Eurographics 2012
Cagliari, Italy May 13-18
33rd ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR COMPUTER GRAPHICS

Mapping images to target devices: spatial, temporal, stereo, tone, and color

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

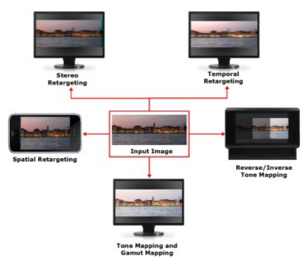
Diego Gutierrez
Universidad de Zaragoza, Spain

Rafał Mantiuk
University of Bangor, UK

Tobias Ritschel
Télécom ParisTech/
MPI Informatik, Germany

Special thanks to:
Karol Myszkowski, MPI Informatik, Germany

Introduction: The Problem



Introduction: Outline

- **Dynamic Range and Color Retargeting (~70 mins):**
  - Rafał Mantiuk, Tobias Ritschel, and Alessandro Artusi
- **Reverse/Inverse Tone Mapping (~45 mins) :**
  - Francesco Banterle
- **Image Spatial Resolution Retargeting (~45 mins) :**
  - Diego Gutierrez
- **Temporal Image Retargeting (~70 mins) :**
  - Tobias Ritschel and Elmar Eisemann
- **Image and Video Quality Assessment(~70 mins) :**
  - Tunç O. Aydın
- **Stereo Content Retargeting (~50 mins):**
  - Piotr Didyk

Eurographics 2012
Cagliari, Italy May 13-18
33rd ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR COMPUTER GRAPHICS

Tone Mapping

Rafal Mantiuk

Bangor University, UK

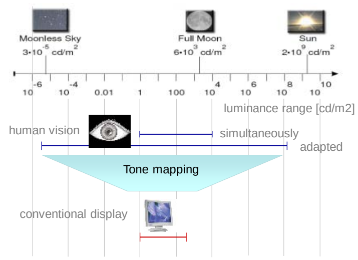
Research Institute of Visual Computing

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

Learning outcomes

- What is tone-mapping?
- What problem(s) does it solve?
- Why is the problem so difficult?
- How do we perceive high dynamic range images?
- What are the major approaches to tone-mapping?
- How to choose a tone-mapping for a particular application?

Tone-mapping problem

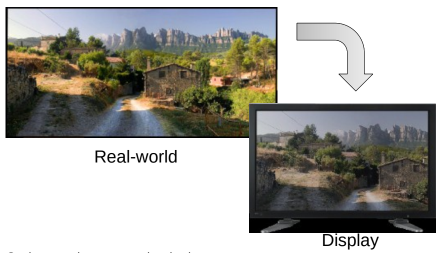


Question to the audience

- Who has never used a tone-mapping operator?

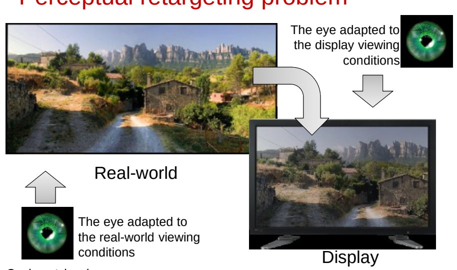
Each camera needs to tone-map a real-world captured light before it can be stored as a JPEG. This is essentially the same process as tone-mapping, although knows as 'color reproduction' or 'color processing'.

Color space retargeting problem



Goal: map colors to a restricted color space

Perceptual retargeting problem



Goal: match color appearance

Tone Mapping?

- HDR ?
- Or something else ?

What is tone-mapping?

Although tone-mapping may have different meanings, this course is about:

A) Transformation of an image from an unrestricted color gamut of real world or an abstract scene to the restricted color gamut of a device

B) Retargeting the perceptual appearance from one viewing conditions to another

Input and output

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

Tone mapping

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

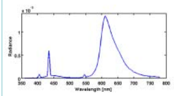
Luminance

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

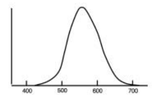
Luminance

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

Light spectrum (radiance)



Luminous efficiency function (weighting)



Eurographics 2012, Cagliari, Italy

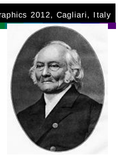
Do HDR images contain luminance values?

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

Eurographics 2012, Cagliari, Italy

Sensitivity to luminance

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



Ernst Heinrich Weber
[From wikipedia]

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

The smallest detectable luminance difference

Background (adapting) luminance

Constant

Typical stimuli:

Consequence of the Weber-law

- Smallest detectable difference in luminance

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

L	ΔL
100 cd/m²	1 cd/m²
1 cd/m²	0.01 cd/m²

- 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!

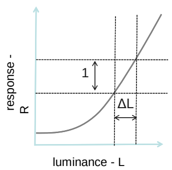
How to make luminance (more) perceptually uniform?

- 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

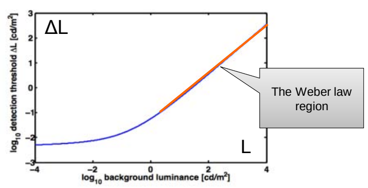
$$R(L) = a \ln(L)$$

- 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]

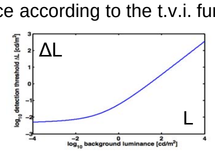
But…the Fechner law does not hold for the full luminance range

- 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:



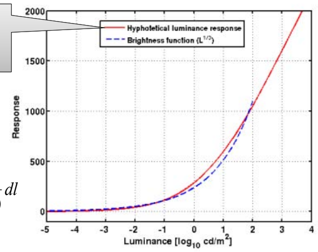
- we can get more accurate estimate of the "response":

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

Fechnerian integration and Steven's law

R(L) - 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 & inverse visual models
- Constraint mapping problem

- This is not a crisp categorization
  - Some operators combine several approaches

Major approaches to tone-mapping

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

Illumination & reflectance separation

Input

Illumination

Reflectance

Illumination and reflectance

Illumination

- Sun ≈ 10^9 cd/m^2
- Lowest perceivable luminance ≈ 10^{-6} cd/m^2
- Dynamic range 10,000:1 or more
- 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

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

Tone-mapped image · Reflectance · Tone-mapping · Illumination

- 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

[Durand & Dorsey, SIGGRAPH 2002]

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

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

Spatially varying smoothing – less smoothing near the edges

- [Farbman et al., SIGGRAPH 2008]

WLS filter

- Stronger smoothing and still distinct edges

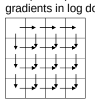
- Can produce stronger effects with fewer artifacts

Tone mapping result

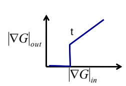
Retinex

- Retinex algorithm was initially intended to separate reflectance from illumination [Land 1964]
  - There are many variations of Retinex, but the general principle is to eliminate from an image small gradients, which are attributed to the illumination

1 step: compute gradients in log domain



2nd step: set to 0 gradients less than the threshold



3rd step: reconstruct an image from the vector field

$$\nabla^2 I = \operatorname{div} G$$

For example by solving the Poisson equation

Retinex examples

From: http://dragon.larc.nasa.gov/retinex/757/

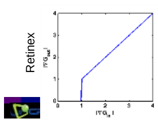
From:http://www.ipol.im/pub/algo/lmps_retinex_poisson_equation/#ref_1

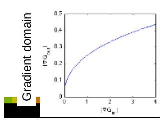
Eurographics 2012, Cagliari, Italy

Gradient domain HDR compression

[Fattal et al., SIGGRAPH 2002]

- Similarly to Retinex, it operates on log-gradients
- But the function amplifies small contrast instead of removing it



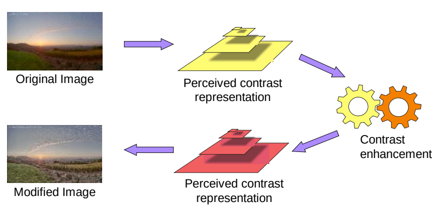


- Contrast compression achieved by global contrast reduction
  - Enhance reflectance, then compress everything

Eurographics 2012, Cagliari, Italy

Contrast domain image processing

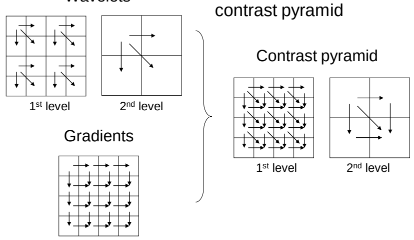
[Mantiuk et al., ACM Trans. Applied Perception, 2006]



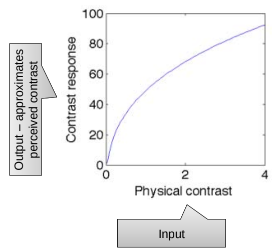
Rationale: Human eye is more sensitive to contrast than luminance

Contrast domain image processing

Image transform: Multi-scale contrast pyramid

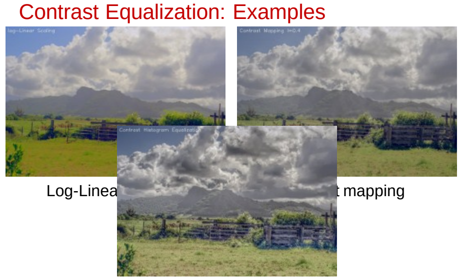


Contrast transducer function



Goal: Transform contrast to the representation that is possibly perceptually uniform.

Contrast Equalization: Examples



Log-Linea

t mapping

Contrast equalization

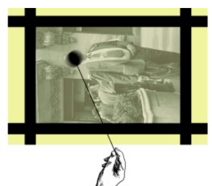
Contrast Equalization: Examples

Log-Linear S   rast mapping

Contrast equalization

Tone mapping in photography

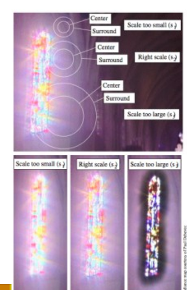
- Dodging and burning
  - Darken on brighten image parts by occluding photographic paper during exposure
  - Ansel Adams, *The print*, 1995
  - Photoshop tool



- Essentially – attenuate low-pass frequencies associated to illumination

Automatic dodging and burning

- Reinhard et al., *Photographic tone reproduction for digital images*. SIGGRAPH 2002
- Choose dodging an burning kernel size adaptively
  - depending on the response of the center-surround filter
  - thus avoid halo artifacts

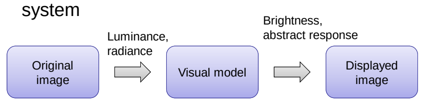


Major approaches to tone-mapping

- Illumination & reflectance separation
- Forward visual model
- Forward & inverse 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

Forward visual model: Retinex

- Remove illumination component from an image
  - Because the visual system also discounts illuminant
- Display ‘reflectance’ image on the screen

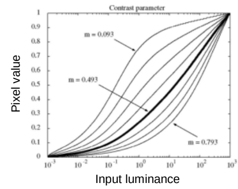
- Assumption:
  - The abstract ‘reflectance’ contains most important visual information
  - Illumination is a distraction for object recognition and scene understanding

Photoreceptor response

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

$$V = \frac{I}{I + \sigma(I_a)} V_{max}$$
$$\sigma(I_a) = (f I_a)^m.$$

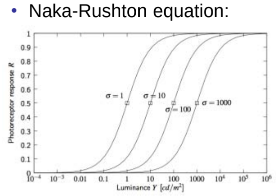
- From gamma to sigmoidal response:

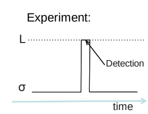


Results: photoreceptor TMO

Photoreceptor models

- Naka-Rushton equation: $\frac{R}{R_{max}} = \frac{Y^n}{Y^n + \sigma^n}$

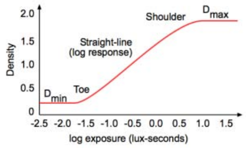




- Response of the photoreceptor to a short flicker of light - less applicable to viewing static images

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!

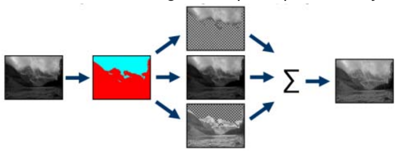


Why sigmoidal tone-curves work

- 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

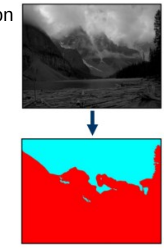
- Lightness perception in tone-reproduction for high dynamic range images [Krawczyk et al. '05]
- Based on Gilchrist lightness perception theory



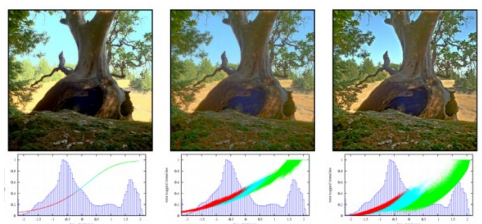
- Perceived lightness is anchored to several frameworks

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

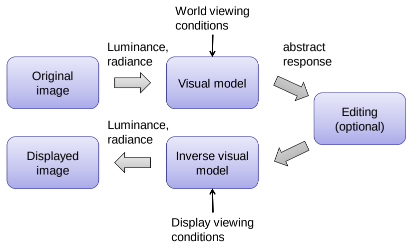


Photographic Tone Reproduction | Bilateral Filtering | Presented Computational Model

Major approaches to tone-mapping

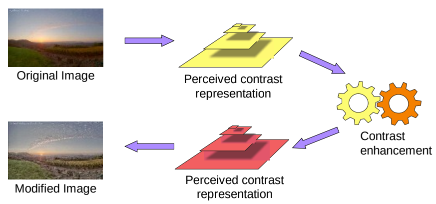
- Illumination & reflectance separation
- Forward visual model
- Forward & inverse visual model
- Constraint mapping problem

Forward and inverse visual model



Contrast domain image processing

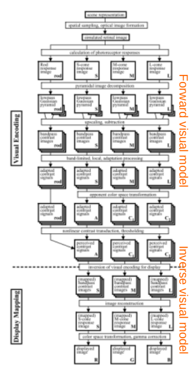
[Mantiuk et al., ACM Trans. Applied Perception, 2006]



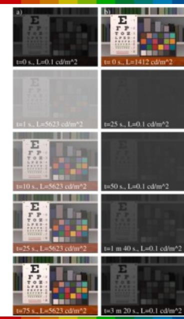
Rationale: Human eye is more sensitive to contrast than luminance

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

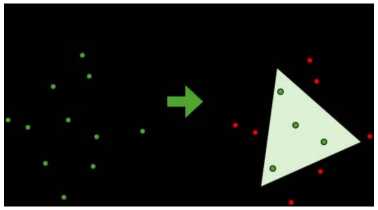
- Advantages of F&I visual models
  - Can render images for different viewing conditions
    - Different state of chromatic or luminance adaptation
  - Physically plausible
    - output in the units of luminance or radiance
- Shortcomings F&I visual models
  - Assume that a standard display can reproduce 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

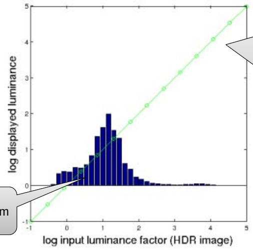
- Illumination & reflectance separation
- Forward visual model
- Forward & inverse visual model
- Constraint mapping problem

Constraint mapping problem

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



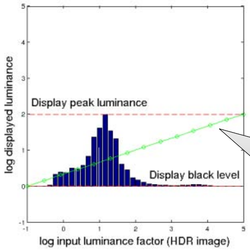
Global tone mapping operator



Best tone-mapping is the one which does not do anything, i.e. slope of the tone-mapping curves is equal to 1.

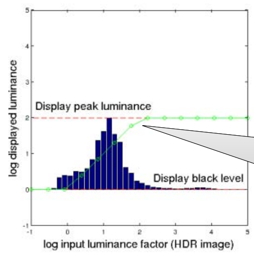
Image histogram

Display limitations



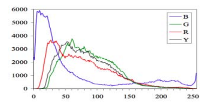
But in practice contrast (slope) must be limited due to display limitations.

Tone mapping



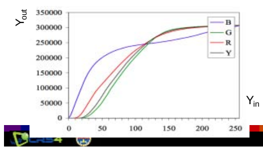
Histogram equalization

- 1. Compute cumulative distribution function:



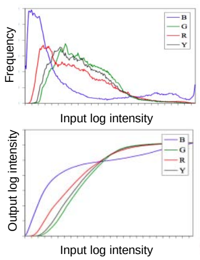
$$c(I) = \frac{1}{N}\sum_{i=0}^{I} h(i) = c(I-1) + \frac{1}{N}h(I)$$

- 2. Use that function to assign new pixel values



$$Y_{out} = c(Y_{in})$$

Histogram equalization



- Steepest slope for strongly represented bins
  - Enhance contrast, if many pixels
  - Reduce contrast, if few pixels
- HE distributes contrast distortions relative to the “importance” of a brightness level

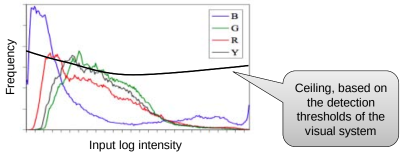
Histogram adjustment with a linear ceiling

- [Larson et al. 1997, IEEE TVCG]

Linear mapping | Histogram equalization | Histogram equalization with ceiling

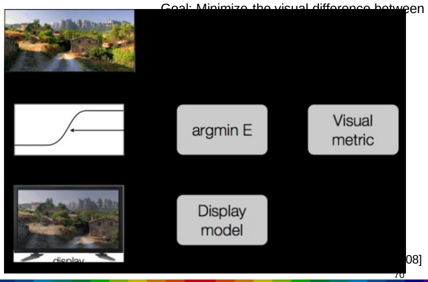
Histogram adjustment with a linear ceiling

- Truncate the bins that exceed the ceiling
- Recompute the ceiling based on the truncated histogram
- Repeat until converges



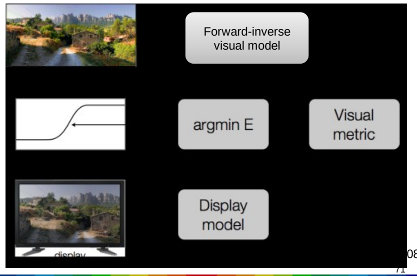
Display adaptive tone-mapping

Goal: Minimize the visual difference between

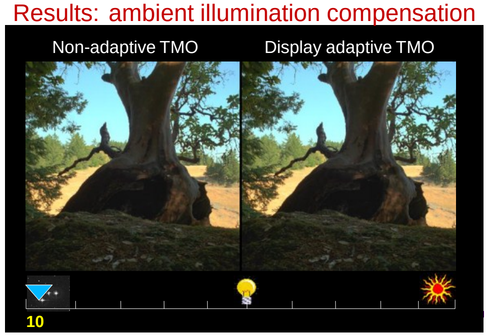


08]

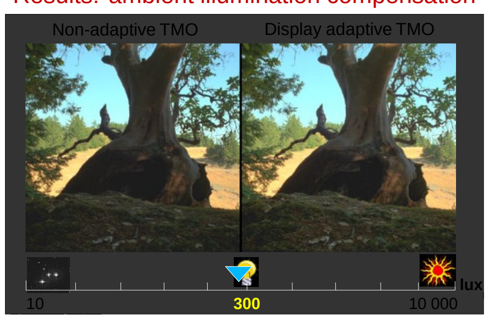
Display adaptive tone-mapping



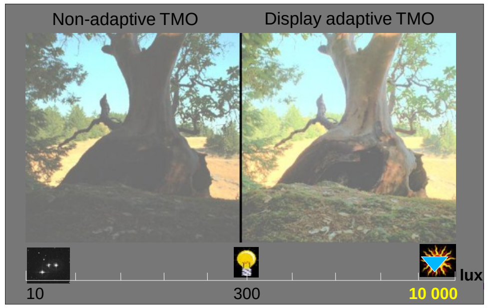
Results: ambient illumination compensation



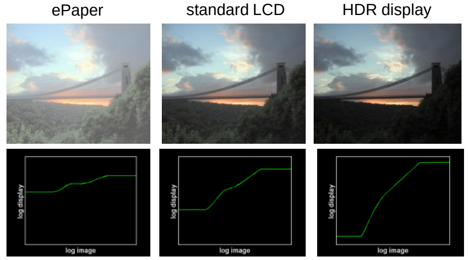
Results: ambient illumination compensation



Results: ambient illumination compensation

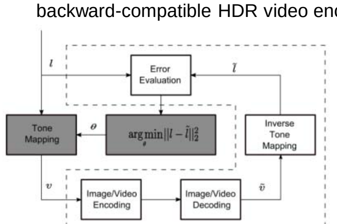


Results: display contrast



Tone-mapping for video compression

- Find the tone-curve that minimizes distortion in a backward-compatible HDR video encoding



Closed-form solution:

$$s_k = \frac{v_{max} \cdot p_k^{1/3}}{\delta \cdot \sum_{k=1}^{N} p_k^{1/3}}$$

[Mai et al., IEEE TIP 2010]

Which tone-mapping to choose?

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

1. Think what is the target application
   - and thus the goal of your tone-mapping
2. Consider which tone-mapping approach(es) will deliver that goal

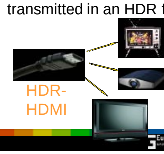
Future of tone-mapping

Tone-mapping of today

- Built into cameras
- Assumes that all displays are the same

Tone-mapping of tomorrow

- Display tone-maps content on demand
- Depending on viewing conditions, viewer, its capabilities
- Content recorded, stored and transmitted in an HDR format



Thank you

Eurographics 2012
Cagliari, Italy May 13-18
33rd ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR COMPUTER GRAPHICS

Apparent Contrast and Brightness Enhancement

Tobias Ritschel
MPI Informatik

Motivation

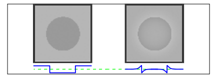
- Image display
  - Dynamic range of existing displays is limited
  - No reproduction real-world contrast/brightness
  - Good image appearance doesn't require that
- Modern tone mapping operators good at optimizing the physical contrast and luminance use

Motivation

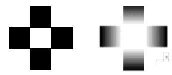
- Human preference
  - Enhanced contrast and brightness improve image appearance
- Can we still boost the contrast and brightness impression?

Human perception

- Spatial vision
- Cornsweet illusion
  Apparent contrast boost



- Glare illusion
  Apparent brightness boost



Eurographics 2012, Cagliari, Italy

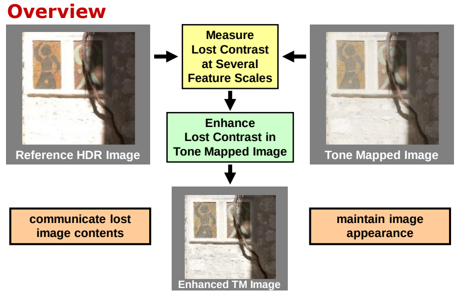
Contrast Enhancement: Motivation

- 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

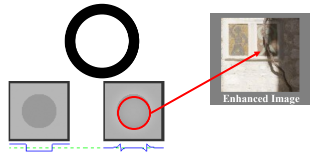
Eurographics 2012, Cagliari, Italy

Overview



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

Details of Contrast Illusion

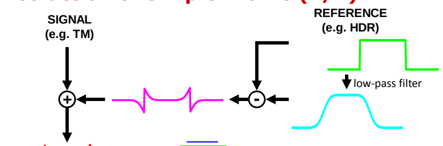
ACTUAL SIGNAL WHAT YOU SEE

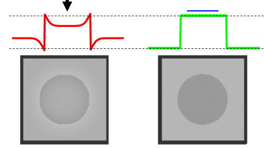
Details of Contrast Illusion

ACTUAL SIGNAL WHAT YOU SEE

- Luminance profiles cause contrast
- Properties:
  - Shape matches shape of the enhanced feature
  - Amplitude defines the perceived contrast
  - Noise (texture) does not cancel the illusion
  - Profiles should not be discernible

Construction of Simple Profile (1/2)

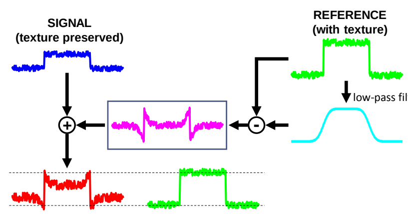




- Profile from low-pass filtered reference
- Size and amplitude adjusted manually
- This is **unsharp masking**

Krawczyk et al. EG2007

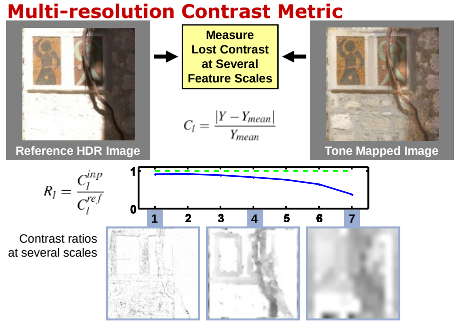
Construction of Simple Profile (2/2)



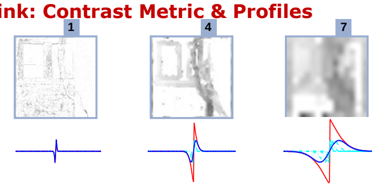
- Well preserved signal is exaggerated by unsharp masking

Krawczyk et al. EG2007

Multi-resolution Contrast Metric



Link: Contrast Metric & Profiles



- Contrast defines the sub-band amplitude
- Contrast for larger scales appears also on smaller scales
  - the full profile is always reconstructed (red)
- Scale of contrast defines the profile size

Eurographics 2012, Cagliari, Italy

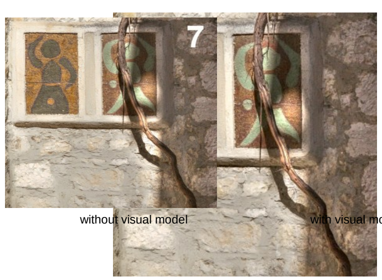
Adaptive Countershading


final contrast restoration

progress of restoration

Eurographics 2012, Cagliari, Italy

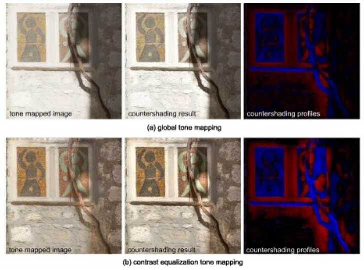
Adaptive Countershading



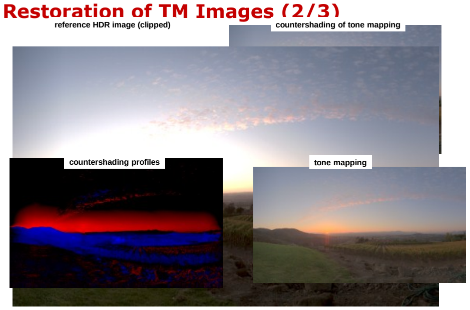
without visual model

with visual model

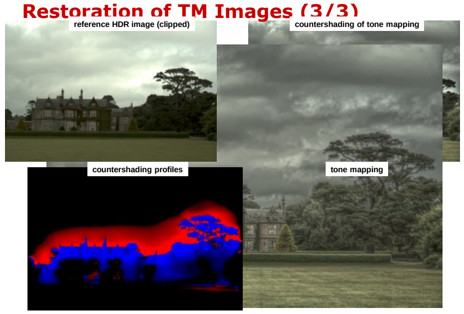
Restoration of TM Images (1/3)



Restoration of TM Images (2/3)



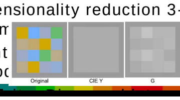
Restoration of TM Images (3/3)



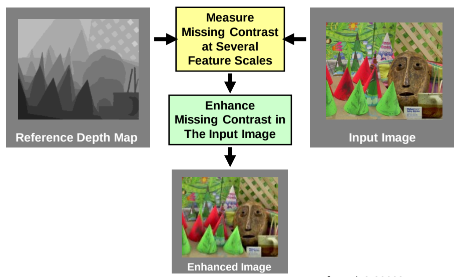
Countershading Variants



- Traditional countershading:
  - Performed in the achromatic cha… …nce perceived luminance contrast
  - Cross-modal approach:
  - Use depth signal to derive counte… …rofile
  - Countershading over chromatic channels enhances the overall image cont…
  - Color2Grey:
  - Dimensionality reduction 3->1: inform…
  - Count… …omatic channel used to reproc… …ntrast

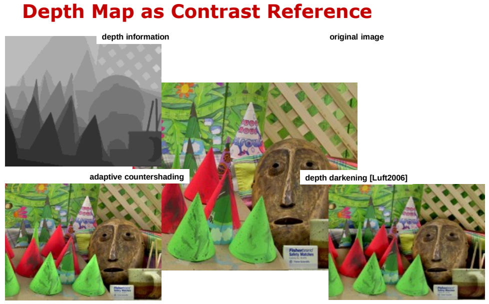


Purpose: Contrast Restoration



Luft et al. SIG2008

Depth Map as Contrast Reference



Luft et al. SIG2008

Colourfulness Countershading

- promotes FG/BG separation
- creates impression of greater dynamic range
- increases impression of depth

Smith et al. EG2006

Countershading Results (original)

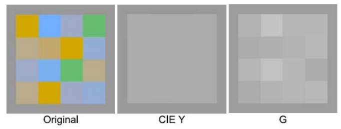
Smith et al. EG2006

Countershading Results (chroma enhancement)

Smith et al. EG2006

Color2Grey Application

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



Eurographics 2012, Cagliari, Italy

Color2Grey Application

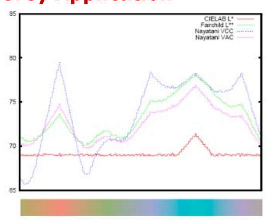


Figure 1: *Lightness values from various H-K effect predictors applied to a spectrum of isoluminant colours, compared to CIE L^*.*

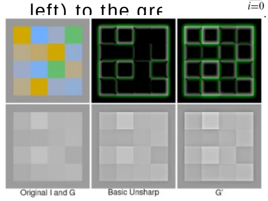
Smith et al. EG2008

Eurographics 2012, Cagliari, Italy

Color2Grey Application

- G′L* : The effe... resolution countershadin... (upper-left) to the gre... lower-left)

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

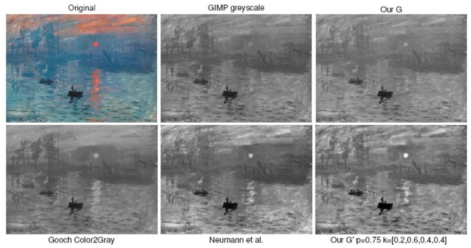


The correction is driven by contrast in chroma channels of the original image I (*upper-left*)

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

Smith et al. EG2008

Color2Grey Application

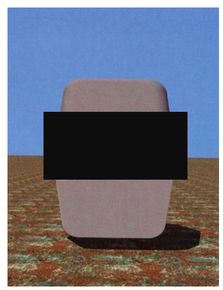


Smith et al. EG2008

Eurographics 2012, Cagliari, Italy

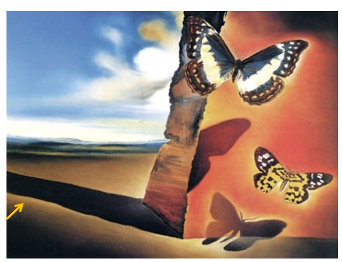
Countershading in 3D?

- Cornsweet in 3D is More plausible → Less of an artefact → Stronger → Better
- **D. Purves, A. Shimpi, R. B. Lotto** *An empirical explanation of the Cornsweet effect.* J. of Neuroscience 19, 1999



Eurographics 2012, Cagliari, Italy

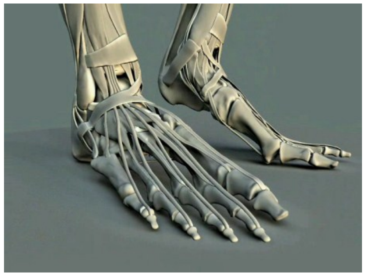
Scene-aligned Countershading



S. Dalí, *Landscape with butterflies*

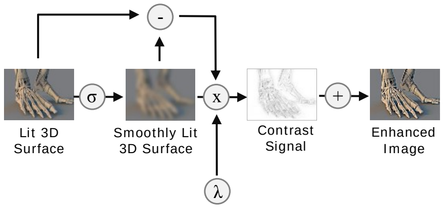
Scene-aligned Countershading

G. Seurat, *Bathers at Asnieres*



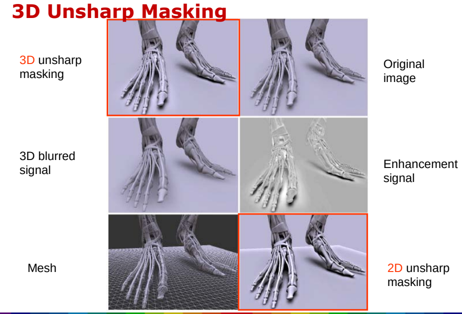
3D Unsharp Masking

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



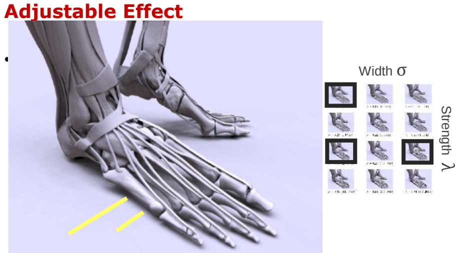
Ritschel et al. SIG2008

3D Unsharp Masking



Eurographics 2012, Cagliari, Italy

Adjustable Effect

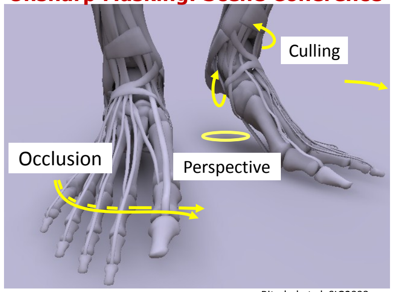


Eurographics 2012, Cagliari, Italy

2D vs. 3D Unsharp Masking Comparison

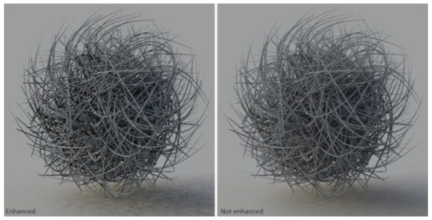
	2D	3D
Signal	Image	Lit Surface
Smoothing	(Gaussian) Image Blur	Laplacian Surface Blur
Representation	Pixels	Lit vertices and pixels
Smoothness σ	Image distance	Geodesic world distance
Strength λ	Factor	Factor

3D Unsharp Masking: Scene Coherence



Ritschel et al. SIG2008

Complex Mesh

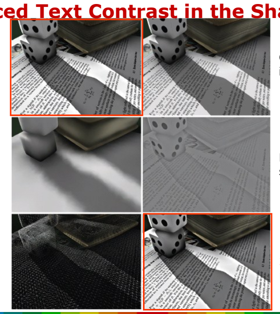


3D unsharp masked rendering

Original rendering

Ritschel et al. SIG2008

Enhanced Text Contrast in the Shadow



3D unsharp masking

Original image

3D blurred signal

Enhancement signal

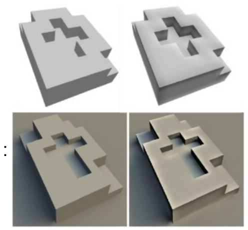
Mesh

2D unsharp masking

Results – Legibility

Normal Enhancement

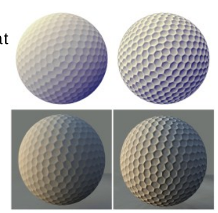
- Only geometric term
  - Shadows ?
  - Hightlights ?
  - Reflectance ?
- Vertex resolution
- 3D unsharp masking: Pixel resolution



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

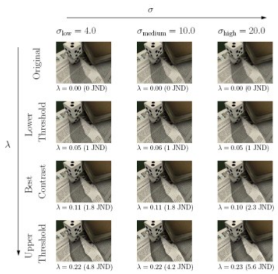
Specular shading

Study

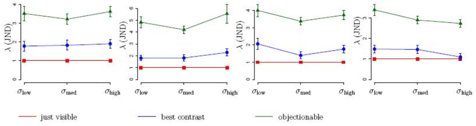
Ihrke et al. SPIE2009

- Goals
  - Find suitable settings
  - See limitations
  - Rank preference
- Method of adjustments
  - Strength λ: adjustable
  - Fixed width σ: low, medium, high
  - 4 scenes,15 participants
  - Task: Find such λ that:
    - Added enhancement is just noticeable
    - Added enhancement becomes objectionable
    - Image appearance is preferred

Results



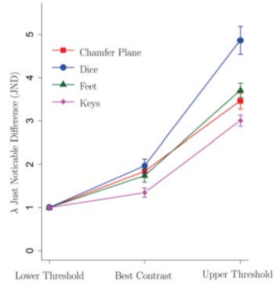
Results



Ihrke et al. SPIE2009

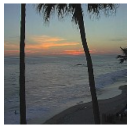
Results

- 2 JND
  - preferred
- 4 JND
  - objectionable

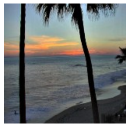


Ihrke et al. SPIE2009

Countershading parameter effect



Original

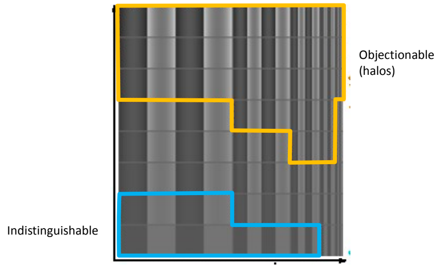


Enhancement

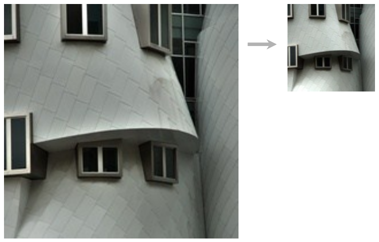
Halo

Unsharp masking, countershading and haloes: Enhancements or artifacts?
M.Trentacoste, R. Mantiuk, W. Heidrich, F. Dufrot
Eurographics 2012

Model of acceptable countershading

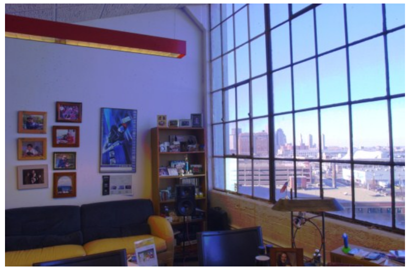


Applications: Image Resizing



Applications: Viewer-adaptive display

Applications: Tone-mapping



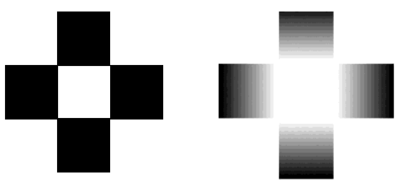
Eurographics 2012, Cagliari, Italy

Summary

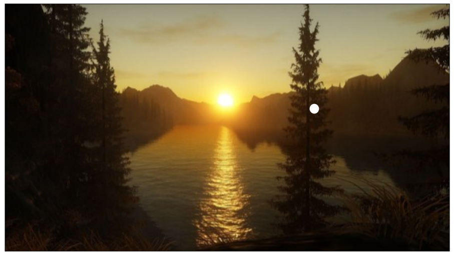
- Better communicate image contents with a minimal change to image appearance
- Application of Cornsweet illusion to image enhancement
  - Generalization of unsharp masking
  - Automatic enhancement given the reference data:
    - HDR image
    - depth information
    - shading in 3D scene
  - Scene consistent 3D unsharp masking leads to even stronger effects

Eurographics 2012, Cagliari, Italy

Glare Illusion [Zavagno and Caputo 2001]



Glare Illusion



"Alan Wake"© Remedy Entertainment

Eurographics 2012, Cagliari, Italy

Glare Illusion in Different Media

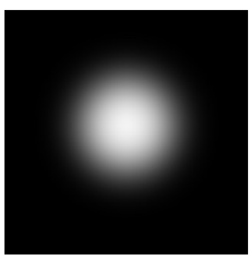

Arts

Computer games

Photography

Eurographics 2012, Cagliari, Italy

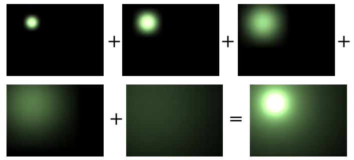
In Games



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

Yoshida *et al.* (2008)

In Games

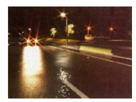


Kawase: Practical Implementation of High Dynamic Range Rendering. **Game Developer's Conference 2004**

Eurographics 2012, Cagliari, Italy

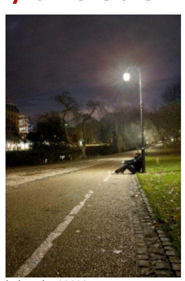
Glare in Realistic Rendering

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



Eurographics 2012, Cagliari, Italy

Dynamic Glare



Ritschel et al. EG2008

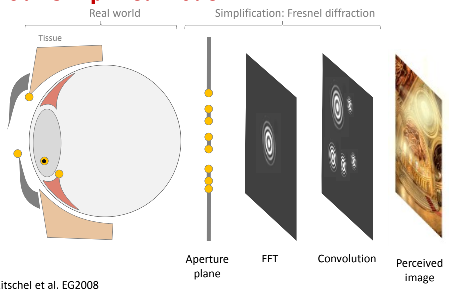
- Realism
  - Movement
  - Colors
- Required
  Model of dynamic human eye to simulate temporal glare
- Study
  Can temporal glare boost even further boost brightness?

Point Spread Function (PSF)

- **P**oint **S**pread **F**unction
- Key to glare modeling
- Describes, how a **pixel** maps to a **pattern** under an **aperture**

Ritschel et al. EG2008

Our Simplified Model



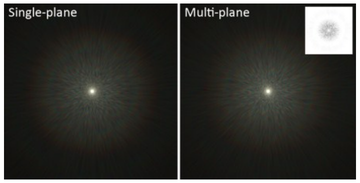
Ritschel et al. EG2008

Diffraction: Single vs. Multi Aperture Planes

$$L_i(x_i, y_i) = K \left| \mathcal{F}\{P(x_p, y_p)E(x_p, y_p)\}_{p=\frac{x_i}{\lambda d}, q=\frac{y_i}{\lambda d}} \right|^2$$

$$K = 1/(\lambda d)^2$$

$$E(x_p, y_p) = e^{i\frac{\pi}{\lambda d}(x_p^2 + y_p^2)}$$



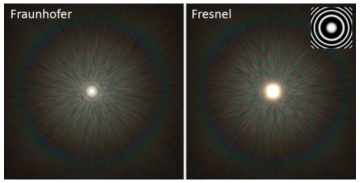
Ritschel et al. EG2008

Diffraction: Fraunhofer vs. Fresnel

$$L_i(x_i, y_i) = K \left| \mathcal{F}\{P(x_p, y_p)E(x_p, y_p)\}_{p=\frac{x_i}{\lambda d}, q=\frac{y_i}{\lambda d}} \right|^2$$

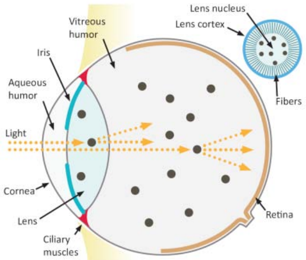
$$K = 1/(\lambda d)^2$$

$$E(x_p, y_p) = e^{i\frac{\pi}{\lambda d}(x_p^2 + y_p^2)}$$



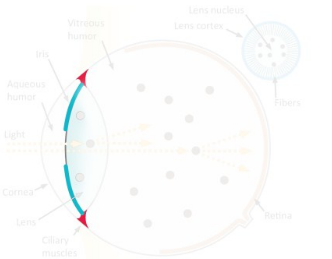
Ritschel et al. EG2008

Temporal Glare Pipeline



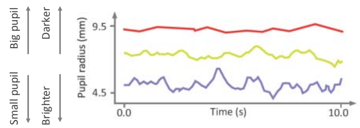
Ritschel et al. EG2008

Aperture: Pupil



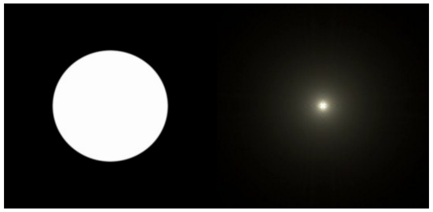
Aperture: Pupil

- Adaptation
- Can convert HDR image into pupil size
- Pupillary hippus



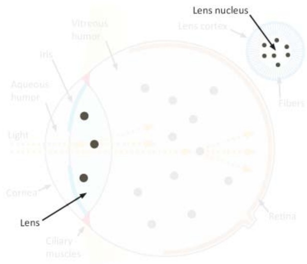
Eurographics 2012, Cagliari, Italy

Aperture: Pupil

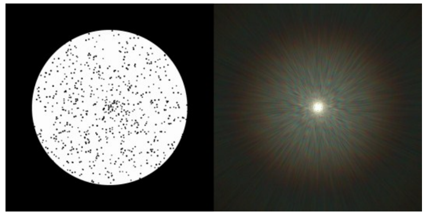


Eurographics 2012, Cagliari, Italy

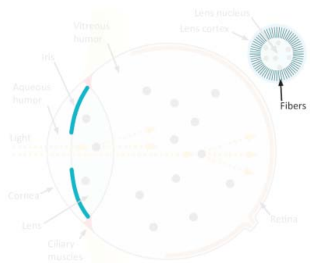
Aperture: Lens



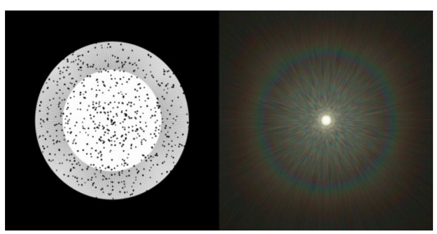
Aperture: Lens



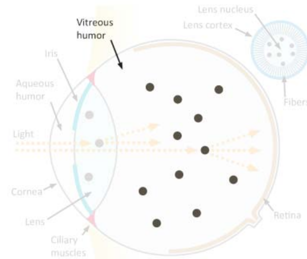
Aperture: Gratings / Lens fibers



Aperture: Gratings / Lens fibers

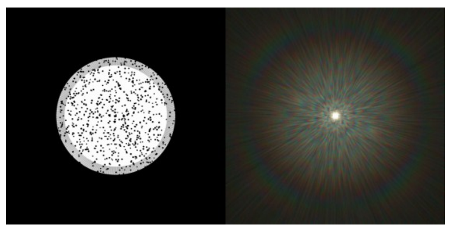


Aperture: Vitreous Humor



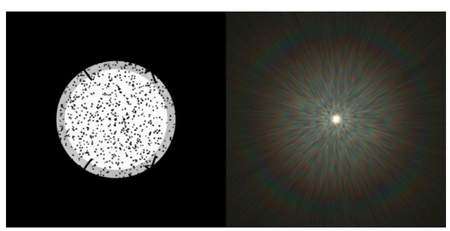
Eurographics 2012, Cagliari, Italy

Aperture: Vitreous Humor



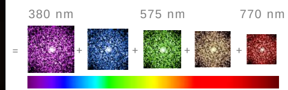
Eurographics 2012, Cagliari, Italy

Aperture: Eyelashes (optional)



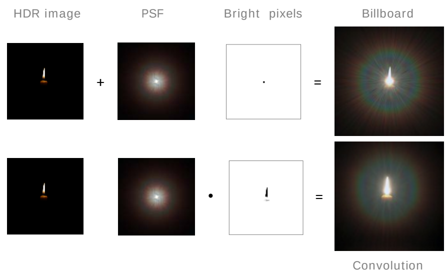
Chromatic Blur

- Compute one wavelength - Get others for free!

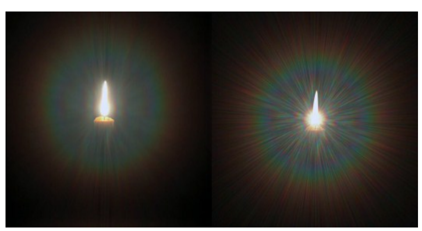


$$\begin{aligned} F_s(\mathbf{x}) &= \sum_{i=0}^{n-1} s(\lambda_i) F_{575\,\mathrm{nm}}(\mathbf{x}_i) \\ \lambda_i &= 380\,\mathrm{nm} + i\,\frac{770\,\mathrm{nm} - 380\,\mathrm{nm}}{n} \\ \mathbf{x}_i &= \mathbf{x}\,\frac{575\,\mathrm{nm}}{\lambda_i} \quad . \end{aligned}$$

Convolution

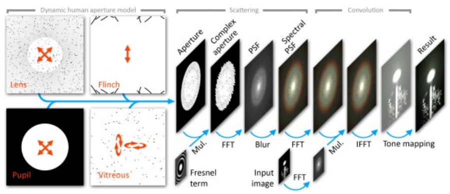


Convolution



Convolution Billboard

Temporal Glare Pipeline



Eurographics 2012, Cagliari, Italy

Eurographics 2012, Cagliari, Italy

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 subjects20 minutes per person

Yoshida et al. APGV2008

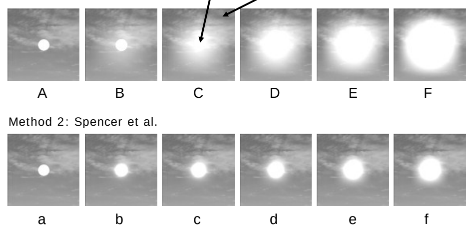
Stimuli

220 cd/m^2

150 cd/m^2

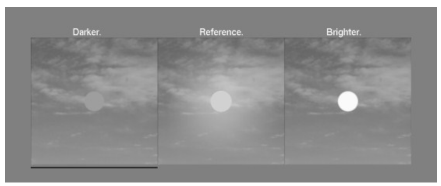
Method 1: Gaussian

Method 2: Spencer et al.



Yoshida et al. APGV2008

Perceptual Experiment

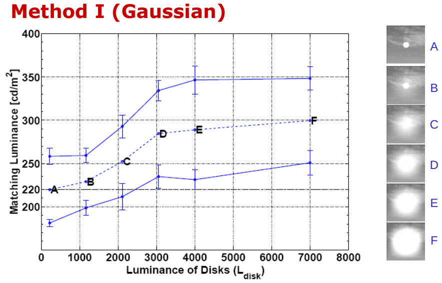


Task:

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

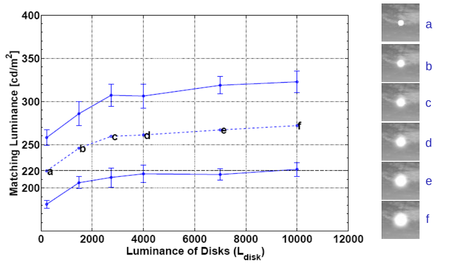
Yoshida et al. APGV2008

Method I (Gaussian)



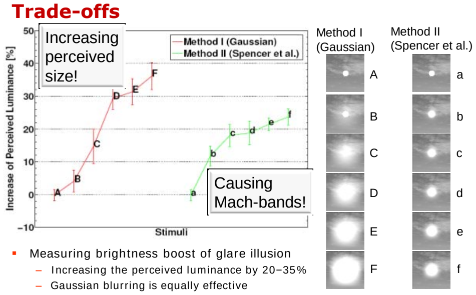
Yoshida et al. APGV2008

Method II (Spencer et al.)



Yoshida et al. APGV2008

Trade-offs



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

Yoshida et al. APGV2008

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
- Other temporal low-level eye physics like
  - Floaters
  - Local adaptation ("After images")

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

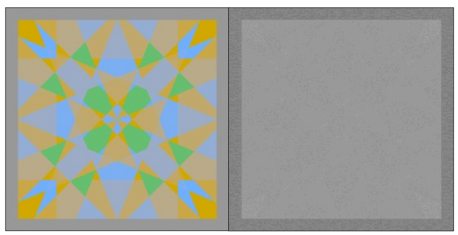
Acknowledgements

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

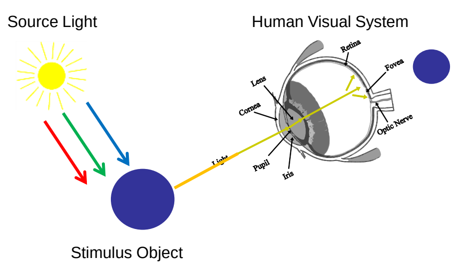
Retargeting Color Content: Color Issues in Tone Mapping

Alessandro Artusi
Ramon Cajal Fellow

Introduction to Color



What is Color?



Quantifying Color

$I(\lambda)$ SPD of the light

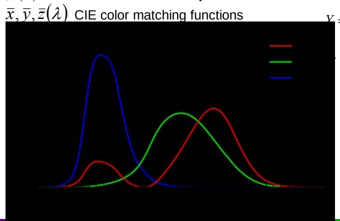
$\rho(\lambda)$ Reflectance of the object

$\bar{x},\bar{y},\bar{z}(\lambda)$ CIE color matching functions

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

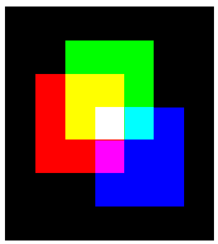
$$Y = \int_0^\infty I(\lambda)\rho(\lambda)\bar{y}(\lambda)d\lambda$$

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

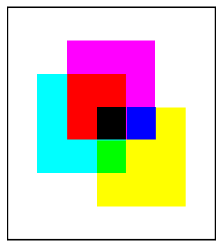


How Color is Produced?

Additive

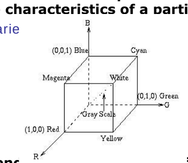


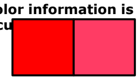
Subtractive



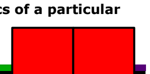
Color Space

- **Device dependent: the description of color information is related to the characteristics of a particu**
  - Set of primarie
  - Technology





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

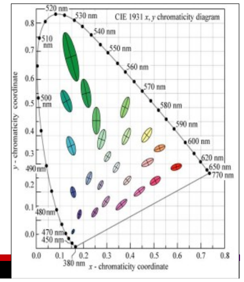


Chromaticity Diagram and MacAdam's Ellipses

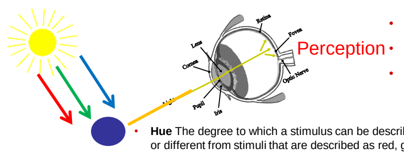
- **MacAdam's Ellipses**
  - contains all colors which are indistinguishable to an human observer from the color at the center of the ellipse
  - the contour of the ellipse represents the just noticeable differences of chromaticity

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

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



Color Attributes by the CIE



- **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

- **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

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

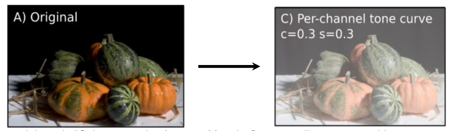
Color in High Dynamic Range

- **Saturation Control (Thumblin and Turk 1999)**

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

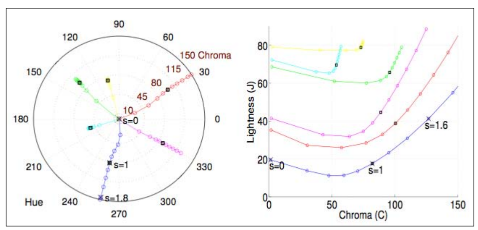
s Saturation Parameter
c Contrast Compression

Under-saturated colors for S=C.



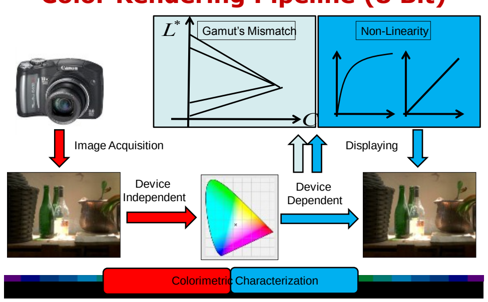
Mantiuk et al.. "Color Correction for Tone Mapping", Proceedings Eurographics 2009.

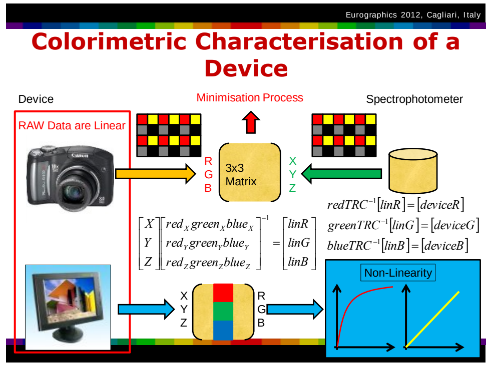
Color in High Dynamic Range

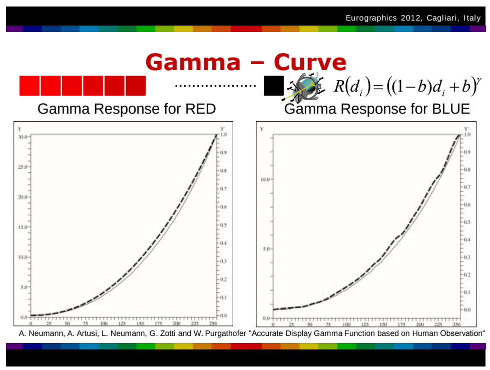


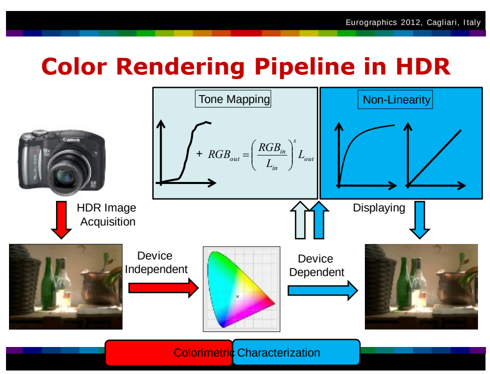
Mantiuk et al.. "Color Correction for Tone Mapping", Proceedings Eurographics 2009.

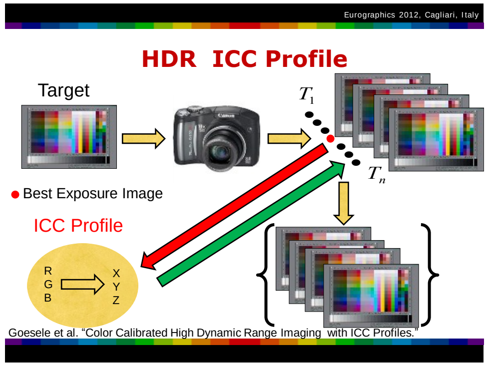
Color Rendering Pipeline (8 Bit)

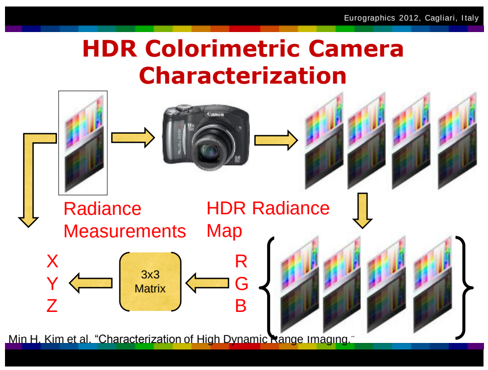


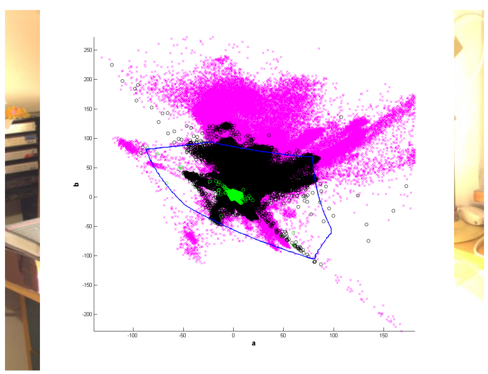

Eurographics 2012, Cagliari, Italy
Colorimetric Characterisation of a Device
Device
Minimisation Process
Spectrophotometer
RAW Data are Linear
R
G
B
3x3 Matrix
X
Y
Z
redTRC^-1[linR] = [deviceR]
greenTRC^-1[linG] = [deviceG]
blueTRC^-1[linB] = [deviceB]
Non-Linearity
X
Y
Z
R
G
B


Eurographics 2012, Cagliari, Italy
Gamma – Curve
R(d_i) = ((1−b)d_i + b)^γ
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"

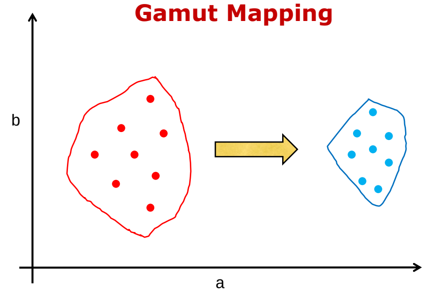

Eurographics 2012, Cagliari, Italy
Color Rendering Pipeline in HDR
Tone Mapping
Non-Linearity
HDR Image Acquisition
Displaying
Device Independent
Device Dependent
Colorimetric Characterization


Eurographics 2012, Cagliari, Italy
HDR ICC Profile
Target
T_1
T_n
Best Exposure Image
ICC Profile
R
G
B
X
Y
Z
Goesele et al. "Color Calibrated High Dynamic Range Imaging with ICC Profiles."

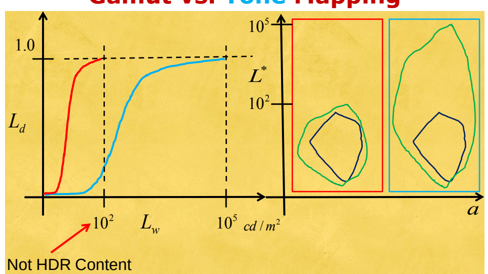

Eurographics 2012, Cagliari, Italy
HDR Colorimetric Camera Characterization
Radiance Measurements
HDR Radiance Map
X
Y
Z
3x3 Matrix
R
G
B
Min H. Kim et al. "Characterization of High Dynamic Range Imaging."



Gamut Mapping

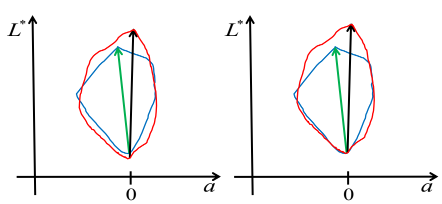


Gamut vs. Tone Mapping



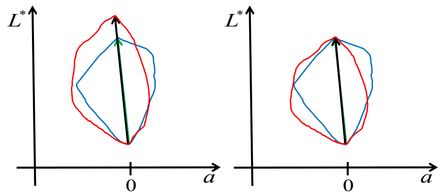
Gamut Mapping Aims (CS)

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



Gamut Mapping Aims (CS)

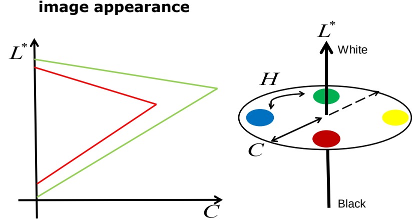
- Gray axes alignment, mapping white to white and black to black



Eurographics 2012, Cagliari, Italy

Gamut Mapping Aims (CS)

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

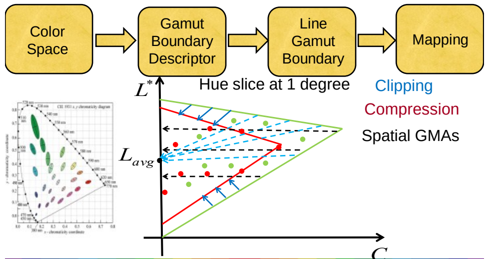


Eurographics 2012, Cagliari, Italy

Gamut Mapping Aims (CS)

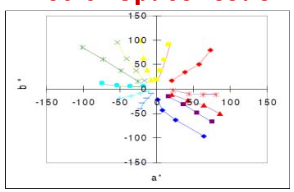
- **Limiting out of gamut colours**
  - Soft clipping can be afterwards adopted to eliminate these extremes
- **Increase Image saturation**
  - Destination gamut has reduced saturation
  - Helps maintaining the original chroma differences of the input Image

Gamut Mapping Pipeline



Eurographics 2012, Cagliari, Italy

Color Space Issue



- **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"

Eurographics 2012, Cagliari, Italy

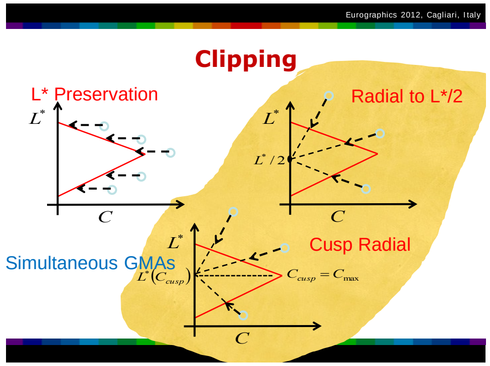
Point-wise Gamut Mapping Techniques

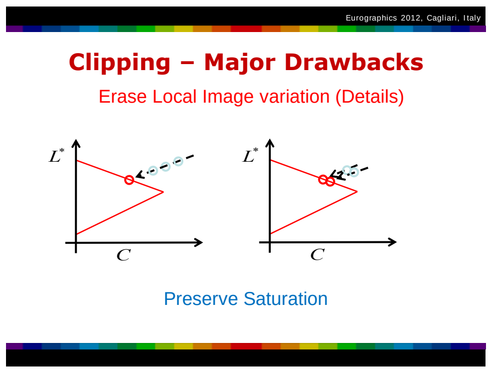
- **Clipping**

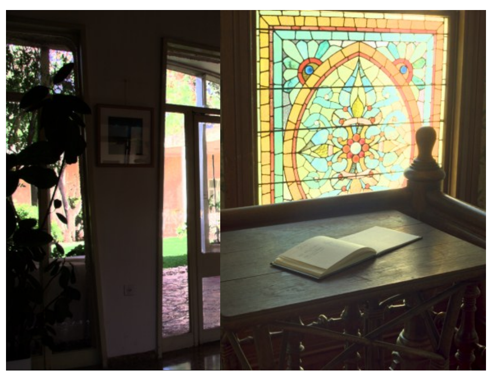
It changes colours which are outside of the destination gamut,

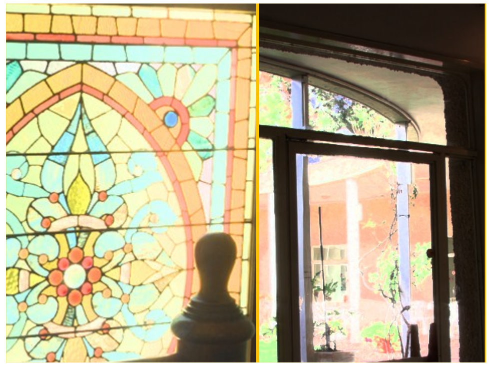
mapping them on the boundaries of the destination gamut

- **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**)


Eurographics 2012, Cagliari, Italy
Clipping
L* Preservation
Radial to L*/2
Cusp Radial
Simultaneous GMAs
L*
C
L*/2
L*(C_cusp)
C_cusp = C_max


Eurographics 2012, Cagliari, Italy
Clipping – Major Drawbacks
Erase Local Image variation (Details)
L*
C
L*
C
Preserve Saturation





Point-wise Gamut Mapping Techniques

- **Compression**

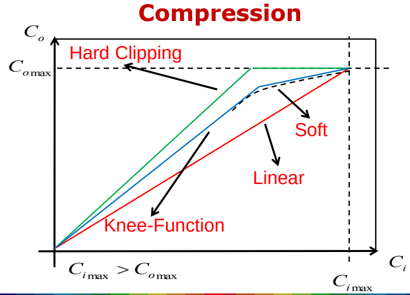
It makes changes to all the colors of the source gamut to be
accommodated into the destination gamut .

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

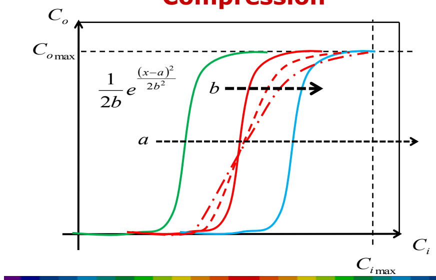
- **Parametric**

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

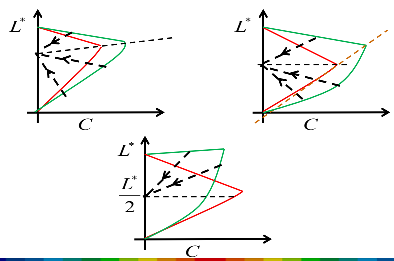


Compression



Eurographics 2012, Cagliari, Italy

Parametric



Eurographics 2012, Cagliari, Italy

Preservation of Spatial Details

- **Optimization**

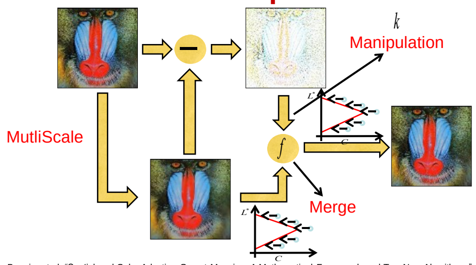
Making use of Human Visual System Models minimize the perceived differences between the input and output image.

- **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."

Mantiuk et al. "Color Correction for Tone Mapping

Automatic estimation of desaturation (s) factor in function of contrast compression (c) (non-linear color co

$$C_{out} = \left(\frac{C_{in}}{L_{in}}\right)^s L_{out} \qquad s(c) = \frac{(1+k_1)\, c^{k_2}}{1+k_1\, c^{k_2}}$$

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 (luminanc preserving solution)

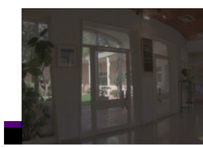
Conclusions

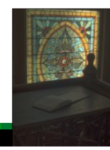
- **Works on high dynamic range imaging have mostly operated on luminance (lightness) information**
  - some works start to appear proposing solution for color saturation, acquisition of colorimetric correct high dynamic range color values, and color appereance
- **In Color Science a lot of works have 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

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 18, 30 and 41) Courtesy of Ela Sikudova**
- **HDR Image s(skide 18) Martin Cadik**





Eurographics 2012
Cagliari, Italy
May 13-18
33rd ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR COMPUTER GRAPHICS

Inverse/Reverse Tone Mapping

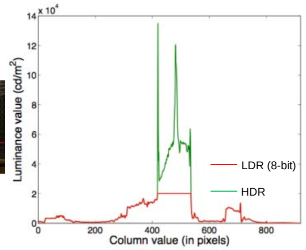
Dr. Francesco Banterle
Visual Computing Lab, ISTI-CNR, Italy

francesco.banterle@isti.cnr.it

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?

- 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:

$$g(I) = \underbrace{\mathbf{D}_i^{w \times h \times c}}_{\text{LDR}} \rightarrow \underbrace{\mathbf{D}_o^{w \times h \times c}}_{\text{HDR}},$$

- 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:

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

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

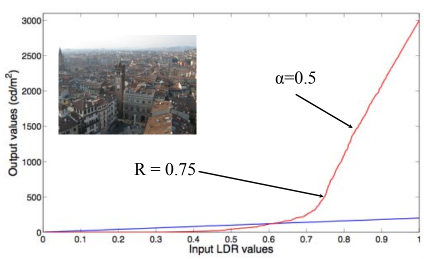

Original LDR EM


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:

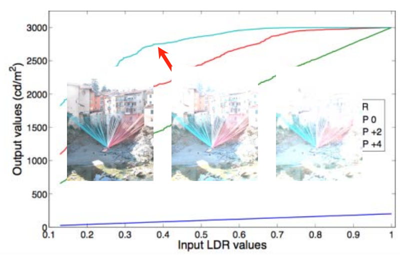
$$L_w(\mathbf{x}) = k\left(\frac{L_d(\mathbf{x}) - L_{d,\,\min}}{L_{d,\,\max} - L_{d,\,\min}}\right)^{\gamma}$$

- 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:

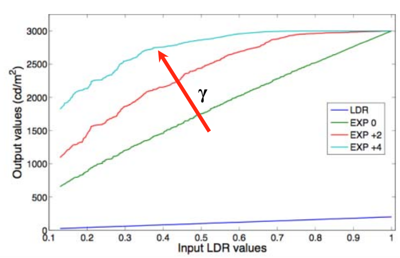
$$L_w(\mathbf{x}) = L_d(\mathbf{x})^{\gamma} \qquad \gamma = 10.44k - 6.282$$

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

Global Methods (IV)



Global Methods (IV)



Classification Methods: Highlights Reproduction on HDR Monitors (I)

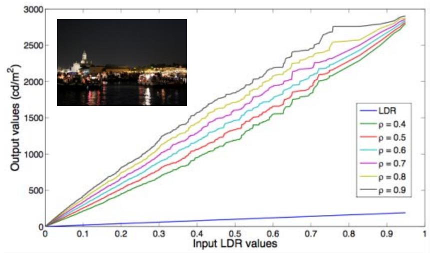
- Meylan et al. [MDDS06, MDS07] present a classification approach:
  - Expand highlights and specular surfaces (ω>0)
  - ω is computed using robust thresholding
  - Expansion using a two-scale model:

$$L_w(\mathbf{x}) = f(L_d(\mathbf{x})) = \begin{cases} s_1 L_d(\mathbf{x}) & \text{if } L_d(\mathbf{x}) \leq \omega \\ s_1\omega + s_2(L_d(\mathbf{x}) - \omega) & \text{otherwise} \end{cases}$$

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

  - To avoid contouring low-pass filtering on expanded regions

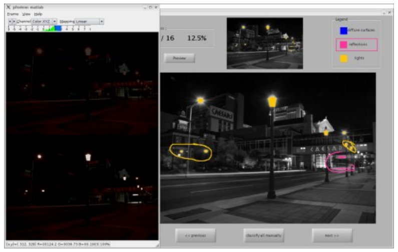
Classification Methods: Highlights Reproduction on HDR Monitors (II)



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 ⇒ 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

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):

$$p = \left(\frac{\exp(v\sin(\pi\frac{z-1}{16}))-1}{\exp(v)-1}\right)^{\frac{1.0}{1.2}} \qquad v = 5.25 \quad z \in [0,9]$$

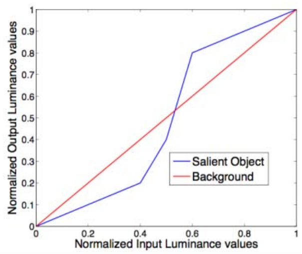
  - **Labeling**:
    - salient objects and background discrimination using different techniques:
      - learning-based saliency detection (Liu et al. [LSZ*07])
      - *saliency cuts* (Fu et al.[FCLL08])
    - Different Labels ⇒ Different RE functions

Classification Methods: Selective Reverse Tone Mapping (II)

Input

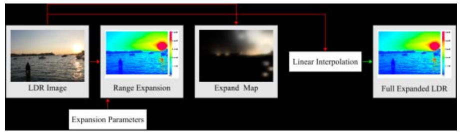
Auto-Labeling

Binary Mask



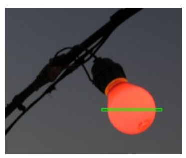
Expand Maps Methods: Non-Linear Expansion using Expand Maps (I)

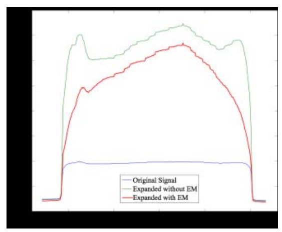
- Banterle et al. [BLDC06,BLDBC07,BLDC08,B09] presented a general and real-time framework:
  - **Range Expansion**: non-linear (inverting an TMO; other functions)
  - **Expand Map**: sampling+density estimation+cross bilateral (avoiding contouring and compression artifacts)



Eurographics 2012, Cagliari, Italy

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





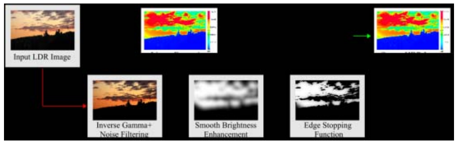
Eurographics 2012, Cagliari, Italy

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



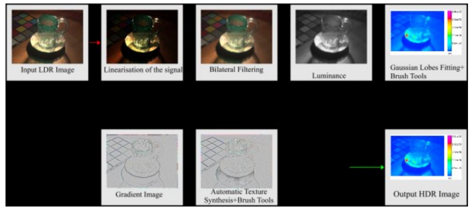
Expand Maps Methods: LDR2HDR (II)

- 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

Evaluation: Why validation

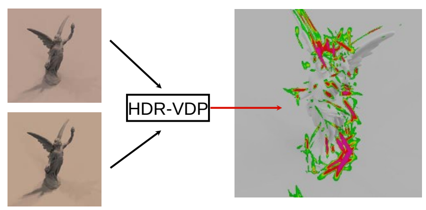
- Need to evaluate different expansion methods against a "*ground truth*". Why?
  - To understand weak features or drawbacks
  - To understand important features
- rTMO/iTMO techniques do not generate exact luminance values
- Evaluation:
  - Perceptual Image Metrics: not exact comparison as in the PSNR,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

Eurographics 2012, Cagliari, Italy

Evaluation: Perceptual Image Metrics (II)



Lucy model is courtesy of the Stanford 3D Scanning Repository

Eurographics 2012, Cagliari, Italy

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, Shai Avidan and Susana Castillo)

The Retargeting Problem

Common solutions

Cropping

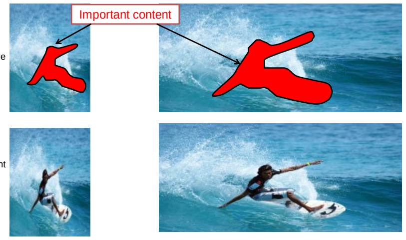
Homogeneous Scaling

Letterboxing

Common solutions

Content-aware

Ignore content



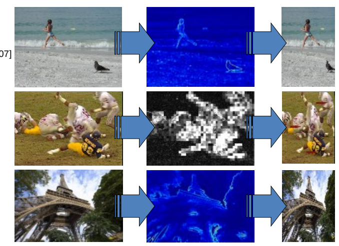
Eurographics 2012, Cagliari, Italy

Common solutions

SC
[Avidan+Shamir 2007]

WARP
[Wolf et al. 2007]

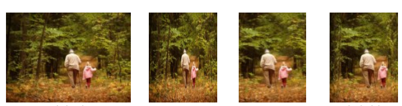
SNS
[Wang et al. 2008]



Eurographics 2012, Cagliari, Italy

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

Scaling

[Avidar & Shamir 2007]

Visual Media Retargeting: Seams

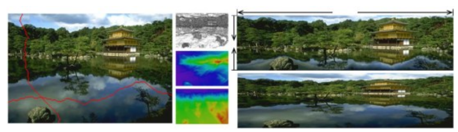
Insert & remove seams

Scaling

[Avidar & Shamir 2007]

Visual Media Retargeting: Energy Concept

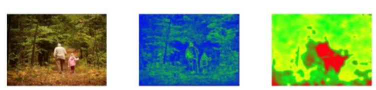
1. Define an energy function **E(I)** (interest, importance, saliency…)	⇨	2. Use some operator(s) to change the image **I**



[Avidar & Shamir 2007]

Visual Media Retargeting: Energy & Saliency

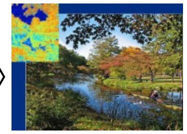
- Magnitude of gradients (simple)
- Saliency (e.g. Itty's method) – multi-res



[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]

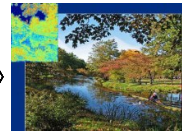
Different energy functions

- 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

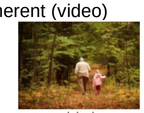
- Crop s.t. important (salient) parts remain
- Use domain-specific tools, such as face detector, gaze estimation… [DeCarlo and Santella 2002; Viola and Jones 2004]

original

crop

Simple operators: scaling

- Can 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

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.*

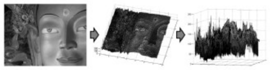


Figure 3: *A digital image as a sampling of a continuous function.*

[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*
- **…and the list goes on**

[Shamir and Sorkine 2009]

Continuous approaches

- **Feature-aware texturing**
  - 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*
- **…and the list goes on**

[Shamir and Sorkine 2009]

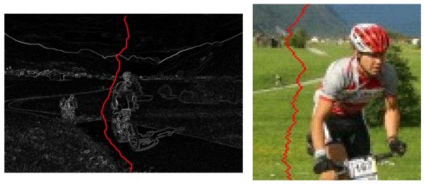
Eurographics 2012, Cagliari, Italy

Discrete example: Seam carving

[Rubinstein, Avidan and Shamir 2007]

Eurographics 2012, Cagliari, Italy

Seam carving



[Rubinstein, Avidan and Shamir 2007]

Seam carving

[Rubinstein, Avidan and Shamir 2007]

Eurographics 2012, Cagliari, Italy

Seam carving

[Rubinstein, Avidan and Shamir 2007]

Eurographics

Eurographics 2012, Cagliari, Italy

Seam carving: problems

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

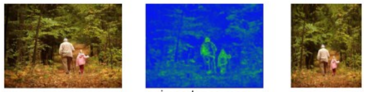
direct SC indirect SC

[Rubinstein, Avidan and Shamir 2007]

Eurographics

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



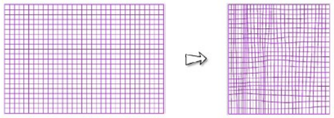
importance map

[Wang, Tai, Sorkine and Lee 2008]

Eurographics 2012, Cagliari, Italy

Continuous example: Warping

- Grid mesh, preserve the shape of the important quads



- Optimize the location of mesh vertices, interpolate image

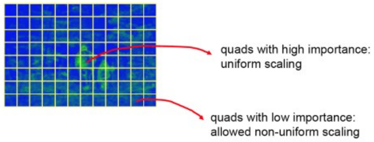
[Wang, Tai, Sorkine and Lee 2008]

Eurographics

Eurographics 2012, Cagliari, Italy

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]

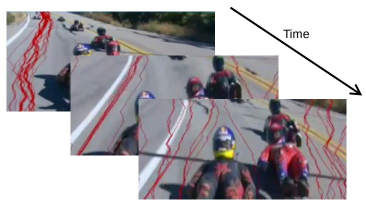
Eurographics

Video?

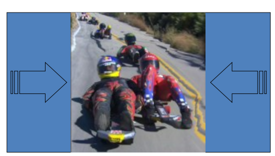
- Naïve… every frame by itself

- Improved seam carving for video resizing [SIGGRAPH 2008]

[Shamir and Sorkine 2010]

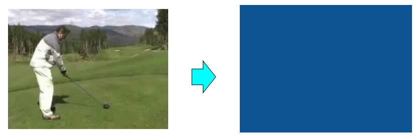


Too jittery!



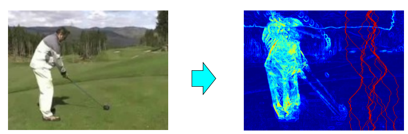
Slightly less naïve…

Reduction of the video problem to image seam carving by using projection of maximum energy through time:

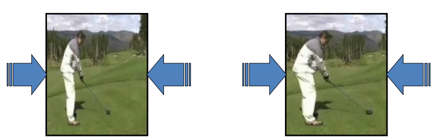


Eurographics 2012, Cagliari, Italy

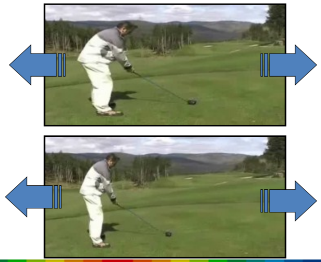
Reduction of the video problem to image seam carving by using projection of maximum energy through time:



Eurographics 2012, Cagliari, Italy



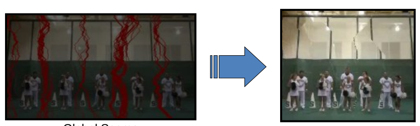
Rescale



Retarget

Problems?

- More complex scenes:
  - Object movement
  - Camera movement



Global Seams

More Complex Scenes

- More complex scenes:
  - Object movement
  - Camera movement

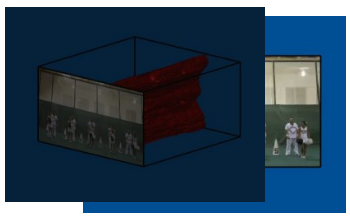
Global Seams

More Complex Scenes

- Seams should adapt and change through time!

Eurographics 2012, Cagliari, Italy

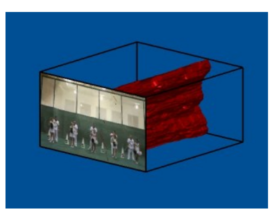
Global Solution

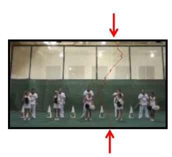


Video Cube

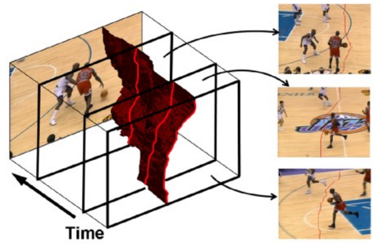
Eurographics 2012, Cagliari, Italy

One 3D Seam



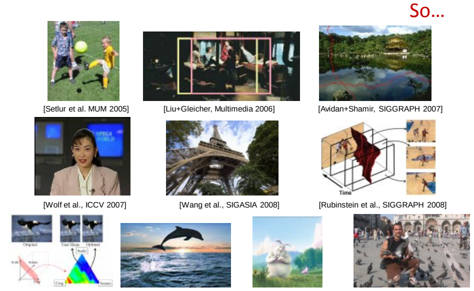


Video?



[Shamir and Sorkine 2010]

So…

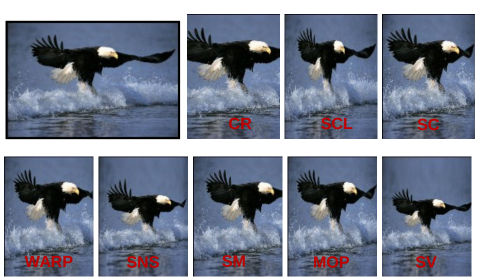


[Setlur et al. MUM 2005] [Liu+Gleicher, Multimedia 2006] [Avidan+Shamir, SIGGRAPH 2007]

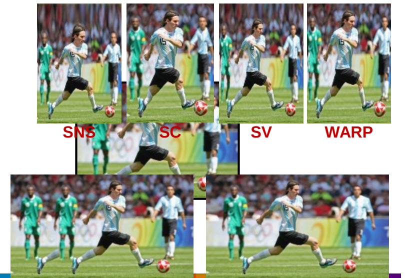
[Wolf et al., ICCV 2007] [Wang et al., SIGASIA 2008] [Rubinstein et al., SIGGRAPH 2008]

[Rubinstein et al., SIGGRAPH 2009] [Karni et al., SGP 2009] [Krähenbühl et al., SIGASIA 2009] [Pritch et al., ICCV 2009]

So…



So…

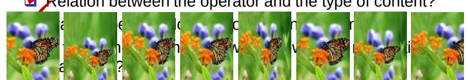


Current State of Retargeting Research

- 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?

A Comparative Study of Image Retargeting
Miki Rubinstein, Diego Gutierrez, Olga Sorkine and Arik Shamir
ACM Transactions on Graphics, Vol. 29(5) (SIGGRAPH Asia 2010)

- Benchmark and evaluation methodology for image retargeting

RetargetMe

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

- Comprehensive perceptual study and analysis of image retargeting

Goals

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


[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Eurographics 2012, Cagliari, Italy

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]

Eurographics 2012, Cagliari, Italy

Constructing the Dataset

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

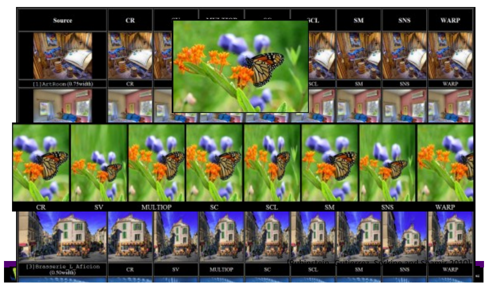
[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Retargeting Operators

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

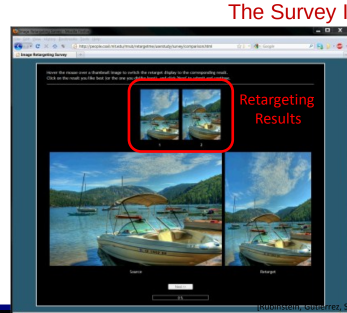
Eurographics 2012, Cagliari, Italy

Comparative Analysis



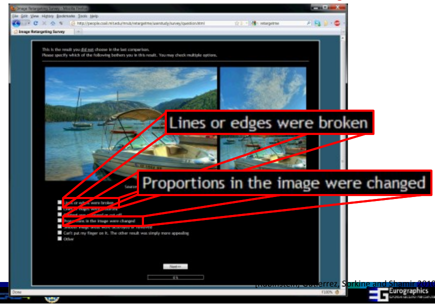
Eurographics 2012, Cagliari, Italy

The Survey Interface



[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

Additional Questions



[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

User Statistics

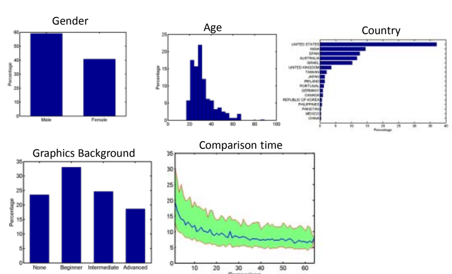
- Each participant performs **12** comparisons over **5** images
- **210** participants; **252** votes per image
  - Half amazon mechanical turk
  - Half (25 cents per completed survey)
- Average time to complete: 20 minutes

"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



User Agreement

- Similarity of votes = consensus on “good” retargeting
- *Coefficient of Agreement* [Kendall 1940]

[redacted]

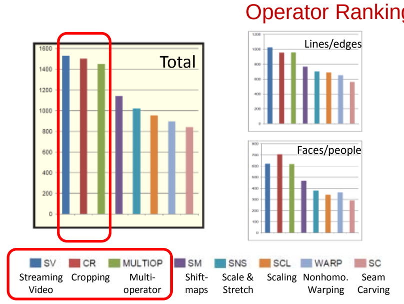
- a_{ij} = # times method i chosen over method j
- m = # participants
- t = 8 (# retargeting operators)
- [redacted]

User Agreement

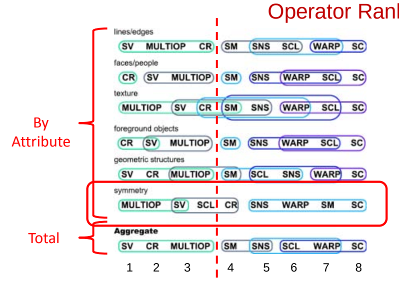
	lines/ edges	faces/ people	Texture	foreground objects	Geometric Structures	Symmetry	Total
u	0.073	**0.166**	0.070	**0.146**	0.084	**0.132**	**0.095**

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

Operator Ranking



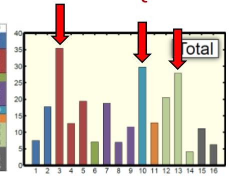
Operator Ranking

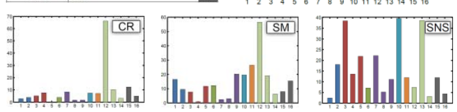


[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

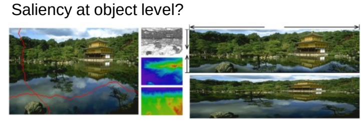
Additional Questions

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

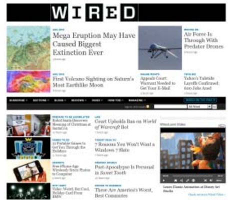




Partial Conclusion

- *(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!



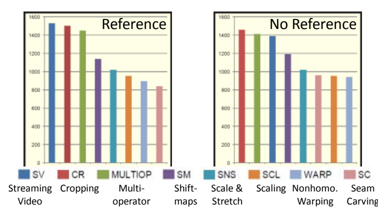
[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

“No Reference” Experiment Results

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

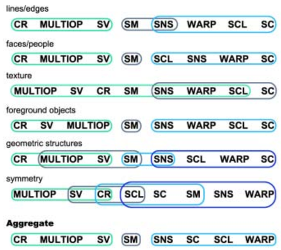
[Rubinstein, Gutierrez, Sorkine and Shamir 2010]

“No Reference” Experiment Results



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

Operator Ranking



[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]

Eurographics 2012, Cagliari, Italy

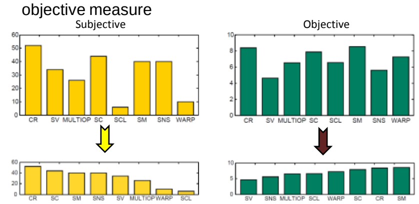
(Non-blind) Retargeting Measures

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

How to Evaluate an Objective Measure?

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



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 ☺

Can computational image distance metrics predict human retargeting perception?

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 ($-1 \leq \tau \leq 1$), 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]

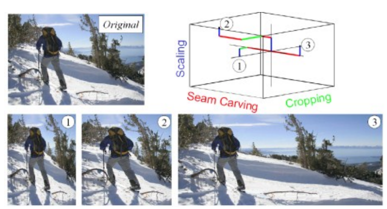
Conclusions

- 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

Image Retargeting Quality Assessment
Computer Graphics Forum, 2011, Vol. 30, No. 2, Eurographics 2011,
Yong-Jin Liu, Xi Luo, Yu-Ming Xuan, Wen-Feng Chen, Xiao-Lan Fu

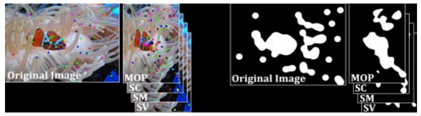
$$ColSim(C^0_{ori}, C^0_{ret}) = w_L SalSim(L^{*0}_{ori}, L^{*0}_{ret}) + w_a SalSim(a^{*0}_{ori}, a^{*0}_{ret}) + w_b SalSim(b^{*0}_{ori}, b^{*0}_{ret})$$

Conclusions

We need **video** analysis and experiments!

Using Eye-Tracking to Assess Different Image Retargeting Methods
Susana Castillo,Tilke Judd and Diego Gutierrez
Applied Perception in Graphics and Visualization 2011

Using Eye-Tracking to Assess Different Image Retargeting Methods



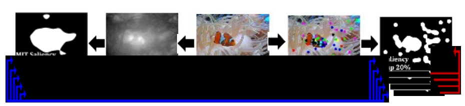
[Castillo, Judd and Gutierrez 2011]

Using Eye-Tracking to Assess Different Image Retargeting Methods
Susana Castillo,Tilke Judd and Diego Gutierrez
Applied Perception in Graphics and Visualization 2011

How to measure the relevance of a retargeting approach?

[Chamaret et al. 2010]

Overview

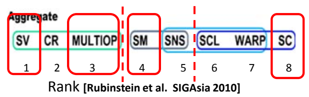


[Castillo, Judd and Gutierrez 2011]

Retargeting Operators

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

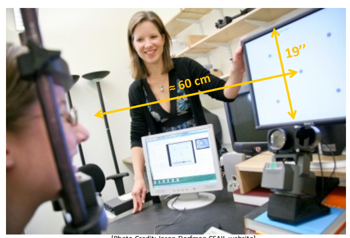
- Multi-Operator **[MULTIOP]** [Rubinstein et al. 2009]
- Streaming Video **[SV]** [Krähenbühl et al. 2009]



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

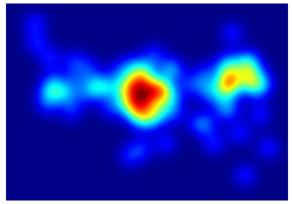
Contextual guidance of eye movements and attention in real-world scenes: The role of global features on object search **[Torralba et al. 2006]**

Fixations for 7 users

[Castillo, Judd and Gutierrez 2011]

Eye tracking data

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



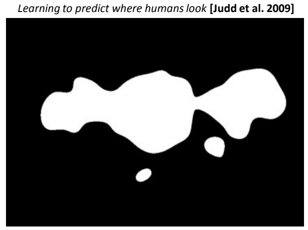
Average fixation locations / continuous saliency map

[Castillo, Judd and Gutierrez 2011]

Eurographics 2012, Cagliari, Italy

Eye tracking data

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

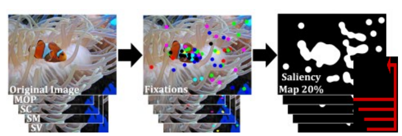


Top 20% salient locations

[Castillo, Judd and Gutierrez 2011]

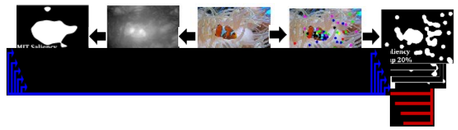
Eurographics 2012, Cagliari, Italy

Eye tracking data



[Castillo, Judd and Gutierrez 2011]

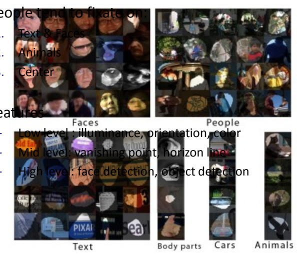
Eye tracking data



[Castillo, Judd and Gutierrez 2011]

MIT Predictive Model of Saliency

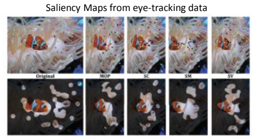
- People tend to fixate on:
  1. Text & Faces
  2. Animals
  3. Center
- Features
  - Low level : illuminance, orientation, color
  - Mid level: vanishing point, horizon line
  - High level: face detection, object detection



[Judd et al. 2009]

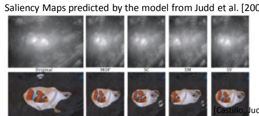
MIT Predictive Model of Saliency

Saliency Maps from eye-tracking data



80% correlation

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



[Castillo, Judd and Gutierrez 2011]

Examples and Discussion

[Castillo, Judd and Gutierrez 2011]

Eurographics 2012, Cagliari, Italy

Conclusions

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

RetargetMe

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

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

Eurographics 2012, Cagliari, Italy

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)*

Spatial Retargeting

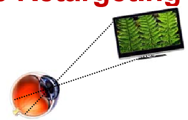
Diego Gutierrez
Universidad de Zaragoza

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

Eurographics 2012
Cagliari, Italy May 13-18
33rd ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR COMPUTER GRAPHICS

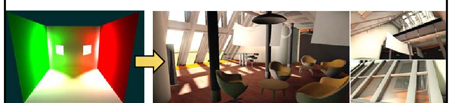
Temporal Image Retargeting

Elmar Eisemann
Telecom ParisTech / CNRS-LTCI



Introduction

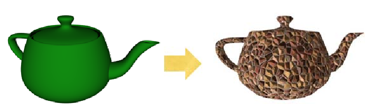
- **Example:**
  2002 (2 hrs) vs. 2012 (13 Hz)



Eurographics 2012, Cagliari, Italy

First Observation

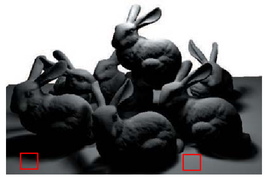
- Today's main cost is **shading**



- **Goal: Reduce shading costs**

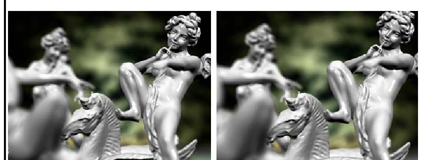
Second Observation

- **Many pixel computations are similar over space and time**



Third Observation

Frames do not differ much…

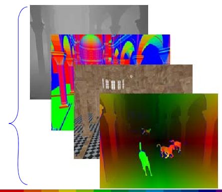


Consequence

- **Given:**
  - Pixel color determination is expensive
  - Computations can be spread over space and time
  - Frames are similar
- **Reuse pixel information over space and time to reduce shading costs**

Remember?

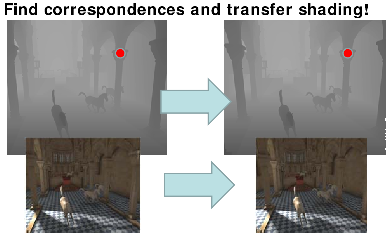
- Observation: shading correlates with geometry
- World information behind pixel is for "free"
  - Depth (position)
  - Normals
  - Materials, Textures
  - Geometric motion flow



Eurographics 2012, Cagliari, Italy

Why does rendering of depth, motion & co. help?

- Find correspondences and transfer shading!



Eurographics 2012, Cagliari, Italy

Forward Reprojection

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

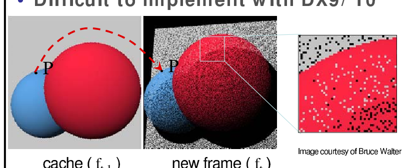
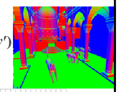


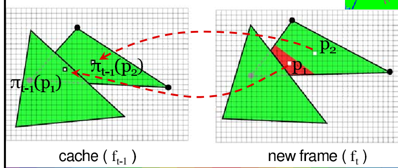
Image courtesy of Bruce Walter

Eurographics 2012, Cagliari, Italy

Reverse Reprojection [Nehab 06/07, Scherzer 07]

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

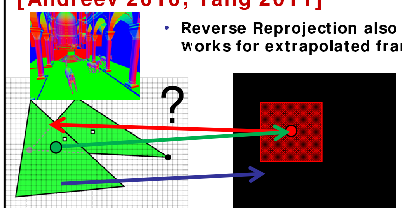




Eurographics 2012, Cagliari, Italy

Reverse Reprojection [Andreev 2010, Yang 2011]

- Reverse Reprojection also works for extrapolated frames

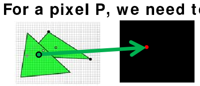


Search in a window around equivalent position for:
Q+Motion(Q) = P

Eurographics 2012, Cagliari, Italy

Reverse Projection - iterative approximation [Andreev 2010, Yang 2011]

- For a pixel P, we need to find a pixel Q:



Q+Motion(Q) = P

Corresponds to a fixpoint:

Q = P – Motion(Q) := F(Q)

Solution: initialize Q0 = P,

iterate Qi = F(Qi-1)

No guarantee for convergence, but often does

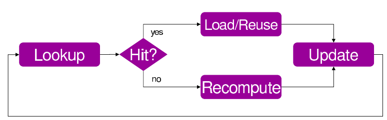
Reality Check – is Reprojection Useful?

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

Eurographics 2012, Cagliari, Italy

Reality Check – is Reprojection Useful?

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

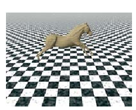


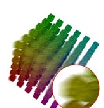
Eurographics 2012, Cagliari, Italy

Many applications of Reprojection


Static procedural texture


Global illumination

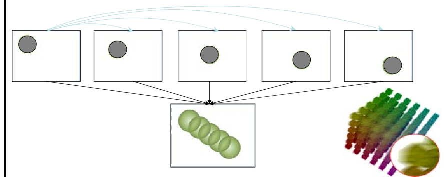

Numerical integral


Multi-pass effects

Eurographics 2012, Cagliari, Italy

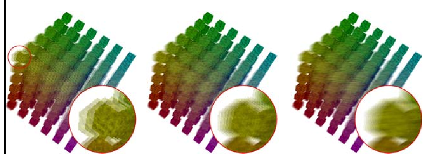
Multi-pass Rendering Effects

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



Eurographics 2012, Cagliari, Italy

Motion Blur



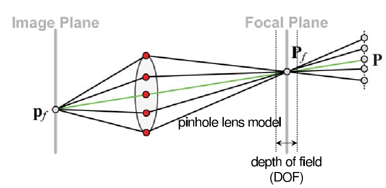
3 time samples
60fps brute-force
60fps RRC

6 time samples
30fps brute-force
60fps RRC

14 time samples
13fps brute-force
30fps RRC

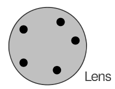
Eurographics 2012, Cagliari, Italy

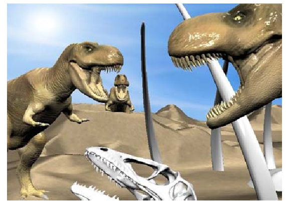
Depth of Field



Depth of Field

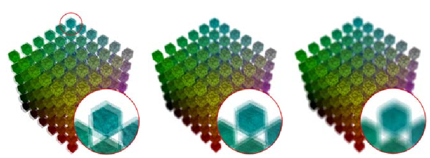
- View synthesis using image-based ray tracing





A few more "tricks" and you get…

Depth of Field



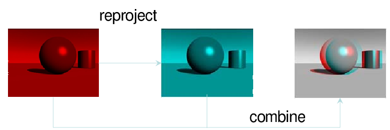
4 aperture samples
45fps brute-force
45fps RRC

9 aperture samples
20fps brute-force
45fps RRC

20 aperture samples
8fps brute-force
20fps RRC

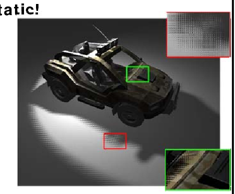
Stereoscopic Rendering

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



This sounds amazing, but…

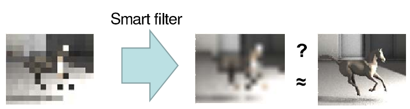
- So far: shading was static!



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

Exploit Spatial Coherence

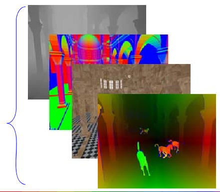
- Idea: use low resolution, then upsample



Eurographics 2012, Cagliari, Italy

Remember?

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



Eurographics 2012, Cagliari, Italy

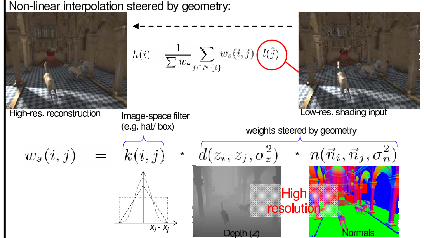
Joint-Bilateral Spatial Upsampling

[Eisemann 2004, Petschnigg 2004, Kopf 2007, Yang 2008]

Non-linear interpolation steered by geometry:

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

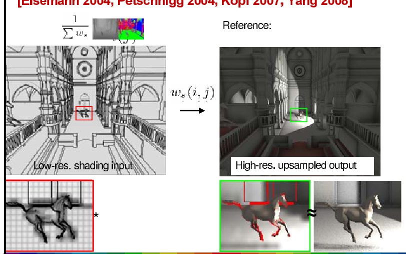
$$w_s(i,j) = k(i,j) \cdot d(z_i, z_j, \sigma_z^2) \cdot n(\vec{n}_i, \vec{n}_j, \sigma_n^2)$$



Eurographics 2012, Cagliari, Italy

Joint-Bilateral Spatial Upsampling

[Eisemann 2004, Petschnigg 2004, Kopf 2007, Yang 2008]



Eurographics 2012, Cagliari, Italy

Spatio-Temporal Upsampling

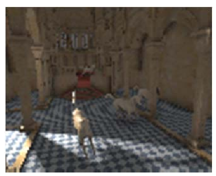
- **Choose preferable method:**

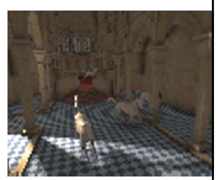
combine spatial upsampling & temporal caching

Eurographics 2012, Cagliari, Italy

Gain information over time?

- **Over time, the same low-res image gives... the same information!**



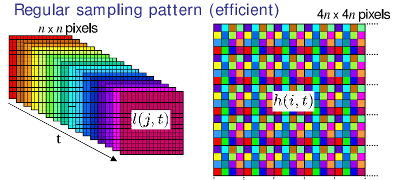


Eurographics 2012, Cagliari, Italy

Temporally Interleaved Sampling

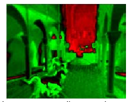
- **Cache different pixel positions to upsample over**
  - → Refresh out-dated pixels (e.g. every **k** x **k** frames)

Regular sampling pattern (efficient)



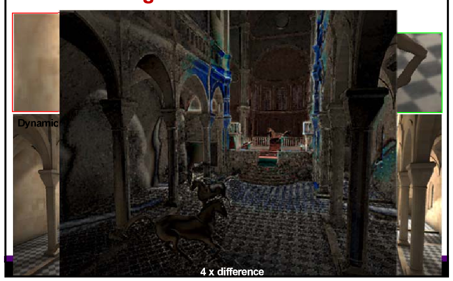
Putting things together:

- **temporal**
  Jittering -> more information for static over time
- **Spatial**
  Bilateral Upsampling (low2high) -> responsiveness


Choose according to change


4x4 upsampled result

What can we gain?

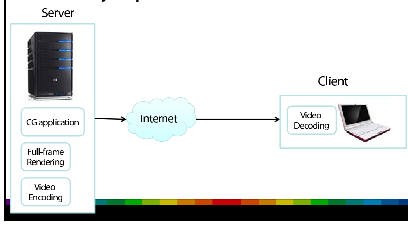


Spatio-Temporal Upsampling [Herzog et al. 2010]

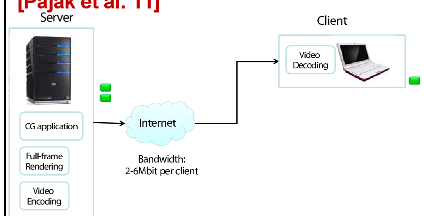
- **Beneficial to use**
  Spatial
  & temporal upsampling
  - Static frame convergence
  - Robustness with respect to changing lighting conditions

Extension: Remote Rendering

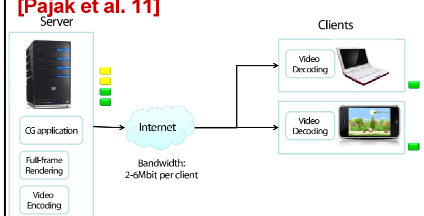
- OnLive, OToy, Gaikai rely on video encoding
  Naturally exploit coherence in video



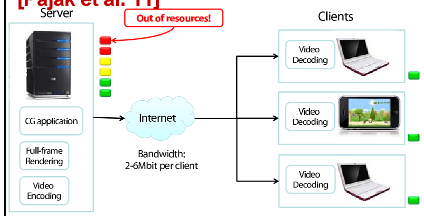
Streaming for Rendered Content [Pajak et al. 11]



Streaming for Rendered Content [Pajak et al. 11]



Streaming for Rendered Content [Pajak et al. 11]

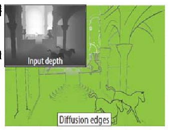


Eurographics 2012, Cagliari, Italy

Streaming for Rendered Content [Pajak et al. 11]

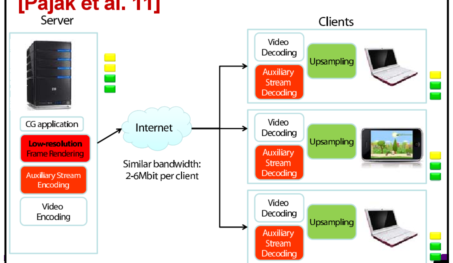
- Rely on **spatio-temporal upsampling** strategies
  - Less bandwidth
  - Less server workload

- Specialized Depth Encod
- encode discontinuities
- exploit smooth variation



Eurographics 2012, Cagliari, Italy

Streaming for Rendered Content [Pajak et al. 11]



Eurographics 2012, Cagliari, Italy

Streaming for Rendered Content [Pajak et al. 11]

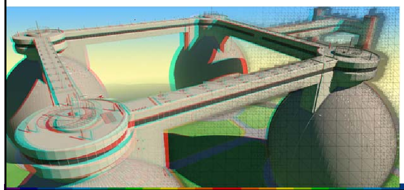
H264

Pajak et al. solution
+ more (depth, motion)

Eurographics 2012, Cagliari, Italy

Extension to Stereo [Didyk et al. VMV'10]

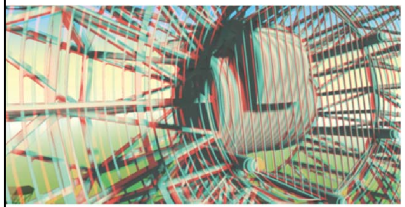
- Adaptive Image-space Stereo View Synthesis
- More sophisticated (adaptive) warping



42

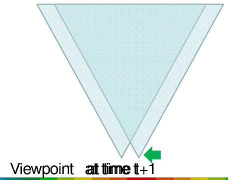
Eurographics 2012, Cagliari, Italy

Extension to Stereo - Results

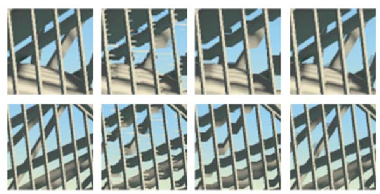


Extension to Stereo

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

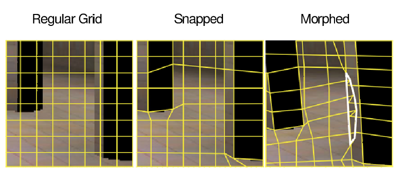


Extension to Stereo - Results



Reference | Previous work | Warping | Temp. Warping

Use a cheap warping technique [Didyk10]



Similar warp as for hold-type blur reduction

Eurographics 2012, Cagliari, Italy

Warping

- **Very efficient**
  - Maps very well to GPU
  - Executes in less than 4ms on a full-HD frame
    - NVIDIA GT 460
- **Easy to implement**
- **Important for streaming architectures**

Eurographics 2012, Cagliari, Italy

Conclusion

- **Spatio-temporal upsampling is very powerful**
- **Extrapolation is possible**
- **Cheap alternatives to rendering all frames**

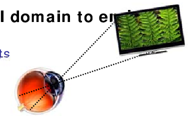
Eurographics 2012, Cagliari, Italy

So far…

- **Different ways rendering (reconstruction, warping, etc.) allow us to produce more efficient high-quality imagery**
- **So far:**
  Have a low computational cost to produce high-quality
- **Now:**
  Make use of temporal domain to improve quality

Exceed display limitations

- **Idea: Exploit temporal domain to en content**
  - Even beyond physical limits



- **Examples:**
  - Color and Brightness: Frame Rate Control + Afterimages
  - Resolution: Apparent Resolution Enhancement

Display Improvement

1990's — 2000's — Today — Future

We are here

1990's	2000's	Today	Future
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	Higher refresh rate Better colors, Better contrast Better brightness ... Less expensive ;)

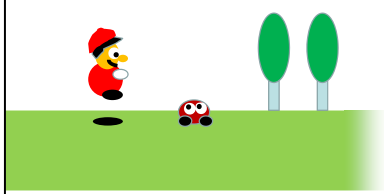
Color Bit Depth: Frame Rate Control [Art04]

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

  - Most screens have 6 bit color depth, but card delivers 8 bit
    -> Flicker different colors and have eye average them

Effect known from older video games

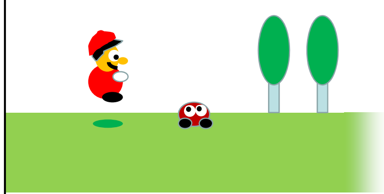
- Virtually augment the color palette



Eurographics 2012, Cagliari, Italy

Effect known from older video games

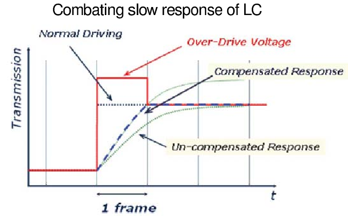
- Virtually augment the color palette



Eurographics 2012, Cagliari, Italy

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

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…**

Based on perception (eye integration)

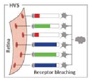
Afterimages [Ritschel&Eisemann2012]

Afterimages [Ritschel&Eisemann2012]

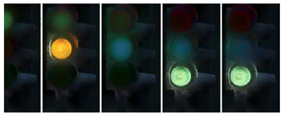
- **The brighter a stimulus, the stronger the afterimage**
- **Bleaching of photoreceptors creates afterimages on retina**

Afterimages [Ritschel&Eisemann2012]

- Computational model for simulation

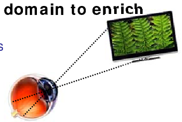


- Can boost perceived brightness similar to glare



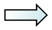
Exceed display limitations

- Idea: Exploit temporal domain to enrich content
  - Even beyond physical limits



- Examples:
  - Color and Brightness: Frame Rate Control + Afterimages
  - Resolution: Apparent Resolution Enhancement

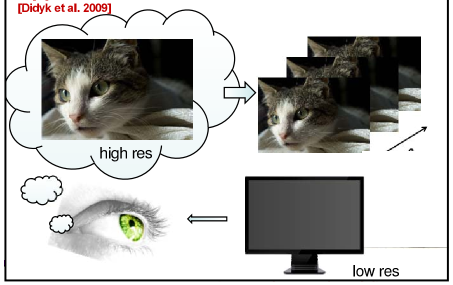
Apparent Resolution Enhancement
[Didyk et al. 2009]



Apparent Resolution Enhancement
[Didyk et al. 2009]

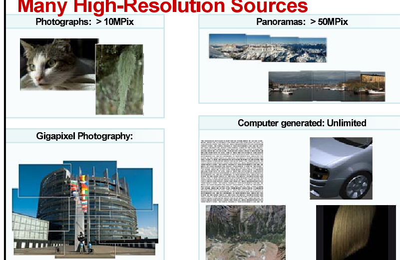
Eurographics 2012, Cagliari, Italy

Apparent Resolution Enhancement
[Didyk et al. 2009]



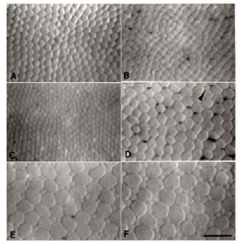
Eurographics 2012, Cagliari, Italy

Many High-Resolution Sources



Eurographics 2012, Cagliari, Italy

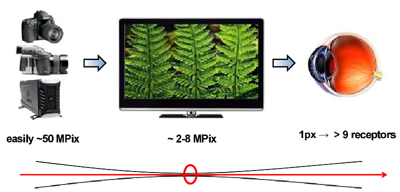
Foveal Photoreceptor Mosaic



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

Eurographics 2012, Cagliari, Italy

Motivation

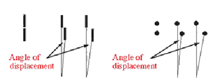


Eurographics 2012, Cagliari, Italy

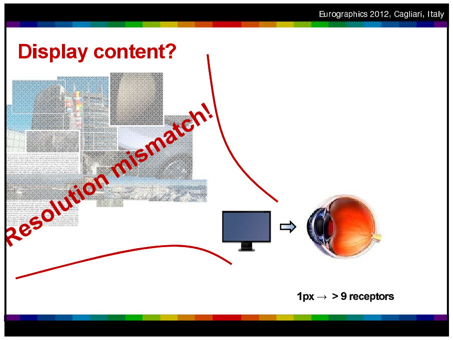
Perception: Spatial Visual Acuity

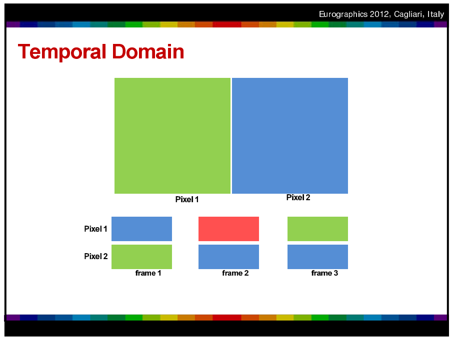
- **Target *resolution* threshold: the smallest angular size at which subjects can discriminate**

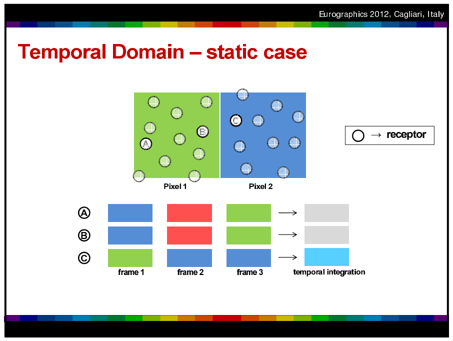
- **Target *localization* threshold: the smallest difference in position which subjects can discriminate (Vernier hyperacuity)**



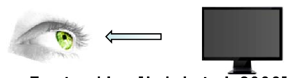
http://webvision.med.utah.edu/book/part-viii-gabac-receptors/visual-acuity/


Eurographics 2012, Cagliari, Italy
Display content?
Resolution mismatch!
1px → > 9 receptors


Eurographics 2012, Cagliari, Italy
Temporal Domain
Pixel 1
Pixel 2
Pixel 1
Pixel 2
frame 1
frame 2
frame 3

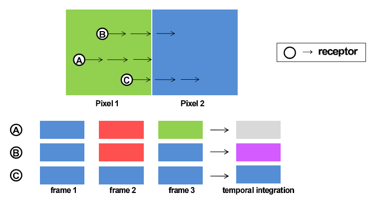

Eurographics 2012, Cagliari, Italy
Temporal Domain – static case
→ receptor
Pixel 1
Pixel 2
A
B
C
frame 1
frame 2
frame 3
temporal integration

Perception: Smooth Pursuit Eye Motion (SPEM)



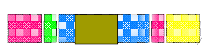
- **Eye tracking [Laird et al. 2006]**
  - Almost perfect tracking and instant for steady linear motion
    - velocities of 0.625 – 80 deg/s
  - HVS compensates for fixational eye movements:
    - tremors, drifts, and microsaccades
- **Reality: Blur from object motion is eliminated**
- **Screen: Blur can be used to increase resolution!**

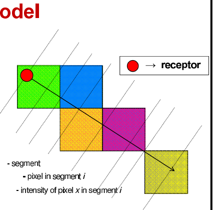
Temporal Domain – dynamic case



Temporal Integration Model

Receptor signal:



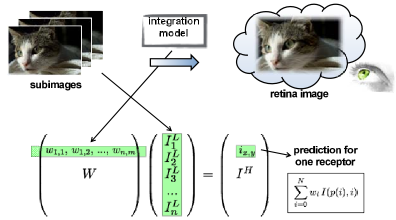


$$\sum_{i=0}^{N} w_i \, I(p(i), i)$$

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

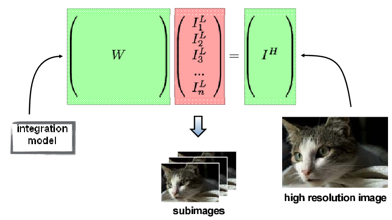
Prediction in Equations



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

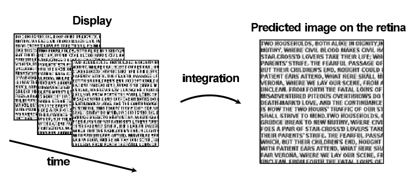
$$\sum_{i=0}^{N} w_i \, I(p(i), i)$$

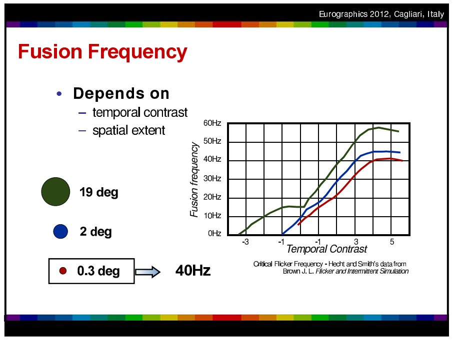
Optimization Problem

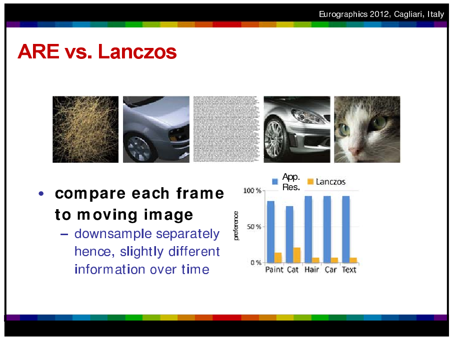


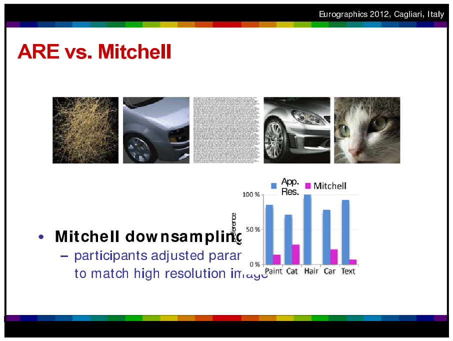
$$\begin{pmatrix} W \end{pmatrix} \begin{pmatrix} I_1^L \\ I_2^L \\ I_3^L \\ \ldots \\ I_n^L \end{pmatrix} = \begin{pmatrix} I^H \end{pmatrix}$$

Optimization Result




Eurographics 2012, Cagliari, Italy
Fusion Frequency
Depends on
temporal contrast
spatial extent
19 deg
2 deg
0.3 deg
40Hz
Fusion frequency
60Hz
50Hz
40Hz
30Hz
20Hz
10Hz
0Hz
Temporal Contrast
Critical Flicker Frequency - Hecht and Smith's data from Brown J. L. Flicker and Intermittent Stimulation

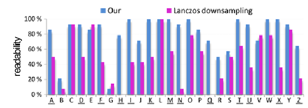

Eurographics 2012, Cagliari, Italy
ARE vs. Lanczos
compare each frame to moving image
downsample separately hence, slightly different information over time
App. Res.
Lanczos
100 %
50 %
0 %
Paint Cat Hair Car Text


Eurographics 2012, Cagliari, Italy
ARE vs. Mitchell
Mitchell downsampling
participants adjusted param to match high resolution image
App. Res.
Mitchell
100 %
50 %
0 %
Paint Cat Hair Car Text

ARE - Alphabet

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Size: 2 x 3 pixels

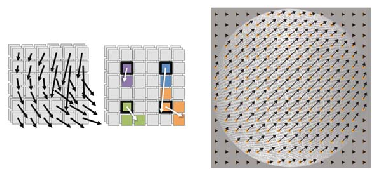


- **Applications:**
  - scrolling text or maps on low resolution devices
  - stock tickers, news headlines

Eurographics 2012, Cagliari, Italy

Recently: Extension to movies

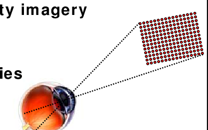
Apparent Resolution Enhancement for Animations [Templin et al. SCCG 2011]



Eurographics 2012, Cagliari, Italy

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…

Eurographics 2012, Cagliari, Italy

Thank you very much for your attention!

Karol Myszkowski
karol@mpi-inf.mpg.de

Elmar Eisemann
eisemann@telecom-paristech.fr

Acknowledgments:
Thank you for support in creating the slides go to
Daniel Scherzer, Robert Herzog and Dawid Pajak

Eurographics 2012
Cagliari, Italy May 13-18
33rd ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR COMPUTER GRAPHICS

Stereo content retargeting

Piotr Didyk
MPI Informatik

Why stereo?

Images are no longer flat

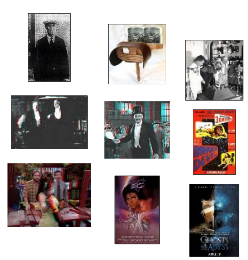
- Improves realism
- Images are not longer flat
- Better layout separation

Reproduced view dependent effects

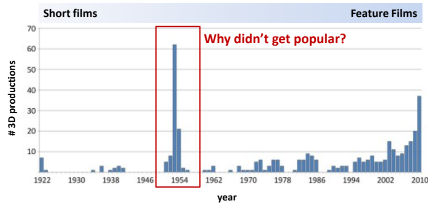
- Improves material perception

History of stereo

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



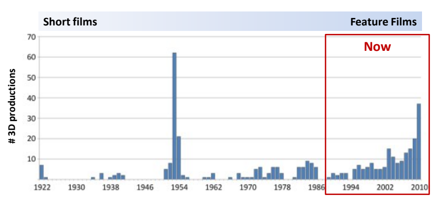
Number of 3D productions



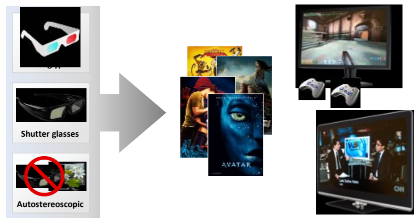
Early 3D production

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

Number of 3D productions



Stereo in daily life



Eurographics 2012, Cagliari, Italy

Current 3D production

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:

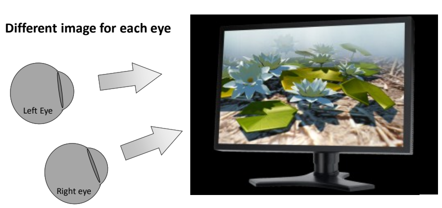
- Shutter glasses
- Polarized glasses
- Autostereoscopic displays are getting better

They are flat!

Eurographics 2012, Cagliari, Italy

Stereo on a flat display

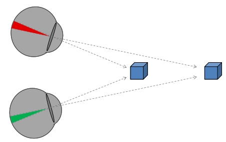
- **Different image for each eye**



Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity



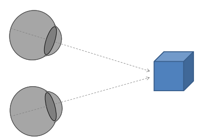
Eurographics 2012, Cagliari, Italy

Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation,



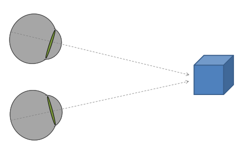
Eurographics 2012, Cagliari, Italy

Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation,

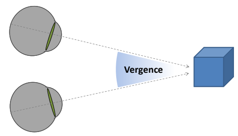


Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence



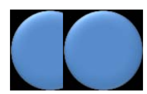
Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
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,

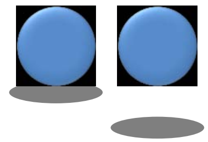
Depth perception

We see depth due to depth cues.

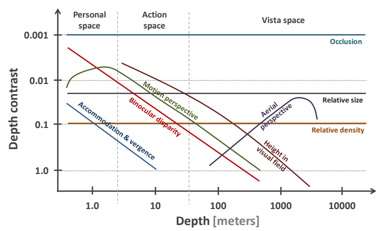
Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
occlusion, size, shadows…



Cues sensitivity



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

Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

Ocular depth cues:
accommodation, vergence

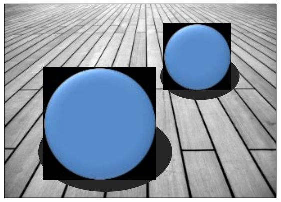
Pictorial depth cues:
occlusion, size, shadows…

Challenge:
Consistency is required!

Simple conflict example

Present cues:

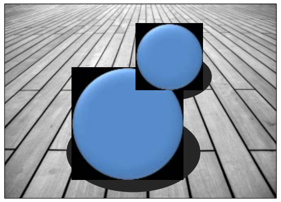
- Size
- Shadows
- Perspective



Simple conflict example

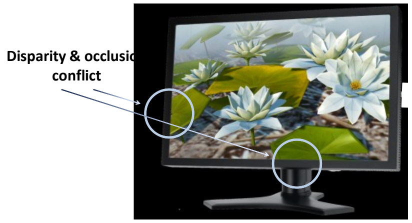
Present cues:

- Size
- Shadows
- Perspective
- Occlusion



Disparity & occlusion conflict

Disparity & occlusion conflict



Depth perception

We see depth due to depth cues.

Stereoscopic depth cues:
binocular disparity

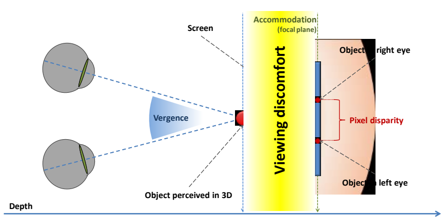
Ocular depth cues:
accommodation, vergence

→ **Require 3D space**
We cheat our Human Visual System!

Pictorial depth cues:
occlusion, size, shadows…

→ **Reproducible on a flat displays**

Cheating our HVS



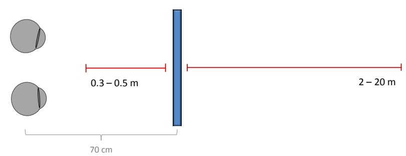
Viewing discomfort

Comfort zones

Comfort zone size depends on:

- Presented content
- Viewing condition

Simple scene



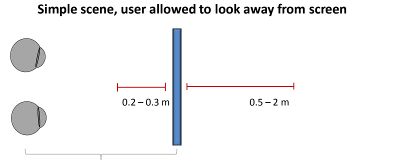
"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

Comfort zones

Comfort zone size depends on:

- Presented content
- Viewing condition

Simple scene, user allowed to look away from screen



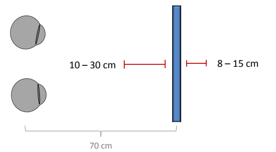
"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

Comfort zones

Comfort zone size depends on:

- Presented content
- Viewing condition

Difficult scene



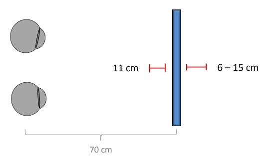
"Controlling Perceived Depth in Stereoscopic Images" by Jones et al. 2001

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

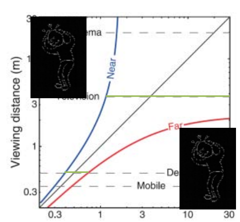
Comfort zones

Comfort zone size depends on:

- Presented content
- Viewing condition
- Screen distance

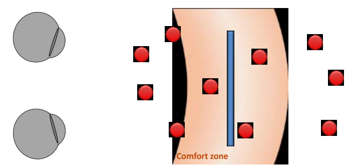
Other factors:

- Distance between eyes
- Depth of field
- Temporal coherence



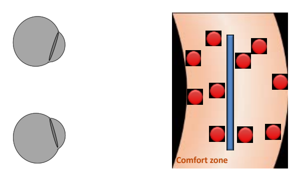
"The zone of comfort: Predicting visual discomfort with stereo displays" by Shibata et al. 2011

Depth manipulation



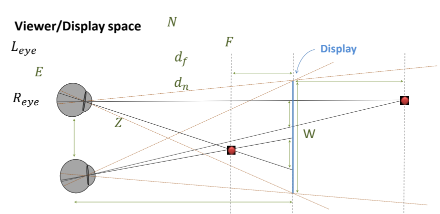
Viewing discomfort

Depth manipulation



Viewing discomfort → Scene manipulation → Viewing comfort

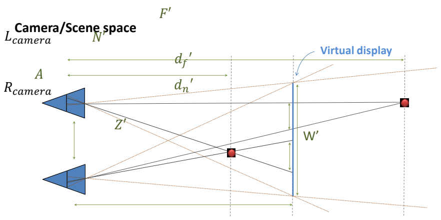
Camera manipulations



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

Camera manipulations

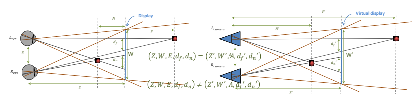
Camera/Scene space



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

Camera manipulations

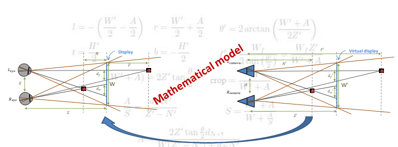
Camera/Scene space



- 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

"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

Game controller:

"OSCAM - Optimized Stereoscopic Camera Control for Interactive 3D" by Oscam et al. 2011

- Calculate appropriate camera parameters
- Adjustment in each frame

"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

General procedure:

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

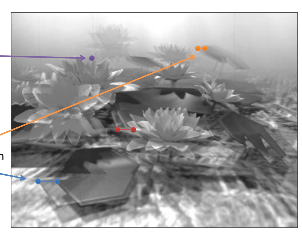
Displaying on different device:
(captured content)

- Potential discomfort
- Recapturing ?

Pixel disparity

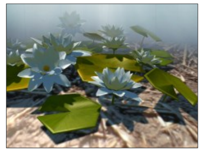
Zero disparity
on the screen plane

Bigger disparities
in front and behind screen



Left + right view

Stereo content

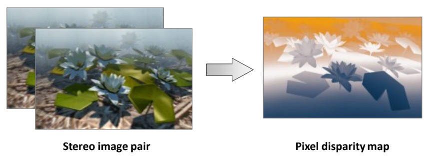

Left view


Right view

Can we have pixel disparity / depth ?

Eurographics 2012, Cagliari, Italy

Sources of pixel disparity

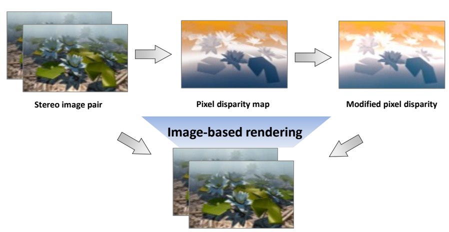


Rendering ⟶ Usually available

Only image pair ⟶ Computer vision technique

Eurographics 2012, Cagliari, Italy

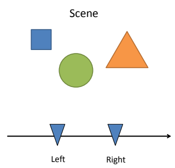
Disparity manipulations



"Adaptive Image-based Stereo View Synthesis" by Didyk et al. 2010
"Nonlinear Disparity Mapping for Stereoscopic 3D" by Lang et al. 2010

Stereoscopy from Light Fields

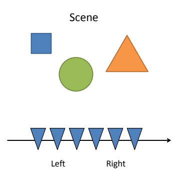
Light Field



"*Multi-Perspective Stereoscopy from Light Fields*" by Kim et al. 2011

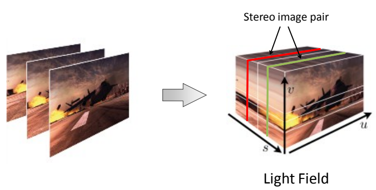
Stereoscopy from Light Fields

Light Field



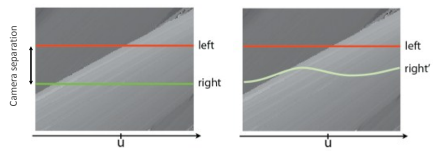
"*Multi-Perspective Stereoscopy from Light Fields*" by Kim et al. 2011

Stereoscopy from Light Fields



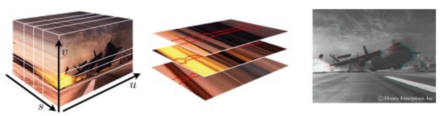
"*Multi-Perspective Stereoscopy from Light Fields*" by Kim et al. 2011

Stereoscopy from Light Fields



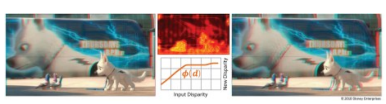
"Multi-Perspective Stereoscopy from Light Fields" by Kim et al. 2011

Stereoscopy from Light Fields



"Multi-Perspective Stereoscopy from Light Fields" by Kim et al. 2011

Disparity manipulations

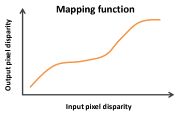


"Nonlinear Disparity Mapping for Stereoscopic 3D" by Lang et al. 2010

Disparity manipulations

Pixel disparity map

Modified pixel disparity



Function:

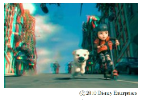
- Liner
- Logarithmic
- Content dependent

Other possibilities:

- Gradient domain
- Local operators

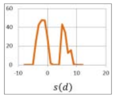
"Nonlinear Disparity Mapping for Stereoscopic 3D" by Lang et al. 2010

Saliency map

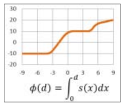


Input stereo image

Saliency map



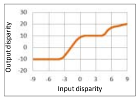
Disparity importance

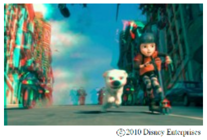


Disparity mapping function

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

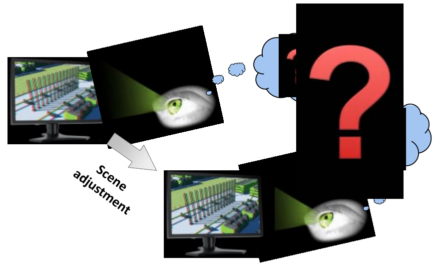
Saliency map



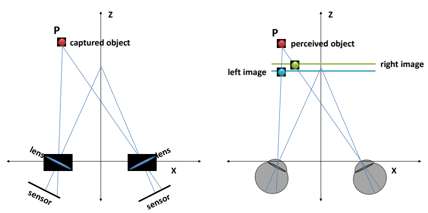


"Nonlinear Disparity Mapping for Stereoscopic 3D" by Lang et al. 2010

Scene manipulation



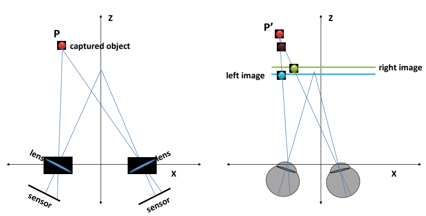
Misperception



Parameters are the same

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

Misperception

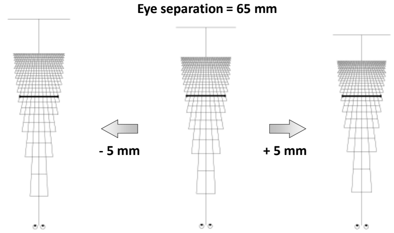


Eyes position and interoccular distance changed

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

Misperception

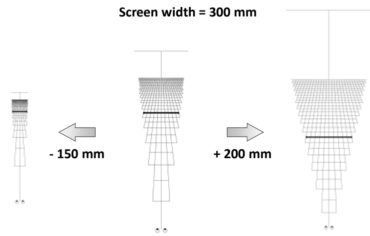
Eye separation = 65 mm



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

Misperception

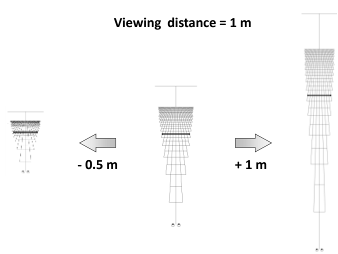
Screen width = 300 mm



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

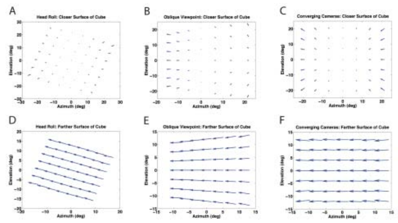
Misperception

Viewing distance = 1 m



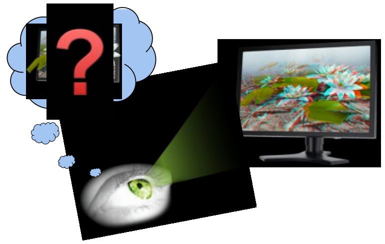
"Image Distortions in Stereoscopic Video Systems" by Woods et al. 1993

Misperception



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

3D image prediction

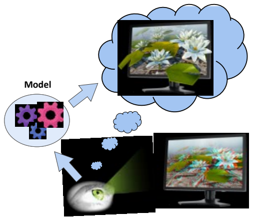


Depth perception

Stereoscopic depth cues:
binocular disparity

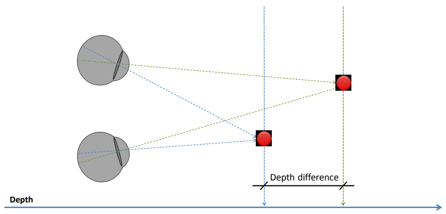
Ocular depth cues:
accommodation, vergence

Pictorial depth cues:
occlusion, size, shadows…

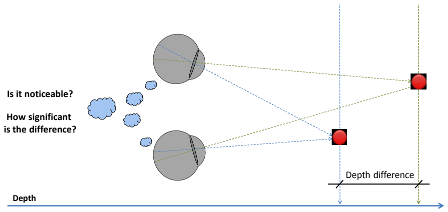


"A perceptual model for disparity" by Didyk et al. 2011

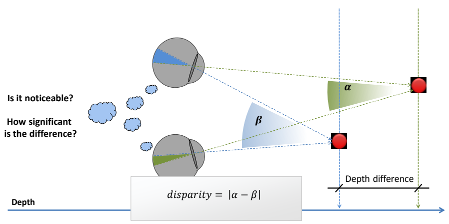
Disparity perception



Disparity perception

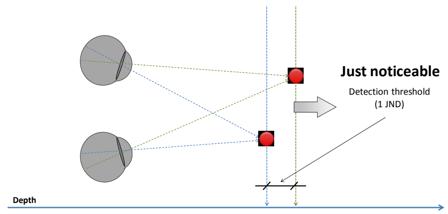


Disparity perception



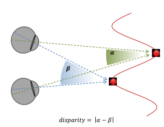
$$disparity = |\alpha - \beta|$$

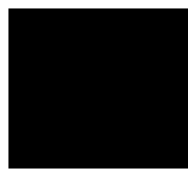
One just-noticeable difference



How big is the detection threshold?

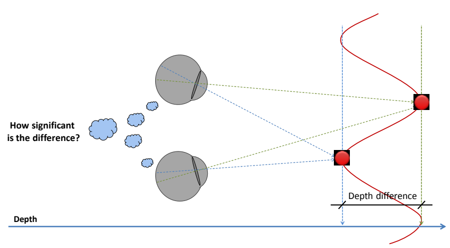
For sinusoidal depth corrugation



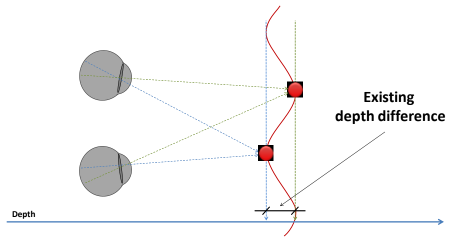


"Sensitivity to horizontal and vertical corrugations defined by binocular disparity."
by Bradshaw et al. 1999

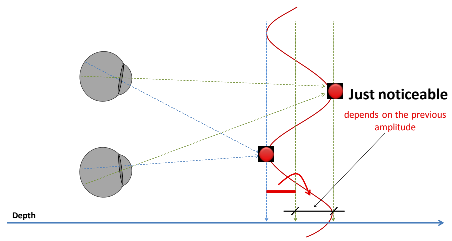
Disparity perception



Discrimination threshold



Discrimination threshold



Disparity perception

Sensitivity to depth changes depends on:

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

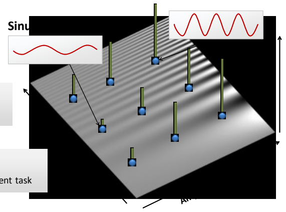
Measurements

Parameter space:

1. Sample the space

3. Measure thresholds

- Experiment with adjustment task

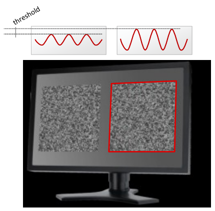


"A perceptual model for disparity" by Didyk et al. 2011

Measurements

Thresholds measurement:

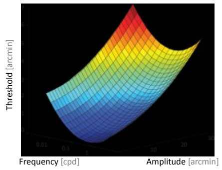
- Two sinusoidal corrugations
- Which has more depth? (left/right)
- Amplitude adjustment (PEST with 2AFC)
- 12 participants → 300+ samples



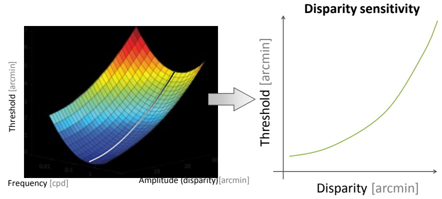
"A perceptual model for disparity" by Didyk et al. 2011

Model

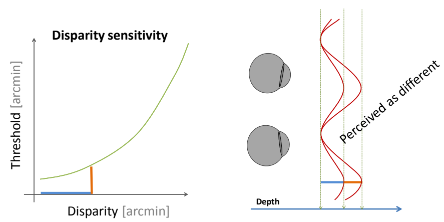
3. Fit analytic function



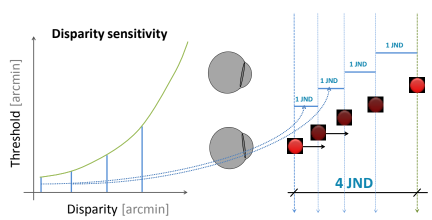
The HVS response



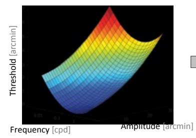
The HVS response

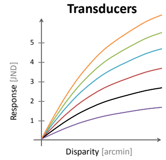


The HVS response



The HVS response





"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

Eurographics 2012, Cagliari, Italy

Perceptual space

We show so far:

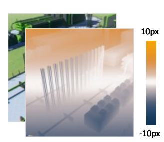
(disparity, frequency)
[arcmin, cpd]

HVS response
[JND]

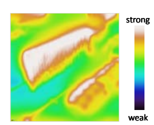
Eurographics 2012, Cagliari, Italy

Perceptual space

Our problem:



3D scene with pixel disparity
[pixels]



Map of HVS response
[JND]

Perceptual space

Our problem:

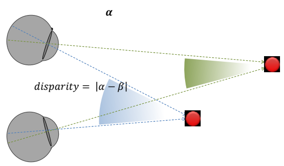
- **Pixel disparity** [pixels] ↔ **Disparity** [arcmin]

3D scene

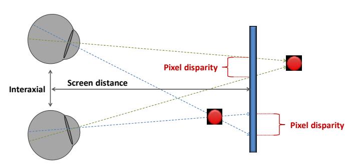
[pixels]

[JND]

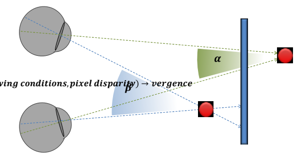
Pixel disparity to disparity



Pixel disparity to disparity



Pixel disparity to disparity



$(viewing\ conditions, pixel\ disparity) \rightarrow vergence$

Vergence to disparity

Vergence [arcmin]

Disparity [arcmin]

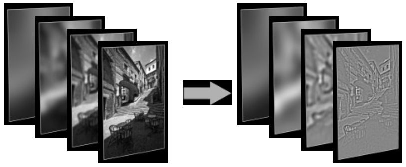
How do people deal with luminance?

Luminance (contrast perception)

Luminance

Perceptual space
(Perceived contrast)

Luminance (contrast perception)



Lowpass filters

Contrast decomposed into frequency bands

Eurographics 2012, Cagliari, Italy

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 into frequency bands

Eurographics 2012, Cagliari, Italy

Luminance (contrast perception)

Luminance ↔ Vergence

Luminance contrast ↔ Disparity

Disparity / Luminance similarity:

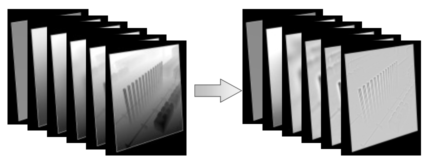
Lowpass filters

Perceptual operations into frequency bands

Vergence to disparity

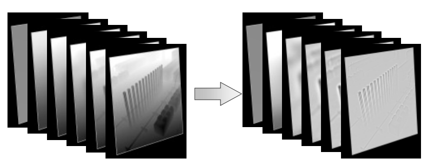
Vergence [arcmin]

Vergence to disparity



Lowpass filters **Differences**

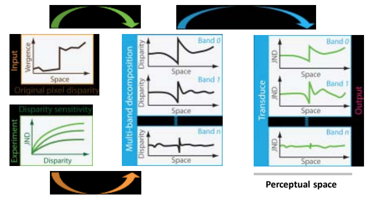
Vergence to disparity



Lowpass filters **Differences**

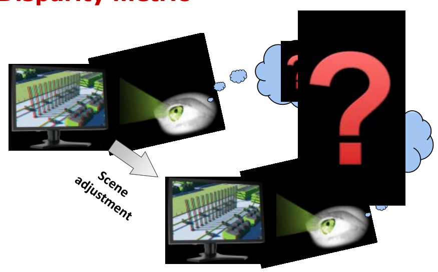
- We can process frequencies independently
- **Vergence → Disparity**

Perceptual model



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

Disparity metric

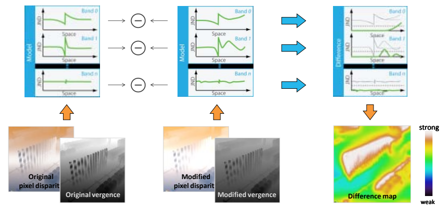


Disparity metric

For Luminance:

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

Disparity metric



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

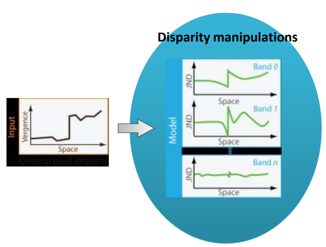
Disparity manipulations

→ The HVS is taken into account
→ Efficient disparity reduction
→ Important disparities preserved

Manipulations in perceptual space:

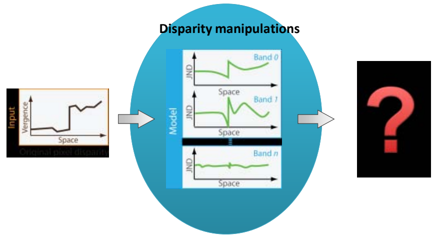
"Nonlinear Disparity Mapping for Stereoscopic 3D"
by Lang et al. 2010

Disparity manipulation



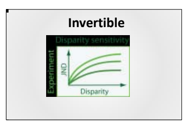
"A perceptual model for disparity" by Didyk et al. 2011

Disparity manipulation

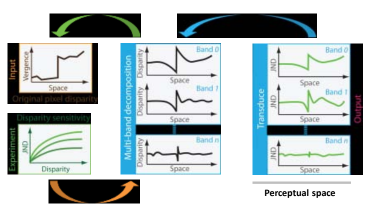


"A perceptual model for disparity" by Didyk et al. 2011

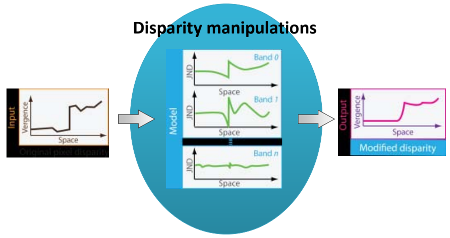
Inverse model



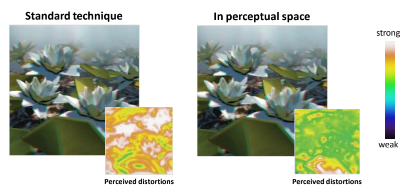
Inverse model



Disparity manipulation

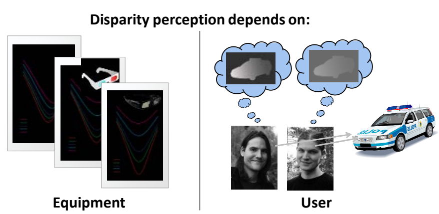


Disparity manipulation

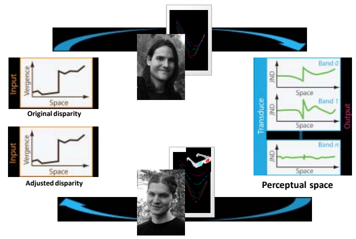


- Important disparities preservation

Personalization

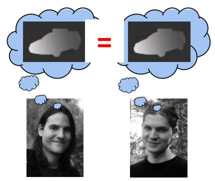


Personalization



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

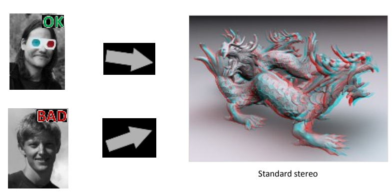
Personalization



All users perceive the same regardless:

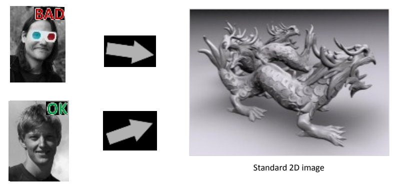
- Equipment
- Disparity sensitivity

Backward-compatible stereo



Standard stereo

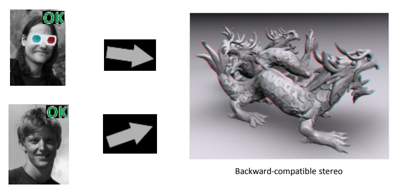
Backward-compatible stereo



Standard 2D image

Eurographics 2012, Cagliari, Italy

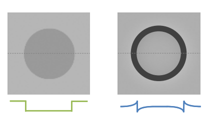
Backward-compatible stereo



Backward-compatible stereo

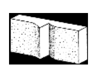
Eurographics 2012, Cagliari, Italy

Cornsweet illusion



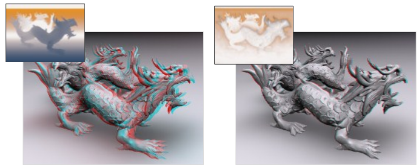
- Similar perceived contrast
- Luminance range reduced

Cornsweet illusion works for depth:



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

Backward-compatible stereo



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

Photo — 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

Conclusions

- Stereo perception is complex phenomena
- Important aspects:
  - Viewing conditions
  - Viewer
  - Equipment
  - Temporal coherence …
- Different adjustment techniques:
  - Camera adjustment
  - Pixel disparity mapping operators
  - Perceptual space

Eurographics 2012
Cagliari, Italy May 13-18
33rd ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR COMPUTER GRAPHICS

Image / Video Quality Assessment

of Retargeted Content

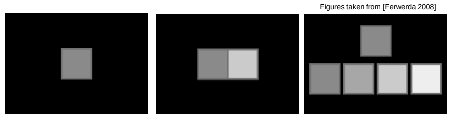
Tunç O. Aydın
Disney Research, Zurich
<tunc@disneyresearch.com>

Problem Definition

Rate the Quality

Subjective Quality Assessment

Figures taken from [Ferwerda 2008]



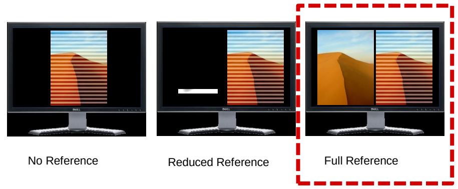
Detection Discrimination Scaling

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

+ Reliable - High Cost

Eurographics 2012, Cagliari, Italy

Objective Quality Assessment



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

Eurographics 2012, Cagliari, Italy

Generic Objective Quality Assessment Workflow

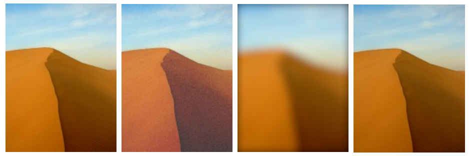
Eurographics 2012, Cagliari, Italy

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!

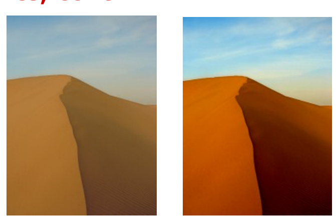
Perception of Distortions

Reference (bmp, 616K)

Compressed (jpg, 48K)

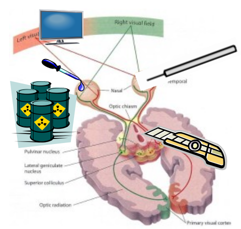
Difference Image (Color coded)

Limitations of Simple Distortion Metrics, cont.



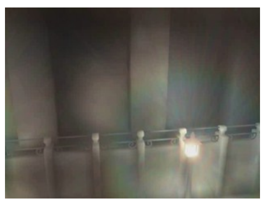
Visible difference doesn't always mean lower quality!

The Human Visual System (HVS)



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

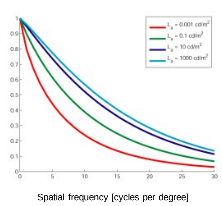
HVS effects: (1) Glare



Video Courtesy of Tobias Ritschel

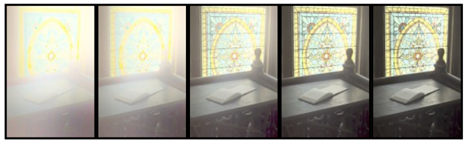
- **Disability Glare (blooming)**

Disability Glare



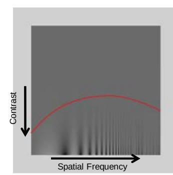
- **Model of Light Scattering**
  - Point Spread Function in spatial domain
  - Optical Transfer Function in Fourier Domain [Deeley et al. 1991]

(2) Light Adaptation



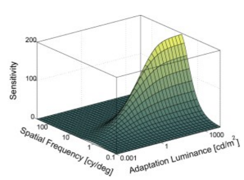
Adaptation Level: 10^{-4} cd/m^2 — Time ⟶ — Adaptation Level: 17 cd/m^2

(3) Contrast Sensitivity



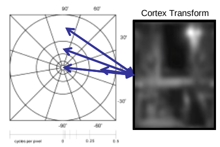
CSF(spatial frequency, adaptation level,
temporal freq., viewing dist, …)

Contrast Sensitivity Function (CSF)



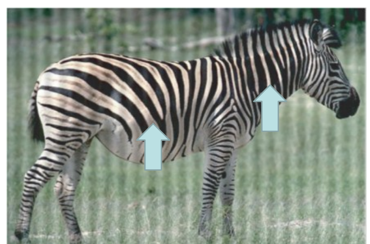
- Steady-state CSF^S: Returns the Sensitivity (1/Threshold contrast), given the adaptation luminance and spatial frequency [Daly 1993, Mantiuk et al. 2011].

(4) Visual Channels



Eurographics 2012, Cagliari, Italy

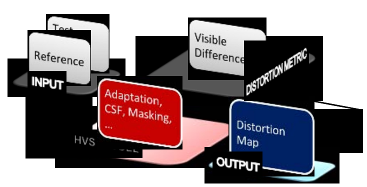
(5) Visual Masking



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

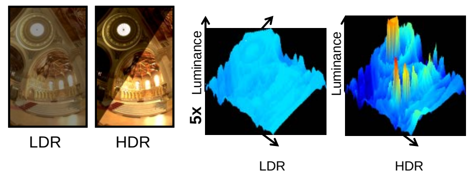
Eurographics 2012, Cagliari, Italy

Generic HVS-based Quality Assessment Workflow



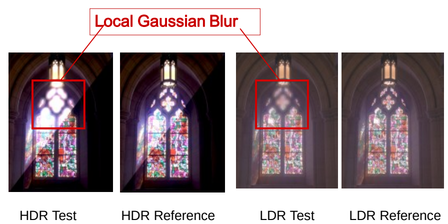
Visible Differences Predictor (VDP) [Daly 93, Mantiuk et al. 05, Mantiuk et al. 11], Visual Discrimination Model (VDM) [Lubin 95]

QA of Retargeted Images? HDR Tone mapping case



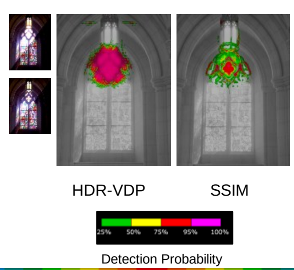
Eurographics 2012, Cagliari, Italy

Case Study

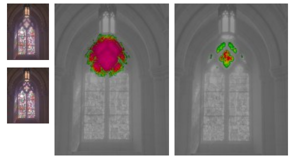


Eurographics 2012, Cagliari, Italy

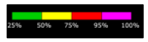
(1) HDR pair



(2) LDR pair

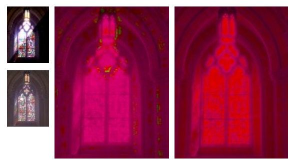


HDR-VDP SSIM

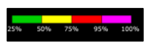


Detection Probability

(3) HDR test, LDR reference

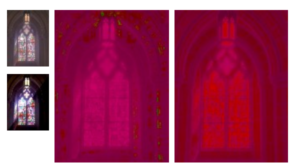


HDR-VDP SSIM

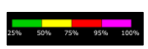


Detection Probability

(4) LDR test, HDR reference

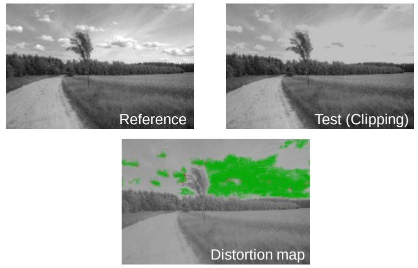


HDR-VDP SSIM



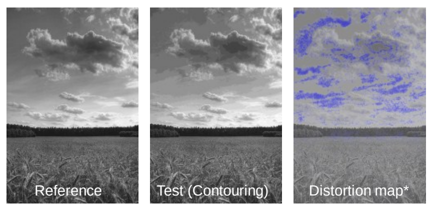
Detection Probability

Loss of Visible Contrast



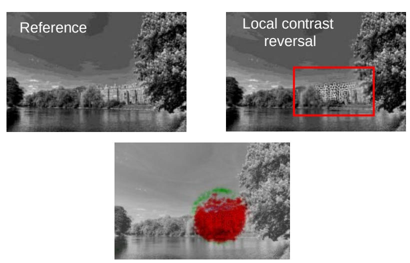
Eurographics 2012, Cagliari, Italy

Amplification of Invisible Contrast

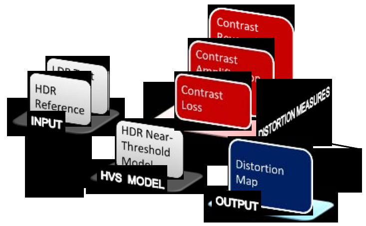


Eurographics 2012, Cagliari, Italy

Reversal of Visible Contrast

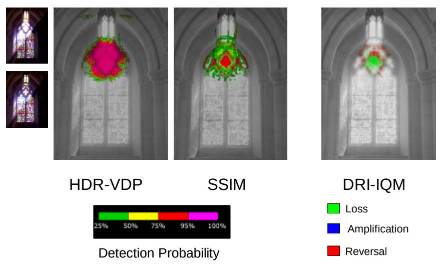


Generic DRI Image Quality Assessment Workflow



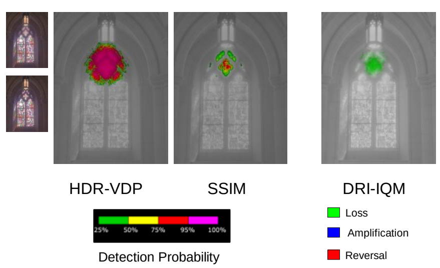
Eurographics 2012, Cagliari, Italy

(1) HDR pair



Eurographics 2012, Cagliari, Italy

(2) LDR pair



(3) HDR test, LDR reference

HDR-VDP SSIM DRI-IQM

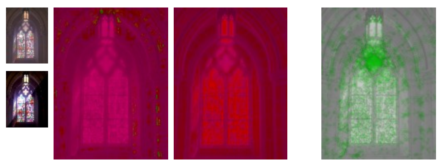
25% 50% 75% 95% 100%

Detection Probability

- Loss
- Amplification
- Reversal

Eurographics 2012, Cagliari, Italy

(4) LDR test, HDR reference



HDR-VDP SSIM DRI-IQM

25% 50% 75% 95% 100%

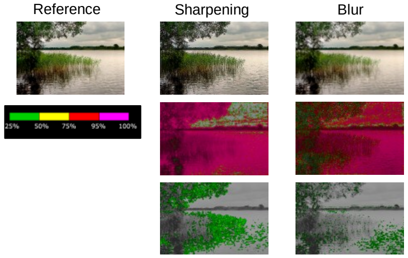
Detection Probability

- Loss
- Amplification
- Reversal

Eurographics 2012, Cagliari, Italy

Detecting distortions

Reference Sharpening Blur

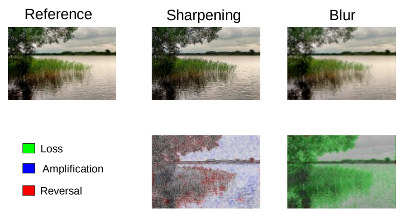


25% 50% 75% 95% 100%

HDR-VDP

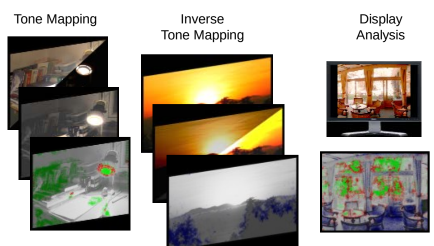
SSIM

Detecting “types” of distortions



Eurographics 2012, Cagliari, Italy

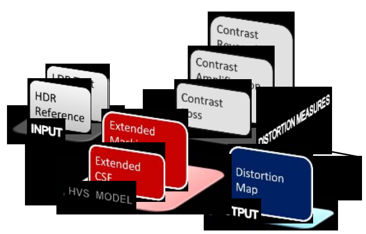
HDR Tone Mapping Evaluation



[Aydin et al. 2008]

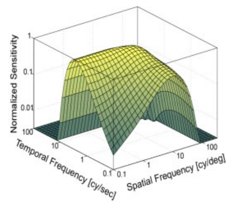
Eurographics 2012, Cagliari, Italy

Generic DRI Video Quality Assessment Workflow



Extended Contrast Sensitivity Function

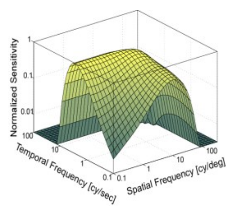
- *CSF: ω,ρ,L_a → S*
  - *ω:* temporal frequency,
  - *ρ:* spatial frequency,
  - *L_a:* adaptation level,
  - *S:* sensitivity.



Spatio-temporal CSF

Extended Contrast Sensitivity Function, cont.

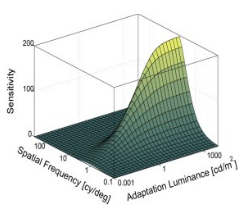
- ***CSF: ω,ρ,L_a → S***
  - *ω:* temporal frequency,
  - *ρ:* spatial frequency,
  - *L_a:* adaptation level,
  - *S:* sensitivity.



Spatio-temporal CSF^T

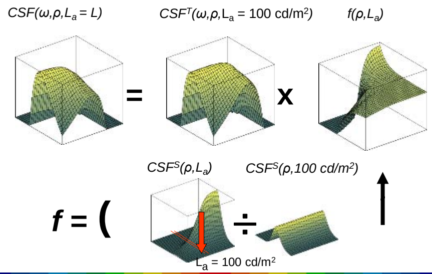
Extended Contrast Sensitivity Function, cont.

- *CSF: ω,ρ,L_a → S*
  - *ω:* temporal frequency,
  - *ρ:* spatial frequency,
  - *L_a:* adaptation level,
  - *S:* sensitivity.



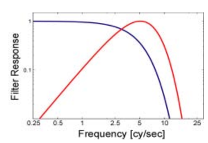
Steady-state CSF^S

Extended Contrast Sensitivity Function, derivation

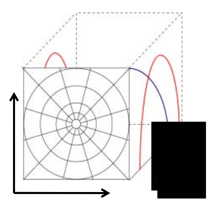


Eurographics 2012, Cagliari, Italy

Extended Cortex Transform

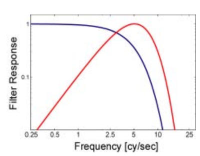


Sustained and **Transient**
Temporal Channels [Winkler 2005]

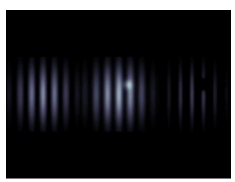


Eurographics 2012, Cagliari, Italy

Temporal Channels



Sustained and **Transient**
Temporal Channels

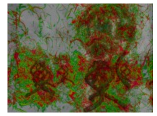


Signal **Sustained** **Transient**

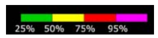
Evaluation of HDR Video Tone Mapping

HDR Reference

LDR Test



Contrast Amplification



Eurographics 2012, Cagliari, Italy

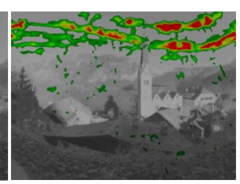
Evaluation of HDR Compression

HDR Reference

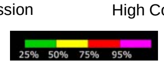
Eurographics 2012, Cagliari, Italy

Evaluation of HDR Compression

Medium Compression



High Compression

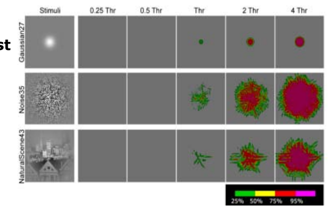


[Aydin et al. 2010]

Eurographics 2012, Cagliari, Italy

Subjective Calibration

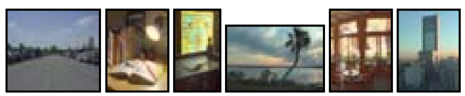
- **Modelfest dataset at five contrast levels**



Eurographics 2012, Cagliari, Italy

Subjective Validation

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



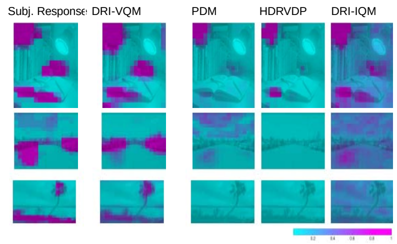
Eurographics 2012, Cagliari, Italy

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. 2011]
  - Smarter distortion measures.