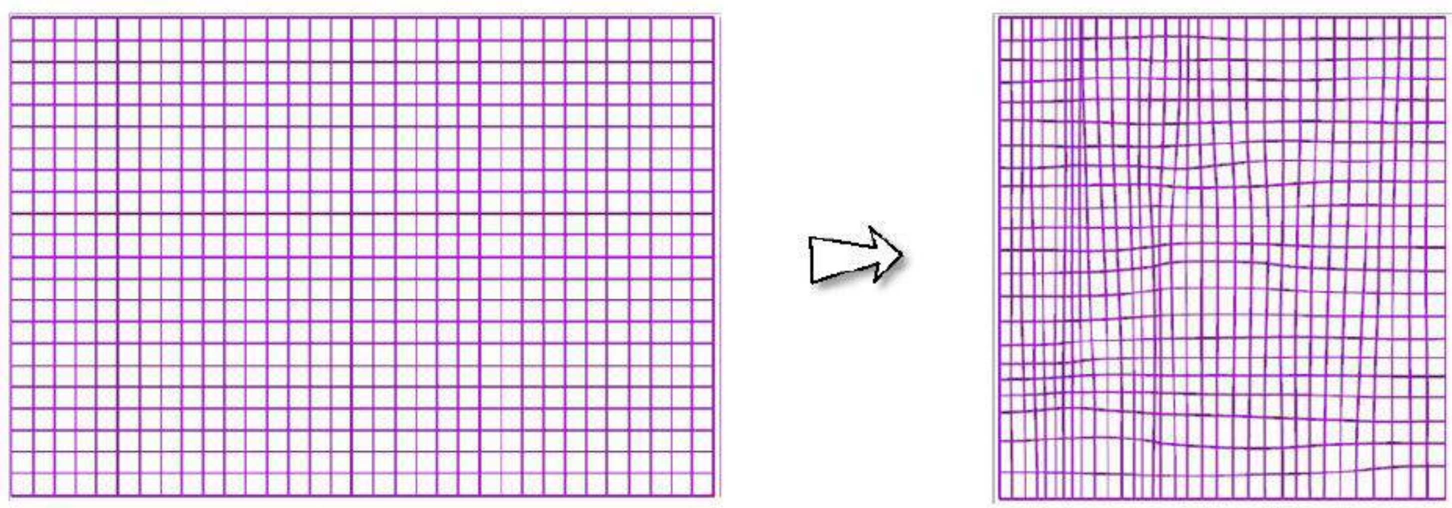
- Allow important regions to **uniformly scale**
- Find **optimal** local scaling factors by global optimization
- Result: preserve the **shape** of important regions, distort non-important ones



[Wang, Tai, Sorkine and Lee 2008]

Continuous example: Warping

- Grid mesh, preserve the shape of the important quads

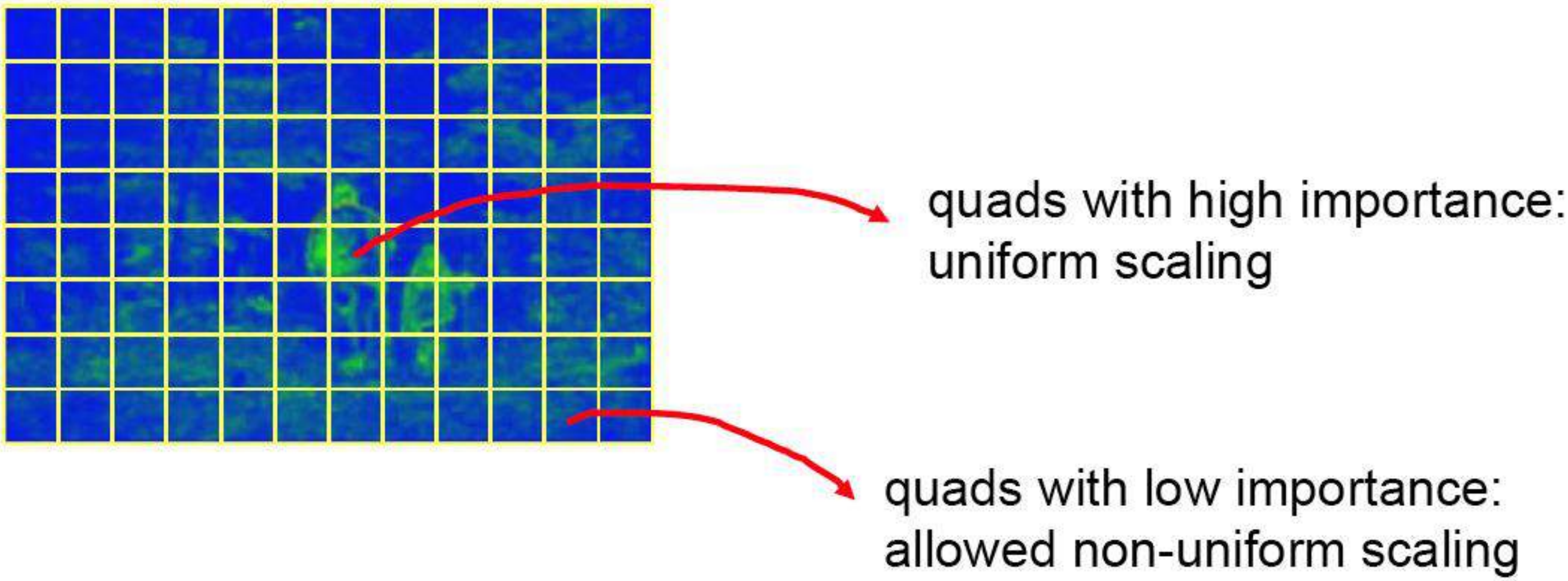


- Optimize the location of mesh vertices, interpolate image

[Wang, Tai, Sorkine and Lee 2008]

Continuous example: Warping

- Grid mesh, preserve the shape of the important quads



- Optimize the location of mesh vertices, interpolate image

[Wang, Tai, Sorkine and Lee 2008]

Video?

- Naïve… every frame by itself

[Shamir and Sorkine 2010]

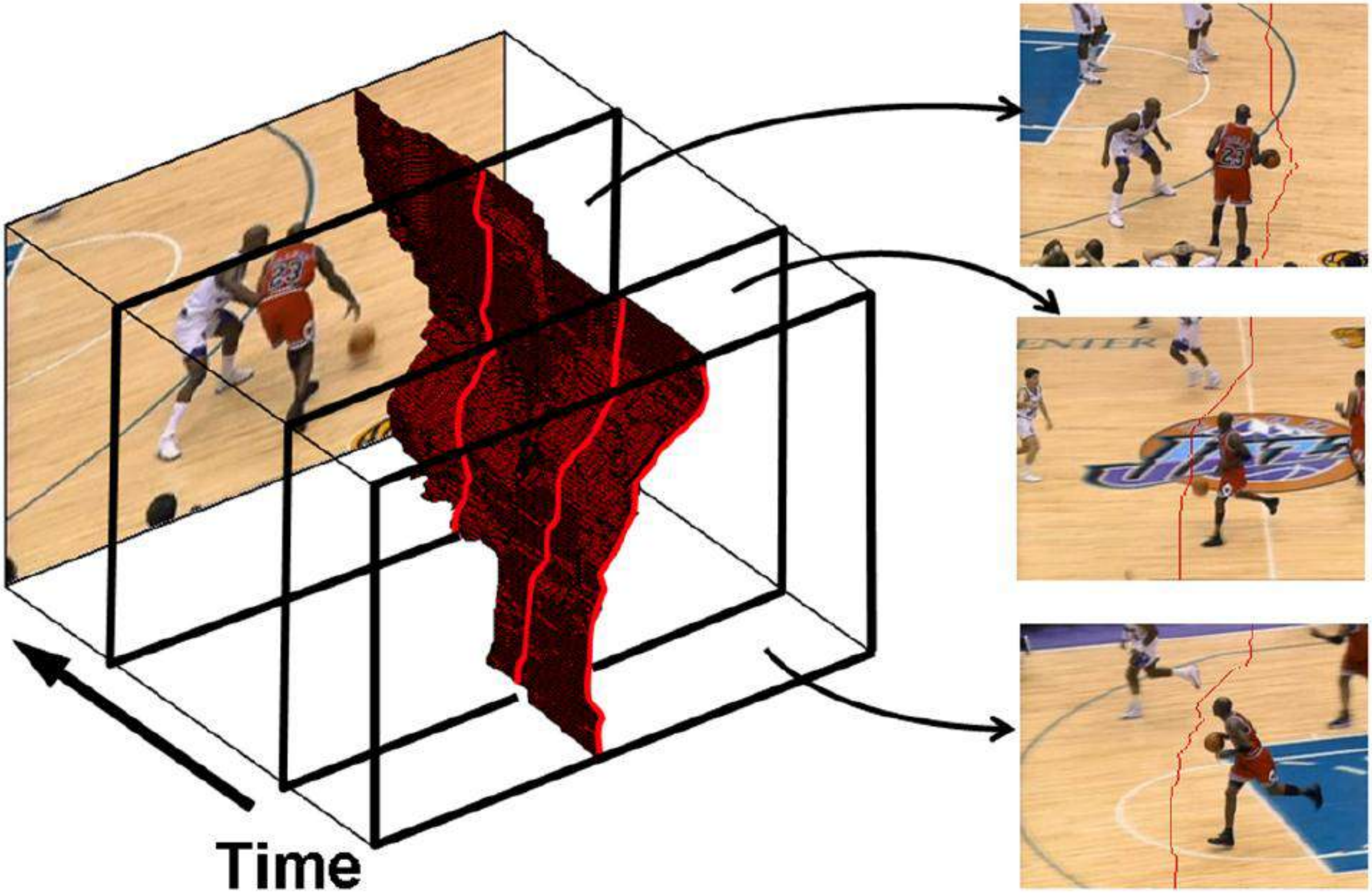
Video?

- Camera movement
- Object movement
- Seams should adapt and change through time!

- ➔ Global Solution (video cube)

[Shamir and Sorkine 2010]

Video?



[Shamir and Sorkine 2010]

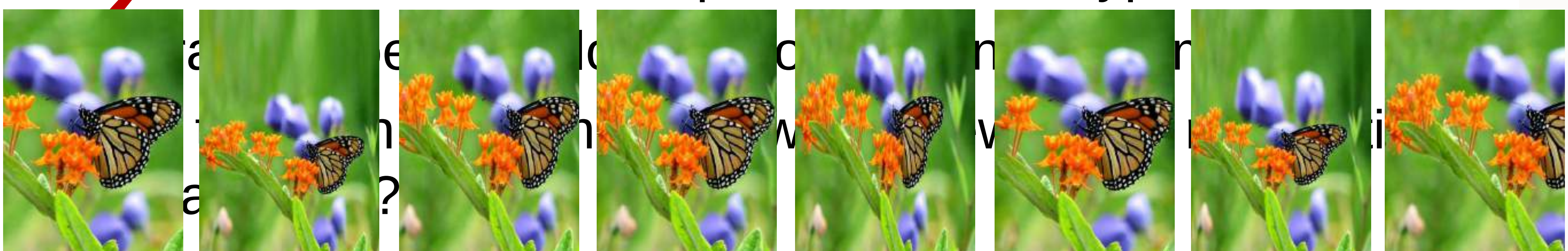
Current State of Retargeting Research

SIGGRAPHASIA2011 HONG KONG

- No clear evaluation methodology!
  - Mostly visual comparison
  - Small subset of previous techniques

Source

- Relation between the operator and the type of content?



- Computational retargeting measure?

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

- Benchmark and evaluation methodology for image retargeting

RetargetMe

http://people.csail.mit.edu/mrub/retargetme/

- Comprehensive perceptual study and analysis of image retargeting

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Goals

- What is the “correct” way to retarget this image?

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Goals

- The dataset and user study
- User response (subjective) analysis
  - Is there consensus between viewers?
  - When is one method better than another?
- Computational (objective) analysis
  - Can an image distance measure predict retargeting quality?

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Constructing the Dataset

- Image Retargeting objectives:
  1. Preserve the important ***content*** and ***structures***
  2. Limit ***artifacts***

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Retargeting Operators

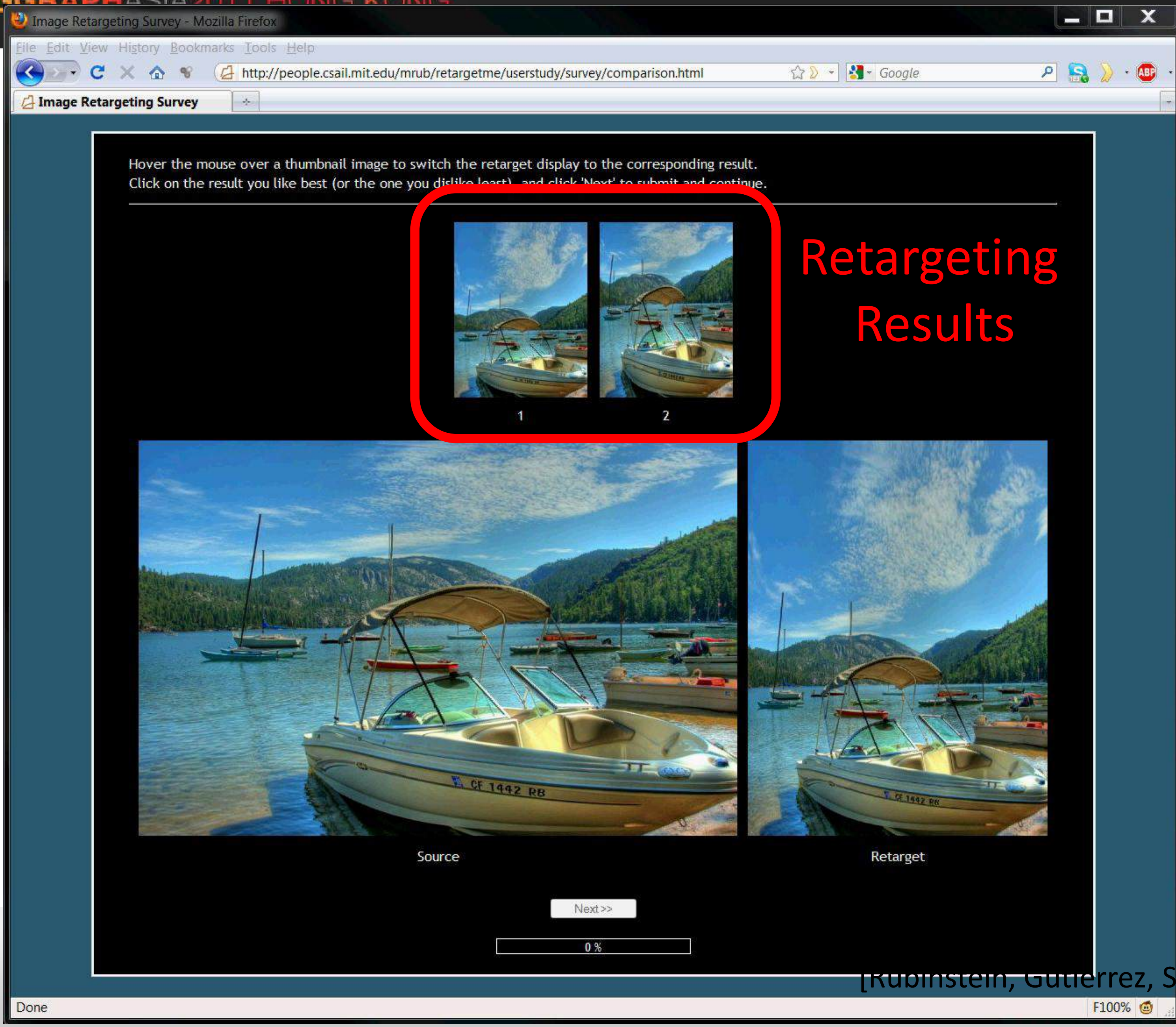
Operator	Reference	Category
Seam Carving **[SC]**	[Rubinstein et al. 2008]	Discrete
Shift Map **[SM]**	[Pritch et al. 2009]	Discrete
Multi-Operator **[MULTIOP]**	[Rubinstein et al. 2009]	Discrete
Warping **[WARP]**	[Wolf et al. 2007]	Continuous
Streaming Video **[SV]**	[Krähenbühl et al. 2009]	Continuous
Scale-and-Stretch **[SNS]**	[Wang et al. 2008]	Continuous
Cropping **[CR]**	[Manual]	Reference
Scaling **[SCL]**	[Cubic interpolation]	Reference

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Comparative Analysis

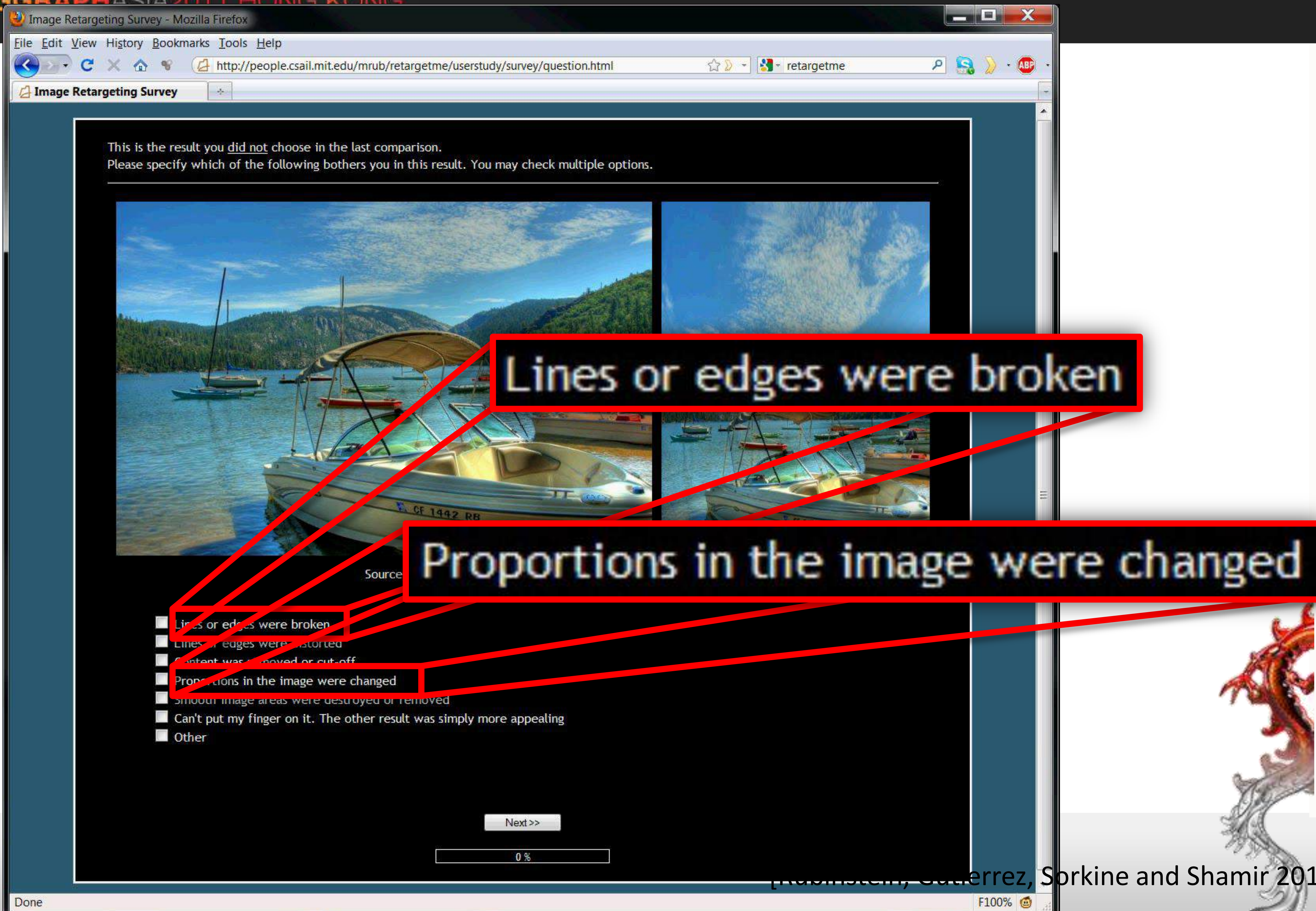
[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

The Survey Interface



[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Additional Questions



[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

User Statistics

- Each participant performs **12** comparisons over **5** images

- **210** participants; **252** votes per image
  - Half amazon mechanical turk Artificial Artificial Intelligence
  - Half (25 cents per completed survey)

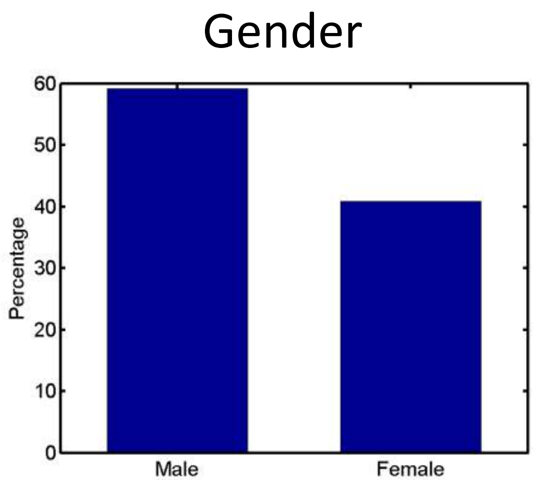
- Average time to complete: 20 minutes

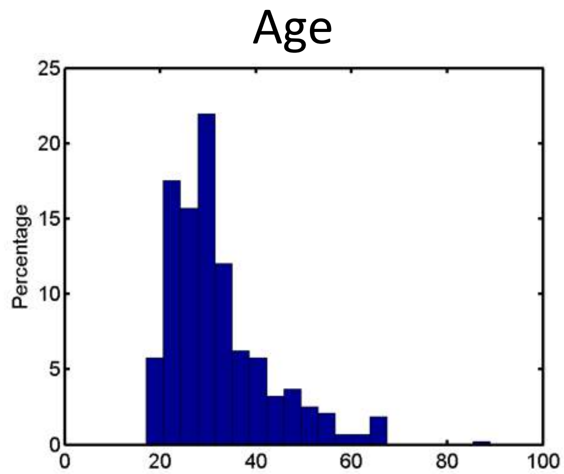
“It was a very interesting survey. Very nice experience”

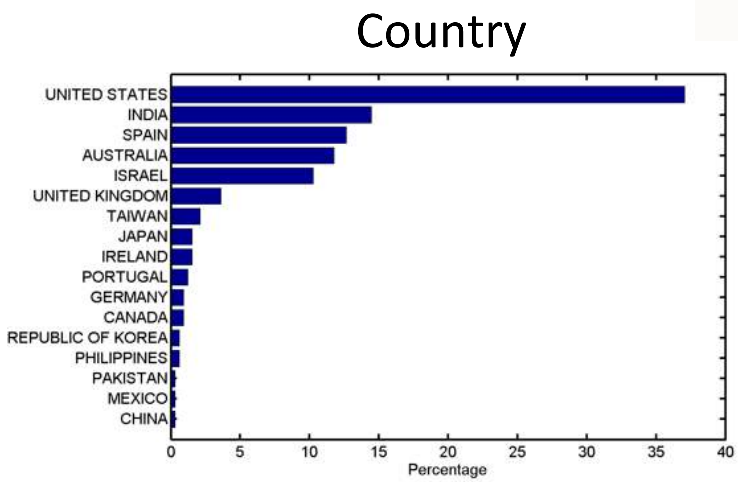
“i need your more survey so that i can help u a lot”

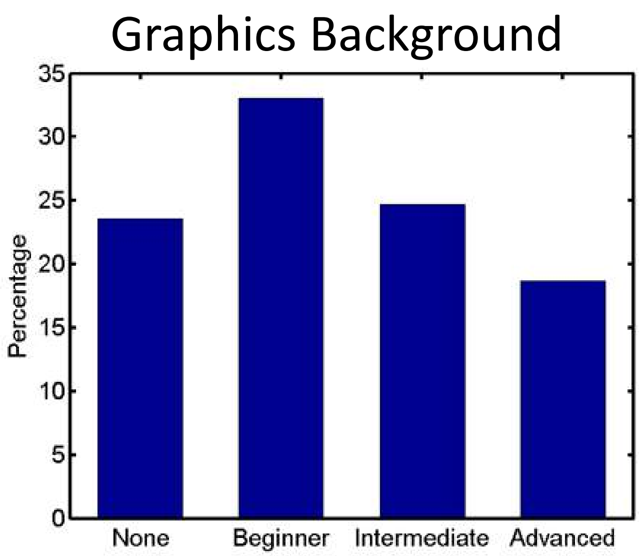
[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

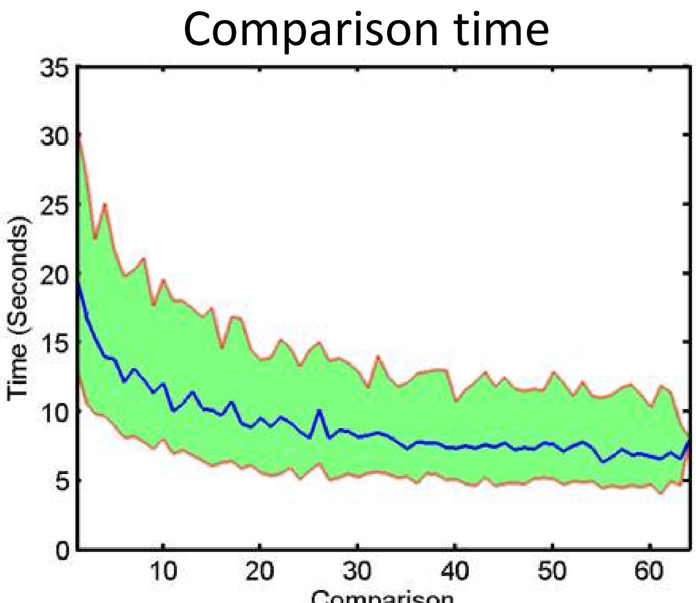
User Statistics











[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

User Agreement

- Similarity of votes = consensus on “good” retargeting

- *Coefficient of Agreement* [Kendall 1940]

$$u = \frac{2\Sigma}{\binom{m}{2}\binom{t}{2}} - 1, \qquad \Sigma = \sum_{i=1}^{t}\sum_{j=1}^{t}\binom{a_{ij}}{2}$$

- a_{ij} = # times method i chosen over method j
- m = # participants
- t = 8 (# retargeting operators)
- $u \in [-\frac{1}{m}, 1]$

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

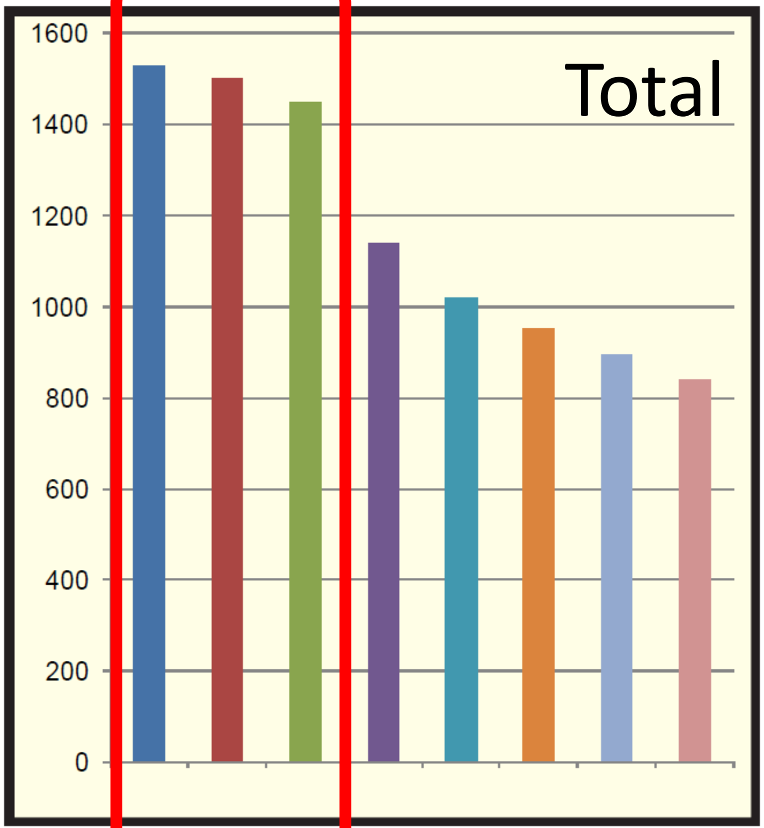
User Agreement

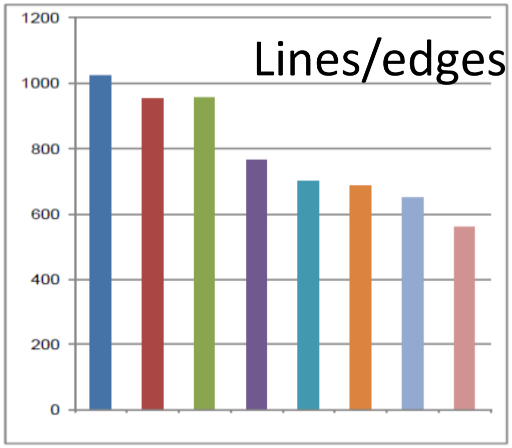
	lines/ edges	faces/ people	Texture	foreground objects	Geometric Structures	Symmetry	Total
u	0.073	**0.166**	0.070	**0.146**	0.084	**0.132**	**0.095**

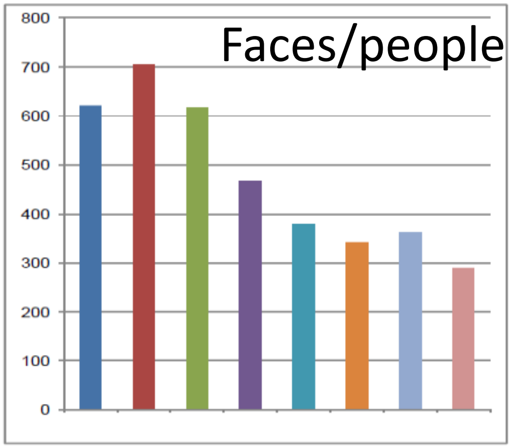
- Low agreement in general
- Greater agreement on images containing faces/people, evident foreground objects and symmetry.

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Operator Ranking



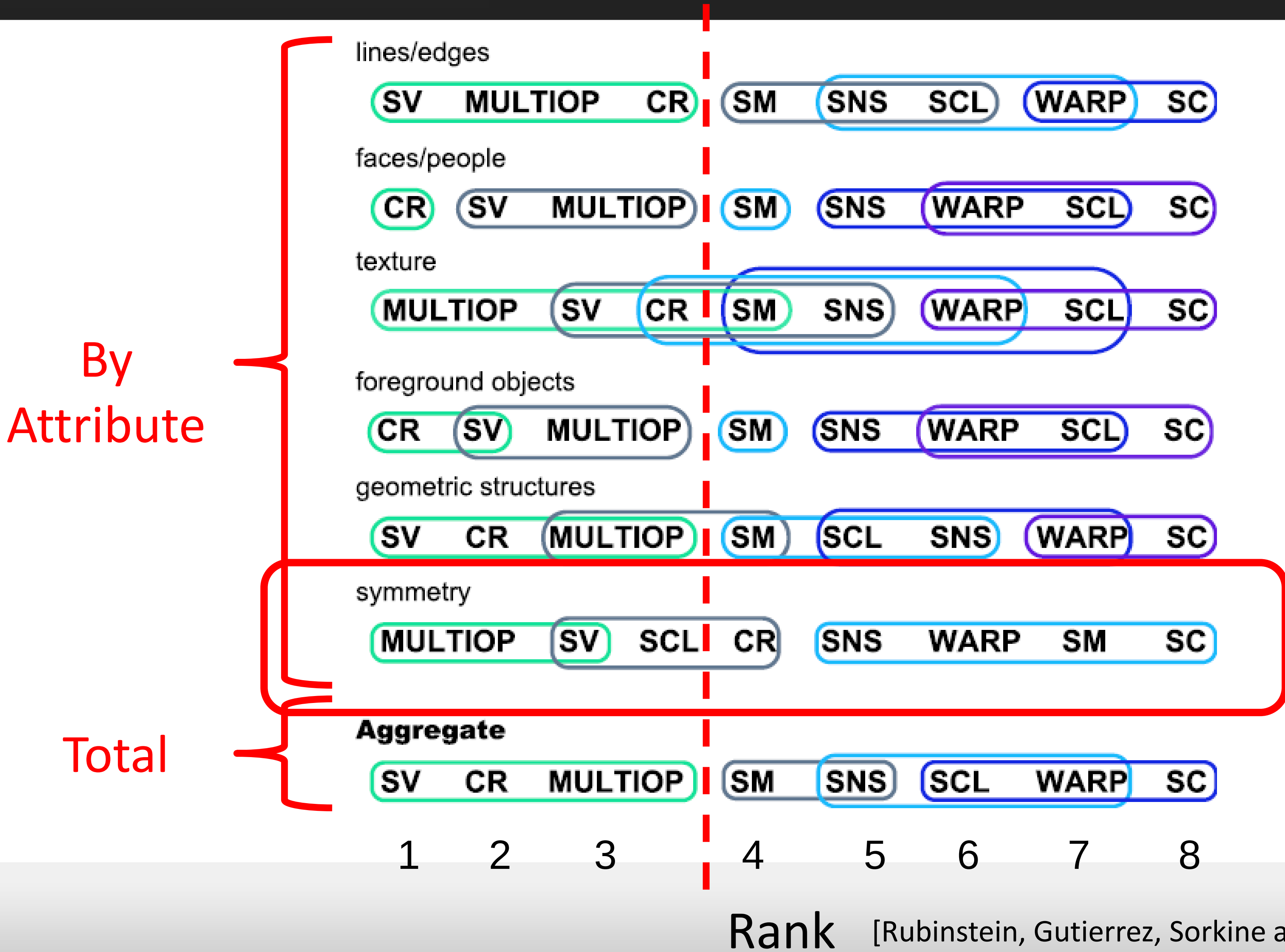




SV	CR	MULTIOP	SM	SNS	SCL	WARP	SC
Streaming Video	Cropping	Multi-operator	Shift-maps	Scale & Stretch	Scaling	Nonhomo. Warping	Seam Carving

SIGGRAPHASIA2011 HONG KONG

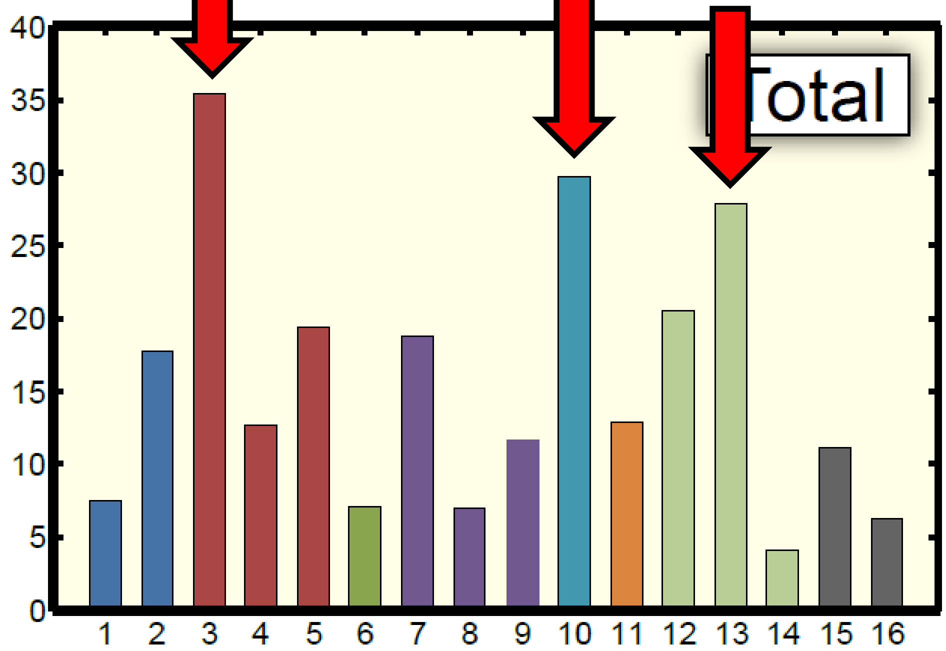
Operator Ranking

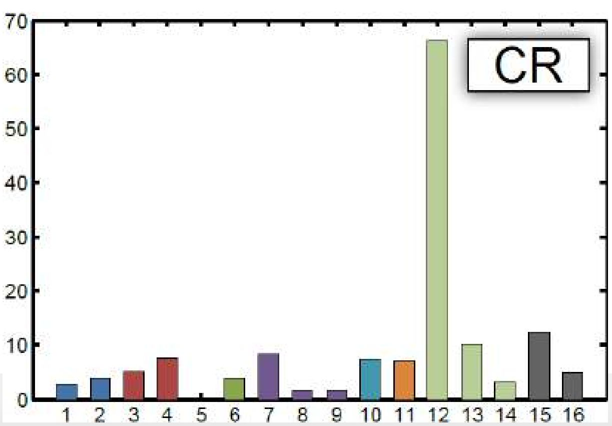


[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

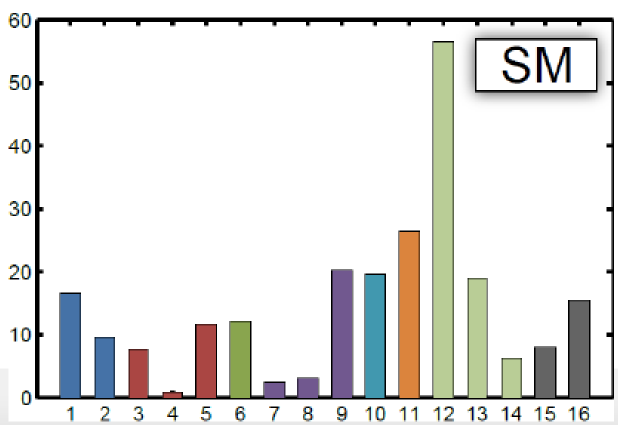
Additional Questions

Attribute	Reason	ID
lines/edges	Lines or edges were broken	1
lines/edges	Lines or edges were distorted	2
faces/people	People or faces were squeezed	3
faces/people	People or faces were stretched	4
faces/people	People or faces were deformed	5
texture	Textures were distorted	6
foreground objects	Foreground objects were squeezed	7
foreground objects	Foreground objects were stretched	8
foreground objects	Foreground objects were deformed	9
geometric structures	Geometric structures were distorted	10
symmetry	Symmetry was violated	11
Common	Content was removed or cut-off	12
Common	Proportions in the image were changed	13
Common	Smooth image areas were destroyed or removed	14
Common	Can't put my finger on it.	
	The other result was simply more appealing	15
Common	Other	16

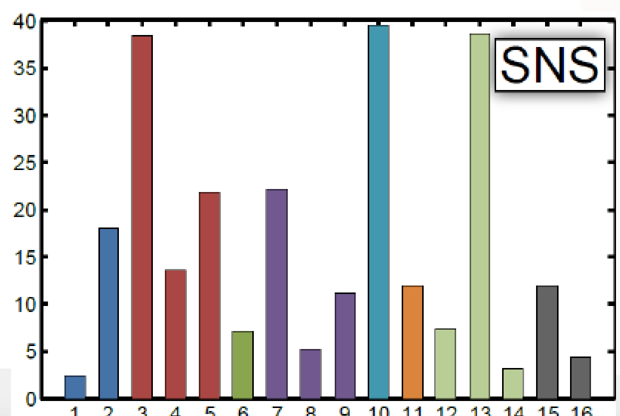




Cropping



Shift-maps



Scale & Stretch

Partial Conclusion

(At least for our retargeted setup)

SUBJECTIVE:

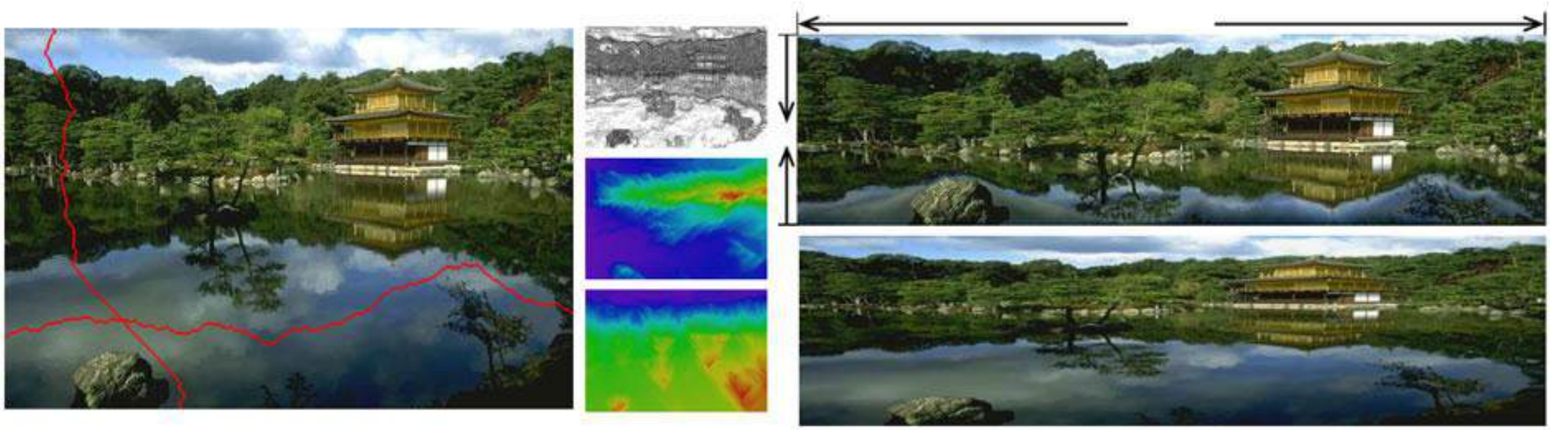
Clear *and consistent* division in groups

CR, SV, MULTIOP: good!

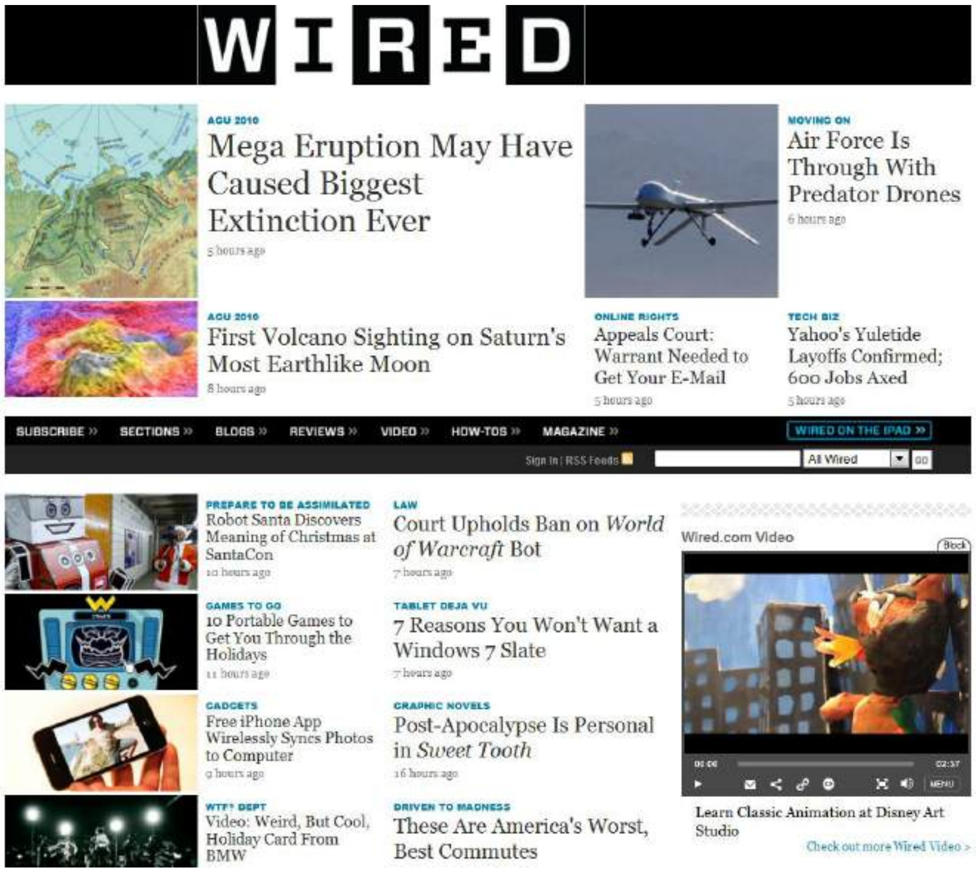
SCL, SC, WARP: not so good

Greater agreement for *faces/people* and *foreground objects:*

Saliency at object level?



Source is Usually Unknown!

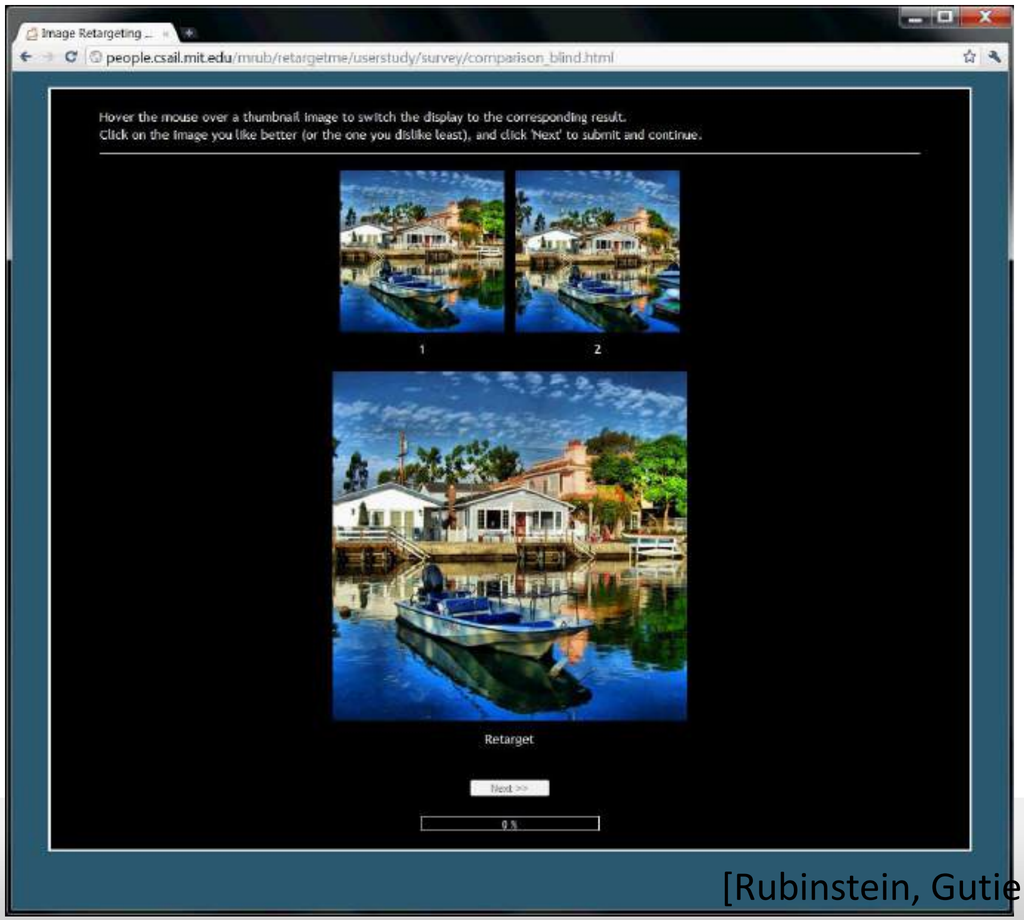


[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

“No Reference” Experiment Results

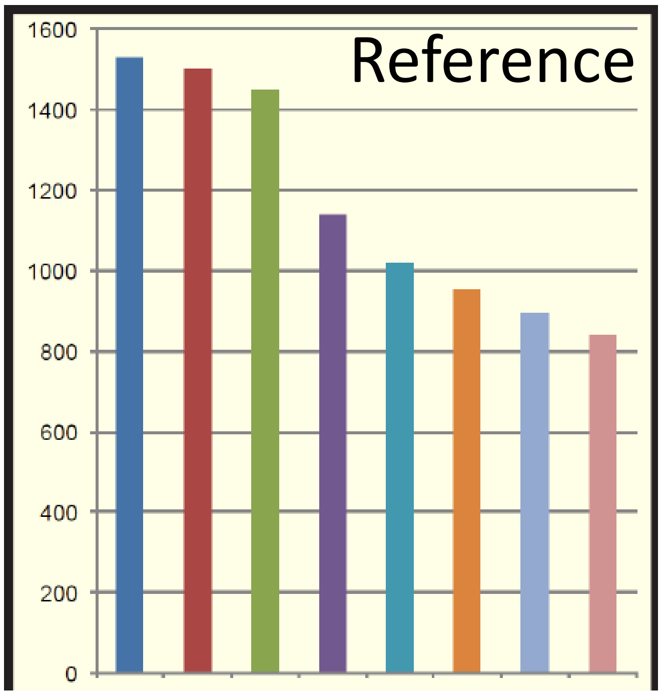
SIGGRAPHASIA2011 HONG KONG

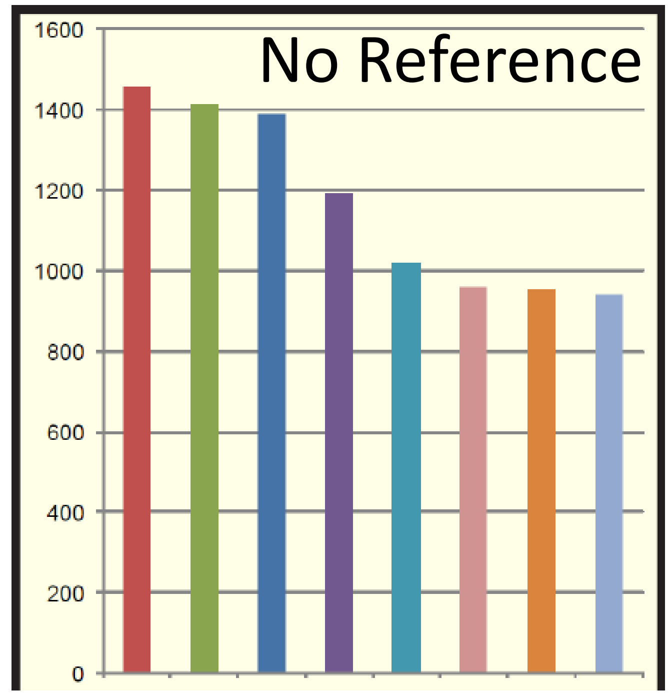
- Similar setup, source image **not shown**
- **New** set of 210 participants



[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

"No Reference" Experiment Results





SV Streaming Video | CR Cropping | MULTIOP Multi-operator | SM Shift-maps | SNS Scale & Stretch | SCL Scaling | WARP Nonhomo. Warping | SC Seam Carving

lines/ edges	faces/ people	texture	foreground objects	geometric structures	symmetry	Aggregate	Rank product
0.964	0.988	0.946	0.737	0.950	0.957	0.978	0.985

Analysis of the users' responses: significance test

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Computational Retargeting Measures

- Goal: can computational image distance measures predict human retargeting preferences?
  - Can be used to evaluate new operators
  - Can be used to develop new operators – [Simakov et al. 2008], [Rubinstein et al. 2009]

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

(Non-blind) Retargeting Measures

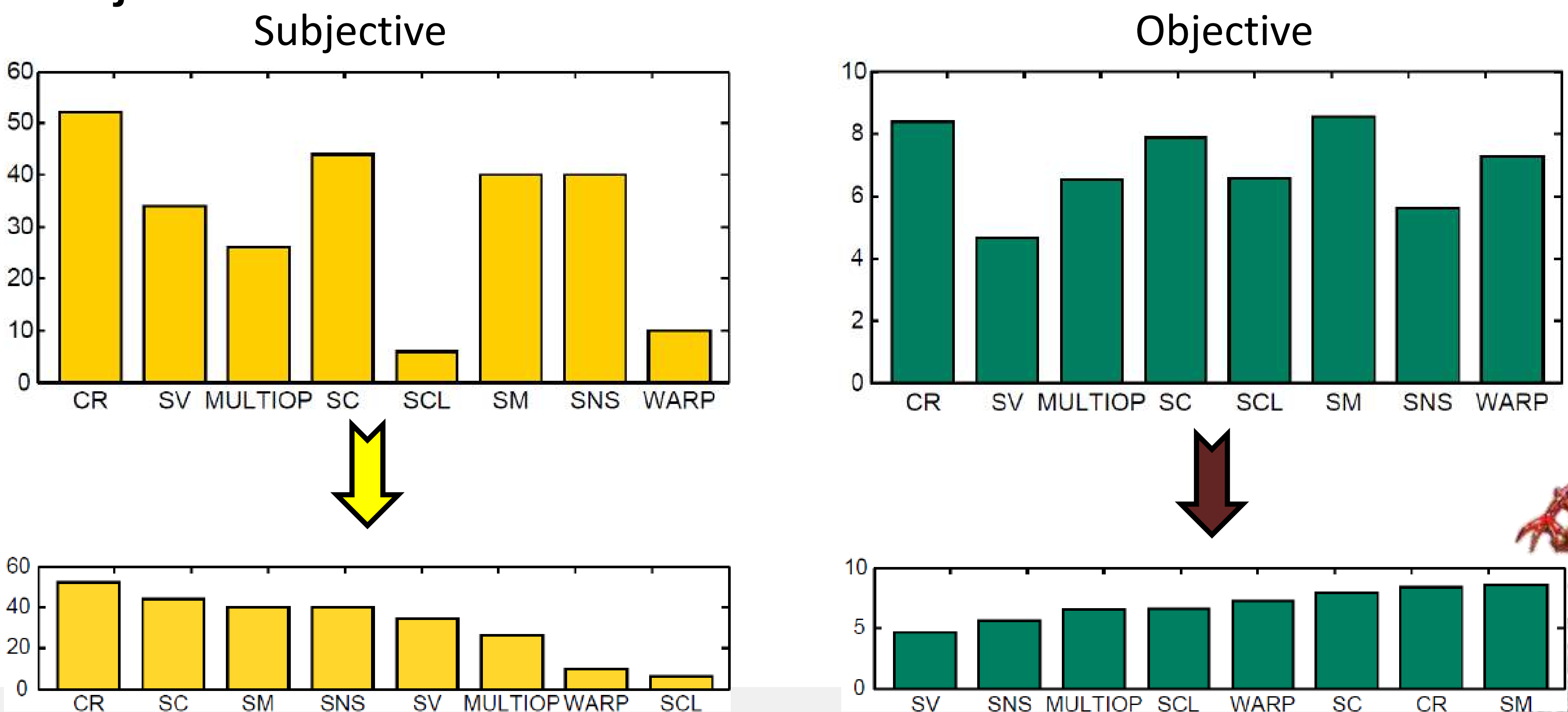
Objective Measures

- High level semantics:
  - Bidirectional Similarity [**BDS**] - Simakov et al. 2008
  - Bidirectional Warping [**BDW**] - Rubinstein et al. 2009
  - SIFT Flow [**SIFTflow**] – Liu et al. 2008
  - Earth Mover's Distance [**EMD**] - Pele and Werman 2009
- Low level features
  - Edge Histogram [**EH**] – Menjunath et al. 2001
  - Color Layout [**CL**] – Kasutani and Yamada 2001
- See dataset website and supplemental material for more details

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

How to Evaluate an Objective Measure?

- Define rate of agreement as the *correlation between rankings* induced by the user responses, and the objective measure



SIGGRAPHASIA2011 HONG KONG

Objective Analysis Results

Metric	lines/ edges	faces/ people	texture	Foreground objects	geometric structures	symmetry	total
BDS	0.04	0.19	0.06	0.17	0.00	-0.01	0.08
BDW	0.03	0.05	-0.05	0.06	0.00	0.12	0.05
EH	0.04	-0.08	-0.06	-0.08	0.10	0.30	0.00
CL	-0.02	-0.18	-0.07	-0.18	-0.01	0.21	-0.07
SIFTflow	0.10	0.25	**0.12**	0.22	0.08	0.07	0.14
EMD	**0.22**	**0.26**	0.11	**0.23**	**0.24**	**0.50**	**0.25**

- The results were spectacular(ly poor!)
- We tried something else:
  - SIFT-flow [Liu et al. 2008]: SIFT
  - Earth mover’s distance [Pele & Werman 2009]: EMD
- Somewhat better ☺

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Can computational image distance metrics predict human retargeting perception?

SIGGRAPHASIA2011 HONG KONG

Metric	Attribute						Total		
	Lines/Edges	Faces/People	Texture	Foreground Objects	Geometric Structures	Symmetry	Mean	std	p-value
BDS	0.040	0.190	0.060	0.167	-0.004	-0.012	0.083	0.268	0.017
BDW	0.031	0.048	-0.048	0.060	0.004	0.119	0.046	0.181	0.869
EH	0.043	-0.076	-0.060	-0.079	0.103	0.298	0.004	0.334	0.641
CL	-0.023	-0.181	-0.071	-0.183	-0.009	0.214	-0.068	0.301	0.384
RAND	-0.046	-0.014	0.048	-0.032	-0.040	0.143	-0.031	0.284	0.693
SIFTflow	0.097	0.252	**0.119**	0.218	0.085	0.071	0.145	0.262	0.031
EMD	**0.220**	**0.262**	0.107	**0.226**	**0.237**	**0.500**	**0.251**	0.272	1e-5

(a) Complete rank correlation ($k = \infty$)

Metric	Attribute						Total		
	Lines/Edges	Faces/People	Texture	Foreground Objects	Geometric Structures	Symmetry	Mean	std	p-value
BDS	0.062	0.280	0.134	0.249	-0.025	-0.247	0.108	0.532	0.005
BDW	0.213	0.141	0.123	0.115	0.212	0.439	0.200	0.395	0.002
EH	-0.036	-0.207	-0.331	-0.177	0.111	0.294	-0.071	0.593	0.013
CL	-0.307	-0.336	-0.433	-0.519	-0.366	0.088	-0.320	0.543	1e-6
SIFTflow	0.241	**0.428**	**0.312**	**0.442**	**0.303**	0.002	0.298	0.483	1e-6
EMD	**0.301**	0.416	0.216	0.295	0.226	**0.534**	**0.326**	0.496	1e-6

(b) Rank correlation with respect to the three highest rank results ($k = 3$).

Table 6: *Correlation of objective and subjective measures for the complete rank (top) and for the three highest ranked results (bottom). In each column the mean τ correlation coefficient is shown ($-1 \le \tau \le 1$), calculated over all images in the dataset with the corresponding attribute. The last three columns show the mean score, standard deviation, and respective p-value over all image types. Highest score in each column appears in bold.*

[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

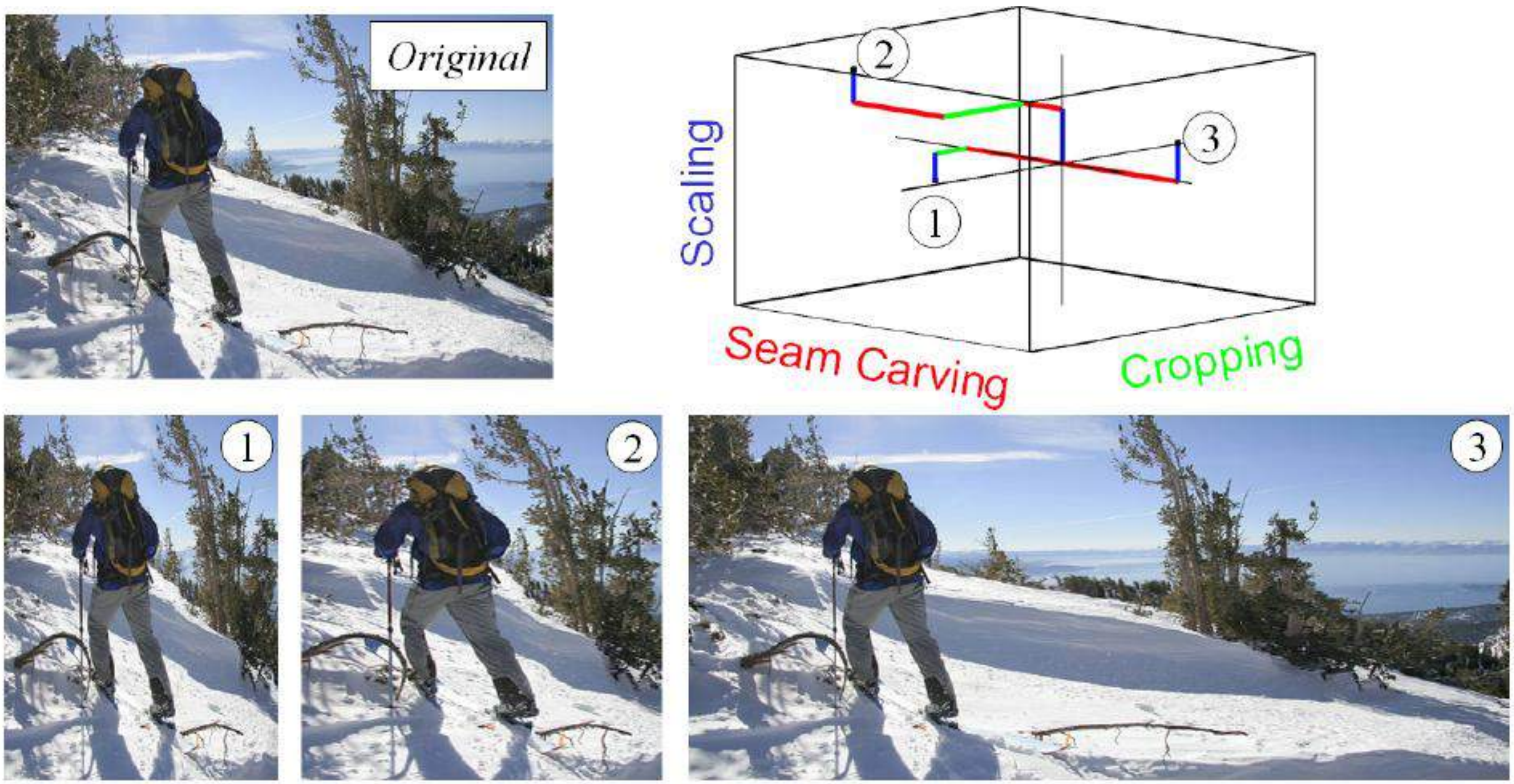
Conclusions

SUBJECTIVE:
More recent algorithms **do** outperform their predecessors in a (surprisingly) consistent way

Cropping is the simplest and one of the best:
- loss of info OK
- distortion **NOT** OK
- bring it back!

Interestingly, scaling and seam carving do not do very well on their own… but are two of the three in MULTIOP:
- *combination* of simple methods?

Conclusions



Conclusions

OBJECTIVE:
We are a long way from predicting human perception

Four similarity image metrics did not perform well at all

Two metrics not originally designed for that purpose did somewhat better

Optimize retargeting wrt those?

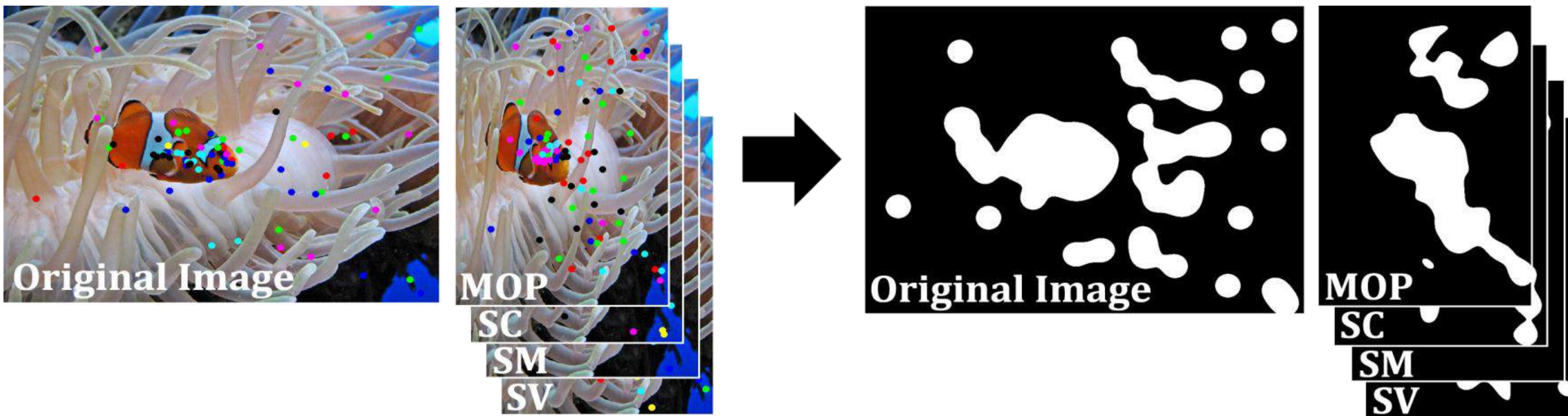
Further research is (badly!) needed

Conclusions

We need **video** analysis and experiments!

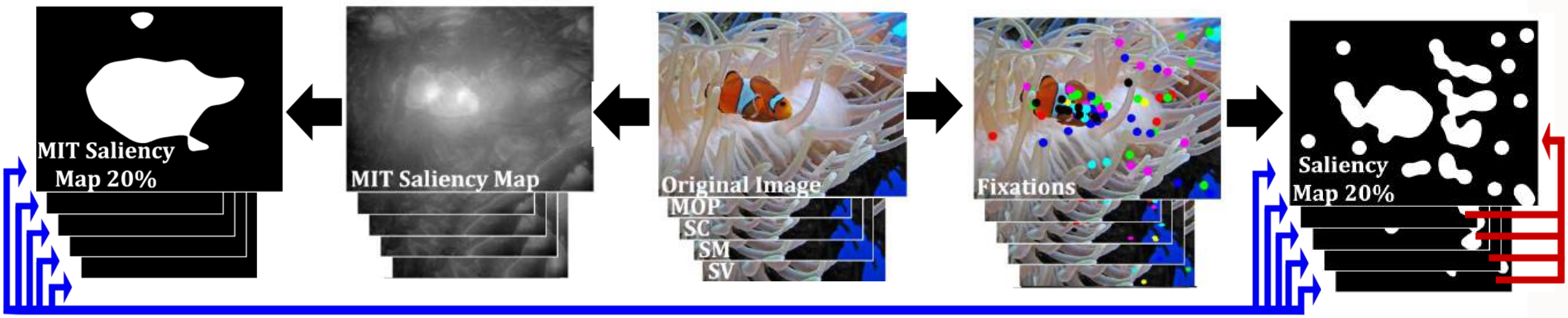
$$ColSim(C^0_{ori}, C^0_{ret}) = w_L SalSim(L^{*0}_{ori}, L^{*0}_{ret}) + w_a SalSim(a^{*0}_{ori}, a^{*0}_{ret}) + w_b SalSim(b^{*0}_{ori}, b^{*0}_{ret})$$

Using Eye-Tracking to Assess Different Image Retargeting Methods



[Castillo, Judd and Gutierrez 2011]

Overview

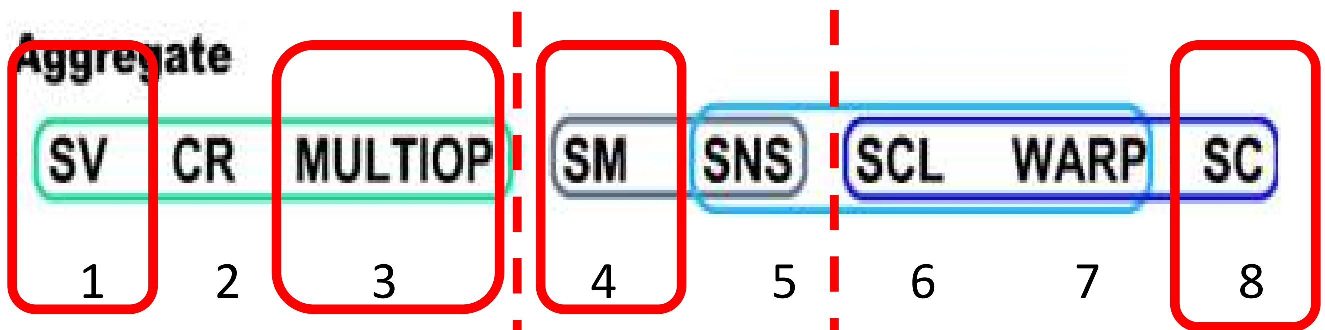


[Castillo, Judd and Gutierrez 2011]

Retargeting Operators

- Seam Carving **[SC]** [Rubinstein et al. 2008]
- Shift Maps **[SM]** [Pritch et al. 2009]

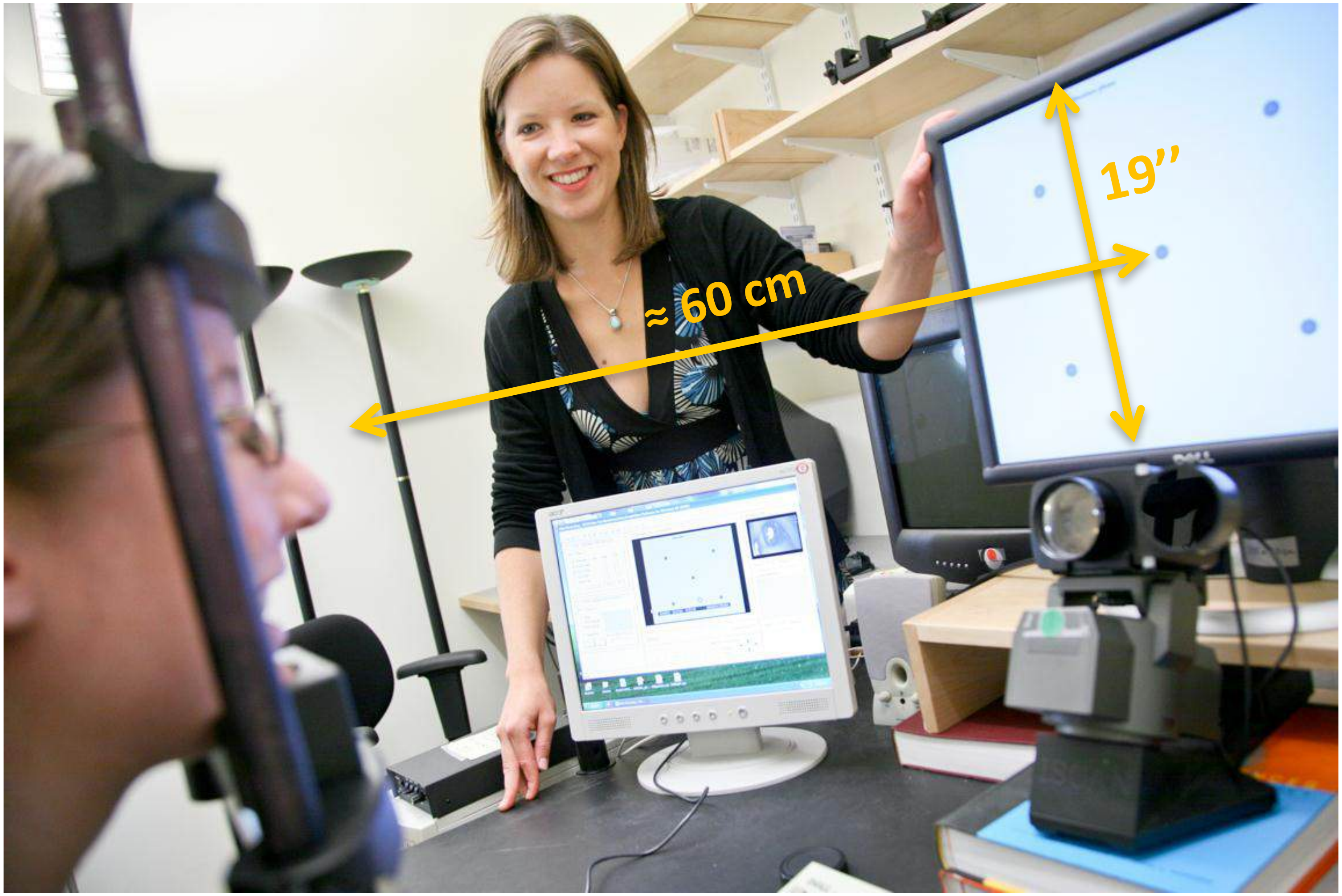
- Multi-Operator **[MULTIOP]** [Rubinstein et al. 2009]
- Streaming Video **[SV]** [Krähenbühl et al. 2009]



Rank **[Rubinstein et al. SIGAsia 2010]**

[Castillo, Judd and Gutierrez 2011]

Collect eye tracking data



[Photo Credit: Jason Dorfman CSAIL website]

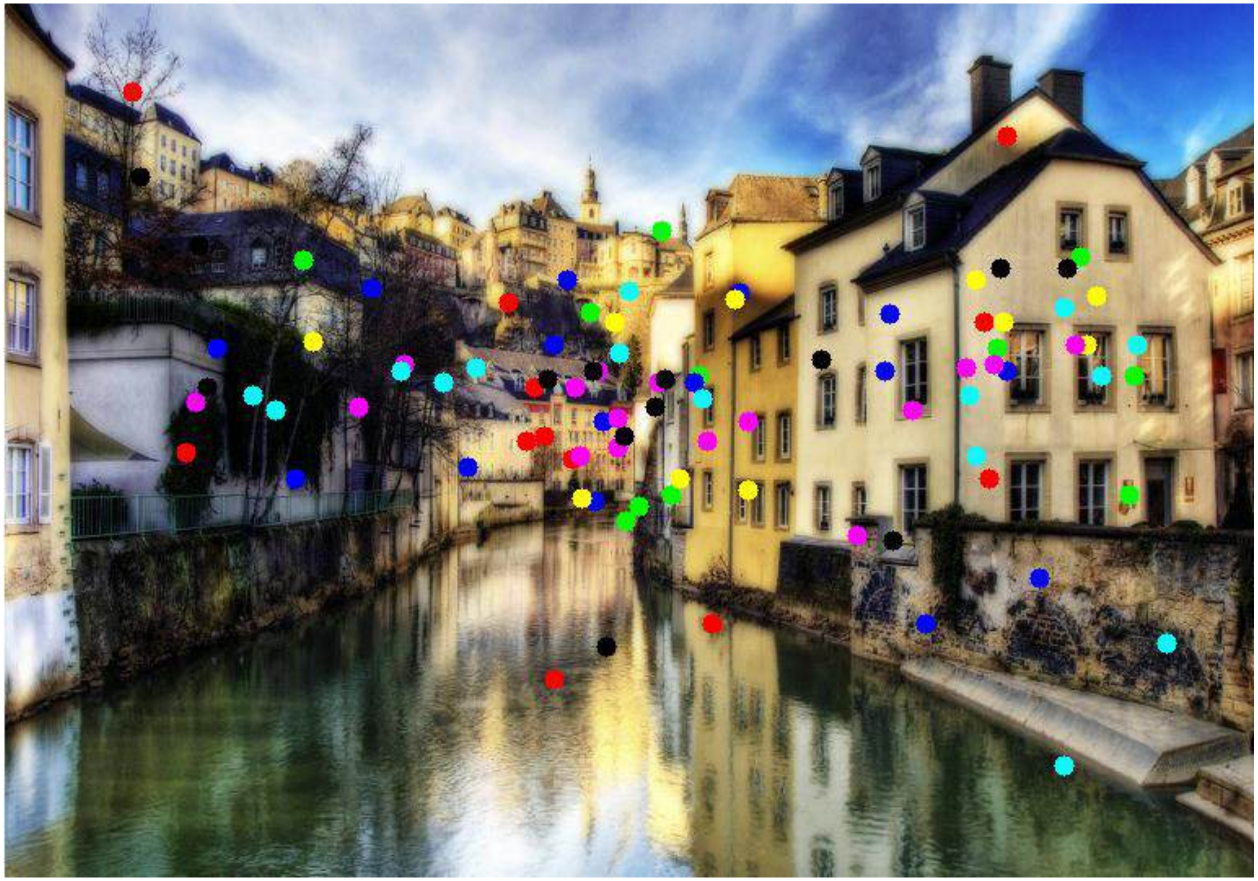
Screen resolution 1280x1024

Each image shown for 5 seconds

[Castillo, Judd and Gutierrez 2011]

Eye tracking data

Contextual guidance of eye movements and attention in real-world scenes: The role of global features on object search **[Torralba et al. 2006]**

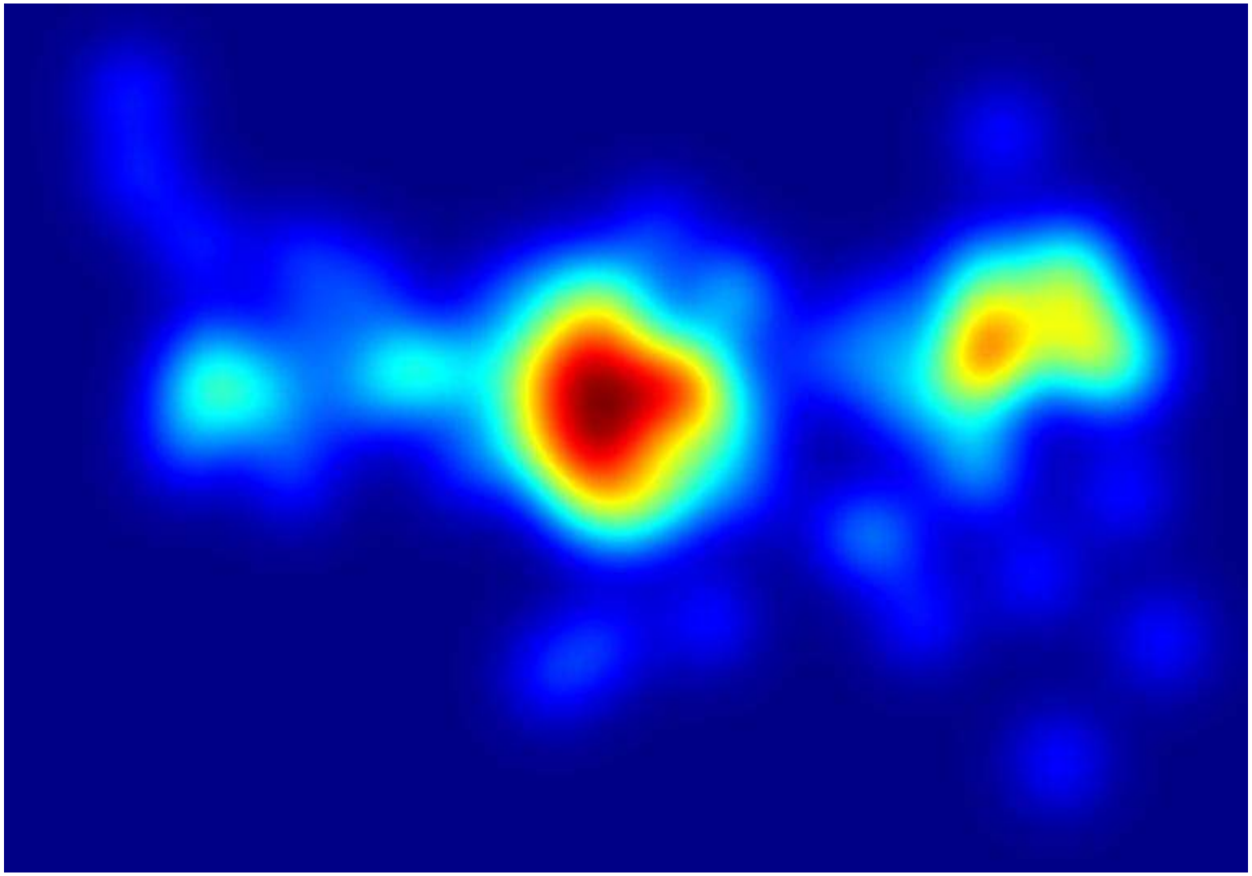


Fixations for 7 users

[Castillo, Judd and Gutierrez 2011]

Eye tracking data

Learning to predict where humans look **[Judd et al. 2009]**

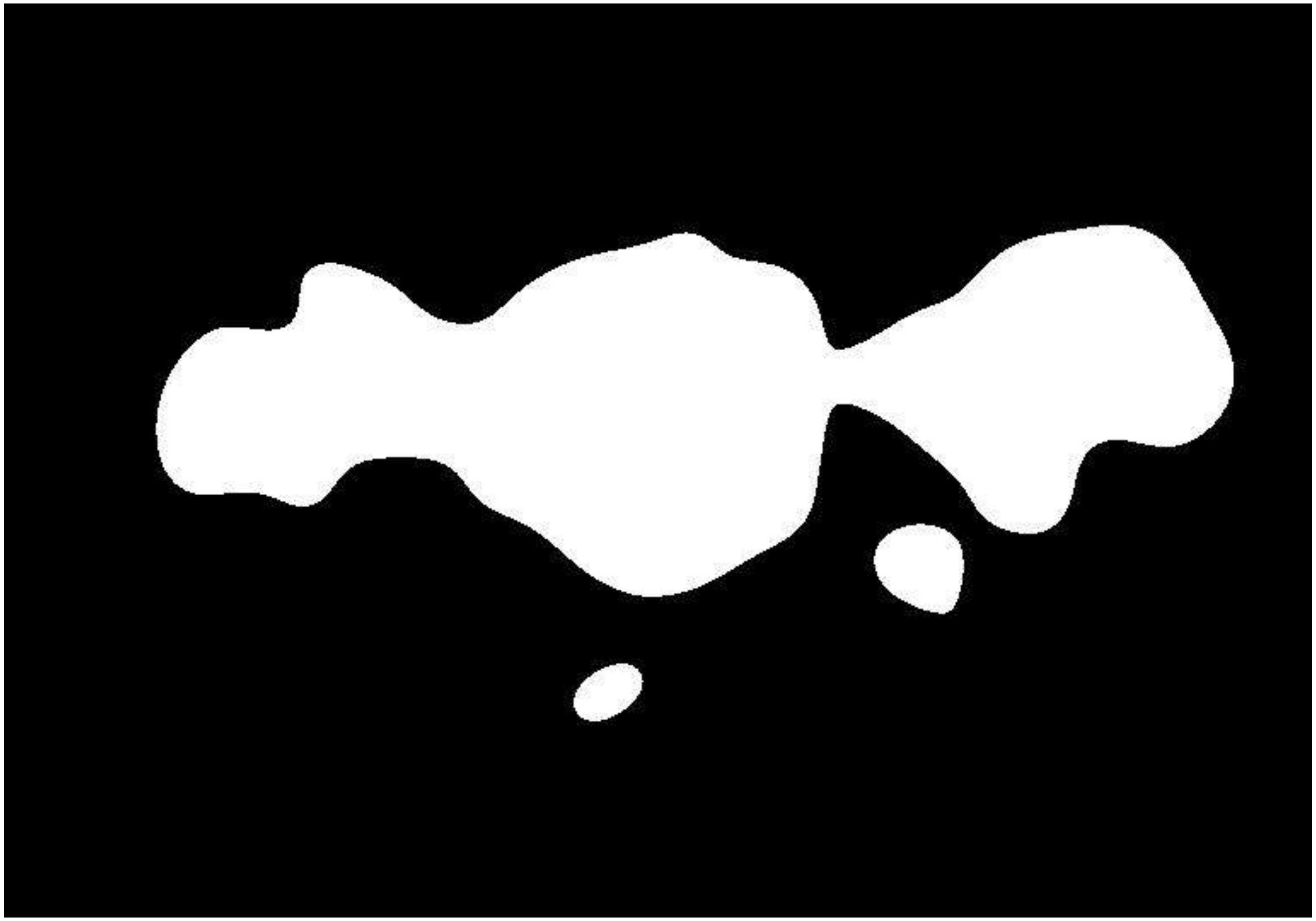


Average fixation locations / continuous saliency map

[Castillo, Judd and Gutierrez 2011]

Eye tracking data

Learning to predict where humans look **[Judd et al. 2009]**

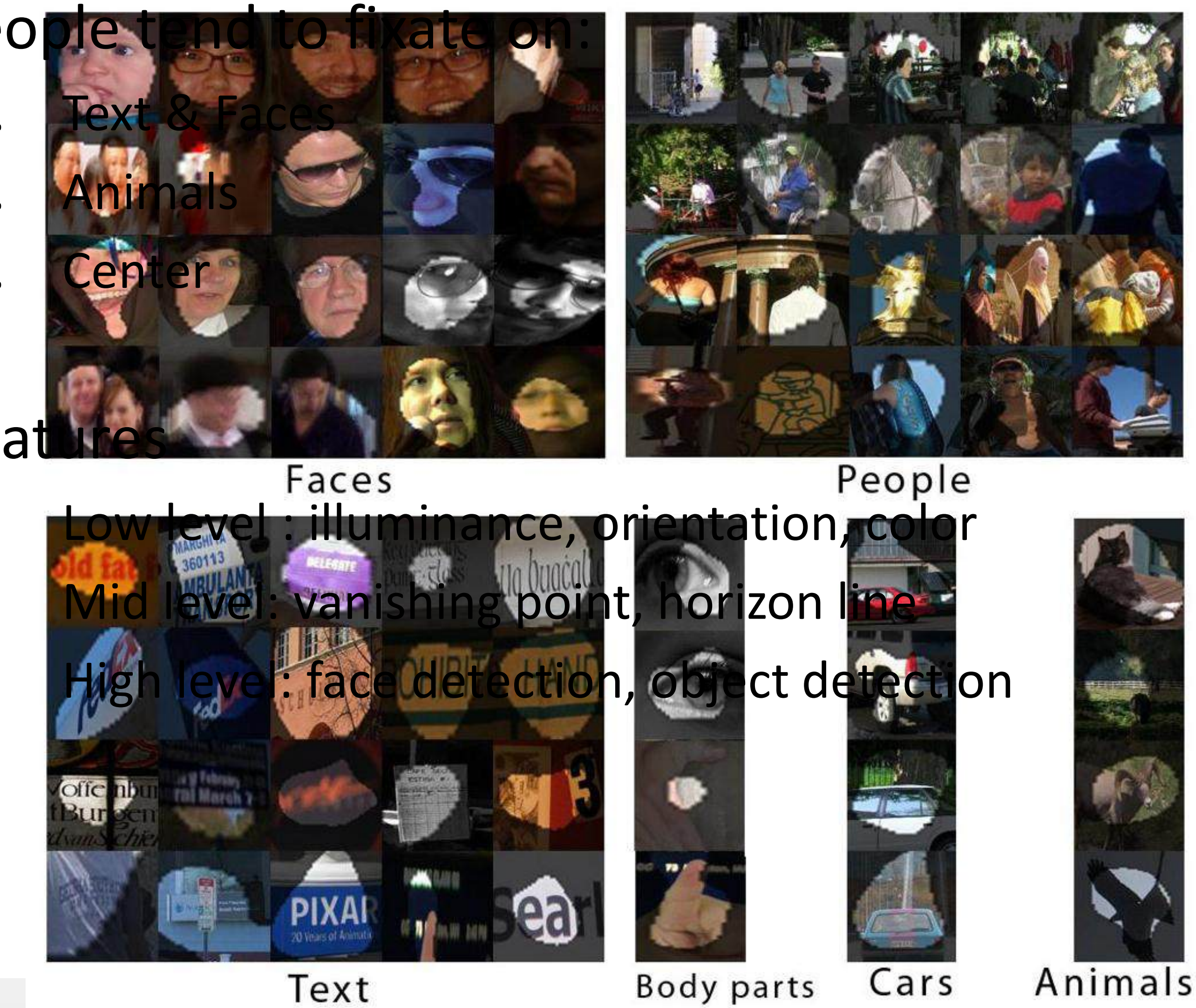


Top 20% salient locations

[Castillo, Judd and Gutierrez 2011]

MIT Predictive Model of Saliency

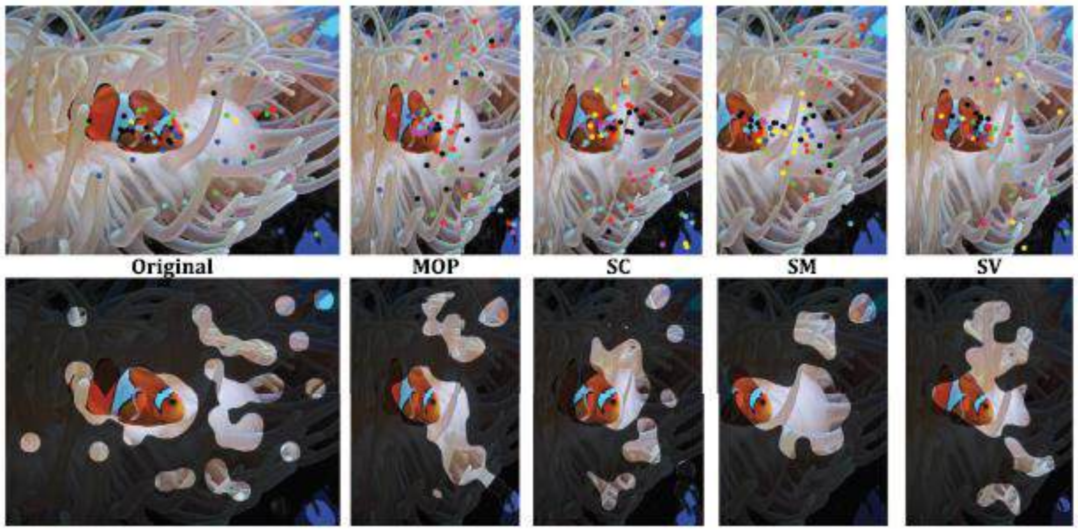
- People tend to fixate on:
  1. Text & Faces
  2. Animals
  3. Center
- Features
  - Low level : illuminance, orientation, color
  - Mid level: vanishing point, horizon line
  - High level: face detection, object detection



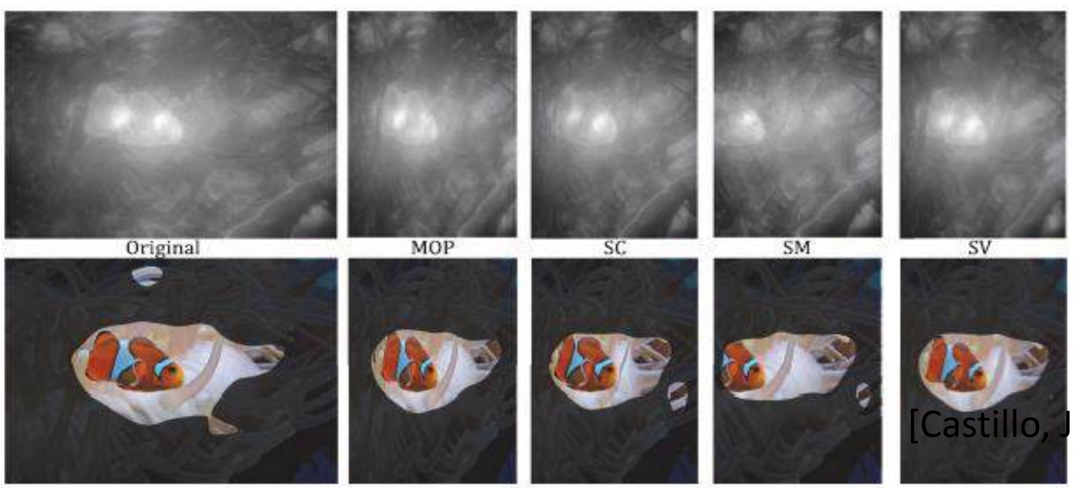
[Judd et al. 2009]

MIT Predictive Model of Saliency

Saliency Maps from eye-tracking data



Saliency Maps predicted by the model from Judd et al. [2009]



[Castillo, Judd and Gutierrez 2011]

Examples and Discussion

[Castillo, Judd and Gutierrez 2011]

Conclusions

- Lots of methods in the past few years, in top-notch places
- Relatively small impact in industry

http://people.csail.mit.edu/mrub/retargetme/
or Google: "retargetme"

- We need more (and better!) metrics
- Does video retargeting *really* work?

Conclusions

- Eye-tracking data framework
- The model of saliency from Judd et al. [2009] can be an useful tool in a retargeting context when using an eye tracker is not feasible
- Analysis of 4 retargeting operators with 6 image distance measures
  - Using eye-tracking data can improve the predicting capabilities of these measures
- Alteration of the image *semantics.*
  - Content removal alters RoIs although the results can be aesthetically pleasing
- **Attentional tension** between RoIs and artifacts
  - Large artifacts can remain unnoticed when not in a RoI *(At least for our 5 second task)*

SIGGRAPHASIA2011 HONG KONG

Temporal Image Retargeting

Karol Myszkowski

Max-Planck-Institut für Informatik

Elmar Eisemann

Telecom ParisTech / CNRS-LTCI

Observations: New Displays

Bigger & brighter

More resolution

Higher refresh rates

3D

Observations: Bigger & Brighter

- Increased role of peripheral vision
  - Higher sensitivity to flickering
  - Lower acuity for high eccentricity

Panasonic 150'' Plasma

Observations: Bigger & Higher resolution

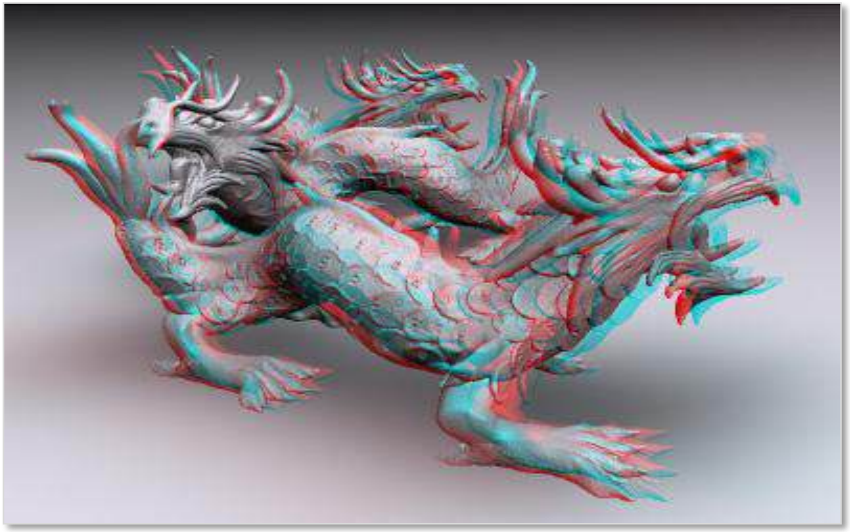
- More pixels to render
- SHD = 2 x HD
- People move closer
  - Higher angular and pixel velocity
  - More perceived **blur** due to smooth pursuit eye motion

Barco Coronis Fusion 6MP DL (MDCC-6130)

Observations: Faster refresh

- **120 Hz** displays (3D stereo applications)
  - LCD displays for gamers: *Samsung*, … (~ $300)
  - DMD projectors: *DepthQ* , … (> $2000)

Observations: 3D is a hot topic



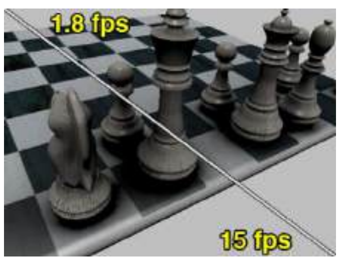
Standard stereo

Backward-compatible stereo

Observations: GPU

- More powerful, multi-core
- More than 50 fps not unusual
- For uncontrolled #fps judder effect
- Advanced per-pixel shaders costly

Super-resolution
[*Yang* et al. EGSR 2008]



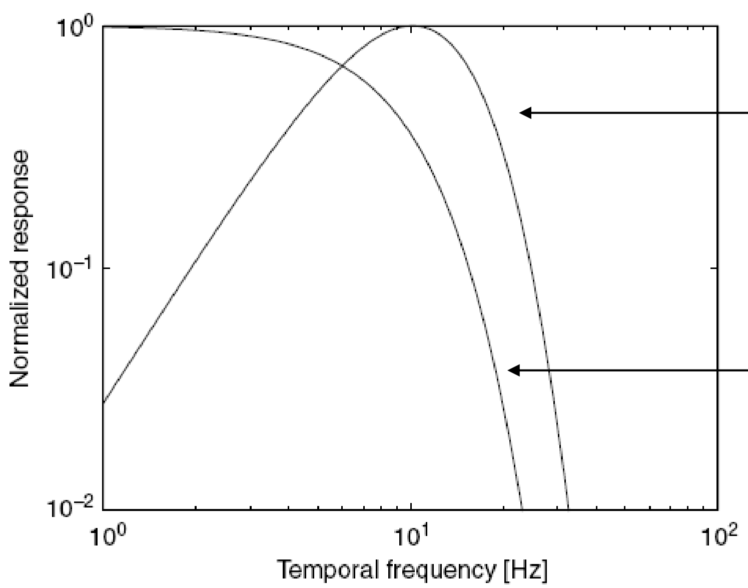
Shader decomposition and caching
[*Sitthi-Amorn* et al.,Siggraph Asia 2008]

Motivation

- More fps help in blur and flicker reduction
  - Adding extra frames in time domain easy
    - TV makers do this using relatively imprecise optical flow computation (100Hz and 200Hz TV sets)
    - In rendering motion flow simulation cheap and precise
    - New opportunities in the design of sharpening filters
      - Take into account perception, image content and display characteristics for rendered frame enhancement
      - So far rendering & enhancement usually separate steps
- Through super-resolution algorithms spatial resolution can be extended
  - Many people in graphics tried this

Basic Psychophysics

- Temporal integration of signal performed by HVS to improve the signal to noise ratio
  - Integration duration up to 120ms
  - Temporal summation faster for low spatial frequencies



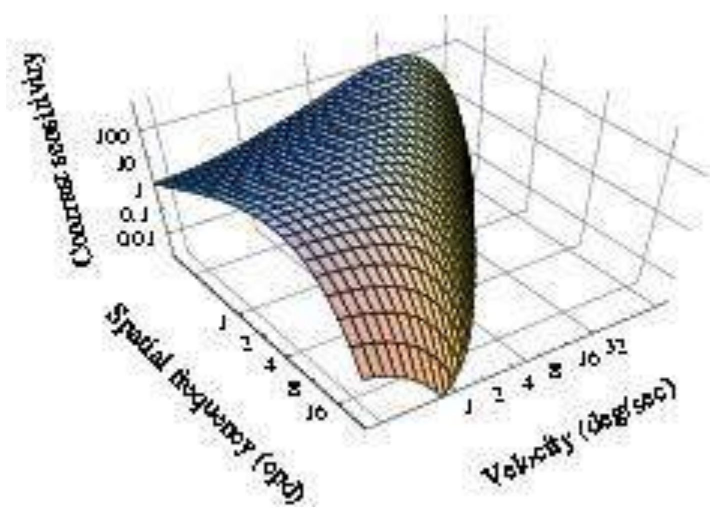
- Temporal frequency responses
  - Band-pass: Fast visual channels tuned to low spatial and high temporal frequencies (***transient*** response) – motion detection
  - Low-pass: Slow visual channels tuned to high spatial and low temporal frequencies (***sustained*** response) – object identification

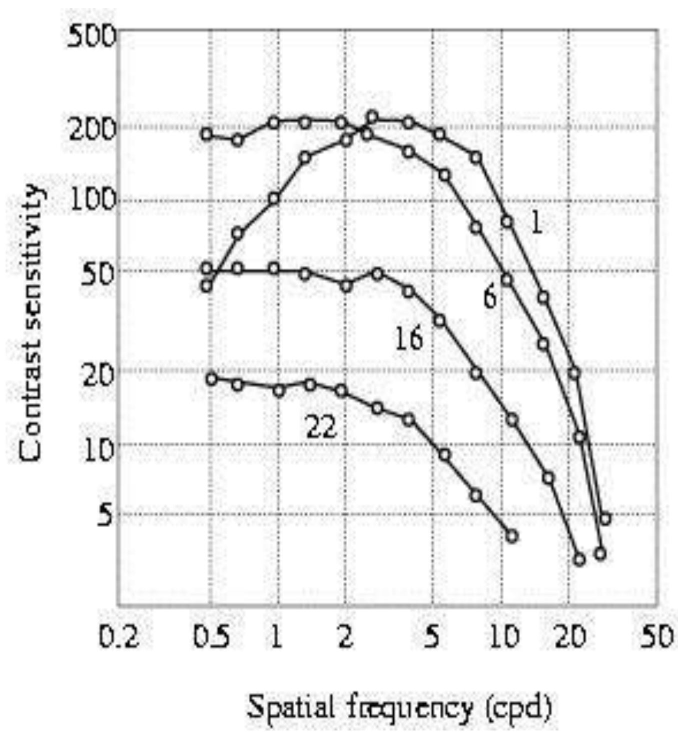
Winkler, Video Quality Metrics, Wiley 2005

Spatio-temporal Contrast Sensitivity Function

- Low sensitivity of HVS to temporal change of high spatial frequencies and high sensitivity to low spatial frequencies

  → high spatial frequencies can be sampled in temporal domain more sparsely





Perception: Flickering

- Critical Flicker Frequency (CFF)
  - Increases with display brightness
    - The Ferry-Porter law:

$$CFF \approx a \cdot \log(luminance) + b$$

  - For bright adaptation conditions and patterns of wide spatial extent the highest flicker sensitivity at the periphery
  - Otherwise, the highest flicker sensitivity at the fovea

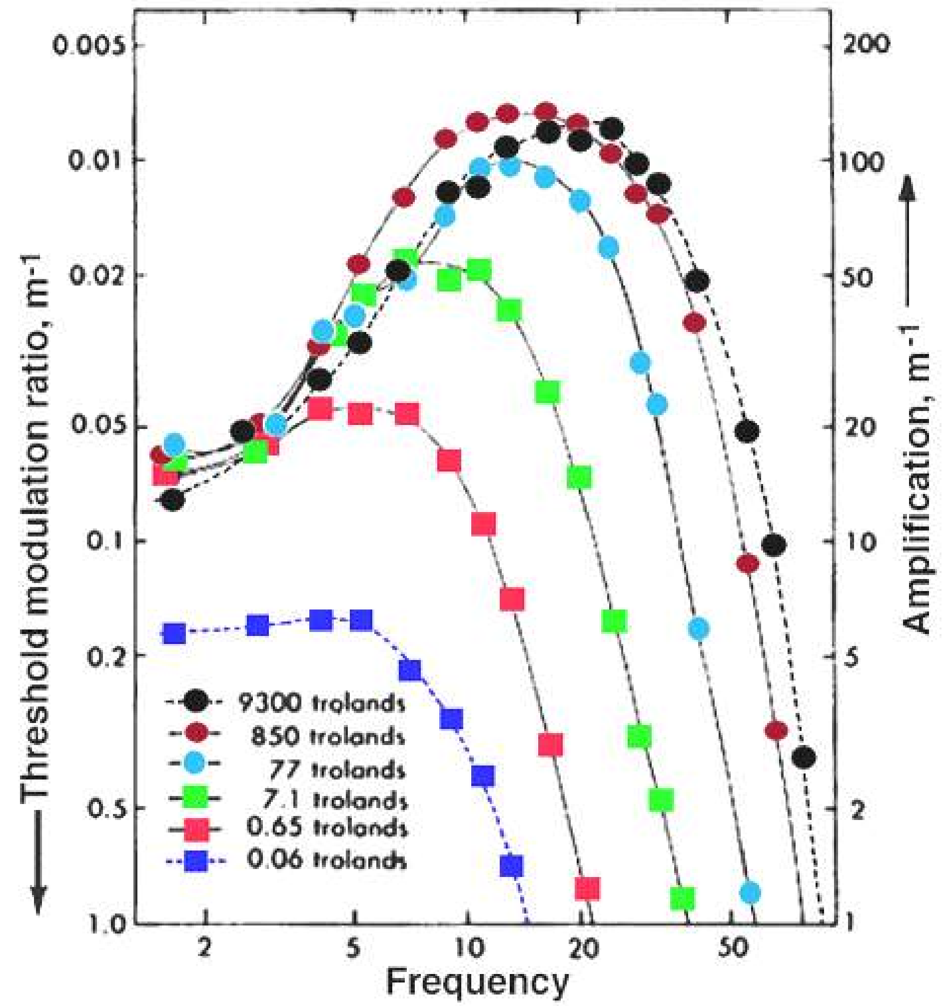
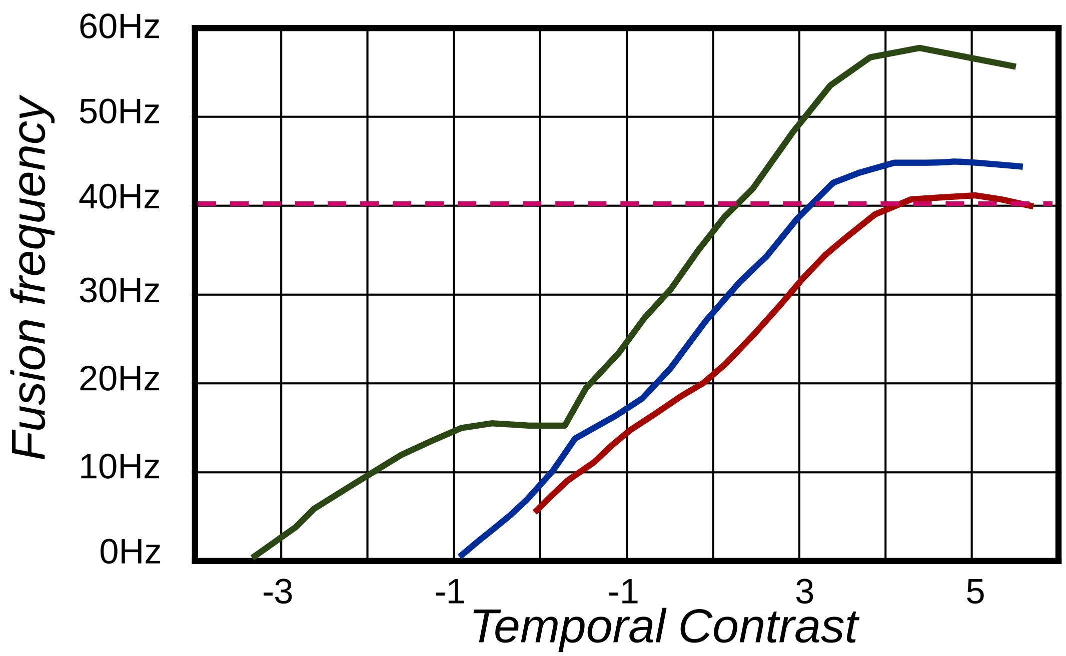


Fig. 11. Temporal Contrast Sensitivity Function (TSF) for various adapting fields. Kelly's data from Hart Jr, W. M., The temporal responsiveness of vision. In: Moses, R. A. and Hart, W. M. (ed) Adler's Physiology of the eye, Clinical Application. St. Louis: The C. V. Mosby Company, 1987.

Perception: Flickering

Fusion frequency vs. temporal contrast & pattern spatial extent



Critical Flicker Frequency - Hecht and Smith's data from Brown J. L. *Flicker and Intermittent Simulation*

19 deg

2 deg

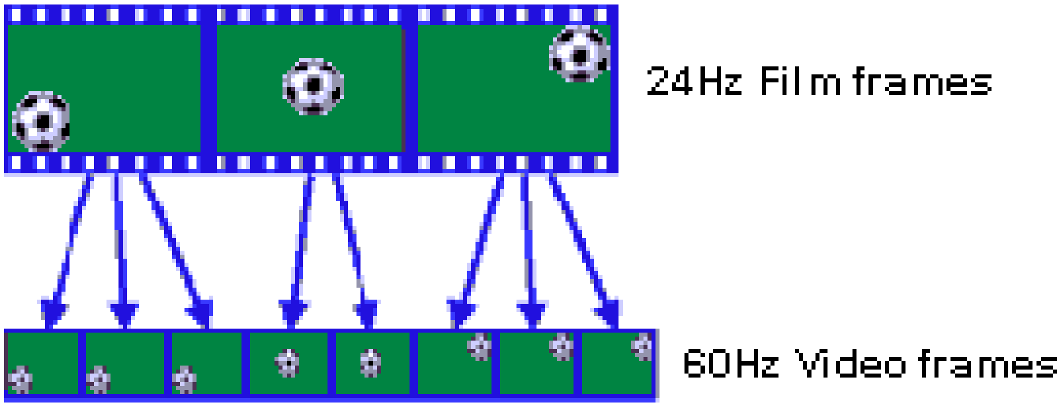
0.3 deg → **40Hz**

Perception: Smooth Pursuit Eye Motion (SPEM)

- Enables to maintain the object of interest in the fovea
- Blur due to object motion is eliminated
- Eye tracking experiment [Laird et al. 2006]
  - Almost perfect tracking for steady linear motion with velocities of 0.625 – 7 deg/s
  - Still possible up to 80 deg/s
- SPEM initialization very fast
  - Good tracking possible in 100ms after switching gaze between objects moving in different directions
- Other fixational eye movements during SPEM: tremors, drifts, and microsaccades similar to static fixations
  - Compensated by HVS contribute little to blur

Perception: Judder

- Repeating the previous frame while the eye is smoothly tracking moving object
- Most noticeable for camera pans, scrolling text, and so on
- 8Hz difference between rendered and displayed frames the most critical, i.e. 42 fps on 50 Hz display
- 3:2 pulldown judder: Converting 24Hz film material to 60Hz

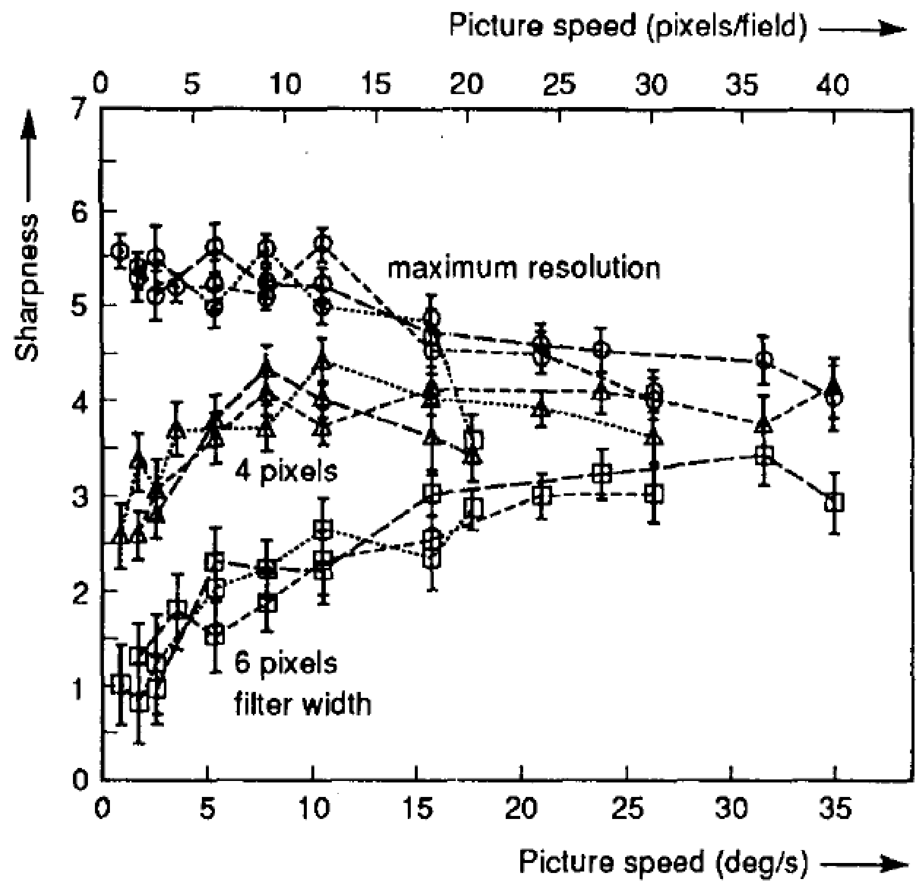


http://msdn.microsoft.com/en-us/windows/hardware/gg463407.aspx

Perception: Blur

- Sharp edges suffer blurring during motion
  - Perceived blur increases with velocity
- Blurred edges appear sharper [Westerink&Teunissen 1994]
  - Apparent sharpening increases with velocity
- Shortly shown blurred edge (7-40ms) appears sharper than the same edge shown for a longer time
- Higher contrast looks sharper
- Adding noise to texture may increase apparent sharpness [*Fairchild and Johnson* 2000, 2005]

Perception: Perceived Sharpness vs. Velocity



J. Westerink, K. Teunissen, Perceived sharpness in complex moving images, Displays 1994

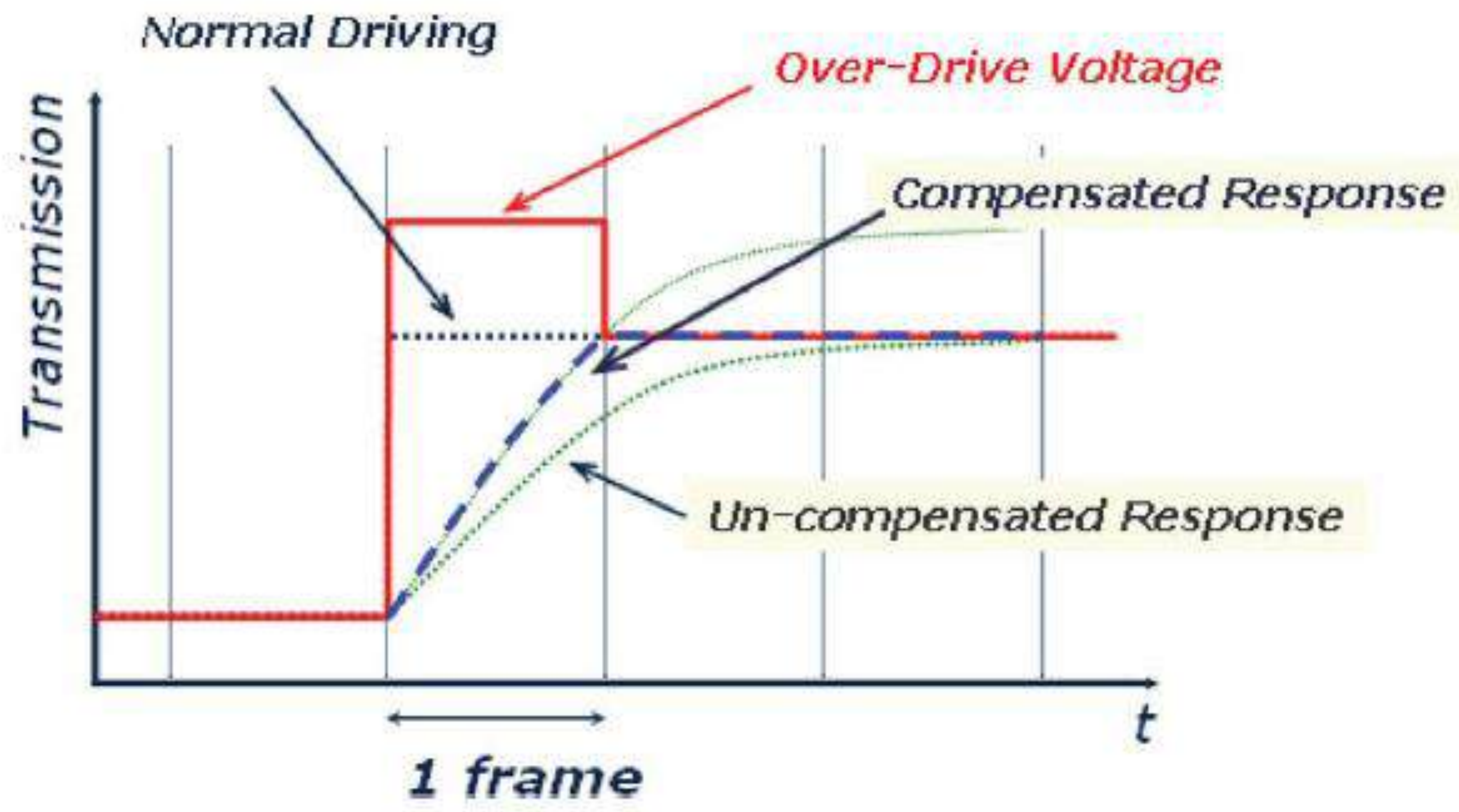
Blur in Hold-type Displays (LCD)

Two main reasons:

- Slow-response of LC
  - 16ms display responsible for only 30% of blur effect
  - Now for 2ms displays mostly negligible
- Image is held while the eye is tracking moving object (*smooth pursuit eye motion* SPEM), which causes blur in the retina image
  - Purely perceptual effect
  - Can be modeled as a box function in temporal domain

Overdriving in LCD TV

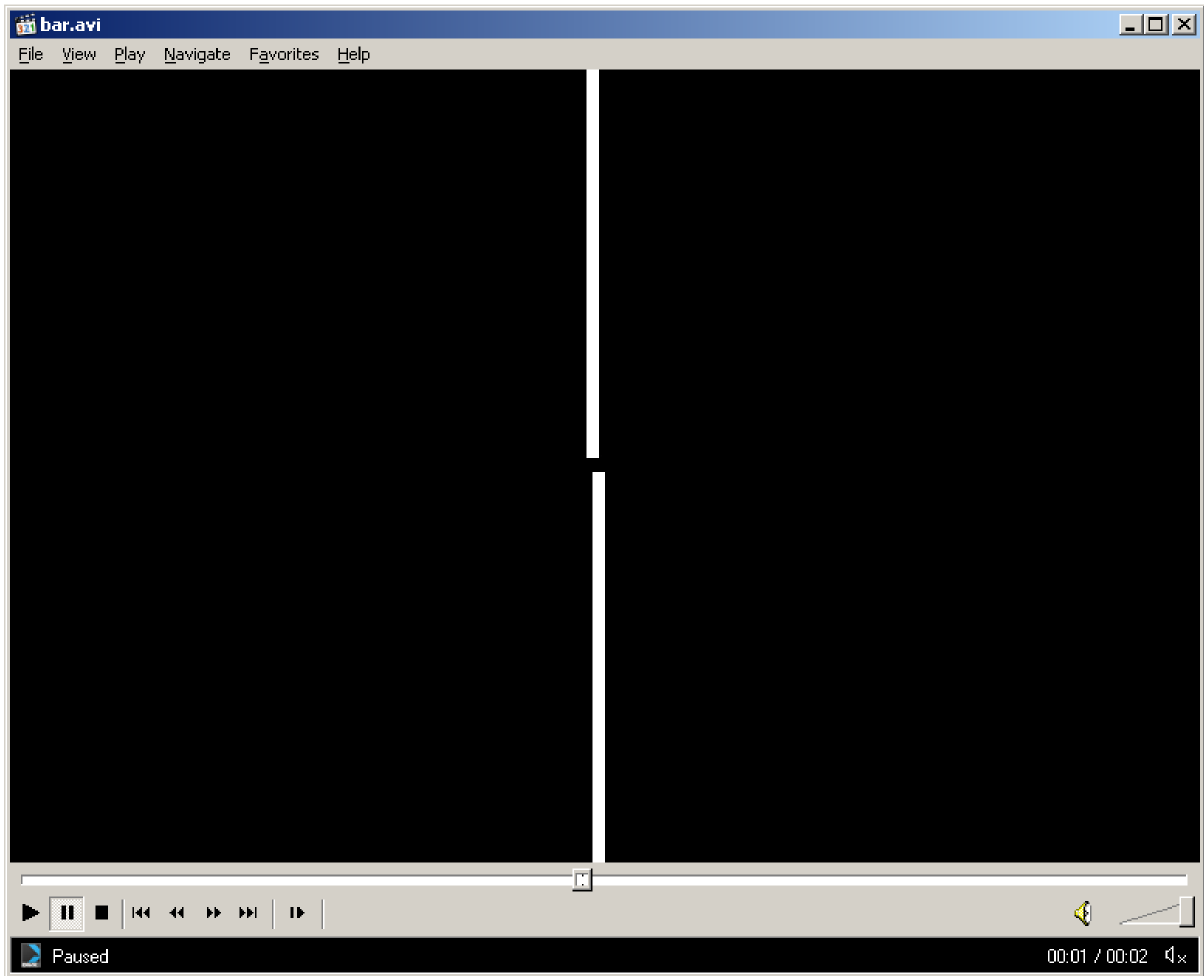
Combating slow response of LC



J.H. Souk, J. Lee, Recent Picture Quality Enhancement Technology Based on Human Visual Perception in LCD TVs, 2007

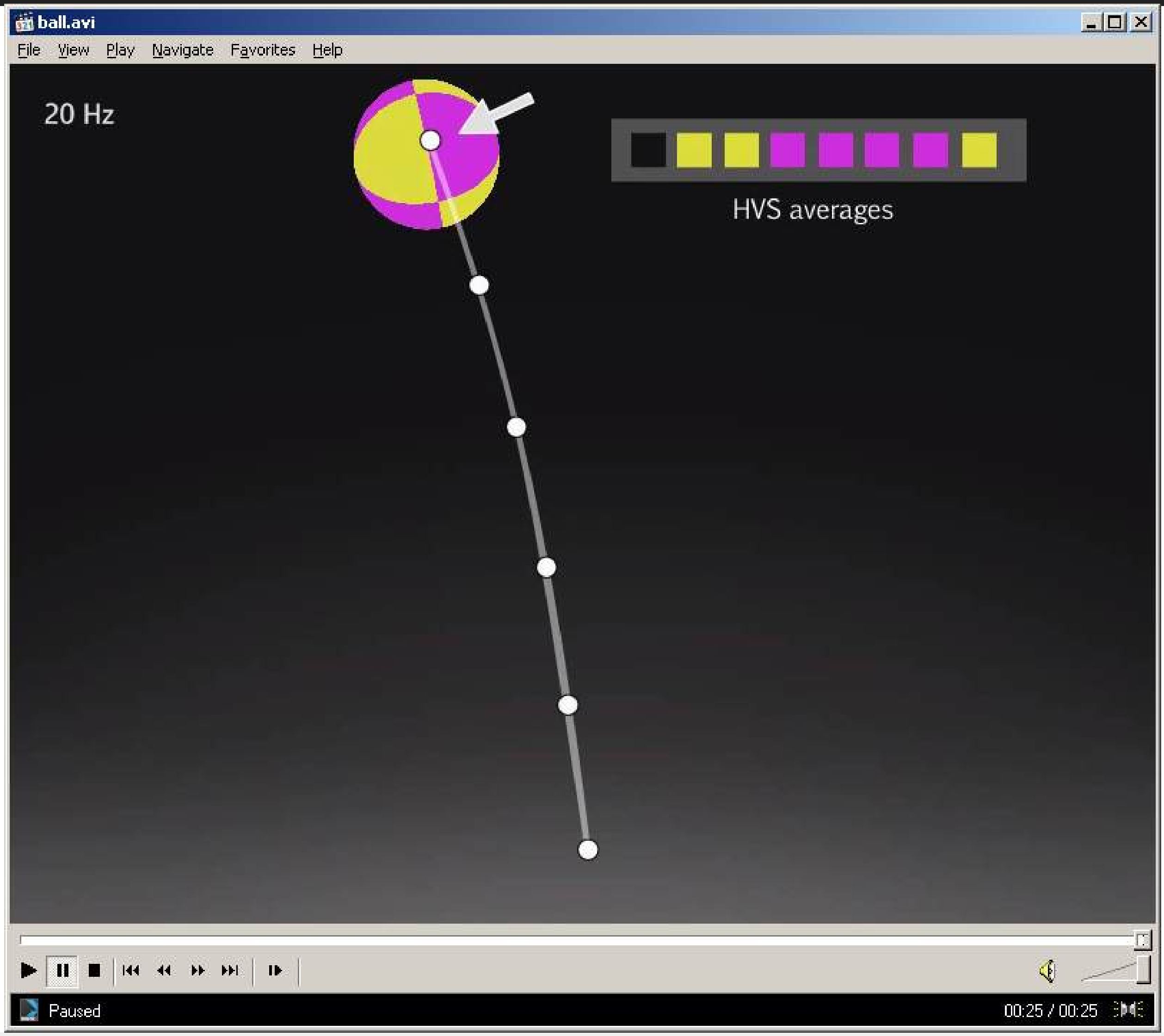
Hold-type Blur Demo: 30Hz vs. 60Hz

30 Hz

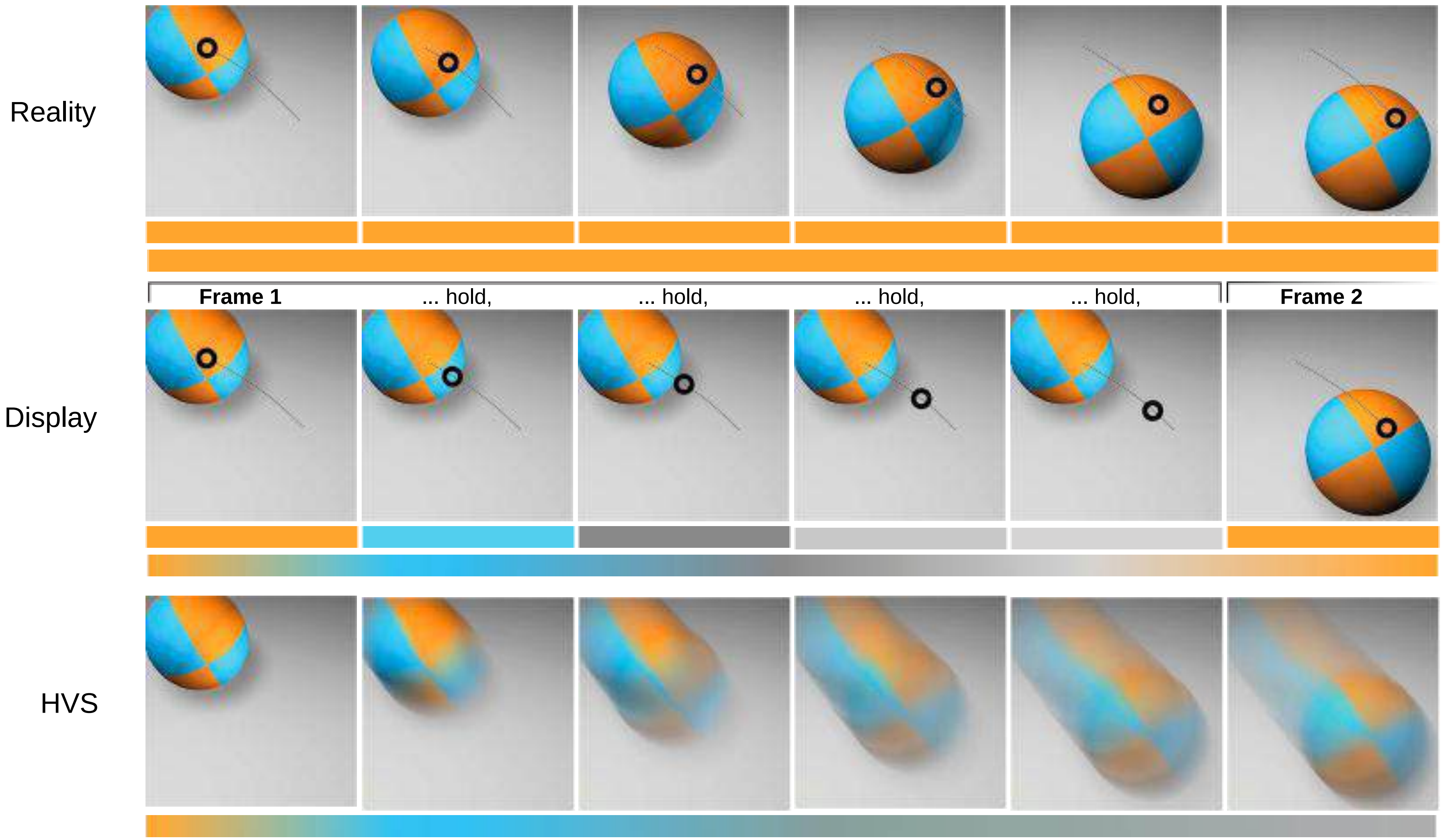


60 Hz

Hold-type Blur Demo: Ball

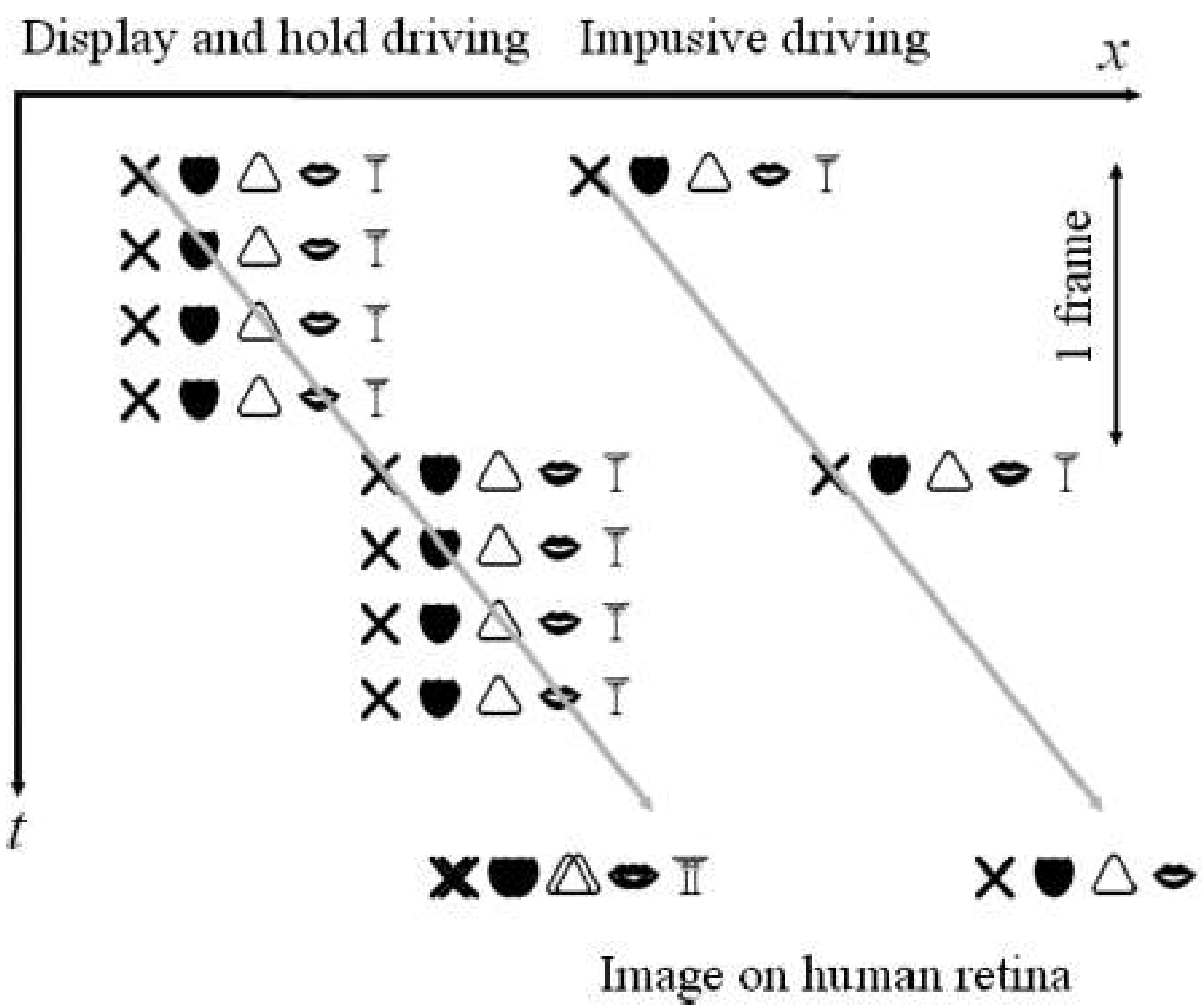


Hold-type Blur Explanation



Demo: Gaze Fixing vs. Dynamic Object Tracking

Hold Effect: LCD vs. CRT Displays

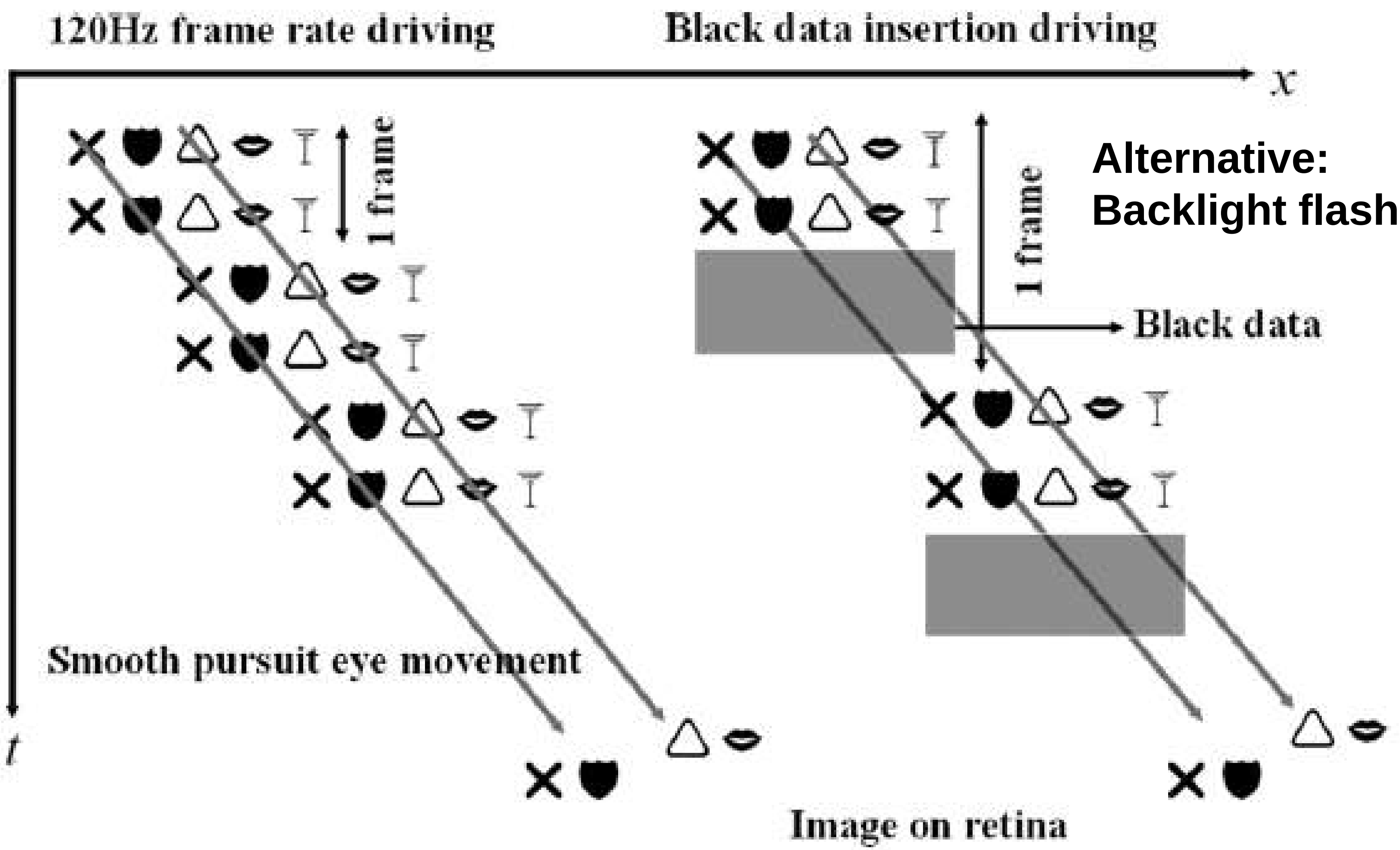


J.H. Souk, J. Lee, Recent Picture Quality Enhancement Technology Based on Human Visual Perception in LCD TVs, 2007

Combating Hold-type Blur in TV Sets

- **Black data insertion (BDI)**
  - Black frames interleaved with the original frames
  - Mimics CRT behavior
- **Frame rate doubling (FRD)**
  - Additional frames are obtained by interpolating pairs of original frames along their optical-flow trajectories
  - Requires introducing latency of one keyframe, which is not a problem in broadcasting applications, but is not suitable for gaming
  - The final effect depends on optical flow accuracy
- **Blurred frame insertion (BFI)**
  - Cheap version of FRD
  - Original frames are replicated and blurred
  - Ghosting for dynamic objects due to lack of motion compensation

Combating Hold-type Blur in TV Sets



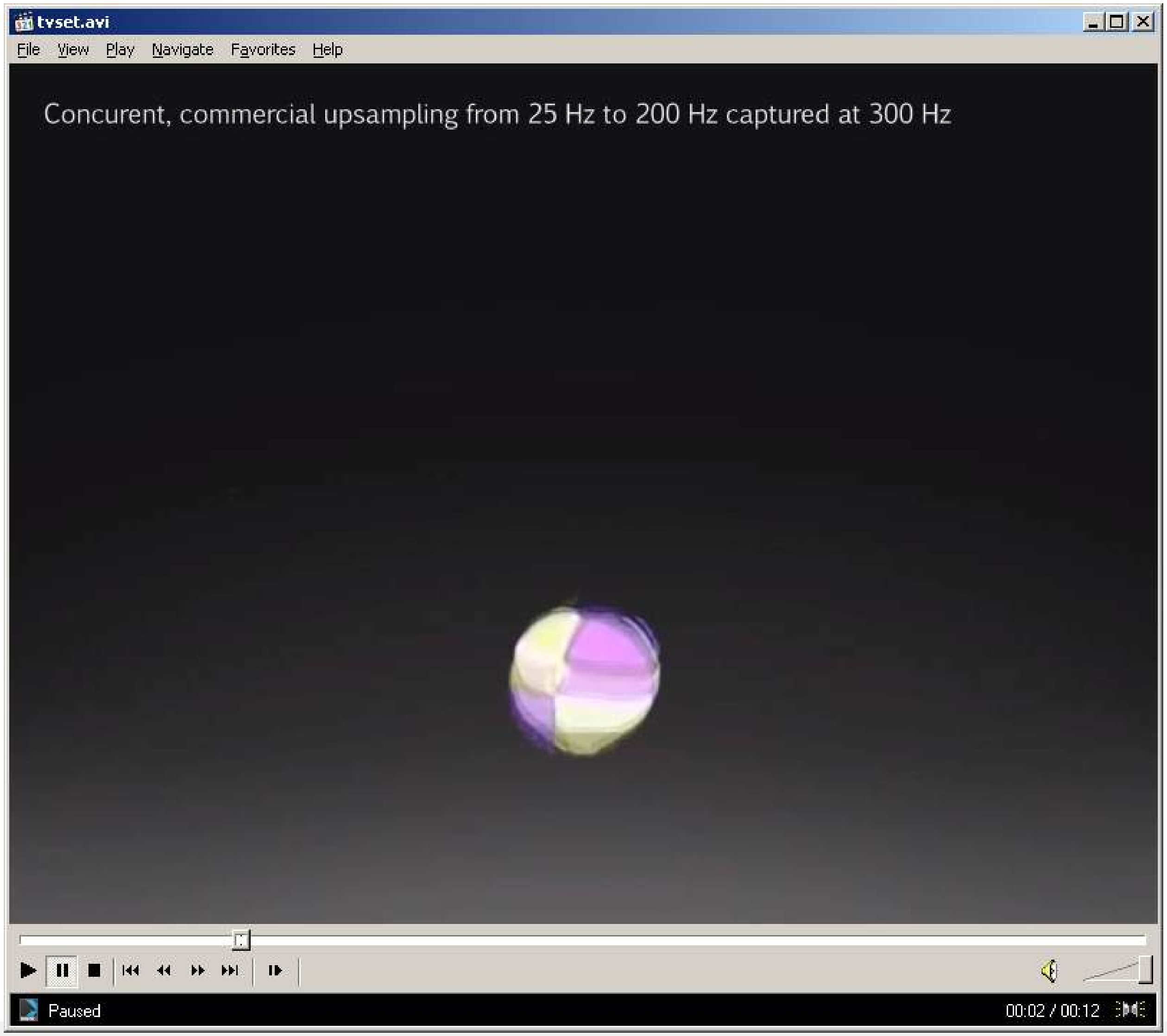
Alternative: Backlight flashing

J.H. Souk, J. Lee, Recent Picture Quality Enhancement Technology Based on Human Visual Perception in LCD TVs, 2007

Combating Hold-type Blur in TV Sets

- **Backlight flashing (BF)**
  - Turning the backlight of an LCD panel on and off
  - LED response is very fast, so flashing 500 Hz and more is possible
  - Flashing on can be synchronized with steady states of LC (reduces ghosting)
- **Motion compensated inverse filtering (MCIF)**
  - Filtering an input image, which aims at inverting hold-type blur
  - Effectively local 1D sharpening filtering, which is computed along the optical flow trajectories
  - Cannot restore frequencies that are completely removed by hold-type blur, but may magnify frequencies that are attenuated
  - Image saturation may cause problems
- **Hybrid Methods**
  - FRD + BF

High-speed Camera Recording: TV-Set



SIGGRAPHASIA2011 HONG KONG

Combating Hold-type Blur in Rendering

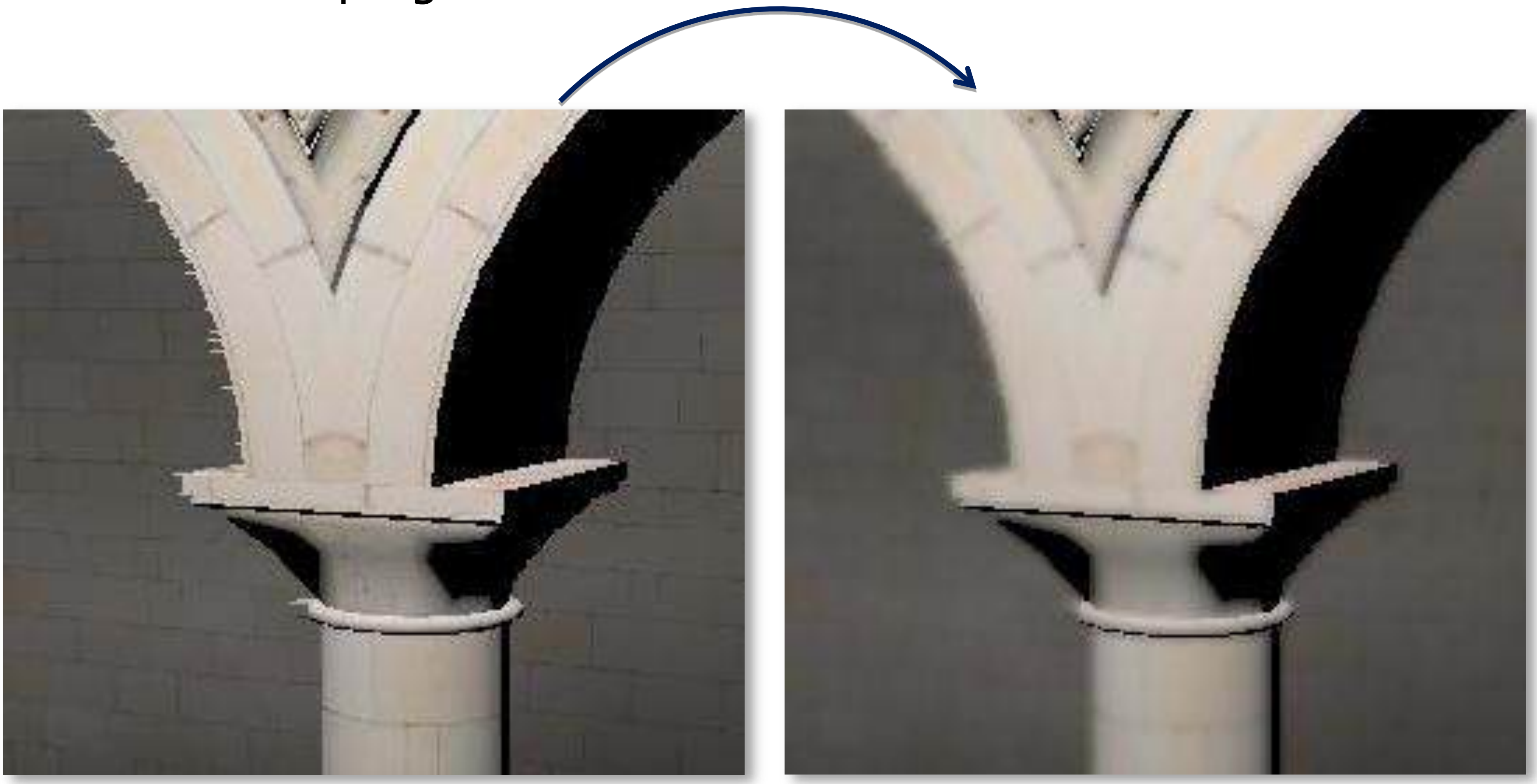
Frame warping up to 120Hz

40 Hz rendering

[Didyk et al. 2010]

Combating Hold-type Blur in Rendering

Blur out warping artifacts



[Didyk et al. 2010]

Combating Hold-type Blur in Rendering

- Interleave blurred and sharp (with doubled high-pass frequencies) frames
  - Energy-wise (brightness) equivalent
  - Blur filter size as a function of retinal velocity
  - Hold effect reduced as high frequencies displayed shorter and low frequencies do not matter for blur

sharpen | blur | blur

120 Hz

[Didyk et al. 2010]

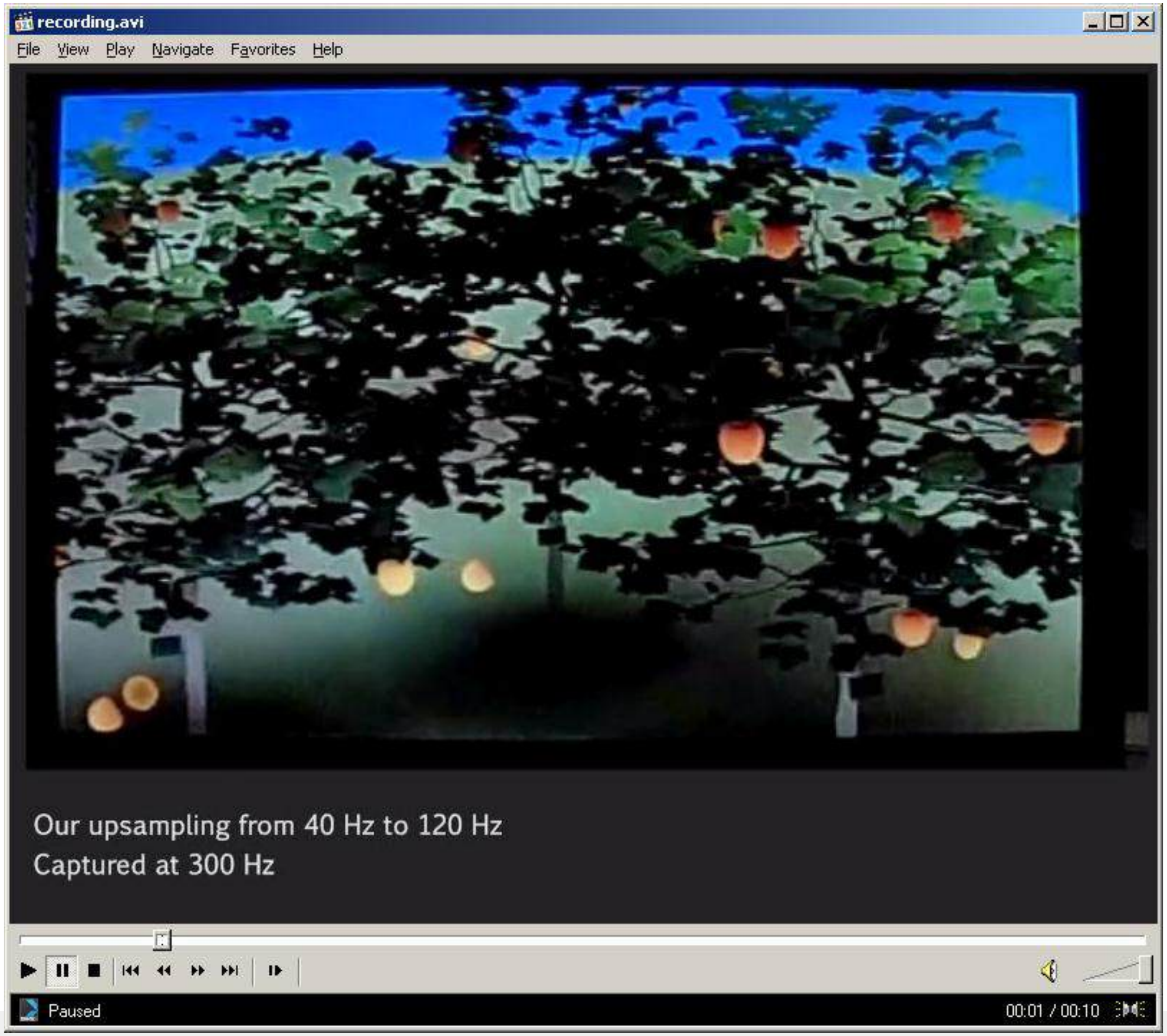
Perceived Blur Reduction Magnitude Estimation

40 Hz

120 Hz

[Didyk et al. 2010]

High-speed Camera Recording: Rendering



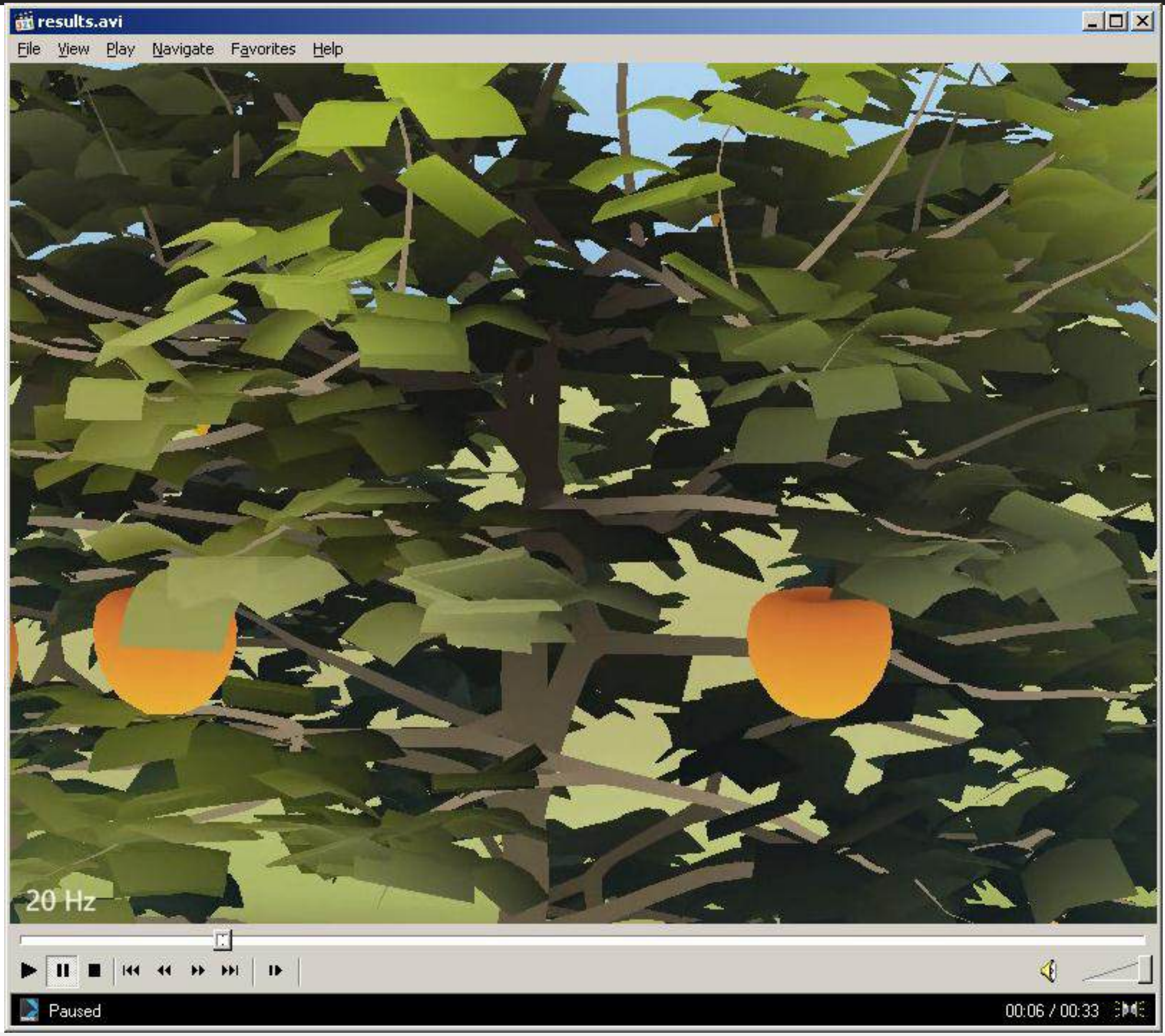
[Didyk et al. 2010]

Comparison

	BDI	BF	BFI	FRD	MCIF	Didyk et al.
LCD response required	High	Moderate	High	High	No	High
Backlight response required	No	High	No	No	No	No
Optical flow quality	No	No	No	High	Moderate	High
Ghosting artifacts	Possible	Possible	Yes	No	No	No
Flickering artifacts	Yes	Yes	No	No	No	No
Luminance reduction	Yes	Yes	No	No	No	No
Limitation of blur reduction	Flickering	Flickering	No	No	Freq. cut-off	No
Other possible artifacts	No	No	No	Fast motion	Oversaturation	No

Rendering Comparison: Animation Examples

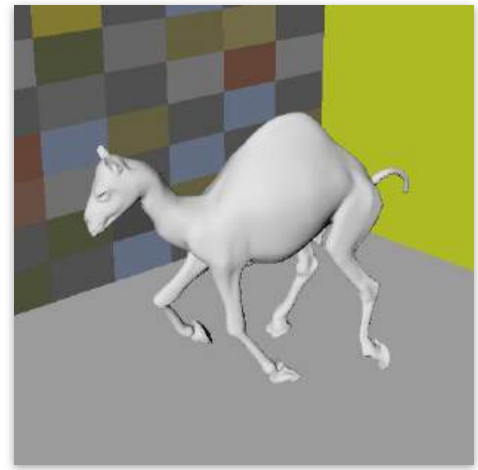
SIGGRAPHASIA2011 HONG KONG

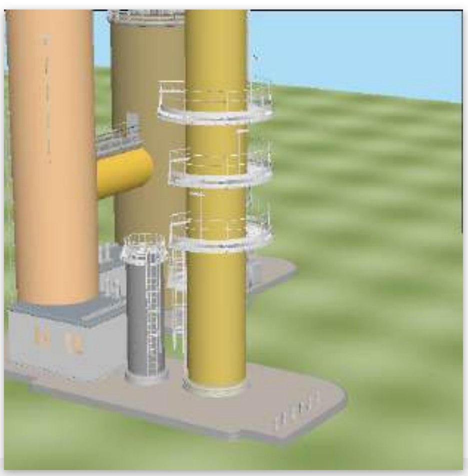


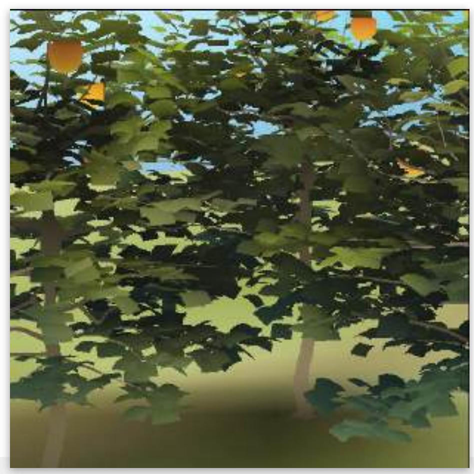
User Study

Pair-wise comparison

- 5 different sequences
- True 40Hz, 120Hz, Our 120Hz
- Blur judgment and artifacts

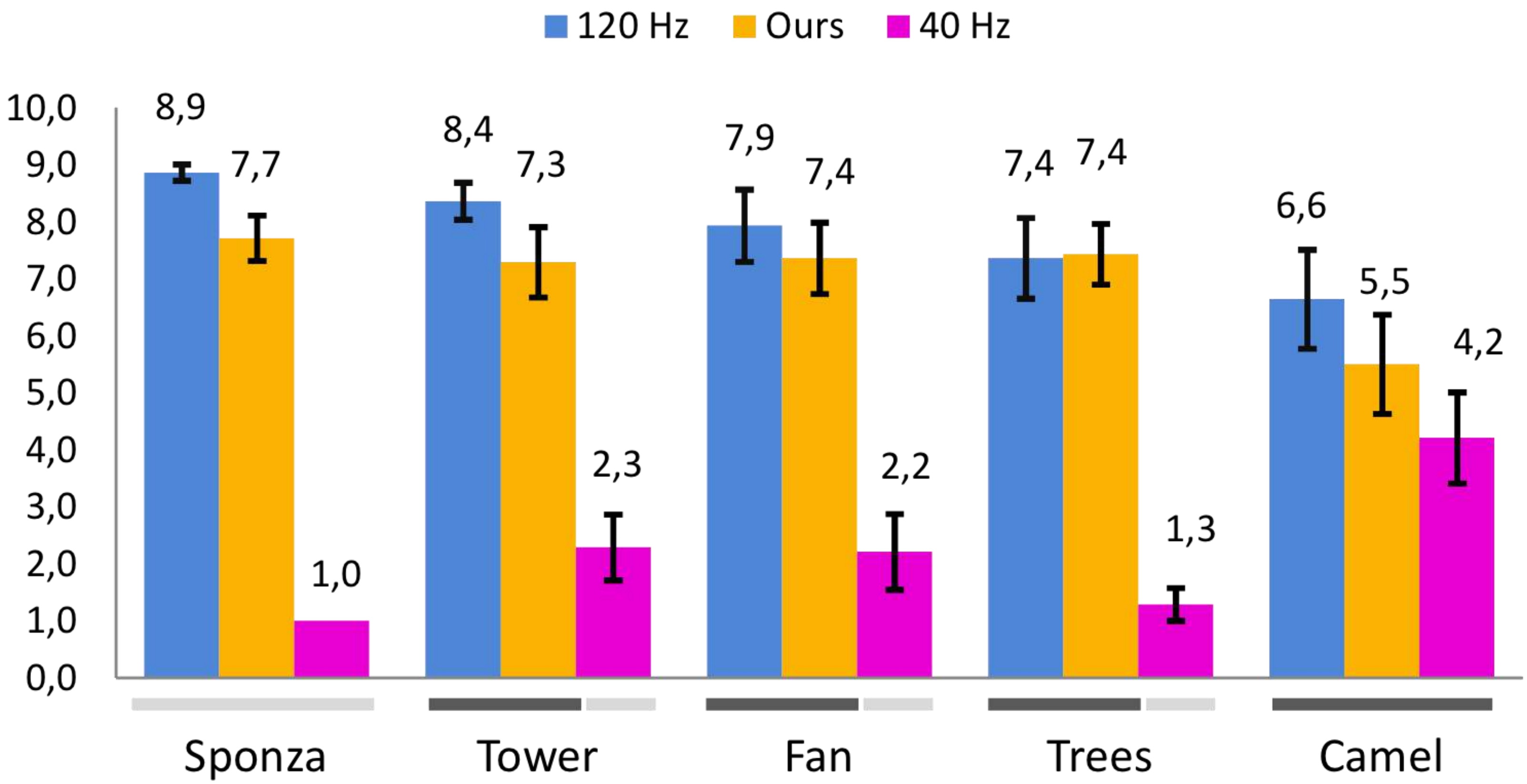






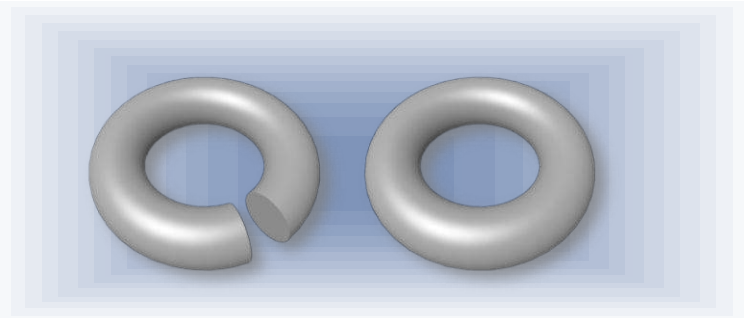
User Study

5 scenes (mean quality score + SEM)



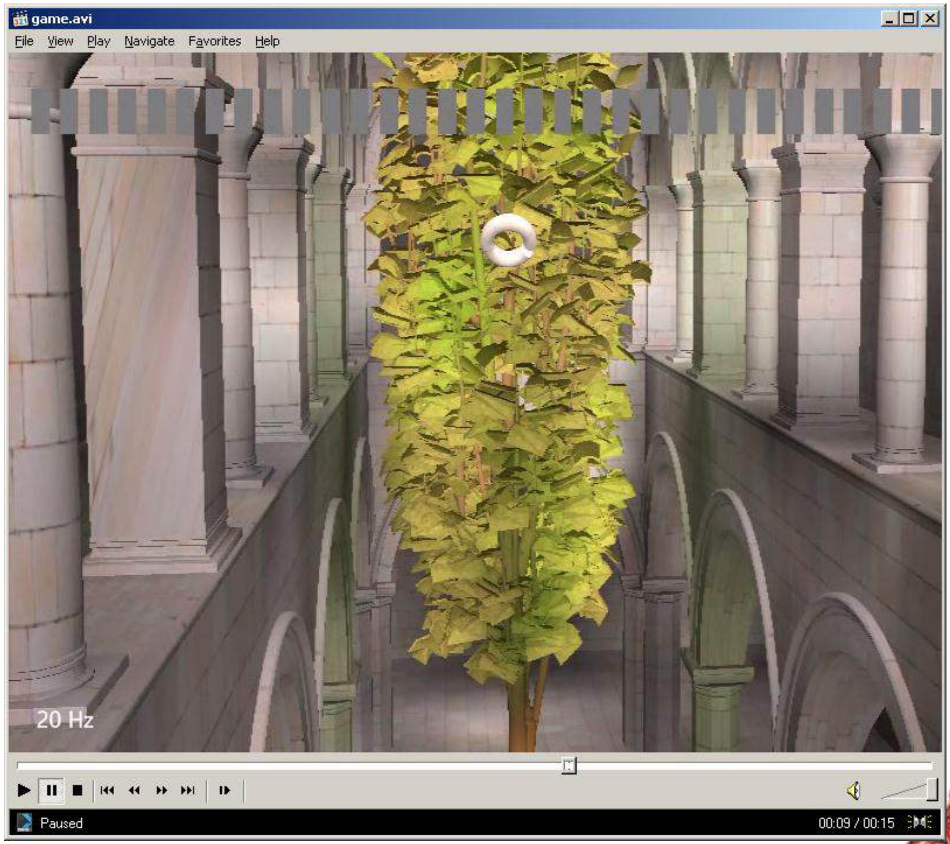
User Study: Game scenario

Targets:

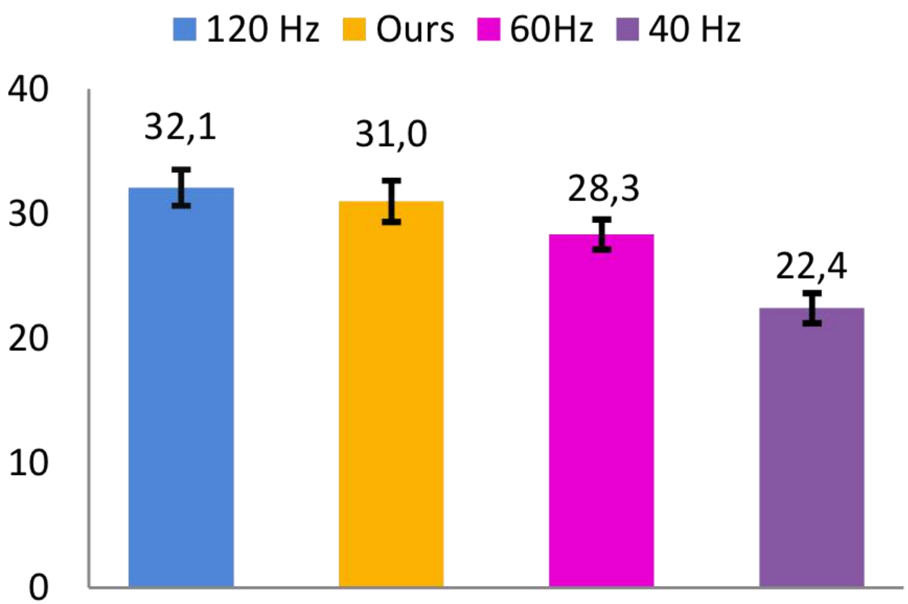


Task:

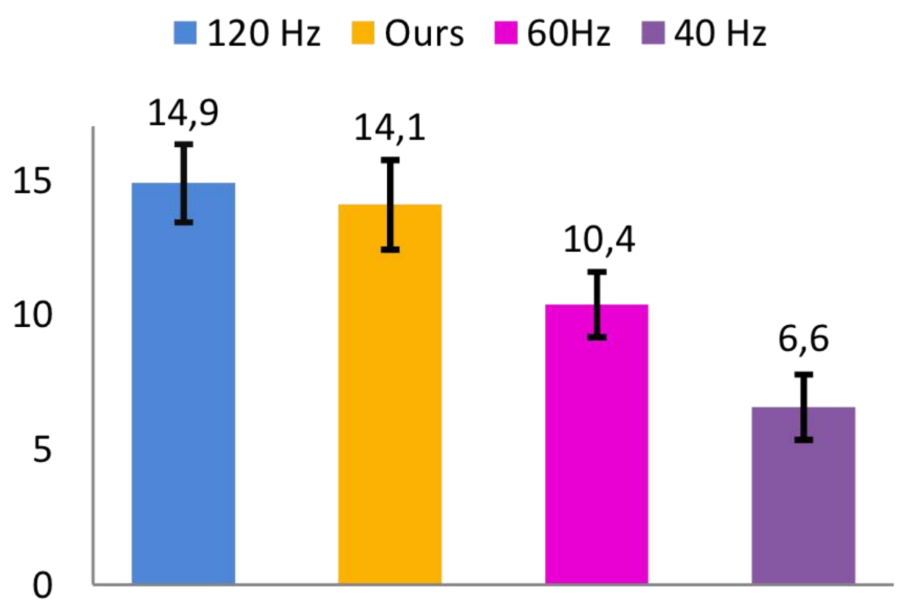
Detect open Landolt shape



User Study: Game scenario



Score



Succeeded Landolt detections

Changing Update Granularity

Why limit to the full frames if the eye can integrate signal @120Hz?

- Possible scenarios: update only 1 color channel, while the other two motion compensated
  - Does it pay off in terms of rendering costs?
- Local dimming behind fast moving and high contrast edges
  - Reduces hold effect
  - Flickering should not be a problem, but lost luminance should be compensated
- For HDR displays we could also control individually time/intensity of local LED backlight:
  - Fast moving objects **shorter**, but **brighter** LED impulses

3D Rendering vs. TV Solutions

- 3D rendering provides a lot of information, which is so difficult to recover based on images only (TV)
  - Precise motion flow, silhouette edges, textures,….
  - This should enable more sophisticated enhancement techniques integrated into rendering
- Perception + display device characteristics can be accounted for at rendering stage
  - Reducing hold effects

Rendering @120Hz

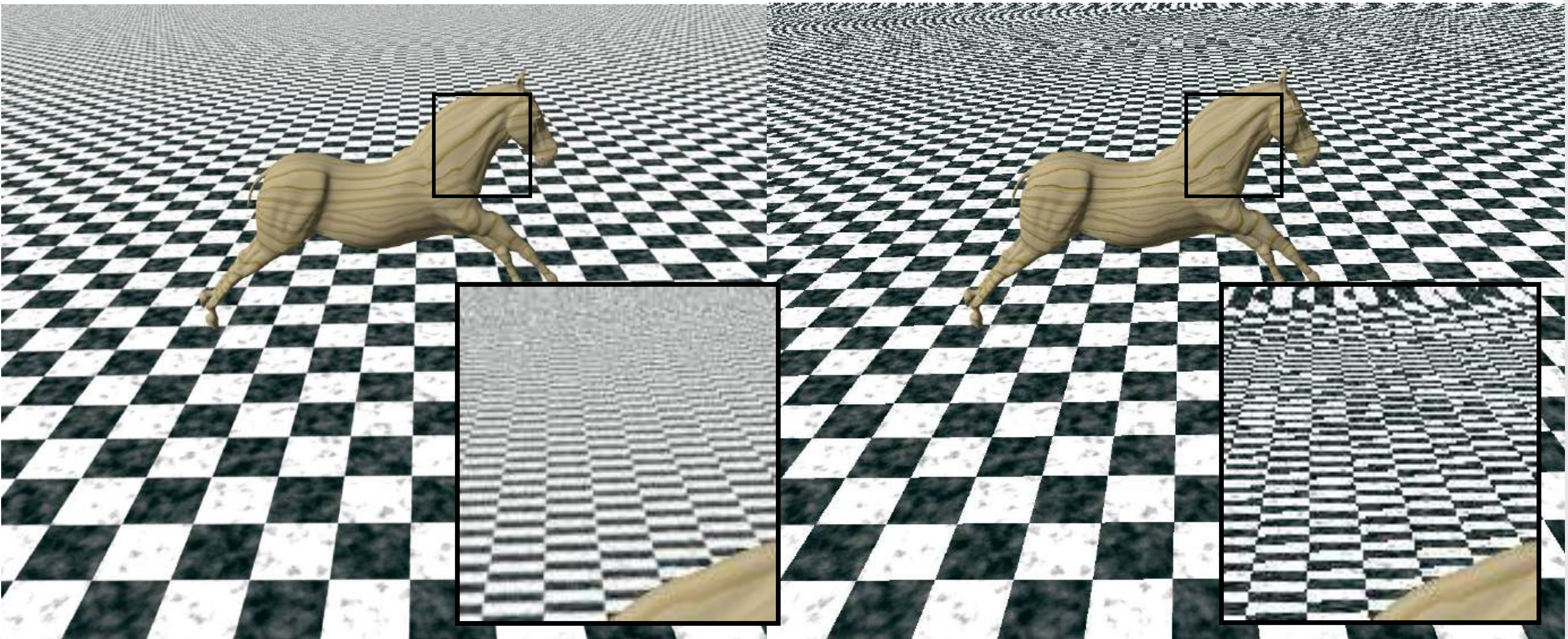
- We hope that the availability of 120 Hz displays can shift accents in rendering
  - More frames of much lower quality
  - Relying on integration in the eye
    - Interleaving such low quality frames at current display frequencies cause flickering, which should be much less visible at 120Hz
  - Extra frames over 60 Hz not wasted anymore

Reuse information

- Speed up: distribute workload over several frames

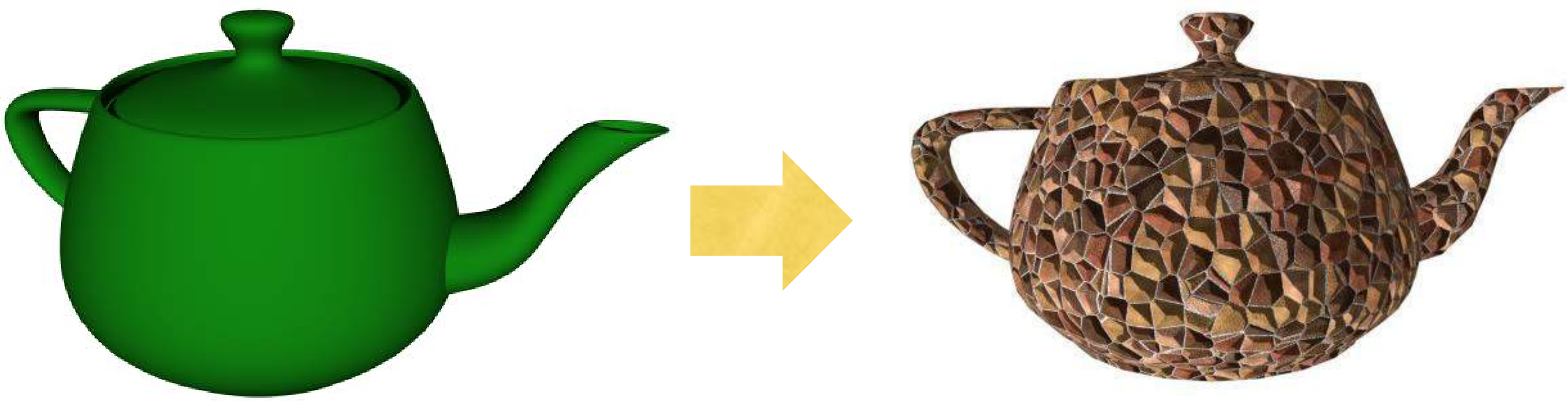
Reuse Information

- Increase in quality
  - Incorporate calculations from previous frames



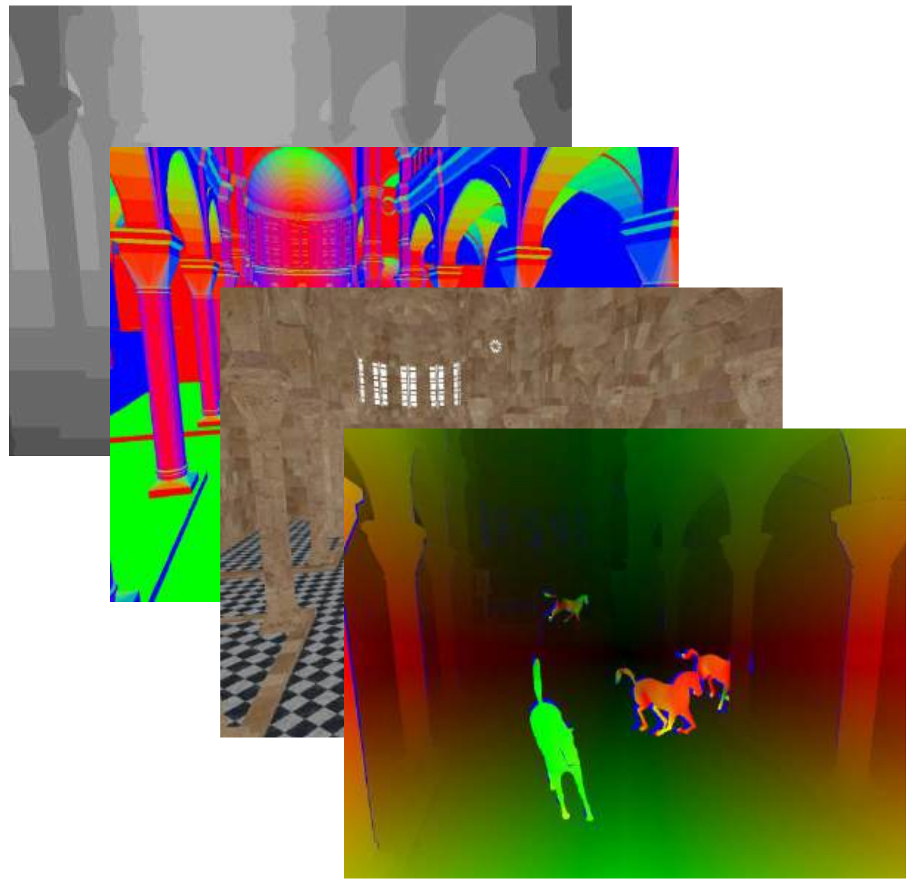
What is actually costly?

- Today’s main cost is **shading**



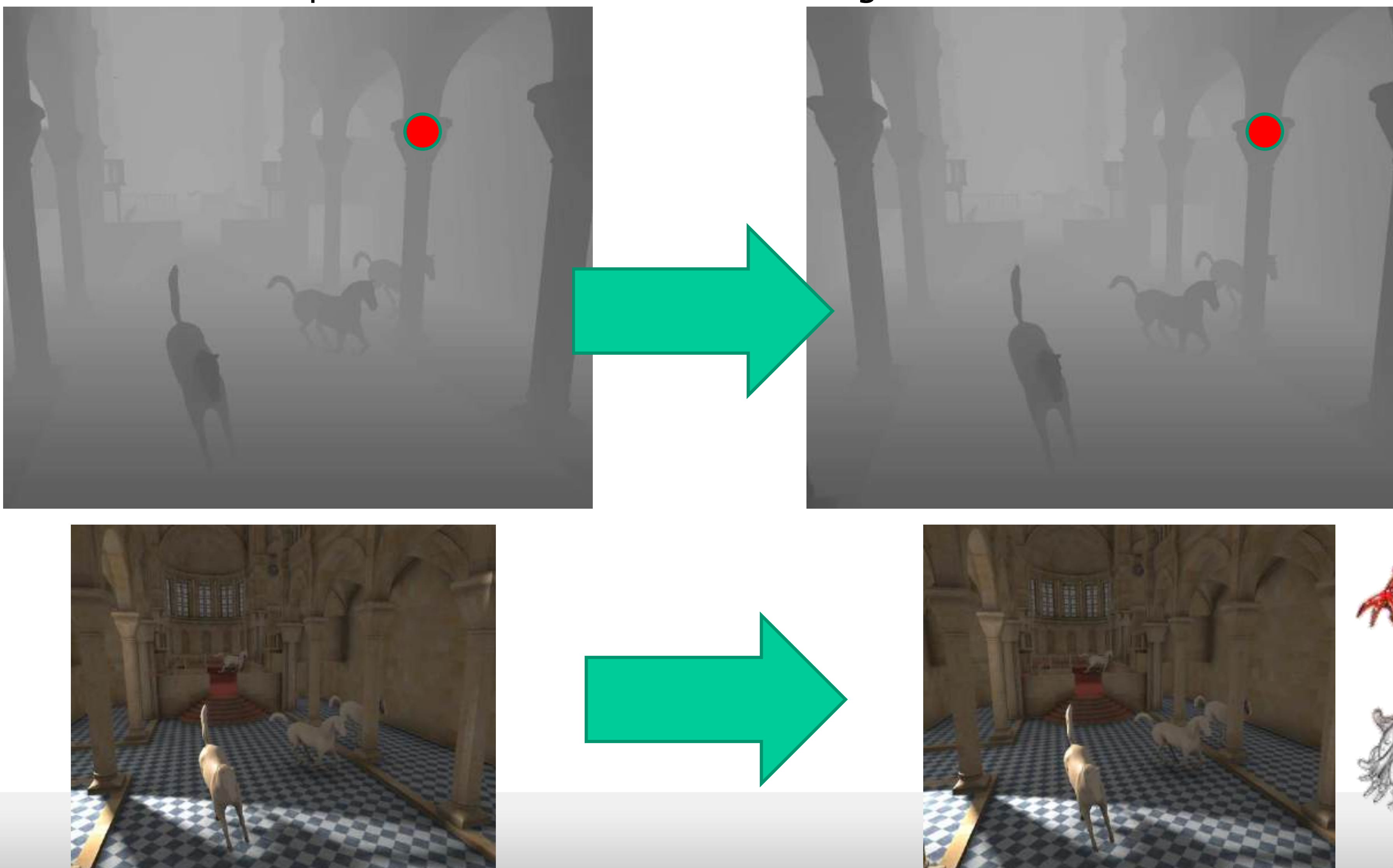
How to reduce shading cost?

- Observation: shading correlates with geometry
- World information behind pixel is for "free"
  - Depth (position)
  - Normals
  - Materials, Textures
  - Geometric motion flow



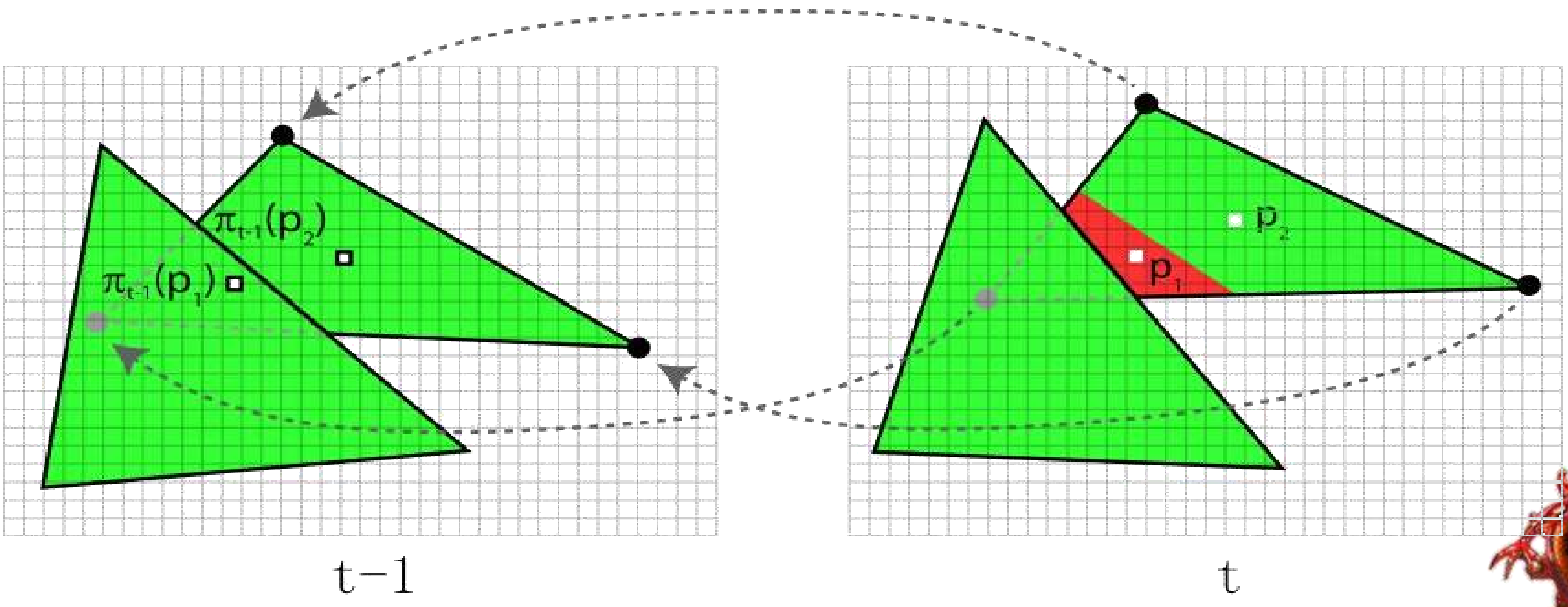
Why does rendering of depth & co. help?

- Find correspondences and transfer shading!



SIGGRAPHASIA2011 HONG KONG

Not that simple...



Forward Reprojection

- Requires forward motion vectors
- Holes and gaps
- Difficult to implement with DX9/10

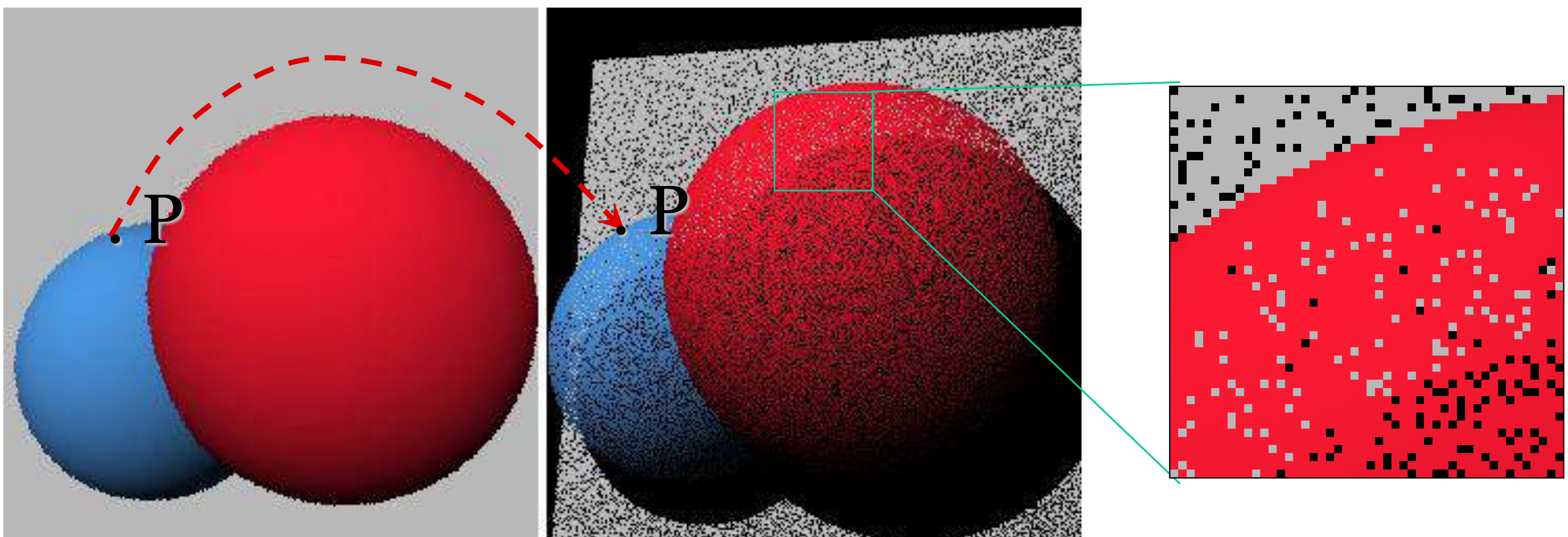


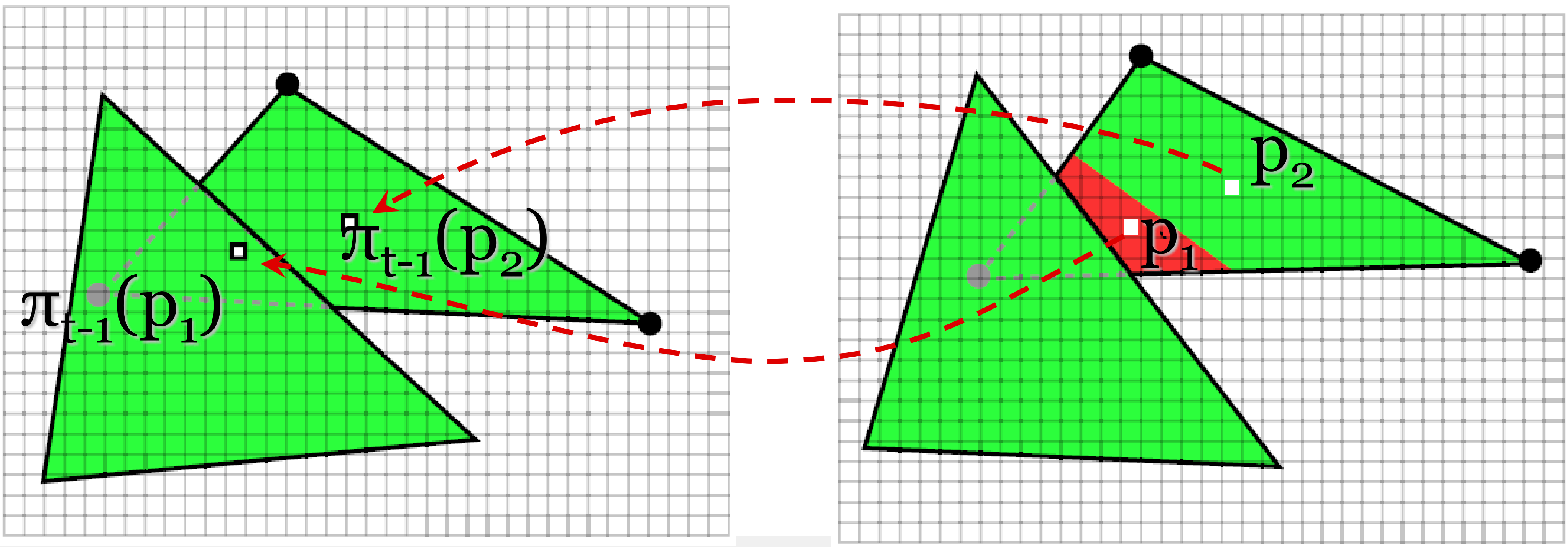
Image courtesy of Bruce Walter

cache (f_{t-1}) new frame (f_t)

SIGGRAPHASIA2011 HONG KONG

Reverse Reprojection
[Nehab 06/07, Scherzer 07]

- Reprojection operator $(x', y', z') = \pi_{t-1}(p)$
- Resolve occlusion: Test if $z' \approx d_{t-1}(x', y')$



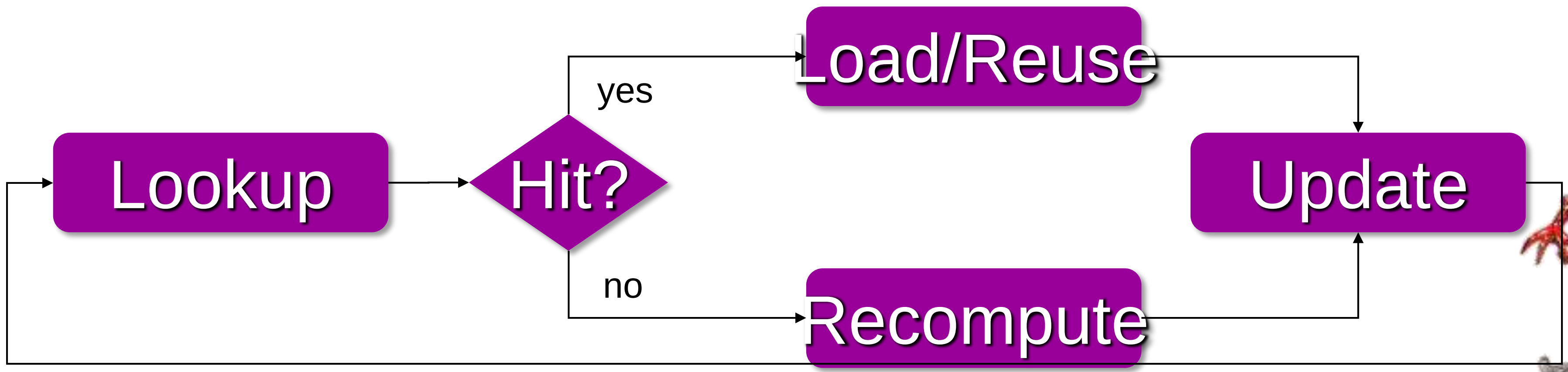
cache (f_{t-1}) new frame (f_t)

Reality Check

- Regular rendering loop (without using TC)
  - Recompute every pixel with original pixel shader

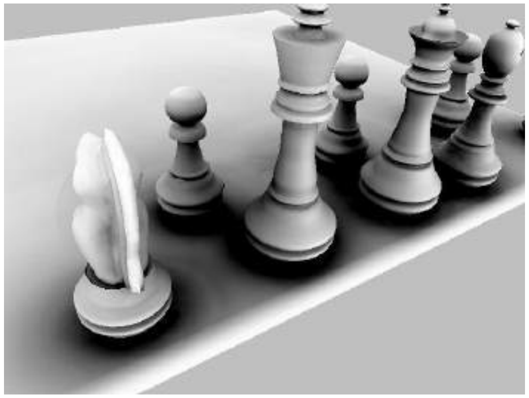
Reality Check

- Reuse previous results using the RRC
  - Reshade on demand
  - Cache reuse path must be cheaper for acceleration

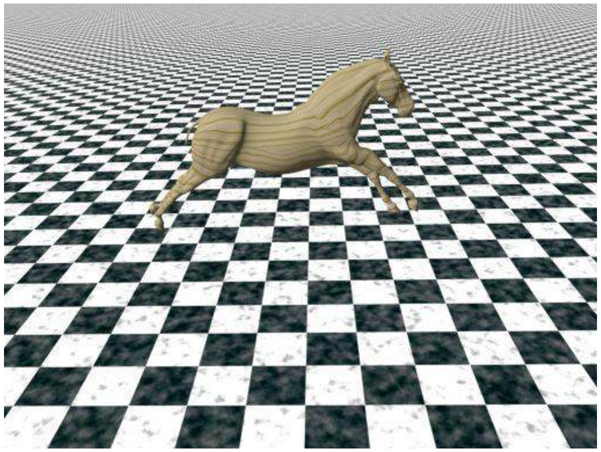


Good Examples to Cache

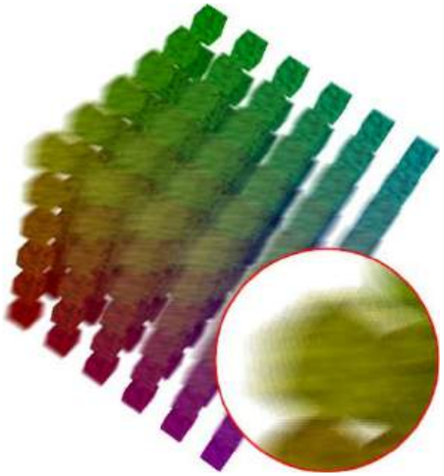
Static procedural texture



Global illumination



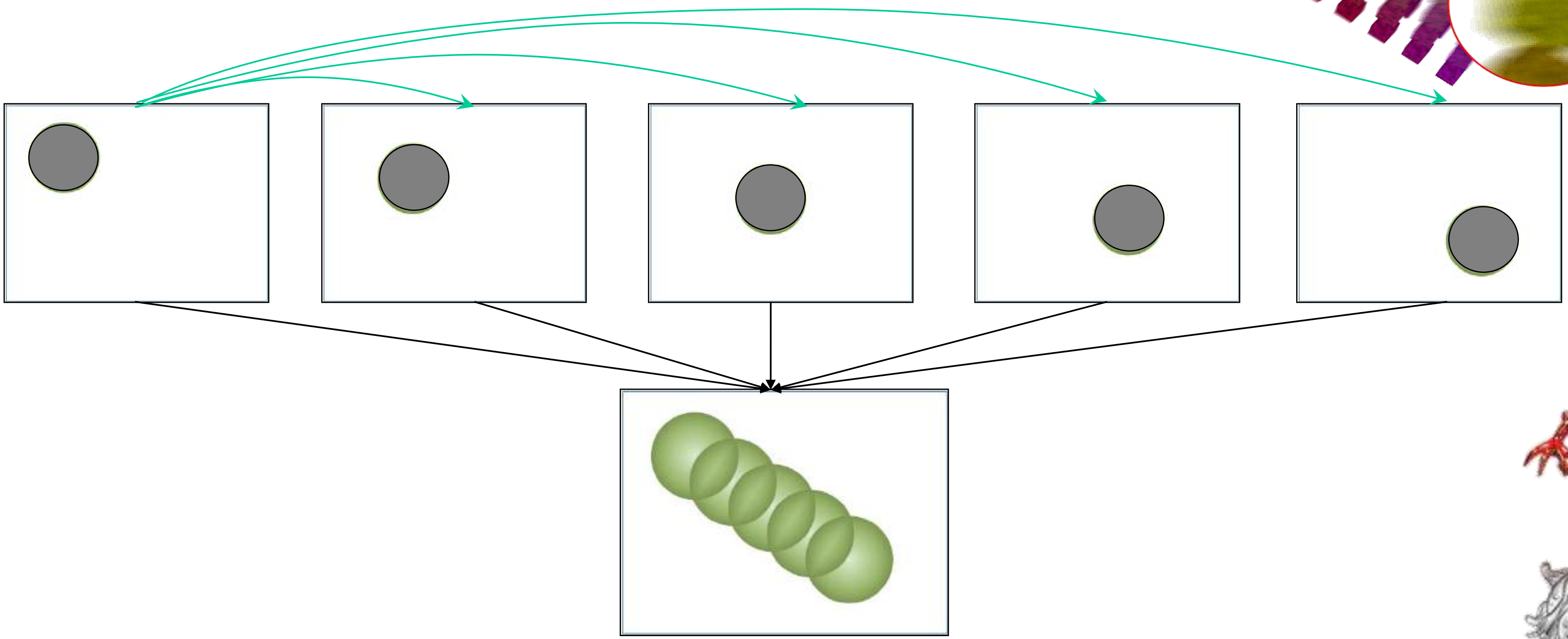
Numerical integral



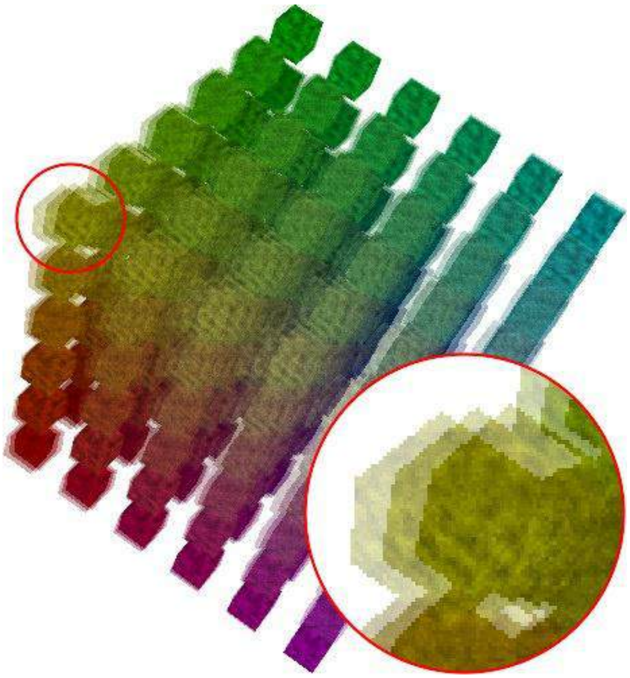
Multi-pass effects

Multi-pass Rendering Effects

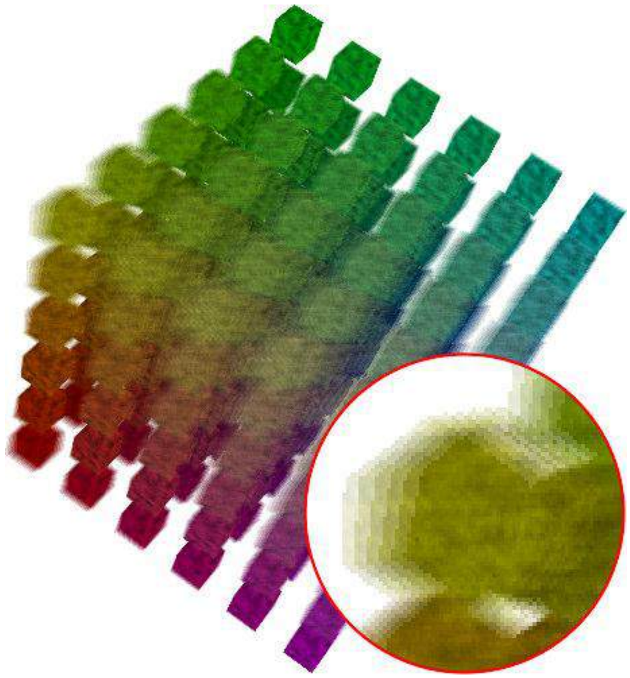
- Render a set of images with similar viewpoints
  - Shade one
  - Shade others via reprojection



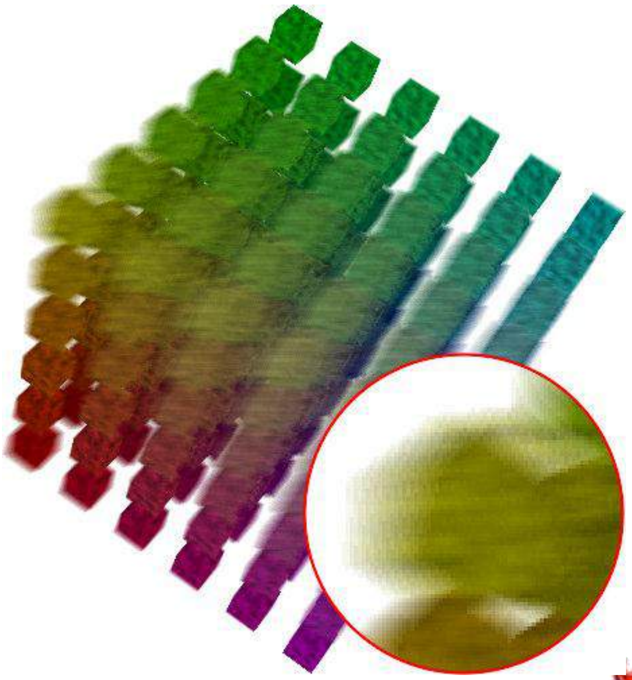
Motion Blur



3 time samples
60fps brute-force
60fps RRC

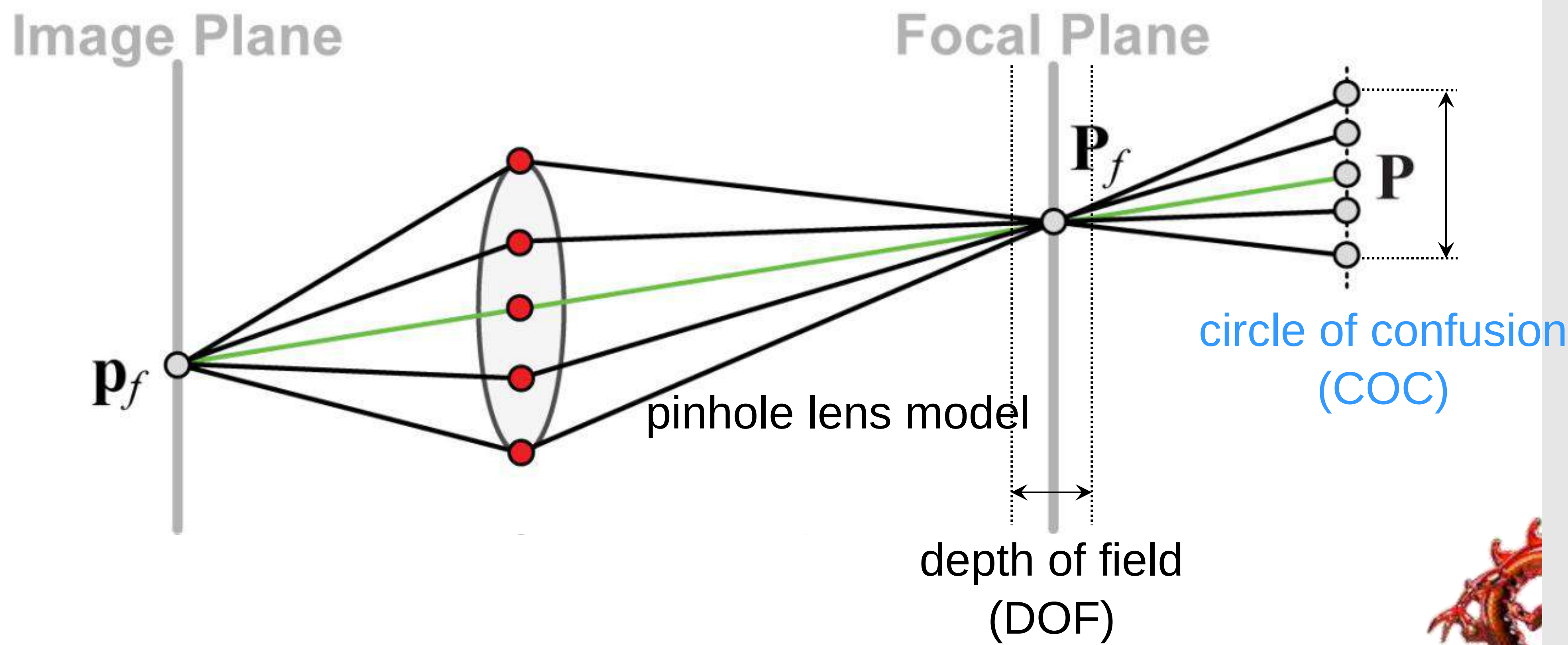


6 time samples
30fps brute-force
60fps RRC



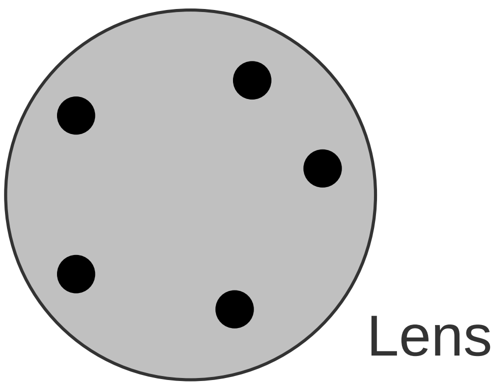
14 time samples
13fps brute-force
30fps RRC

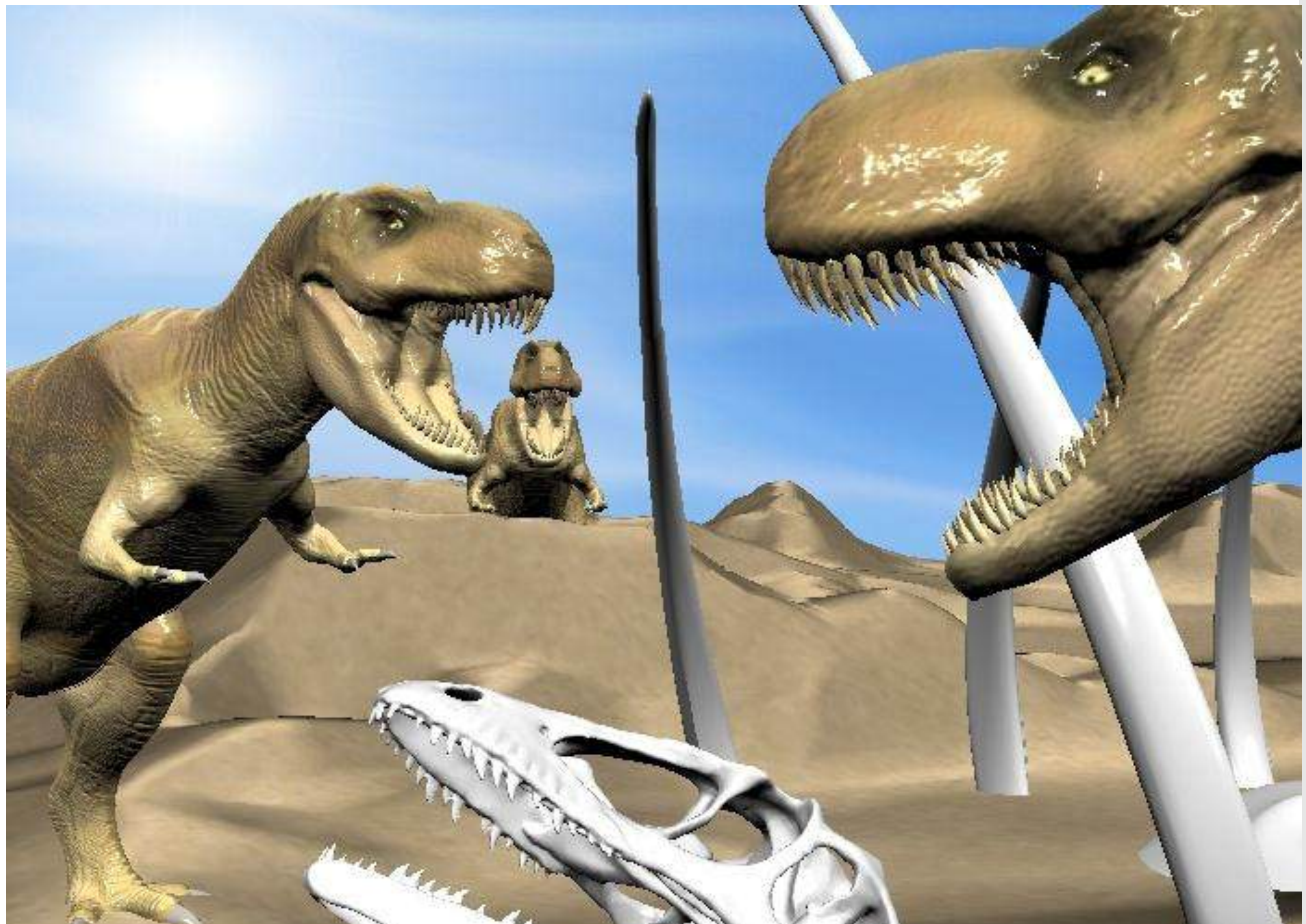
Example: Depth of Field



Our Algorithm

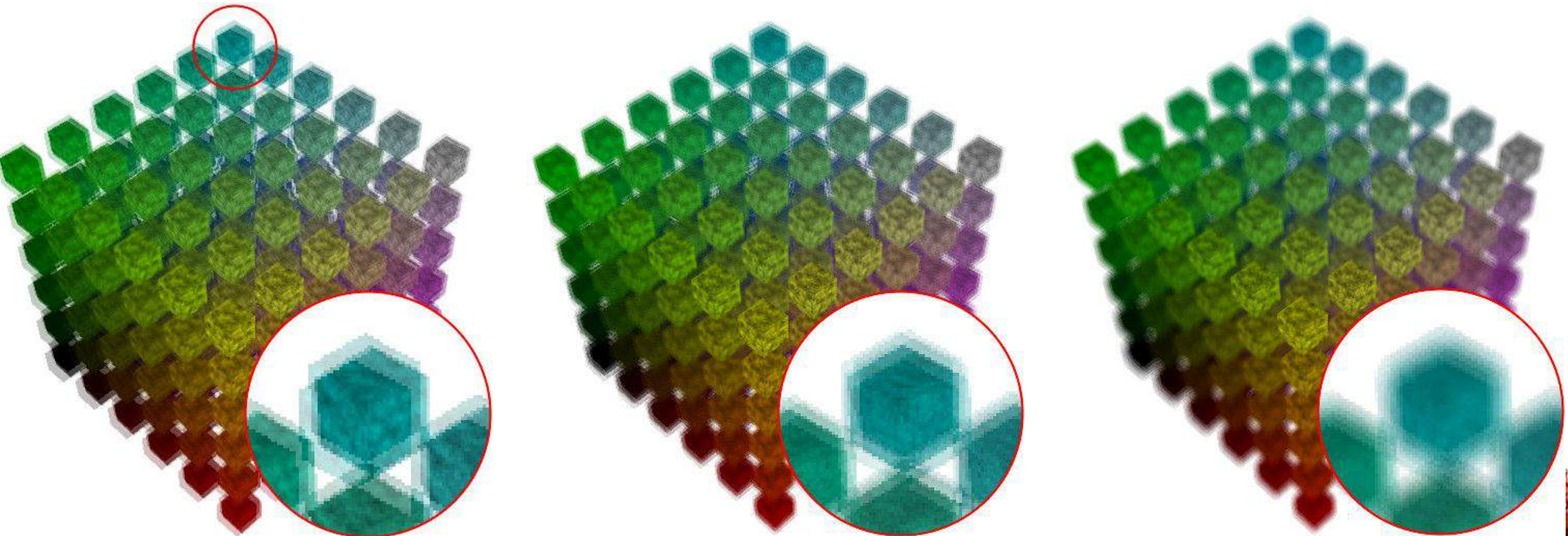
- View synthesis using image-based ray tracing





A few more "tricks" and you get…

Depth of Field



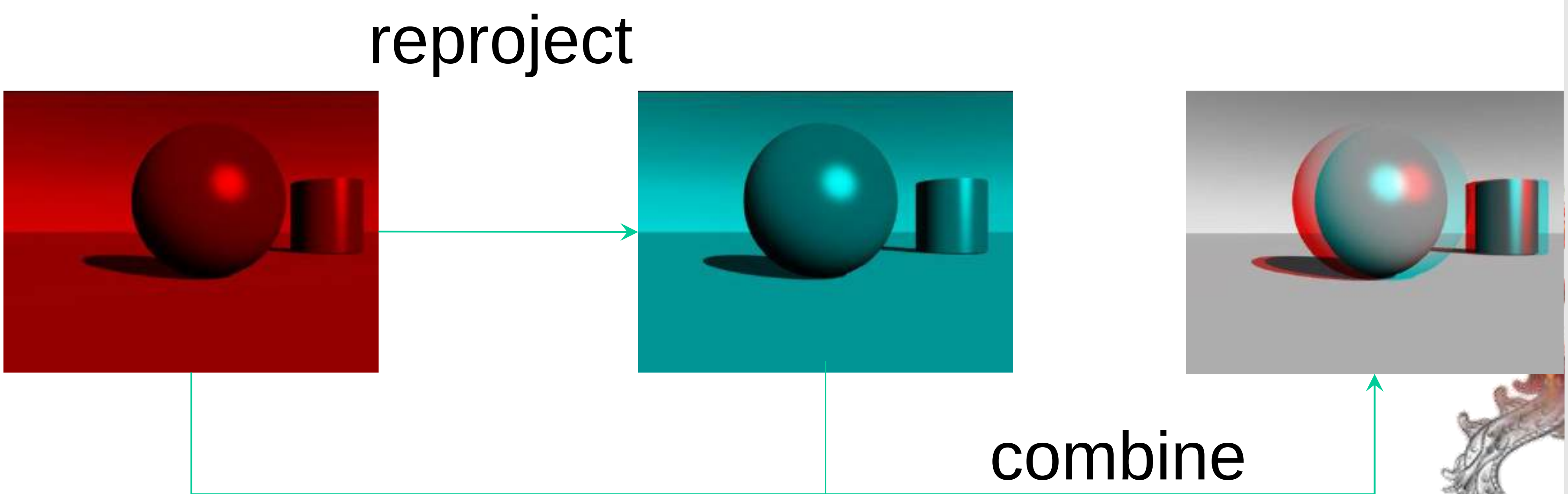
4 aperture samples
45fps brute-force
45fps RRC

9 aperture samples
20fps brute-force
45fps RRC

20 aperture samples
8fps brute-force
20fps RRC

Stereoscopic Rendering

- Generate images from two nearby views
  - Render the left eye normally
  - Render right eye with reprojection



This sounds amazing, but…

- So far: everything was static!
- Nothing moved… !

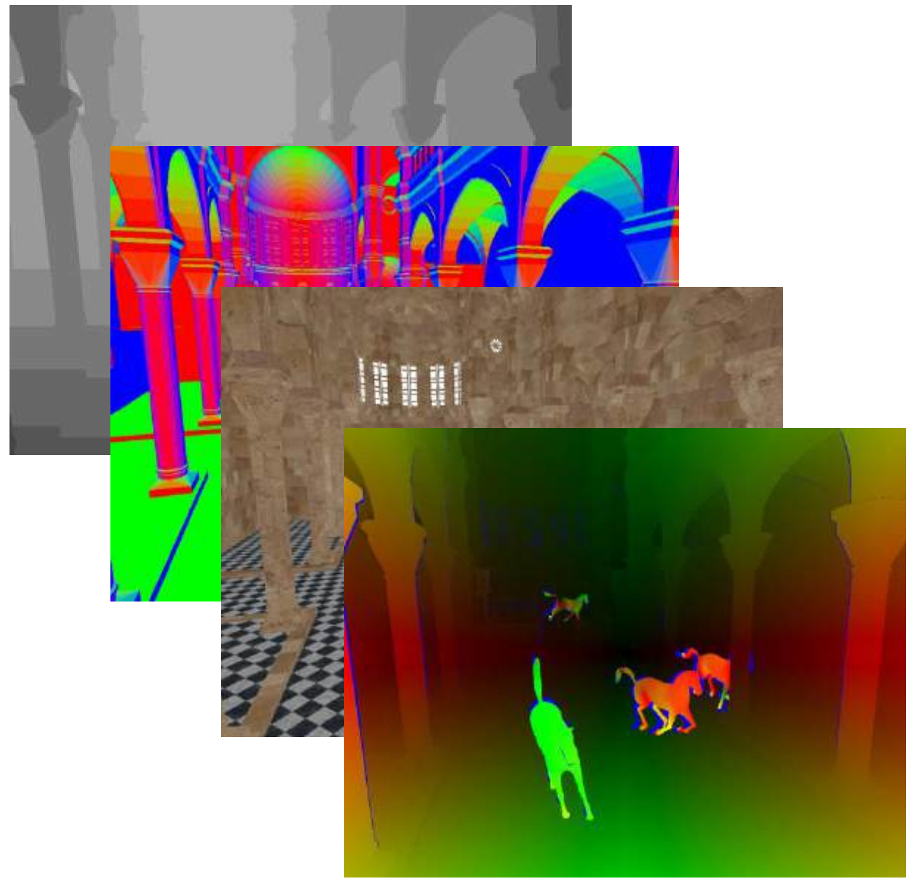
- How to deal with temporal changes?
  - Can we exploit spatial coherence?

Idea: use low resolution, then upsample

- Exploit spatial coherence:

Remember?

- Observation: shading correlates with geometry
- World information behind pixel is for "free"
  - Depth (position)
  - Normals
  - Materials, Textures
  - Geometric motion flow



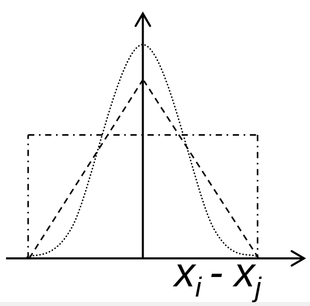
Joint-Bilateral Spatial Upsampling

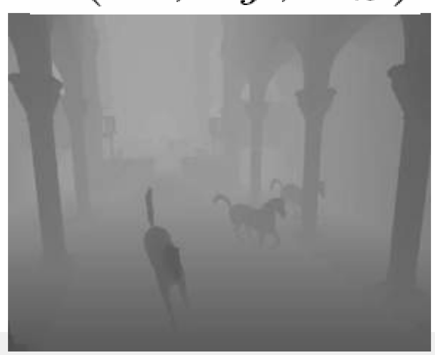
Non-linear interpolation steered by geometry:

$$h(i) = \frac{1}{\sum w_s} \sum_{j \in N\{i\}} w_s(i,j) \cdot l(\hat{j})$$

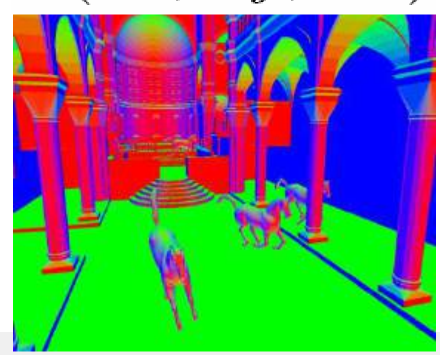
Low-res. shading input

$$w_s(i,j) = \underbrace{k(i,j)}_{\text{Image-space filter (e.g. hat/ box)}} * \underbrace{d(z_i, z_j, \sigma_z^2) * n(\vec{n}_i, \vec{n}_j, \sigma_n^2)}_{\text{weights steered by geometry}}$$





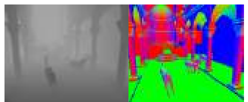
Depth (z)



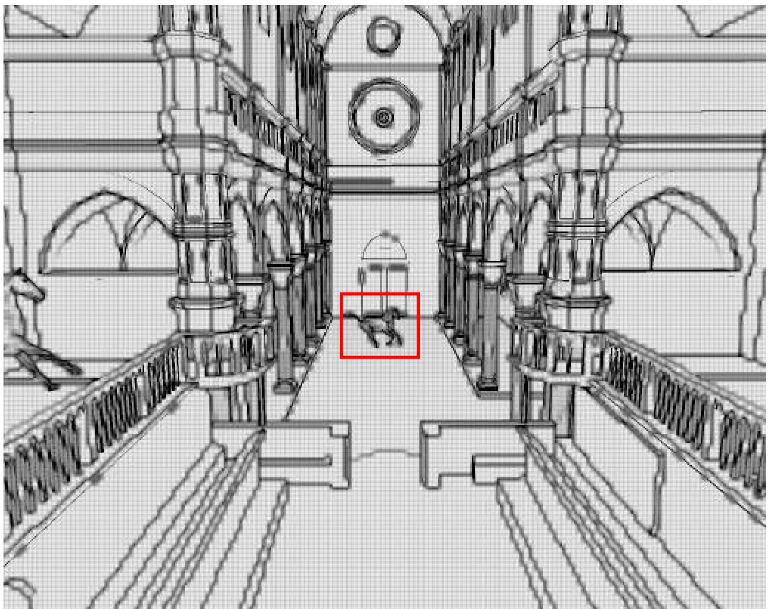
Normals

Joint/Cross-Bilateral Upsampling Revisited

$$\frac{1}{\sum w_s}$$

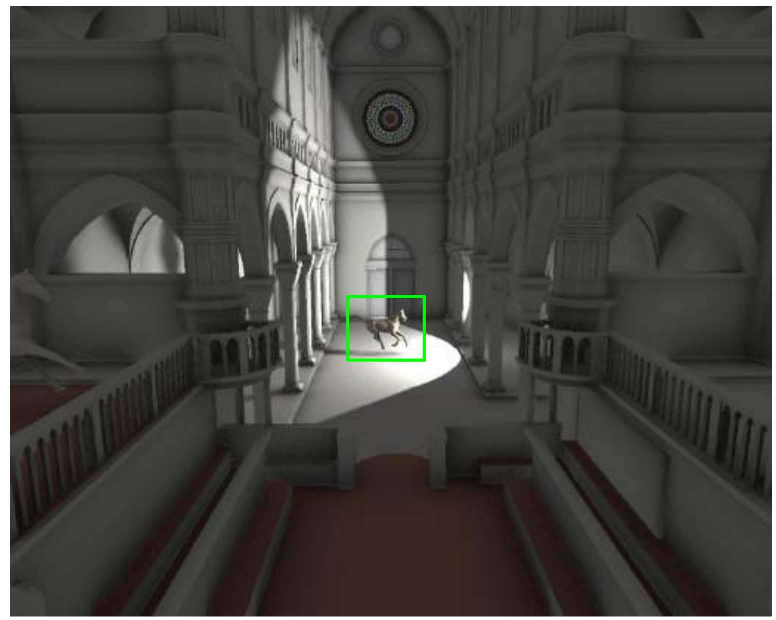


Reference:

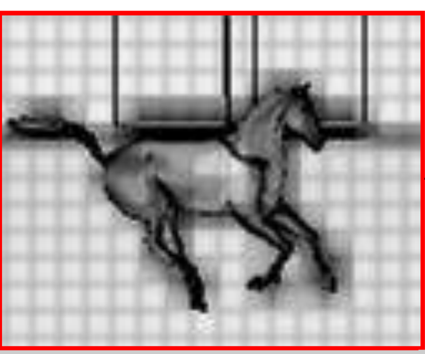


Low-res. shading input

$$w_s(i,j) \longrightarrow$$



High-res. upsampled output



*

≈

Spatio-Temporal Upsampling

- Choose preferable method:

combine spatial upsampling

& temporal caching

Gain information over time?

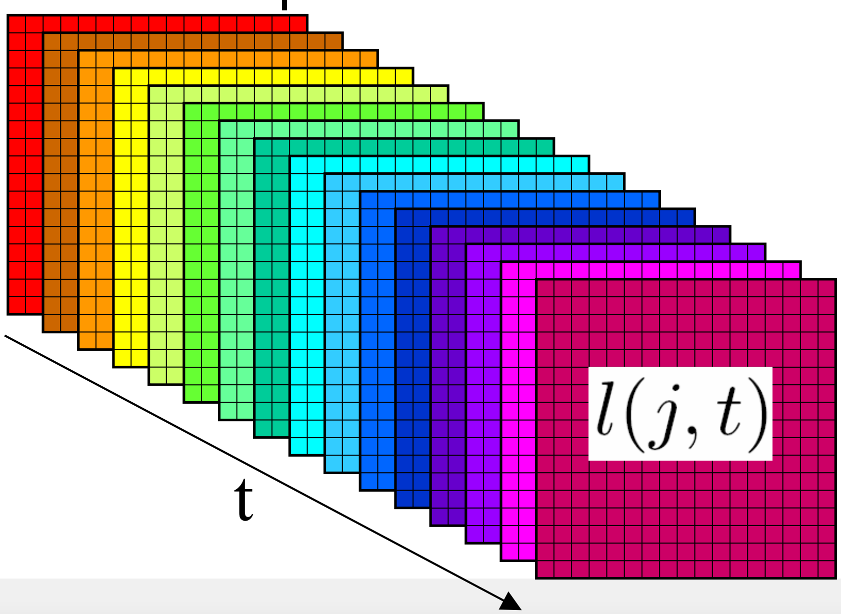
- The same low-res image gives the same information…

Temporally Interleaved Sampling

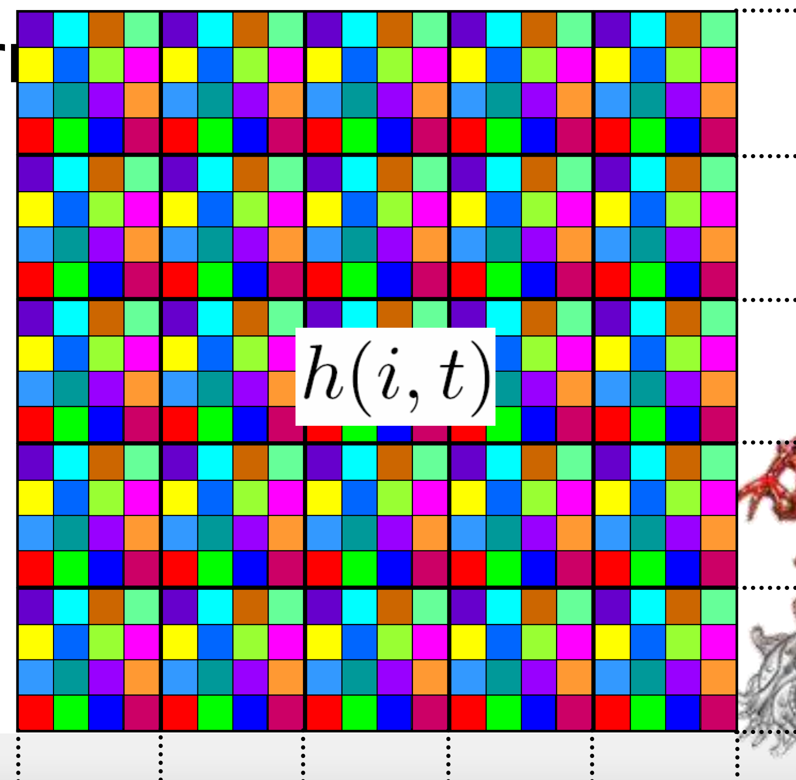
- Cache different pixel positions to upsample over time
  - → Refresh out-dated pixels (e.g. every **k** × **k** frames)

Regular sampling pattern

n x *n* pixels



4*n* x 4*n* pixels



SIGGRAPHASIA2011 HONG KONG

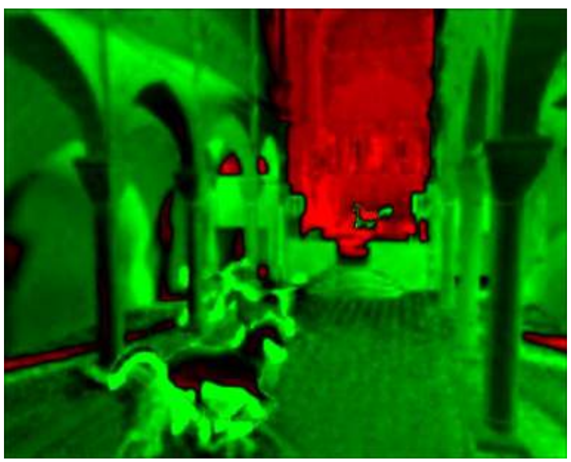
Putting things together:

- temporal

  Jittering -> more information for static over time

- Spatial

  Bilateral Upsampling (low2high) -> responsiveness

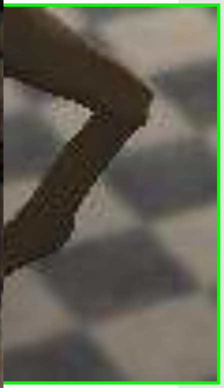


Choose according to change

4x4 upsampled result

Static Frame Convergence

SIGGRAPHASIA2011 HONG KONG



Spatio-Temporal Upsampling [Herzog et al. 2010]

- Beneficial to use

    Spatial

  & temporal upsampling

  - Static frame convergence

  - Robustness with respect

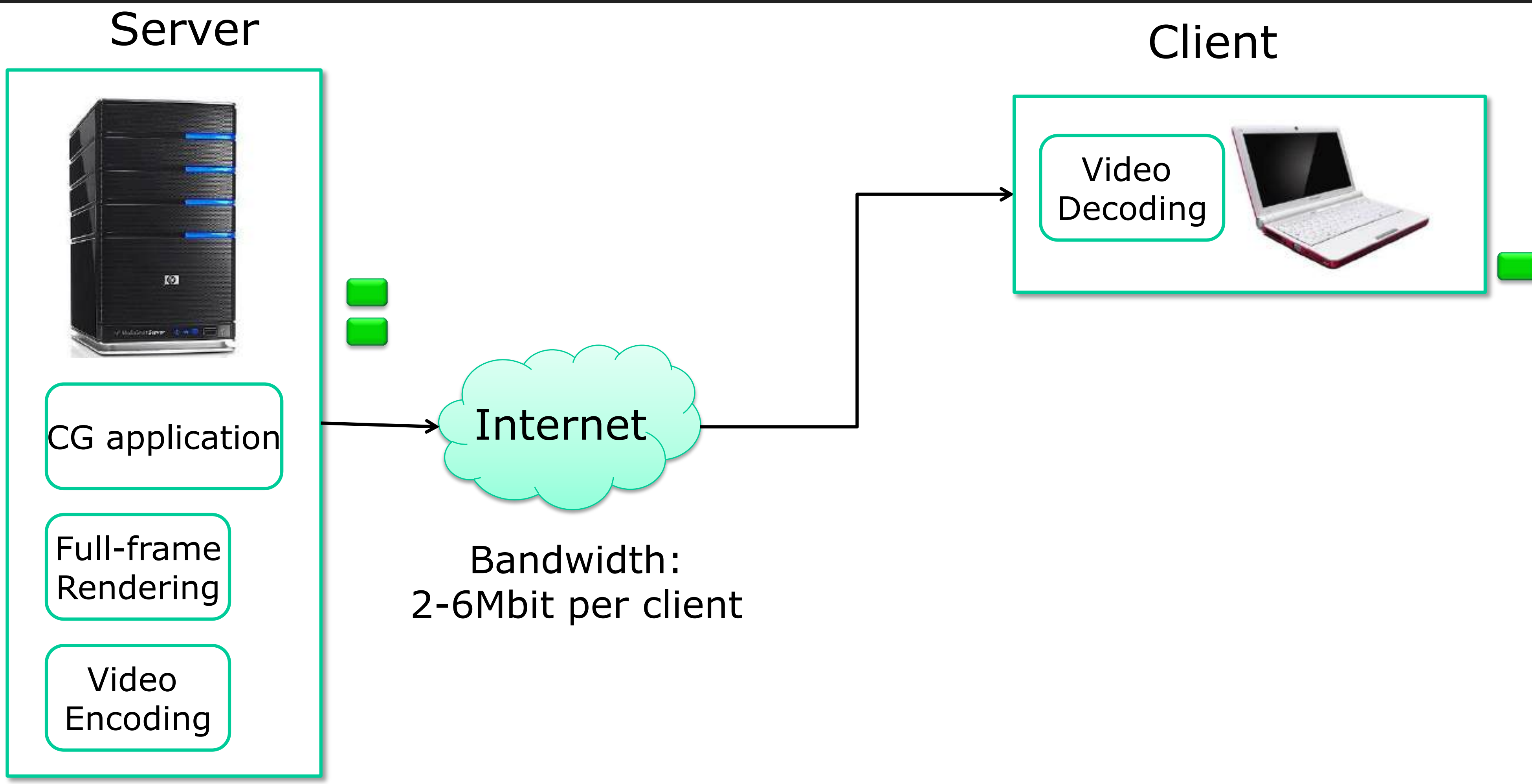
    to changing lighting conditions

Extension: Remote Rendering

- OnLive, OToy, Gaikai rely on video encoding

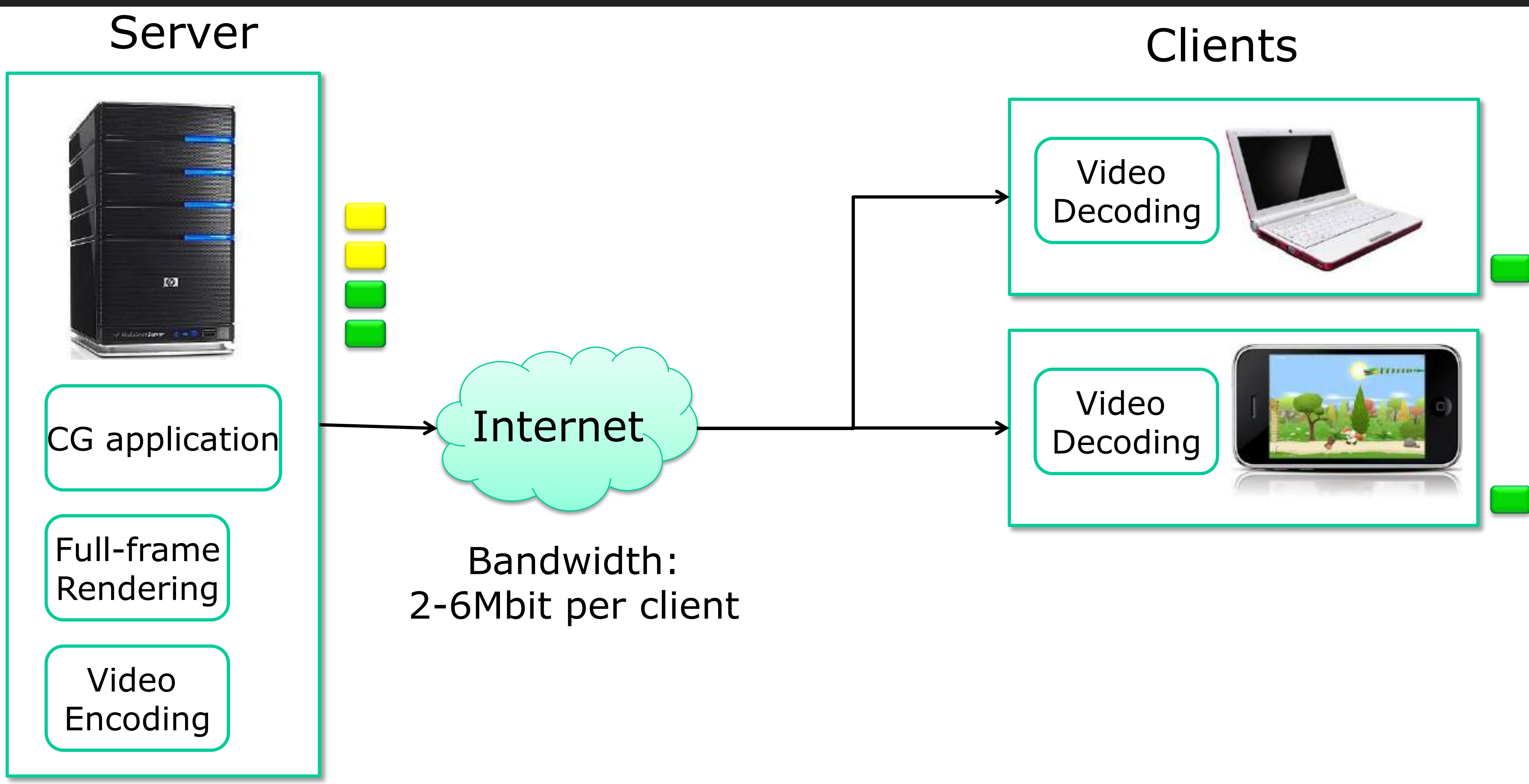
  Naturally exploit coherence in video

Streaming for Rendered Content
[Pajak et al. 11]



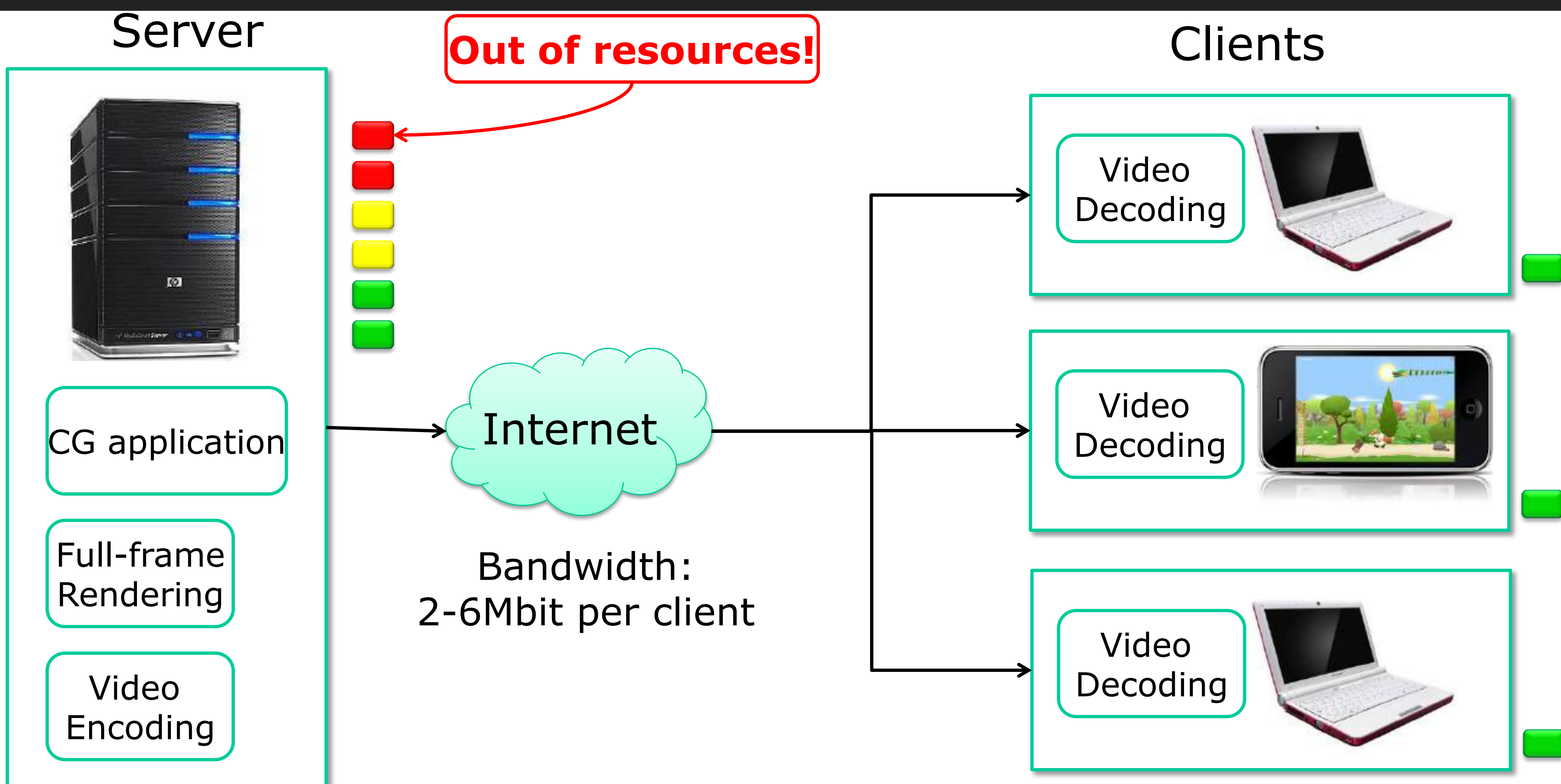
Streaming for Rendered Content

[Pajak et al. 11]



Streaming for Rendered Content
[Pajak et al. 11]

SIGGRAPHASIA2011 HONG KONG



Design similar to current commercial solutions

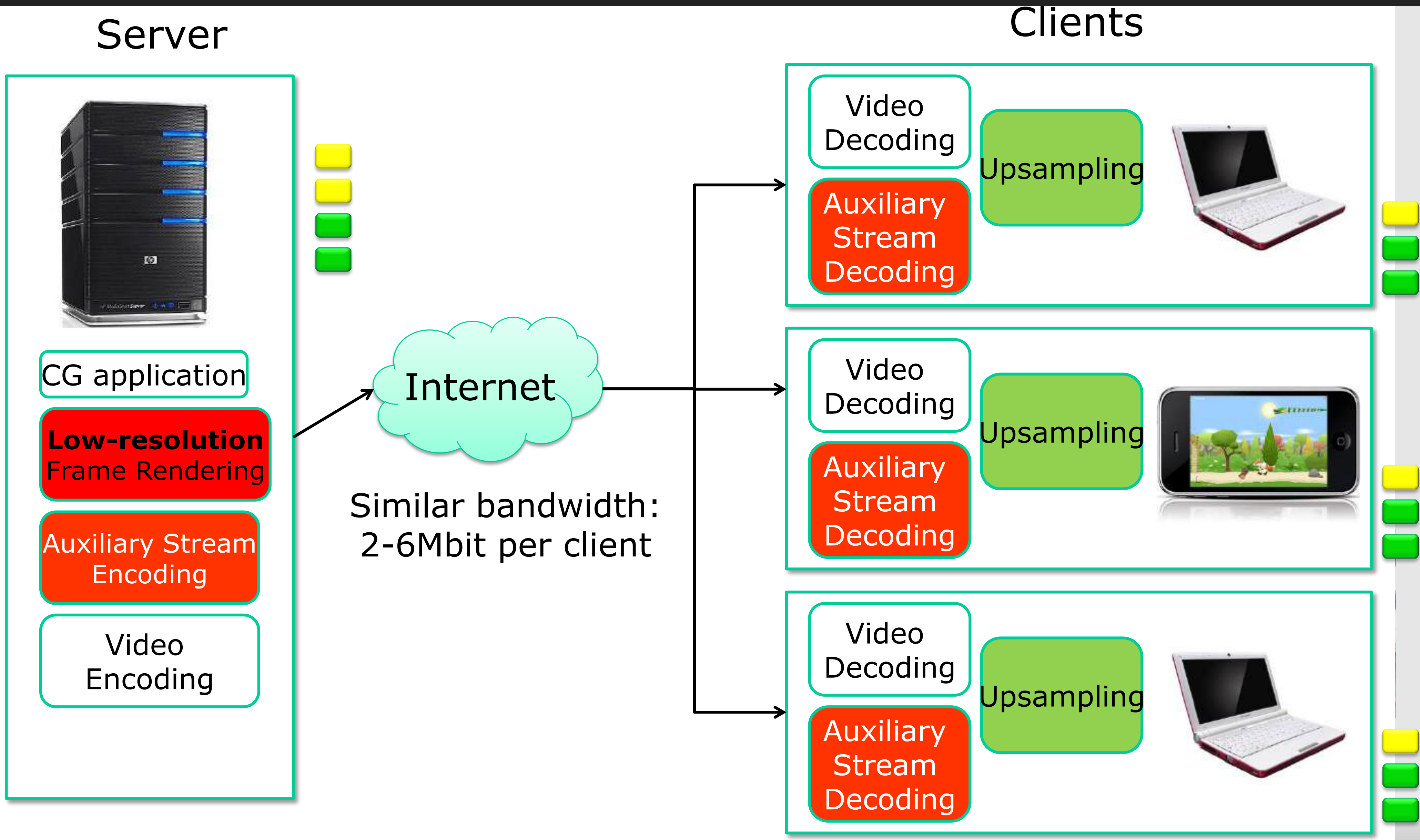
Streaming for Rendered Content
[Pajak et al. 11]

- Rely on spatio-temporal upsampling strategies
  - Less bandwidth
  - Less server workload

- Specialized Encoding

Streaming for Rendered Content

[Pajak et al. 11]



Streaming for Rendered Content

[Pajak et al. 11]

H264

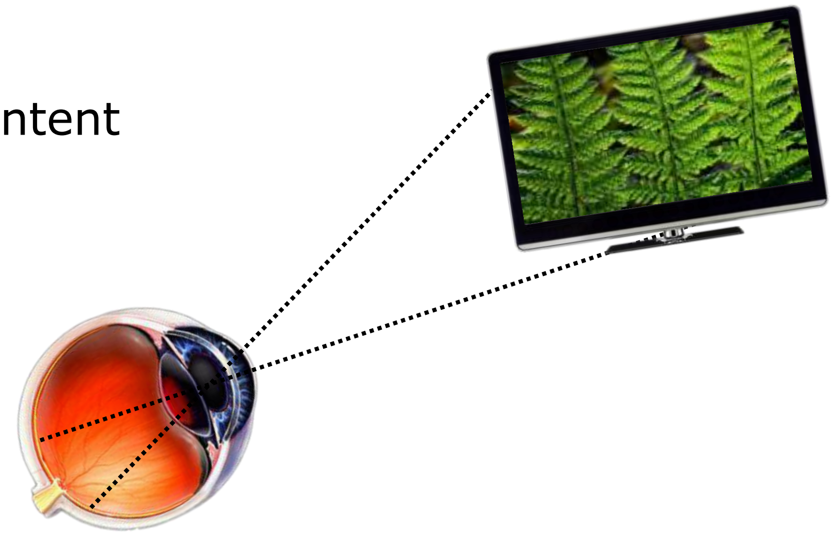
Pajak et al. solution + more

Image-Space Coherence

- Very efficient
- Easy to implement
- Adapted to Graphics pipeline
- Important for streaming architectures

Exceed display limitations

- Idea: Temporal coherence to enrich content
  - Even beyond physical limits



- Examples:
  - Color bit depth: Frame Rate Control
  - Hold-type effect reduction: Temporal Upsampling
  - Resolution: Apparent Resolution Enhancement

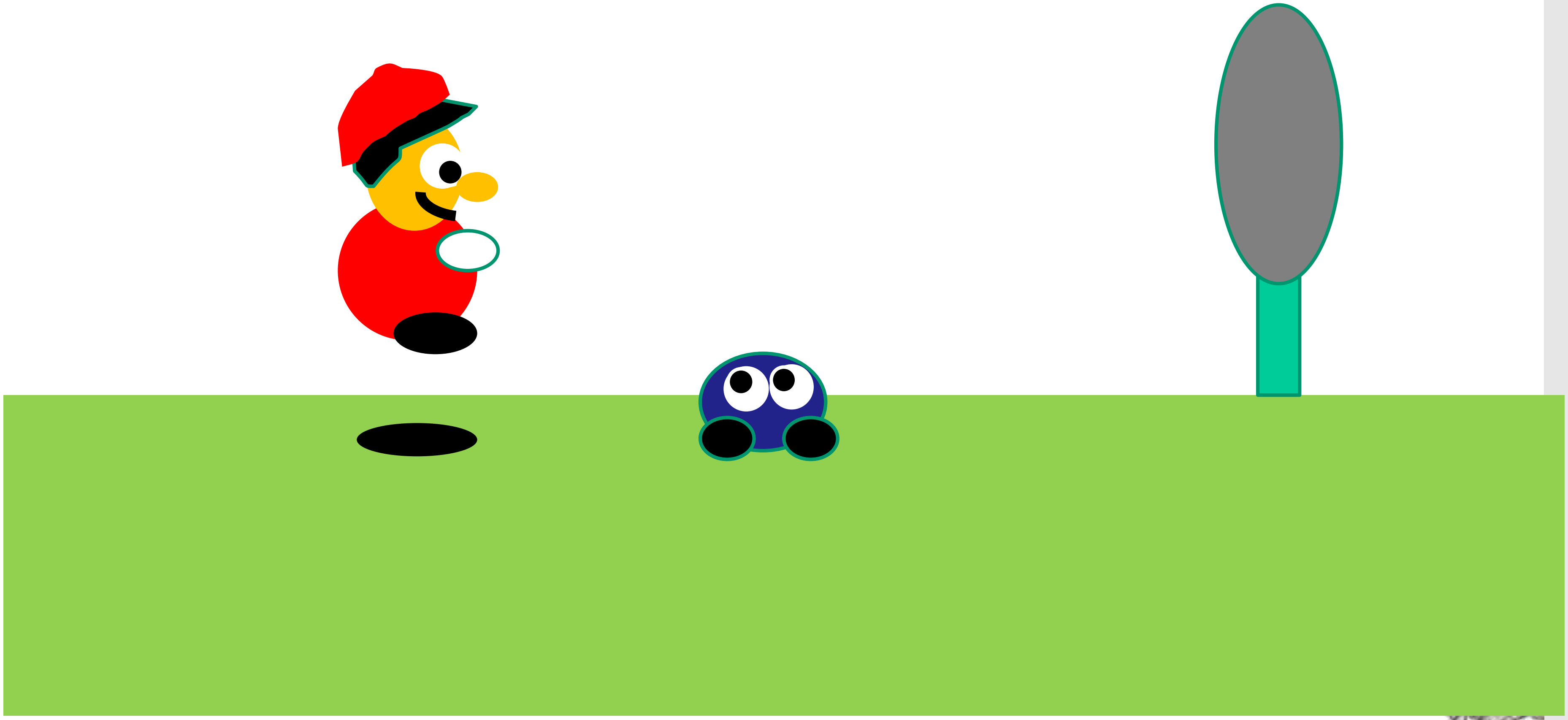
Color Bit Depth: Frame Rate Control [Art04]

- Use eye latency to integrate color sequences
  - Similar principle as DLP projectors

-> Flicker different colors and have eye average them

Effect known from older video games

- Virtually augment the color palette



Flickering even works for >8 bit

- Fight mach banding artifacts
- Manually:
  - Switch last color bit
- Useful for HDR imagery,

  but very high refresh rates needed…

SIGGRAPHASIA2011 HONG KONG

Display Improvement

1990's | 2000's | Today | Future

We are here

High refresh rate

more than 120Hz

Low brightness

Flicker for low rates

No flickering

Higher level of luminance

Low refresh rate - ~60Hz

Long response time

Brighter

Better contrast

Low response time

High refresh rate

Exploit HVS to improve quality

Small response time

Higher refresh rate

Better colors,

Better contrast

Better brightness

…

Less expensive ;)

Hold-Type Blur Reduction [Didyk10]

- Exploit limitations of the HVS

original frames
+ motion flow & depth
(40Hz)

Efficient
Perceptually-motivated
GPU interpolation

Reduced blur
(120 Hz)

High-Frequency propagation

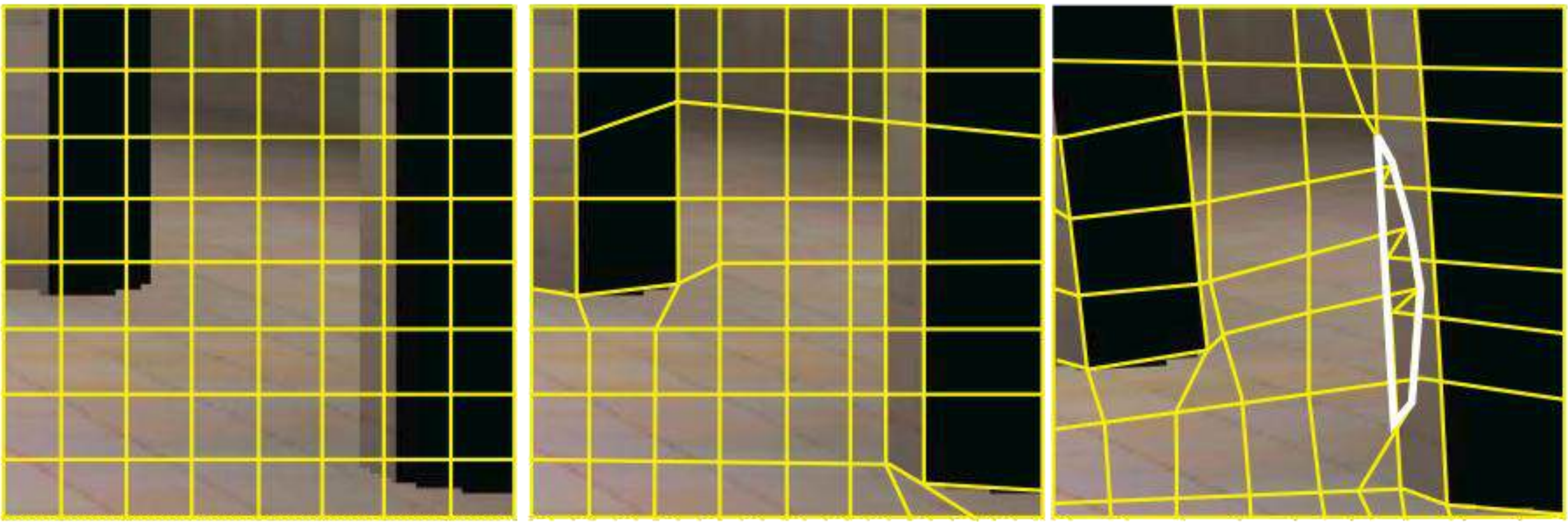
- High-frequency information

  is spread across time at 120Hz

-> Idea: Increase high-frequency in first frame

hide artifacts in extrapolation via blur

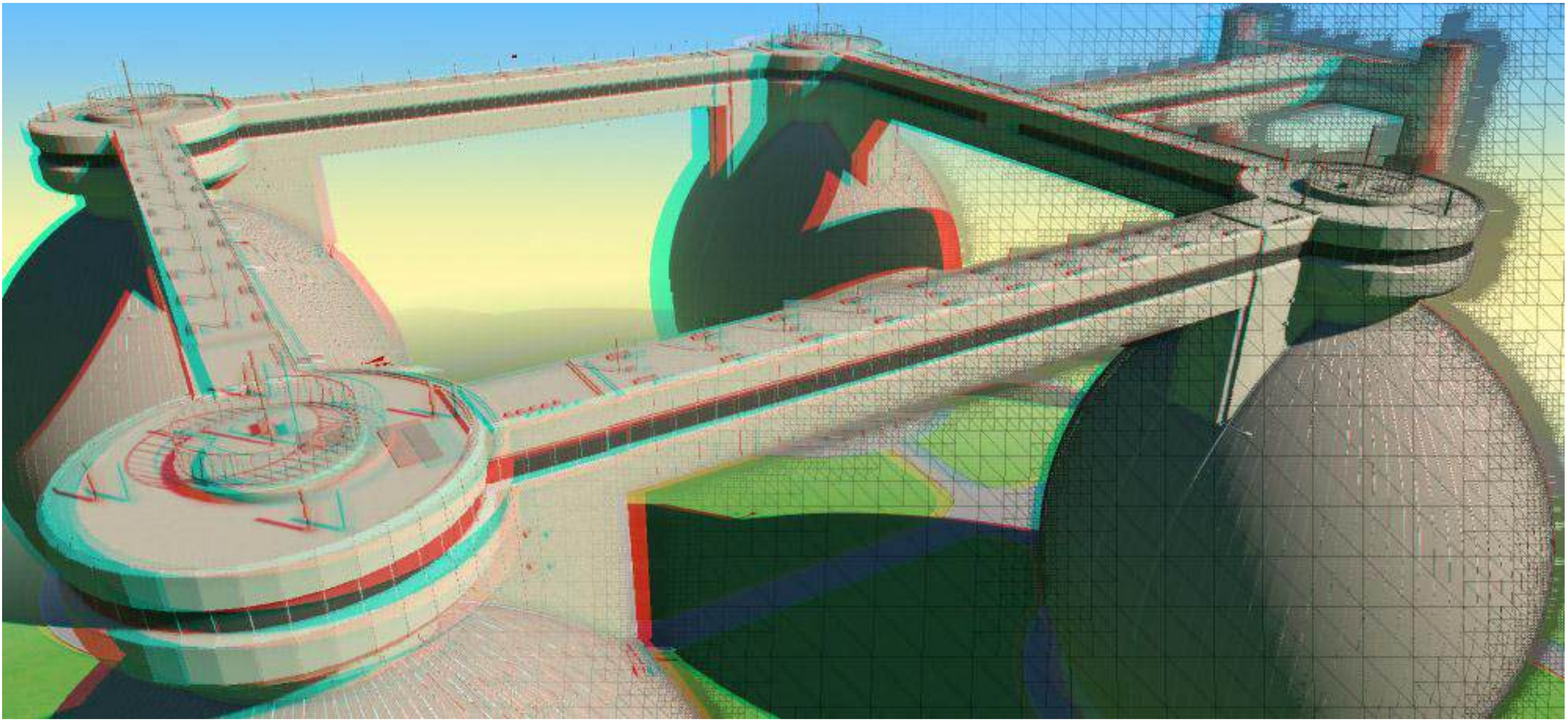
Use a cheap extrapolation technique



- Artifacts will be hidden by blur

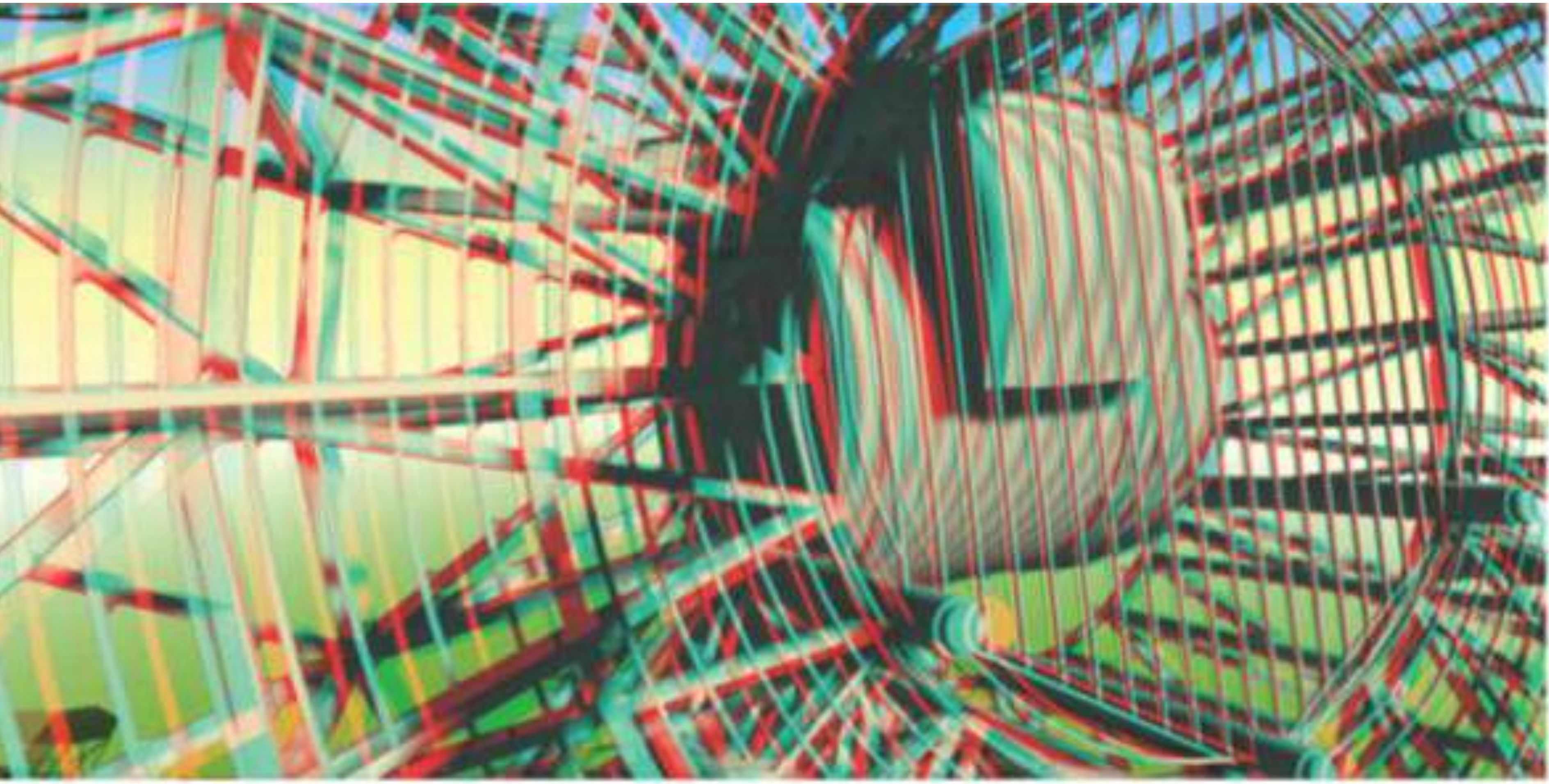
Extension to Stereo

- Adaptive Image-space Stereo View Synthesis [Didyk et al. VMV'10]
- More sophisticated (adaptive) warping



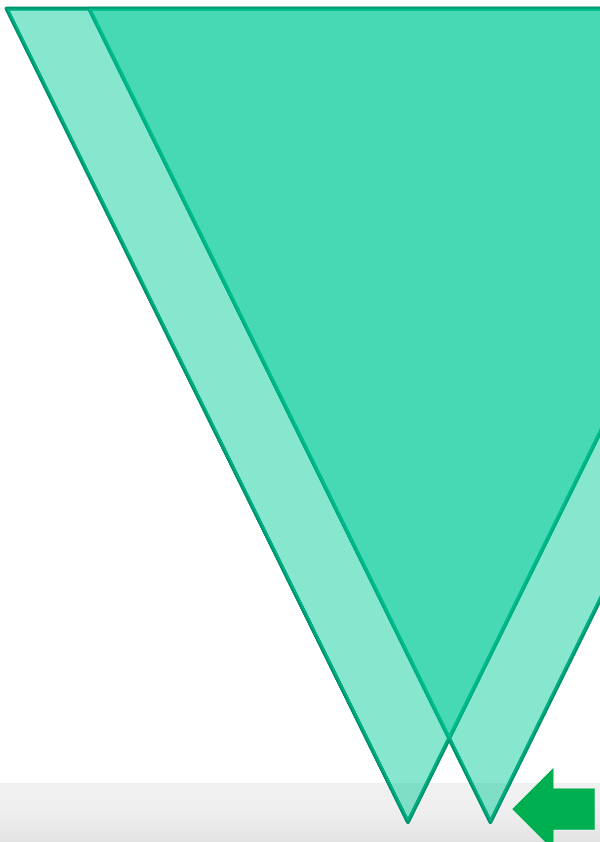
SIGGRAPHASIA2011 HONG KONG

Extension to Stereo - Results



Extension to Stereo

- Temporal coherence of viewpoint
  - Reuse nearby view from previous frame
  - Only render one new view and rely on warping



Viewpoint at time t+1

Extension to Stereo - Results

Reference Previous work Warping Temp. Warping

Warping

- Very cheap alternative to complex methods
- Maps very well to GPU
- Executes in less than 4ms on a full-HD frame
  - NVIDIA GT 460
- Two applications, others exist
  - Hold-type blur reduction and Stereo

Combating Hold-type Blur [DER*10]

- Many advantages:
  - Crispness
  - Quality
  - Task-performance

  - Low overall cost

Can we push blur reduction even further?

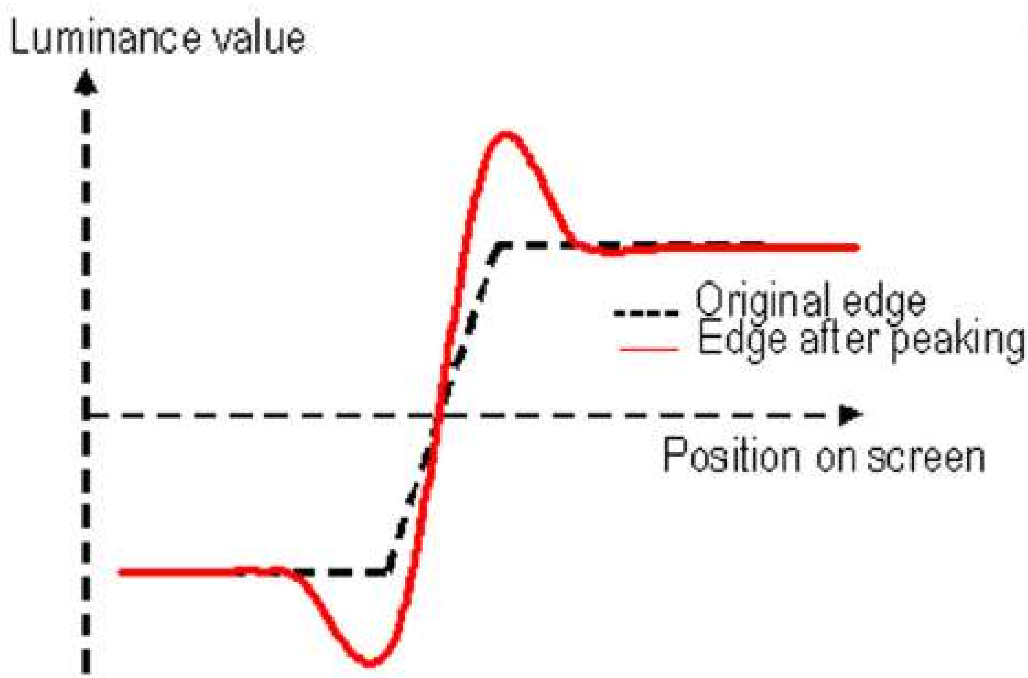
SIGGRAPHASIA2011 HONG KONG

Super-resolution

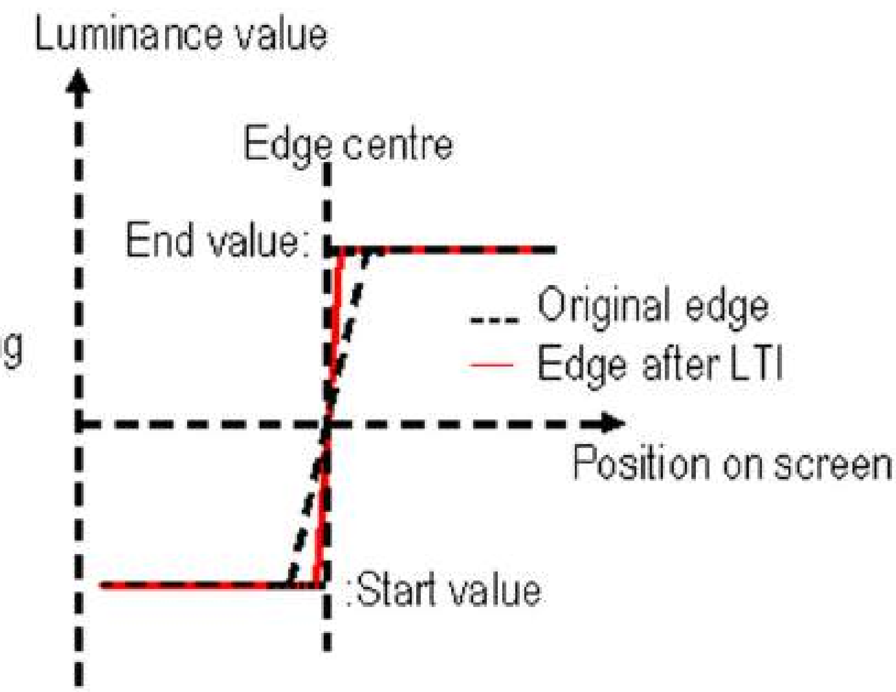
- Upscaling, solved problem, ICs at all PC
  - Does not add new frequencies
- Super-resolution goal: restore high frequencies
  - De-interlacing: images show alias
  - In graphics it is easy to get aliasing
- Typical sharpening algorithms used in TV sets
  - Peaking
  - Luminance Transient Improvement (LTI)
- Temporal domain can also be exploited

Sharpening Filters

Peaking



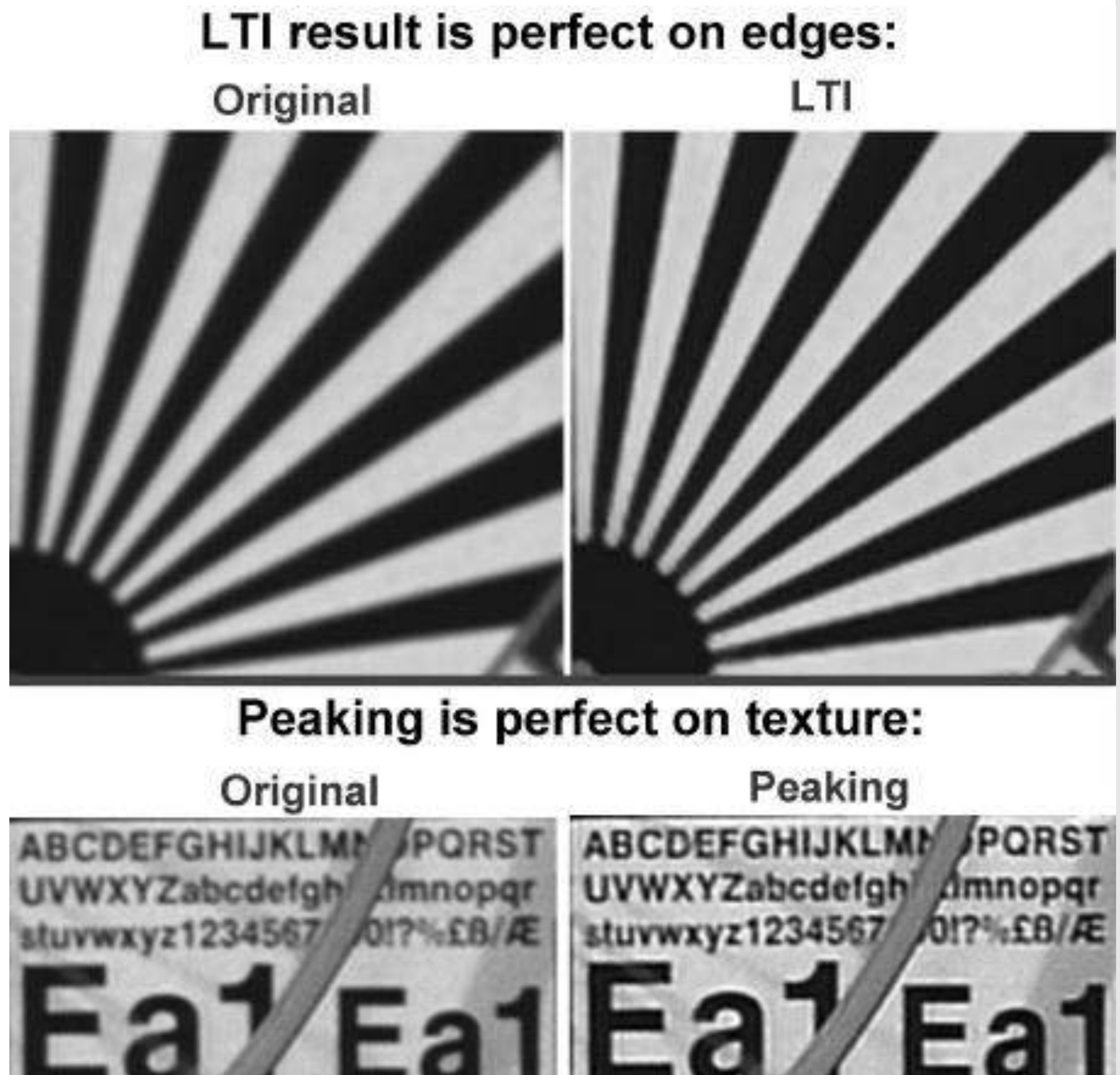
Luminance Transient Improvement (LTI)



M. Zhao, M. Bosma, G. de Haan, Making the best of legacy video on modern displays, 2006 Society for Information Display

Sharpening Filters: Results

- Peaking similar to unsharp masking
- In 3D rendering enhancement of noise signal is not a problem
- In 3D rendering we can better detect object silhouettes
- LTI ~ velocity

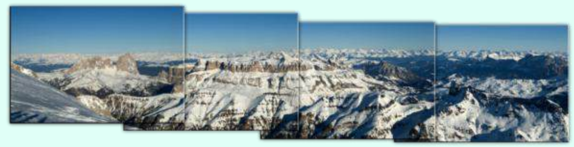


M. Zhao, M. Bosma, G. de Haan, Making the best of legacy video on modern displays, 2006 Society for Information Display

Many High-Resolution Sources

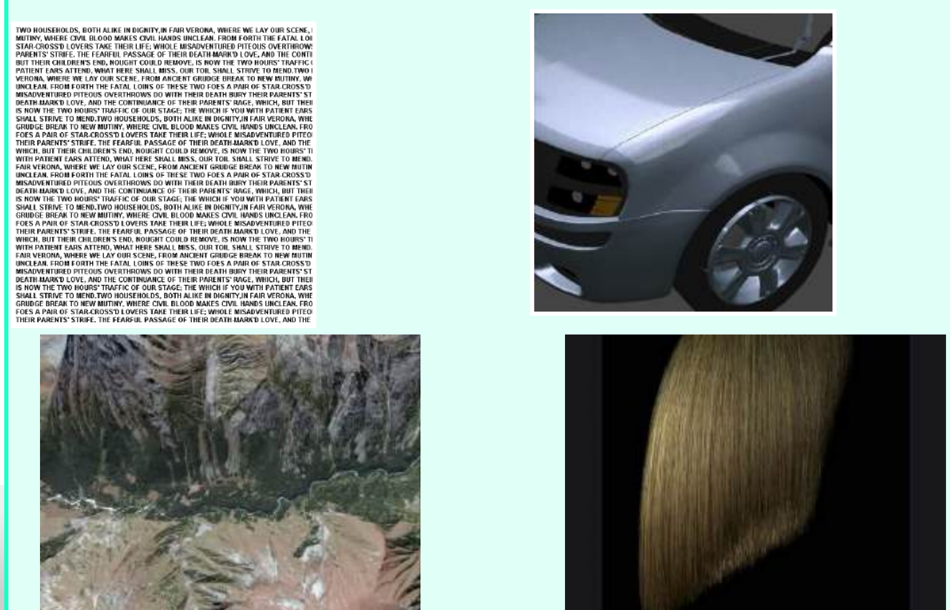
Photographs: > 10MPix

Panoramas: > 50MPix

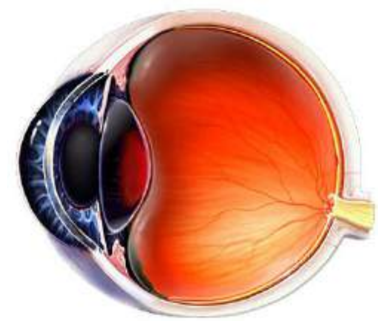


Gigapixel Photography:

Computer generated: Unlimited



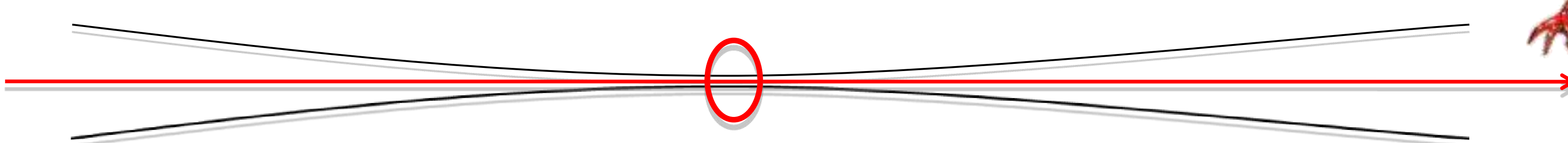
Motivation



easily ~50 MPix

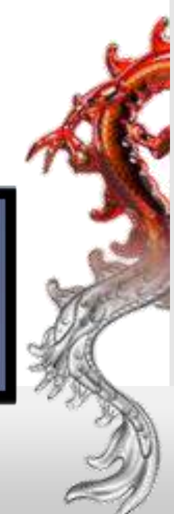
~ 2-8 MPix

1px → > 9 receptors



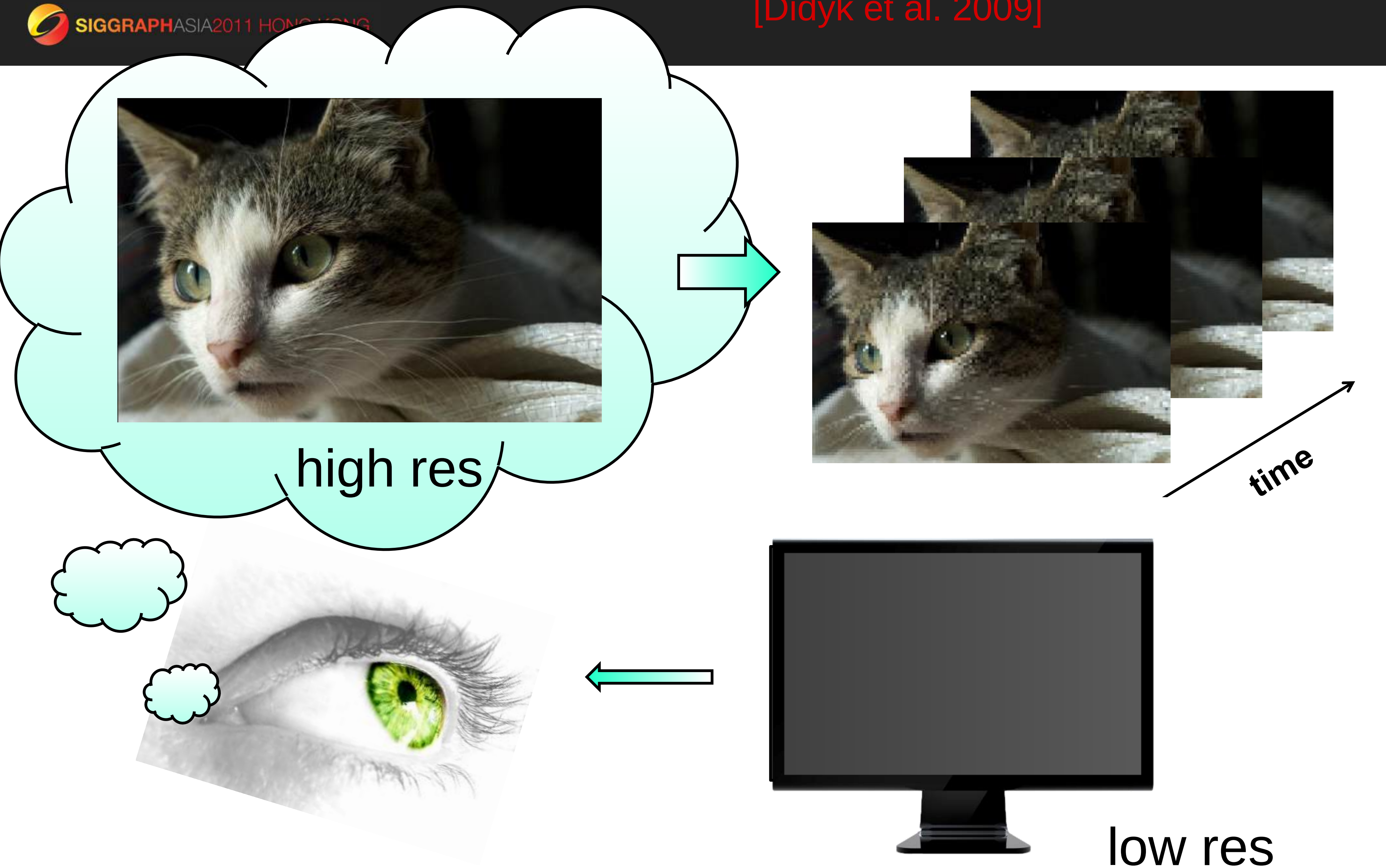
Display content?

Resolution mismatch!



Apparent Resolution Enhancement

[Didyk et al. 2009]



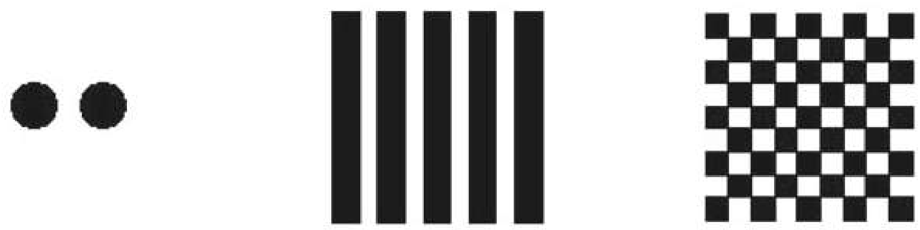
increased apparent resolution

Perception: Spatial Visual Acuity

- Cone density in the fovea may reach 28" (arc seconds)
  - Nyquist's theorem: then 1D sine gratings of 60cycles/deg can be resolved
  - Low-pass filtering in the eye optics removes higher frequencies causing aliasing
- Pixel size at a full-HD desktop display observed from 50cm distance spans 1.5' (arc minutes)
  - In such observation conditions 1 pixel covers roughly 9 cones
  - Estimation valid only for the central fovea region
- Visual *hyperacuity* enables to locate slightly shifted lines in the Vernier acuity task with precision higher than 5" (arc seconds)
  - This more a *localization* task than a *resolution* task

Perception: Spatial Visual Acuity

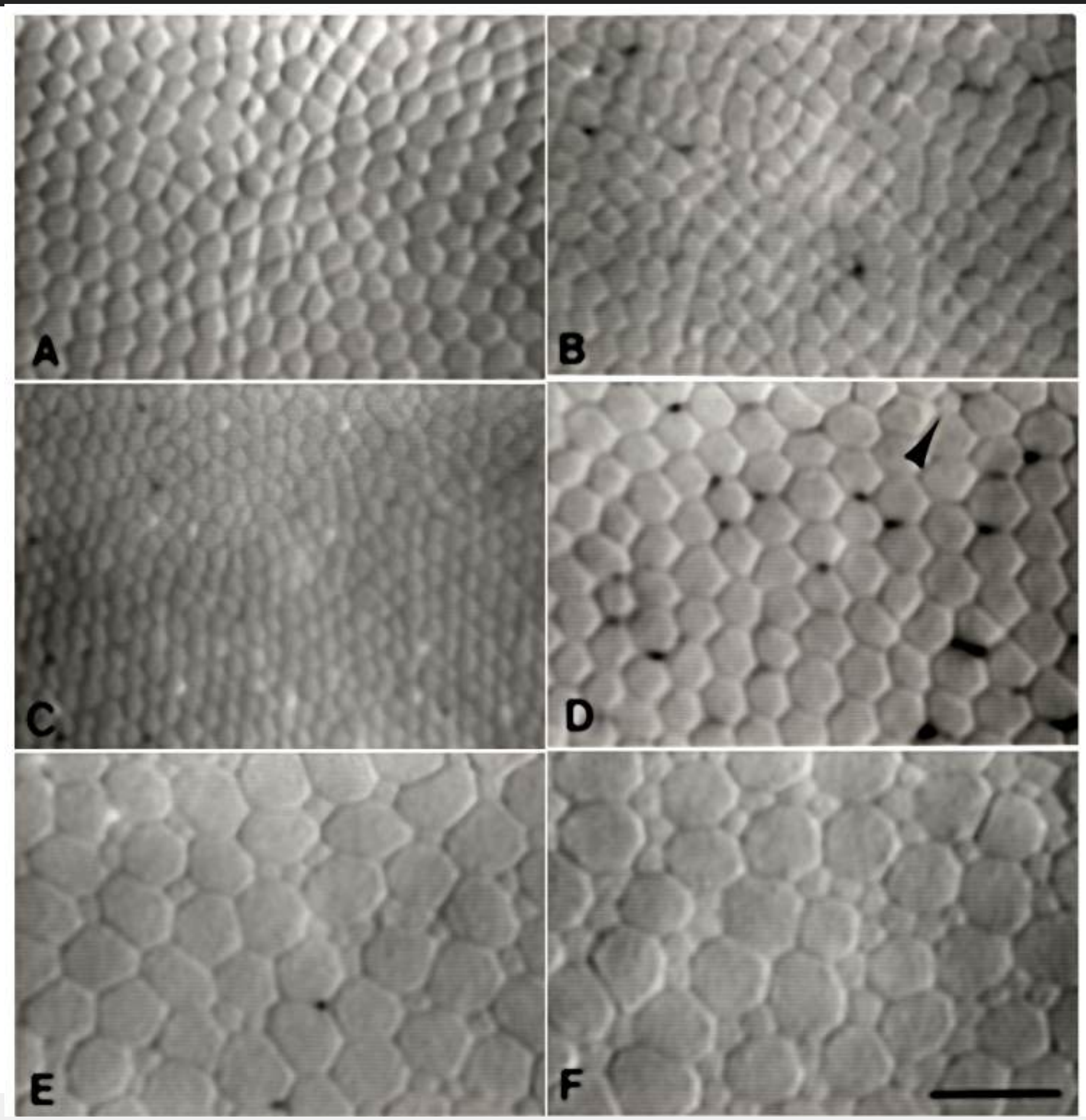
- Target *resolution* threshold: the smallest angular size at which subjects can discriminate



- Target *localization* threshold: the smallest difference in position which subjects can discriminate (Vernier hyperacuity)

http://webvision.med.utah.edu/book/part-viii-gabac-receptors/visual-acuity/

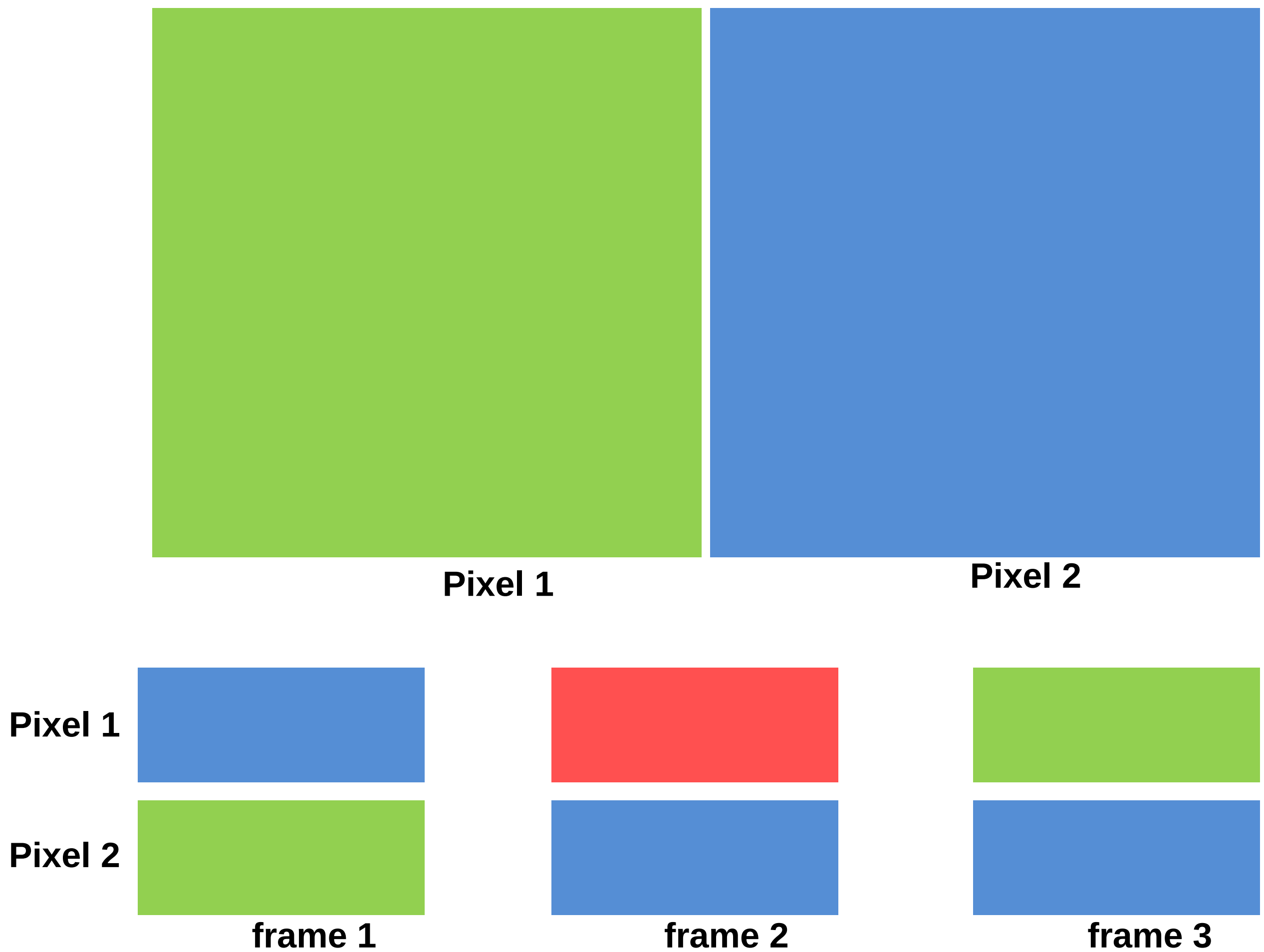
Foveal Photoreceptor Mosaic



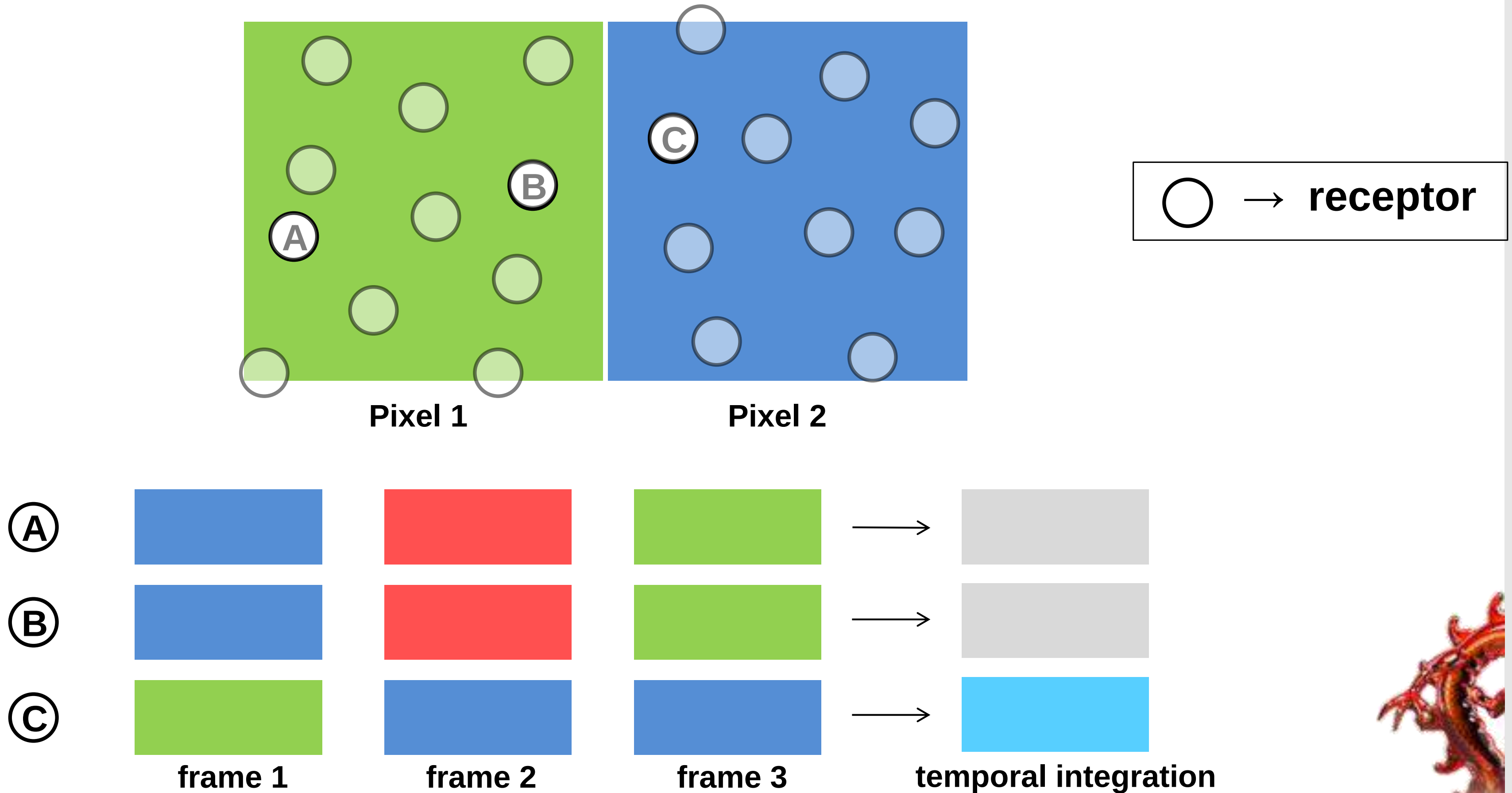
A-C fovea center - cones only
D rod-free region boundary, the arrow shows rod
E cones-rods balanced
F rods outnumber cones

Curcio, C. A., Sloan, K. R., Kalina, R. E., Hendrickson, A. E., 1990. Human photoreceptor topography. J Comp Neurol 292, 497-523

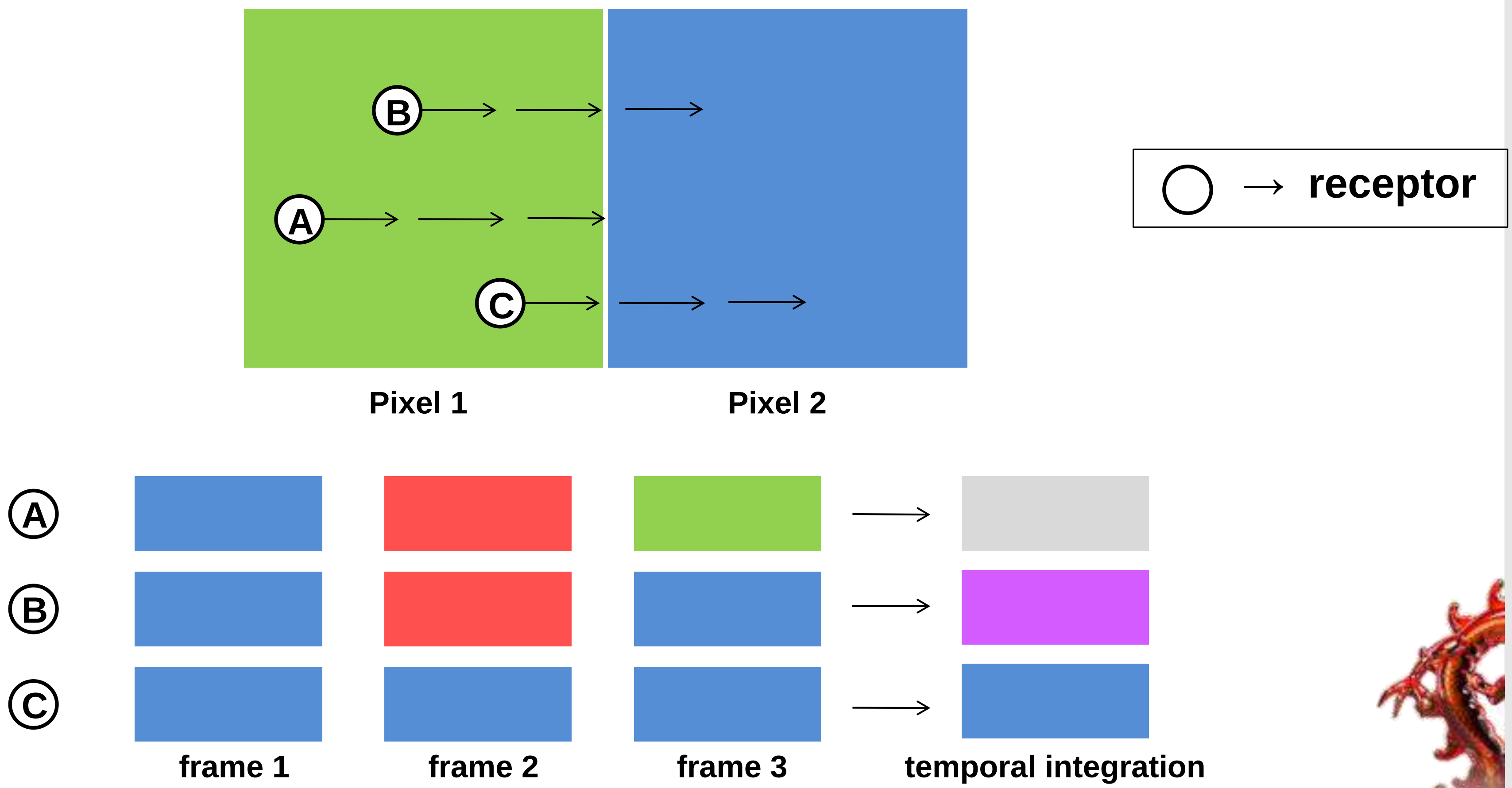
Temporal Domain



Temporal Domain – static case

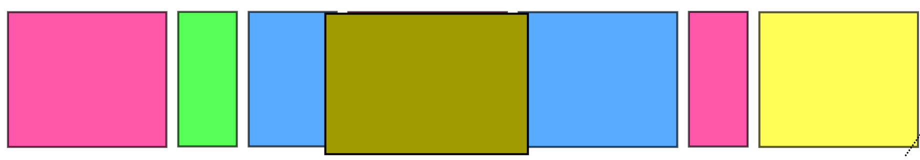


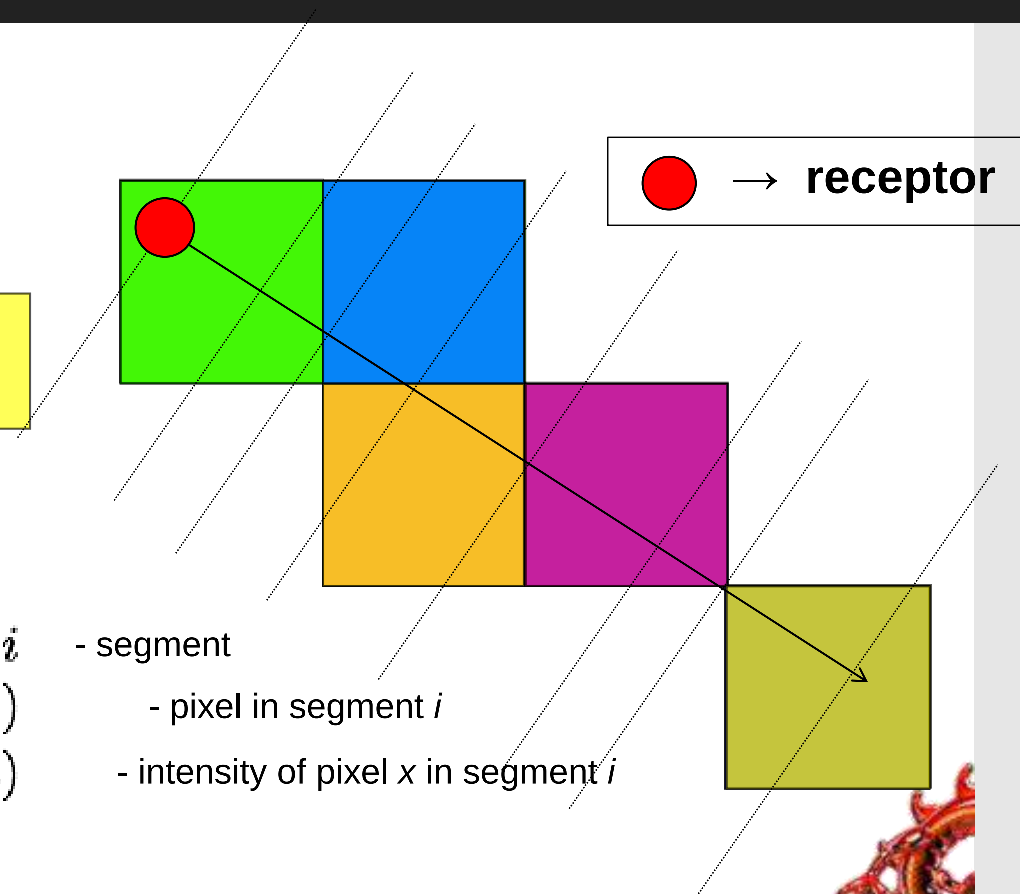
Temporal Domain – dynamic case



Temporal Integration Model

Receptor signal:





$$\sum_{i=0}^{N} w_i \, I(p(i), i)$$

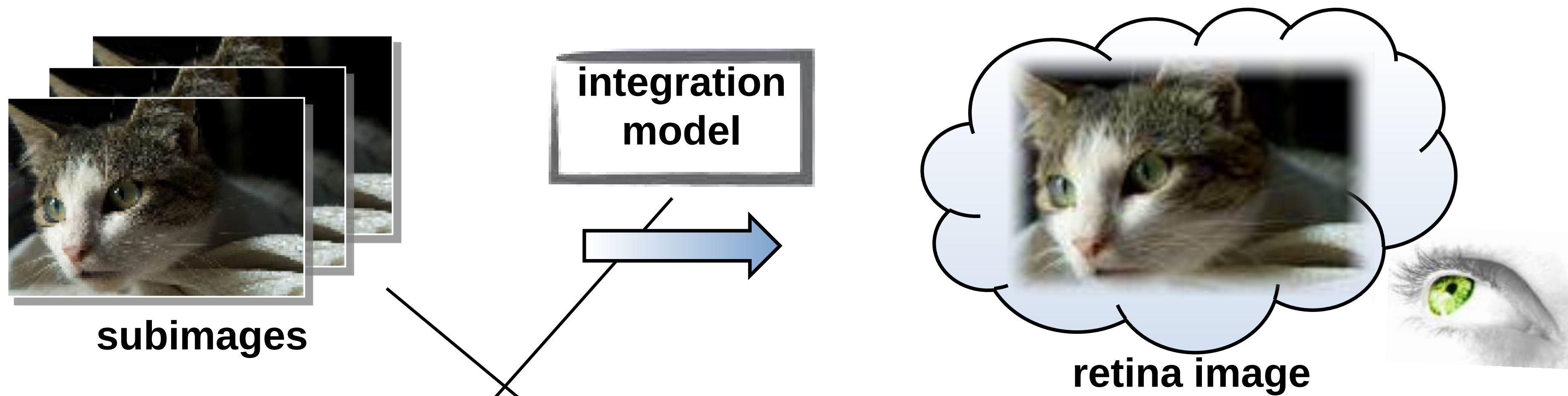
i - segment

$p(i)$ - pixel in segment *i*

$I(x, i)$ - intensity of pixel *x* in segment *i*

w_i - weights proportional to the length of the segment

Prediction in Equations



$$\begin{pmatrix} w_{1,1},\ w_{1,2},\ \ldots,\ w_{n,m} \\ W \\ \\ \end{pmatrix} \begin{pmatrix} I_1^L \\ I_2^L \\ I_3^L \\ \ldots \\ I_n^L \end{pmatrix} = \begin{pmatrix} i_{x,y} \\ I^H \\ \\ \end{pmatrix}$$

prediction for one receptor

$$\boxed{\sum_{i=0}^{N} w_i\, I(p(i), i)}$$

Optimization Problem

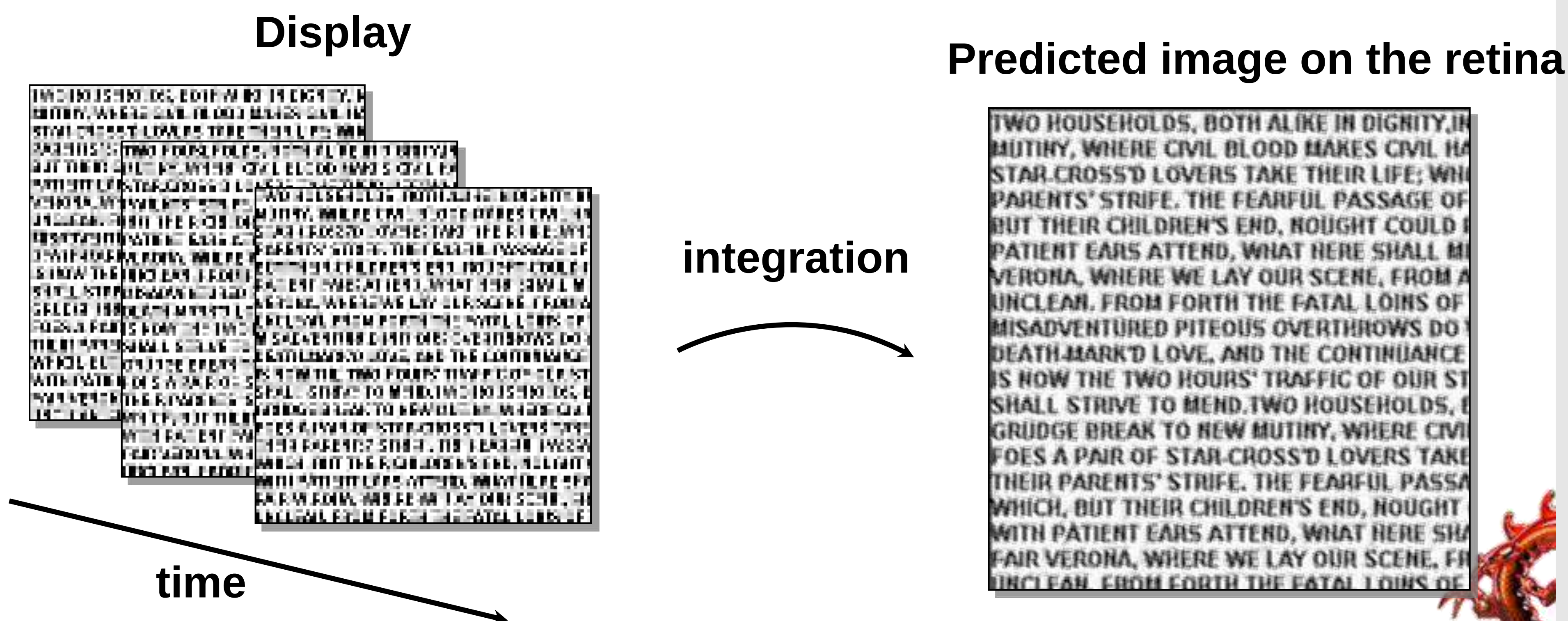
$$\begin{pmatrix} & & \\ & W & \\ & & \end{pmatrix} \begin{pmatrix} I_1^L \\ I_2^L \\ I_3^L \\ \dots \\ I_n^L \end{pmatrix} = \begin{pmatrix} I^H \end{pmatrix}$$

integration model

subimages

high resolution image

Optimization Result



ARE vs. Lanczos

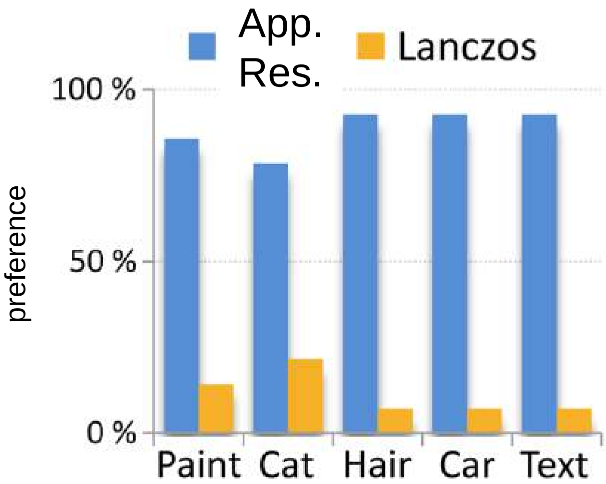
- compare each frame

  to moving image

  - downsample separately

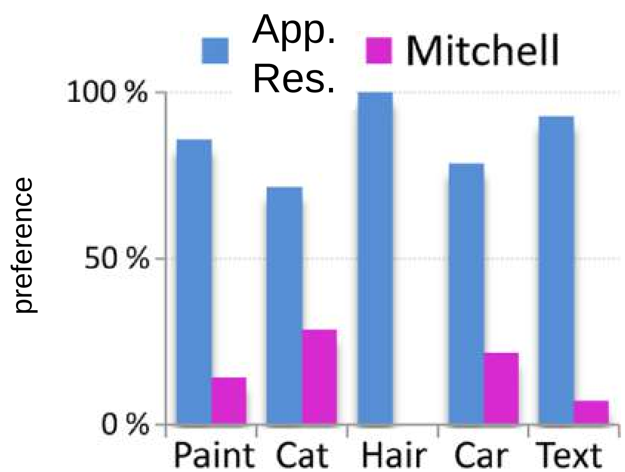
    hence, slightly different

    information over time



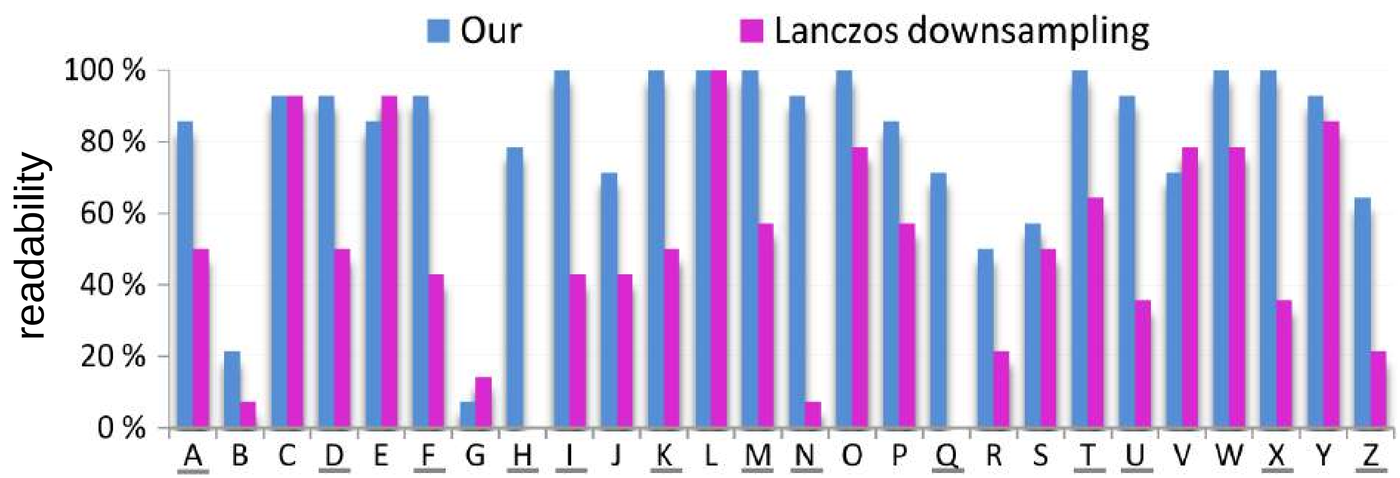
ARE vs. Mitchell

- Mitchell downsampling
  - participants adjusted parameters to match high resolution image



ARE - Alphabet

Size: 2 x 3 pixels

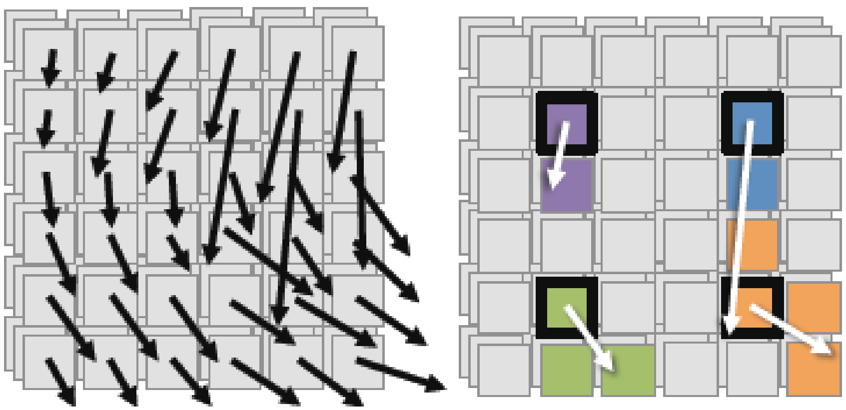


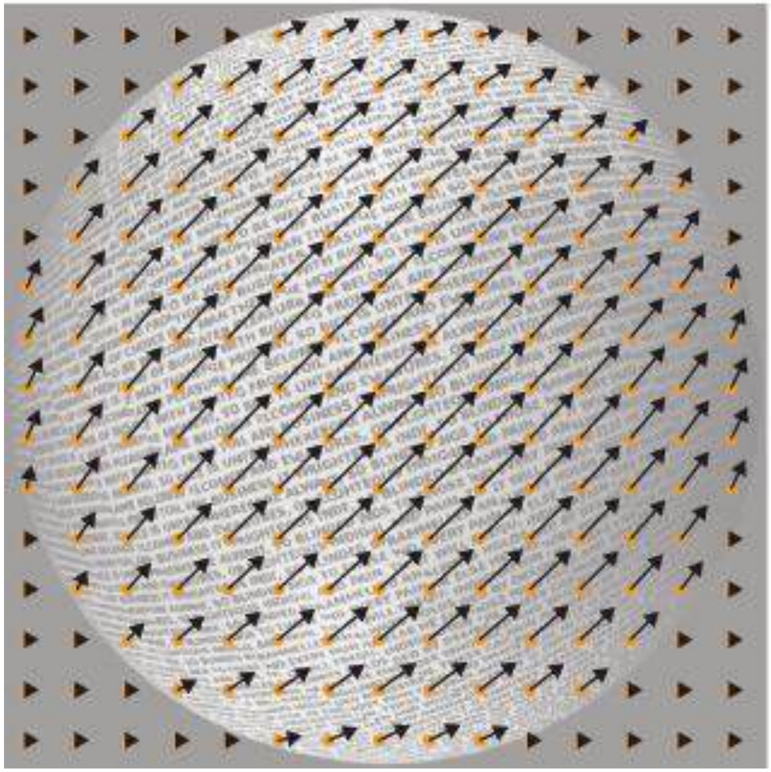
- Applications:
  - scrolling text or maps on low resolution devices
  - stock tickers, news headlines

Recently: Extension to movies

Apparent Resolution Enhancement for Animations

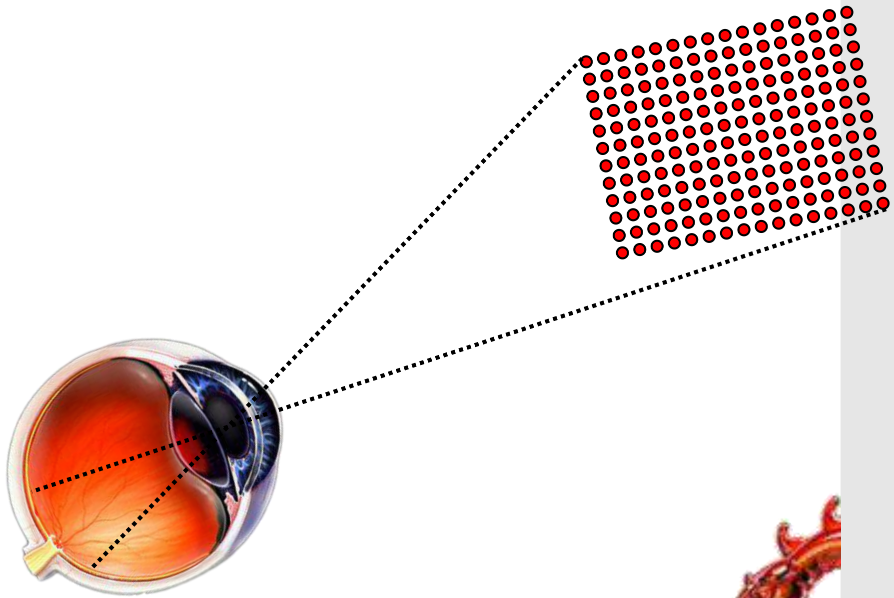
[Templin et al. SCCG 2011]





Conclusions

- Human perception is a crucial component to high-quality imagery
- Resolution & Colors
  physical screen capabilities
- Works for large range
  of commonly used display devices



Future?

- Bigger,

  better,

  faster…
  - More realism
  - More details
  - More effects

- Higher quality beyond physical limitations
  - Only first steps in this direction
  - More to come…

Thank you very much for your attention!

Karol Myszkowski
karol@mpi-inf.mpg.de

Elmar Eisemann
eisemann@telecom-paristech.fr

Acknowledgment:
Daniel Scherzer, Robert Herzog and Dawid Pajak

Image / Video Quality Assessment

Tunç O. Aydın

Disney Research, Zurich

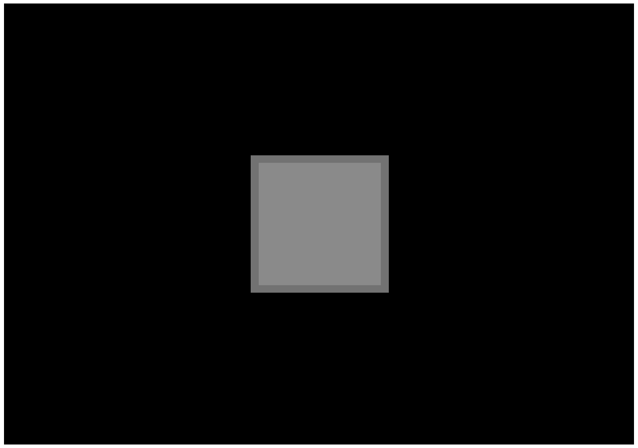
<tunc@disneyresearch.com>

Problem Definition

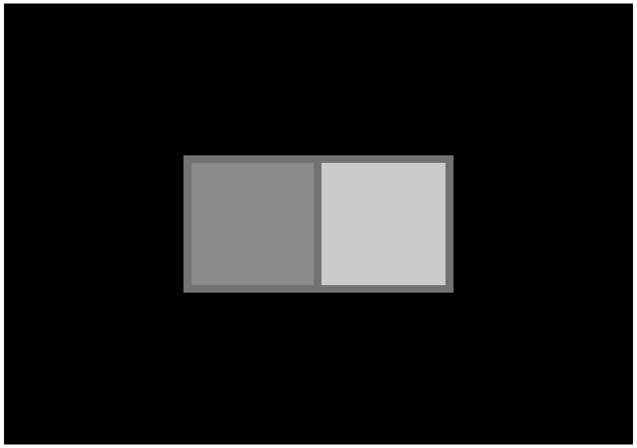
Rate
the
Quality

Subjective Quality Assessment

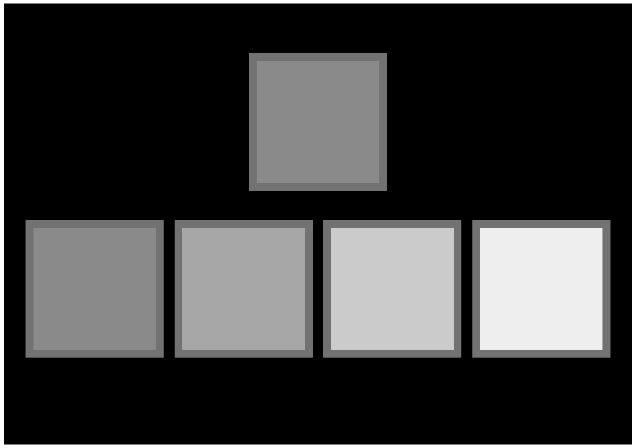
Figures taken from [Ferwerda 2008]



Detection



Discrimination



Scaling

Refer to: [James Ferwerda, Psychophysics 101: How to Run Perception Experiments in Computer Graphics, SIGGRAPH 2008].

+ Reliable **- High cost**

Objective Quality Assessment

No Reference

Reduced Reference

Full Reference

Refer to: [Wang & Bovik, Modern Image Quality Assessment, 2008].

Generic Quality Assessment Workflow

Reference

Test

Distortion Map

Simple Distortion Metrics

- **Mean Squared Error (MSE)** $$MSE(x,y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

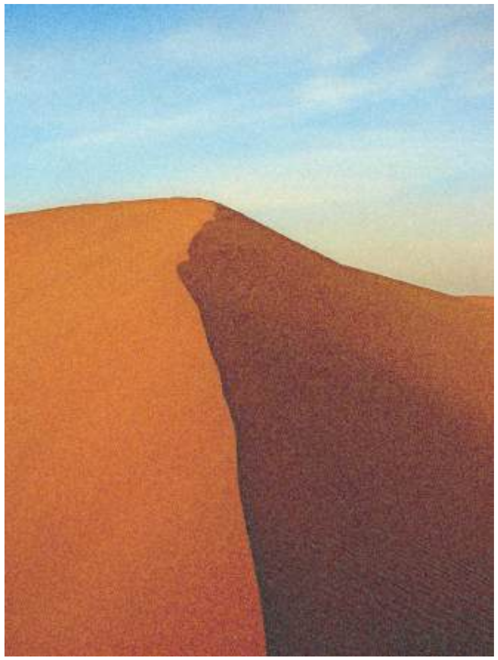
- **Peak Signal to Noise Ratio (PSNR)** $$PSNR(x,y) = 10\log_{10}\frac{L^2}{MSE}$$

- **Structural Similarity Index Metric (SSIM):** More sophisticated, accounts for luminance contrast and structural distortions

$$SSIM(x,y) = l(\mu_x, \mu_y)^{\alpha} c(\sigma_x, \sigma_y)^{\beta} s(\sigma_x, \sigma_y)^{\gamma}$$

Limitations of Simple Distortion Metrics

Reference



Random Noise

Blur

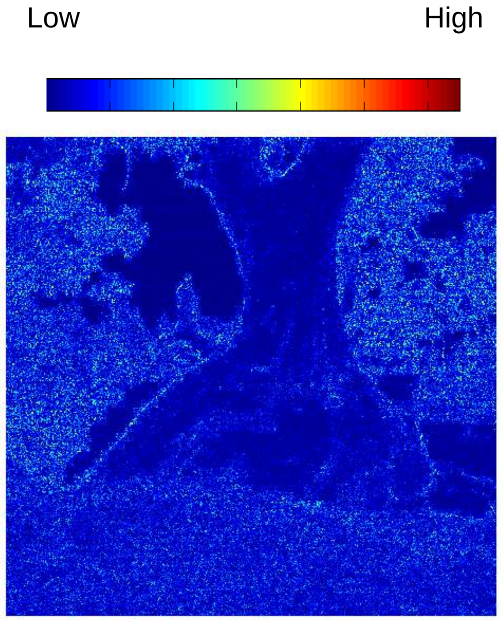
~15% Decreased Luminance

Same MSE for all three images!

Perception of Distortions

Reference
(bmp, 616K)

Compressed
(jpg, 48K)

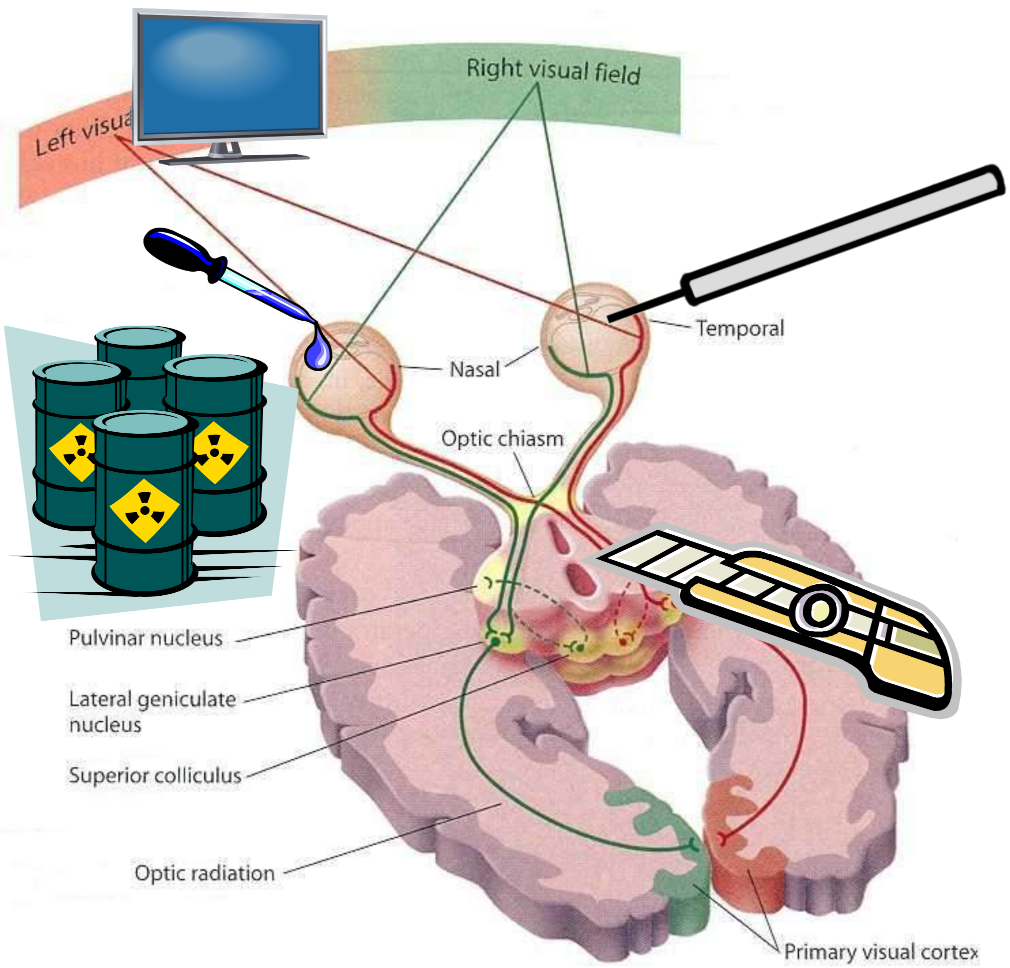


Difference Image
(Color coded)

Limitations of Simple Distortion Metrics, cont.

Visible difference doesn't always mean lower quality!

The Human Visual System (HVS)



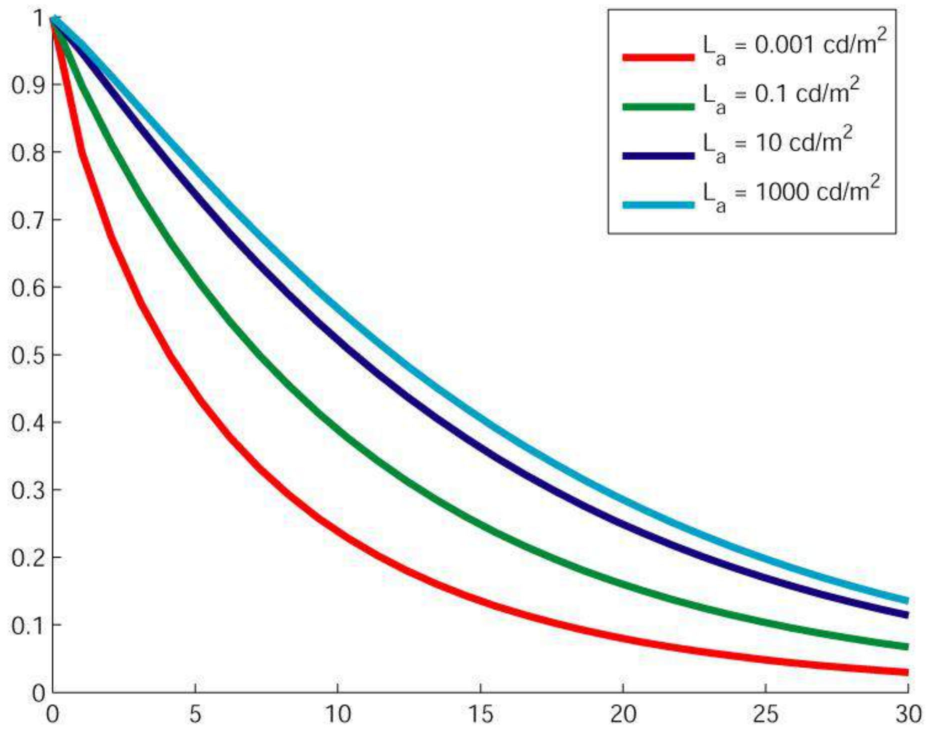
- Experimental Methods of Vision Science
  - Micro-electrode
  - Radioactive Marker
  - Vivisection
  - Psychophysical Experimentation

HVS effects: (1) **Glare**

Video Courtesy of Tobias Ritschel

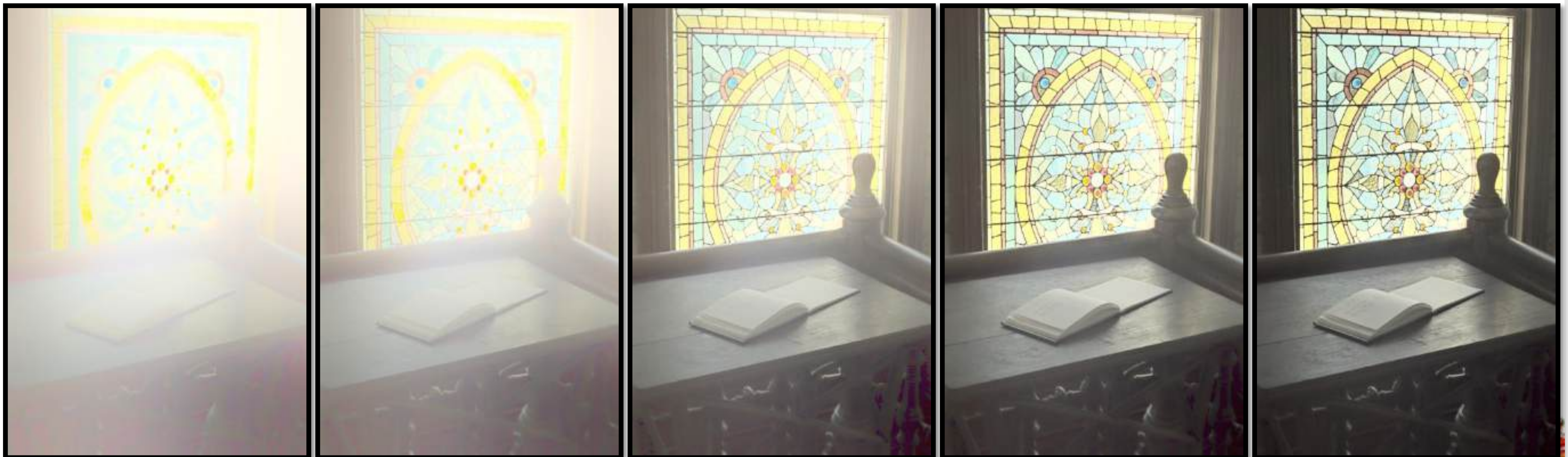
- **Disability Glare** (blooming)

Disability Glare



- Model of Light Scattering
  - Point Spread Function in spatial domain
  - Optical Transfer Function in Fourier Domain [Deeley et al. 1991]

(2) Light Adaptation

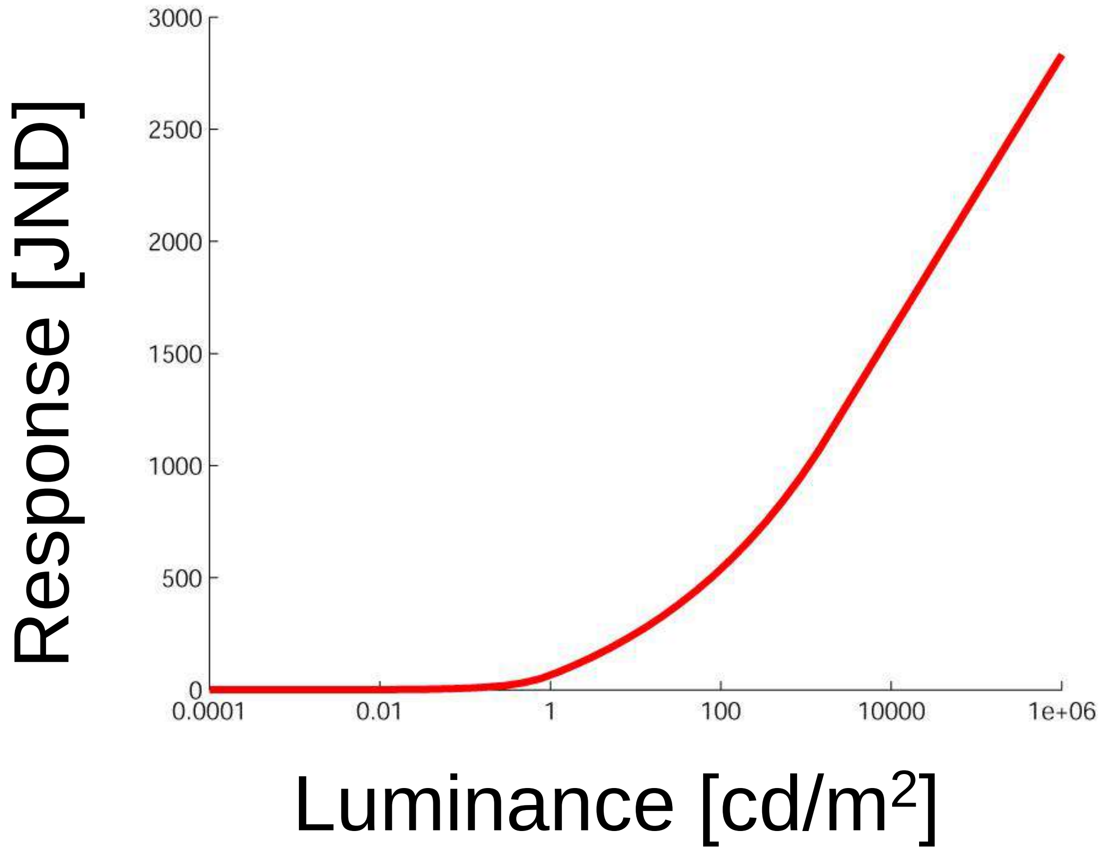


Adaptation Level:
10^{-4} cd/m^2

Time $\longrightarrow$

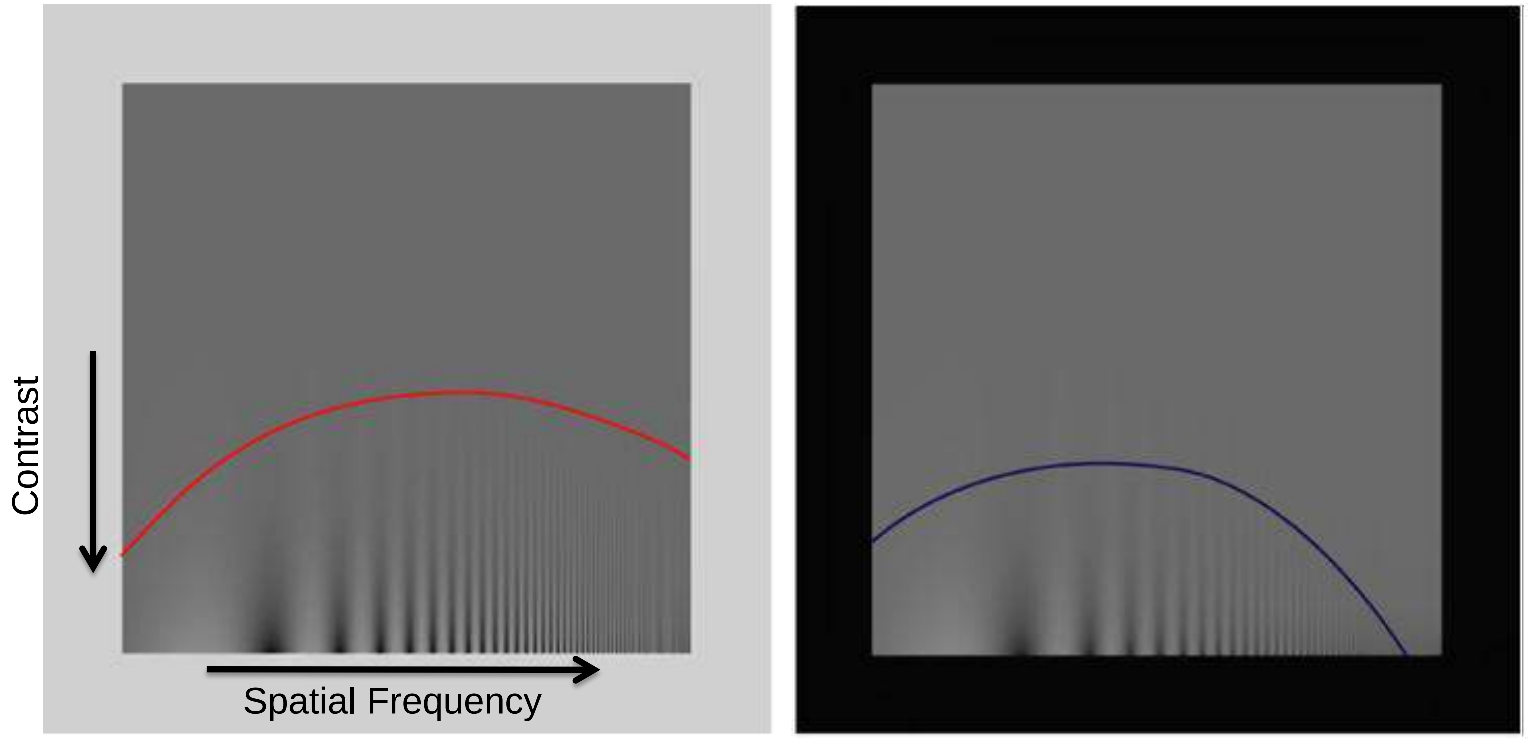
Adaptation Level:
17 cd/m^2

Perceptually Uniform Space



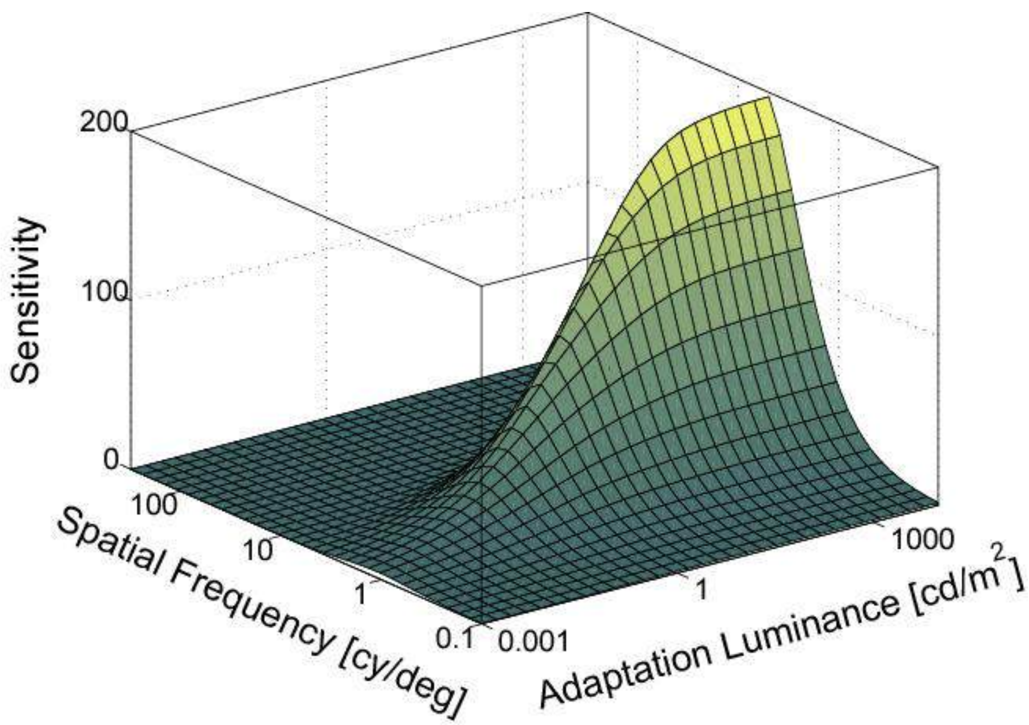
- Transfer function: Maps **Luminance** to **Just Noticeable Differences (JNDs)** in Luminance. [Mantiuk et al. 2004, Aydın et al. 2008]

(3) Contrast Sensitivity



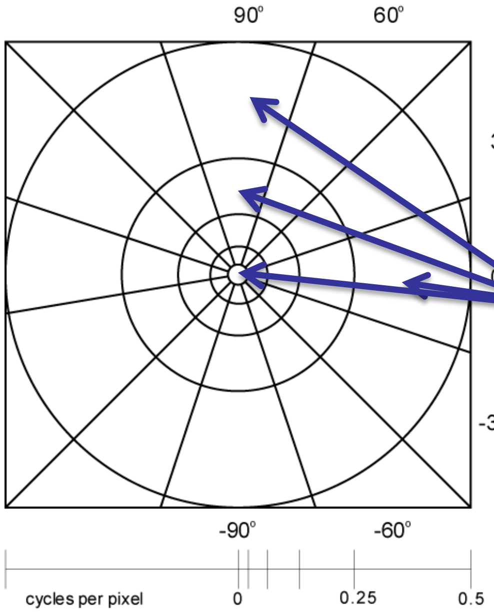
CSF(spatial frequency, adaptation level, temporal freq., viewing dist, …)

Contrast Sensitivity Function (CSF)



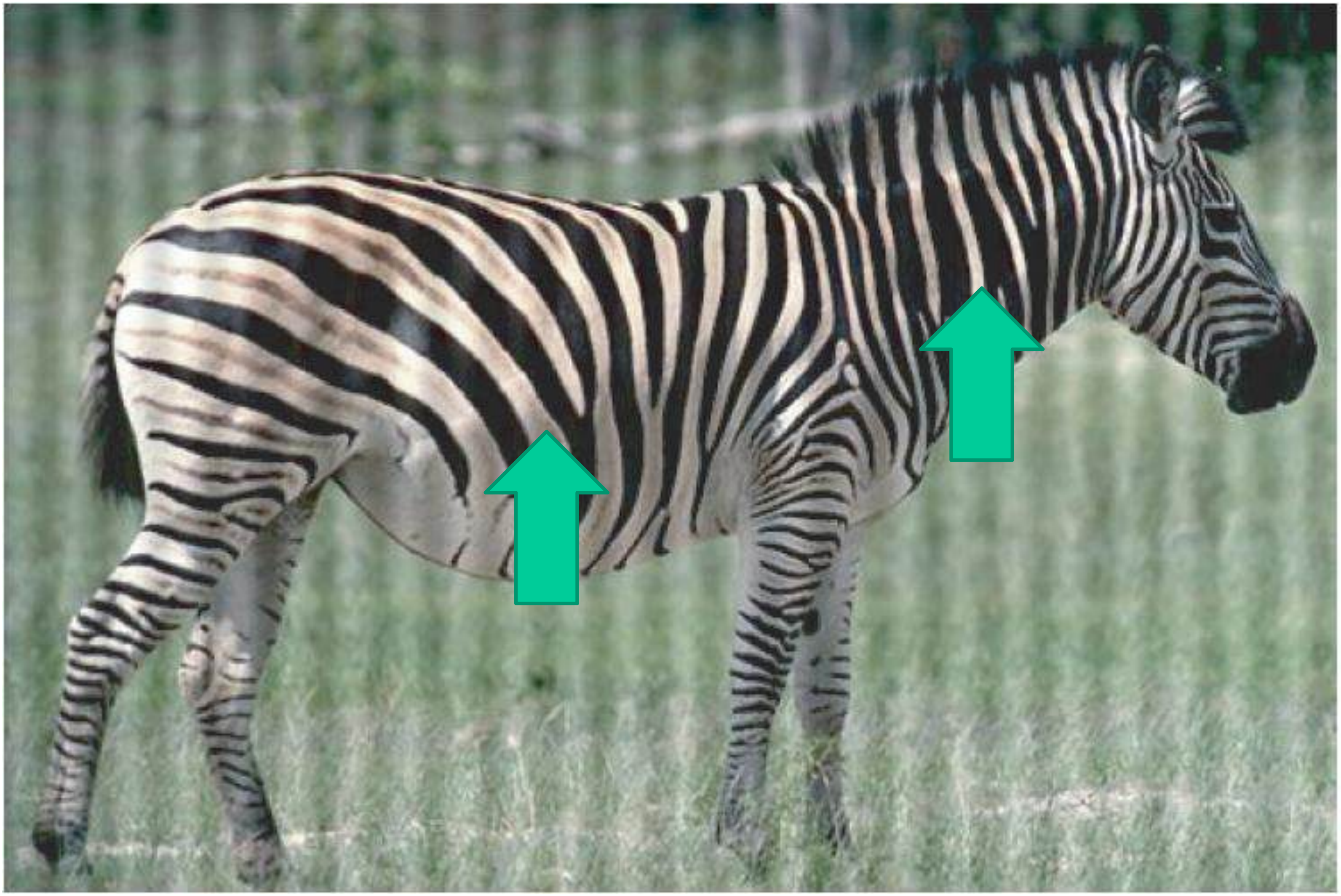
- Steady-state CSF^S: Returns the Sensitivity (1/Threshold contrast), given the adaptation luminance and spatial frequency [Daly 1993, Mantiuk et al. 2011].

(4) Visual Channels



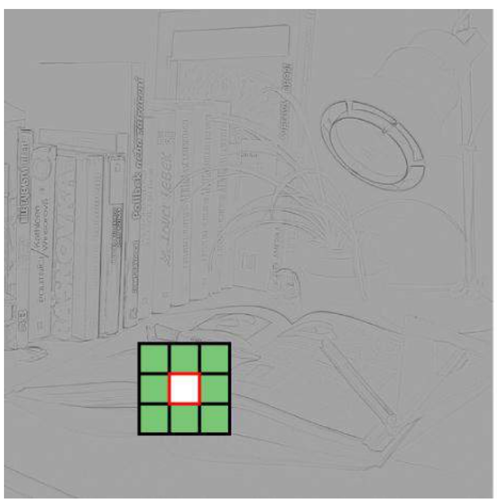
Cortex Transform

(5) Visual Masking

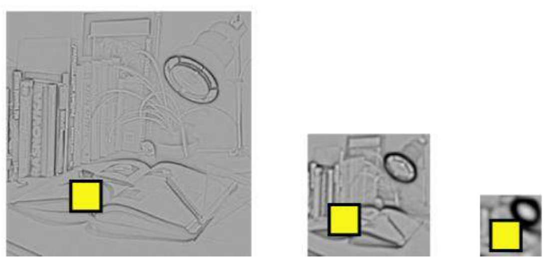


Loss of sensitivity to a signal in the presence of a "similar frequency" signal "nearby".

Visual Masking Models



- Masked coefficient
- Intra-channel neighborhood
- Inter-channel neighborhood

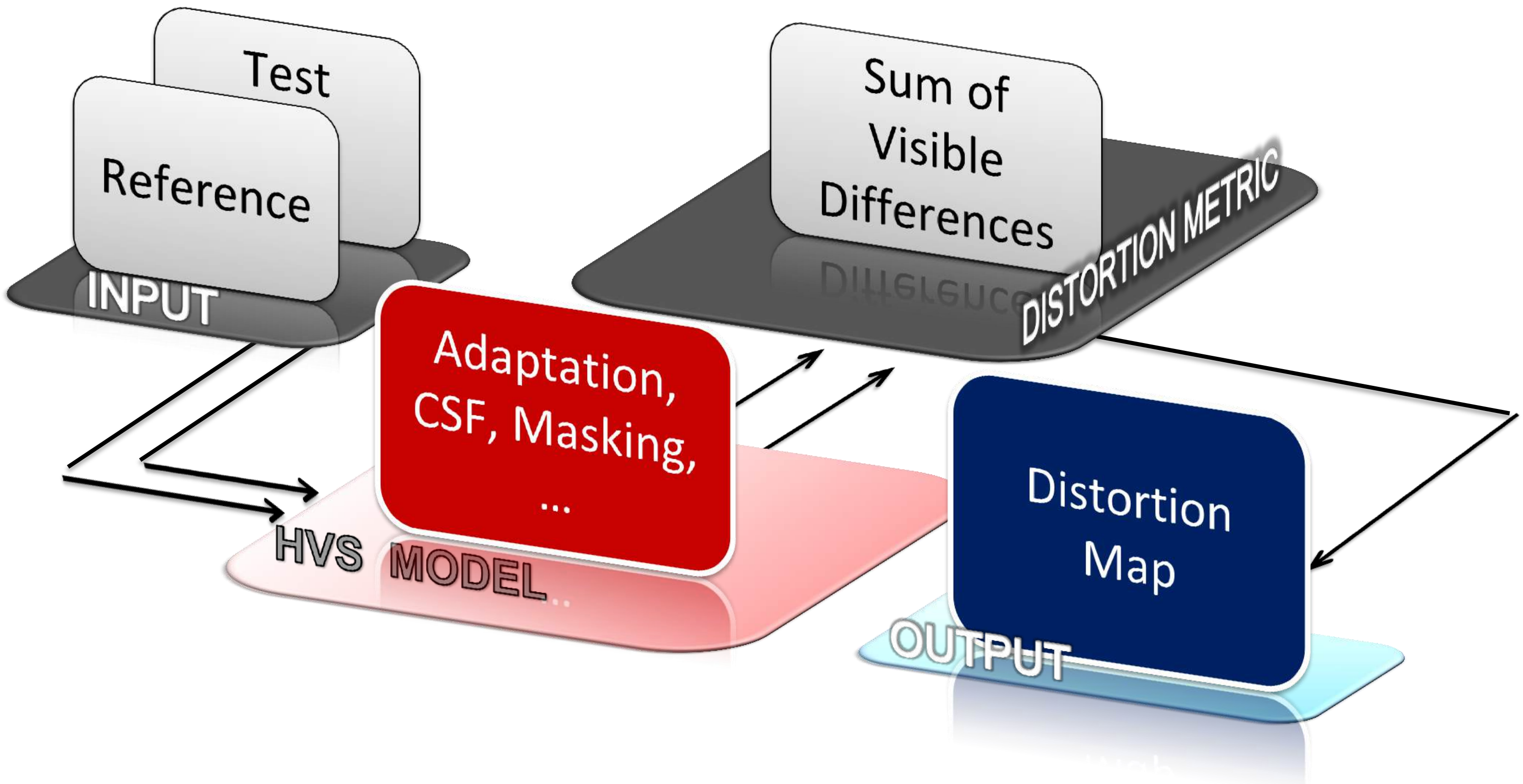


- **Example:** JPEG's pointwise extended masking:

$$R = \frac{sign(C')|C'|^{0.5}}{(1 + \sum_K |C'_k|^{0.2})}$$

C': Normalized Contrast

Generic HVS-based Quality Assessment Workflow

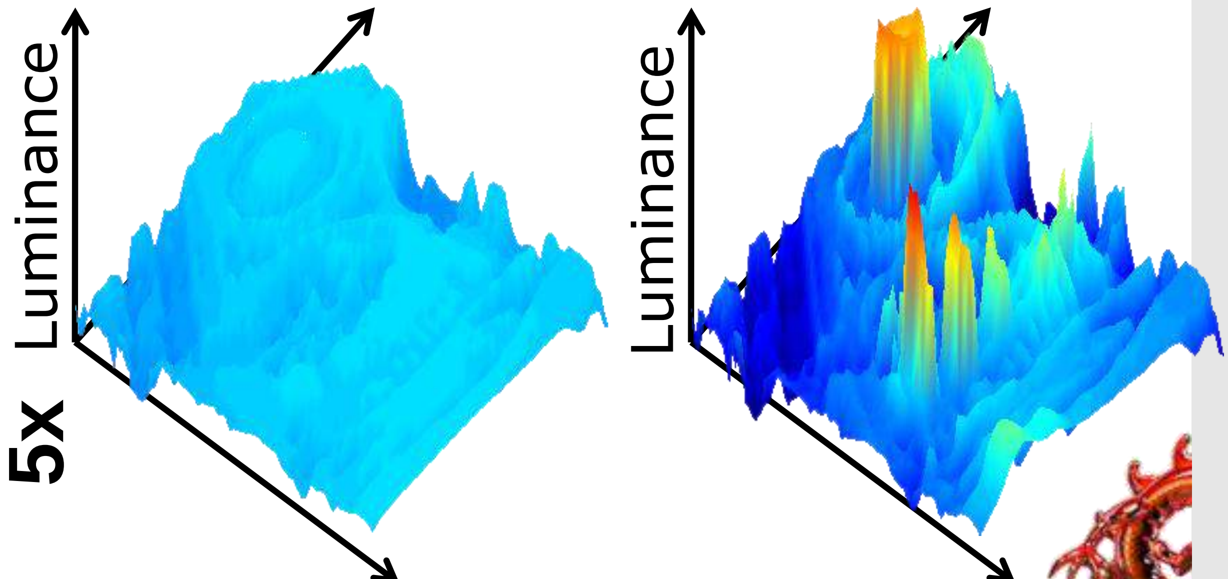


Visible Differences Predictor (VDP) [Daly 93, Mantiuk et al. 05, Mantiuk et al. 11],

Visual Discrimation Model (VDM) [Lubin 95]

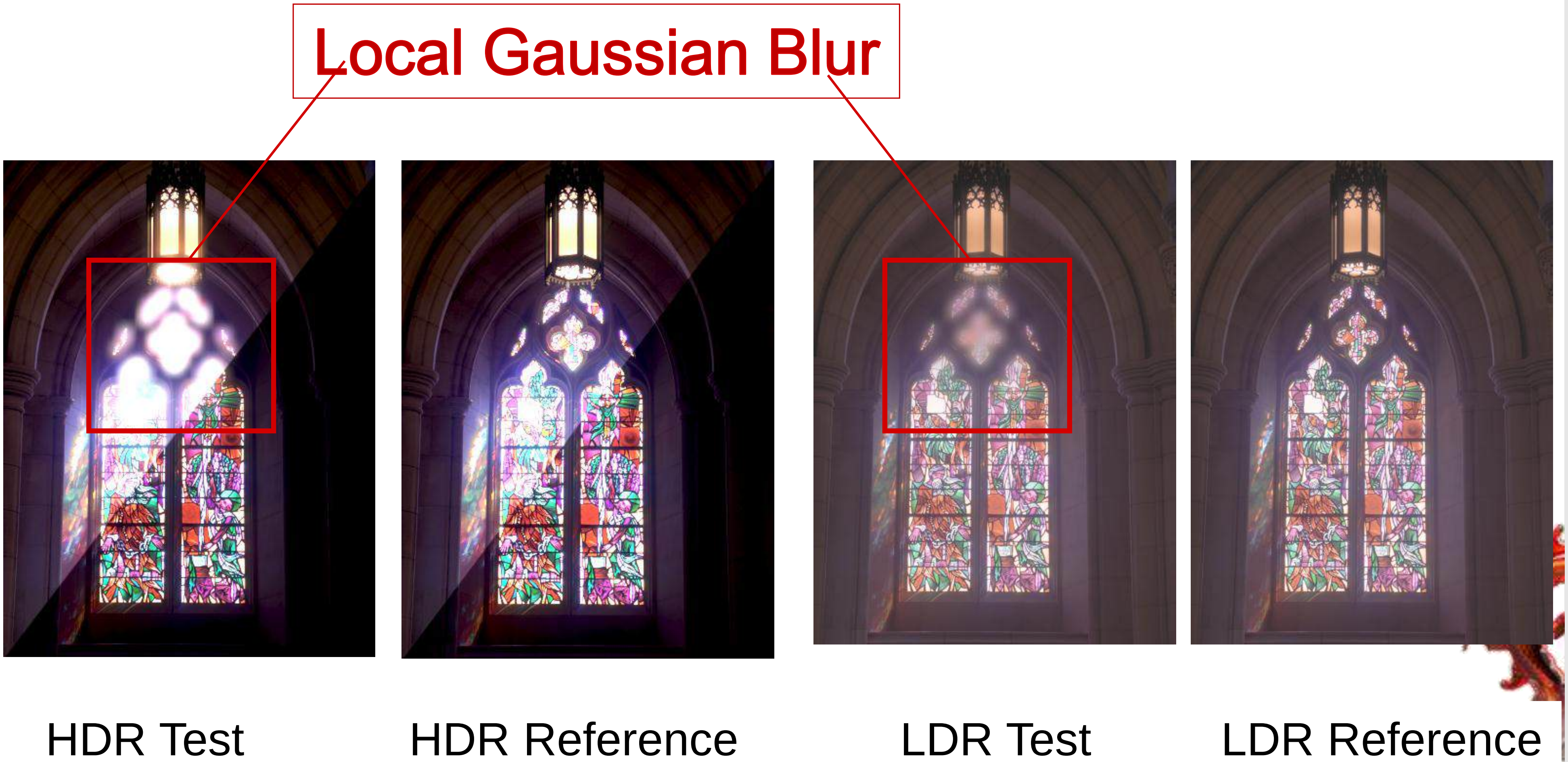
QA of Retargeted Images? **HDR Tone mapping** case

LDR HDR



LDR HDR

Case Study

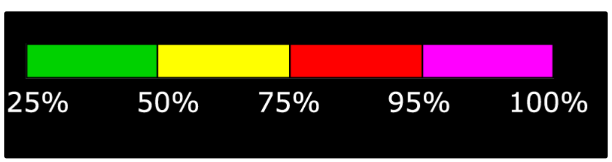

Local Gaussian Blur
HDR Test
HDR Reference
LDR Test
LDR Reference

(1) HDR pair

HDR-VDP

SSIM

DRI-IQM



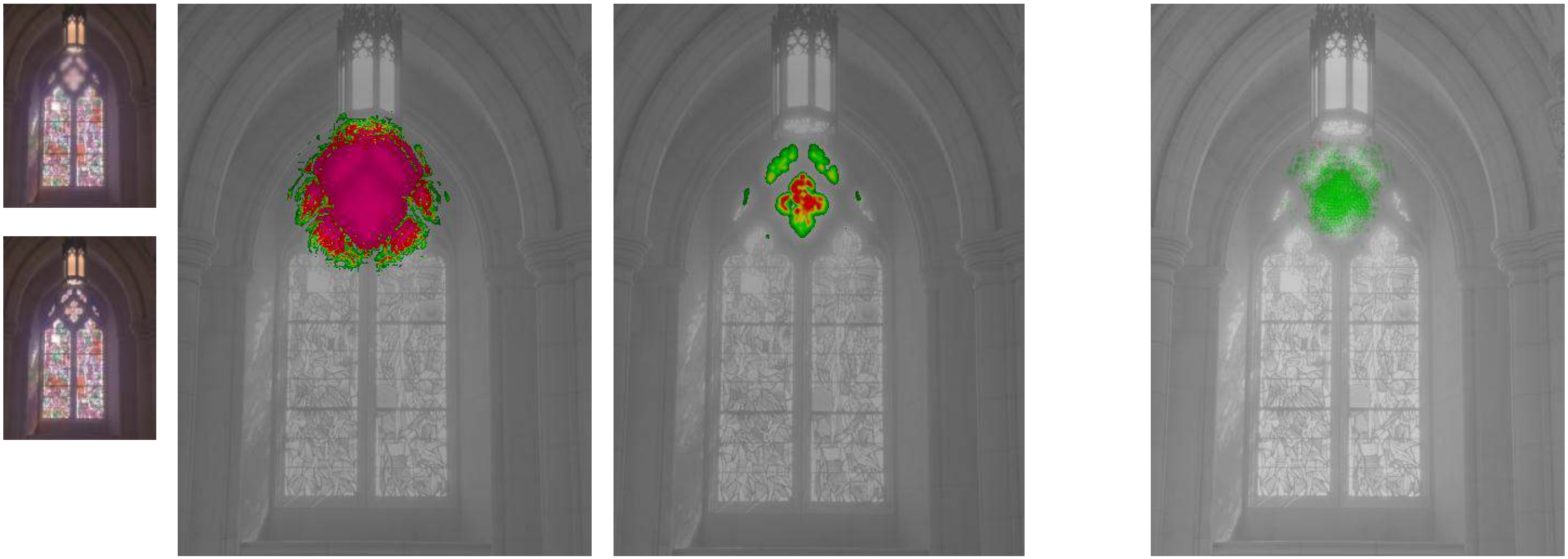
Detection Probability

Loss

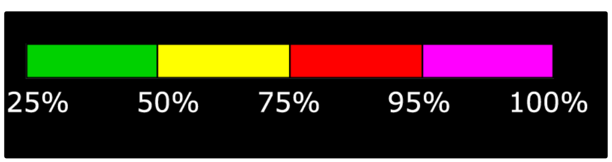
Amplification

Reversal

(2) LDR pair



HDR-VDP SSIM DRI-IQM

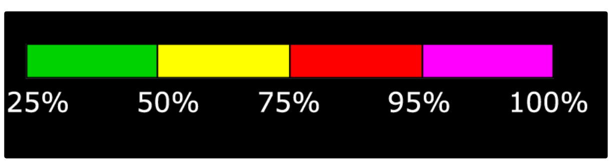


Detection Probability

- Loss
- Amplification
- Reversal

(3) HDR test, LDR reference

HDR-VDP SSIM DRI-IQM



Detection Probability

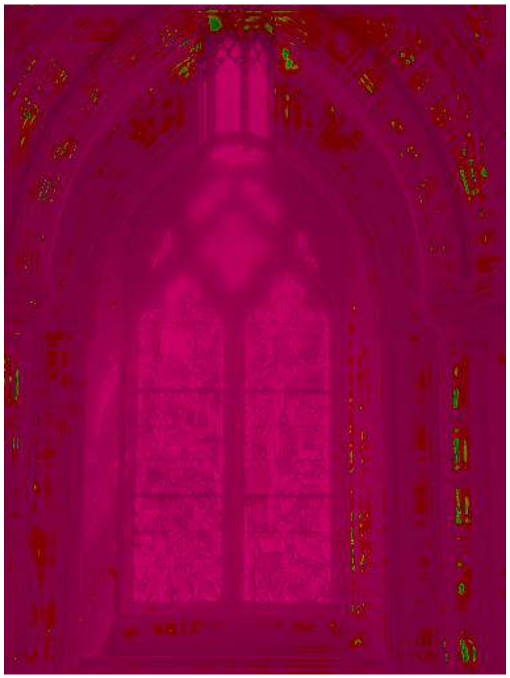
Loss

Amplification

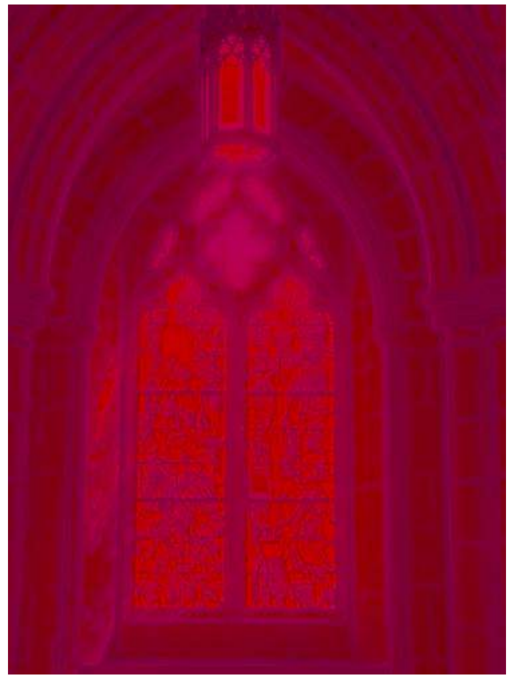
Reversal

SIGGRAPH ASIA 2011 HONG KONG

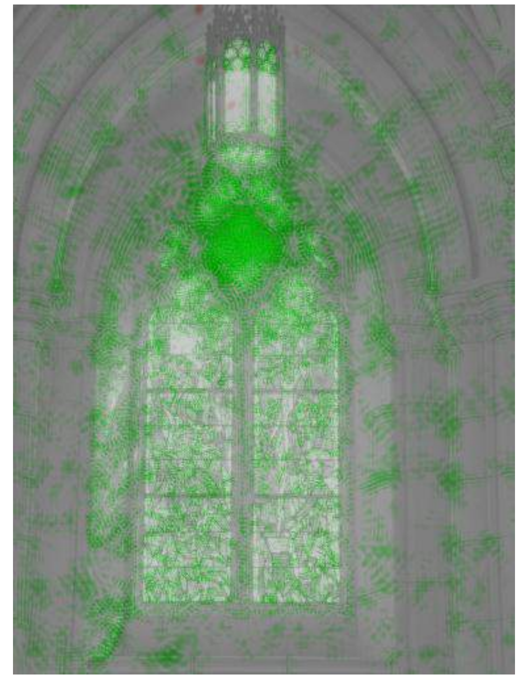
(4) LDR test, HDR reference



HDR-VDP



SSIM



DRI-IQM

25% 50% 75% 95% 100%

Detection Probability

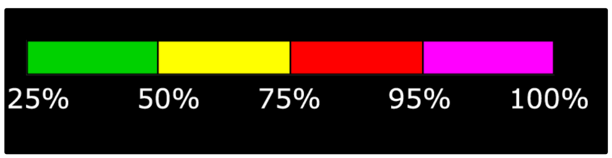
Loss

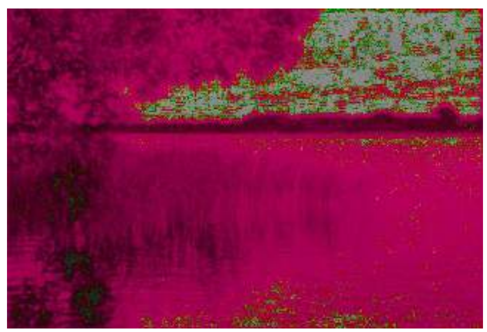
Amplification

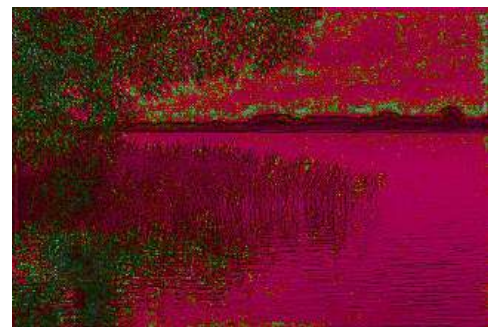
Reversal

Detecting distortions

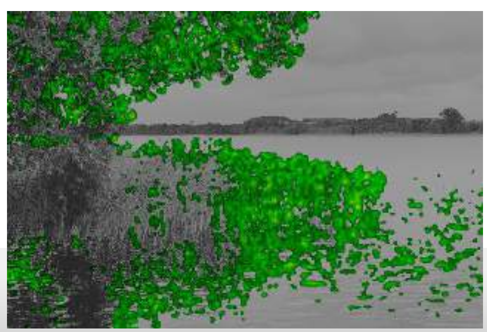
Reference	Sharpening	Blur







HDR-VDP

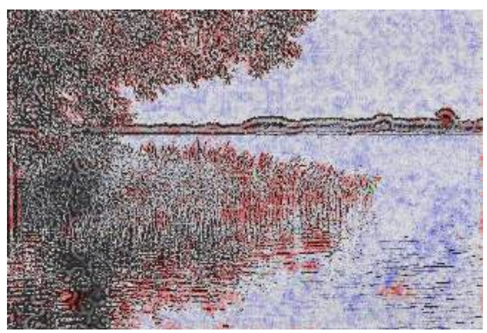


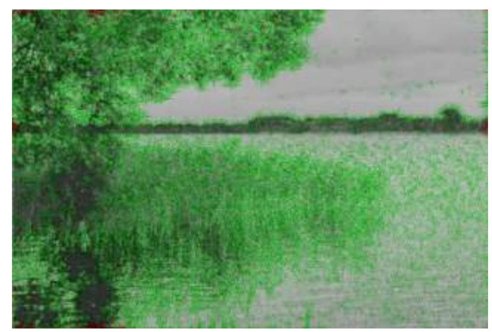
SSIM

Detecting “types” of distortions

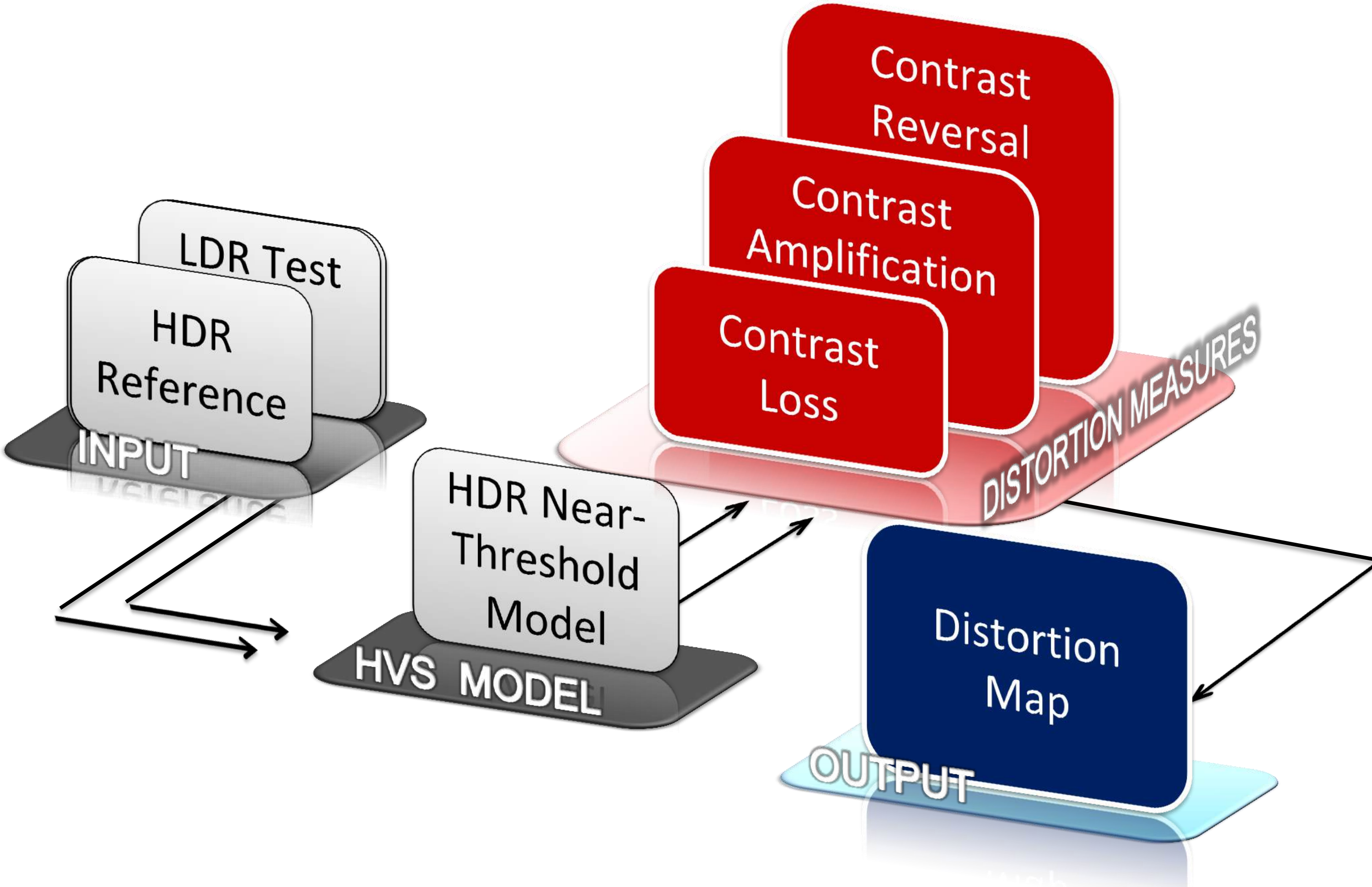
Reference	Sharpening	Blur

- Loss
- Amplification
- Reversal



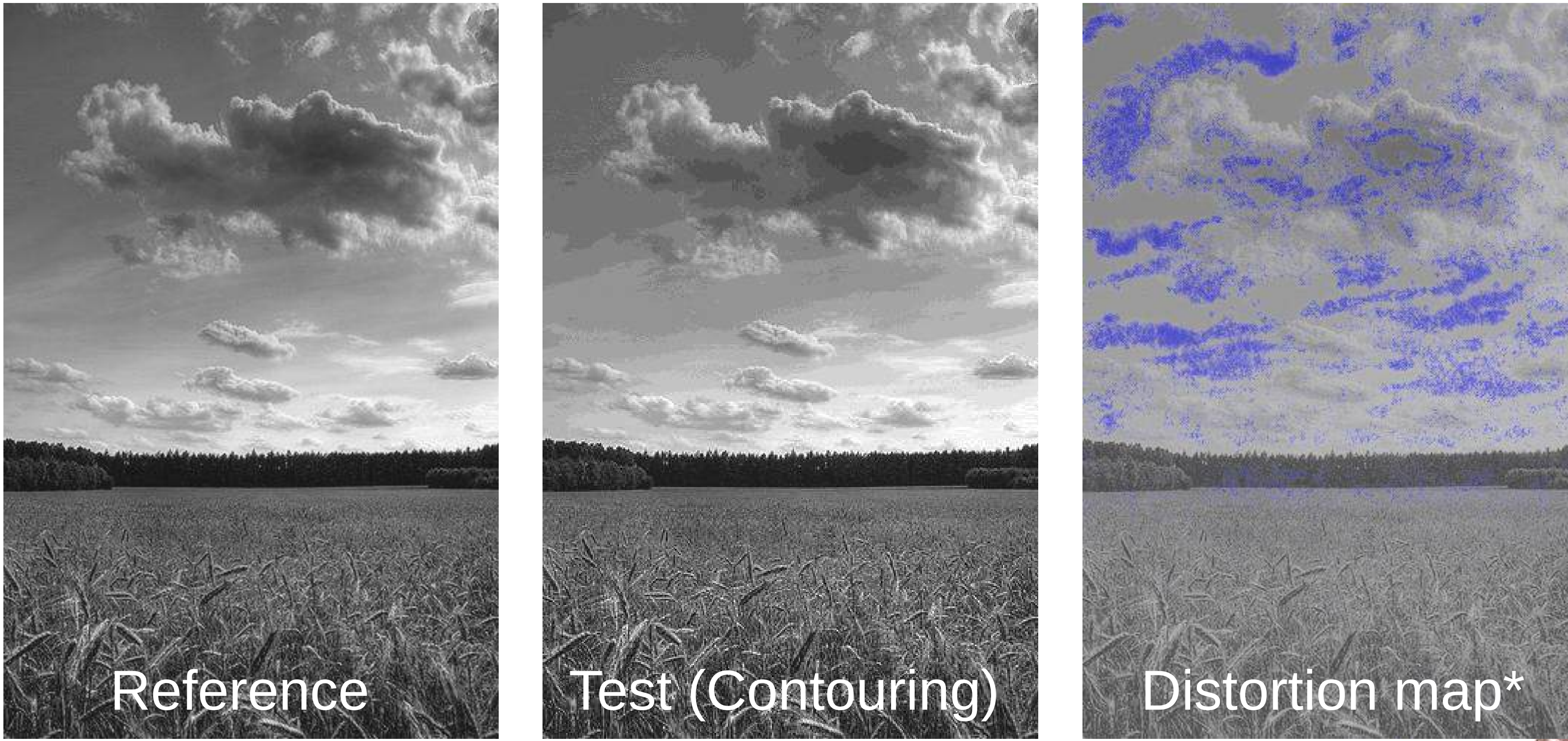


Generic DRI Image Quality Assessment Workflow

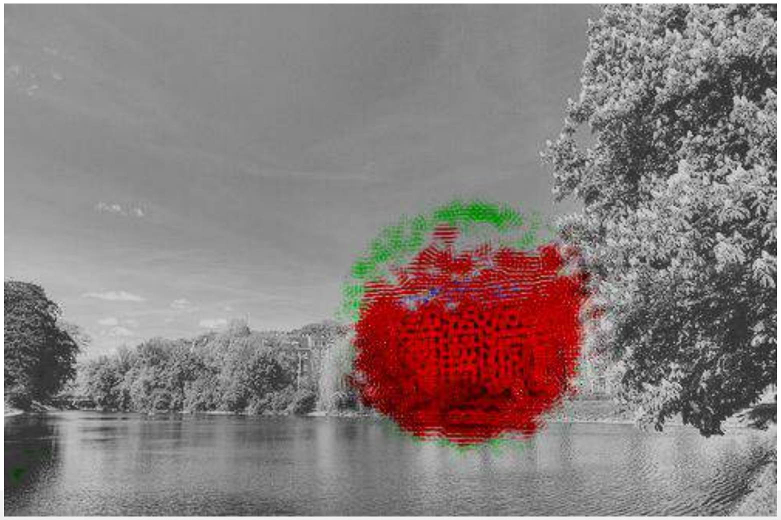


Loss of Visible Contrast

Amplification of Invisible Contrast



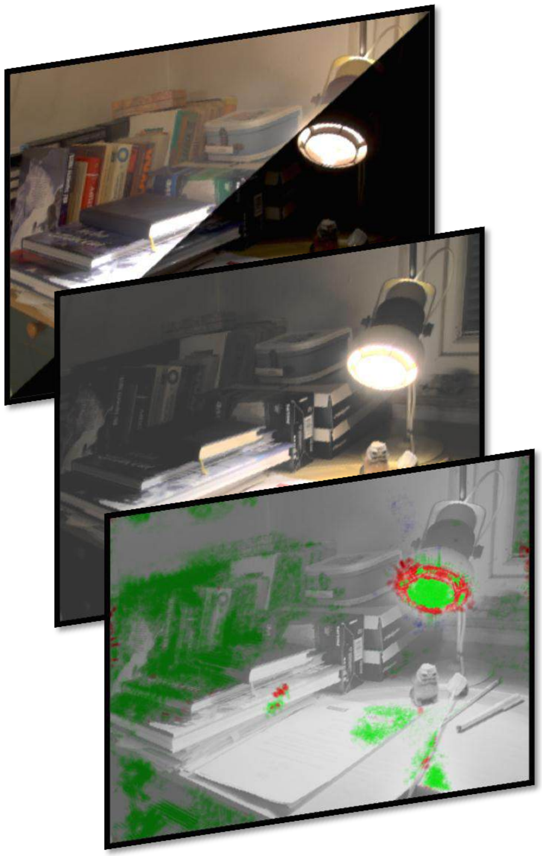
Reversal of Visible Contrast



HDR Tone Mapping Evaluation

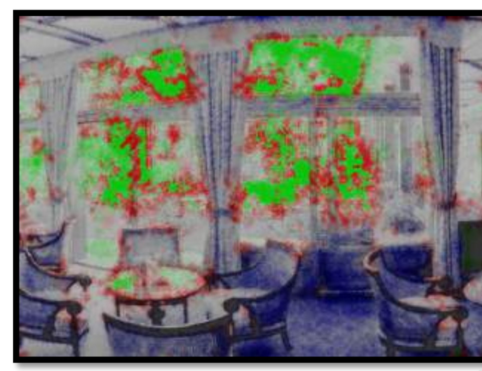
SIGGRAPHASIA2011 HONG KONG

Tone Mapping



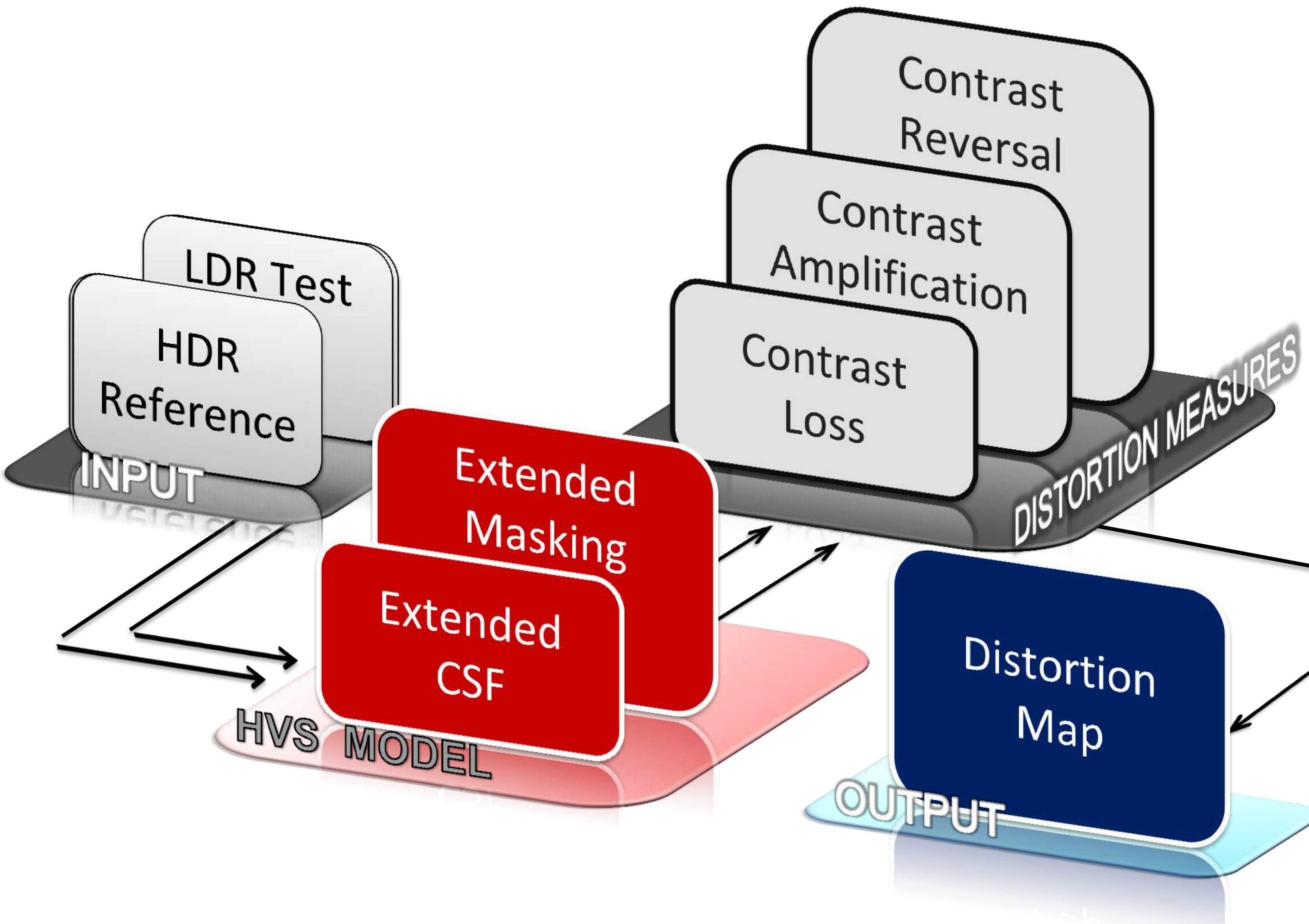
Inverse
Tone Mapping

Display
Analysis



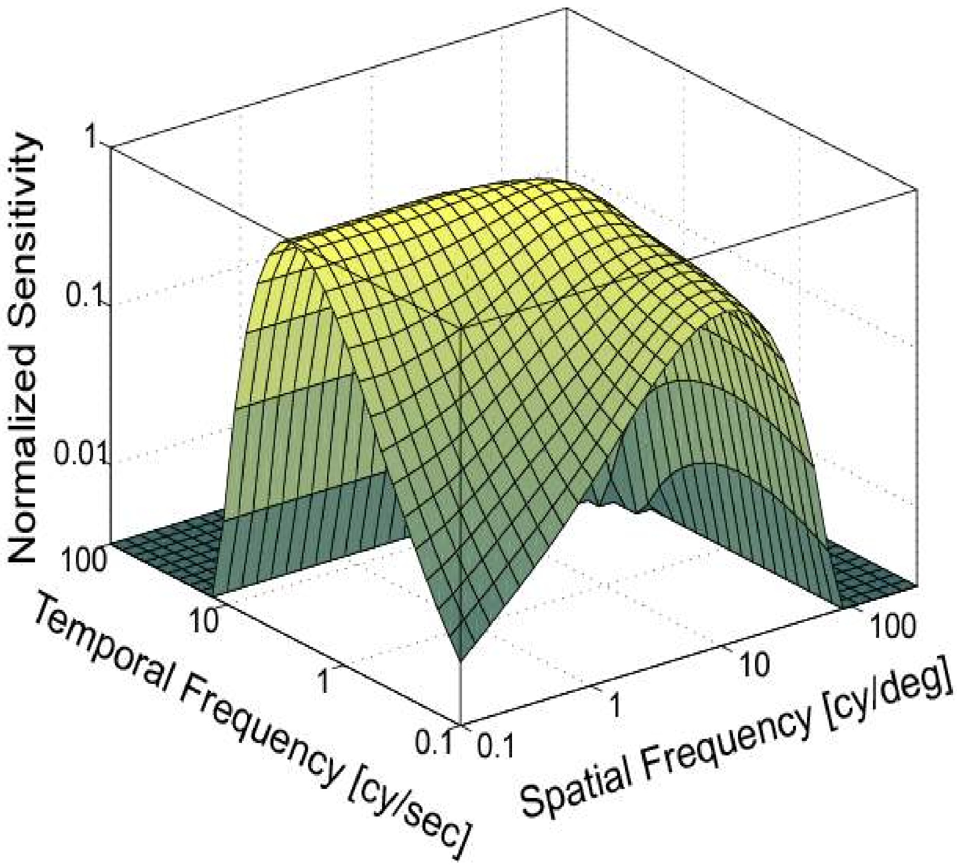
Generic DRI Video Quality Assessment Workflow

SIGGRAPHASIA2011 HONG KONG



Extended Contrast Sensitivity Function

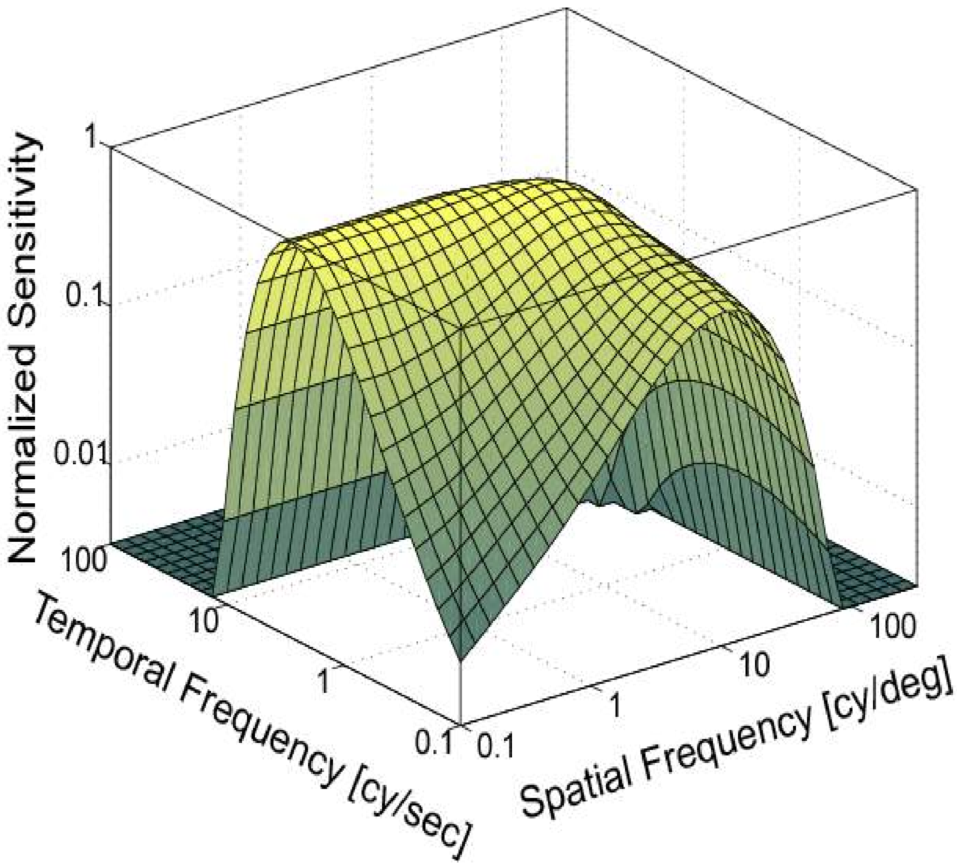
- *CSF:* $\omega,\rho,L_a \rightarrow S$
  - ω: temporal frequency,
  - ρ: spatial frequency,
  - L_a: adaptation level,
  - S: sensitivity.



Spatio-temporal CSF

Extended Contrast Sensitivity Function, cont.

- *CSF:* $\omega, \rho, L_a \rightarrow S$
  - *ω:* temporal frequency,
  - *ρ:* spatial frequency,
  - L_a*:* adaptation level,
  - *S:* sensitivity.

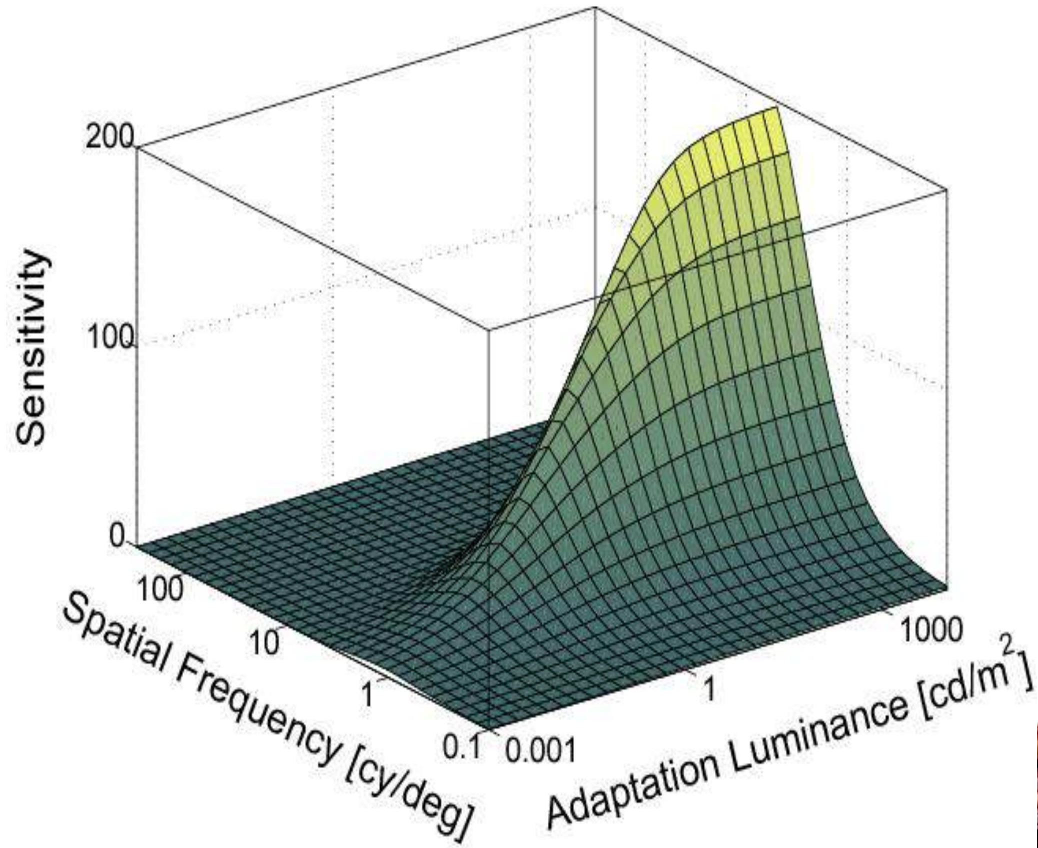


Spatio-temporal CSF[T]

Extended Contrast Sensitivity Function, cont.

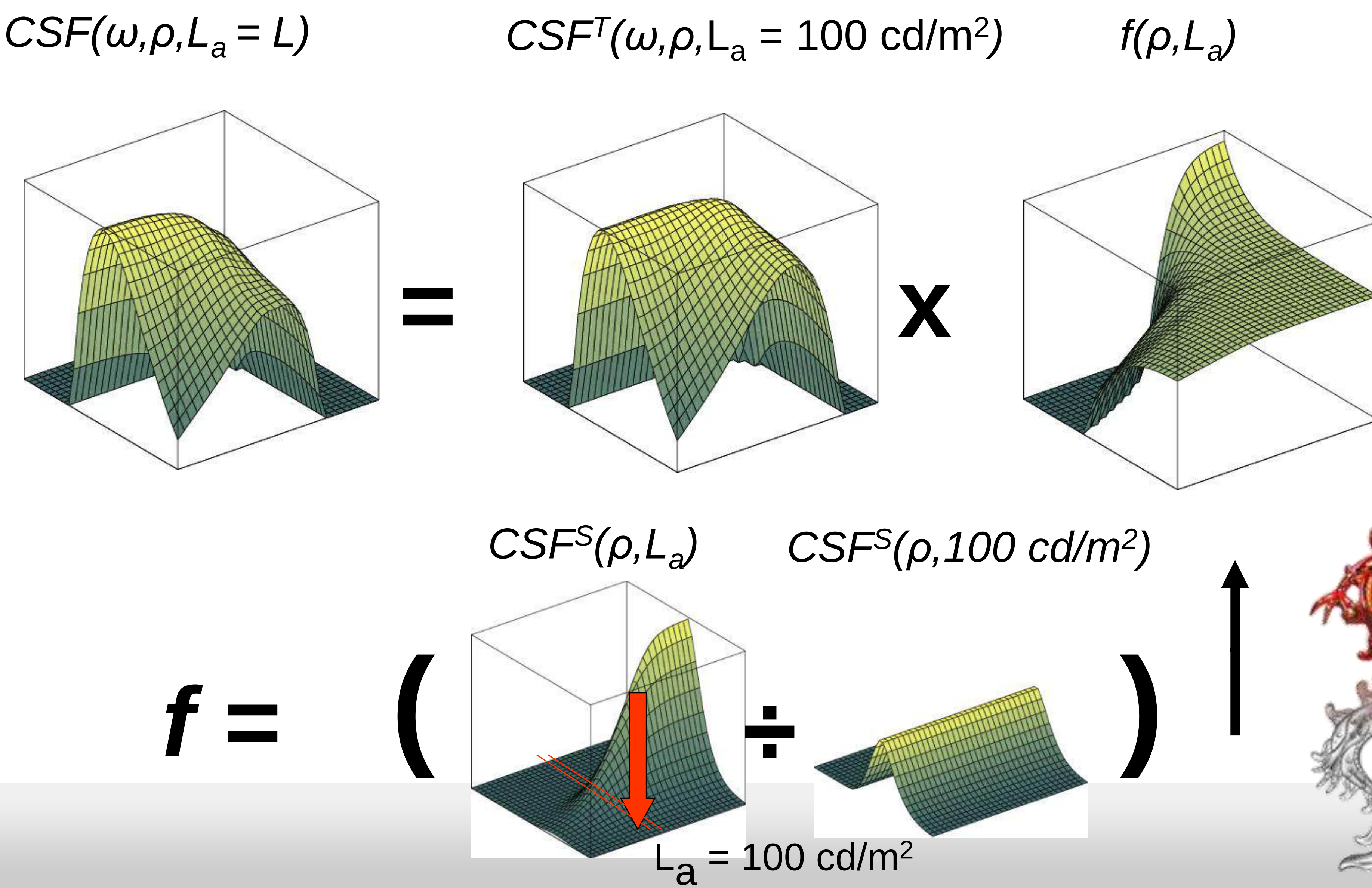
SIGGRAPHASIA2011 HONG KONG

- *CSF:* $\omega, \rho, L_a \rightarrow S$
  - *ω:* temporal frequency,
  - *ρ:* spatial frequency,
  - *L_a:* adaptation level,
  - *S:* sensitivity.

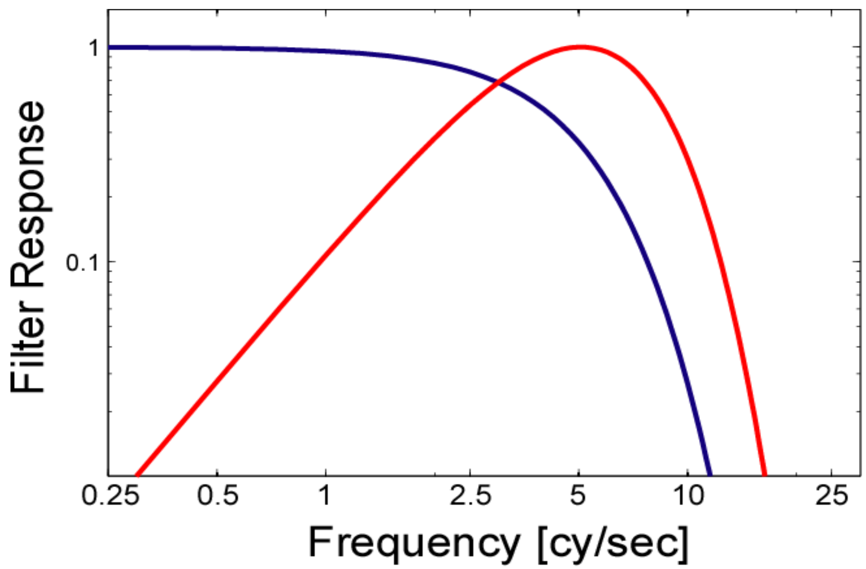


Steady-state CSF^S

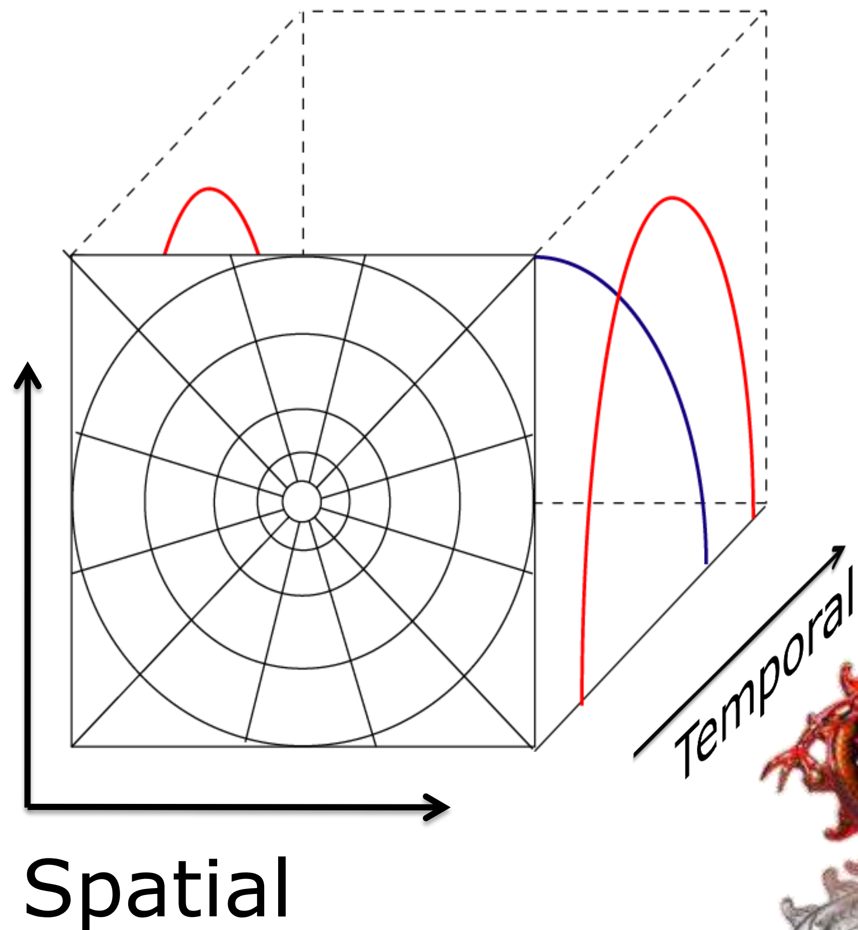
Extended Contrast Sensitivity Function, derivation



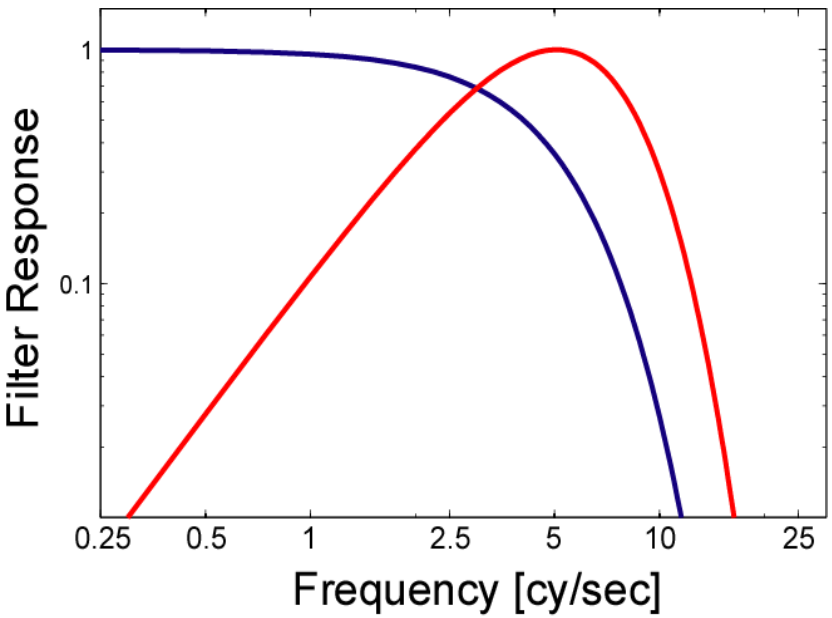
Extended Cortex Transform



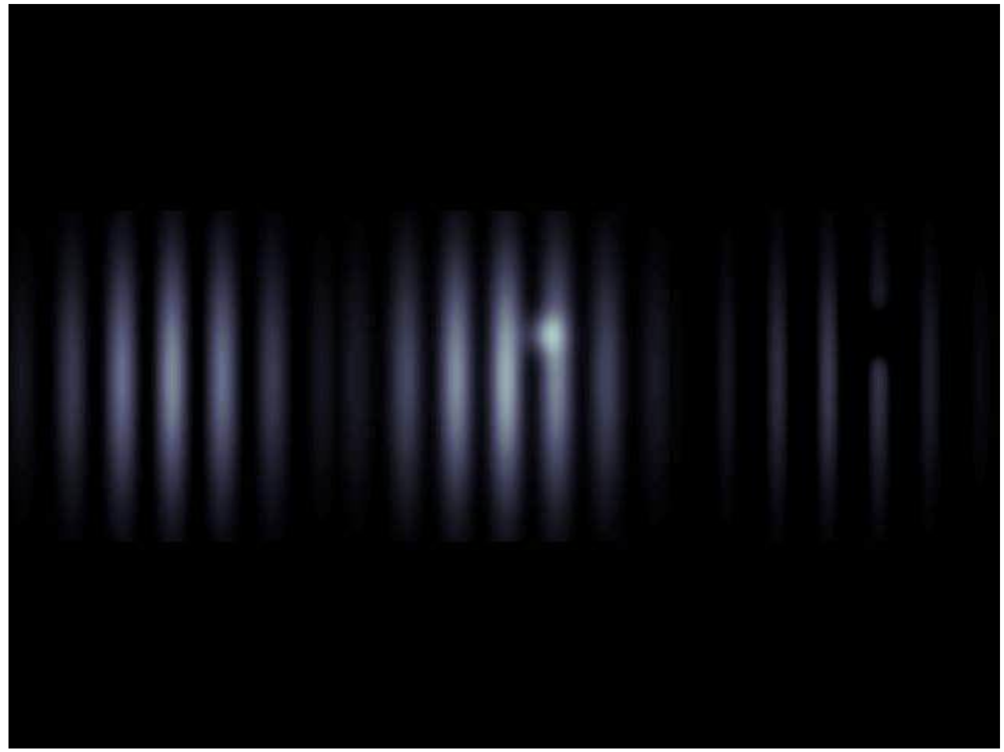
Sustained and **Transient**
Temporal Channels [Winkler 2005]



Temporal Channels



Sustained and **Transient**
Temporal Channels



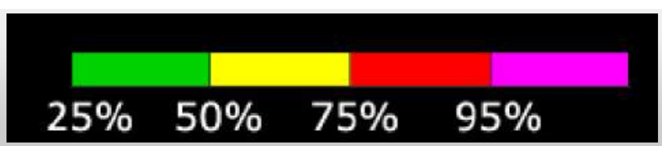
Signal **Sustained** **Transient**

Evaluation of Rendering Methods

With temporal filtering [Herzog et al. 2010]

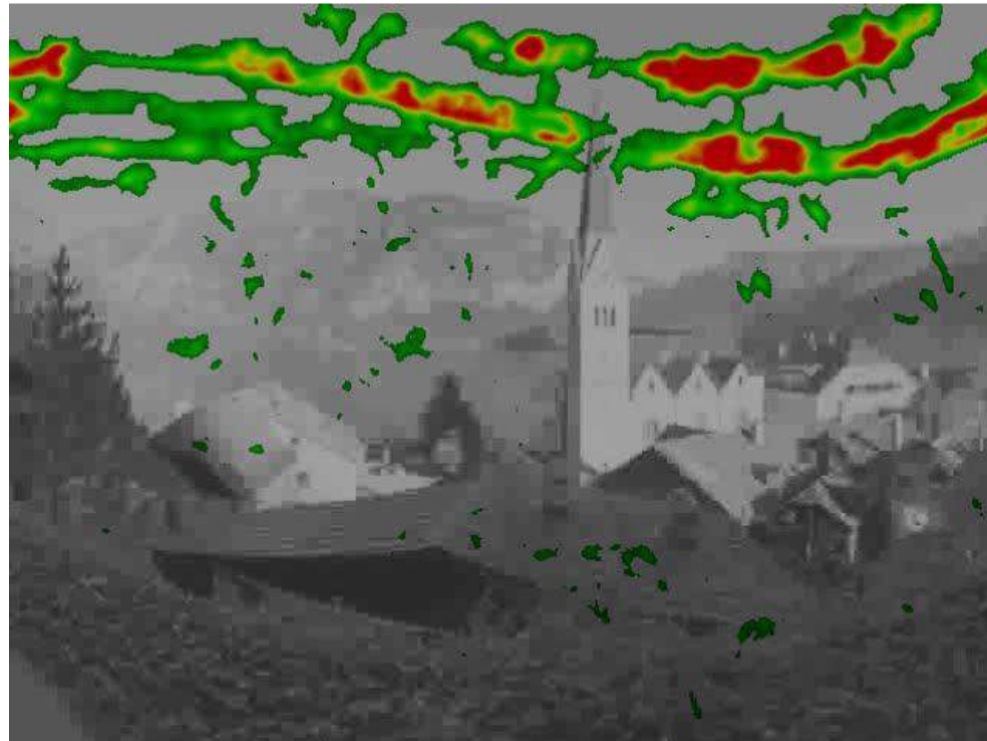
No temporal filtering

Predicted distortion map

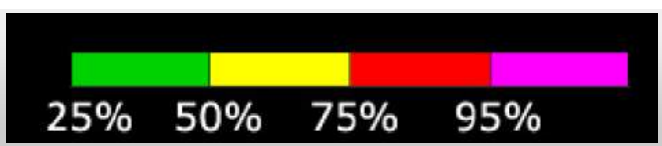


Evaluation of HDR Compression

Medium Compression

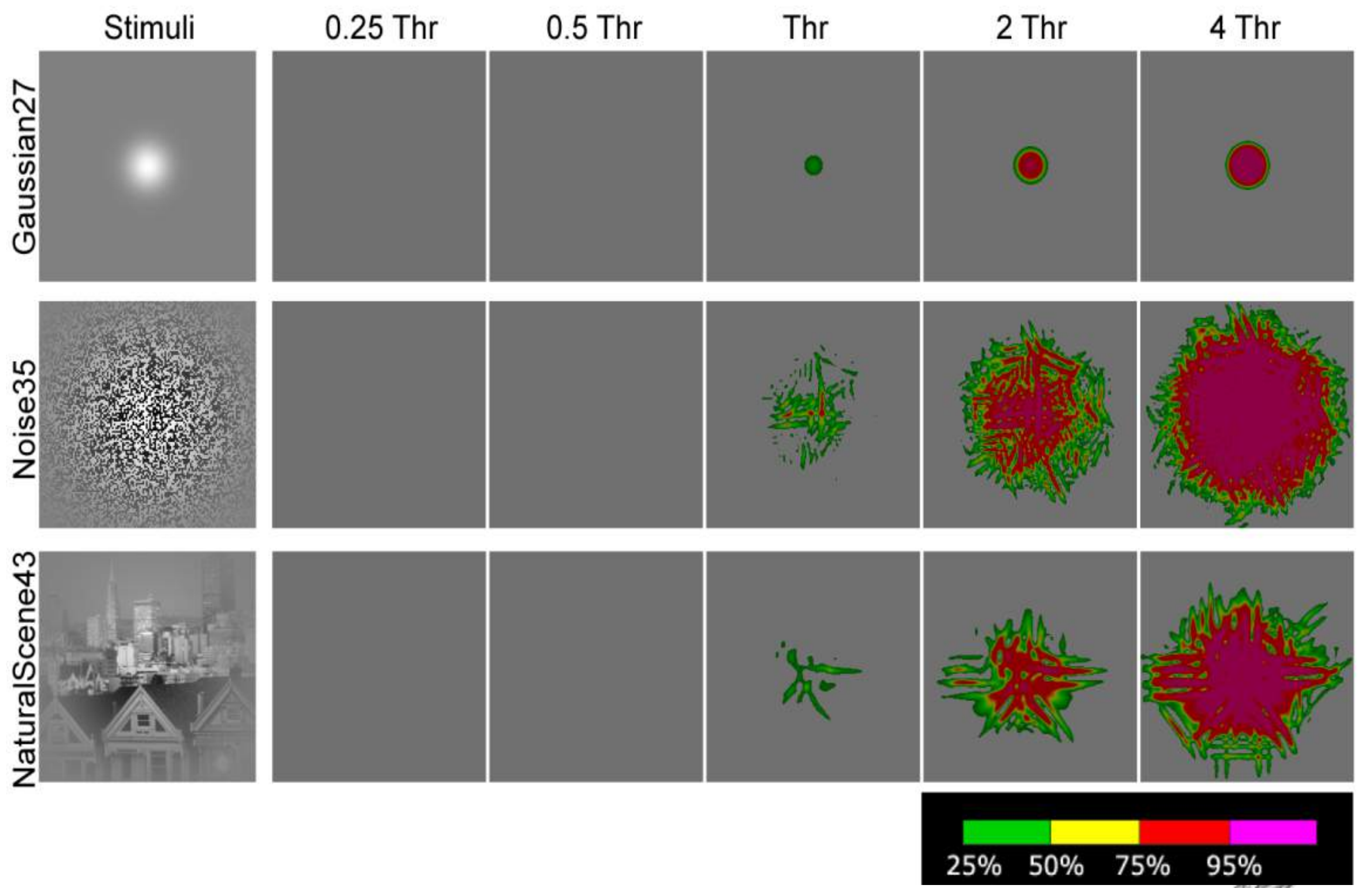


High Compression



Subjective Calibration

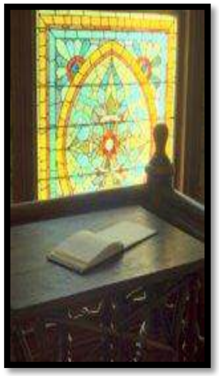
- Modelfest dataset at five contrast levels



Subjective Validation

- Example [Aydın et al. 2010, Čadík et al. 2010]
- Noise, HDR video compression, tone mapping
- “2.5D videos”
- LDR-LDR, HDR-HDR, HDR-LDR



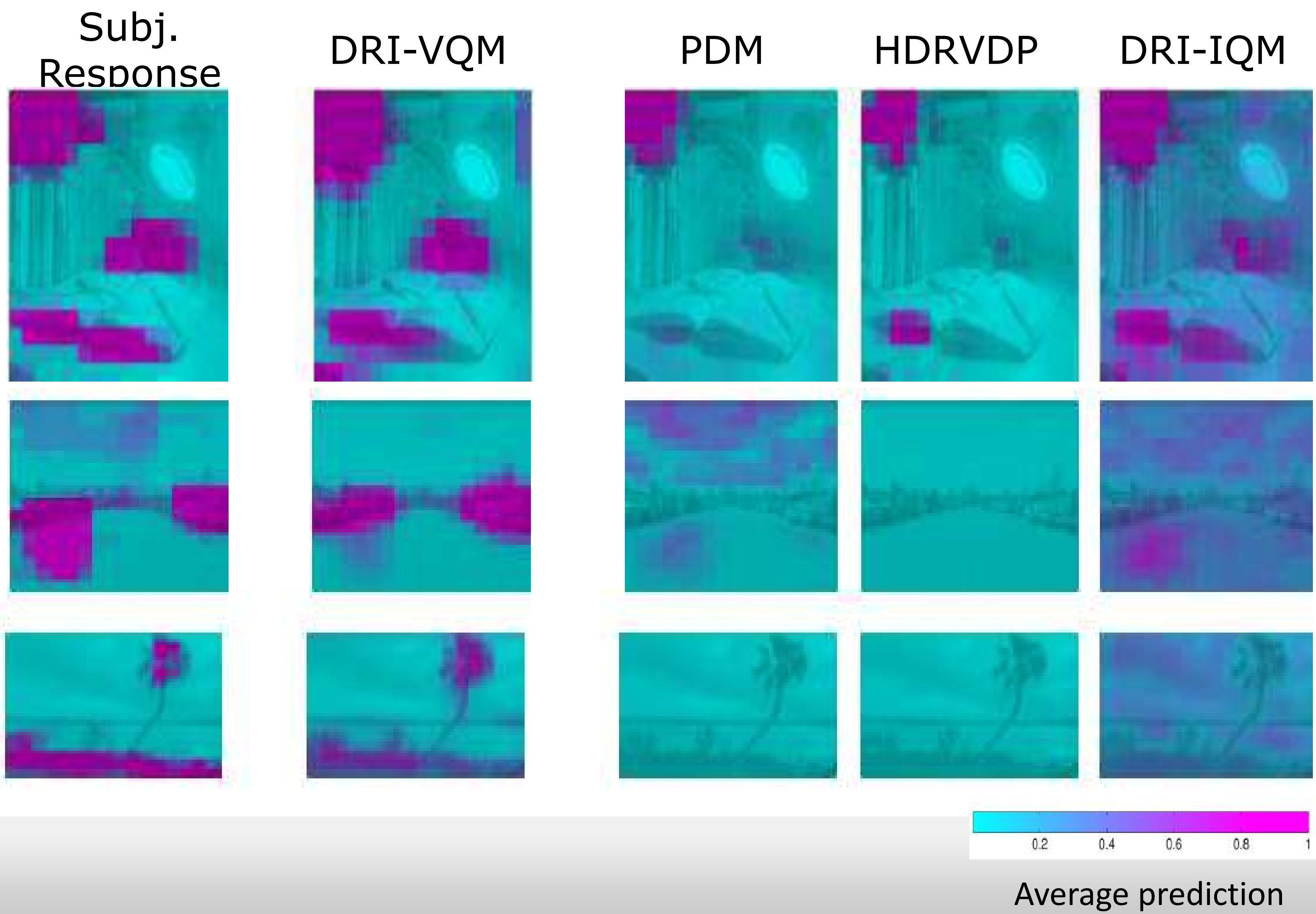



Subjective Validation, cont.

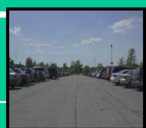
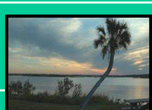
(1) Show videos side-by-side on a HDR Display

(2) Subjects mark regions where they detect differences

Subjective vs. Objective Results



Subjective Validation, cont.

Stimulus	DRI-VQM	PDM	HDRVDP	DRIVDP
1 	0.765	-0.0147	0.591	0.488
2 	0.883	0.686	0.673	0.859
3 	0.843	0.886	0.0769	0.865
4 	0.815	0.0205	0.211	-0.0654
5 	0.844	0.565	0.803	0.689
6 	0.761	-0.462	0.709	0.299
7 	0.879	0.155	0.882	0.924
8 	0.733	0.109	0.339	0.393
9 	0.753	0.368	0.473	0.617
Average	0.809	0.257	0.528	0.563

- [Čadík et al. 2010] Data available at: http://www.mpi-inf.mpg.de/resources/hdr/quality

Conclusions

- A number of established metrics are available as source code or web service
  - SSIM: https://ece.uwaterloo.ca/~z70wang/research/ssim/
  - HDRVDP : http://sourceforge.net/projects/hdrvdp/files/hdrvdp/
  - DRI-IQM and DRI-VQM:

    http://drim.mpi-inf.mpg.de/
- Researchers are starting using these metrics instead of user studies.
- Future directions:
  - Metrics for retargeted images [Liu et al. 2011]
  - Better HVS models [Mantiuk et al. 2001]
  - Smarter distortion measures.

Stereo content retargeting

Piotr Didyk

MPI Informatik

Why stereo?

Images are no longer flat

- Improves realism
- Images are not longer flat
- Better layout separation

Reproduced view dependent effects

- Improves material perception

Stereo in daily life

History of stereo

SIGGRAPH ASIA 2011 HONG KONG

1838: different images for both eyes

1890: patent on 3D movies

1900: tripod for taking 3D pictures

1915: exhibition of 3D images

1922: 3D movie

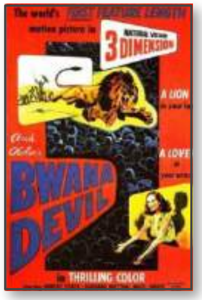
1923: 3D movie with stereo sound

1952: 3D movie in color

90's: IMAX cinemas, TV series

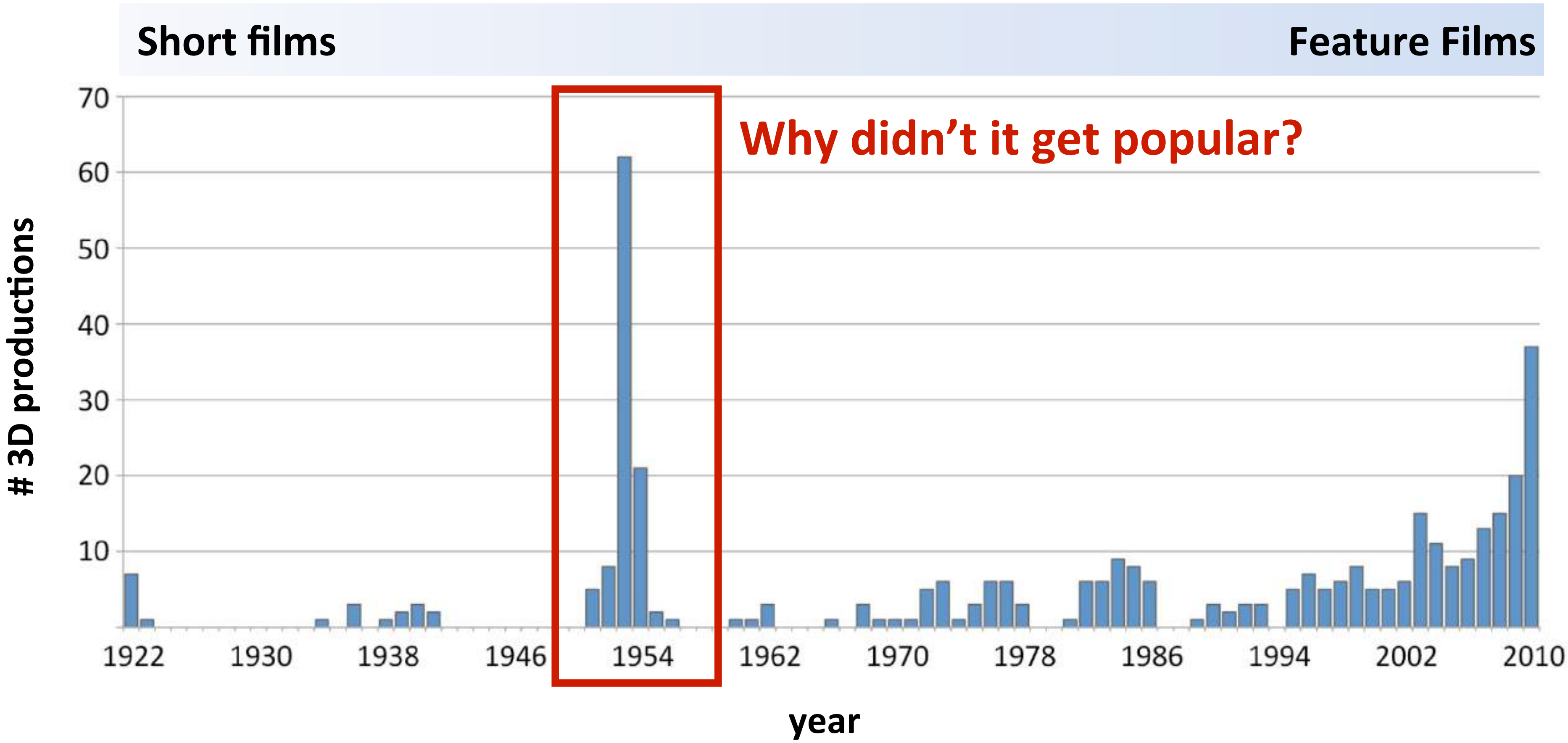
2003: feature film in 3D for IMAX

2009 - now: became very popular



Number of 3D productions

SIGGRAPHASIA2011 HONG KONG



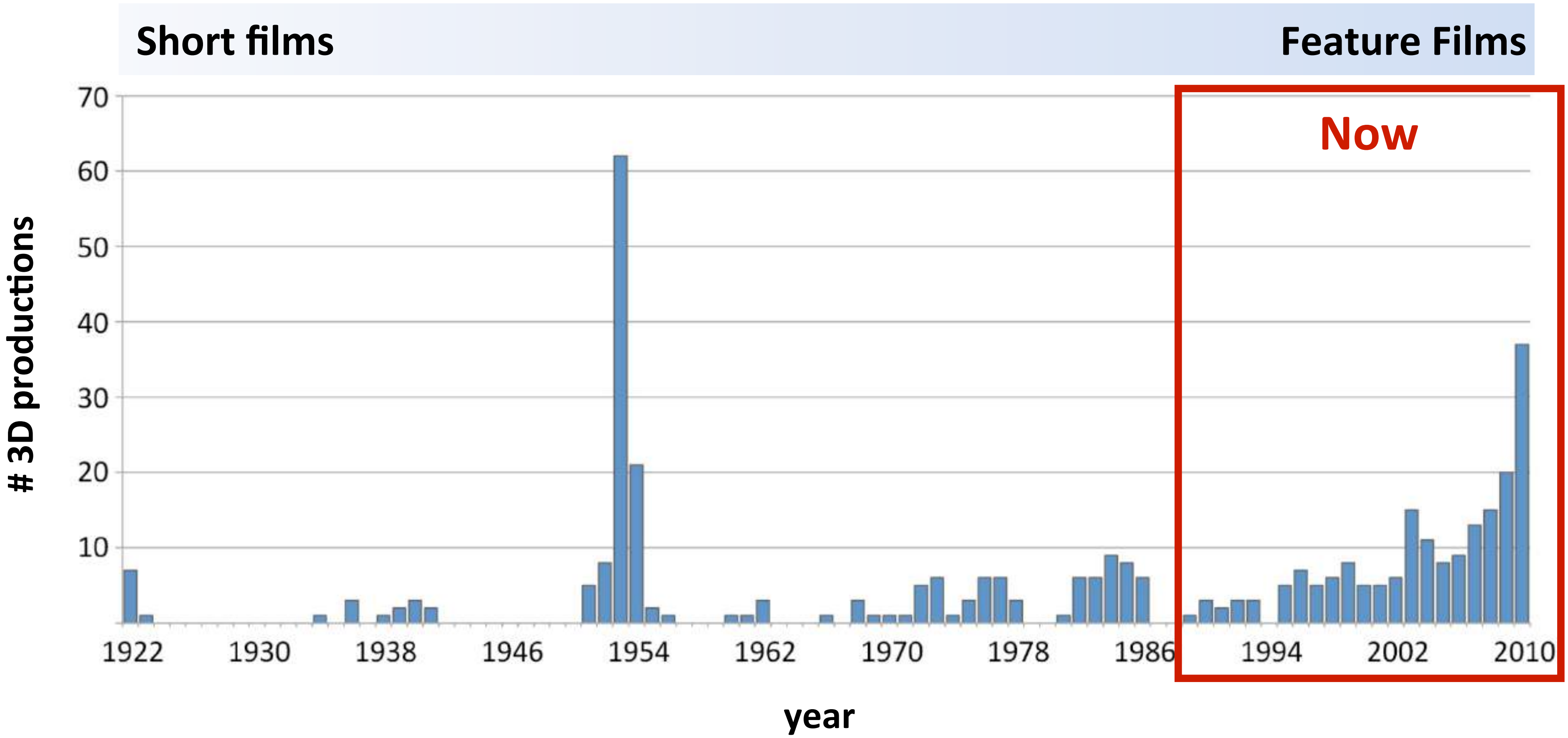
Early 3D production

SIGGRAPH ASIA 2011 HONG KONG

- Expensive hardware
- Lack of standardized format
- Impossible at home
- Lack of interesting content

Number of 3D productions

SIGGRAPHASIA2011 HONG KONG



Current 3D production

SIGGRAPHASIA2011 HONG KONG

Great content:

- Beautiful shots with complex depth
- Computer generated special effects

3D is coming to our homes:

- Equipment is getting less expensive
- 3D games / TV

New better 3D equipment:

- Shutter glasses
- Polarized glasses
- Autostereoscopic displays are getting better

Current 3D production

SIGGRAPHASIA2011 HONG KONG

Great content:

- Beautiful shots with complex depth
- Computer generated special effects

3D is coming to our homes:

- Equipment is getting less expensive
- 3D games / TV

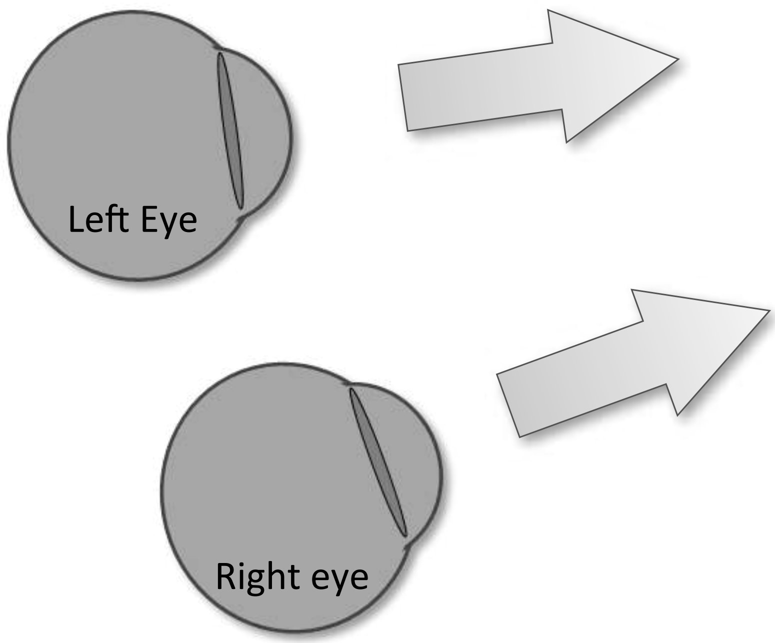
New better 3D equipment:

- Shutter glasses
- Polarized glasses
- Autostereoscopic displays are getting better

They are flat!

Stereo on a flat display

- **Different image for each eye**

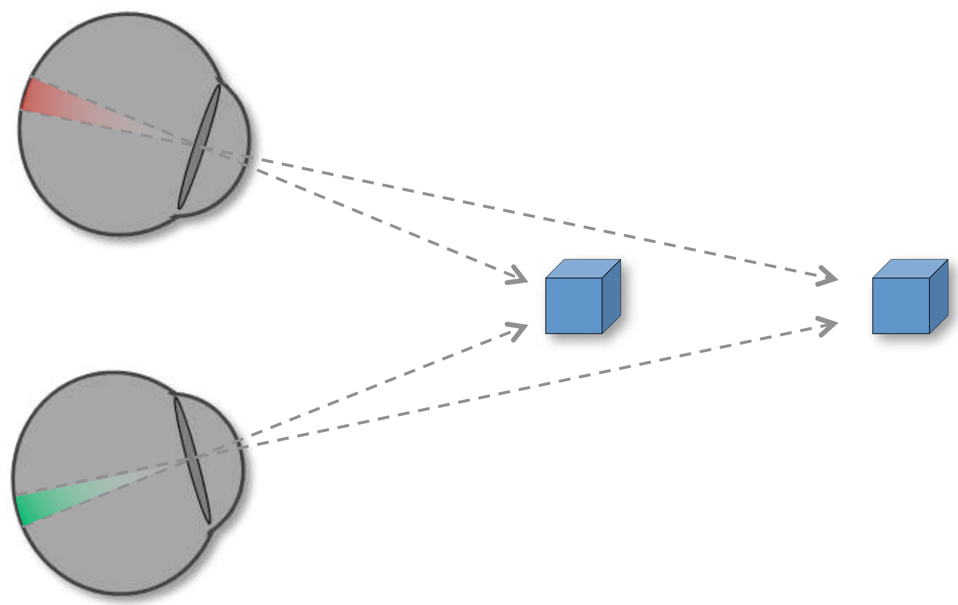


Depth perception

SIGGRAPHASIA2011 HONG KONG

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

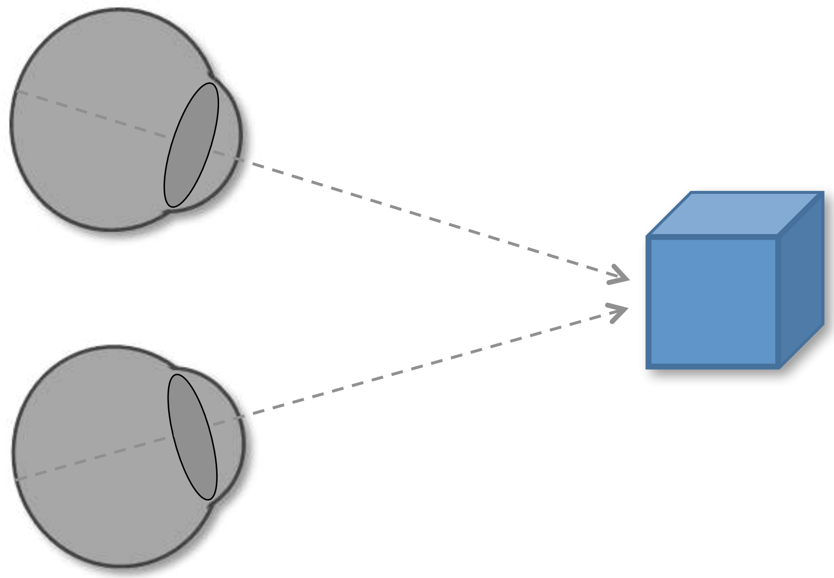


Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation,



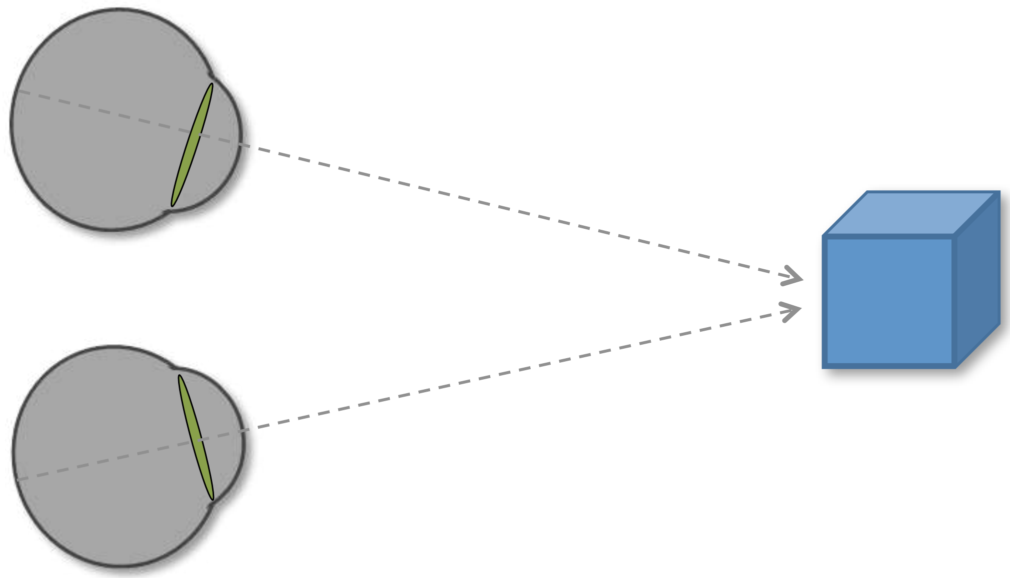
Depth perception

SIGGRAPHASIA2011 HONG KONG

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation,



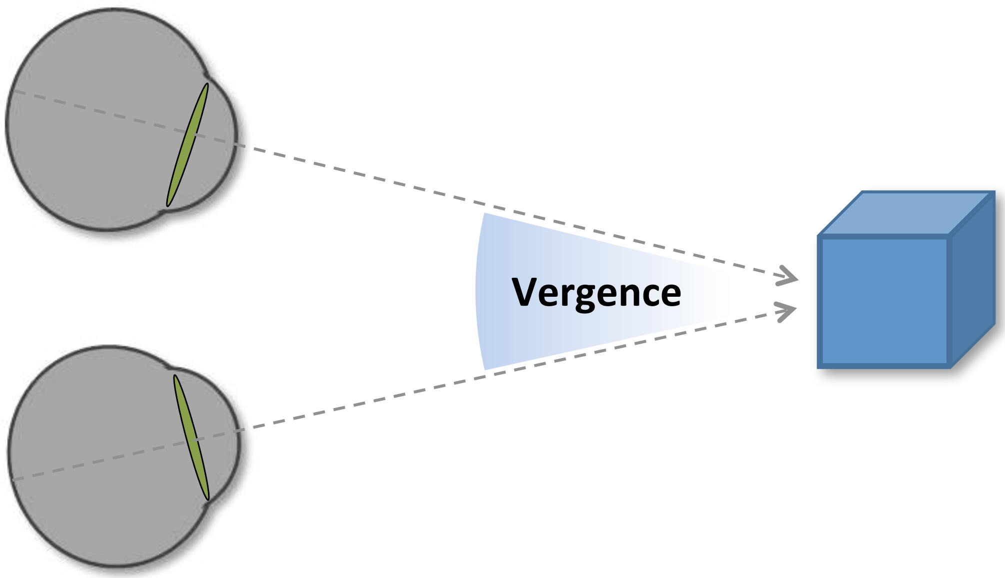
Depth perception

SIGGRAPHASIA2011 HONG KONG

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence



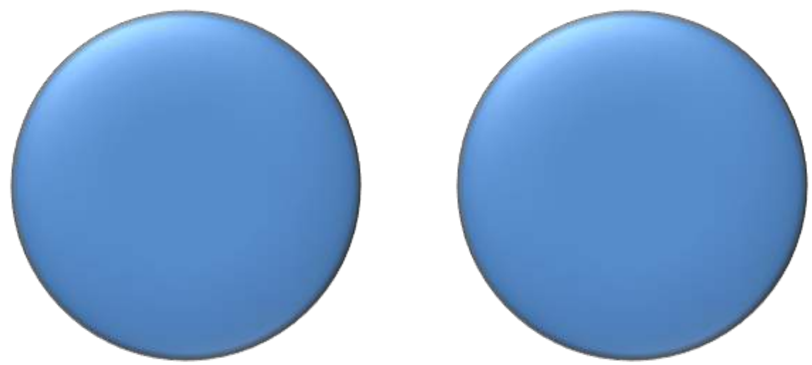
Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

Pictorial depth cues:



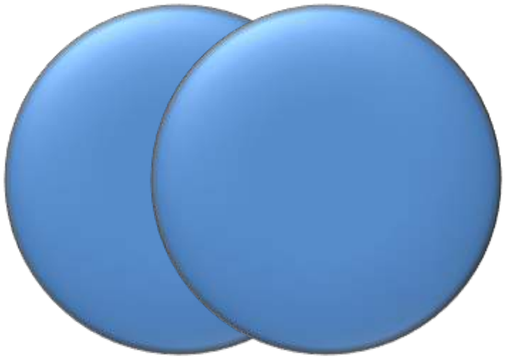
Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
occlusion,



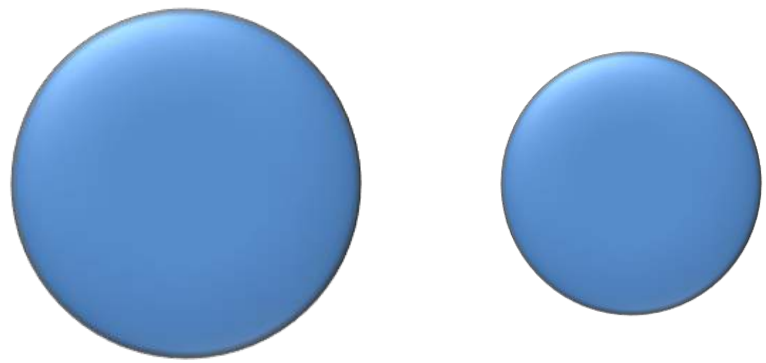
Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
occlusion, size,



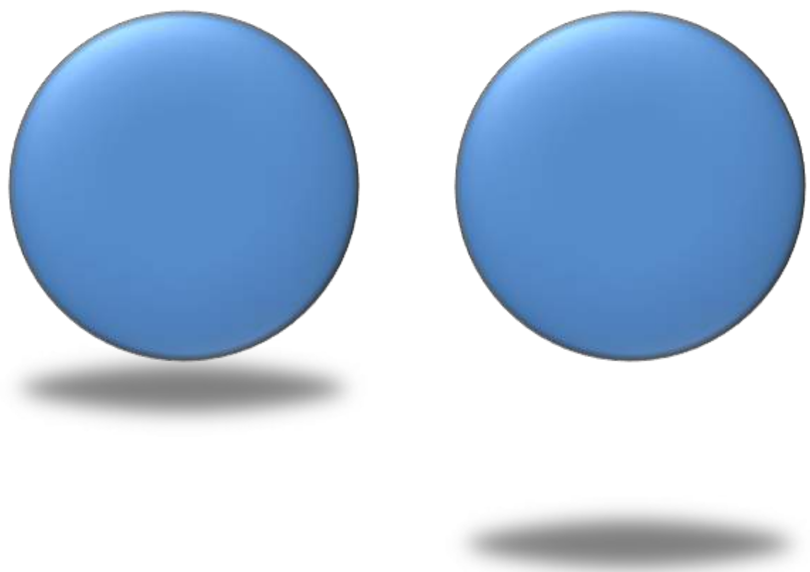
Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

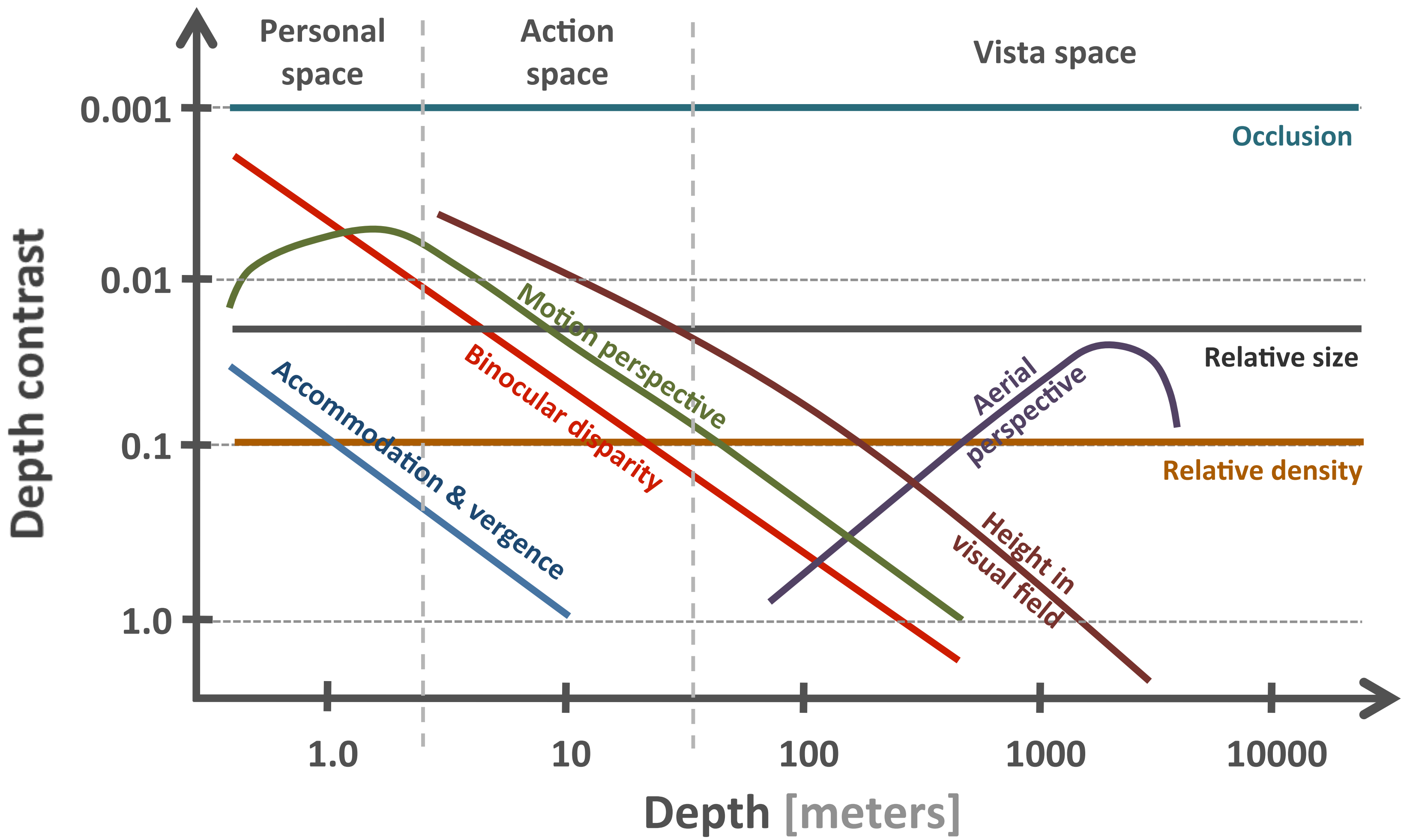
Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
occlusion, size, shadows…



Cues sensitivity

SIGGRAPHASIA2011 HONG KONG



"Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth"
by Cutting and Vishton [1995]

Depth perception

SIGGRAPHASIA2011 HONG KONG

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
occlusion, size, shadows…

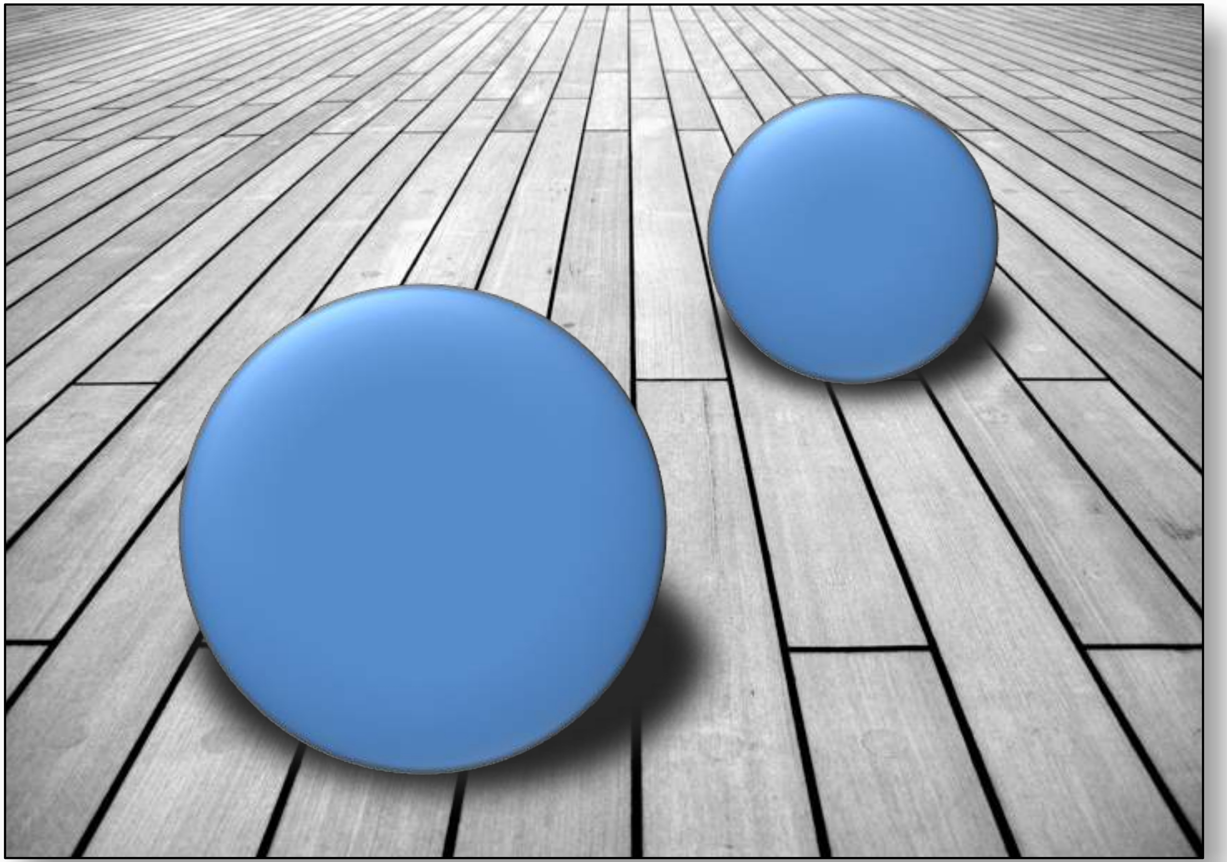
Challenge:
Consistency is required!

Simple conflict example

SIGGRAPHASIA2011 HONG KONG

Present cues:

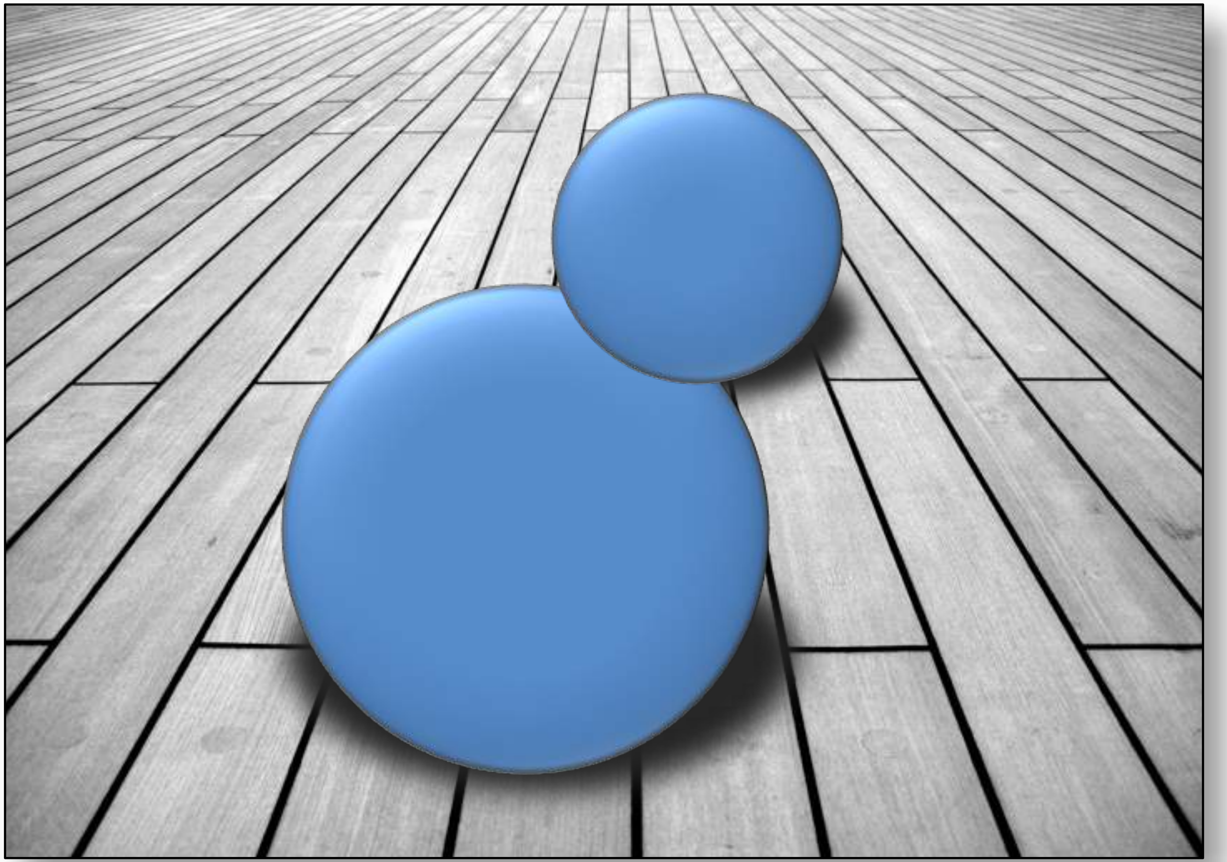
- Size
- Shadows
- Perspective



Simple conflict example

Present cues:

- Size
- Shadows
- Perspective
- Occlusion



Disparity & occlusion conflict

Objects in front

Disparity & occlusion conflict

Disparity & occlusion conflict

Disparity & occlusion conflict

Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

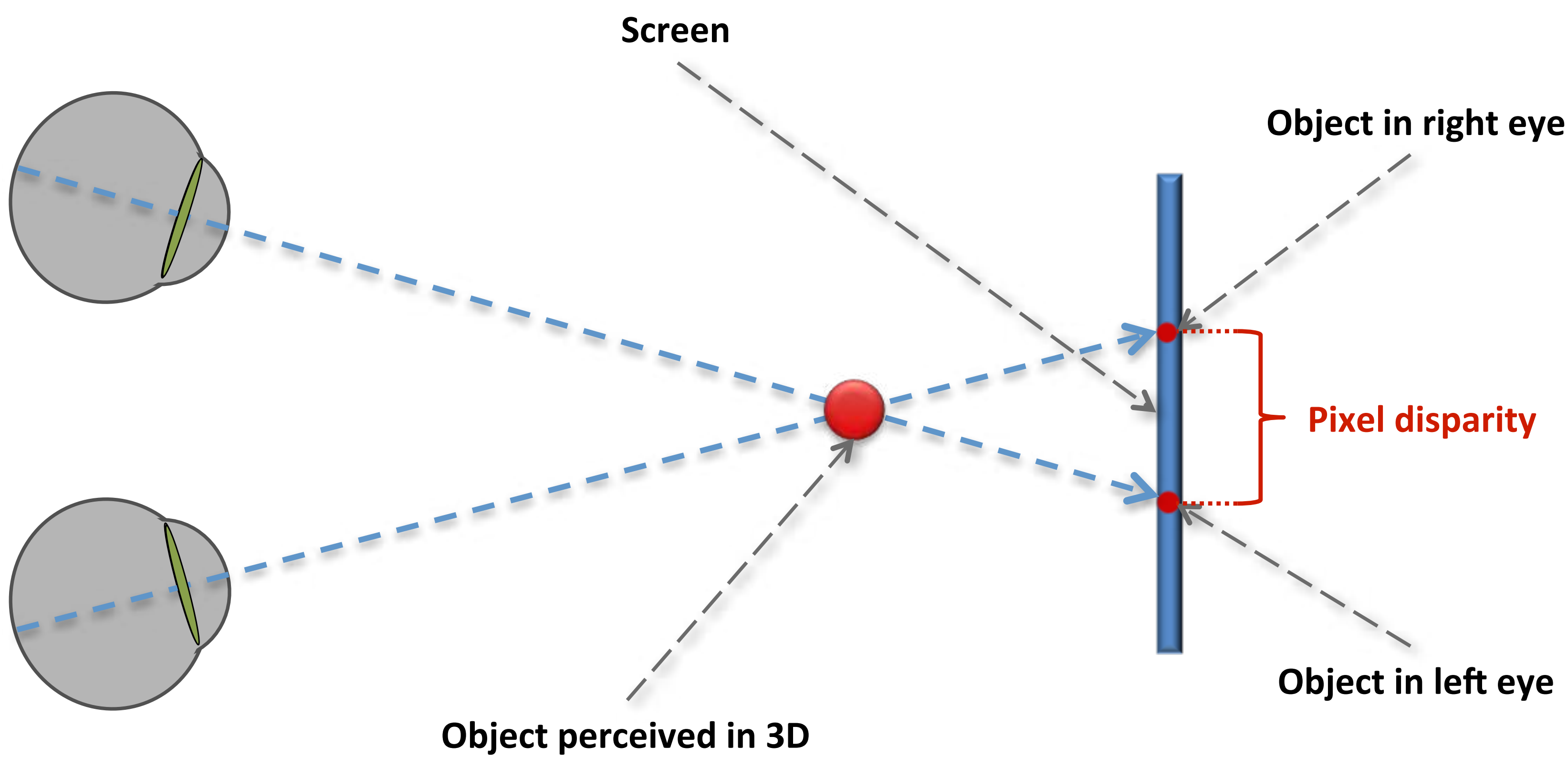
→ **Require 3D space**

We cheat our Human Visual System!

Pictorial depth cues:
occlusion, size, shadows…

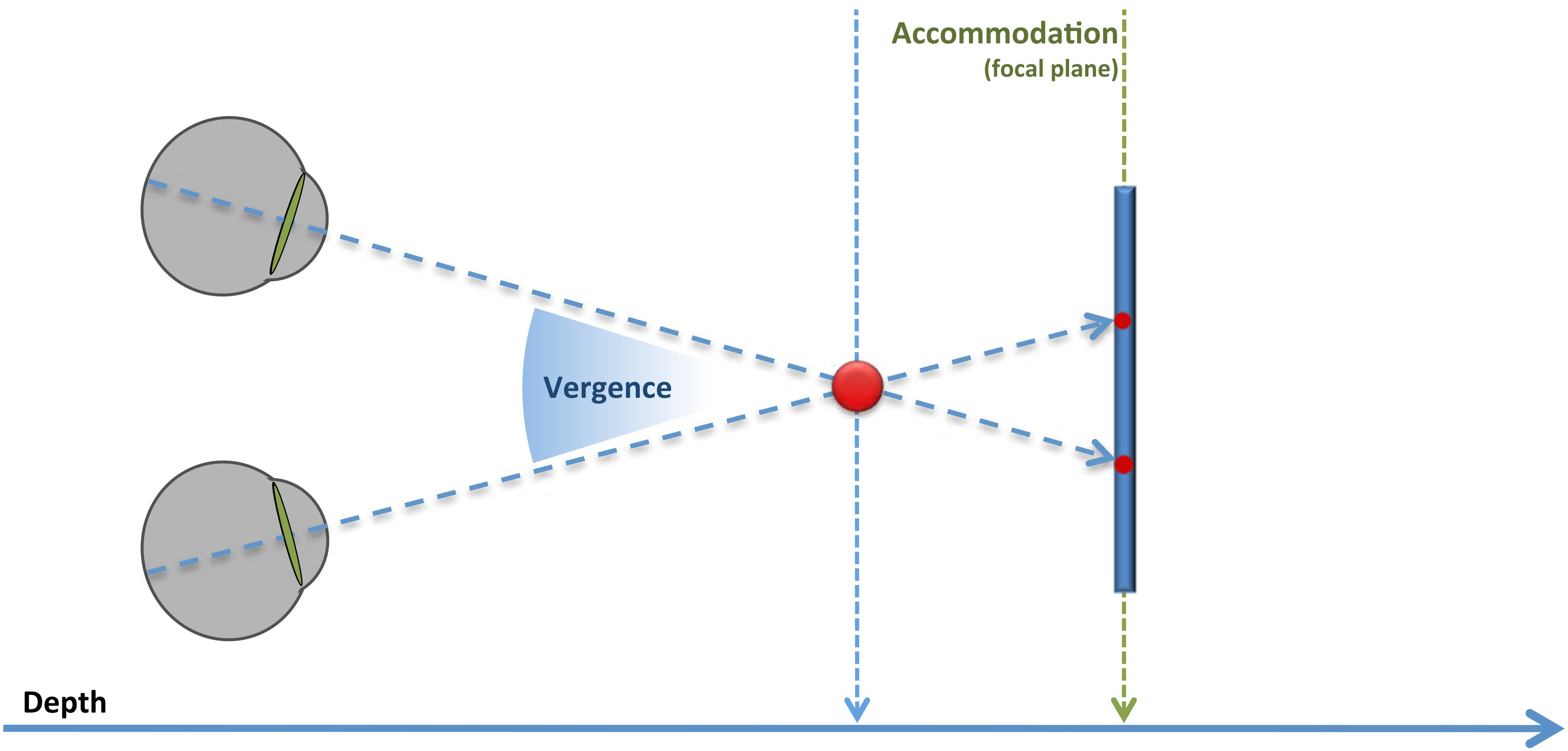
→ **Reproducible on flat displays**

Cheating our HVS



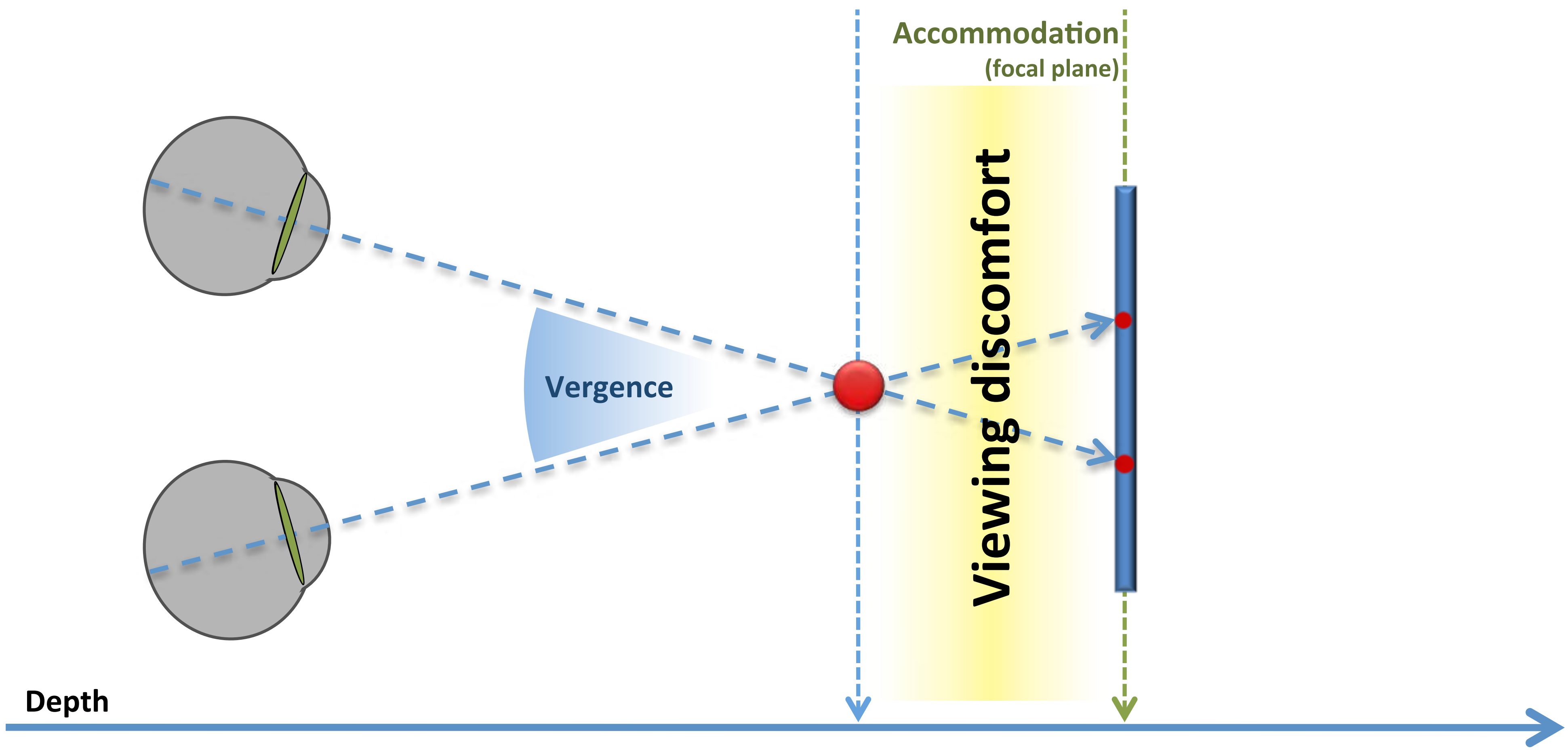
Cheating our HVS

SIGGRAPH ASIA 2011 HONG KONG

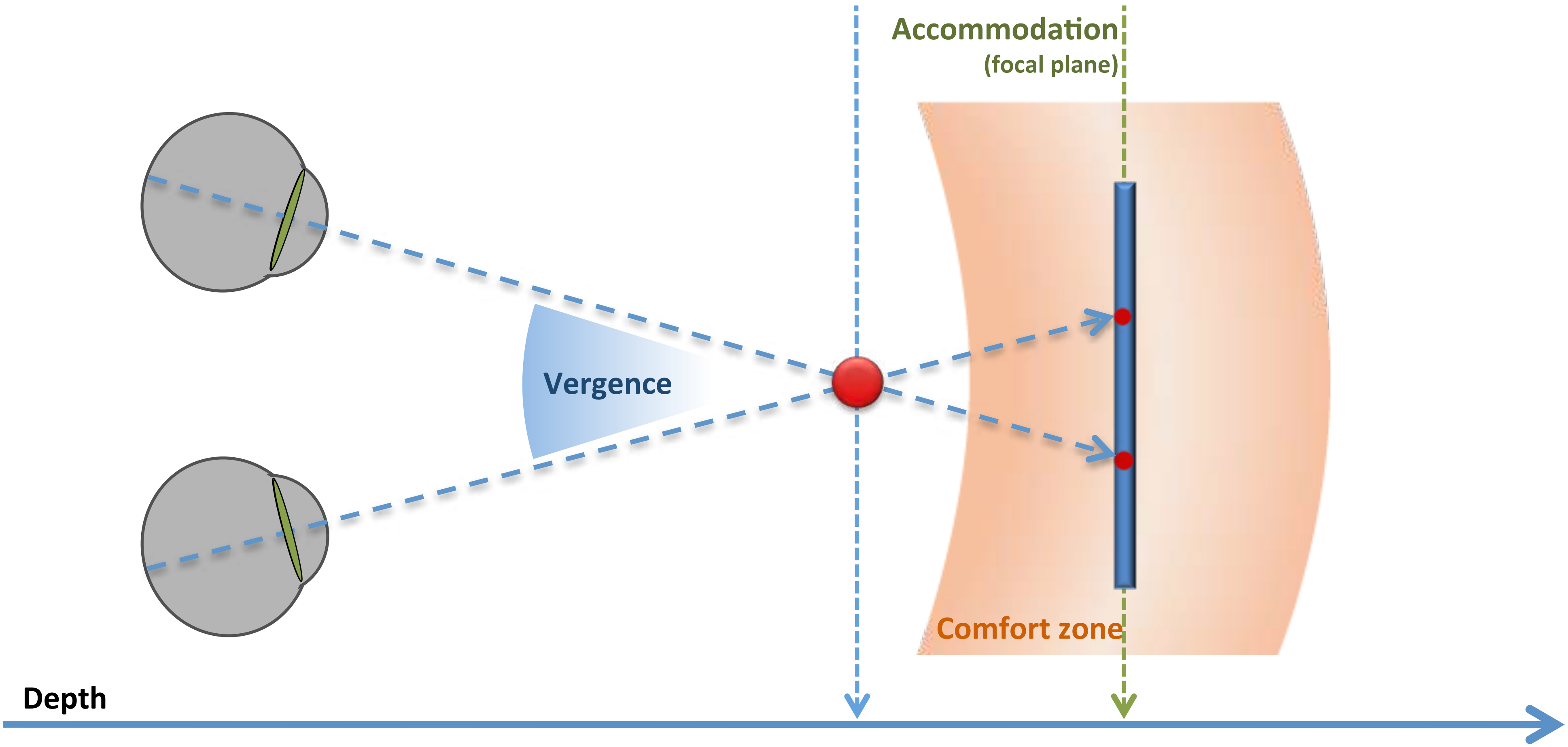


Cheating our HVS

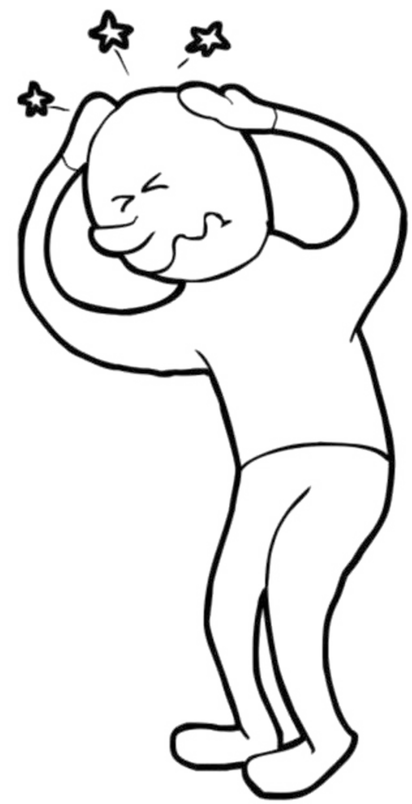
SIGGRAPHASIA2011 HONG KONG



Cheating our HVS



Viewing discomfort

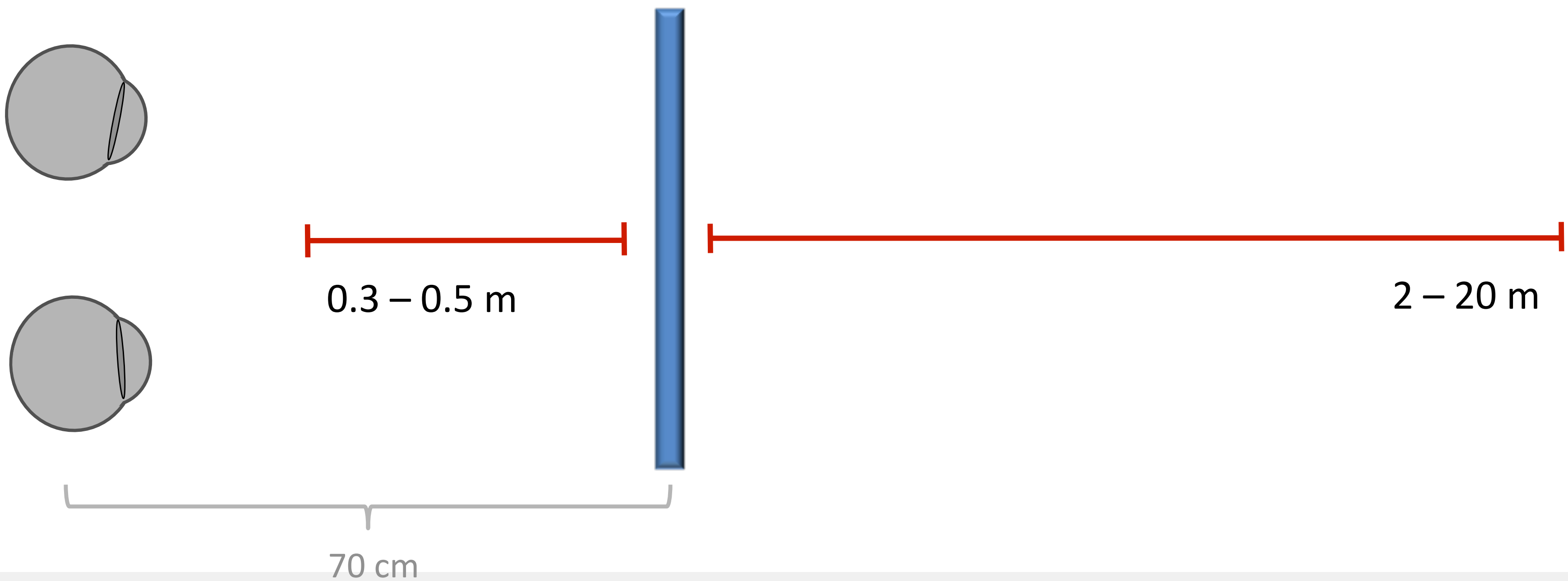


Comfort zones

Comfort zone size depends on:

- Presented content
- Viewing condition

Simple scene



"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

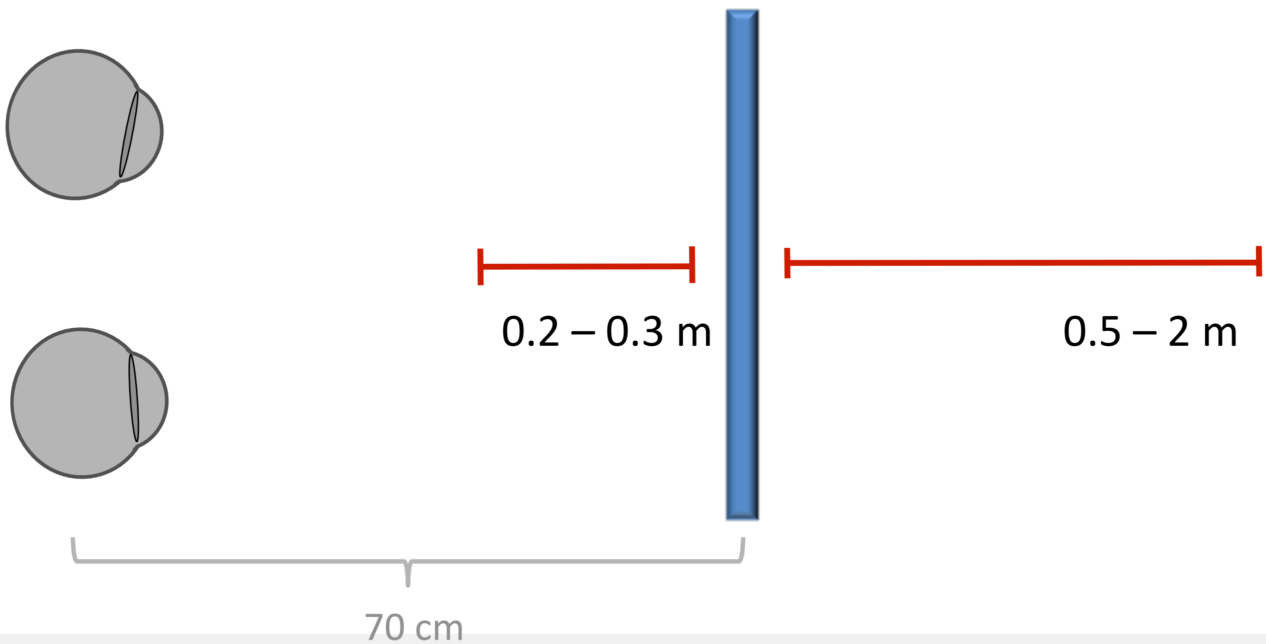
Comfort zones

SIGGRAPH ASIA 2011 HONG KONG

Comfort zone size depends on:

- Presented content
- Viewing condition

Simple scene, user allowed to look away from screen



"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

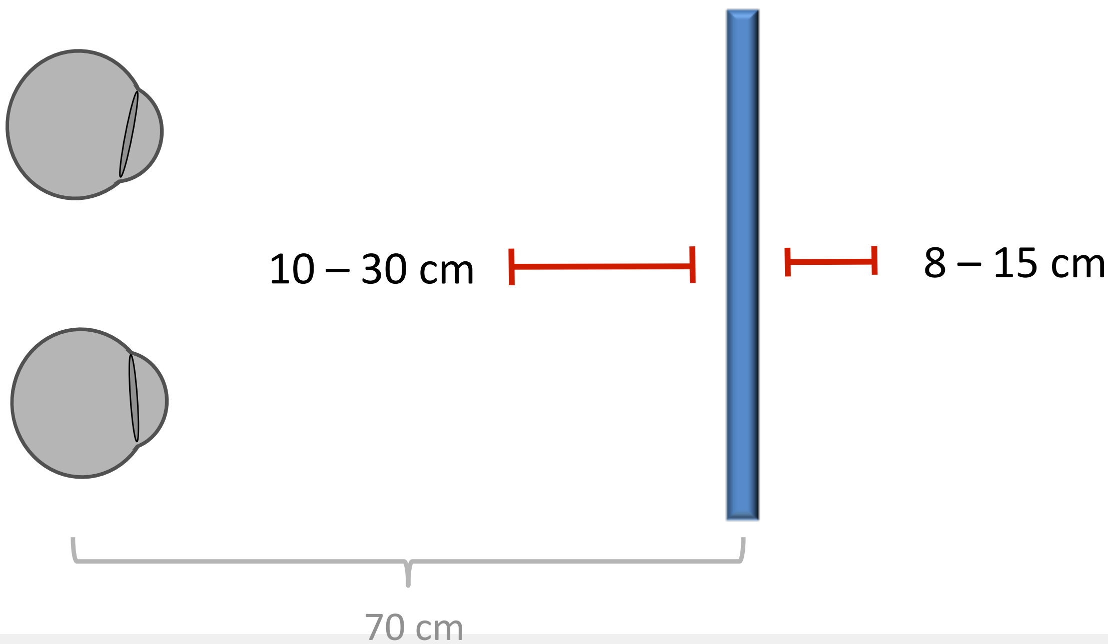
Comfort zones

SIGGRAPHASIA2011 HONG KONG

Comfort zone size depends on:

- Presented content
- Viewing condition

Difficult scene



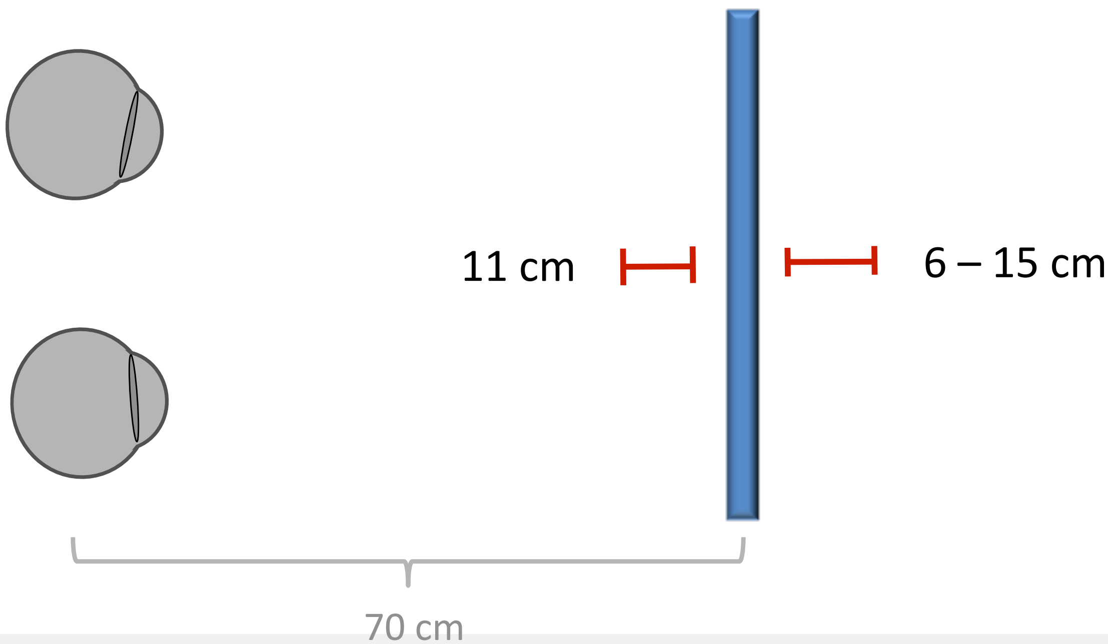
"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

Comfort zones

Comfort zone size depends on:

- Presented content
- Viewing condition

Difficult scene, user allowed to look away from screen



"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

Comfort zones

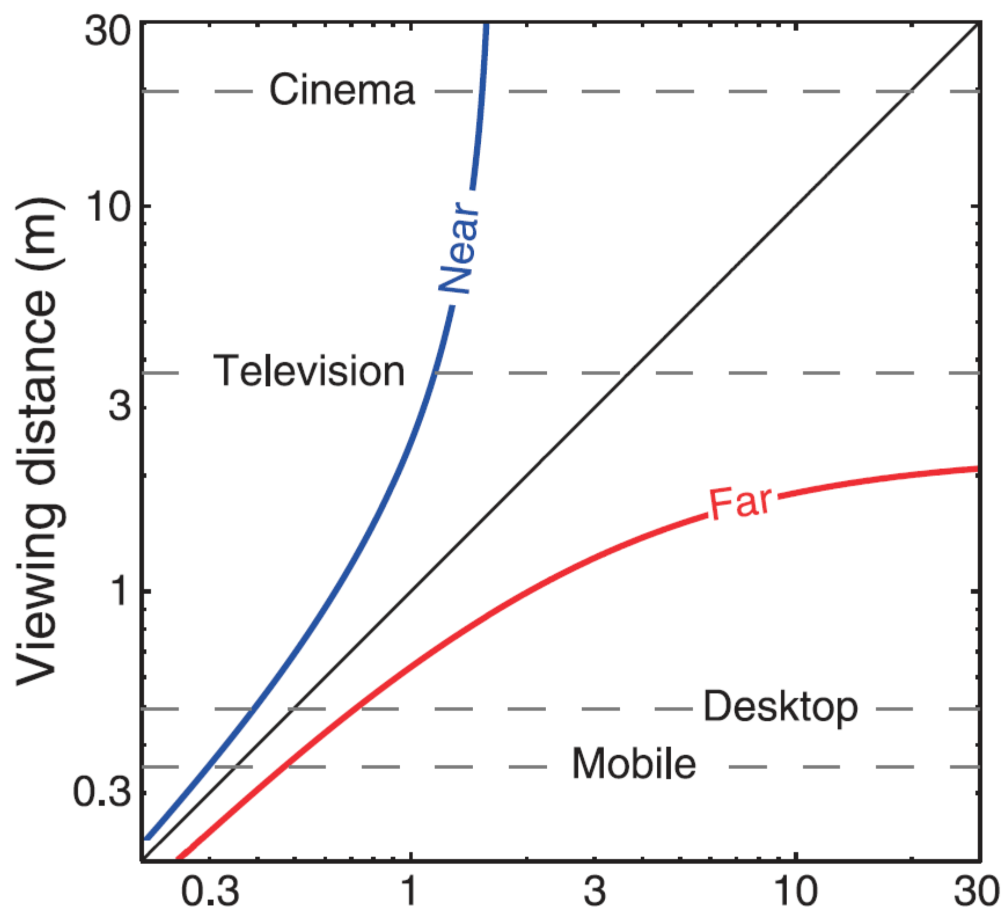
SIGGRAPHASIA2011 HONG KONG

Comfort zone size depends on:

- Presented content
- Viewing condition
- Screen distance

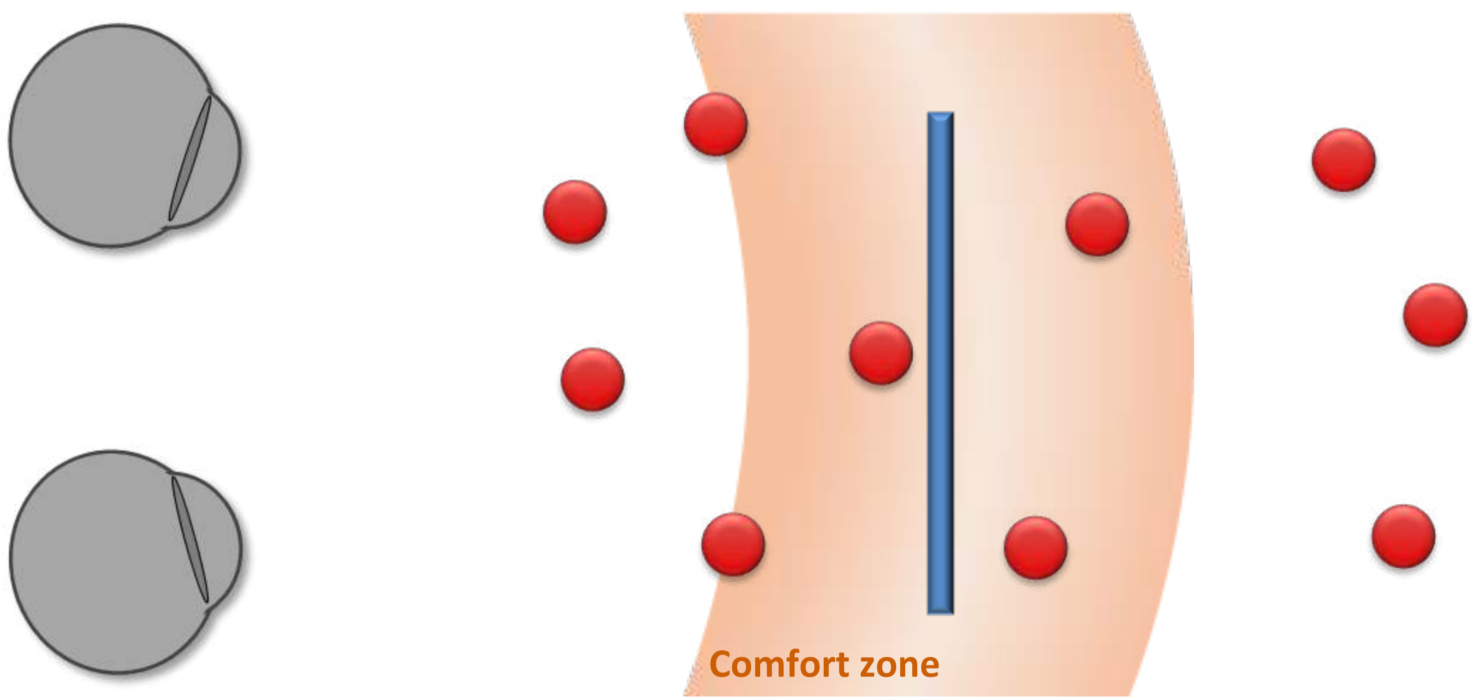
Other factors:

- Distance between eyes
- Depth of field
- Temporal coherence



"The zone of comfort: Predicting visual discomfort with stereo displays" by Shibata et al. 2011

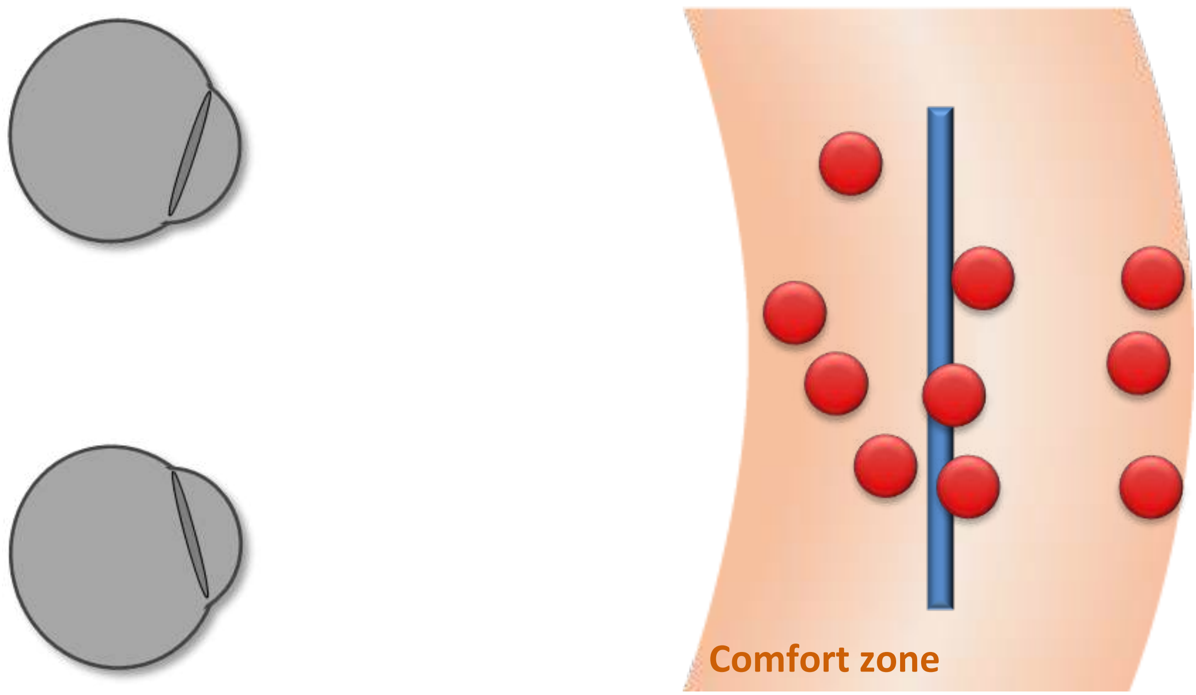
Depth manipulation



Viewing discomfort

Depth manipulation

SIGGRAPHASIA2011 HONG KONG

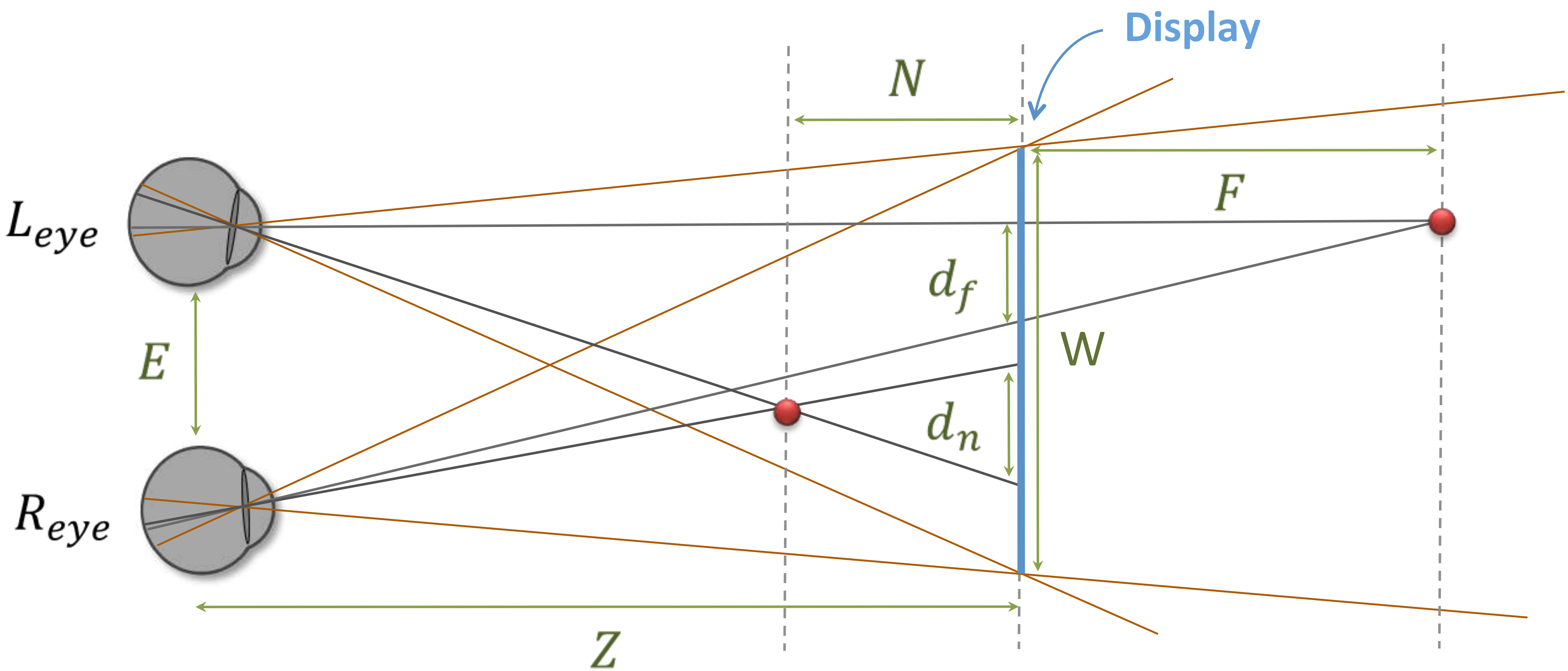


Viewing discomfort $\xrightarrow{\text{Scene manipulation}}$ **Viewing comfort**

Camera manipulations

SIGGRAPHASIA2011 HONG KONG

Viewer/Display space

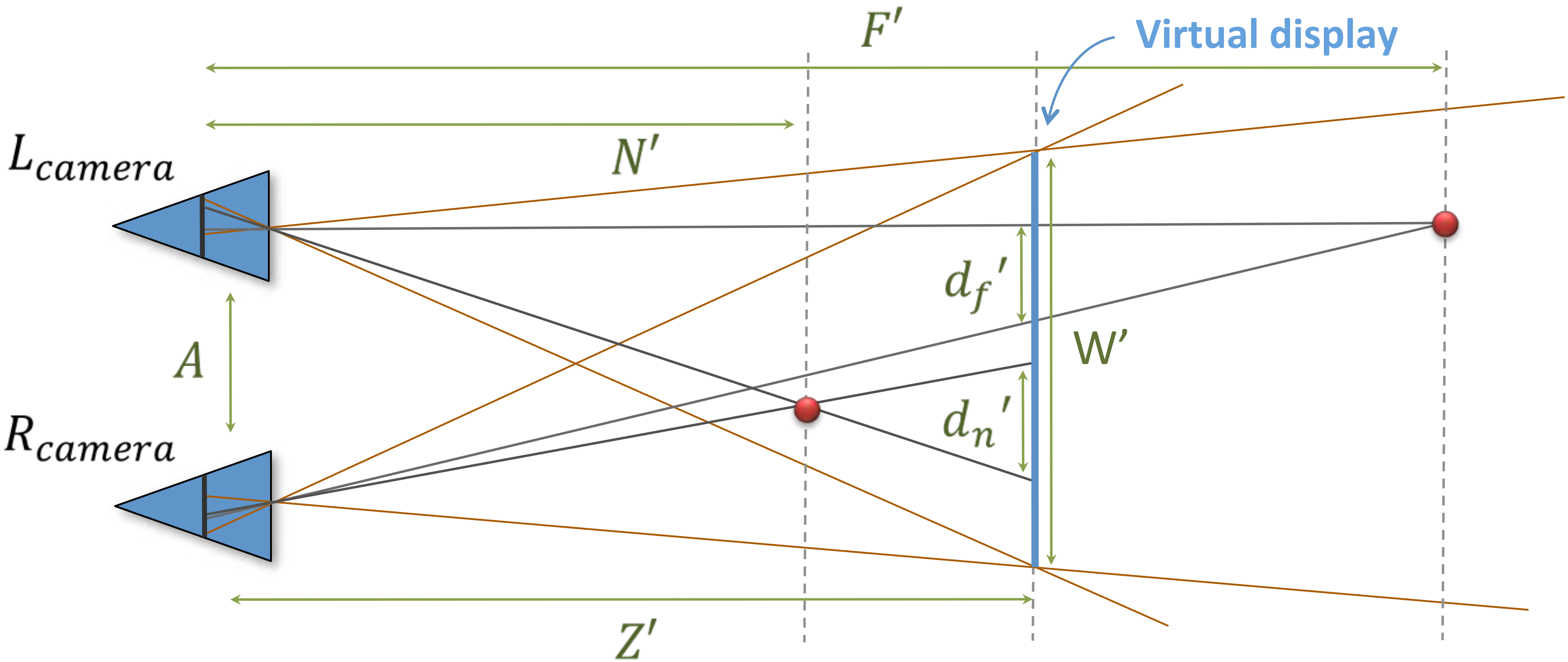


"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

Camera manipulations

SIGGRAPHASIA2011 HONG KONG

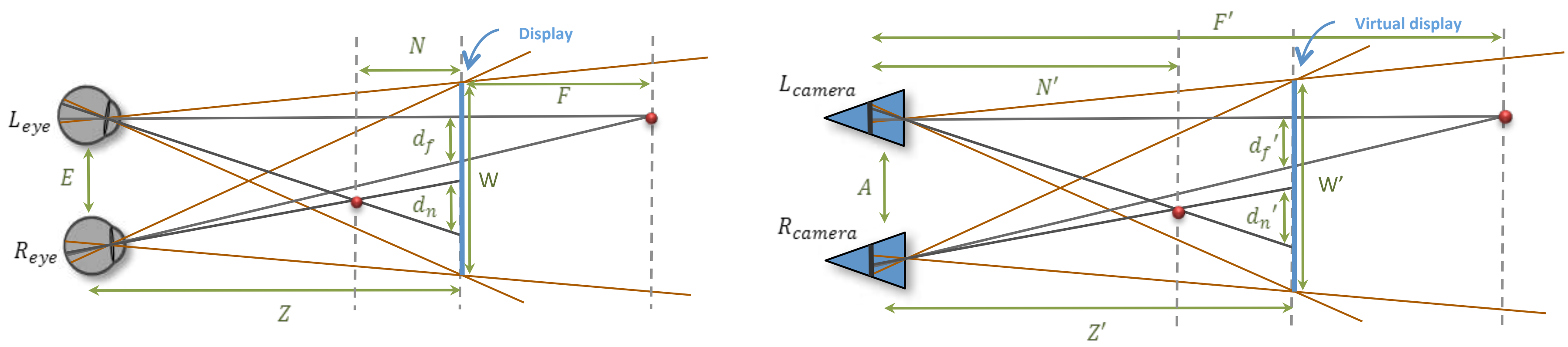
Camera/Scene space



"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

Camera manipulations

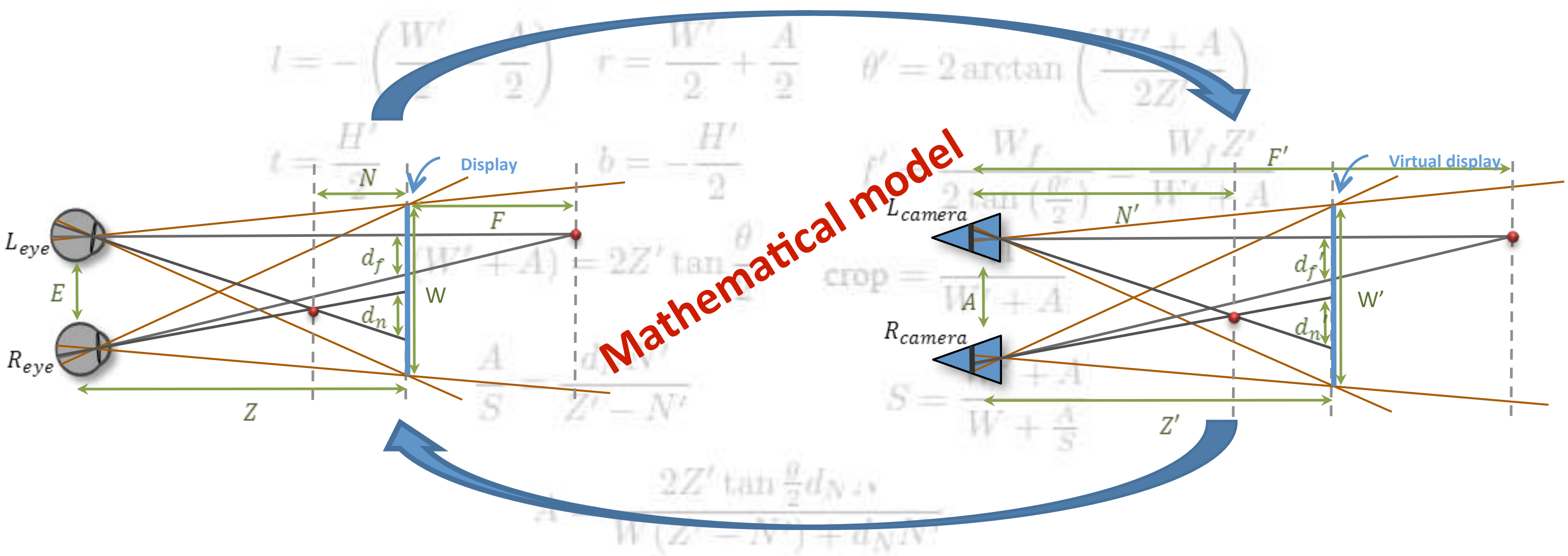
Camera/Scene space



- The parameters can be the same
  - ➢ may cause discomfort
- Different parameters for capturing the scene

"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

Camera manipulations



- Define the disparity limits
- Calculate appropriate camera parameters
- Adjustment in each frame
- Compensation for viewer motion

"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001
"Evaluating methods for controlling depth perception in stereoscopic cinematography" by Sun et al. 2009

Camera manipulations

SIGGRAPHASIA2011 HONG KONG

General procedure:

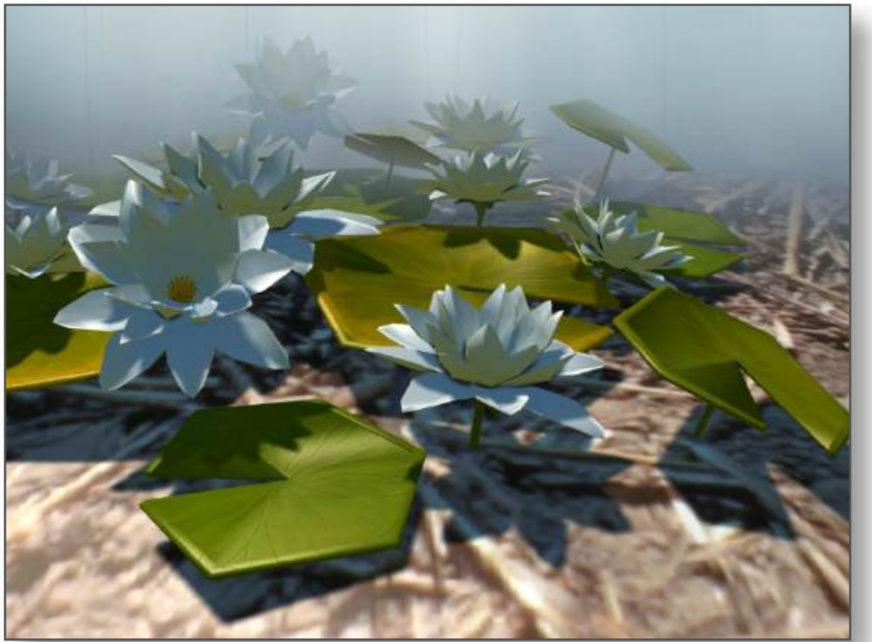
1. Define viewing condition
2. Adjust cameras parameters
3. Capturing

Displaying on different device:

- Potential discomfort
- Recapturing ?

Stereo content

SIGGRAPH ASIA 2011 HONG KONG



Left view

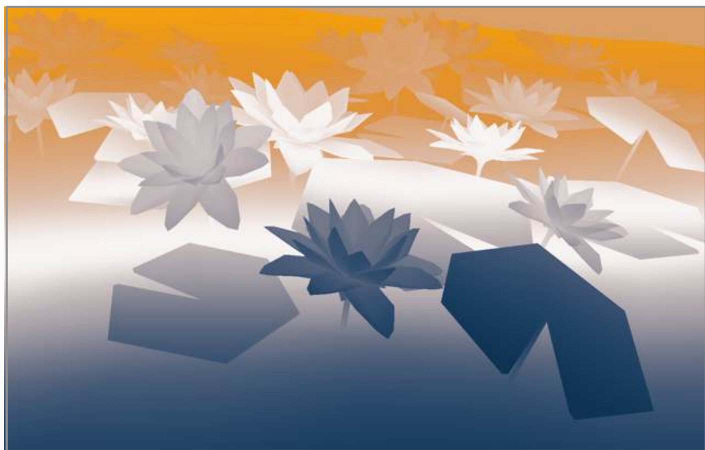
Right view

Can we have pixel disparity / depth ?

Sources of pixel disparity

SIGGRAPHASIA2011 HONG KONG

Stereo image pair



Pixel disparity map

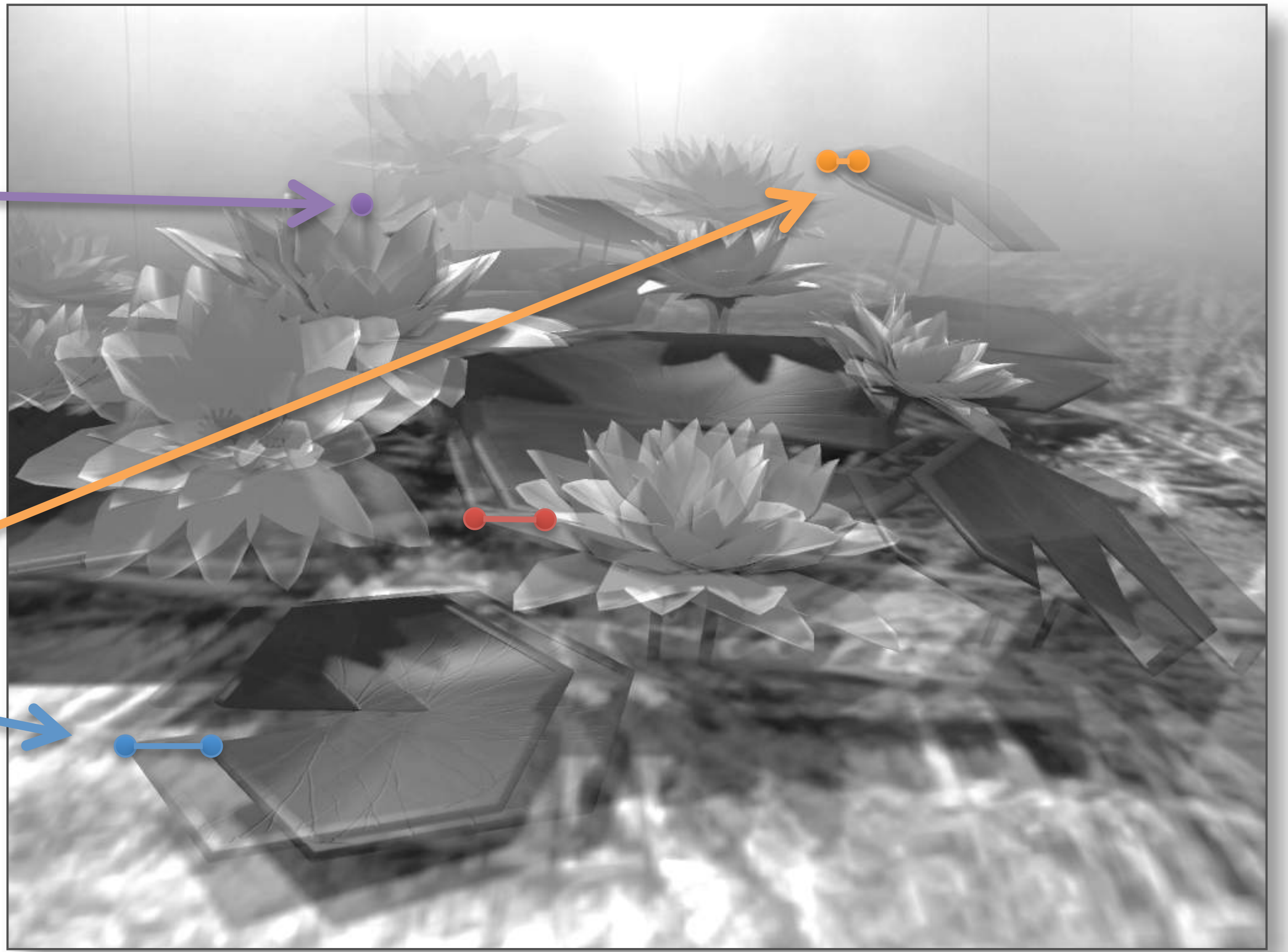
Rendering ⟶ **Usually available**

Only image pair ⟶ **Computer vision technique**

Pixel disparity

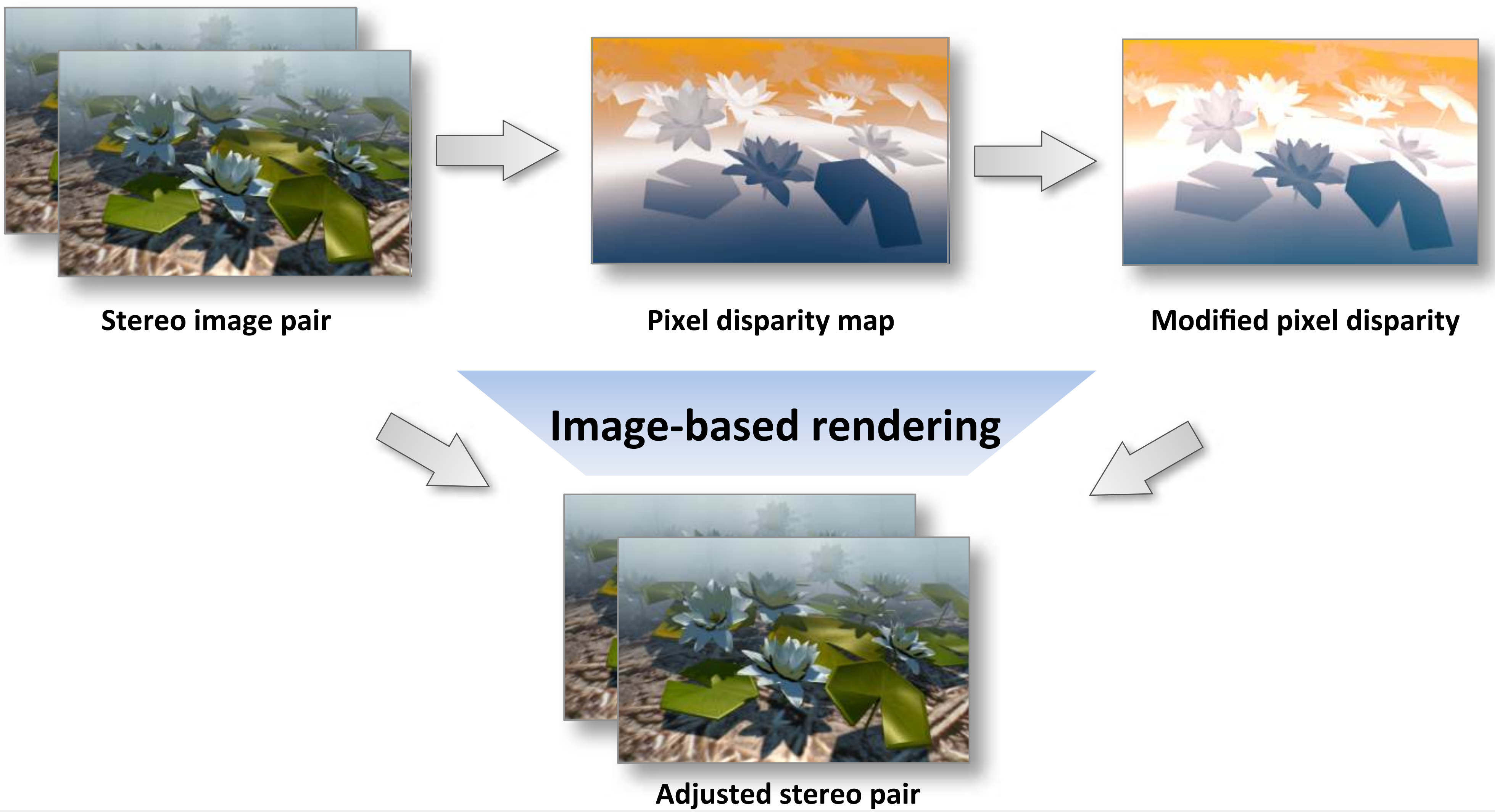
Zero disparity
on the screen plane

Bigger disparities
in front and behind screen



Left + right view

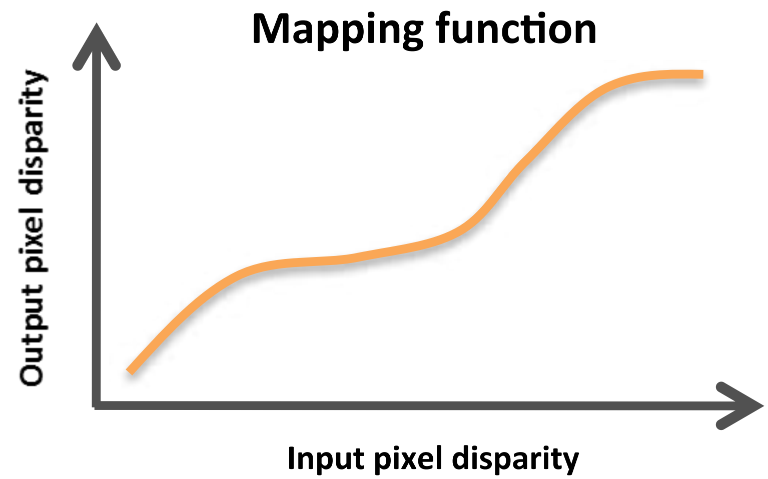
Disparity manipulations



Disparity manipulations

Pixel disparity map

Modified pixel disparity



Function:

- Liner
- Logarithmic
- Content dependent

Other possibilities:

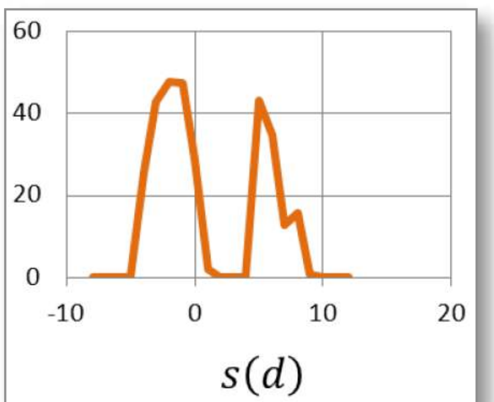
- Gradient domain
- Local operators

"Nonlinear Disparity Mapping for Stereoscopic 3D" by Lang et al. 2010

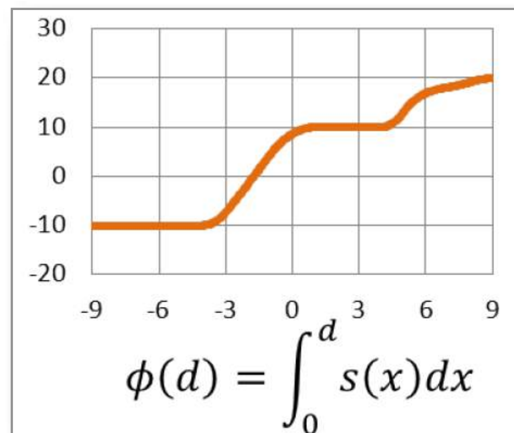
Saliency map

Input stereo image

Saliency map



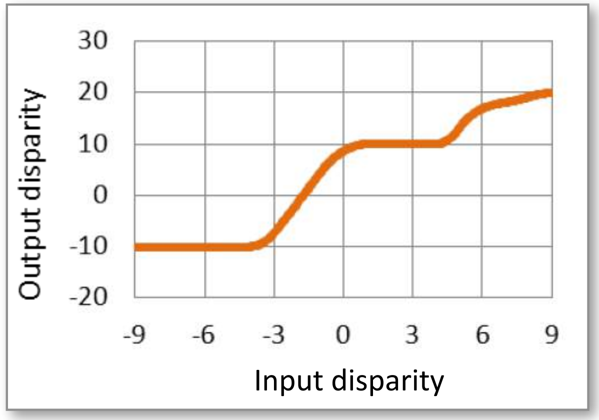
Disparity importance

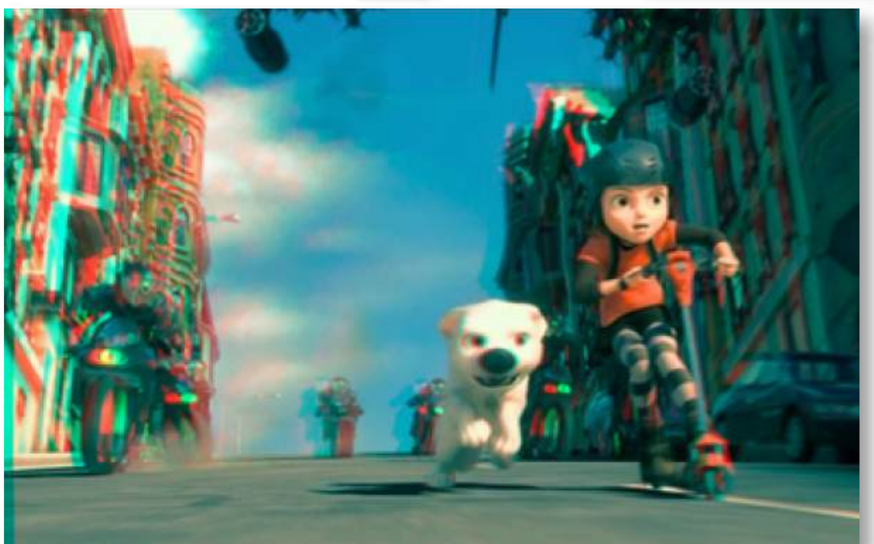


Disparity mapping function

"Nonlinear Disparity Mapping for Stereoscopic 3D" by Lang et al. 2010

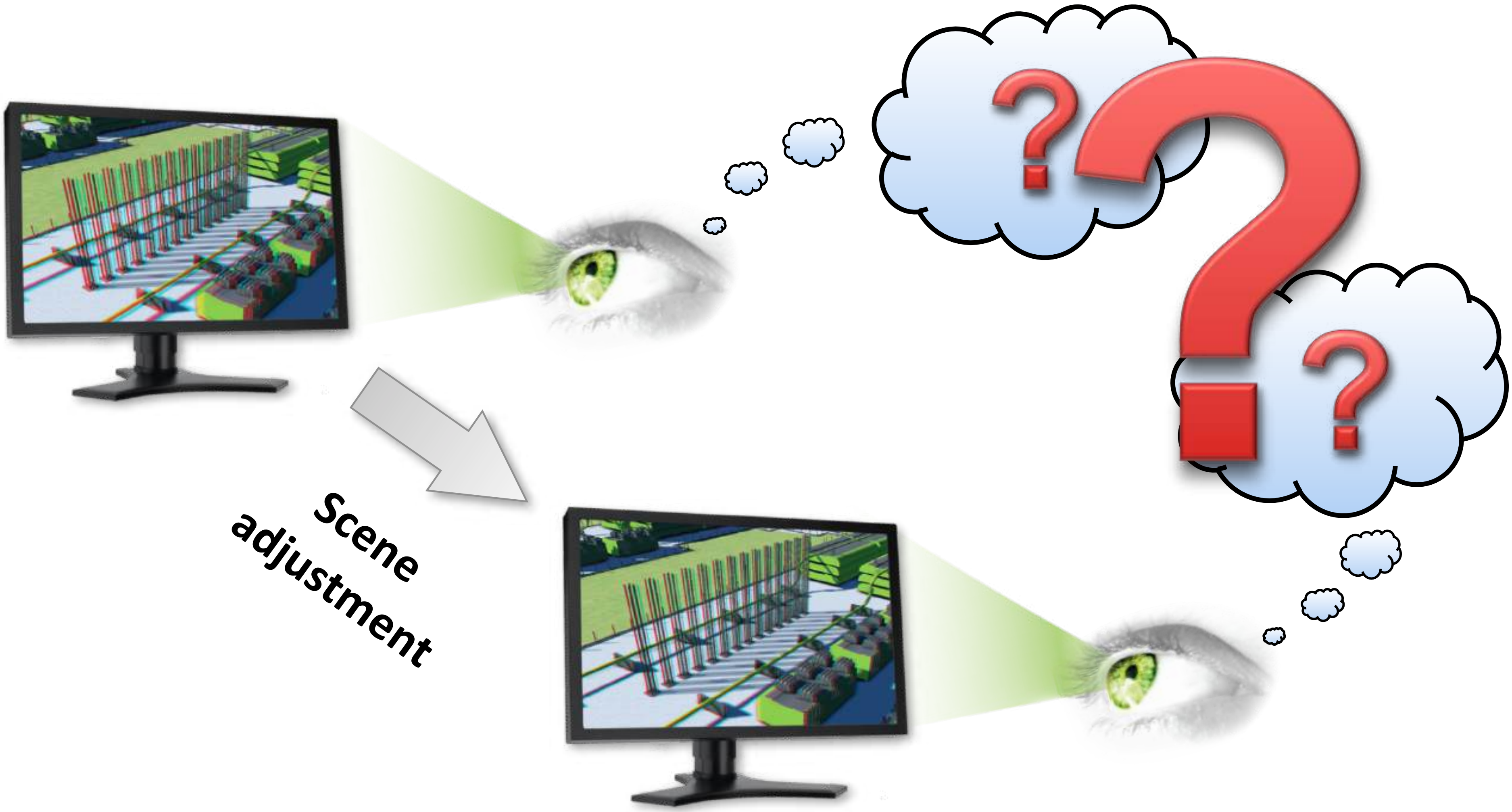
Saliency map



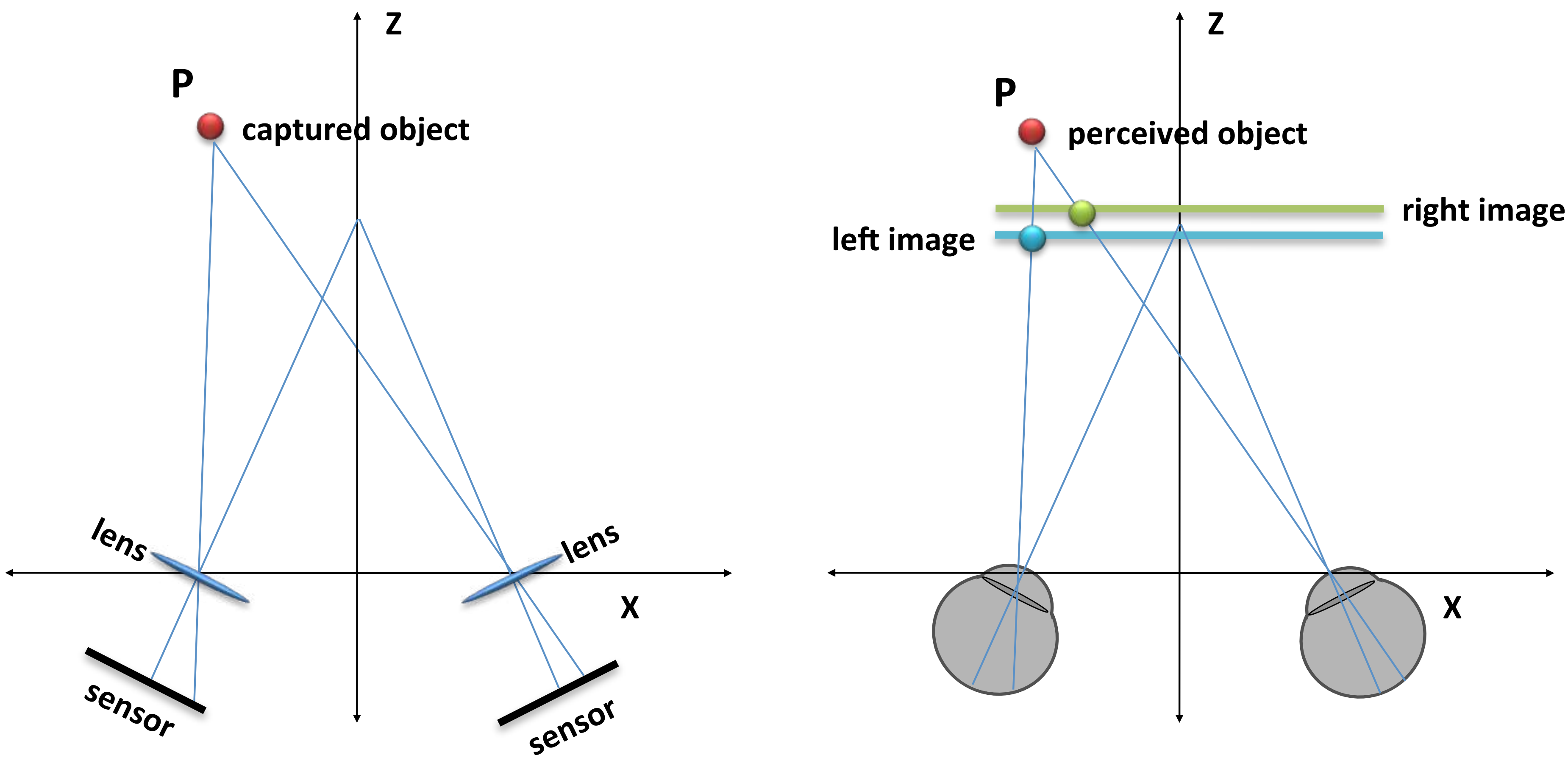


"Nonlinear Disparity Mapping for Stereoscopic 3D" by Lang et al. 2010

Scene manipulation



Misperception

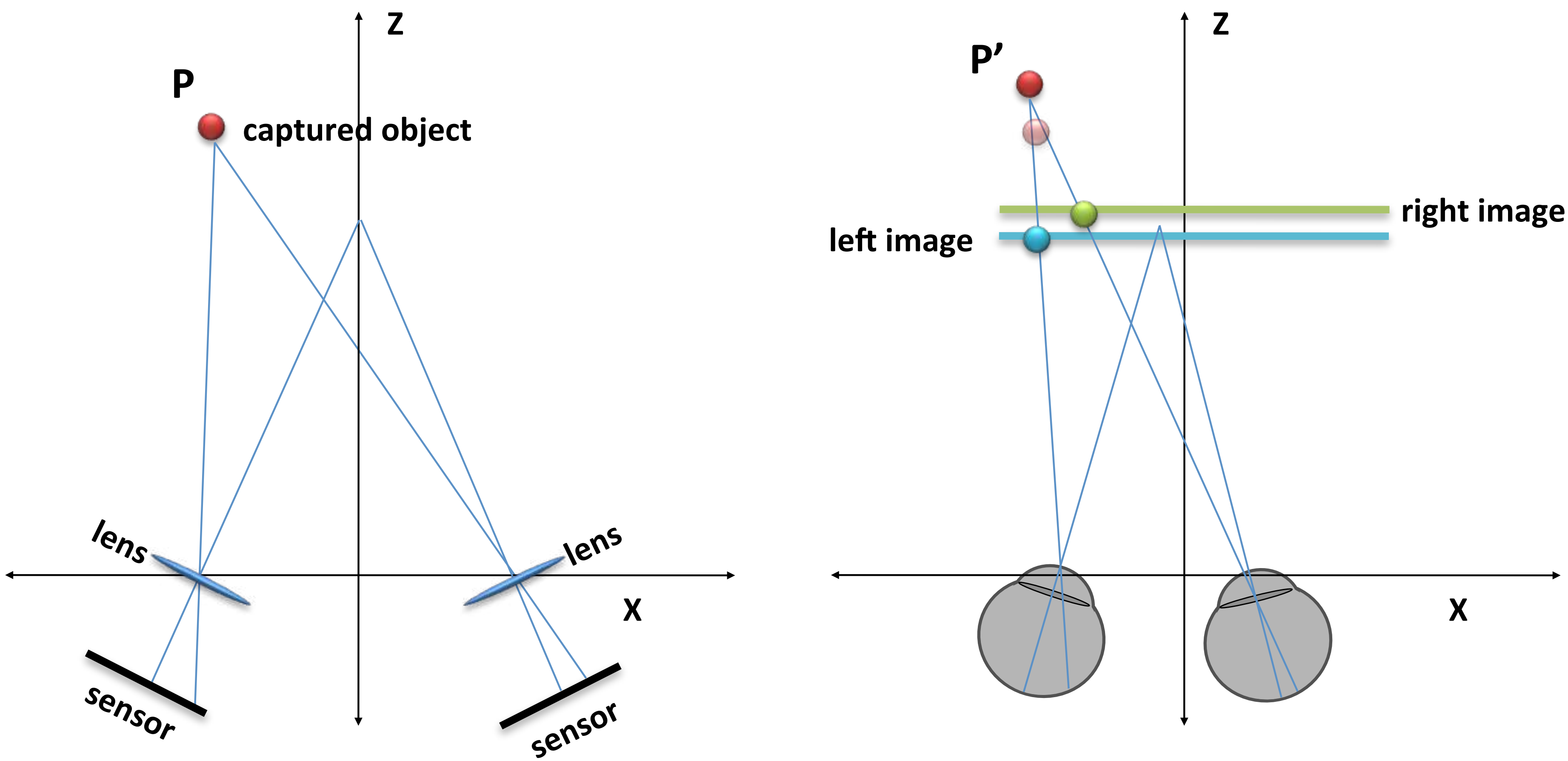


Parameters are the same

"Image Distortions in Stereoscopic Video Systems" by Woods et al. 1993

Misperception

SIGGRAPH ASIA 2011 HONG KONG

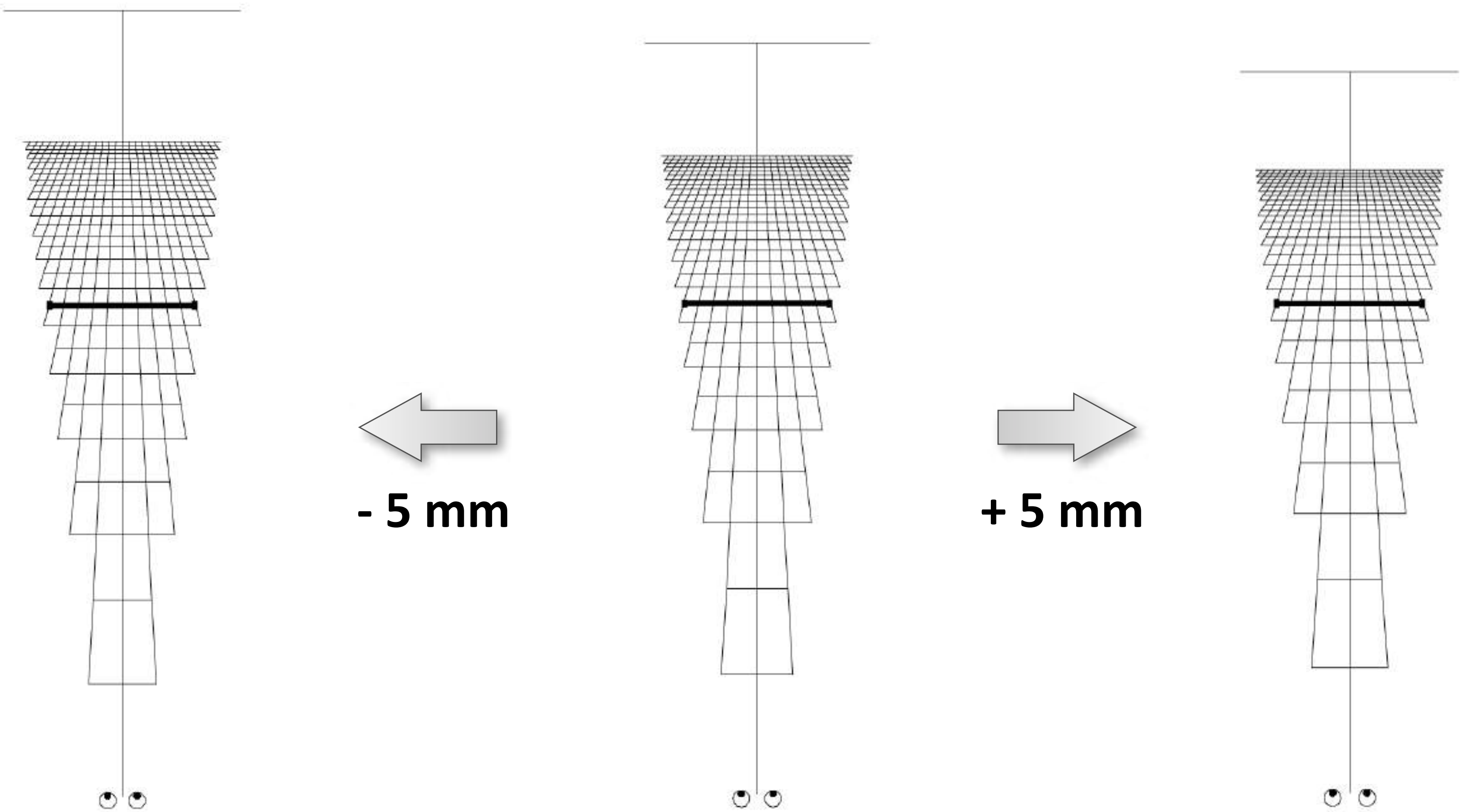


Eyes position and interocular distance changed

“Image Distortions in Stereoscopic Video Systems” by Woods et al. 1993

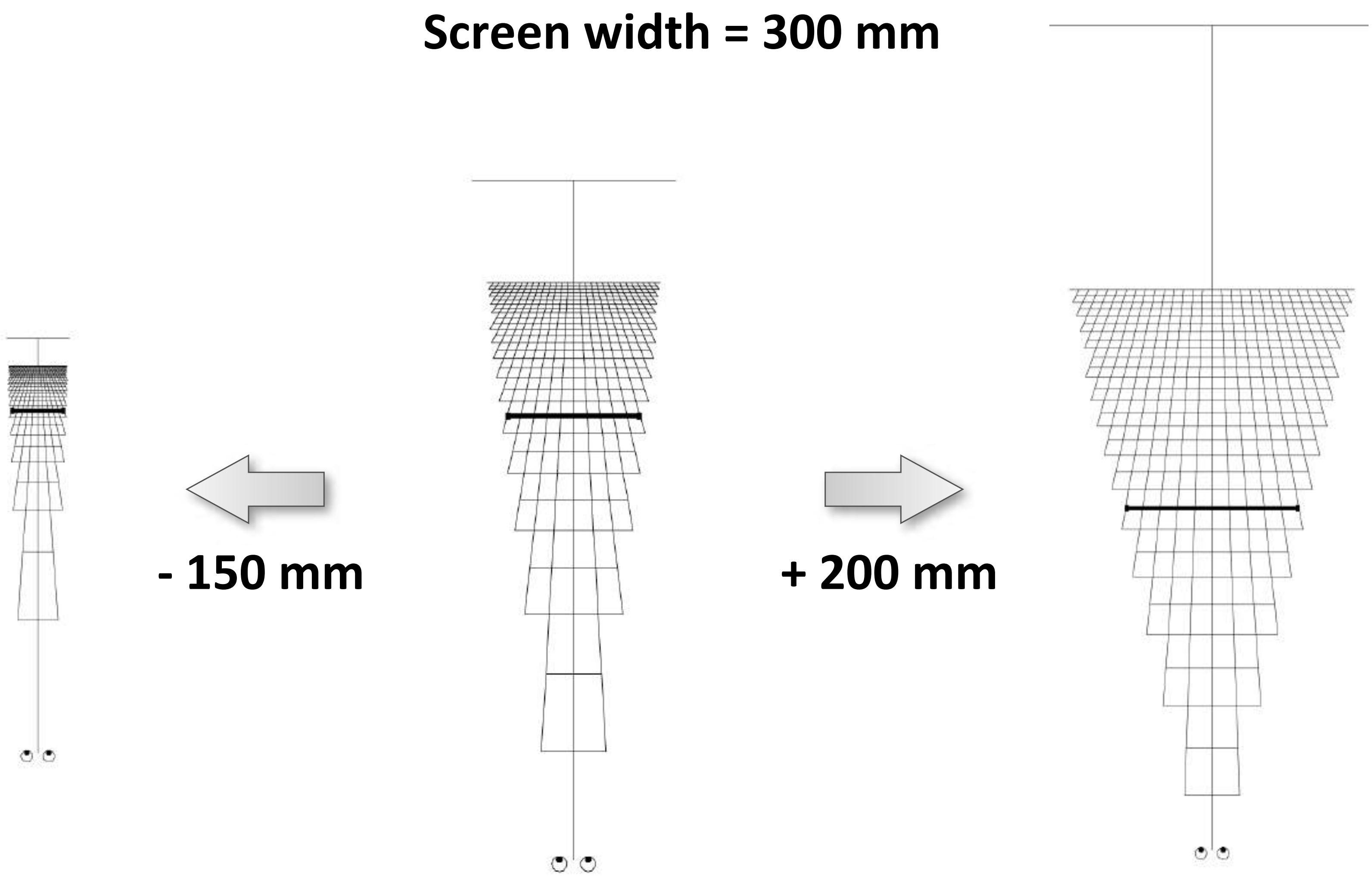
Misperception

Eye separation = 65 mm



"Image Distortions in Stereoscopic Video Systems" by Woods et al. 1993

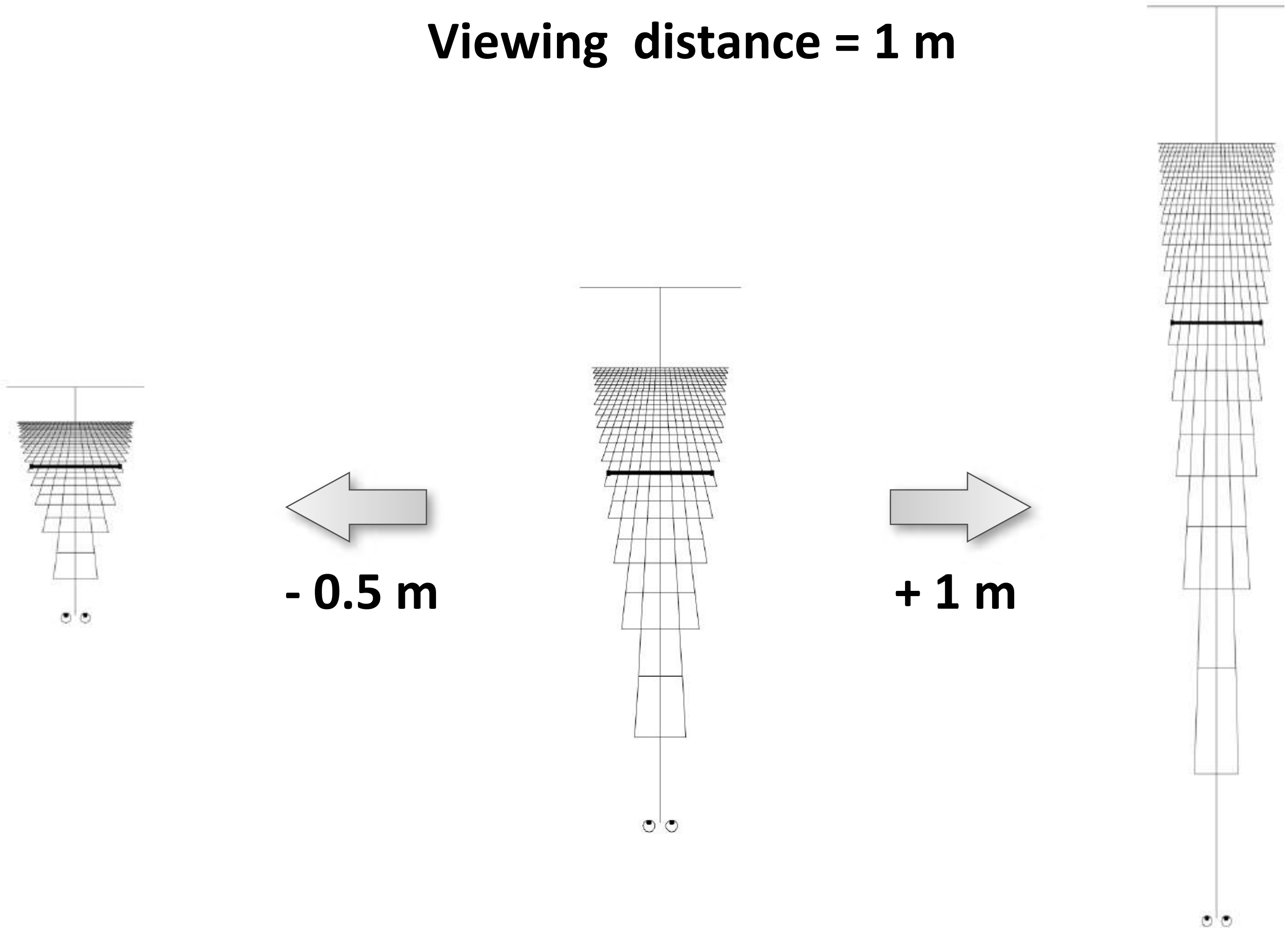
Misperception



"Image Distortions in Stereoscopic Video Systems" by Woods et al. 1993

Misperception

Viewing distance = 1 m



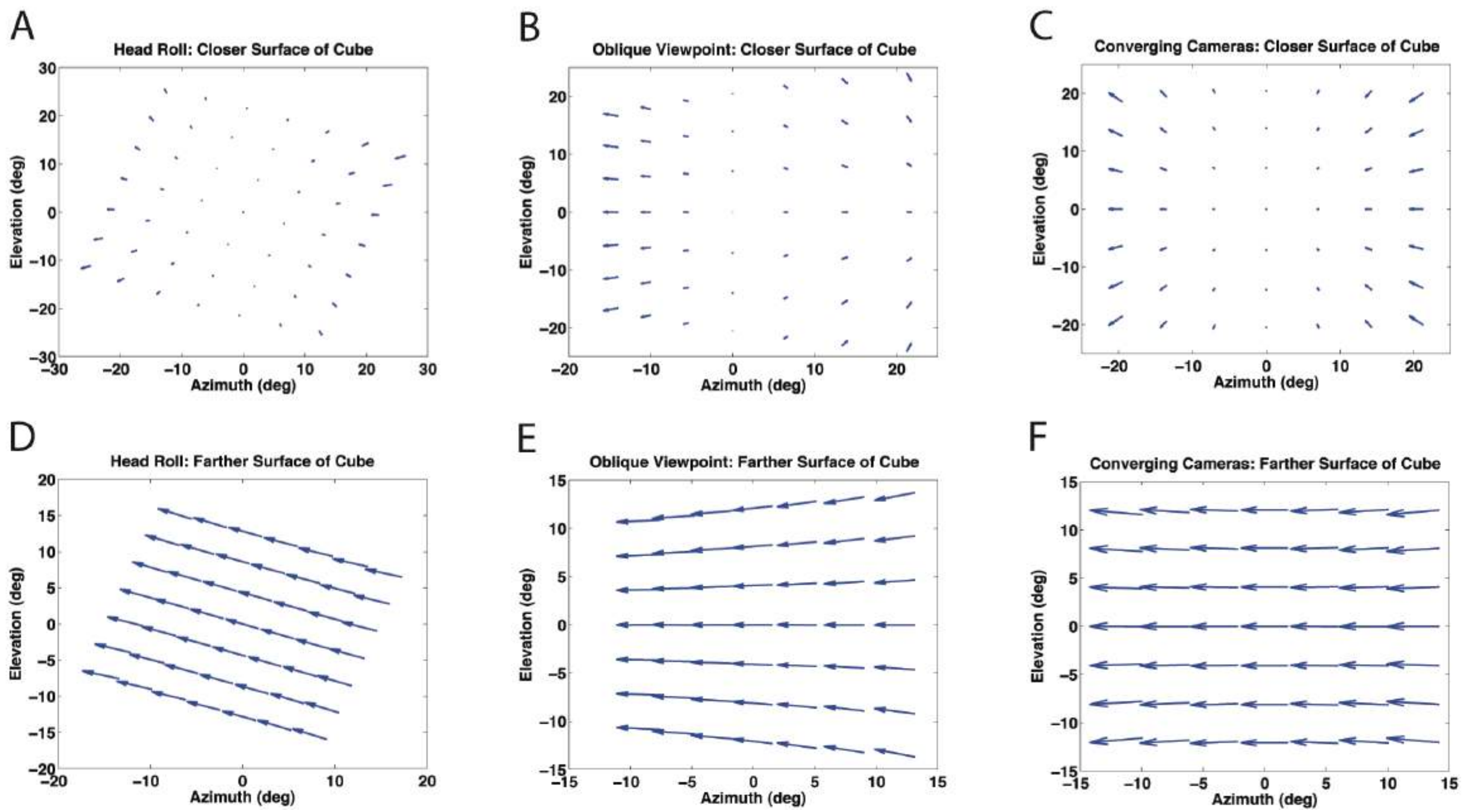
- 0.5 m

+ 1 m

"Image Distortions in Stereoscopic Video Systems" by Woods et al. 1993

Misperception

SIGGRAPHASIA2011 HONG KONG



"Misperceptions in Stereoscopic Displays: A Vision Science Perspective" by Held et al. 2008

3D image prediction

Depth perception

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
occlusion, size, shadows…

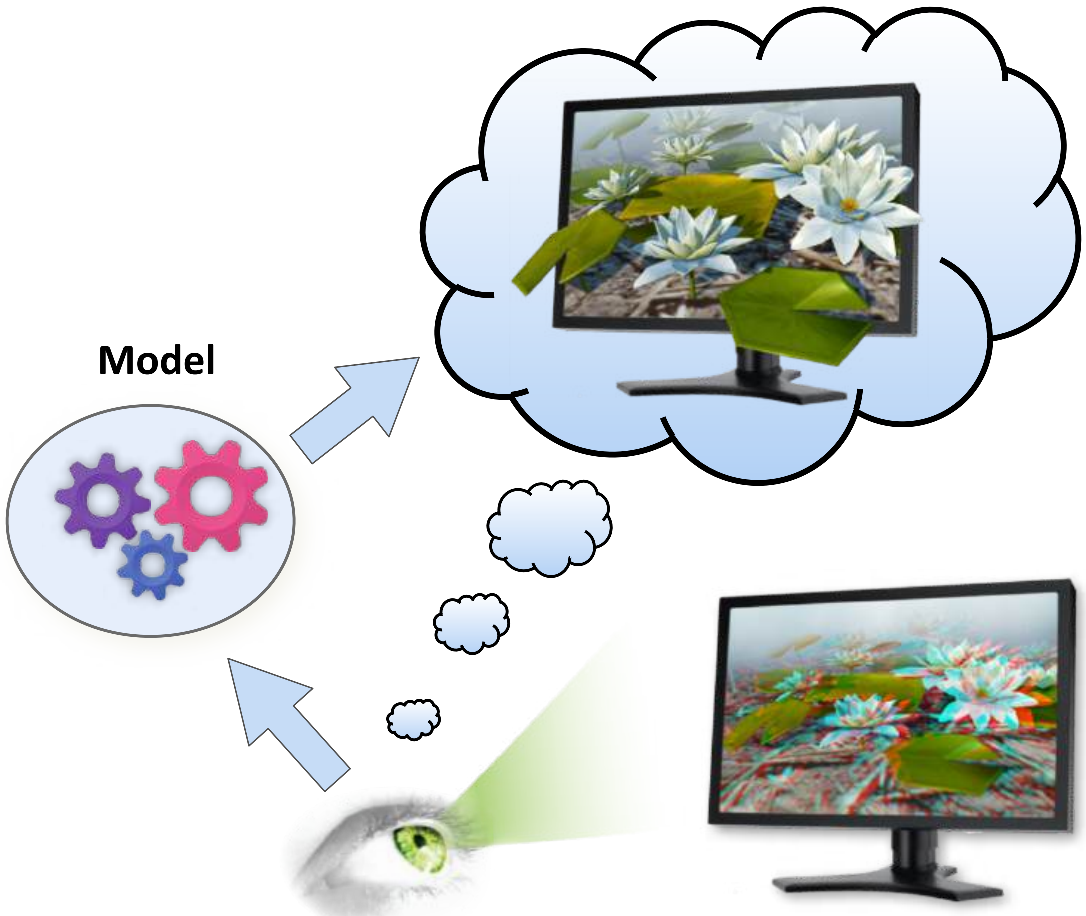
"A perceptual model for disparity" by Didyk et al. 2011

Depth perception

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

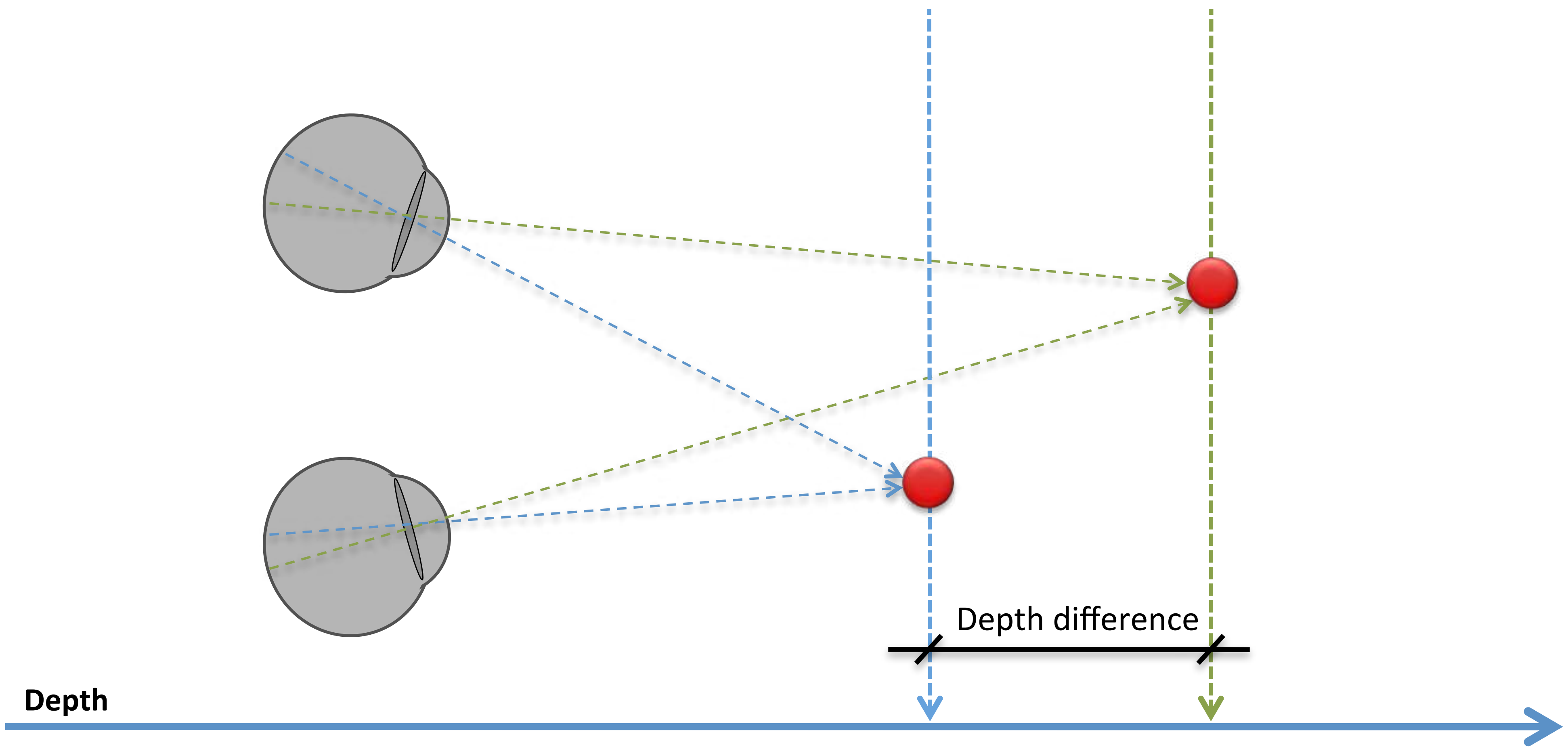
Pictorial depth cues:
occlusion, size, shadows…



"A perceptual model for disparity" by Didyk et al. 2011

Disparity perception

SIGGRAPHASIA2011 HONG KONG

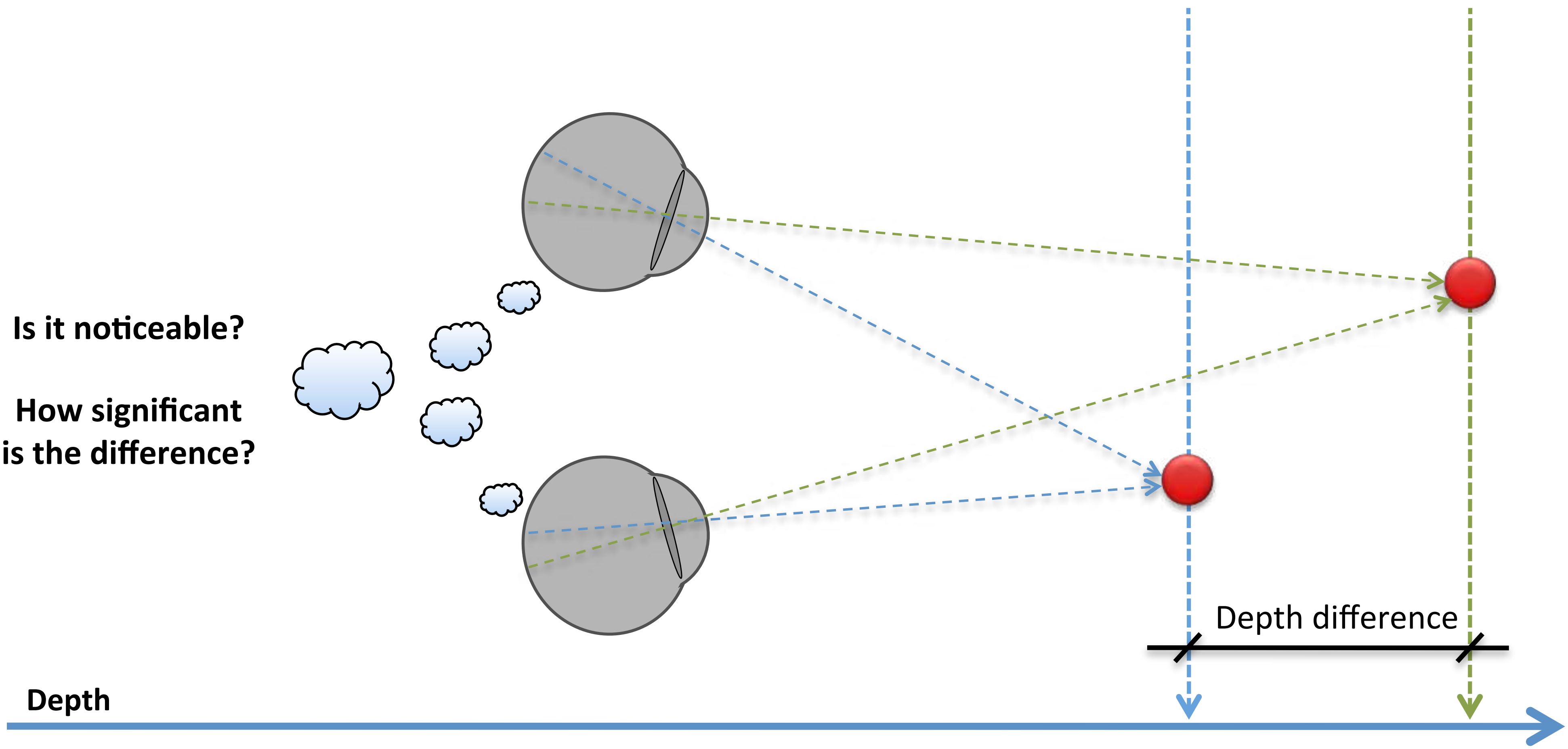


Disparity perception

SIGGRAPHASIA2011 HONG KONG

Is it noticeable?

How significant is the difference?

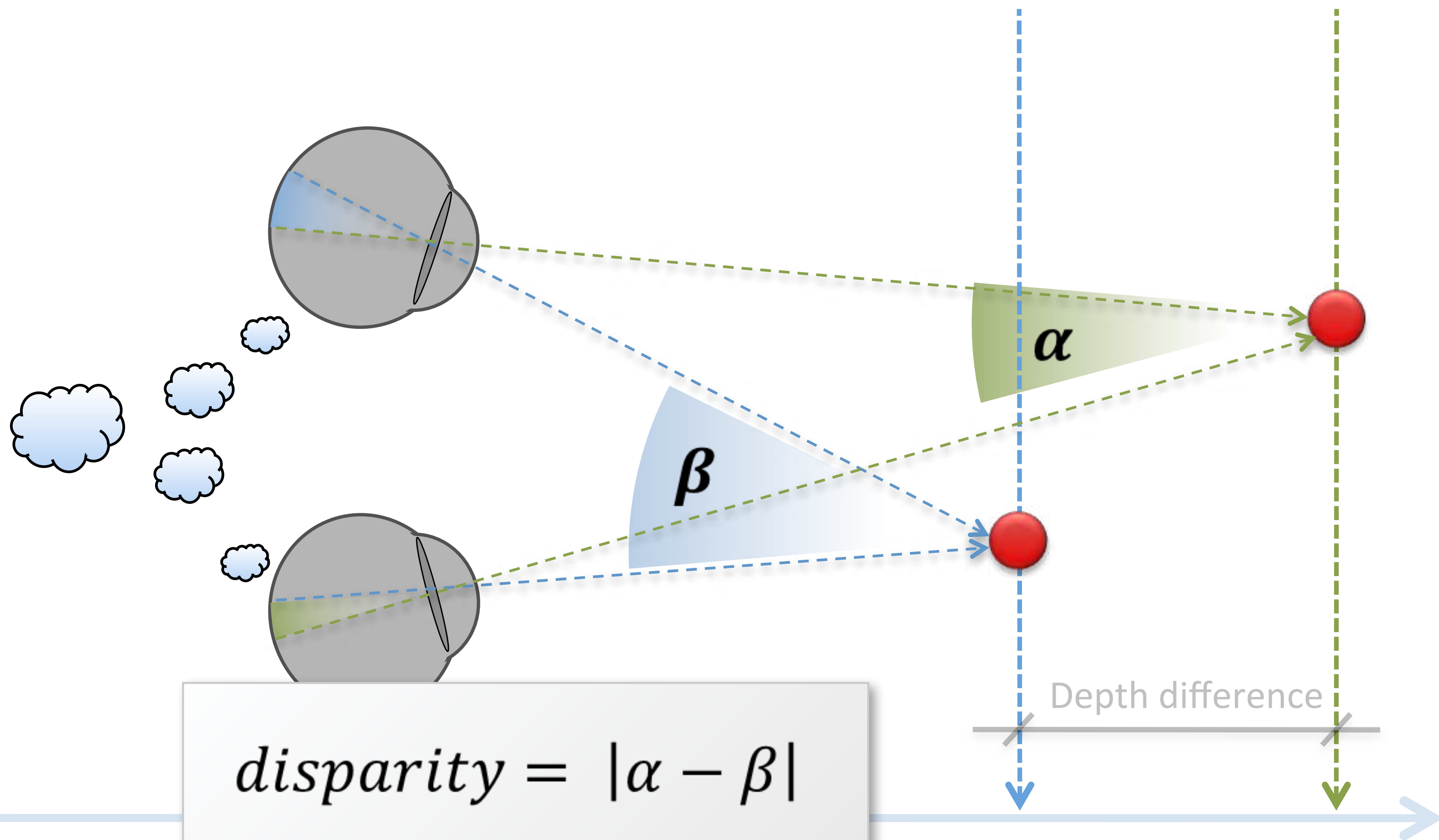


Disparity perception

SIGGRAPH ASIA 2011 HONG KONG

Is it noticeable?

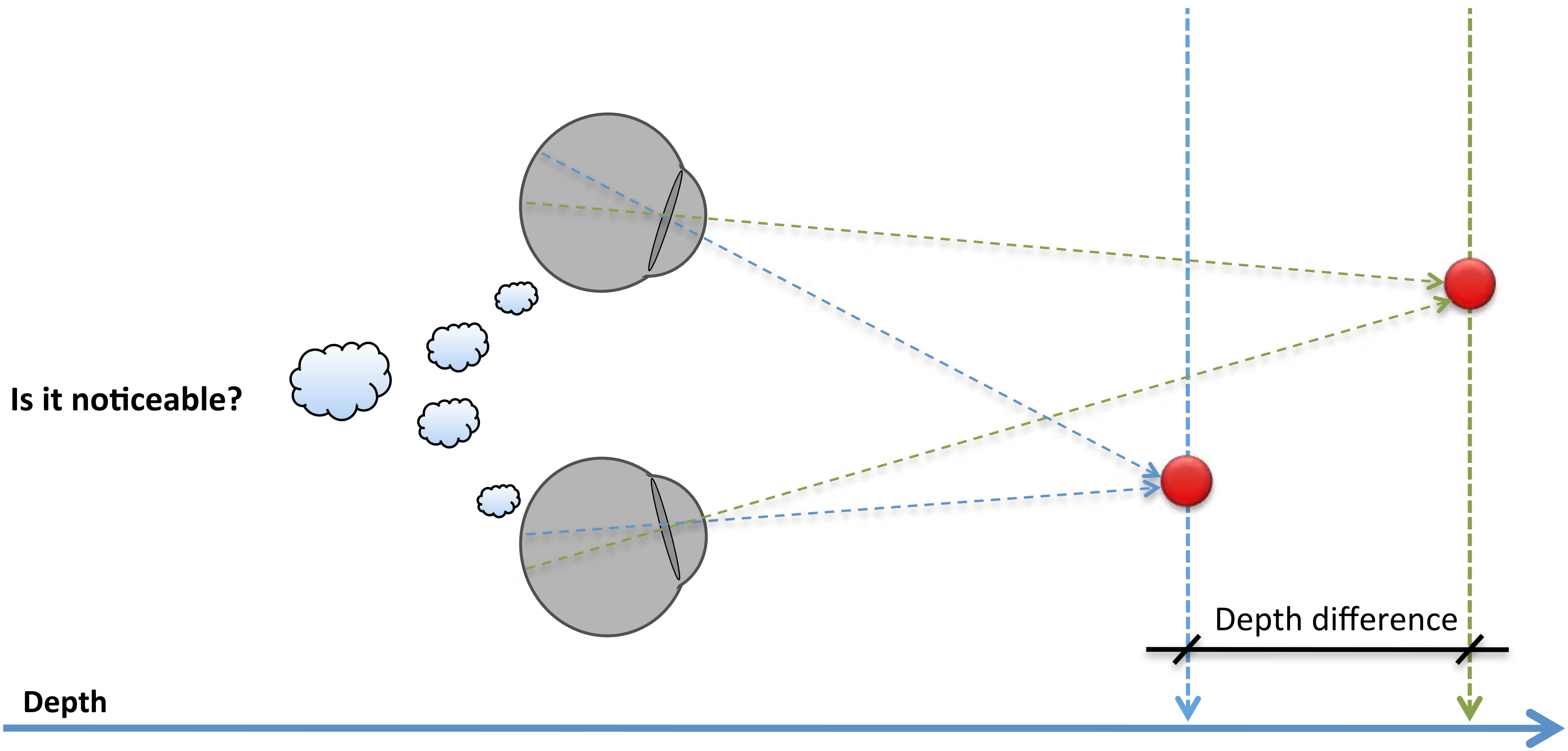
How significant is the difference?



$$disparity = |\alpha - \beta|$$

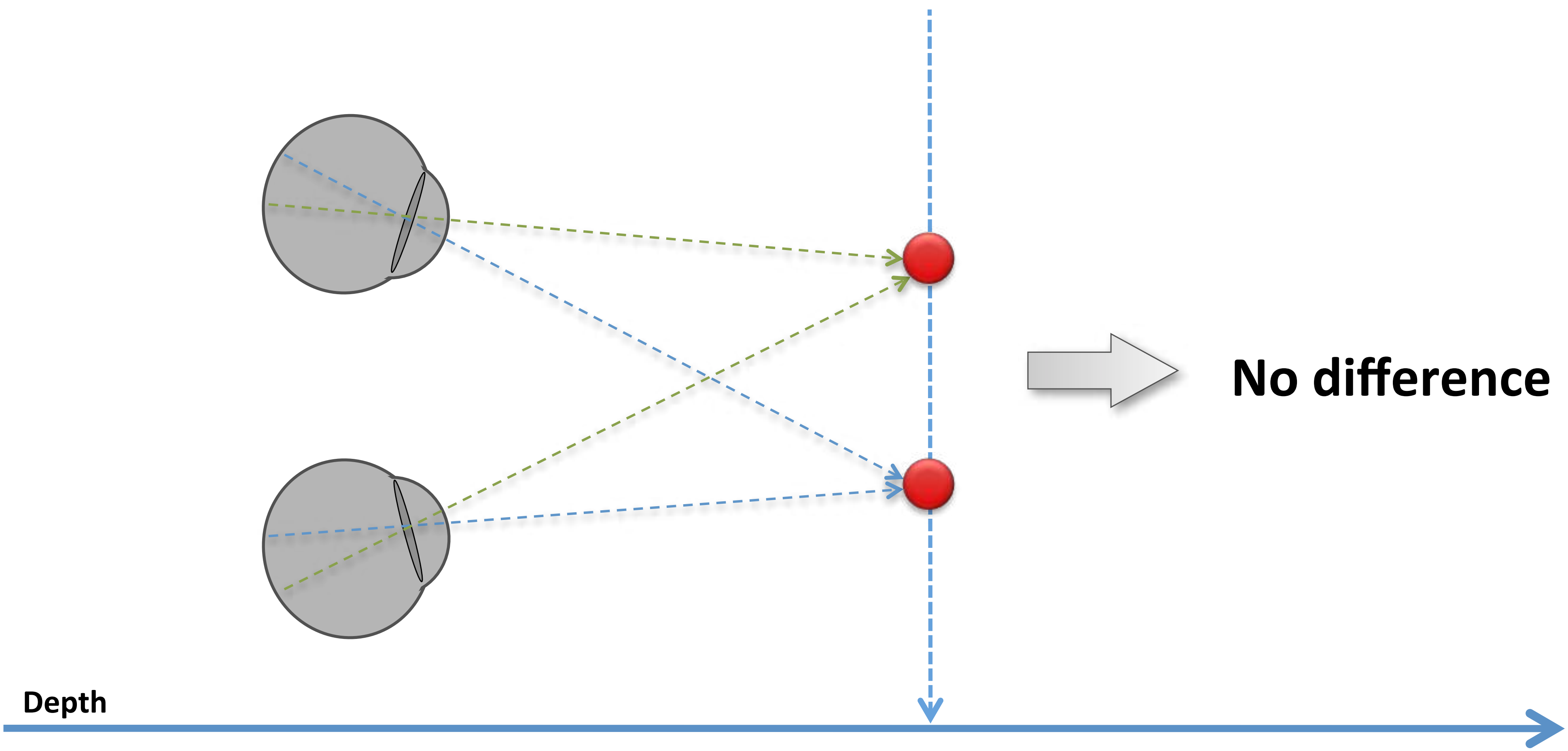
Disparity perception

SIGGRAPHASIA2011 HONG KONG

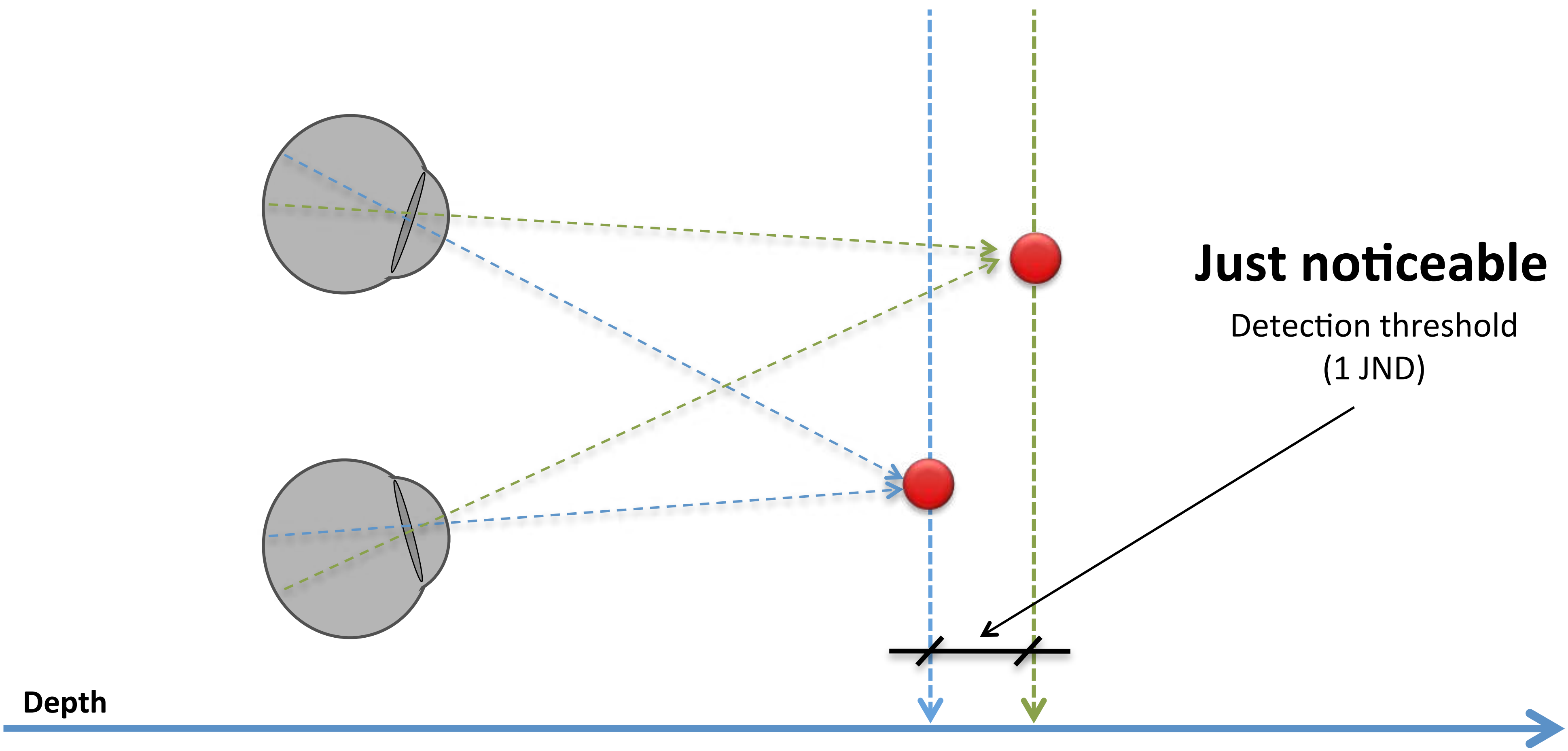


One just-noticeable difference

SIGGRAPHASIA2011 HONG KONG



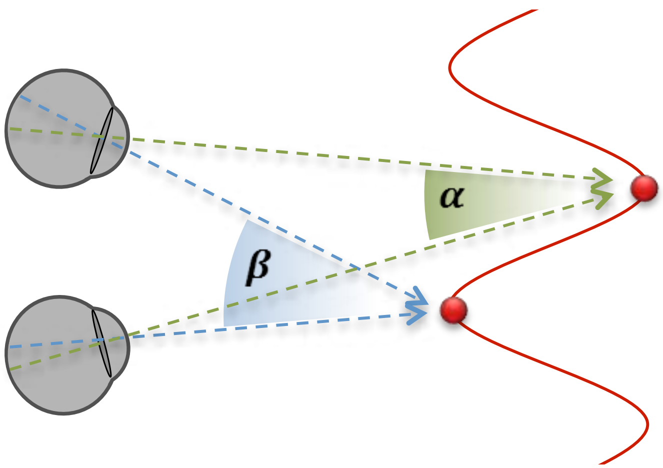
One just-noticeable difference



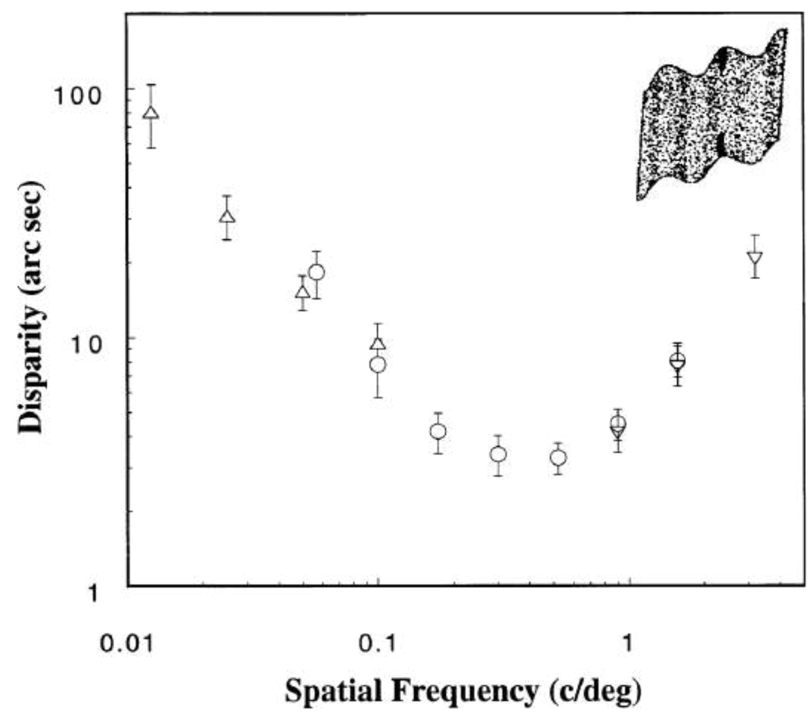
How big is the detection threshold?

SIGGRAPHASIA2011 HONG KONG

For sinusoidal depth corrugation



$$disparity = |\alpha - \beta|$$



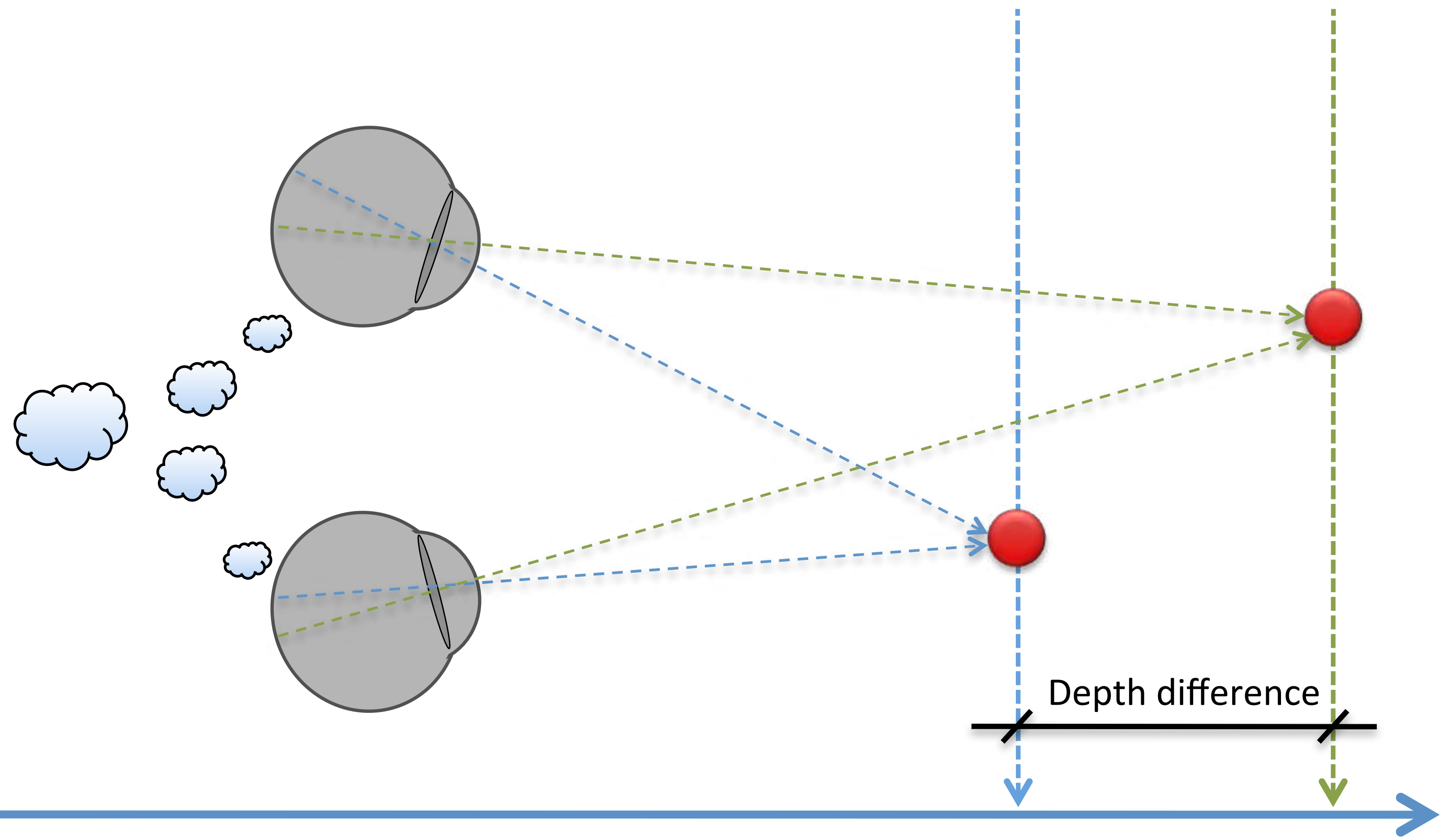
"Sensitivity to horizontal and vertical corrugations defined by binocular disparity."
by Bradshaw et al. 1999

Disparity perception

SIGGRAPHASIA2011 HONG KONG

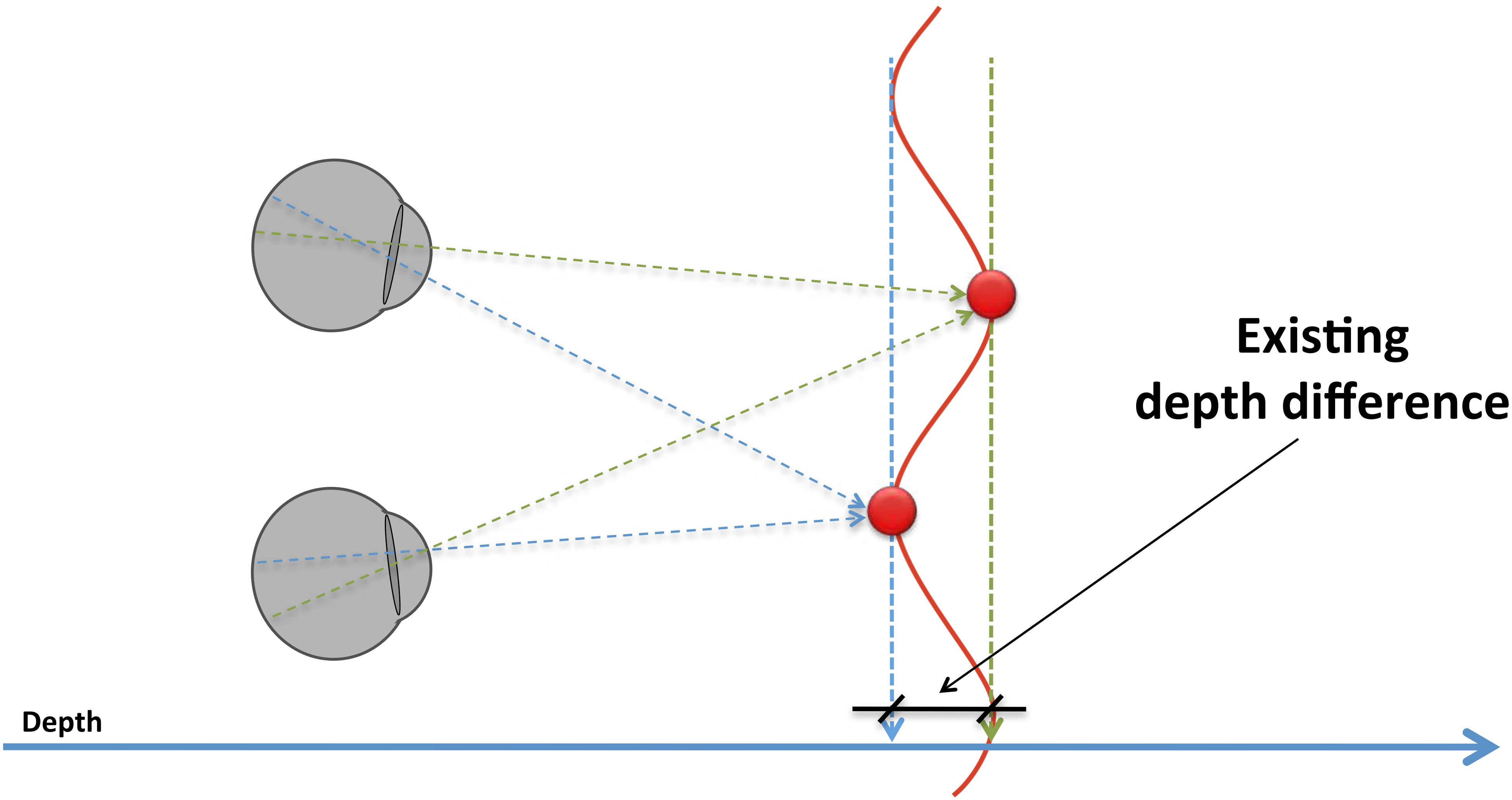
Is it noticeable?

How significant is the difference?



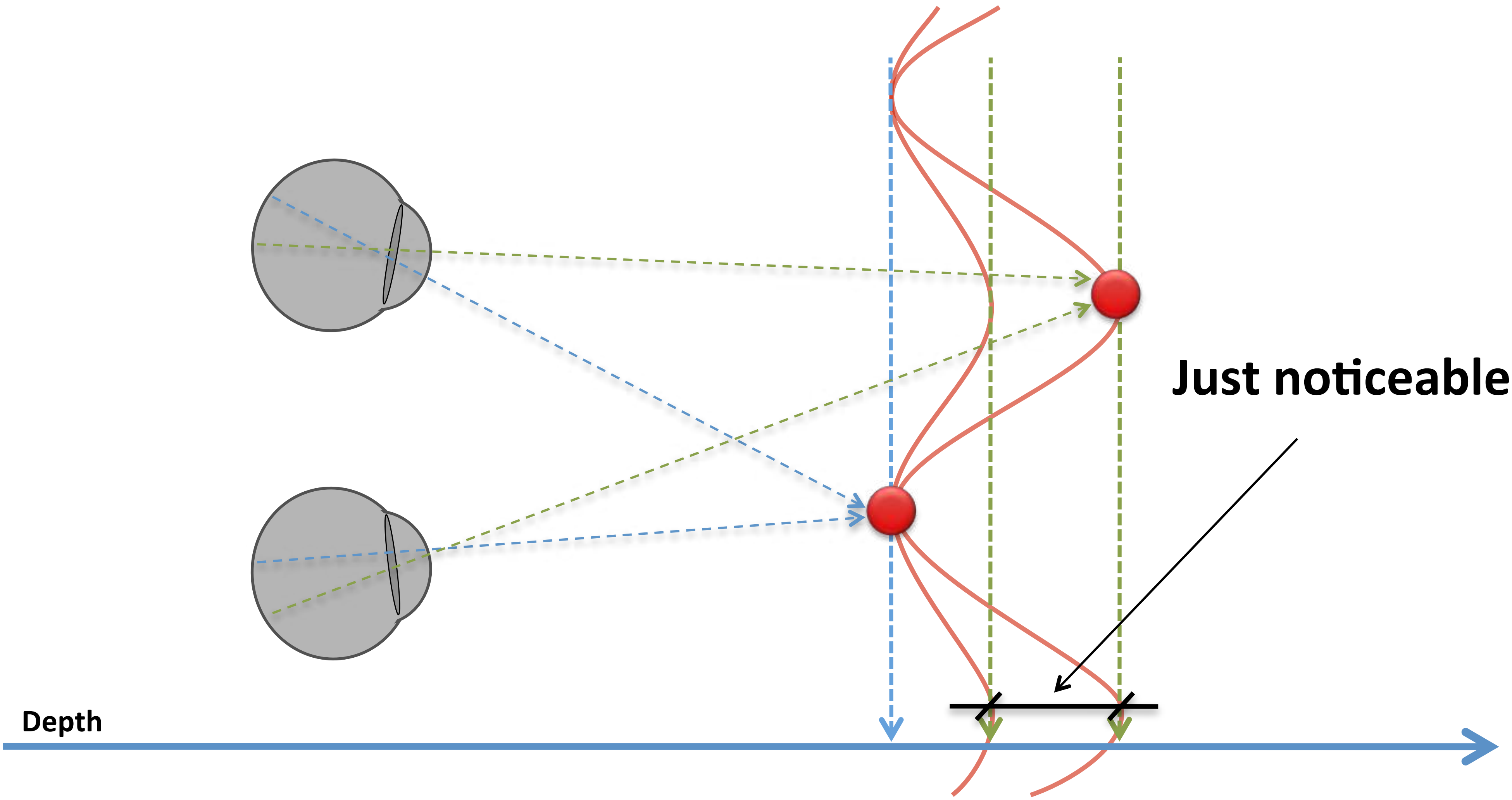
Discrimination threshold

SIGGRAPH ASIA 2011 HONG KONG

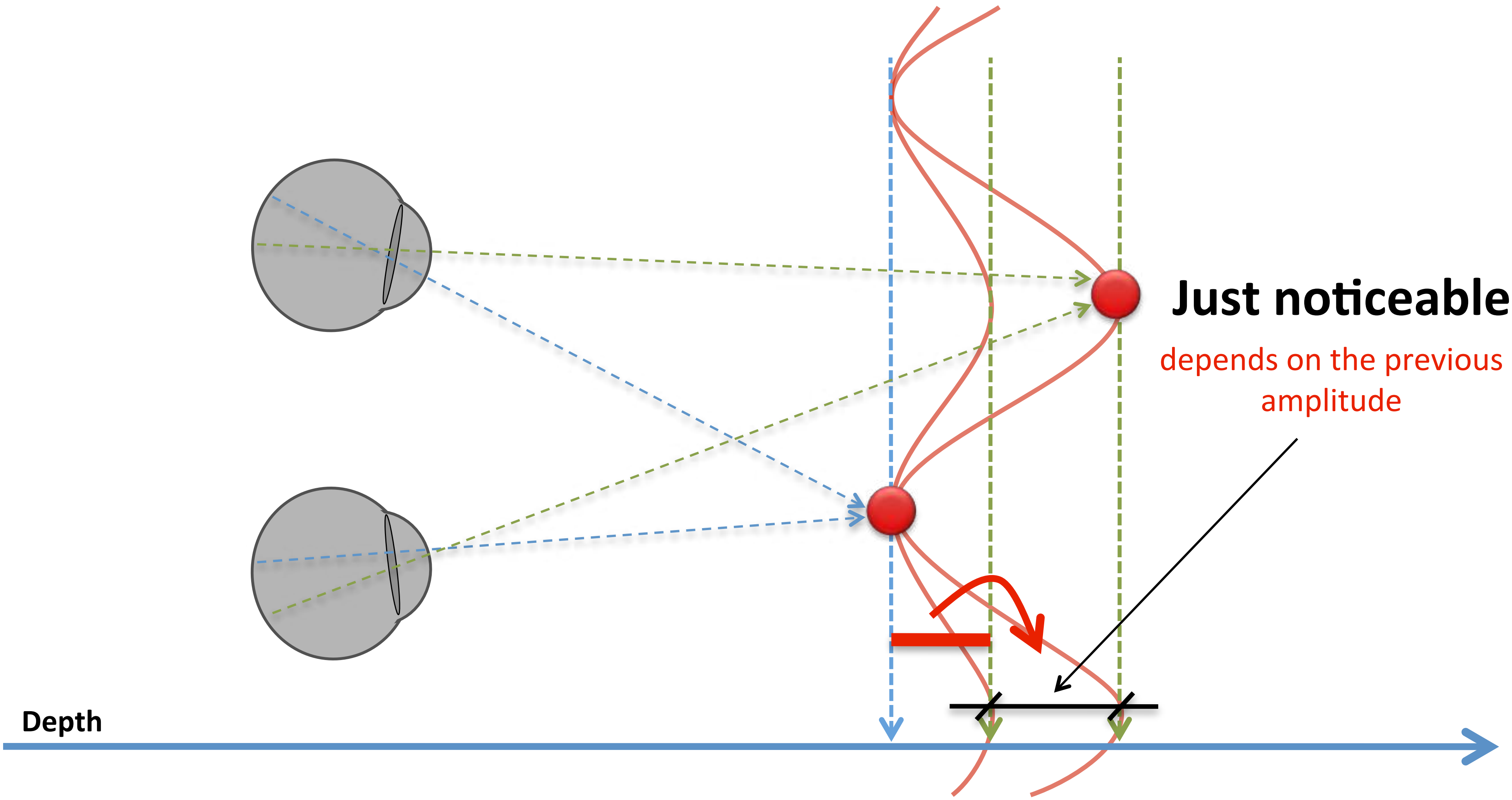


Discrimination threshold

SIGGRAPHASIA2011 HONG KONG



Discrimination threshold



Disparity perception

SIGGRAPHASIA2011 HONG KONG

Sensitivity to depth changes depends on:

- Spatial frequency of disparity corrugation
- Existing disparity (sinusoid amplitude)

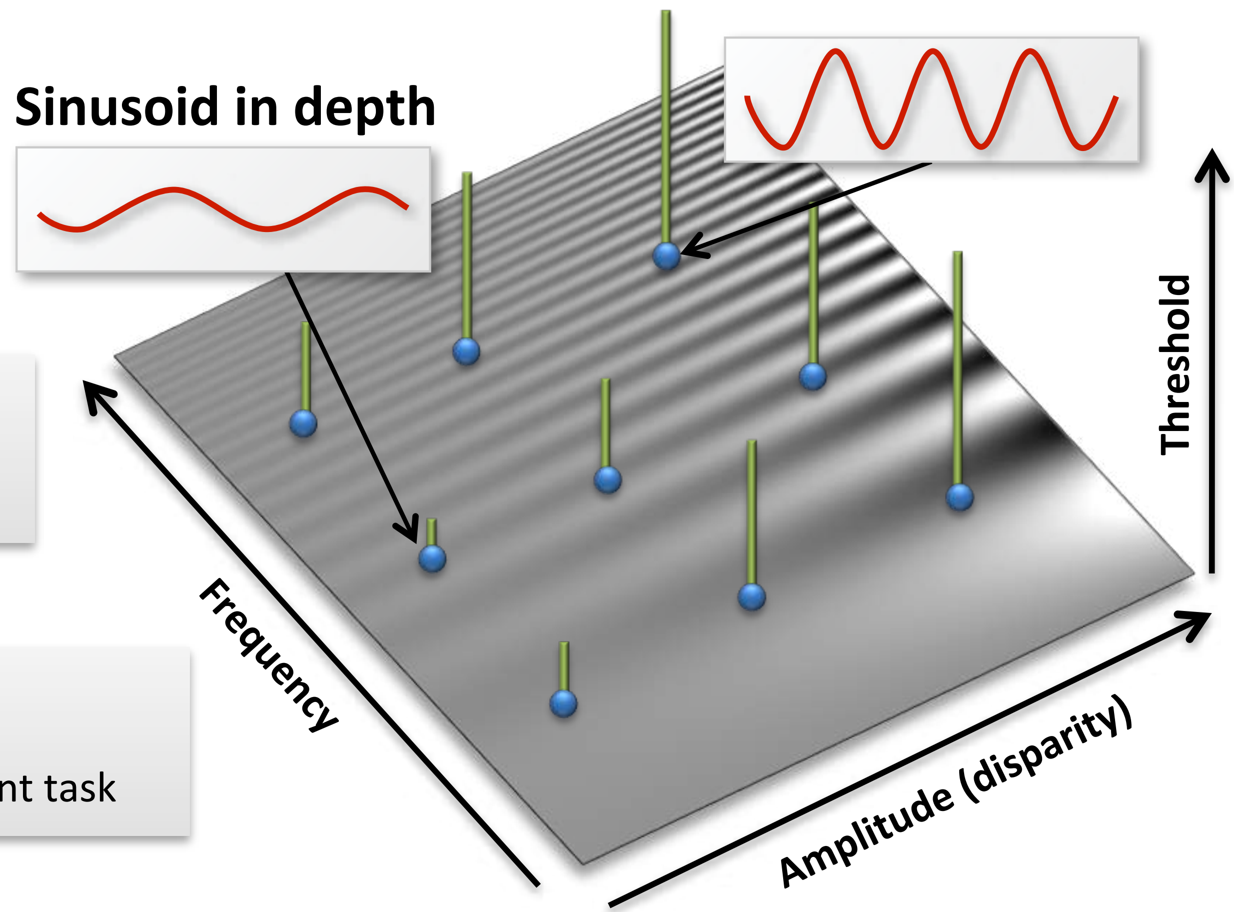
Measurements

Parameter space:

1. Sample the space

3. Measure thresholds

- Experiment with adjustment task

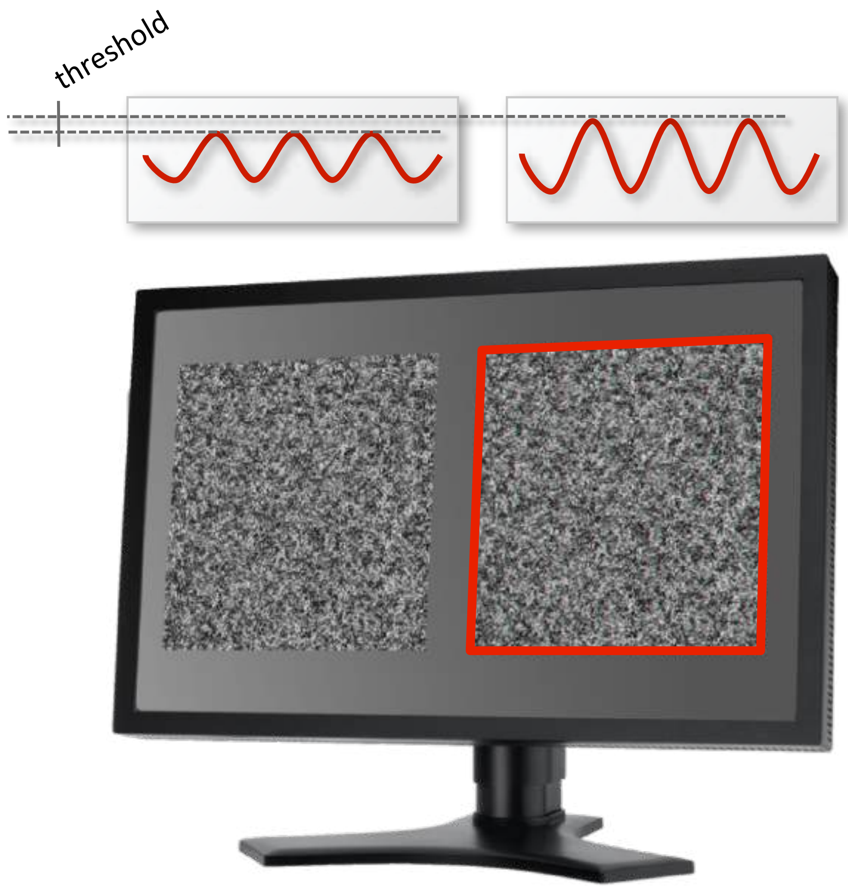


"A perceptual model for disparity" by Didyk et al. 2011

Measurements

Thresholds measurement:

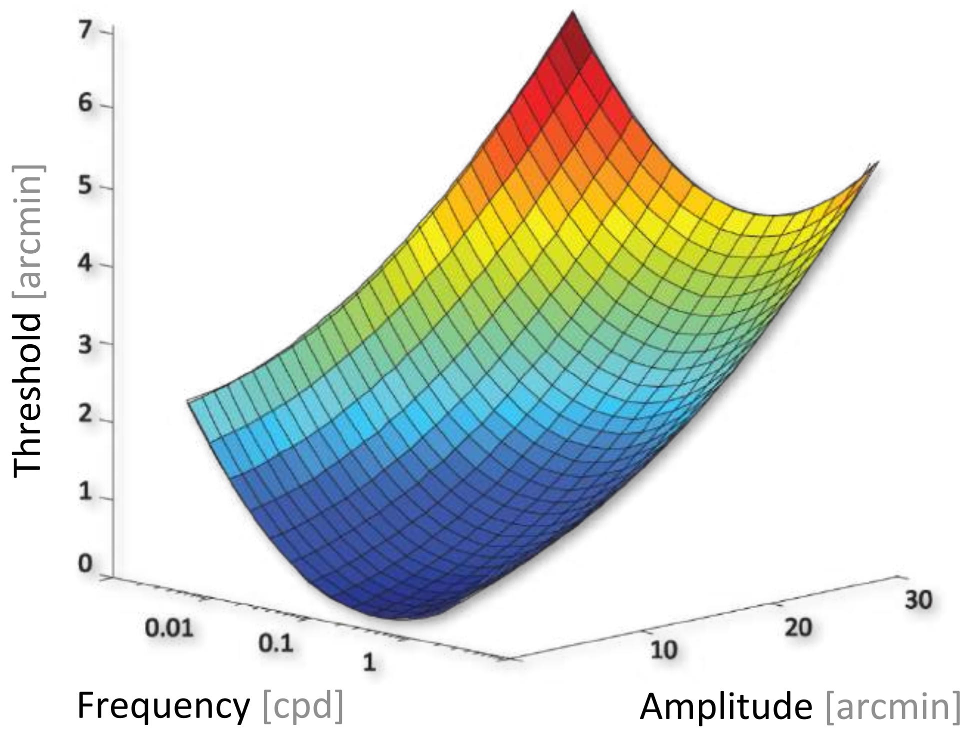
- Two sinusoidal corrugations
- Which has more depth? (left/right)
- Amplitude adjustment (PEST with 2AFC)
- 12 participants → 300+ samples



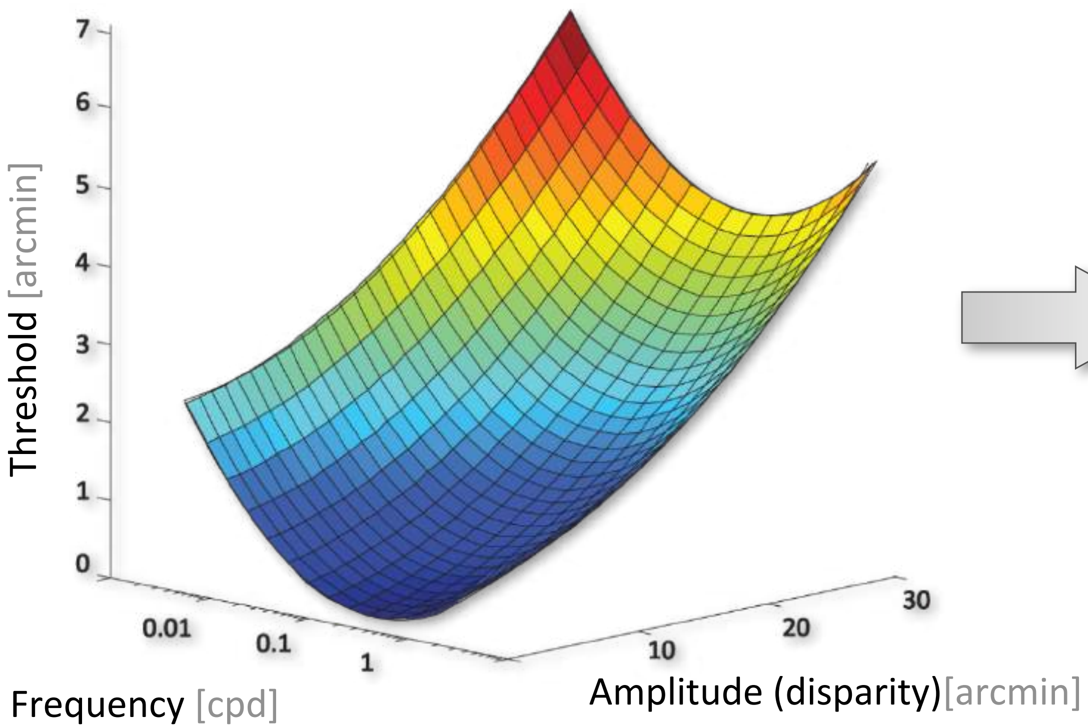
"A perceptual model for disparity" by Didyk et al. 2011

Model

3. Fit analytic function

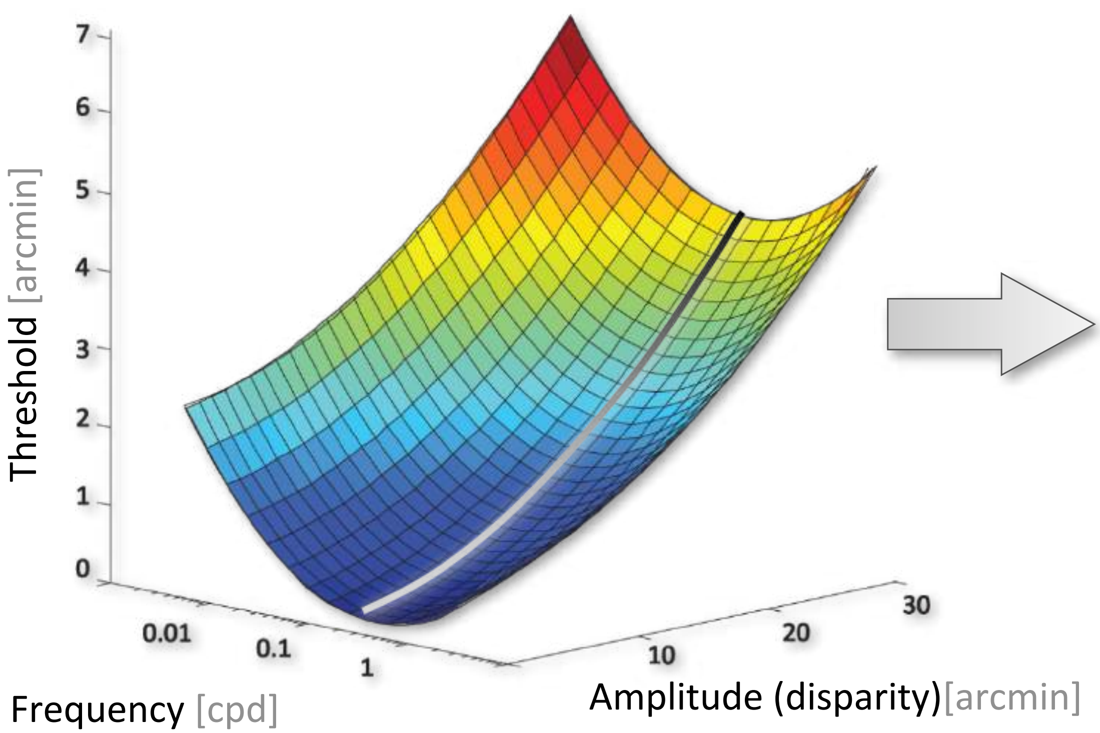


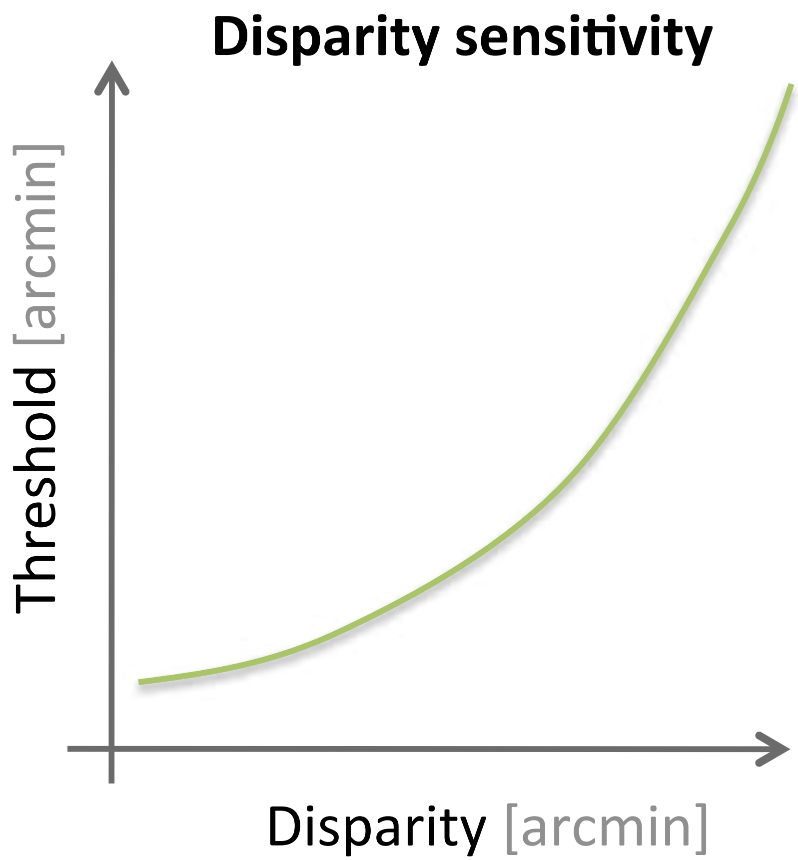
The HVS response



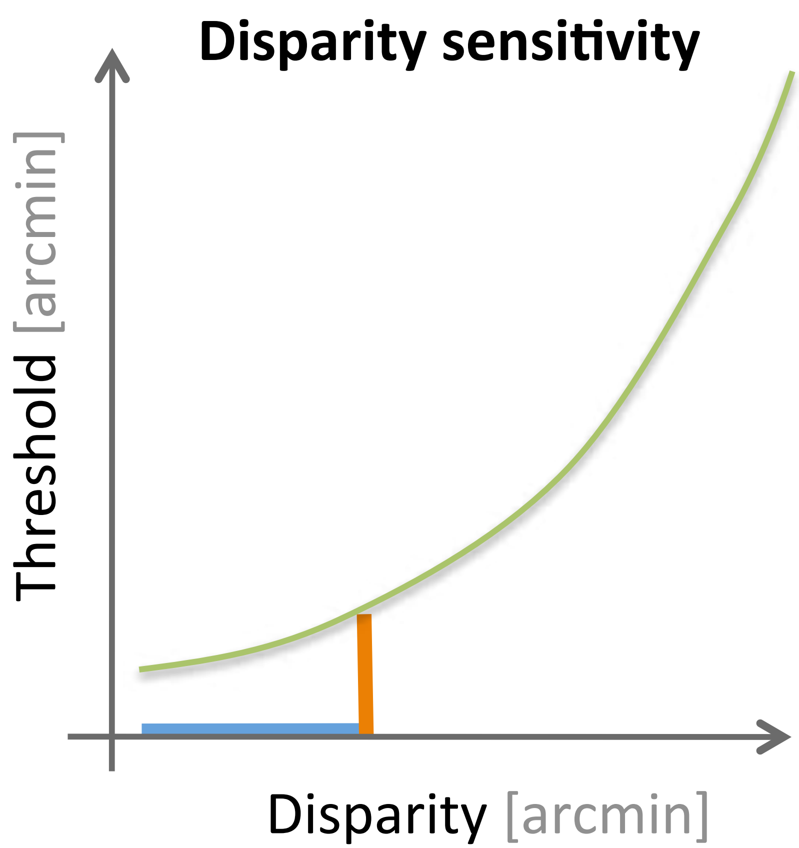
The HVS response?

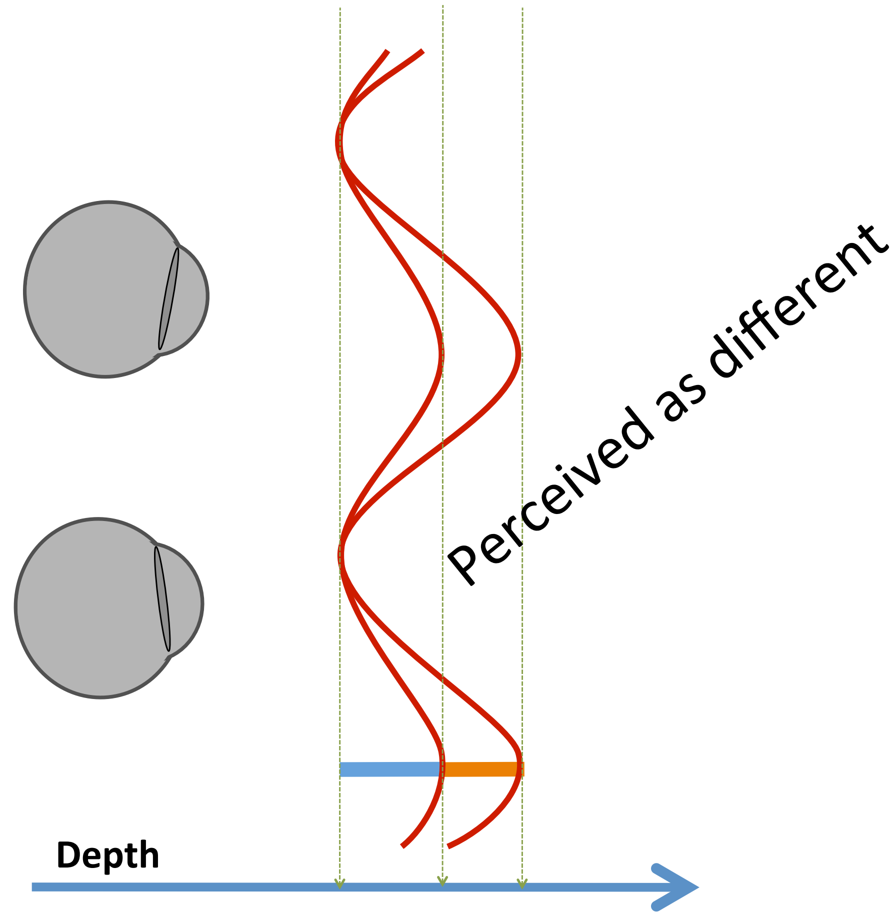
The HVS response





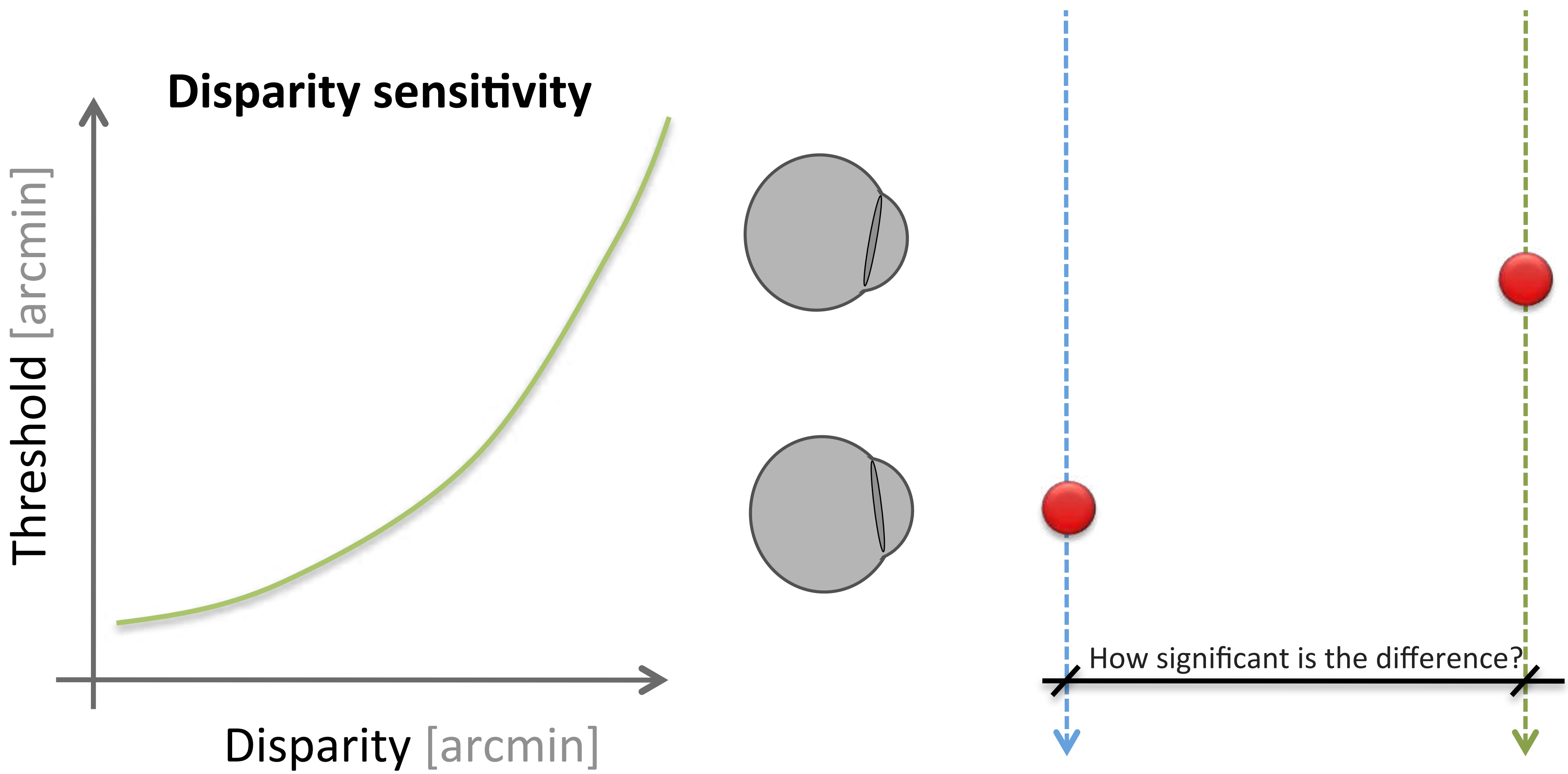
The HVS response





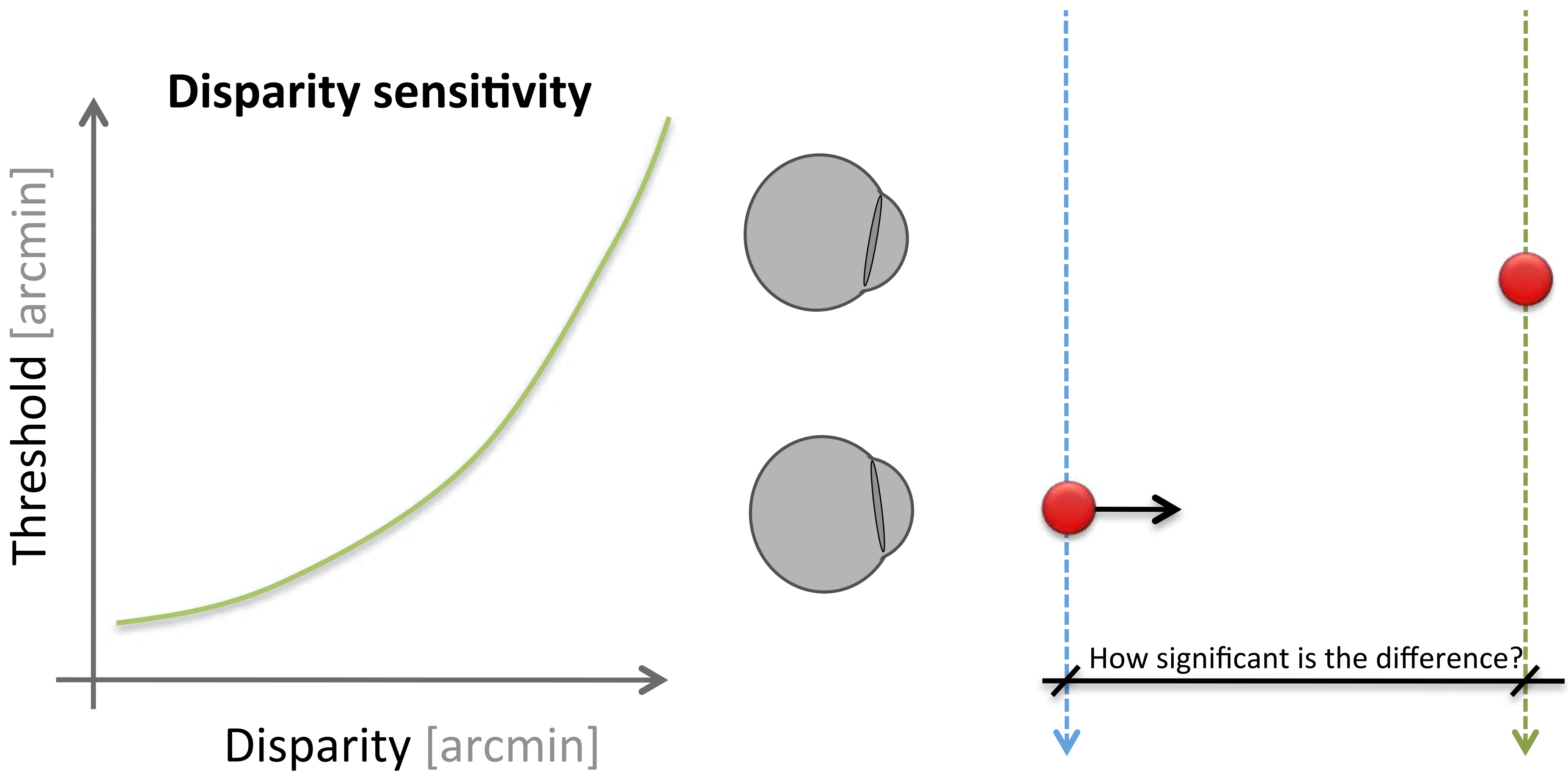
The HVS response

SIGGRAPHASIA2011 HONG KONG

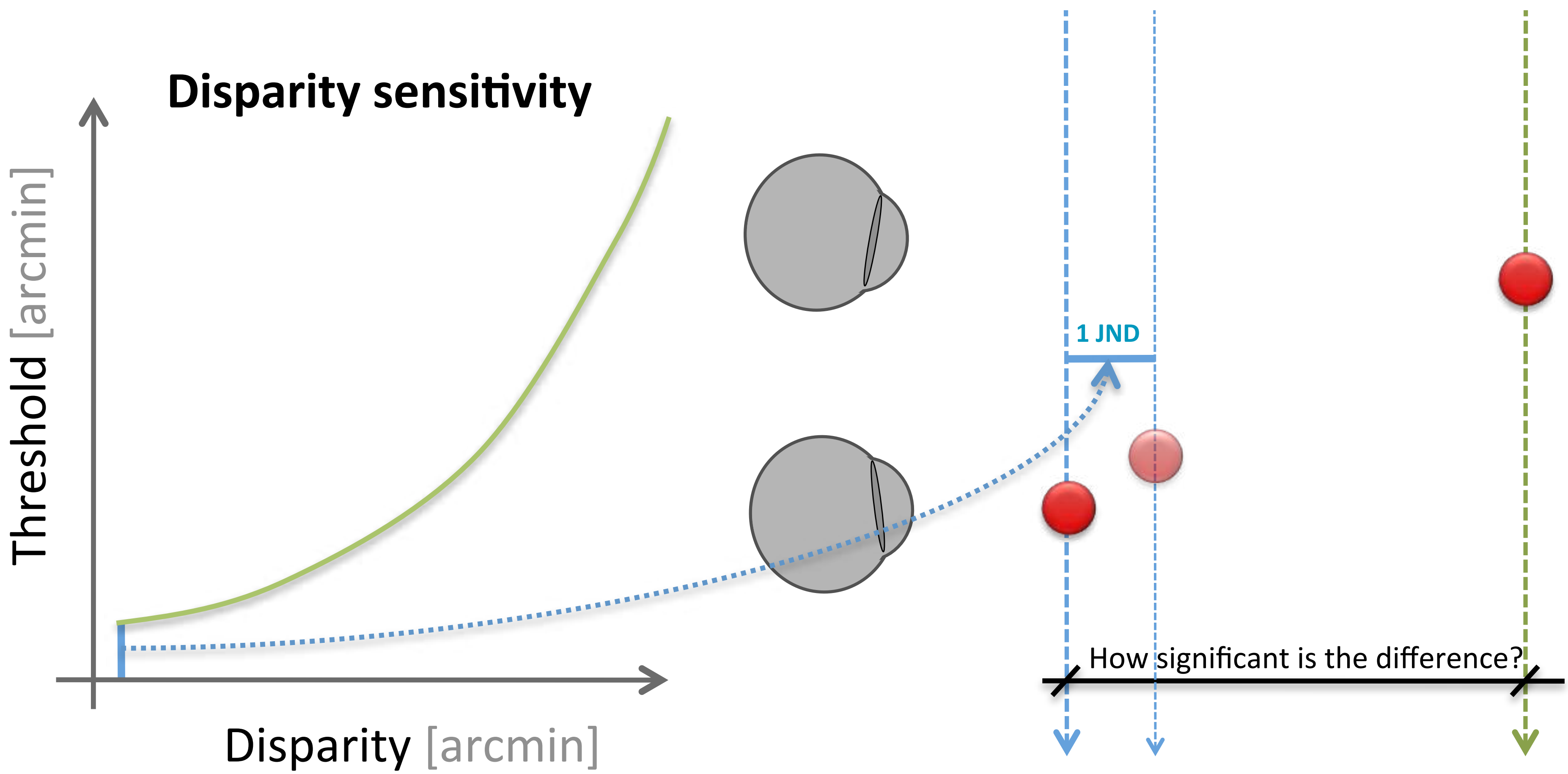


The HVS response

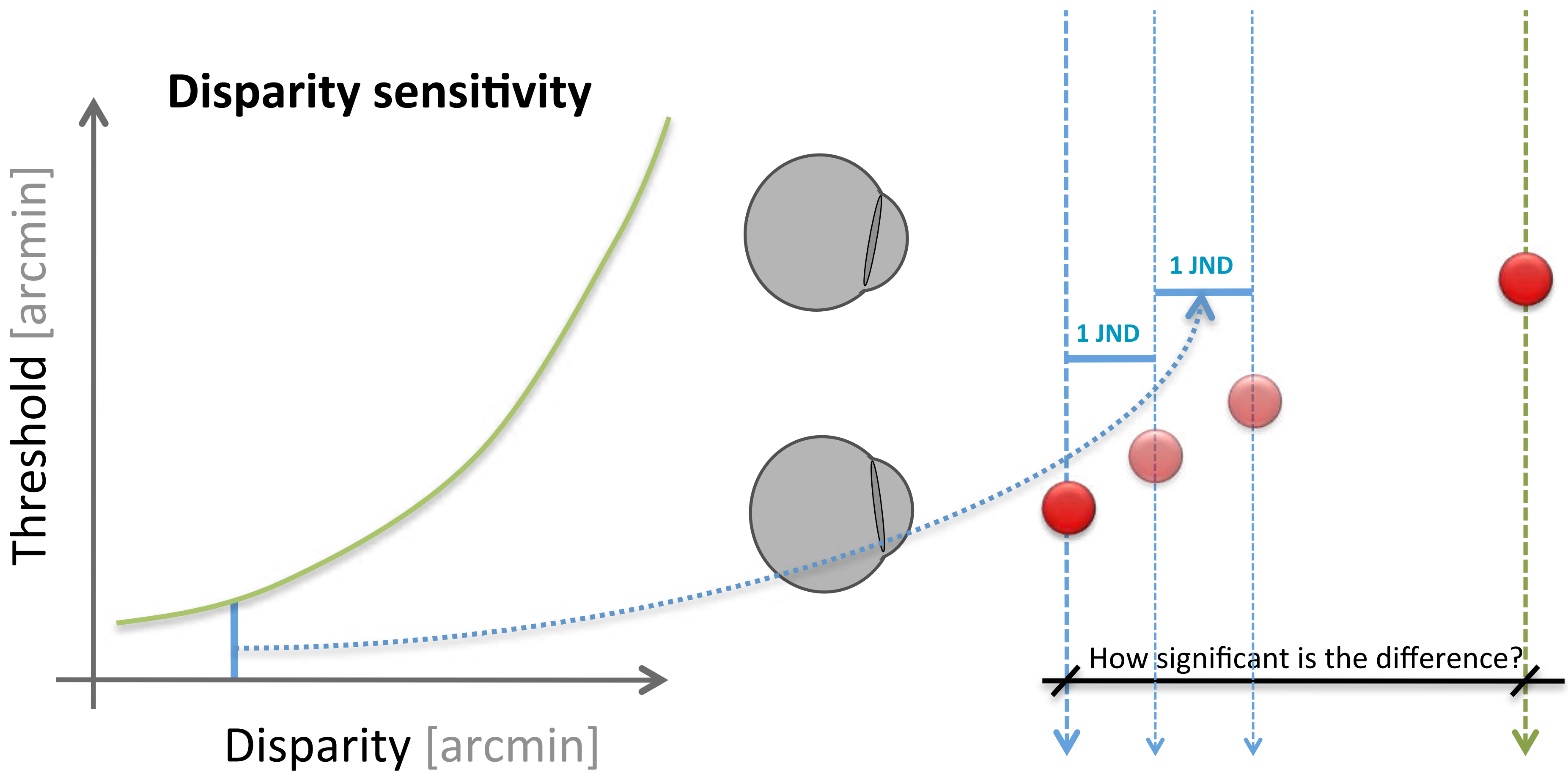
SIGGRAPHASIA2011 HONG KONG



The HVS response

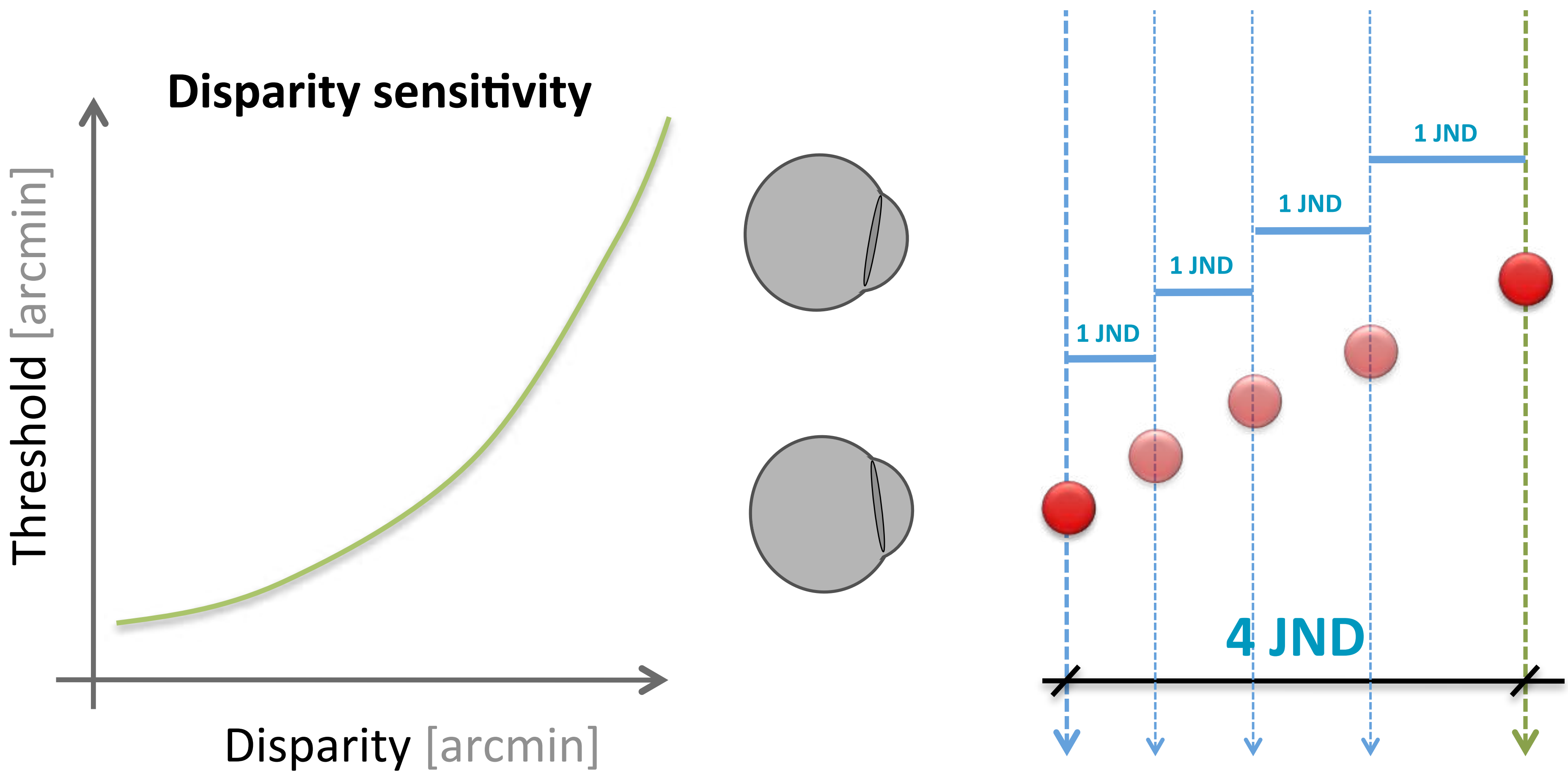


The HVS response

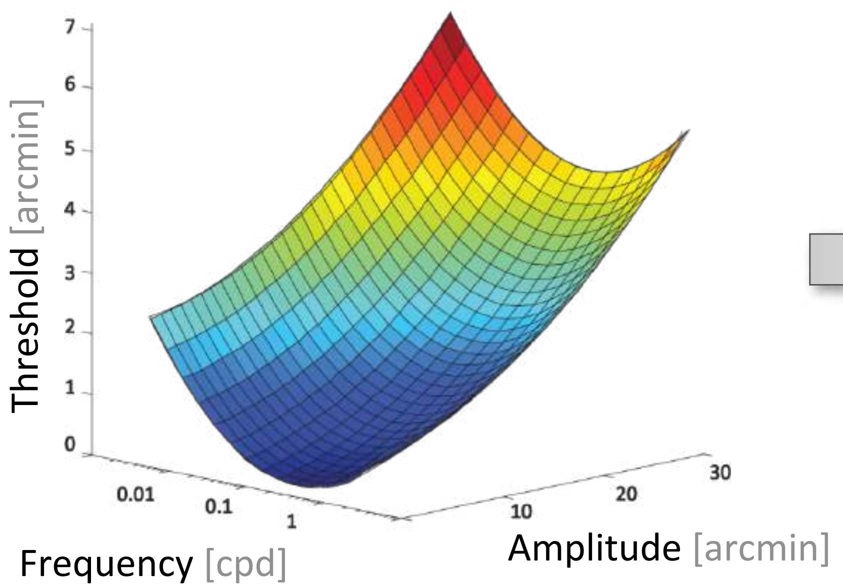


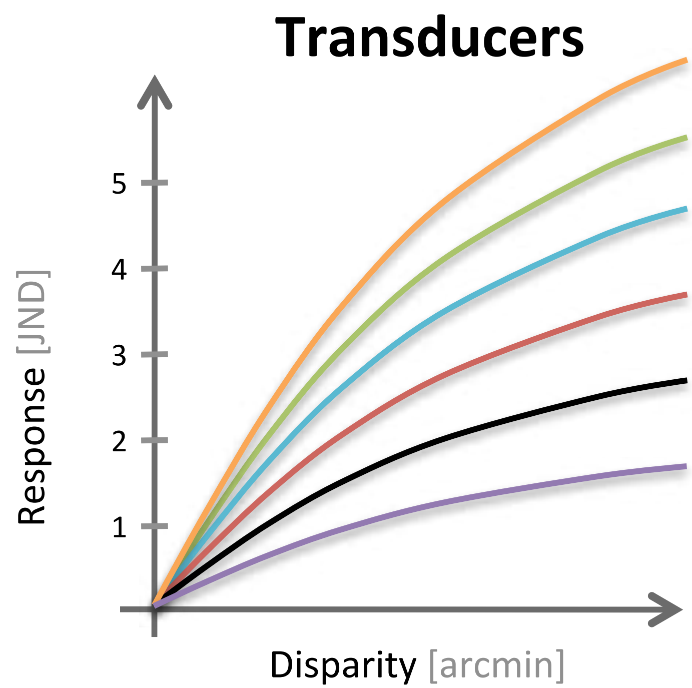
The HVS response

SIGGRAPHASIA2011 HONG KONG



The HVS response



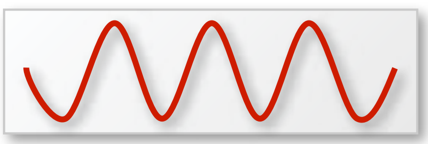


"A transducer function for threshold and suprathreshold human vision" by Wilson 1980

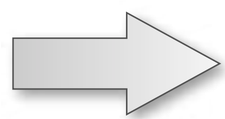
"A perceptual framework for contrast processing of high dynamic range images" by Mantiuk et al. 2005

Perceptual space

We show so far:



(disparity, frequency)
[arcmin, cpd]

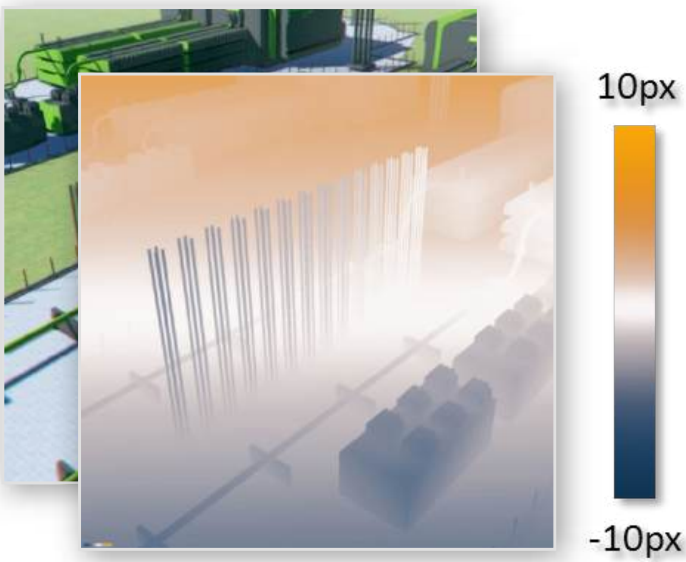


HVS response
[JND]

Perceptual space

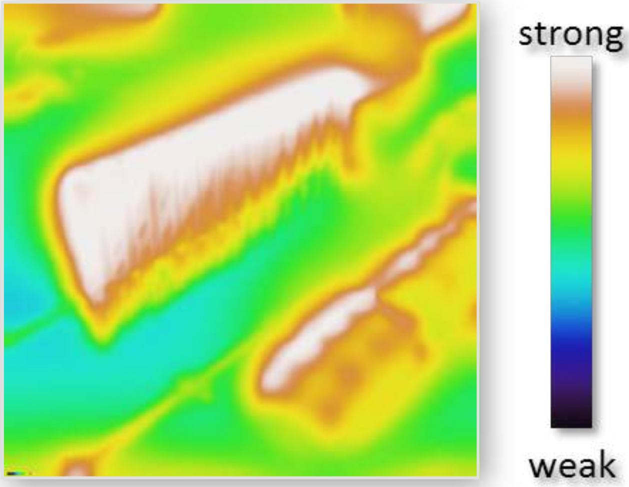
SIGGRAPH ASIA 2011 HONG KONG

The reality is more complex:



3D scene with pixel disparity

[pixels]

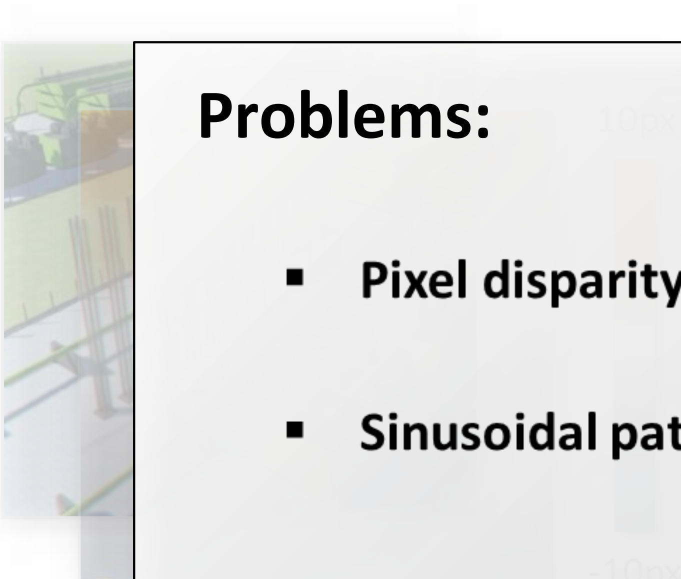


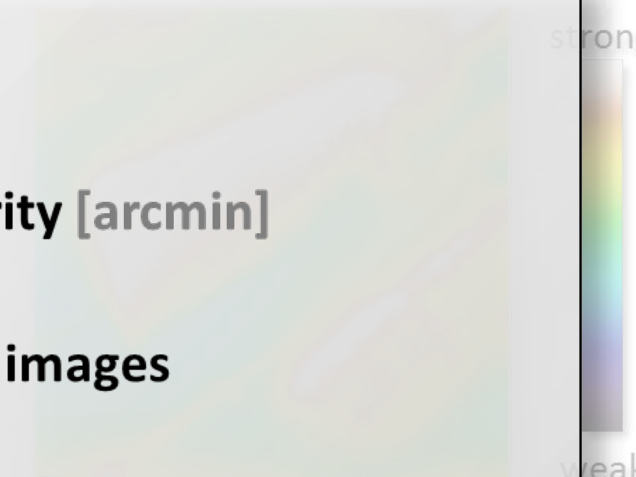
Map of HVS response

[JND]

Perceptual space

The reality is more complex:





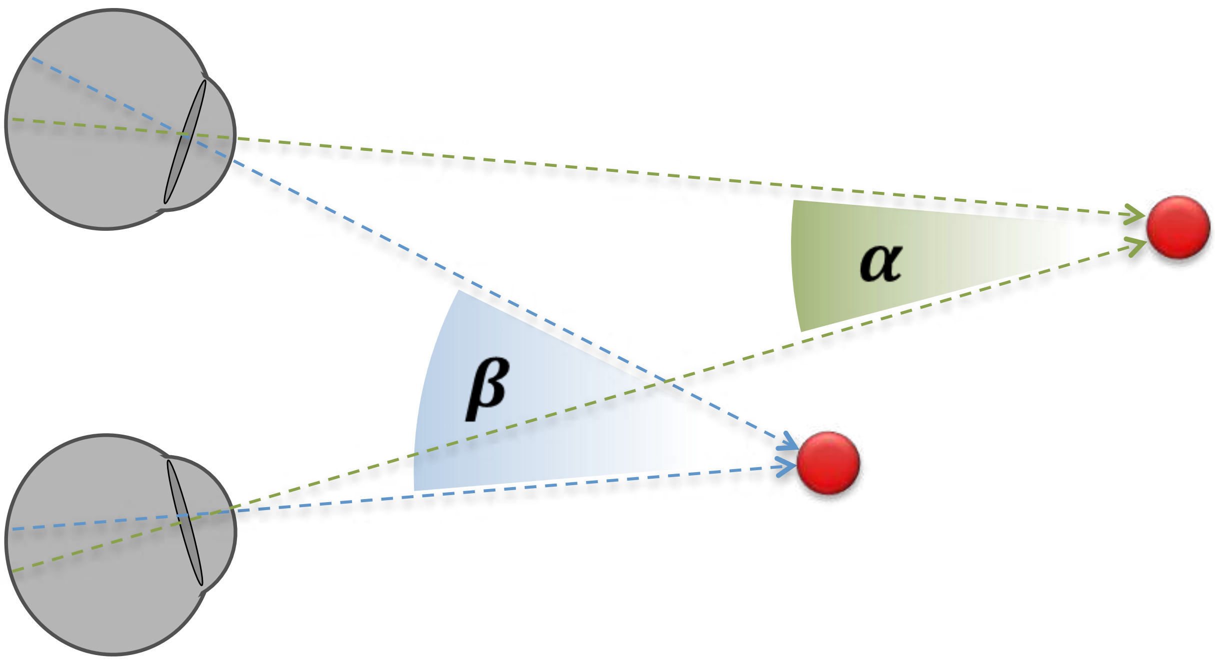
Problems:

- **Pixel disparity** [pixels] ↔ **Disparity** [arcmin]
- **Sinusoidal patterns → Complex images**

3D scene with pixel disparity [pixels]

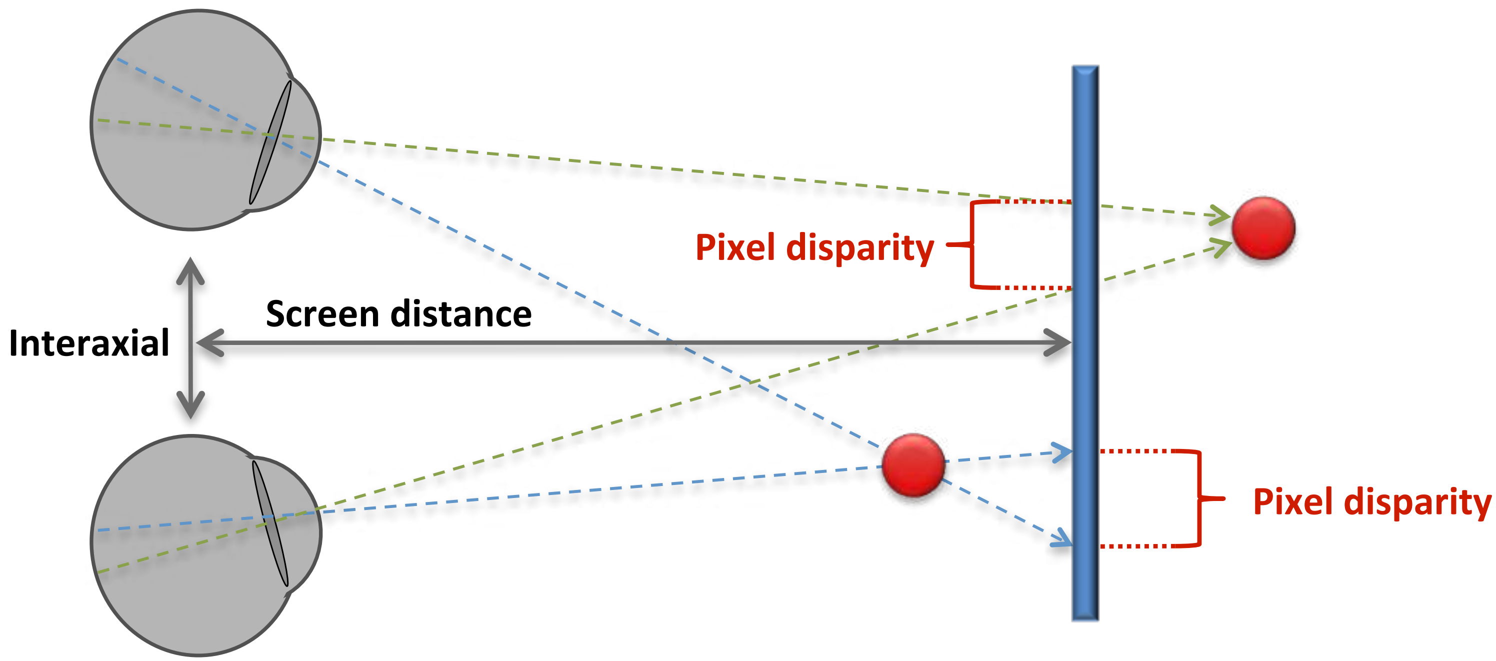
Map of HVS response [JND]

Pixel disparity to disparity



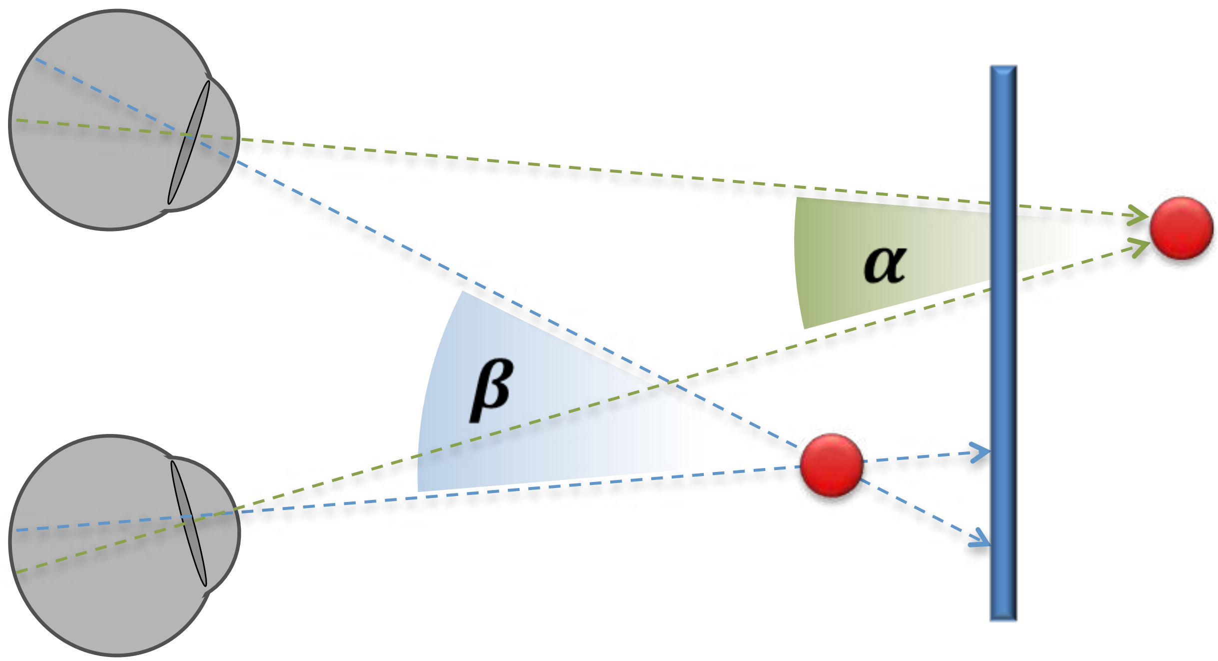
$$disparity = |\alpha - \beta|$$

Pixel disparity to disparity



Pixel disparity to disparity

SIGGRAPHASIA2011 HONG KONG



$$(\boldsymbol{viewing\ conditions, pixel\ disparity}) \rightarrow \boldsymbol{vergence}$$

Vergence to disparity

Vergence [arcmin]

Disparity [arcmin]

Vergence to disparity

How do people deal with luminance?

Luminance (contrast perception)

SIGGRAPHASIA2011 HONG KONG

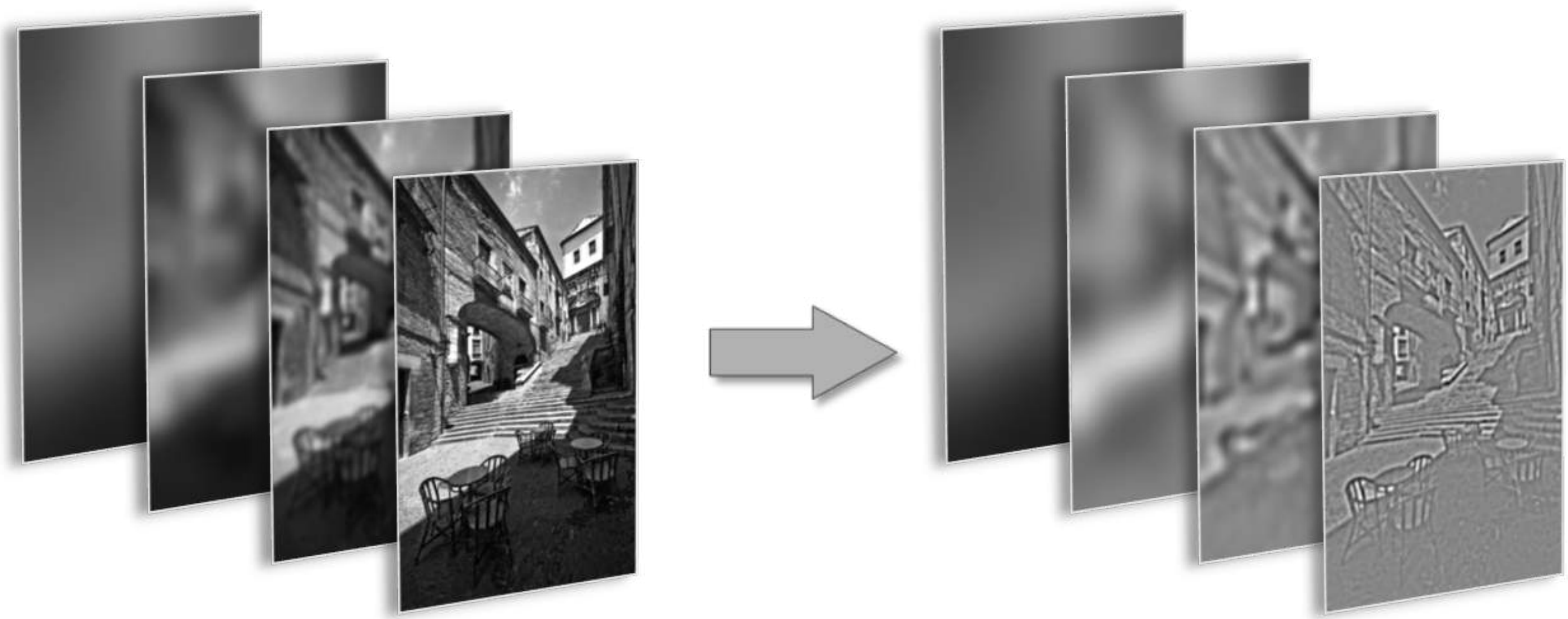
Luminance

Perceptual space

(Perceived contrast)

Luminance (contrast perception)

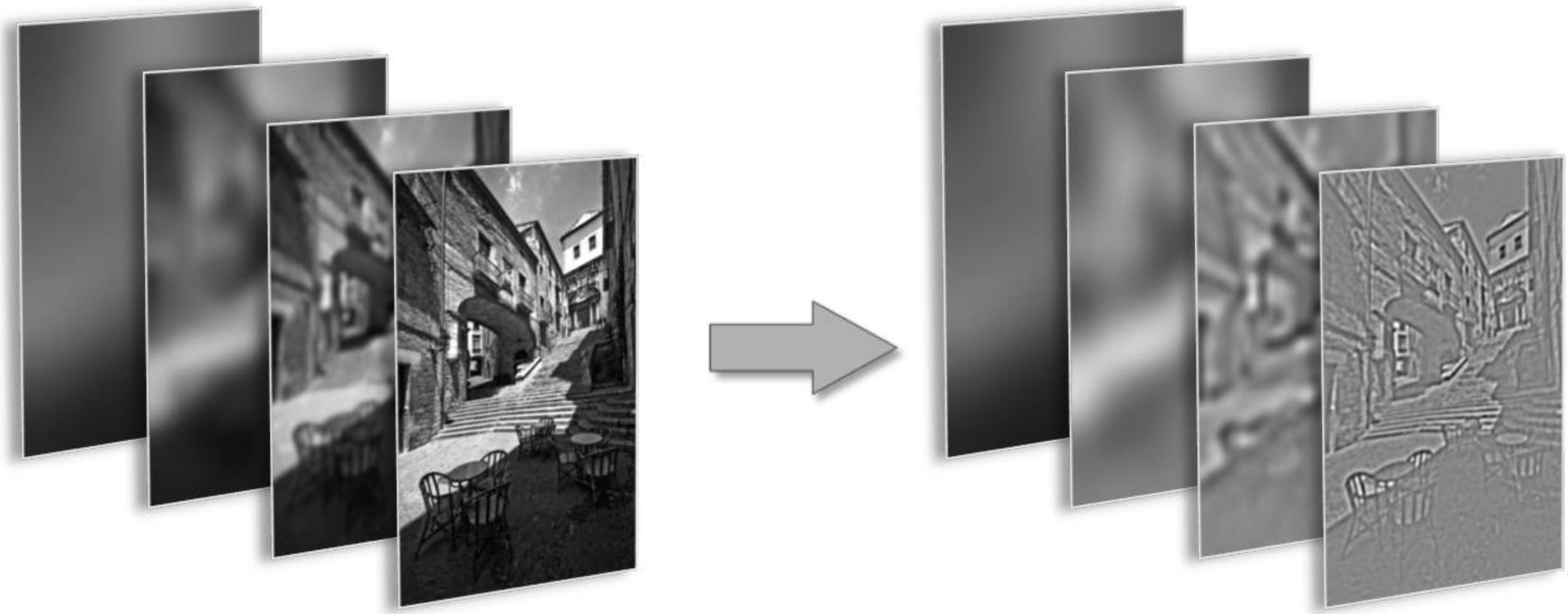
SIGGRAPHASIA2011 HONG KONG



Lowpass filters

Contrast decomposed into frequency bands

Luminance (contrast perception)



Lowpass filters

Perceptual operations

Luminance (contrast perception)

Works because:

Different frequencies are processed separately.

For disparity is similar.

Disparity is processed in independent channels.

"Seeing in depth" by Howard and Rogers 2002

Lowpass filters

Perceptual operations

Luminance (contrast perception)

Disparity / Luminance similarity:

Luminance → Vergence

Luminance contrast → Disparity

Lowpass filters

Perceptual operations

Vergence to disparity

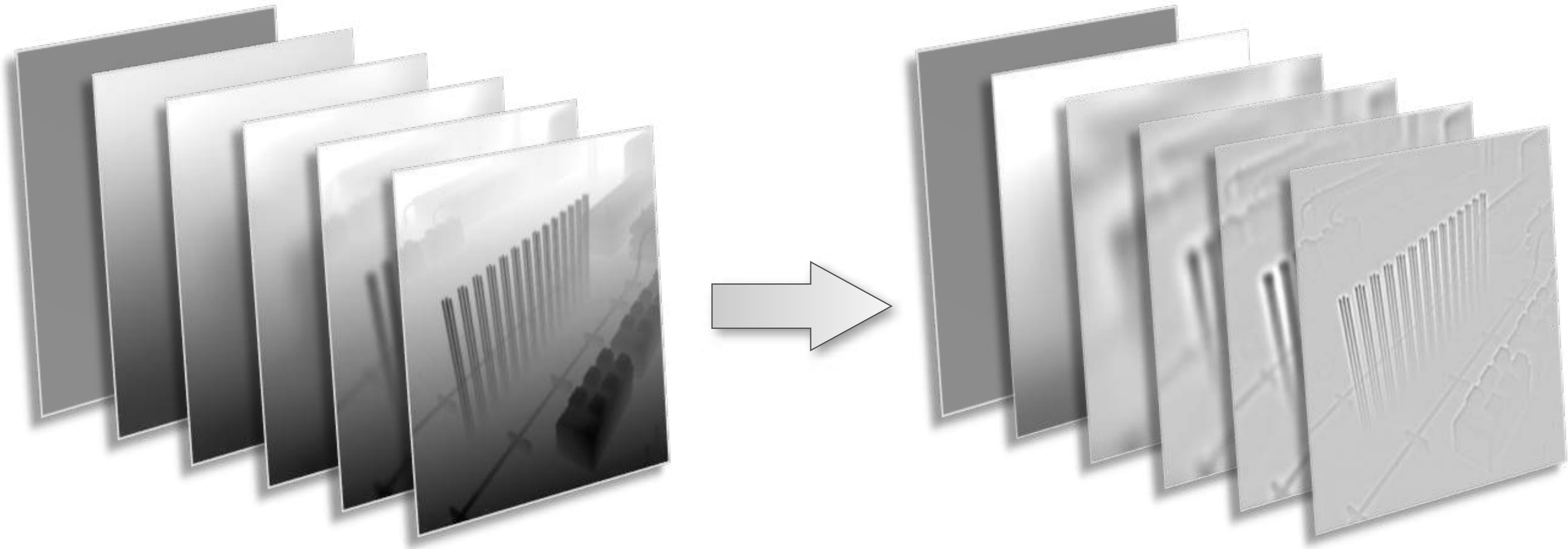
SIGGRAPHASIA2011 HONG KONG

Vergence [arcmin]

Vergence to disparity

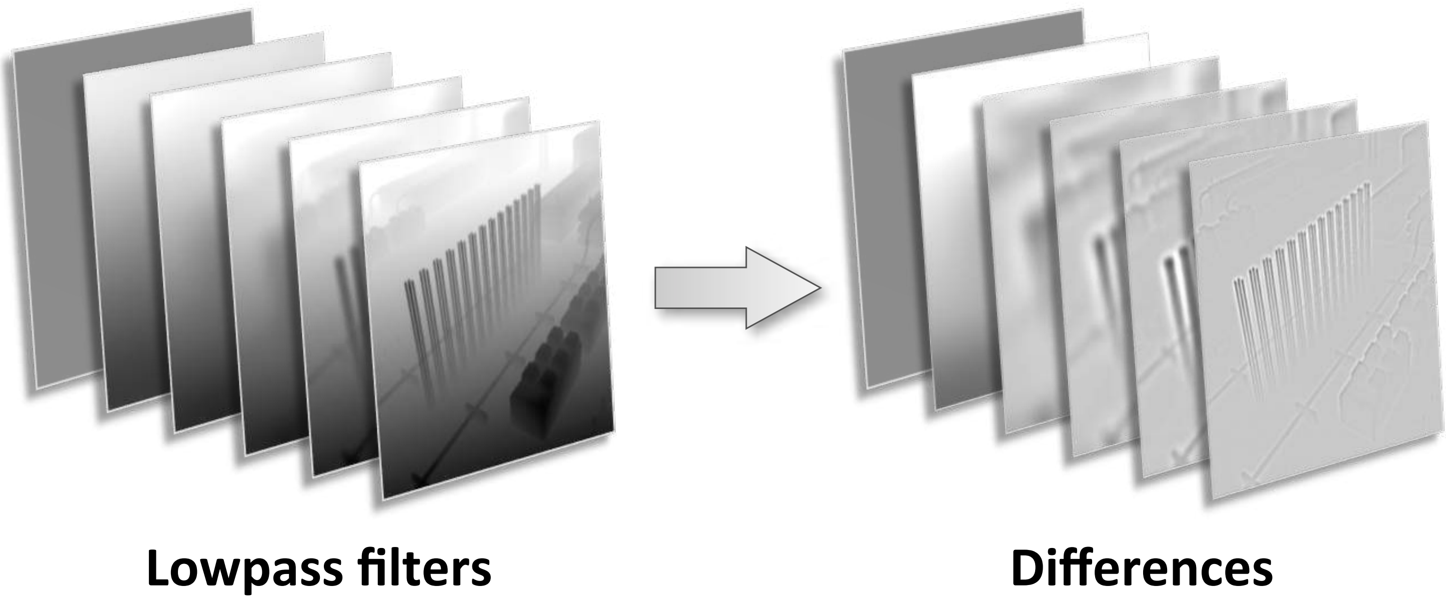
Lowpass filters

Vergence to disparity



Lowpass filters **Differences**

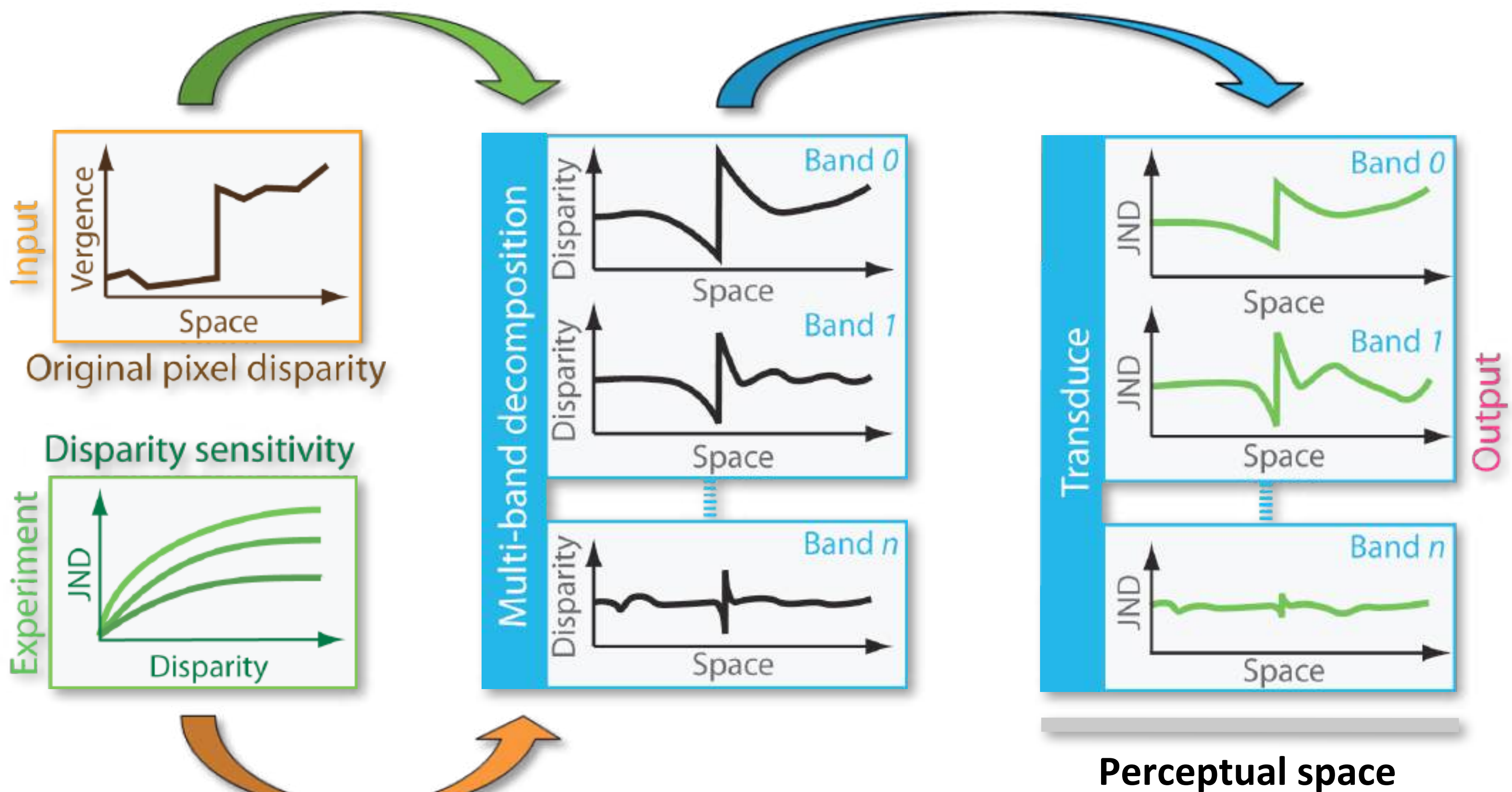
Vergence to disparity



- We can process frequencies independently
- Vergence → Disparity

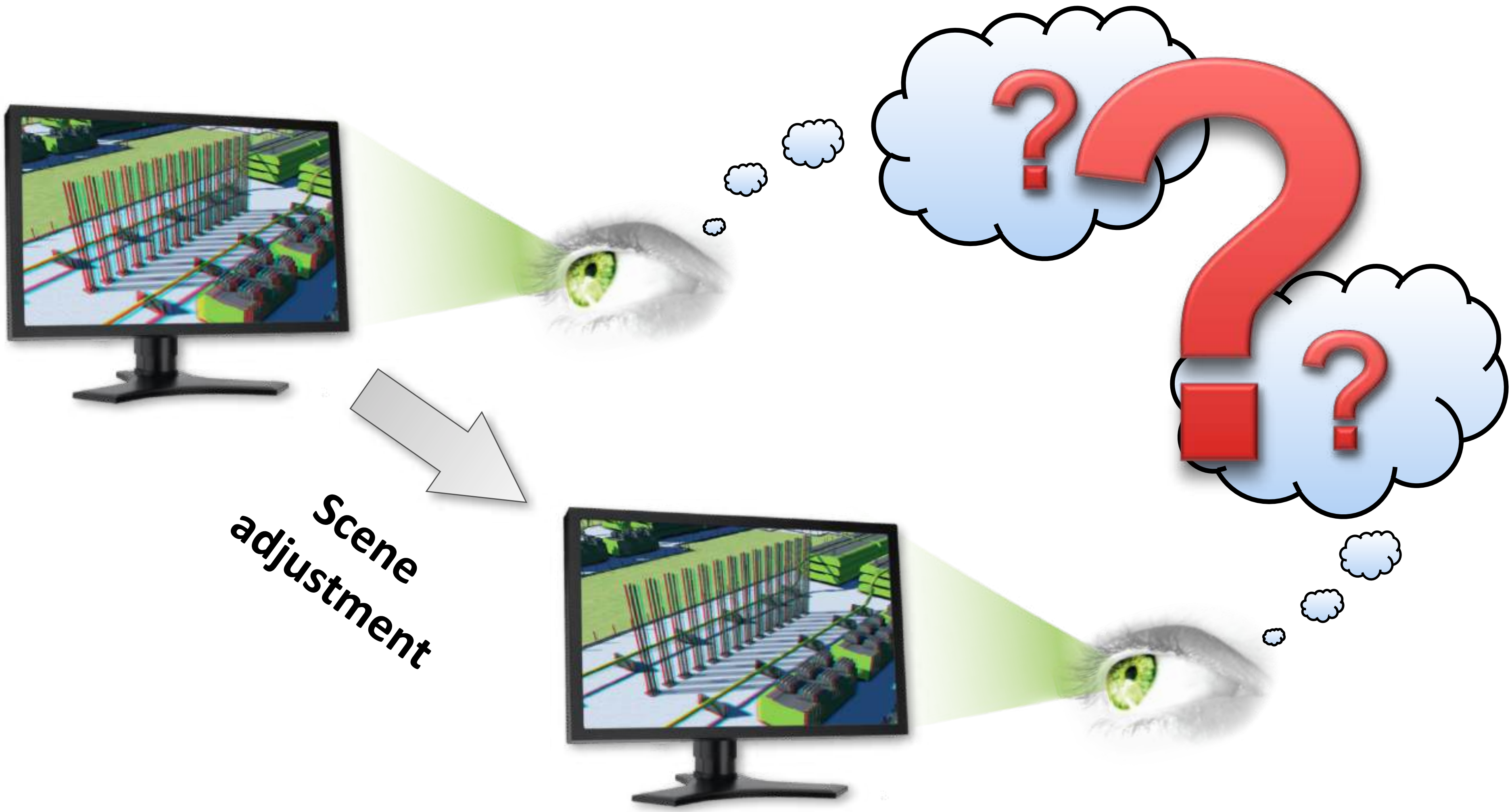
Perceptual model

SIGGRAPHASIA2011 HONG KONG



"A perceptual model for disparity" by Didyk et al. 2011

Disparity metric



Disparity metric

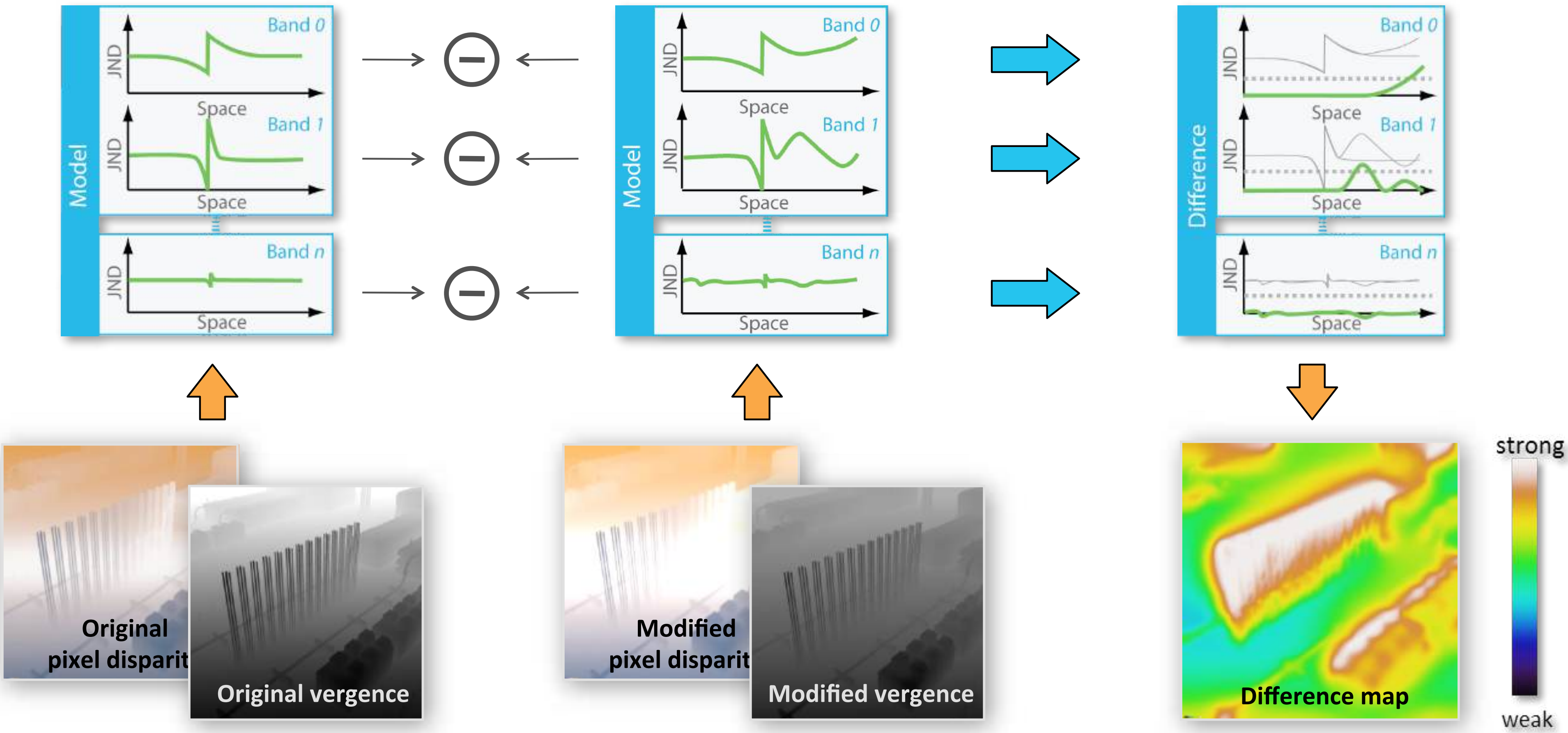
SIGGRAPH ASIA 2011 HONG KONG

For Luminance:

"A visual discrimination model for imaging system design and development"
by Lubin 1995

Disparity metric

SIGGRAPH ASIA 2011 HONG KONG



"A perceptual model for disparity" by Didyk et al. 2011

Disparity manipulations

SIGGRAPHASIA2011 HONG KONG

Manipulations in perceptual space:

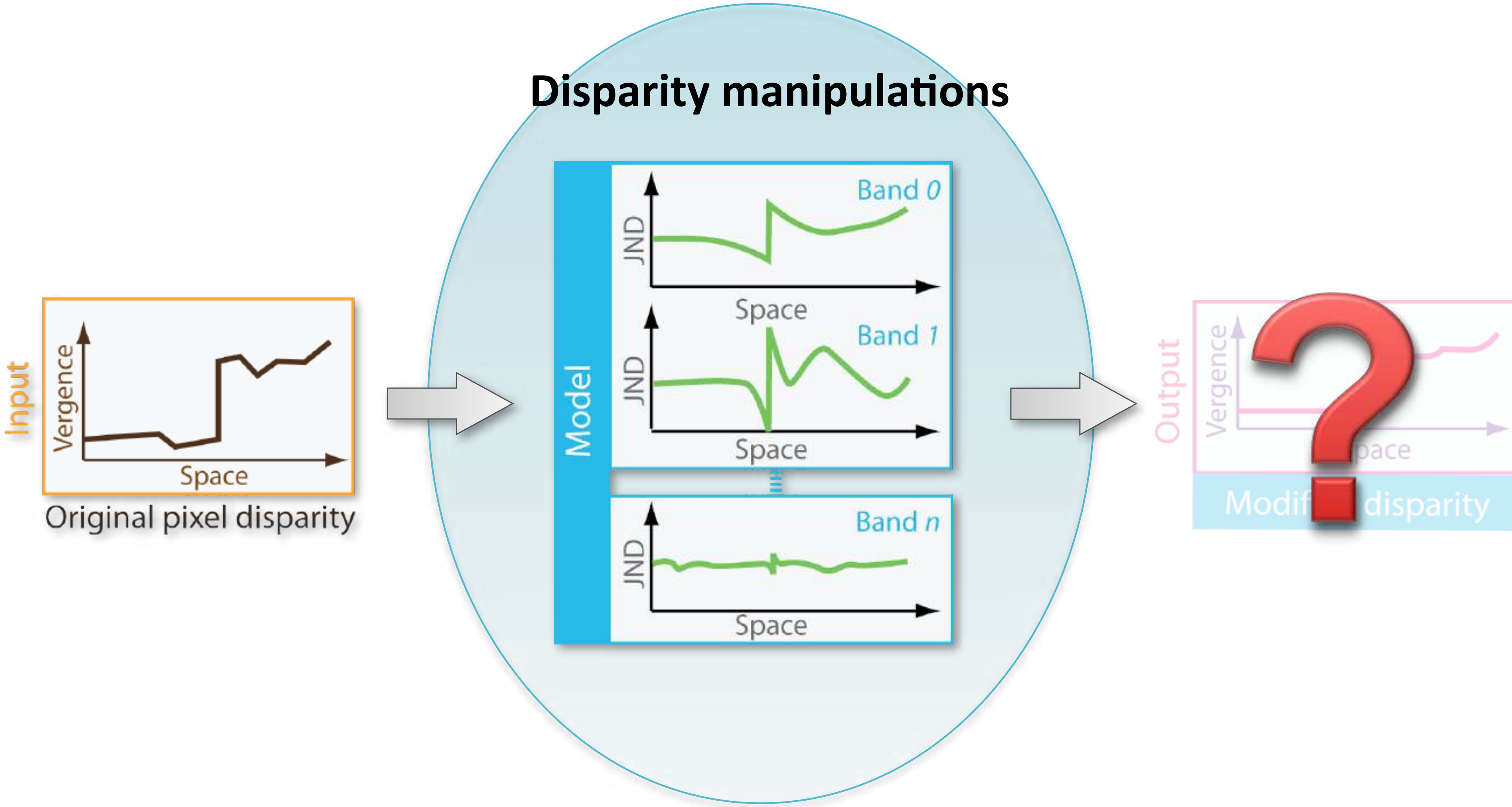
→ The HVS is taken into account

→ Efficient disparity reduction

→ Important disparities preserved

"Nonlinear Disparity Mapping for Stereoscopic 3D"
by Lang et al. 2010

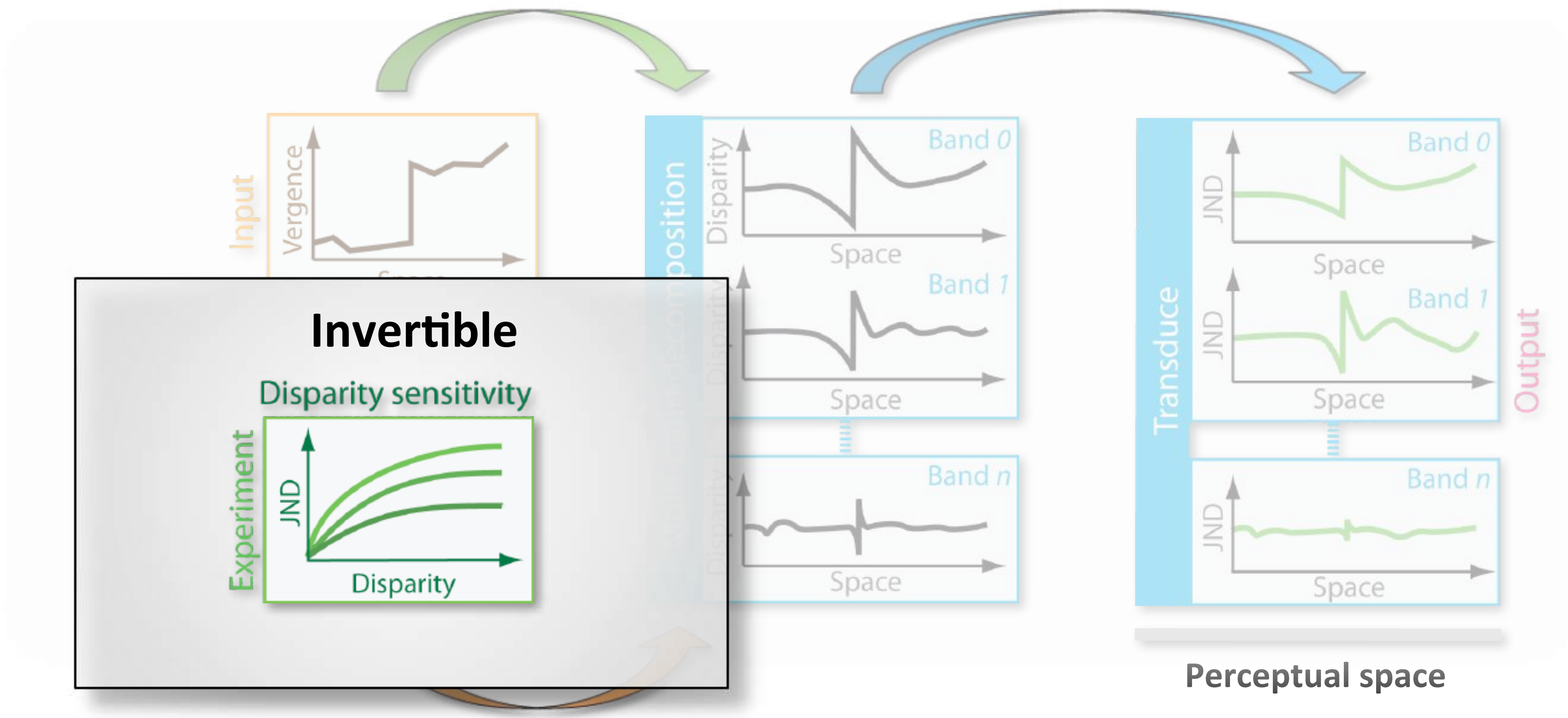
Disparity manipulation



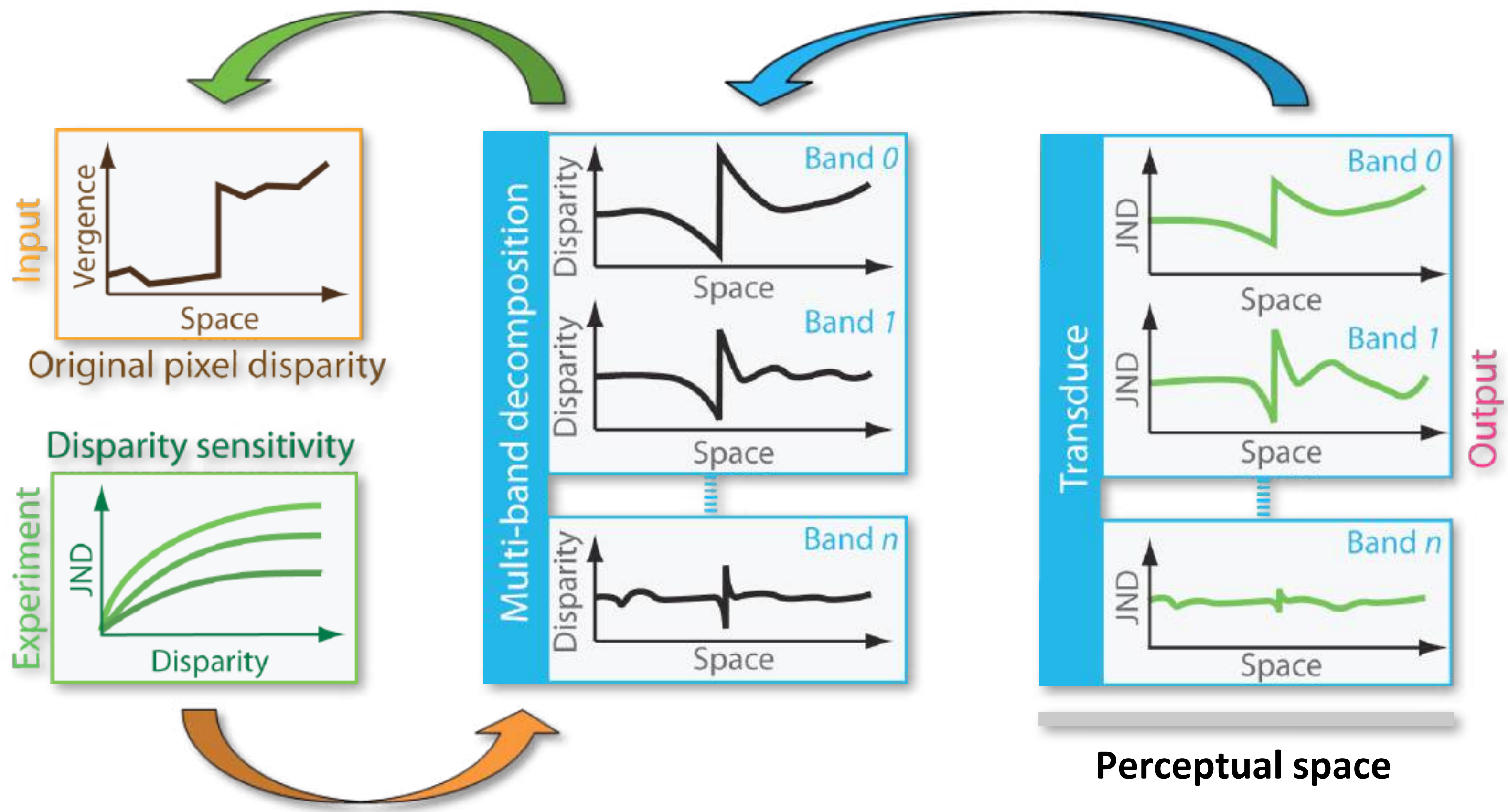
"A perceptual model for disparity" by Didyk et al. 2011

Inverse model

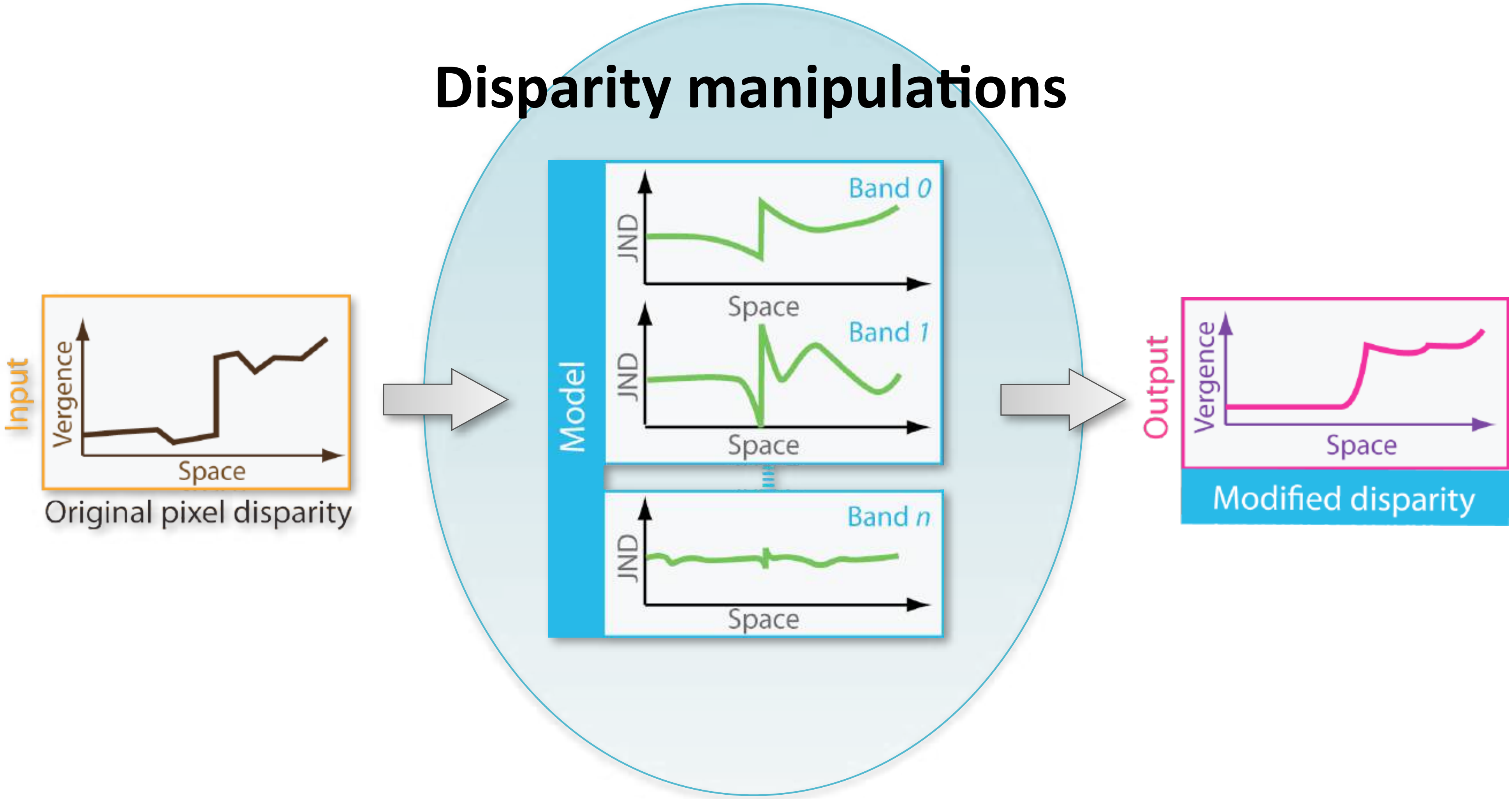
SIGGRAPHASIA2011 HONG KONG



Inverse model

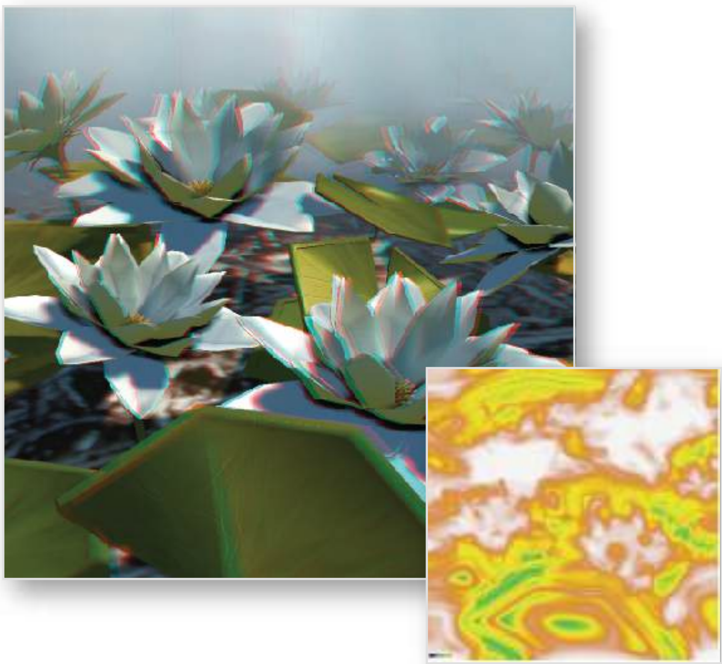


Disparity manipulation



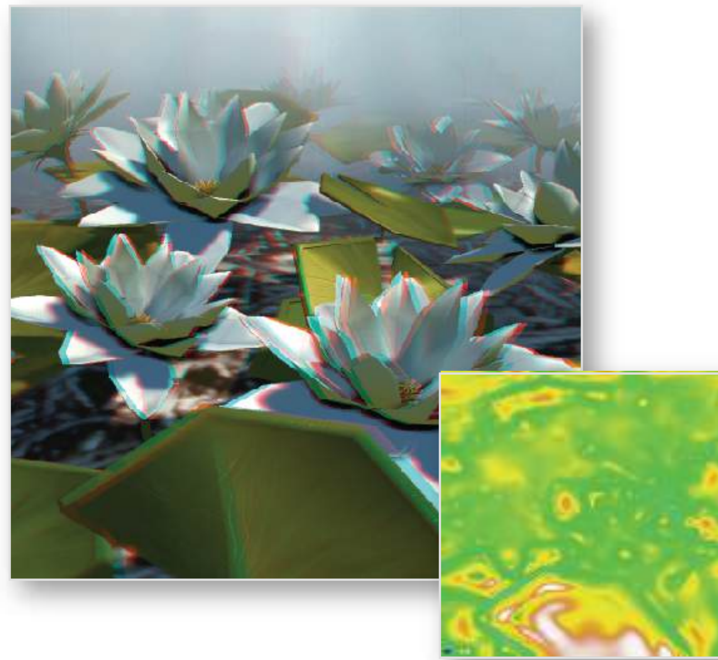
Disparity manipulation

Standard technique



Perceived distortions

In perceptual space



Perceived distortions

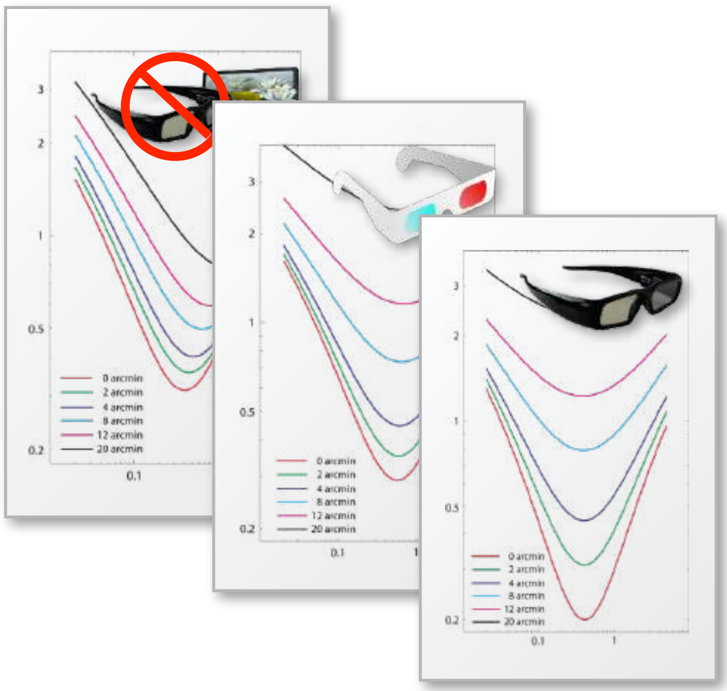
strong

weak

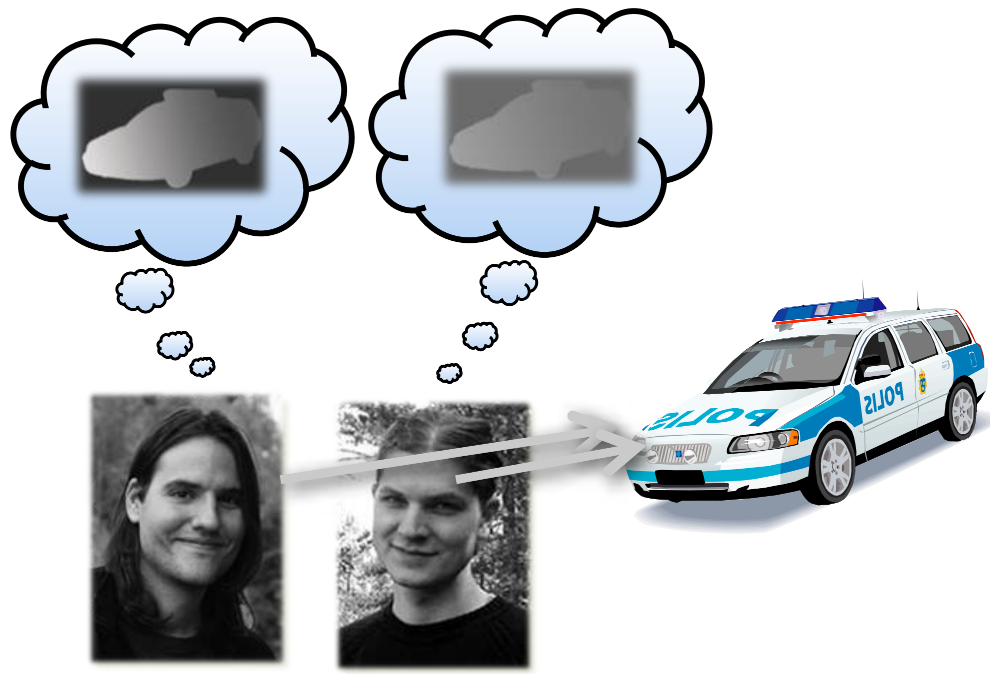
- Important disparities preserved

Personalization

Disparity perception depends on:

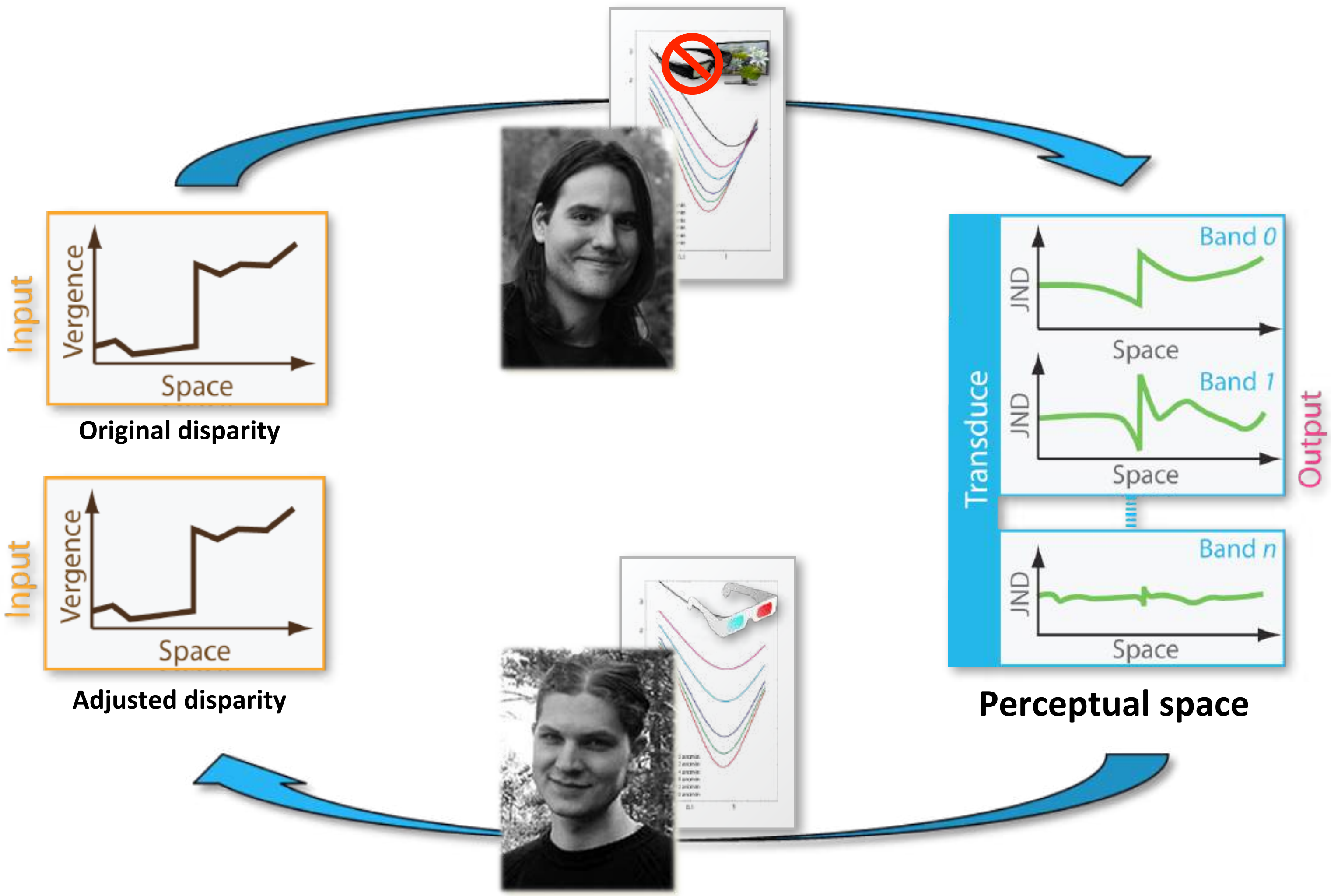


Equipment



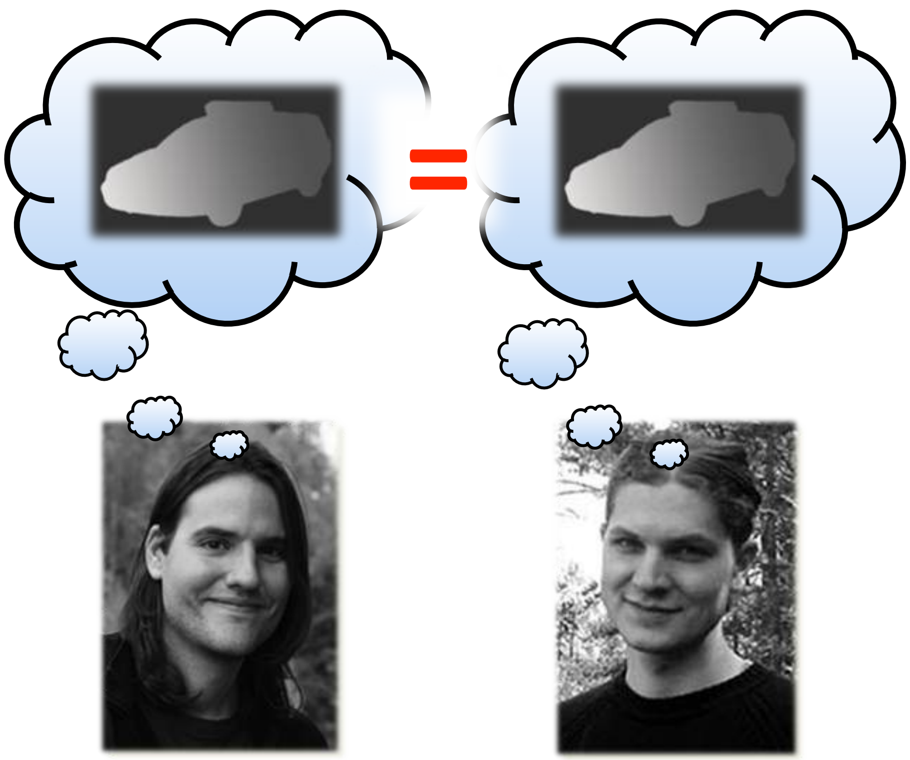
User

Personalization



"A perceptual model for disparity" by Didyk et al. 2011

Personalization



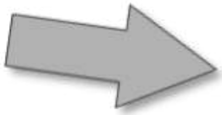
All users perceive the same regardless:

- Equipment
- Disparity sensitivity

Backward-compatible stereo

Standard stereo

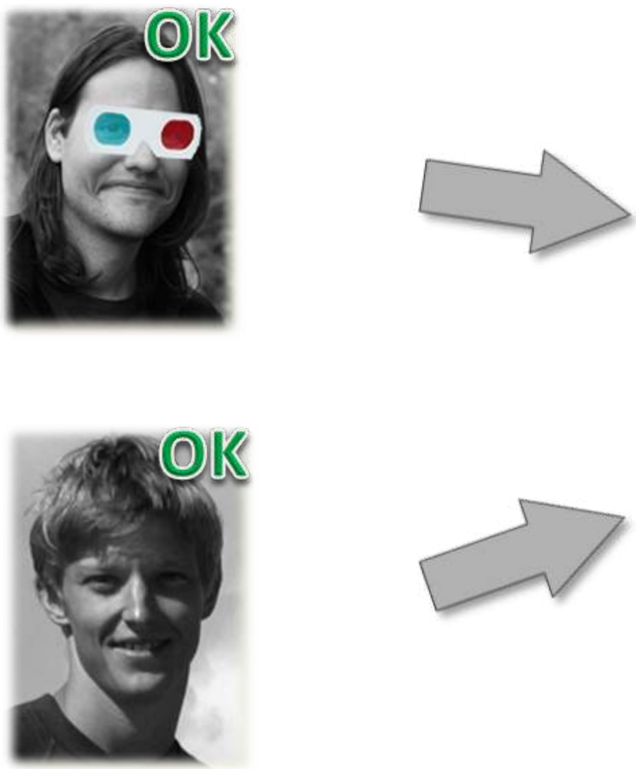
Backward-compatible stereo



Standard 2D image

Backward-compatible stereo

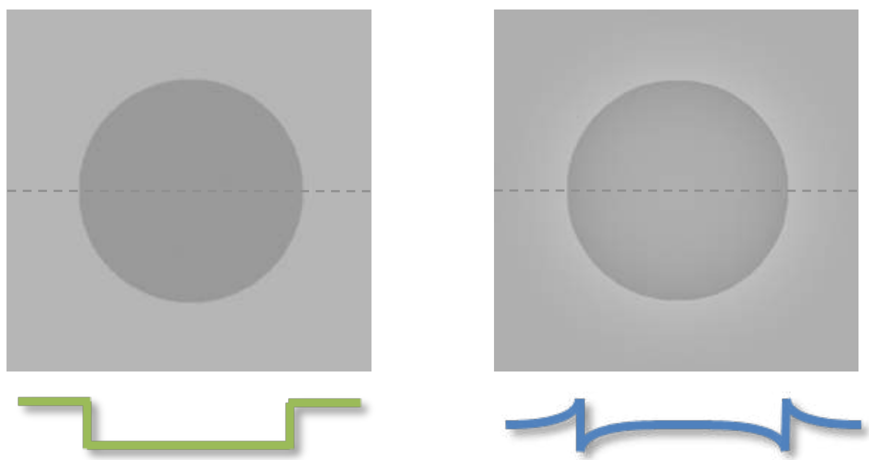
SIGGRAPHASIA2011 HONG KONG



Backward-compatible stereo

Cornsweet illusion

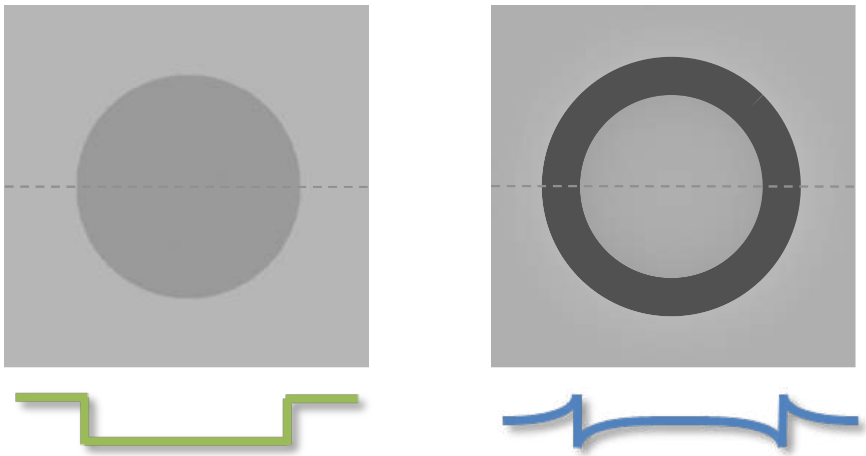
SIGGRAPHASIA2011 HONG KONG



- Similar perceived contrast

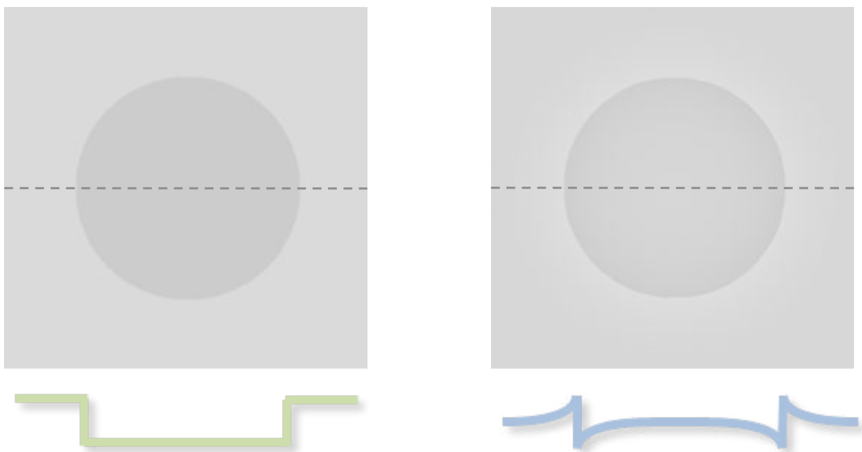
Cornsweet illusion

SIGGRAPHASIA2011 HONG KONG



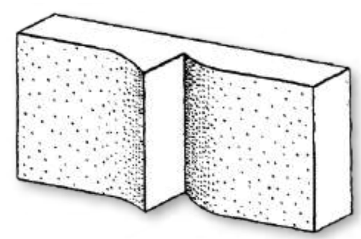
- Similar perceived contrast
- Luminance range reduced

Cornsweet illusion



- Similar perceived contrast
- Luminance range reduced

Cornsweet illusion works for depth:



“A Craik-O'Brien-Cornsweet illusion for visual depth” by Anstis et al. 1997

Backward-compatible stereo

SIGGRAPHASIA2011 HONG KONG

Standard stereo

Backward-compatible stereo

- 3D impression preserved
- No artifacts when special equipment is unavailable

"A perceptual model for disparity" by Didyk et al. 2011

Backward-compatible stereo

- 3D impression preserved
- No artifacts when special equipment is unavailable

"A perceptual model for disparity" by Didyk et al. 2011

Conclusions

SIGGRAPHASIA2011 HONG KONG

- Stereo is a hot topic
- Stereo perception is complex phenomenon
- Stereo content should be adjusted according to:
  - Viewing conditions, viewer, equipment
- Different ways of stereo content adjustment exist:
  - Camera adjustment
  - Pixel disparity mapping operators
  - Perceptual space
- Predicting perceived distortions is important for 3D content preparation