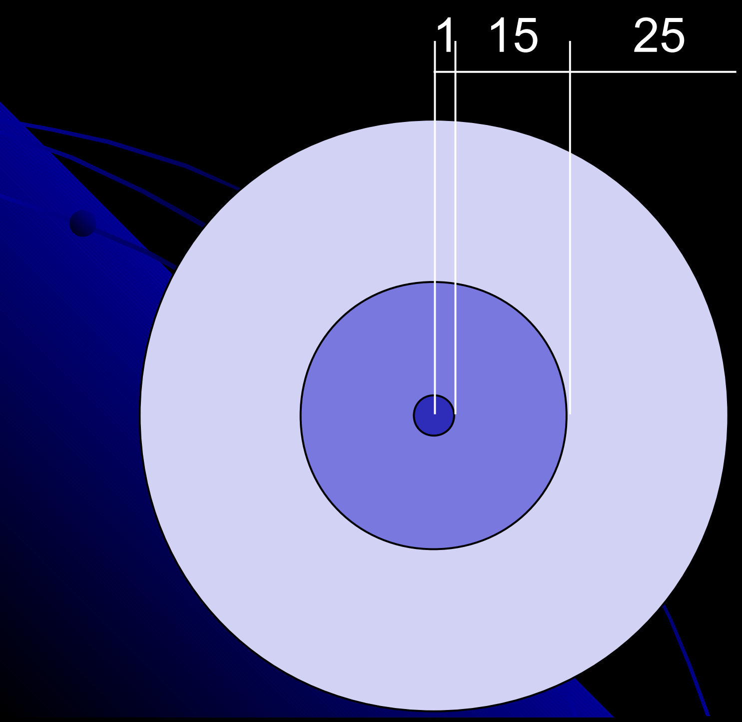
Sledovanie pohybu očí

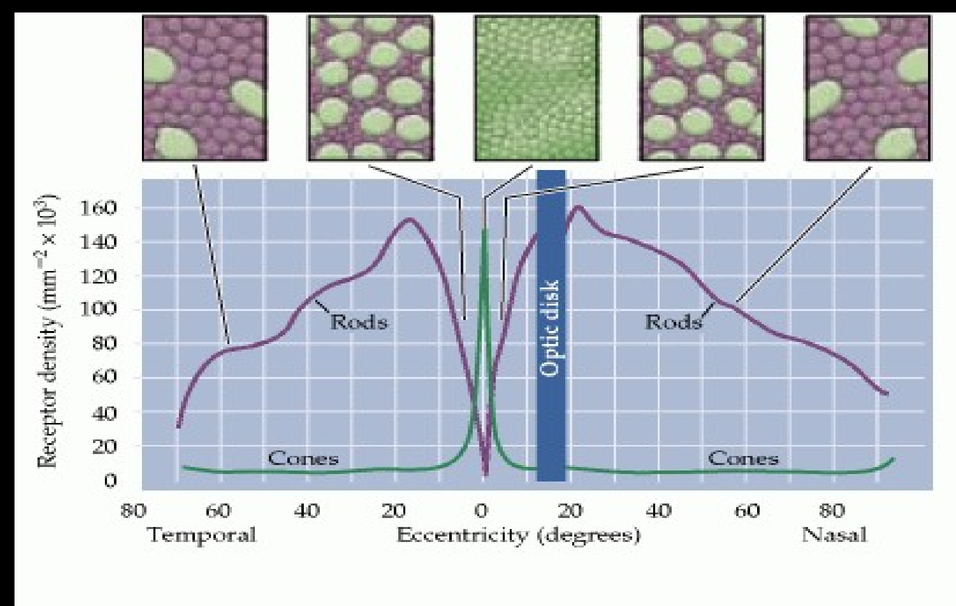
Zorné pole:

Centrálne videnie – najostrejšie (1-2 stupne)

Mimocentrálne videnie – zaisťuje priestorové videnie (30 stupňov)

Periférne videnie – reaguje na blikajúce objekty a náhly pohyb (15-50% ostrosti centrálneho videnia, nerozpoznáva farby)





Pohyby očí

- Oči sa hýbu stále (aj počas spánku)
- Rôzne typy pohybov
  - Sledovanie
  - Chvenie
  - Otáčanie
  - Drift
- Najzaujímavejšie sú
  - **Fixácie**
  - **Kmity**

Fixácie

Obdobie relatívnej stability, počas ktorého vnímame objekty

Relatívna stabilita: oko nie je nehybné, ale mierne driftuje (v rámci 1 stupňa), vracia sa skokmi – mikrosakádami

Informáciu o scéne dostávame len počas počas fixácií

Trvanie 120-1000 ms, väčšinou okolo 200-600 ms

Nasledované kmitmi – sakádami

Kmity - sakády

Skoky spájajúce fixácie

Veľmi krátke – trvajú 40-120 ms

Veľmi rýchle (až do 600o/s), vizuálny systém je nefunkčný počas pohybu

V priebehu kmitu sa nedá meniť jeho smer alebo koncový bod, tieto sú určené pred začatím kmitu

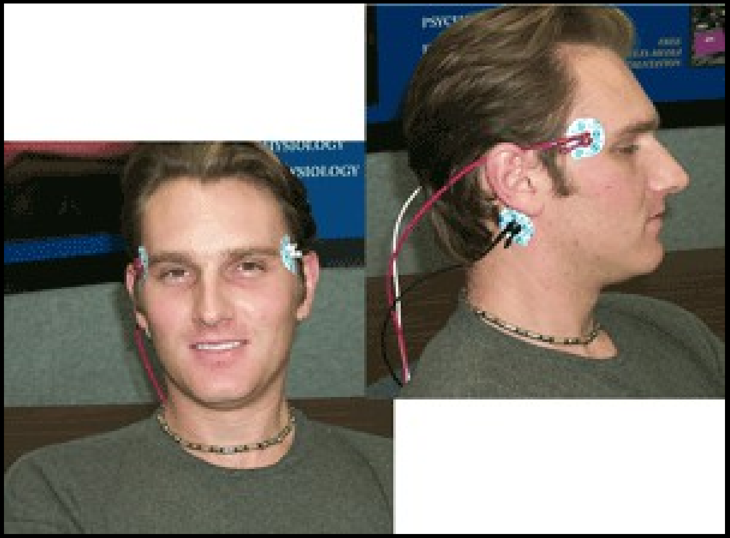
Pri kmitoch väčších ako 30 stupňov sa hýbe aj hlava

Metódy sledovania pohybov oka

- Elektronické
- Mechanické
- Videometódy
  - Jednobodové
  - Dvojbodové

Elektronické metódy

- Elektródy na koži okolo oka merajú potenciálové zmeny v oku
- Slabá presnosť
- Zachytáva lepšie relatívne pohyby
- Použitie pri neurologickej diagnostike - elektrookulografia



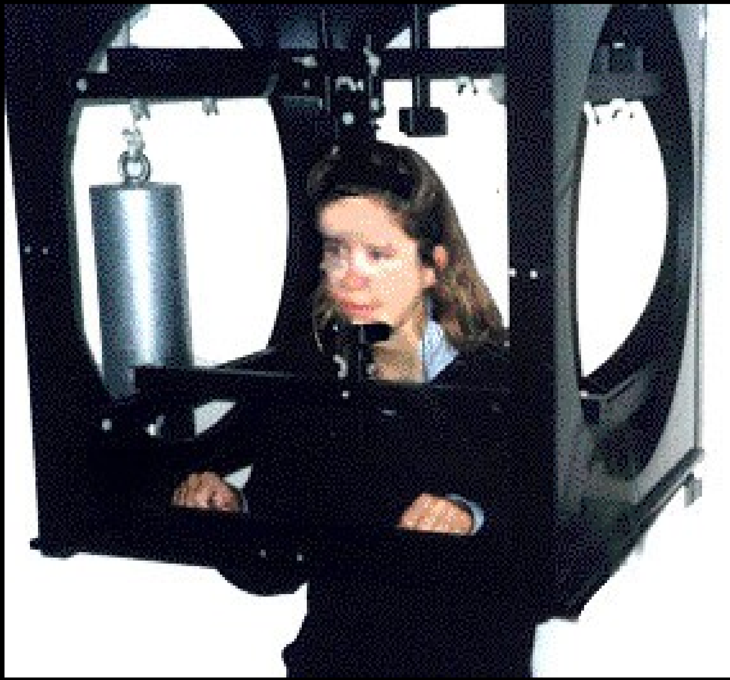
Mechanické metódy

Kontaktné šošovky

Cievka v magnetickom poli

Veľmi presné

Veľmi nepríjemné



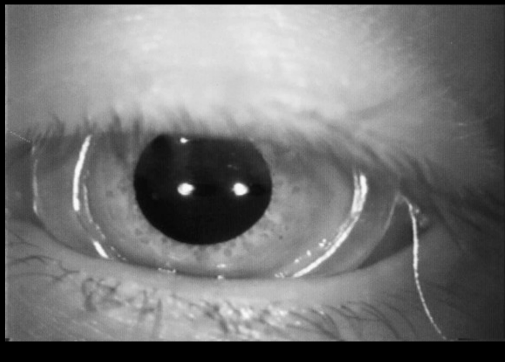
Rozdiel medzi MM a VM 0.56°, 0.78° a 0.18° v troch smeroch

Comparing the accuracy of video-oculography and the scleral search coil system in human eye movement analysis

Takao Imai[a,*], Kazunori Sekine[b], Kousuke Hattori[c], Noriaki Takeda[b], Izumi Koizuka[c], Koji Nakamae[d], Katsuyoshi Miura[d], Hiromu Fujioka[d], Takeshi Kubo[a]

[a] *Department of Otolaryngology and Sensory Organ Surgery, Osaka University, Graduate School of Medicine, Osaka, Japan*
[b] *Department of Otolaryngology, University of Tokushima, School of Medicine, Tokushima, Japan*
[c] *Department of Otolaryngology, St. Marianna University, School of Medicine, Kanagawa, Japan*
[d] *Department of Information Systems Engineering, Osaka University, Graduate School of Engineering, Japan*

Received 25 May 2004; received in revised form 29 September 2004; accepted 26 November 2004
Available online 18 January 2005



Jednobodové videometódy

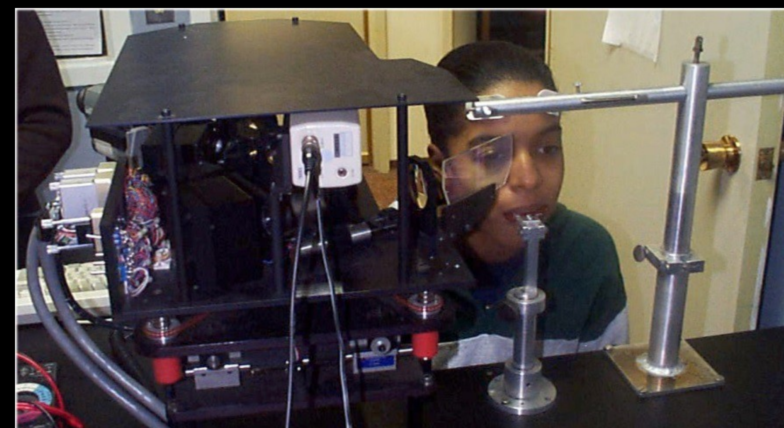
Sledovanie jednoho objektu vo videu:

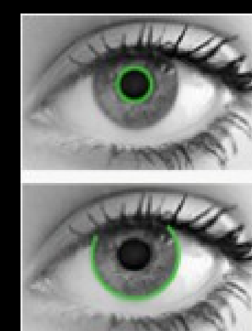
- Zrenička
- Dúhovka

SW analyzuje videosignál a sleduje vybraný objekt

Kalibračná procedúra

Prenosný, neprenosný

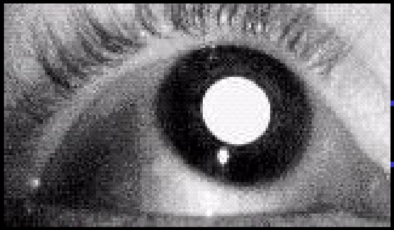




Dvojbodové videometódy

Sledujú dva objekty

Zrenička + odraz na rohovke



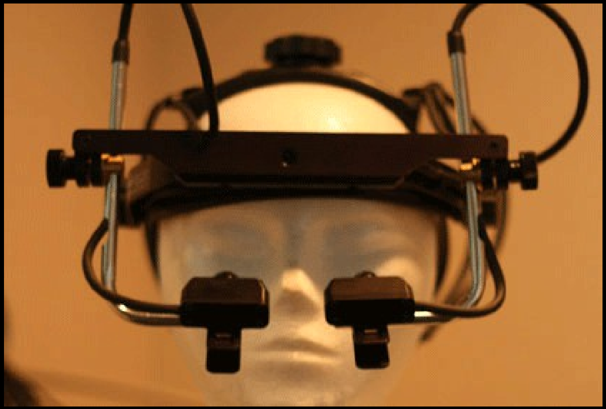
Výrazná zrenička

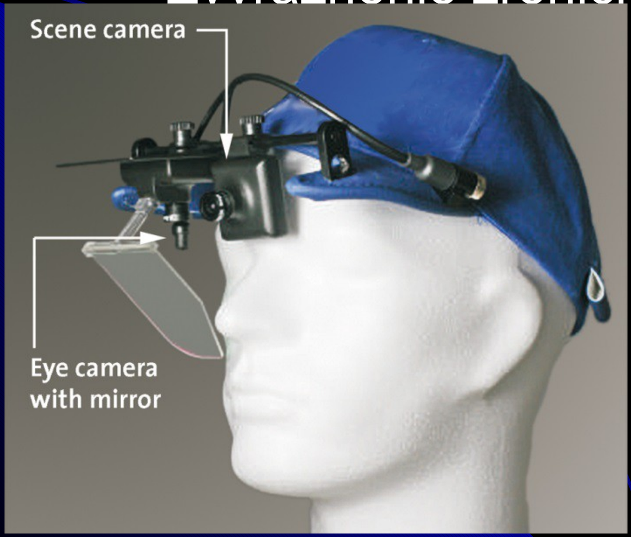
Odraz na rohovke

IR svetlo

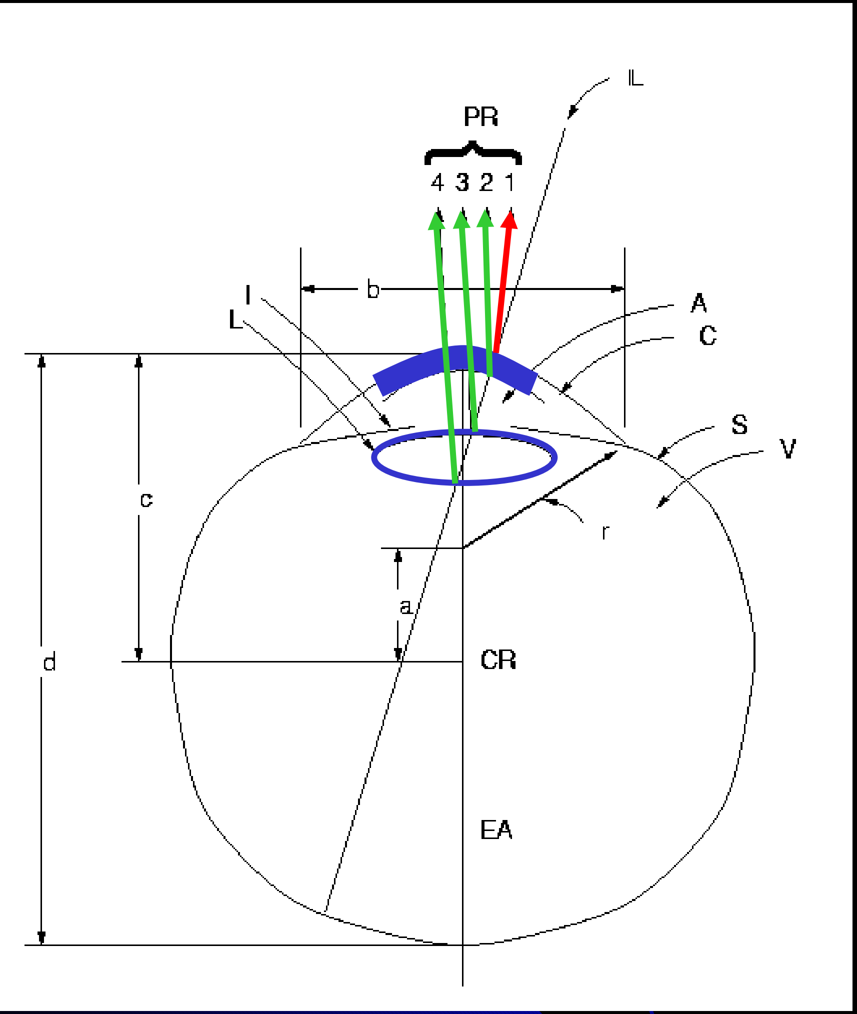
Odraz na rohovke

Zvýraznenie zreničky





Odraz na rohovke



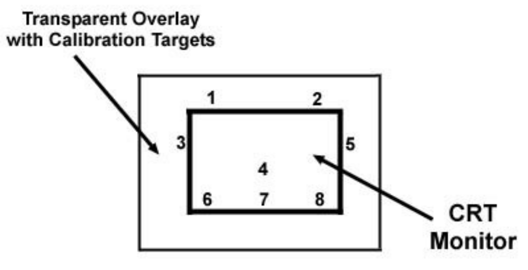
Purkiňove obrázky

Odraz na

1. Prednej strane rohovky
2. Zadnej strane rohovky
3. Prednej strane šošovky
4. Zadnej strane šošovky

Väčšina systémov sleduje 1.

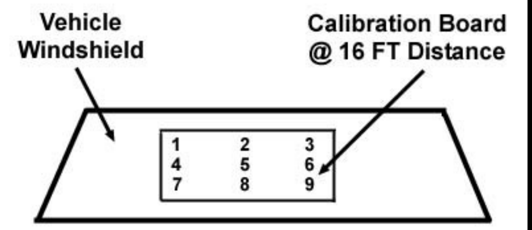
Odraz na rohovke



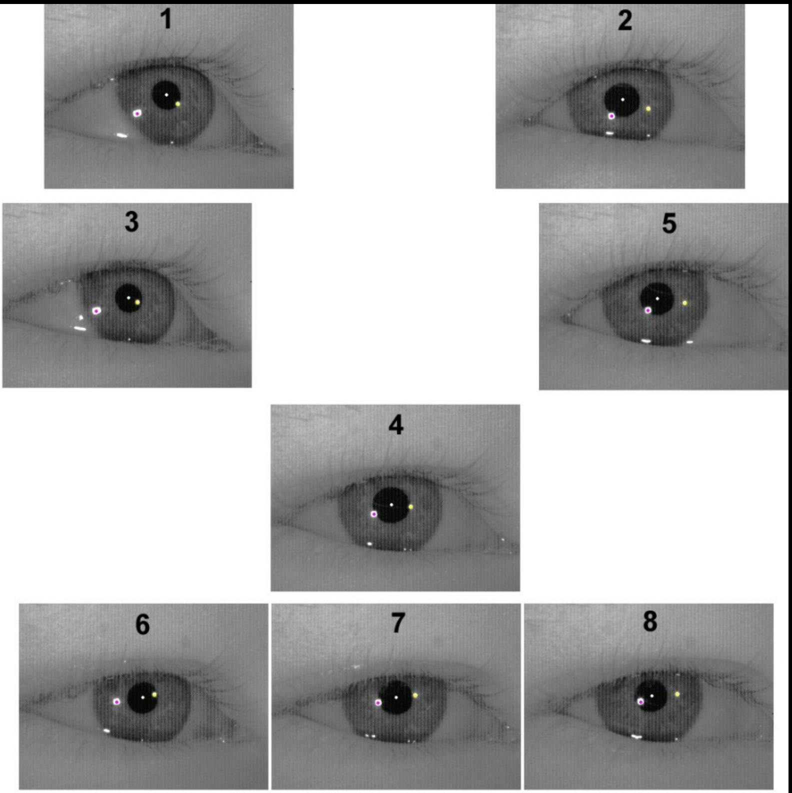
Calibration Target #1.
Computer Display

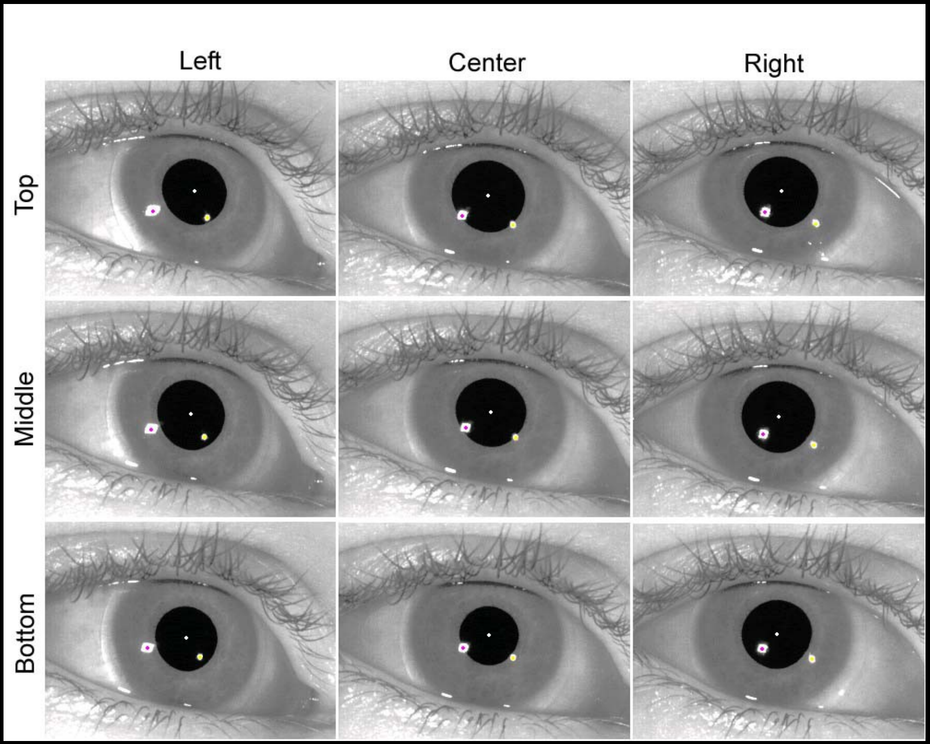
Poloha odrazu voči zreničke určuje smer pohľadu

kalibrácia

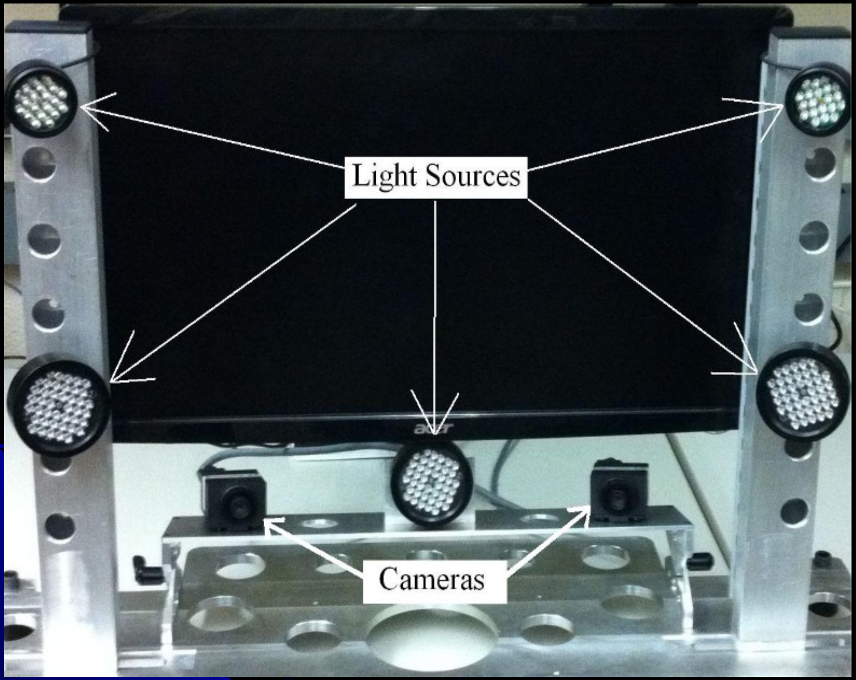


Calibration Target #2.
Reading Distant Highway Signs





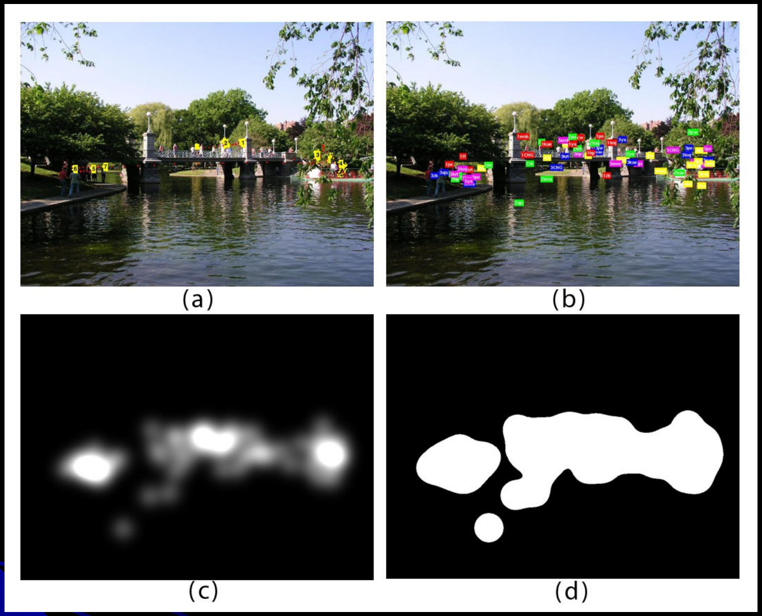
Bez kalibrácie



Viac kamier, viac svetiel

- Pozeráme sa tam kam klikáme
- Pozeráme sa na významné oblasti
- Video z modelovanou saliency
- Vzor pozornosti od iného používateľa
- Stereo sledovanie

Top down Saliency



Saliency model trénovaný na dátach z eyetrackera.

Nízka: intenzita, farba, orientácia

Stredná: horizont, Center prior

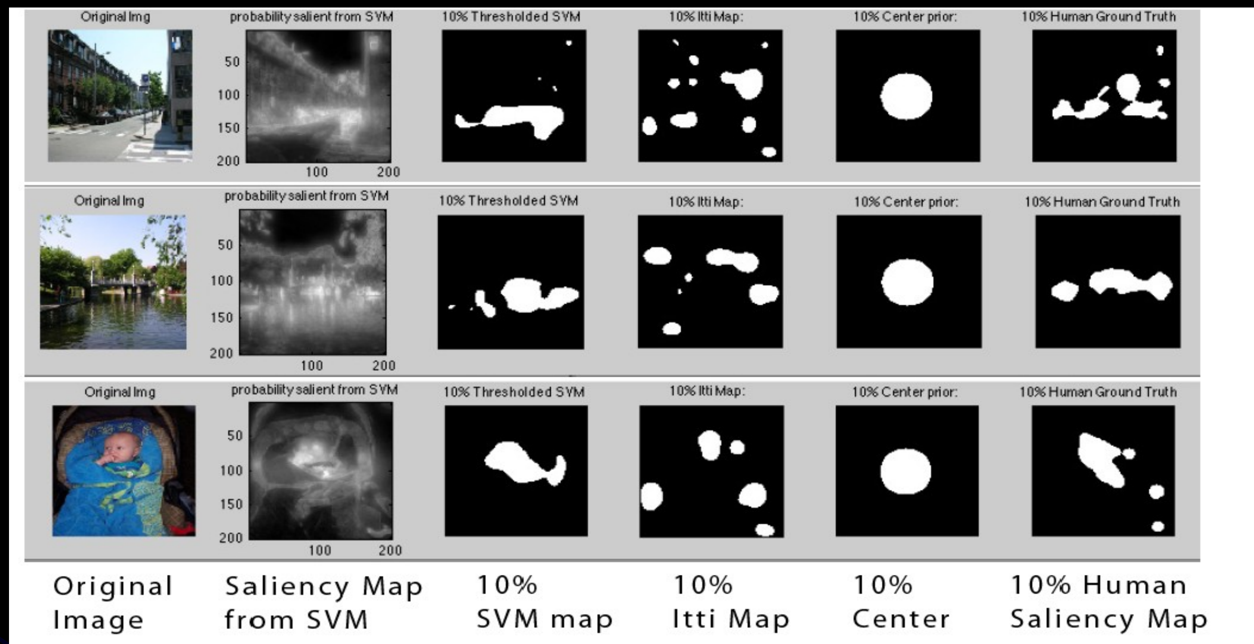
Vysoká: tváre, ľudia

SVM na prahovaných gaze mapách

Learning to Predict Where Humans Look

Tilke Judd tjudd@mit.edu
Krista Ehinger kehinger@mit.edu
Frédo Durand fredo@csail.mit.edu
Antonio Torralba torralba@csail.mit.edu
MIT Computer Science Artificial Intelligence Laboratory and MIT Brain and Cognitive Sciences

Top down Saliency



Learning to Predict Where Humans Look

Tilke Judd tjudd@mit.edu
Krista Ehinger kehinger@mit.edu
Frédo Durand fredo@csail.mit.edu
Antonio Torralba torralba@csail.mit.edu
MIT Computer Science Artificial Intelligence Laboratory and MIT Brain and Cognitive Sciences

Termíny

Presnosť

- Očakávaný rozdiel medzi skutočnou pozíciou oka a vypočítanou
- Väčšinou do 1°

Priestorové rozlíšenie

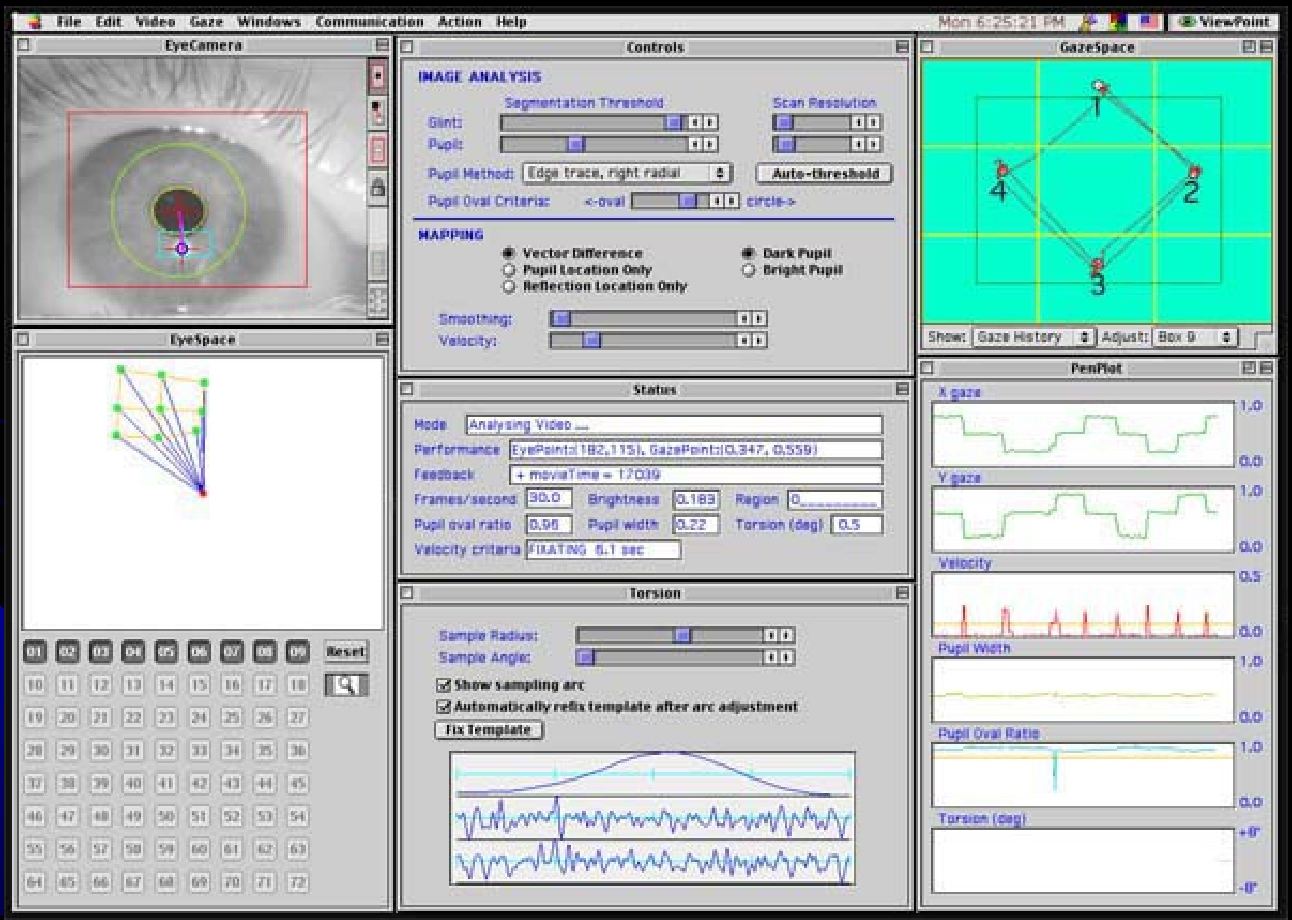
- Najmešia zachytiteľná zmena pozície oka (0.03-0.2°)

Časové rozlíšenie

- Počet zachytených pozícií oka za sekundu (30-500Hz)

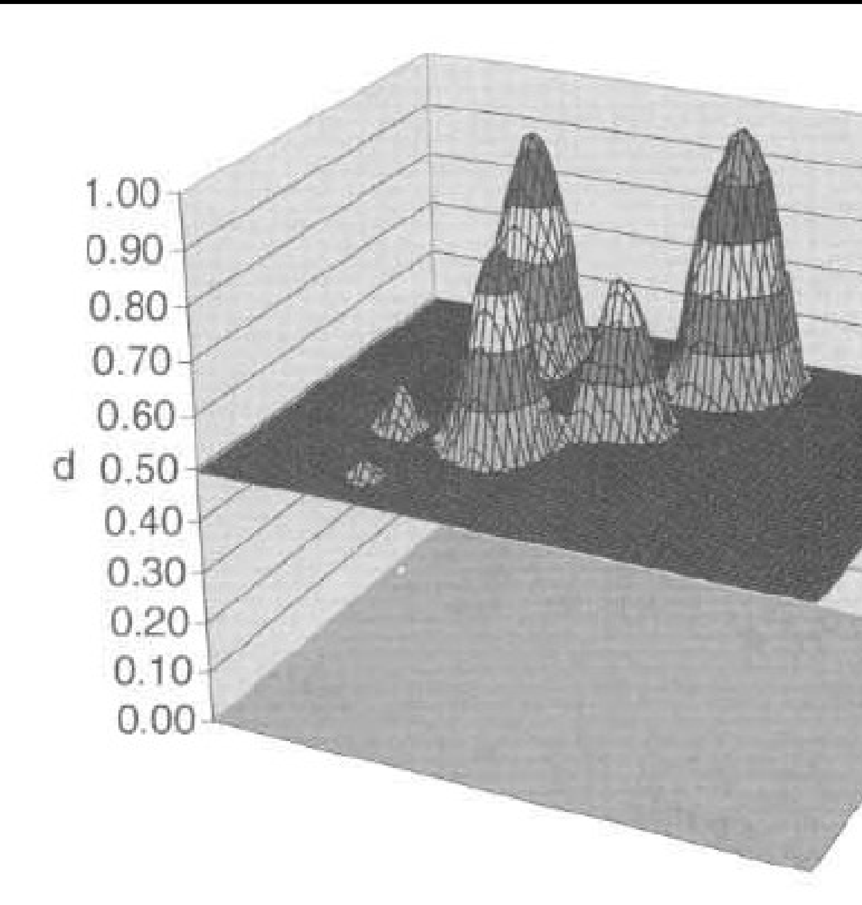
Spracovanie dát

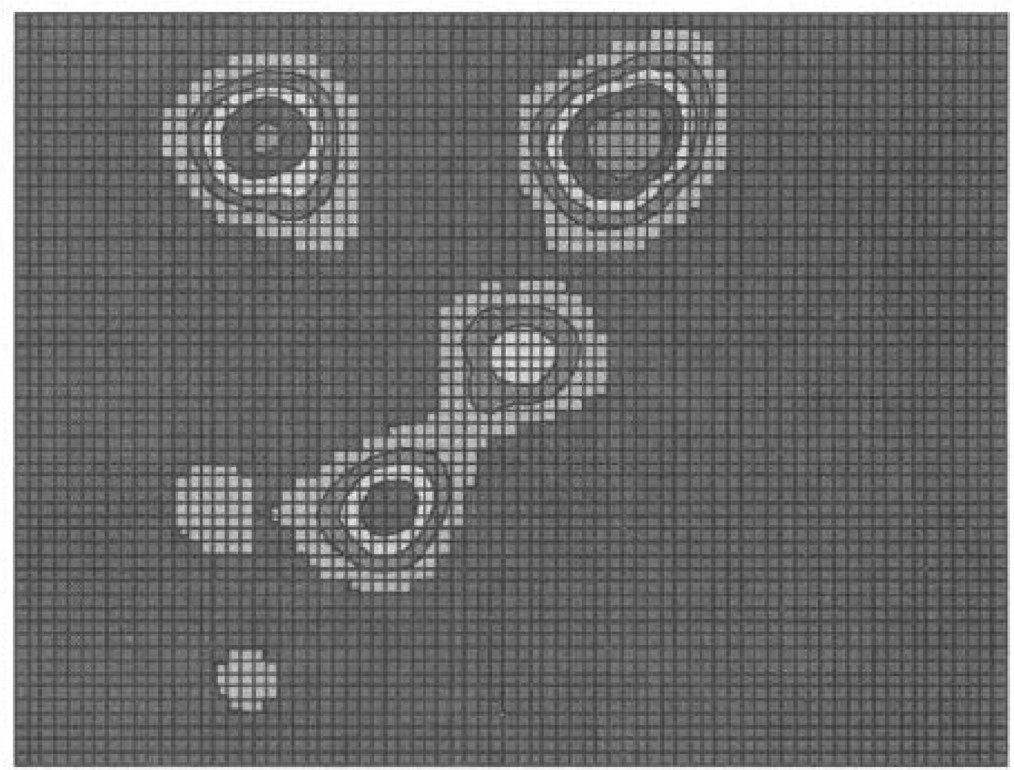
Dáta priamo z eye-trackera nepoužiteľné
Treba ich analyzovať, vizualizovať



Vizualizácia

Teplotné mapy





Fixation Maps: Qu
movement Traces
David S. Wooding

Teplotné mapy

Teplotné mapy

Paul Cairns, York Univers

http://www.gro

Miera rozdielnosti teplotných máp

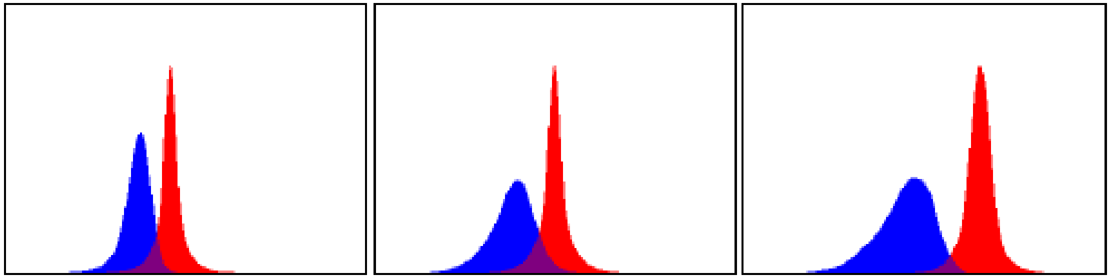
Kullback-Leiblerova divergencia

Teplotná mapa = hustota pravdepodobnosti

$D(f||g) \geq 0;$

$D(f||g) = 0$ iff $f(x) = g(x)$ for all $x \in X;$

$D(f||g) \neq D(g||f);$

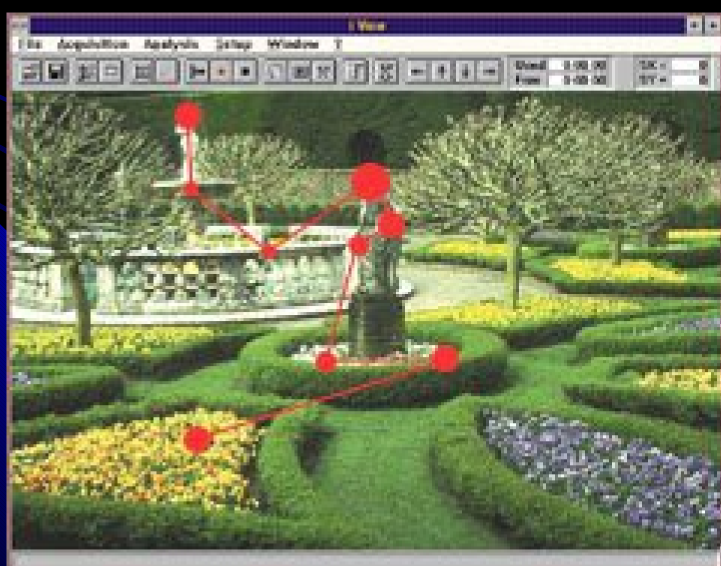


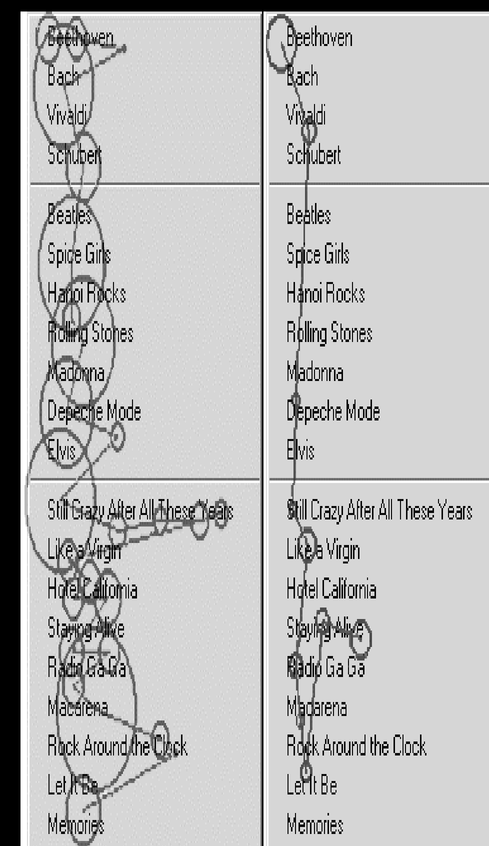
(a) KL=2.944 (b) KL=3.246 (c) KL=10.967

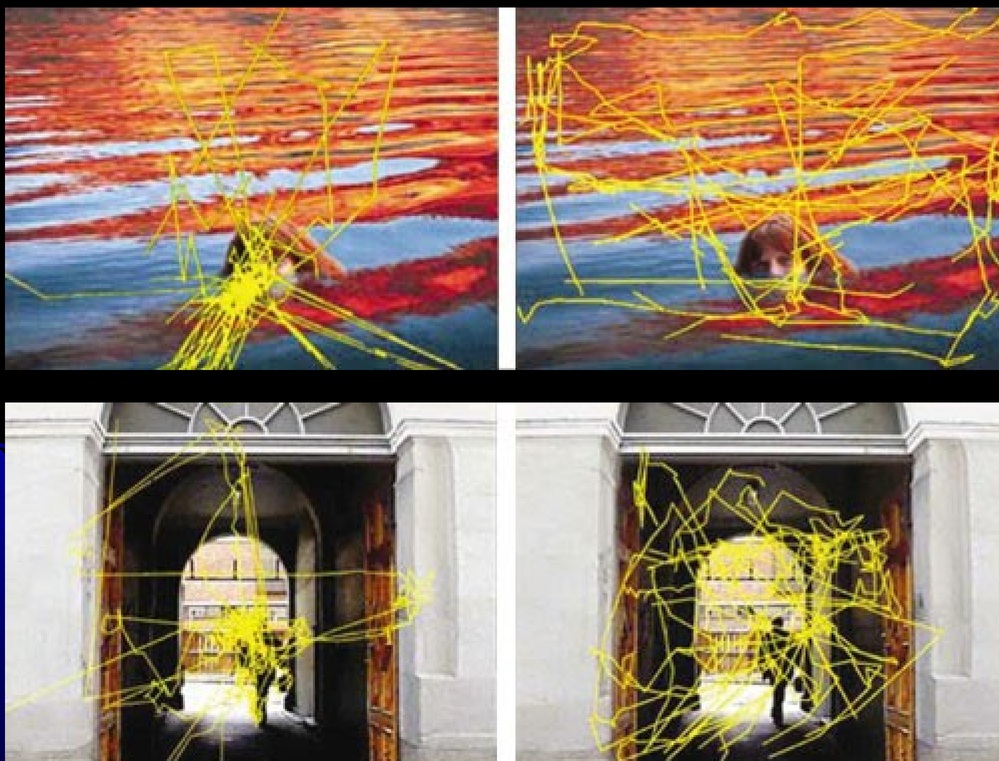
$$KLD(p,q) = \sum_{x \in Img} (p(x) - q(x)) * \log(p(x)/q(x))$$

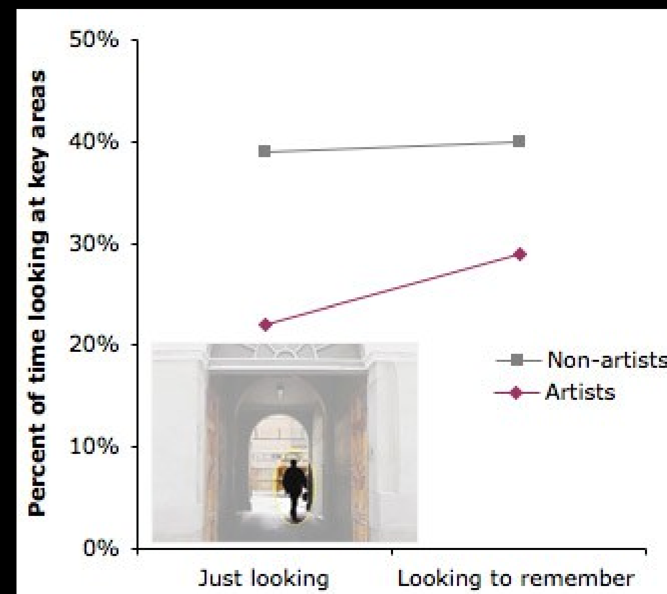
Trajektória pohľadu s fixáciami

Kruhy označujú fixácie
Polomer určuje dobu strávenú pri danej fixácii
Čiarou spojené kruhy – trajektória, časová postupnosť fixácií, určuje sakadický pohyb









http://gurneyjourney.blogspot.sk/2009/11/do-artists-see-differently.html

Fixácie určujú, kam sa človek pozerá, ale nie, či daný obraz vníma

HDR

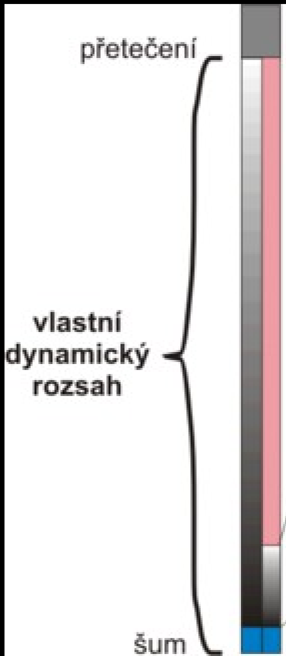
Dynamický rozsah

Pomer maximálnej a minimálnej hodnoty merania

Scéna: pomer najjasnejšieho a najtmavšieho miesta v scéne

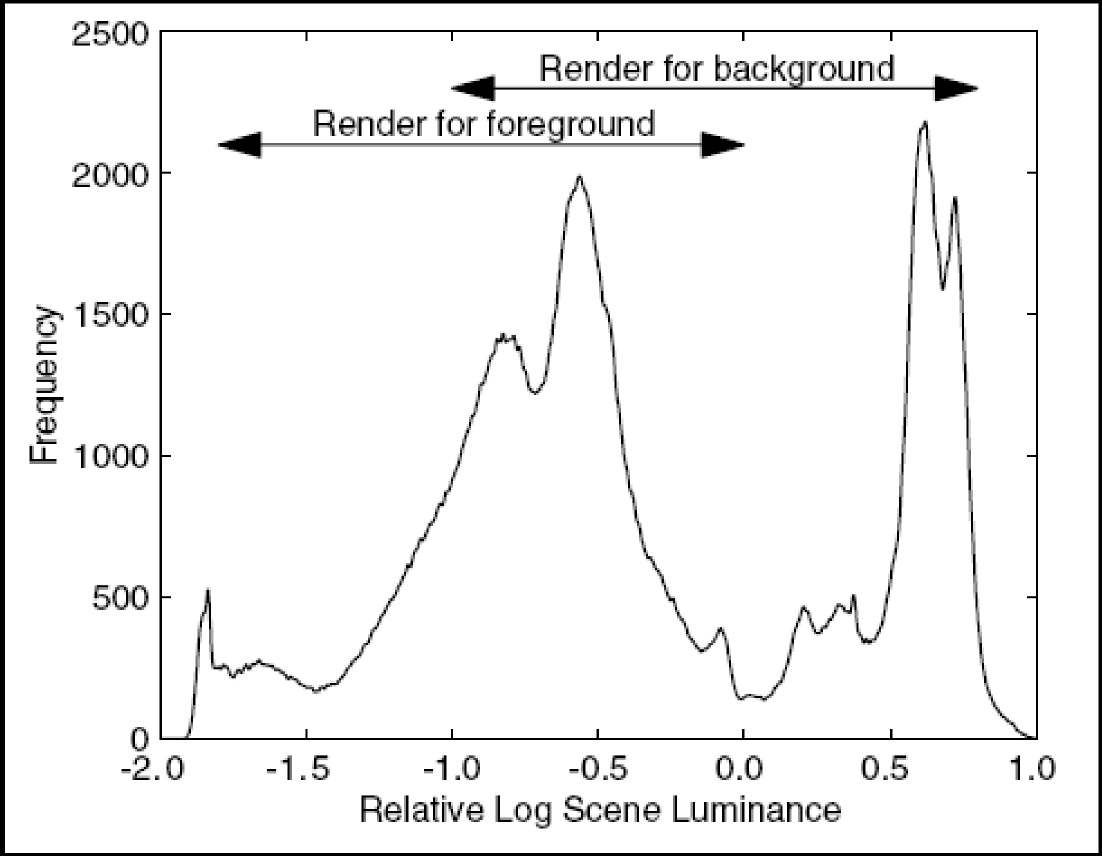
Fotoaparát: pomer nasýtenia (pretečenia senzoru) a šumu

Displej: pomer maximálnej a minimálnej vyžarovanej intenzity

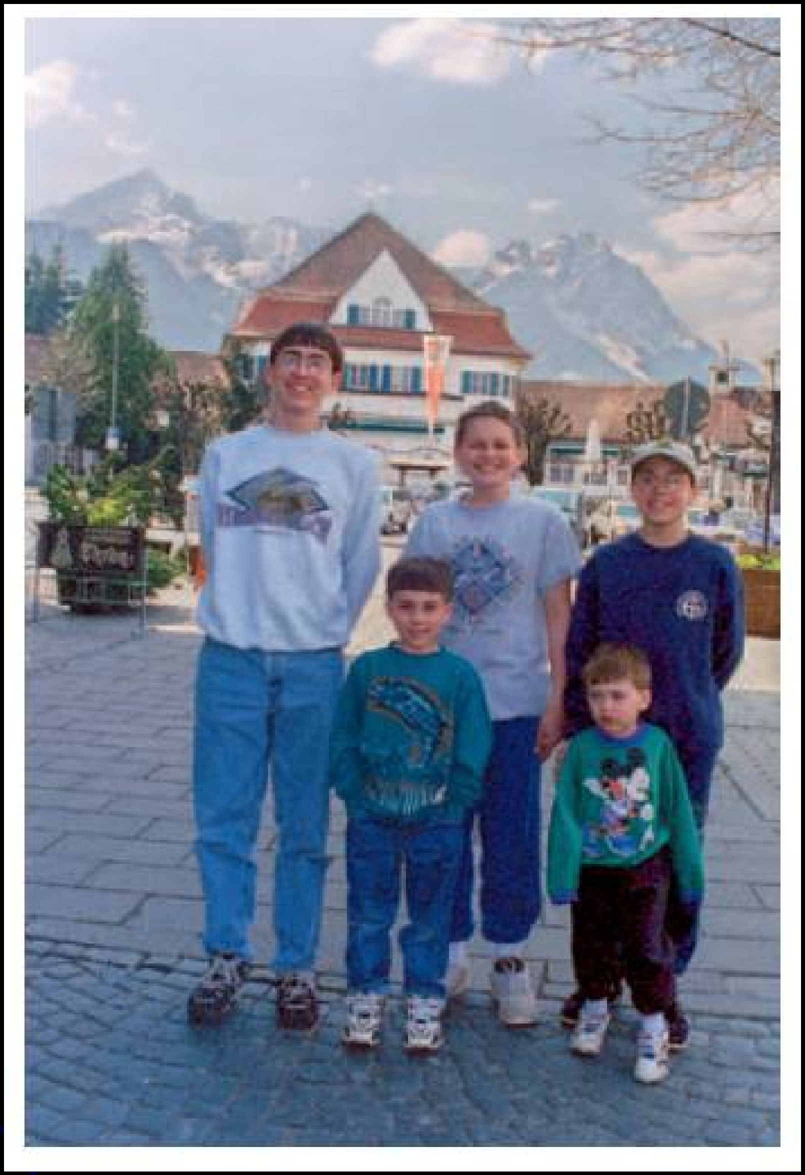


Správna expozícia scény záleží na obsahu scény a dynamickom rozsahu

Sce



Histogram log-intenzít scény
Na fotku dokážeme preniesť 1.8 log-jednotiek



Digitálna úprava “dodge-and-burn”

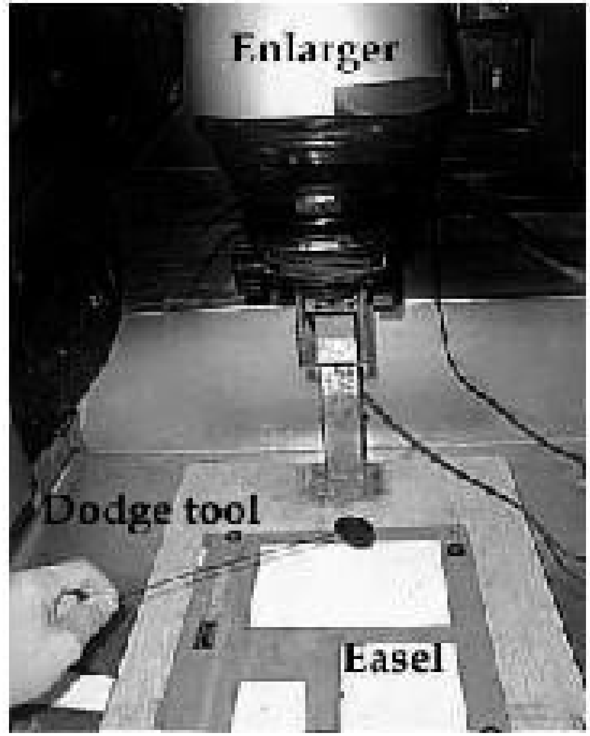
1850 Gustave Le Gray

Kombinácia dvoch negatívov

Ansel Adams
Clearing Winter Storm,
Yosemite National Park

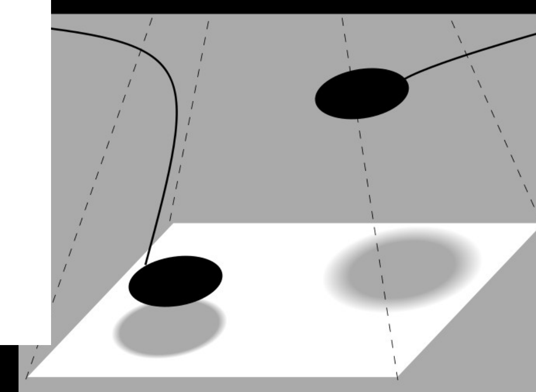
Burning In
stmavenie časti fotografie
Dodging
zjasnenie časti fotografie

prekrytie miest, na ktoré nemá dopadať toľko svetla pri vyvolaní



The **Burn in Mask** is always kept in motion while making an exposure. The light is ìpaintedî on the photo to make a portion of the print darker.

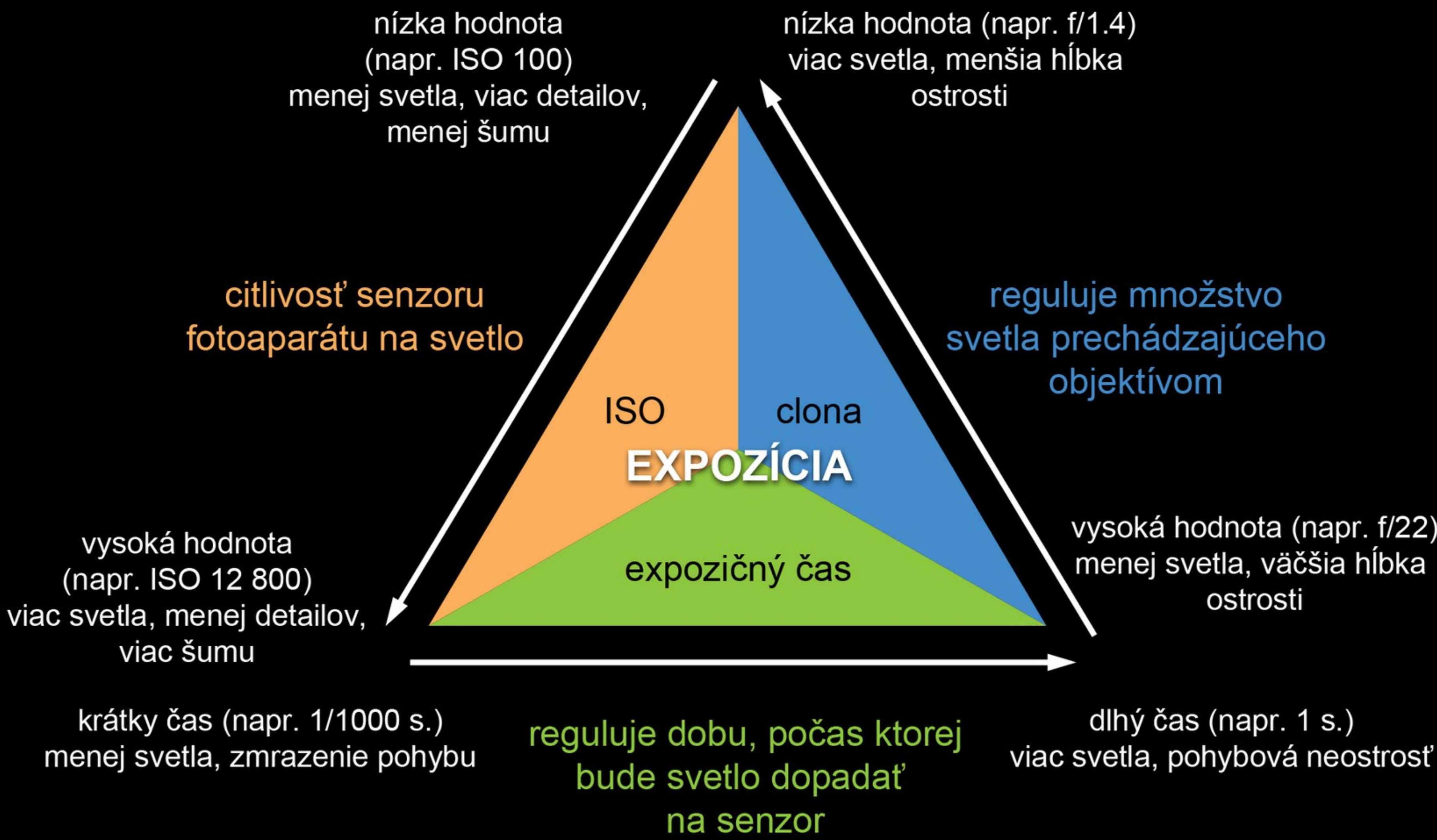
The **dodge tool** is a piece of plastic or cardboard on the end of a metal wand. This tool is also kept in motion during exposure. It is used to hold back light from a portion of a print to make it lighter.



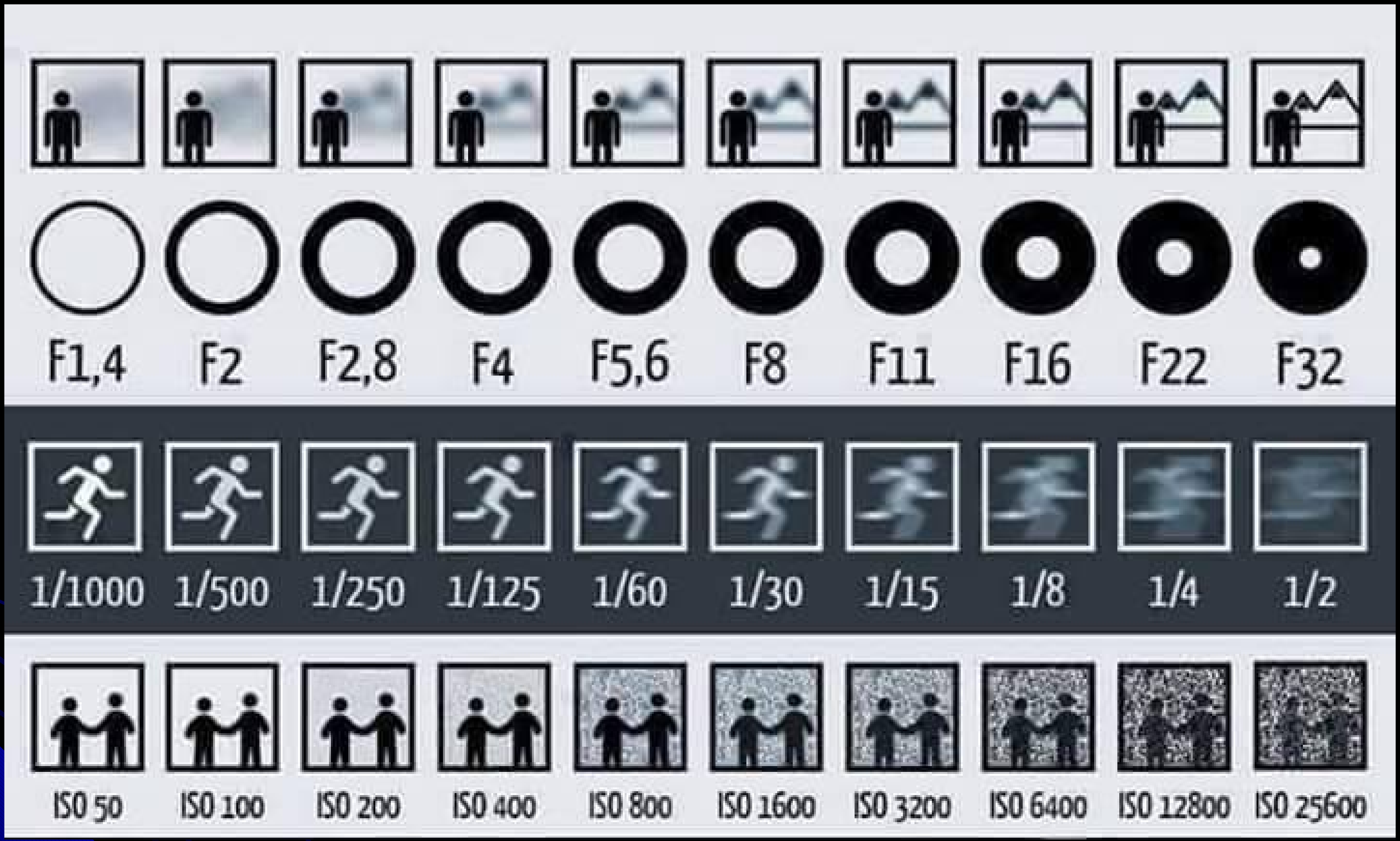
Dynamický rozsah fotoaparátu

CCD chip	1:1000
čas	1:100
clona	~1:100
citlivosť (ISO)	~1:10
Celkový operačný rozsah	1:100,000,000

Jediný obrázok 1:1000



http://www.ephoto.sk/fotoskola/cla

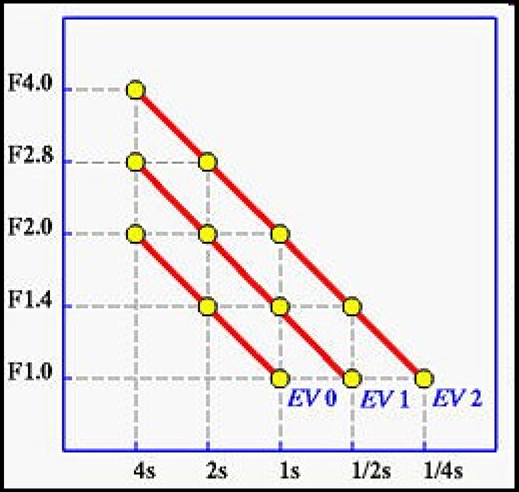


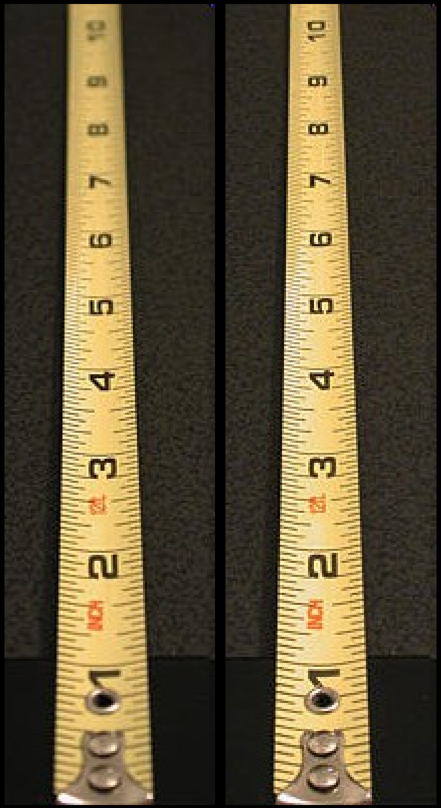
http://www.lifehack.org/articles/lifestyle/picture-show-you-clearly-the-effects-aperture-shutter-speed-and-iso-image

$$EV = \log_2 \left(\text{Aperture}^2 * \frac{1}{\text{Shutter speed}} * \frac{\text{ISO}}{100} \right)$$

EV+1 … s/2 alebo a*√2

EV-1 … 2*s alebo a/√2





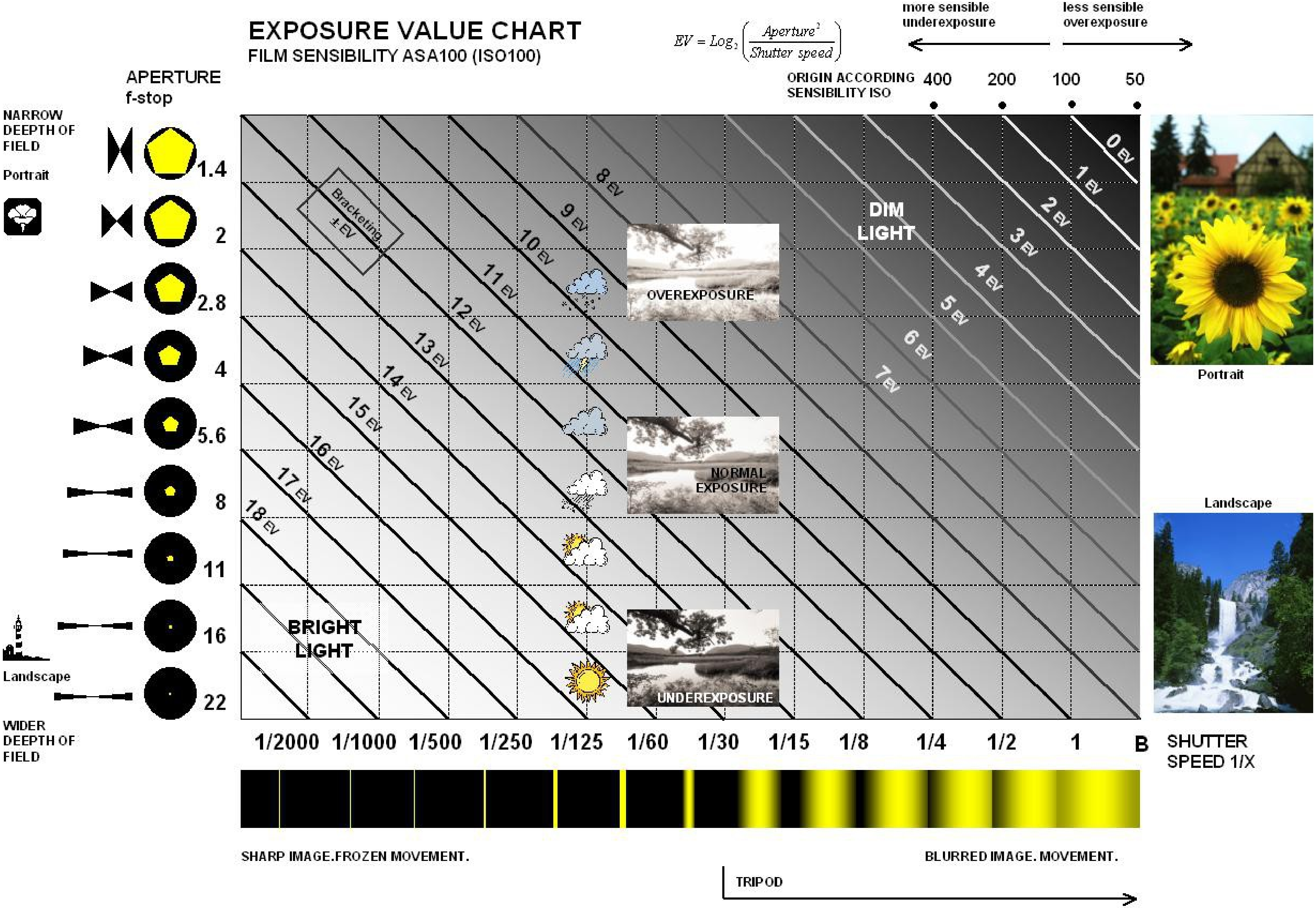
Fixná clona a => rovnaká hĺbka ostrosti

Expozičný čas:

¼, 1/8, 1/15, 1/30, 1/60, 1/125, 1/250, 1/500, 1/1000 sec

V skutočnosti:

¼, 1/8, 1/16, 1/32, 1/64, 1/128, 1/256, 1/512, 1/1024 sec

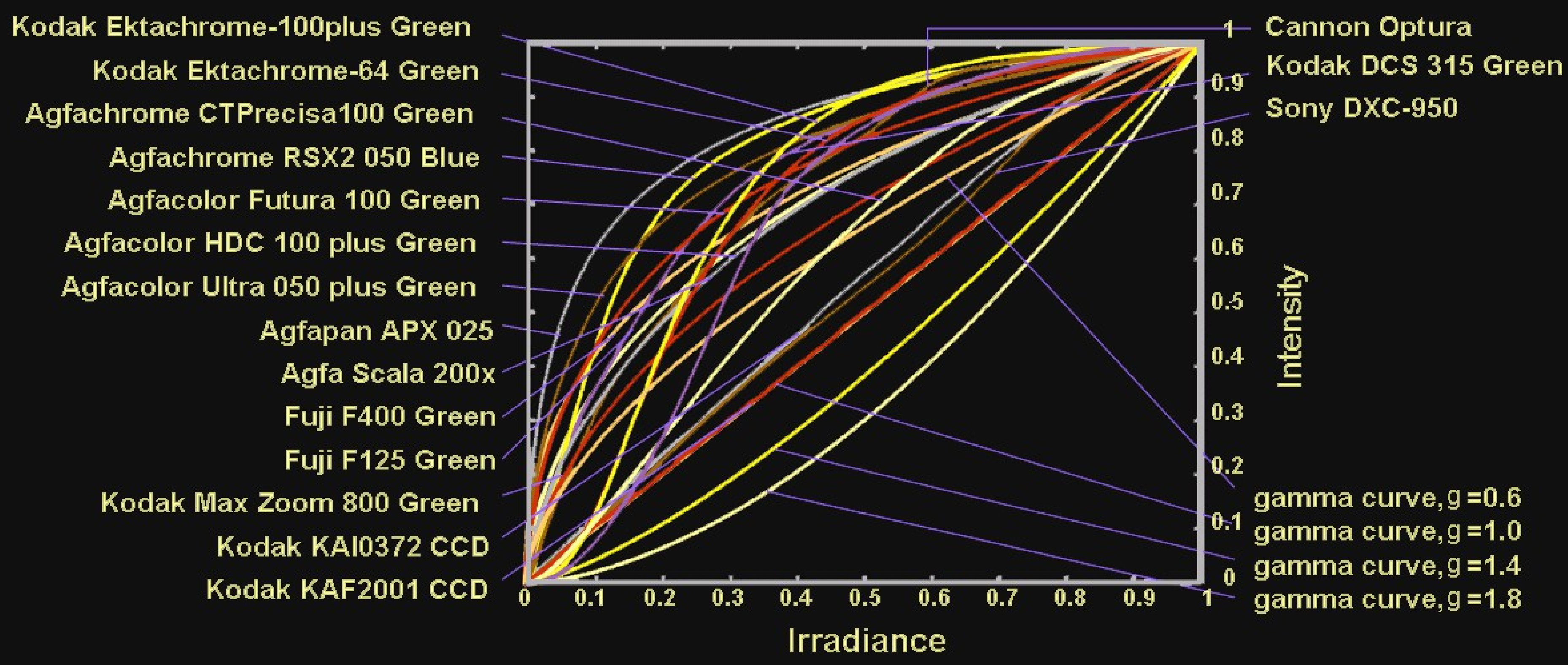


http://www.ephoto.sk/fotoskola/cla

Figure 3: Sequence of exposure-bracketed image, separated by one f-stop. The shortest shown exposure time is 1/60 sec (top left), the longest 15 sec (bottom right). Aperture fixed at 2.6, ISO 100. Taken with Nikon CoolPix 995. *Source*: Axel Jacobs

Mapovanie osvetlenia scény na hodnoty intenzity – krivka reakcie na svetlo

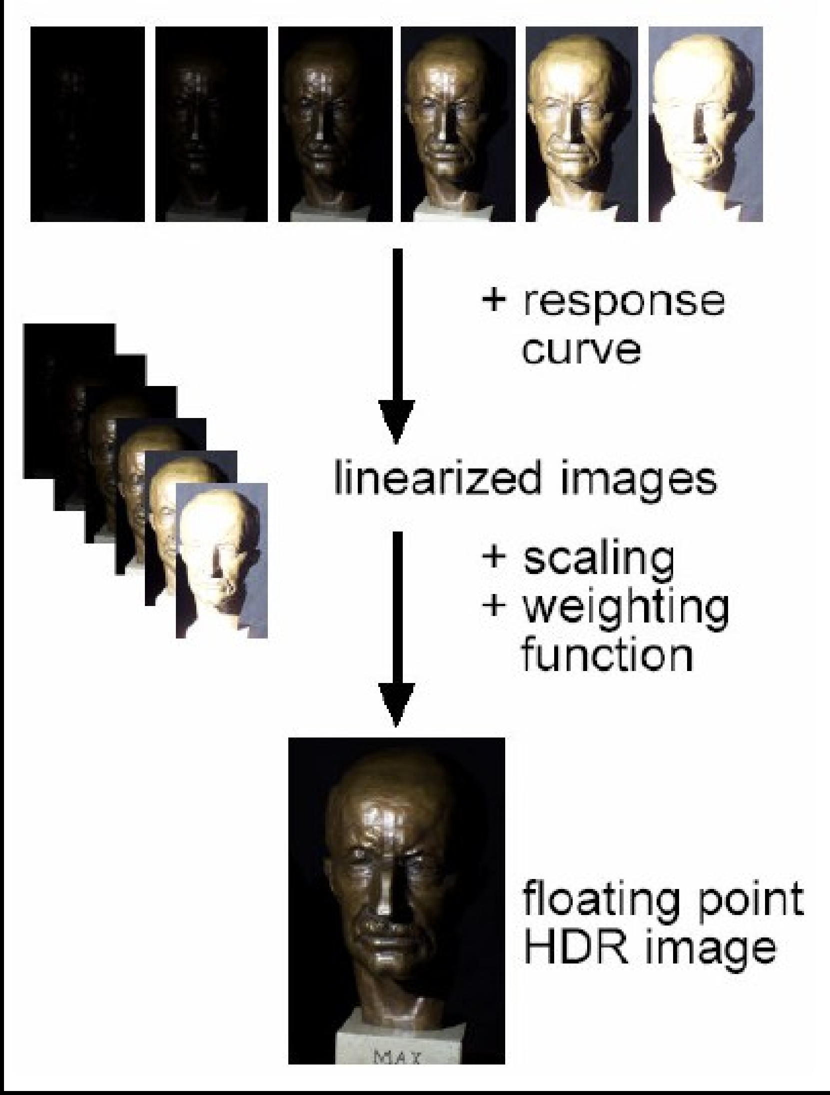
Nie je známa, líši sa podľa výrobcu, modelu, fotoaparátu...



Zvýšenie DR

Technika viacerých expozícií
- využíva operačný rozsah
- kombinuje viacero fotiek do jednej

Problémy:
- neznáma reakcia fotoaparátu na svetlo
- neznáma šumová charakeristika



Recovering High Dynamic Range Radiance Maps from Photographs
Paul E. Debevec, Jitendra Malik, 1997

vstup: fotky scény z rovnakého miesta s rôznymi časmi expozície Δtj

predpoklad: scéna je statická, proces snímania je rýchly => osvetlenie scény v bode i: Ei je konštantné

Zij – hodnota intenzity v bode i pri čase Δtj

$$Z_{ij} = f(E_i \Delta t_j)$$

$$f^{-1}(Z_{ij}) = E_i \Delta t_j$$

$$\ln f^{-1}(Z_{ij}) = \ln E_i + \ln \Delta t_j$$

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j$$

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j$$

minimalizujem

$$\mathcal{O} = \sum_{i=1}^{N} \sum_{j=1}^{P} [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} g''(z)^2$$

$$g''(z) = g(z-1) - 2g(z) + g(z+1)$$

kvôli hladkosti krivky g(z)

$$w(z) = \begin{cases} z - z_{\min} & z \le (z_{\min} + z_{\max})/2 \\ z_{\max} - z & z > (z_{\min} + z_{\max})/2 \end{cases}$$

Dôraz na hladkosť v strede krivky

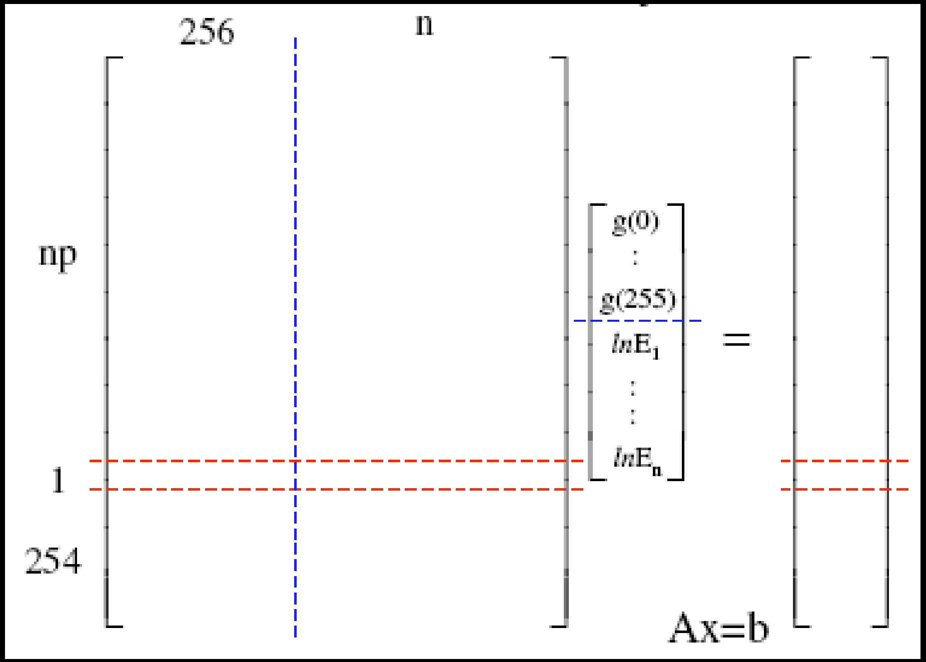
$$\mathcal{O} = \sum_{i=1}^{N} \sum_{j=1}^{P} \{w(Z_{ij}) [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]\}^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} [w(z) g''(z)]^2$$

Neznáme:
Ei
g(zij)

$$g(Z_{mid}) = 0, \text{ where } Z_{mid} = \tfrac{1}{2}(Z_{min} + Z_{max})$$

stredná intenzita zodpovedá jednotkovej expozícii

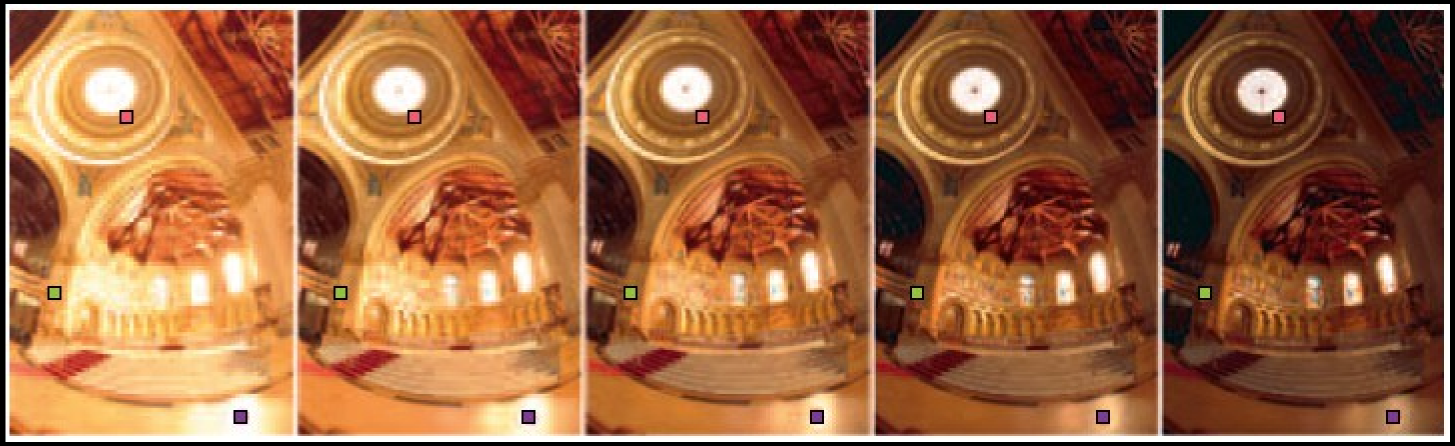
len N bodov
v P fotografiách

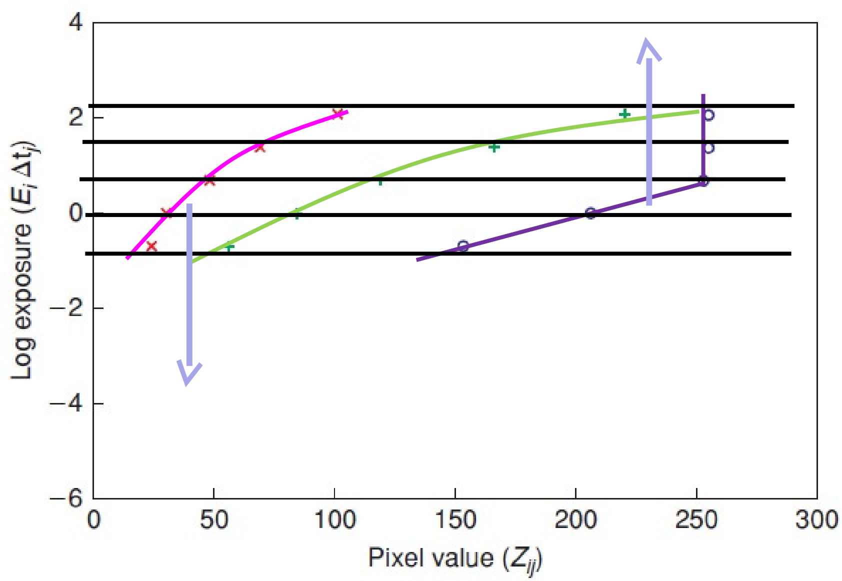


Správne určený systém rovníc
N(P - 1) > (Zmax – Zmin)

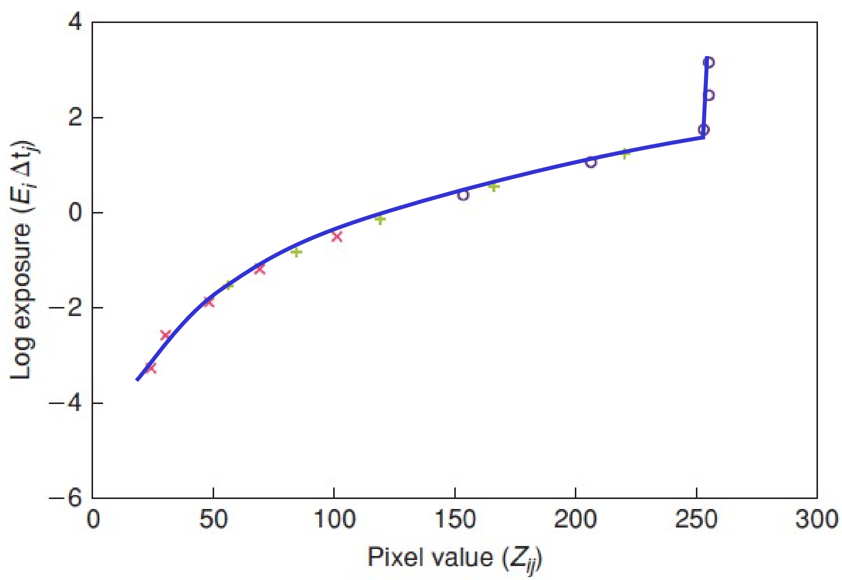
ak (Zmax - Zmin) = 255, P = 11
N ≥ 50

Výber bodov: vhodné rozloženie intenzít z intervalu Zmin to Zmax
zároveň vhodne priestorovo rozložené v obraze v oblastiach relatívne konštantnej intenzity





Log exposure (E_i Δt_j)
Pixel value (Z_{ij})
4
2
0
−2
−4
−6
0
50
100
150
200
250
300



Log exposure (E_i Δt_j)
Pixel value (Z_{ij})
4
2
0
−2
−4
−6
0
50
100
150
200
250
300

Figure 3: **(a)** *Eleven grayscale photographs of an indoor scene acquired with a Kodak DCS460 digital camera, with shutter speeds progressing in 1-stop increments from* $\frac{1}{30}$ *of a second to 30 seconds.*

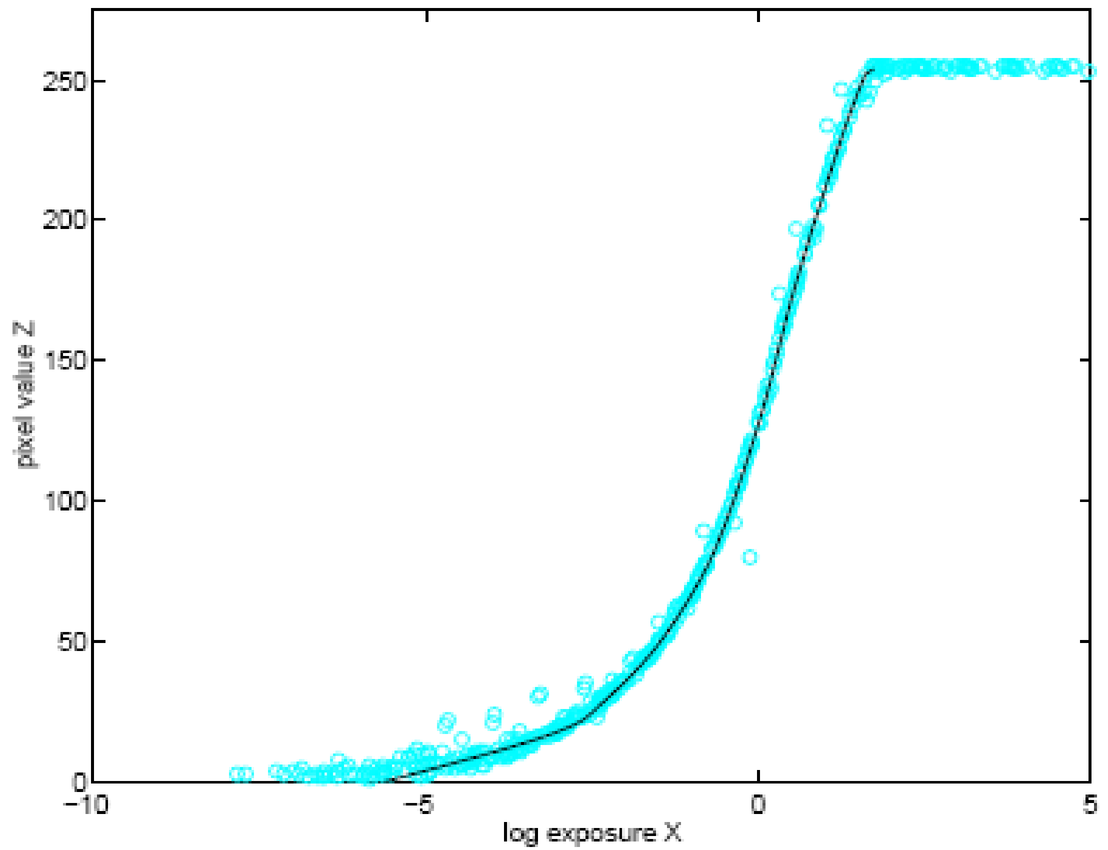


Figure 4: *The response function of the DCS460 recovered by our algorithm, with the underlying* $(E_i \Delta t_j, Z_{ij})$ *data shown as light circles. The logarithm is base* e.

$$\ln E_i = g(Z_{ij}) - \ln \Delta t_j$$

$$\ln E_i = \frac{\sum_{j=1}^{P} w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^{P} w(Z_{ij})}$$

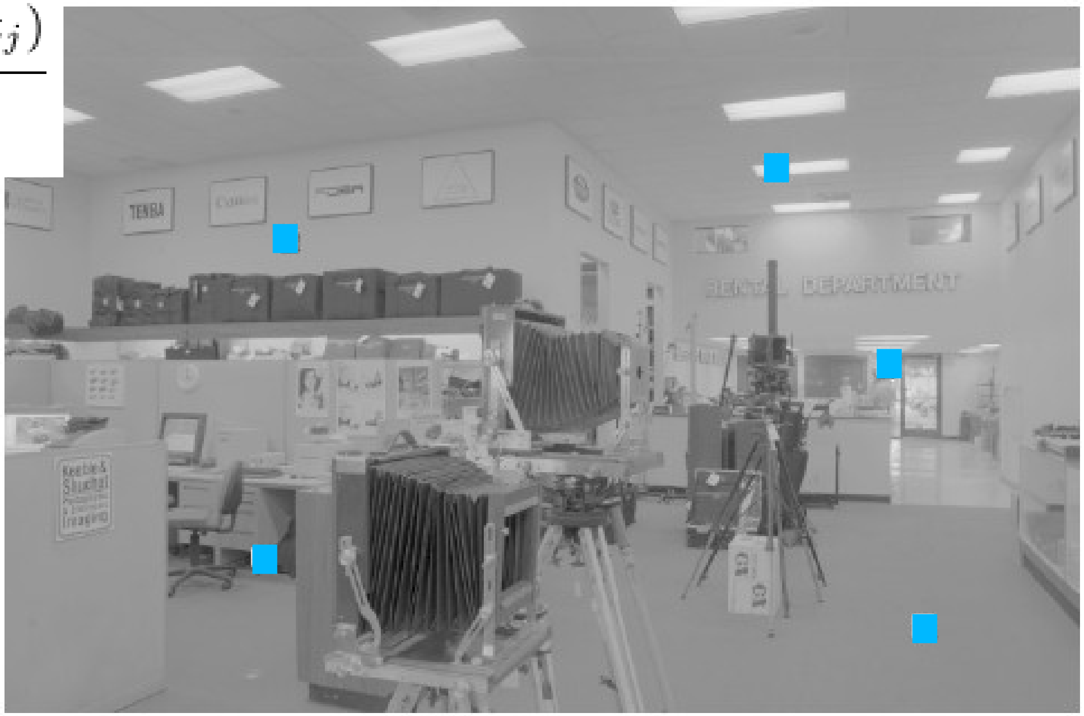
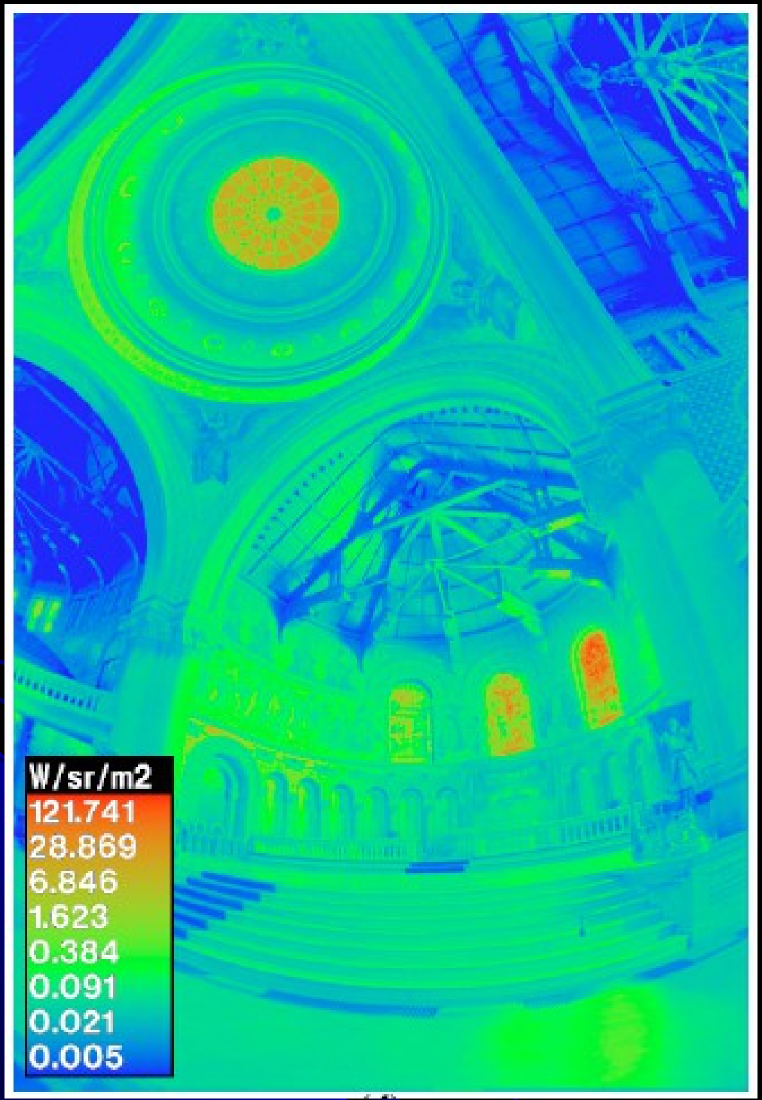
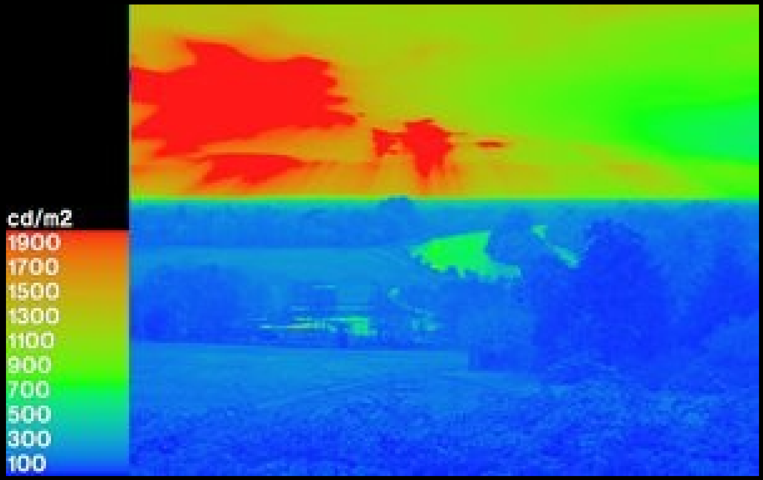


Figure 5: *The reconstructed high dynamic range radiance map, mapped into a grayscale image by taking the logarithm of the radiance values. The relative radiance values of the marked pixel locations, clockwise from lower left: 1.0, 46.2, 1907.1, 15116.0, and 18.0.*

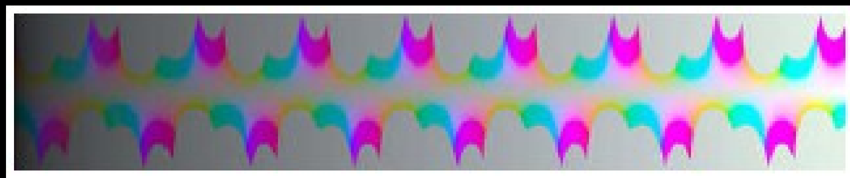

W/sr/m2
121.741
28.869
6.846
1.623
0.384
0.091
0.021
0.005


cd/m2
1900
1700
1500
1300
1100
900
700
500
300
100


cd/m2
1900
1700
1500
1300
1100
900
700
500
300
100

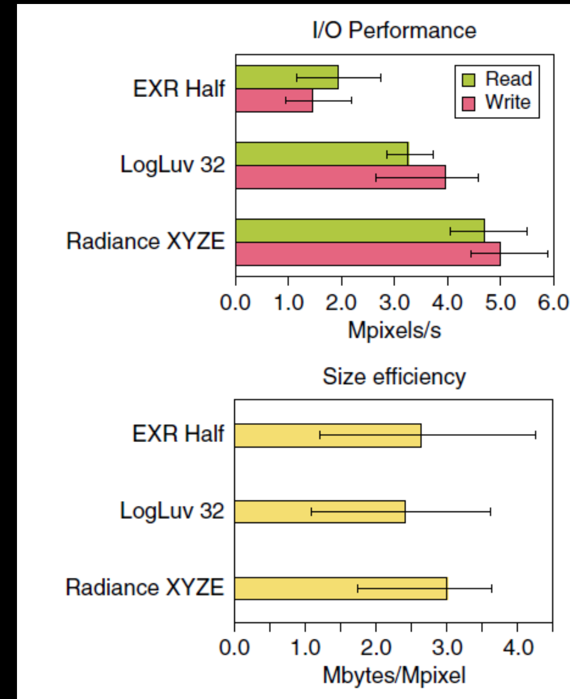
Encoding	Color Space	Bits/pixel	Dynamic Range ($\log_{10}$)	Relative Error (%)
sRGB	RGB in [0,1] range	24	1.6 orders	variable
RGBE	positive RGB	32	76 orders	1.0
XYZE	(CIE) XYZ	32	76 orders	1.0
IEEE RGB	RGB	96	79 orders	0.000003
LogLuv24	$\log Y + (u', v')$	24	4.8 orders	1.1
LogLuv32	$\log Y + (u', v')$	32	38 orders	0.3
Half RGB	RGB	48	10.7 orders	0.1

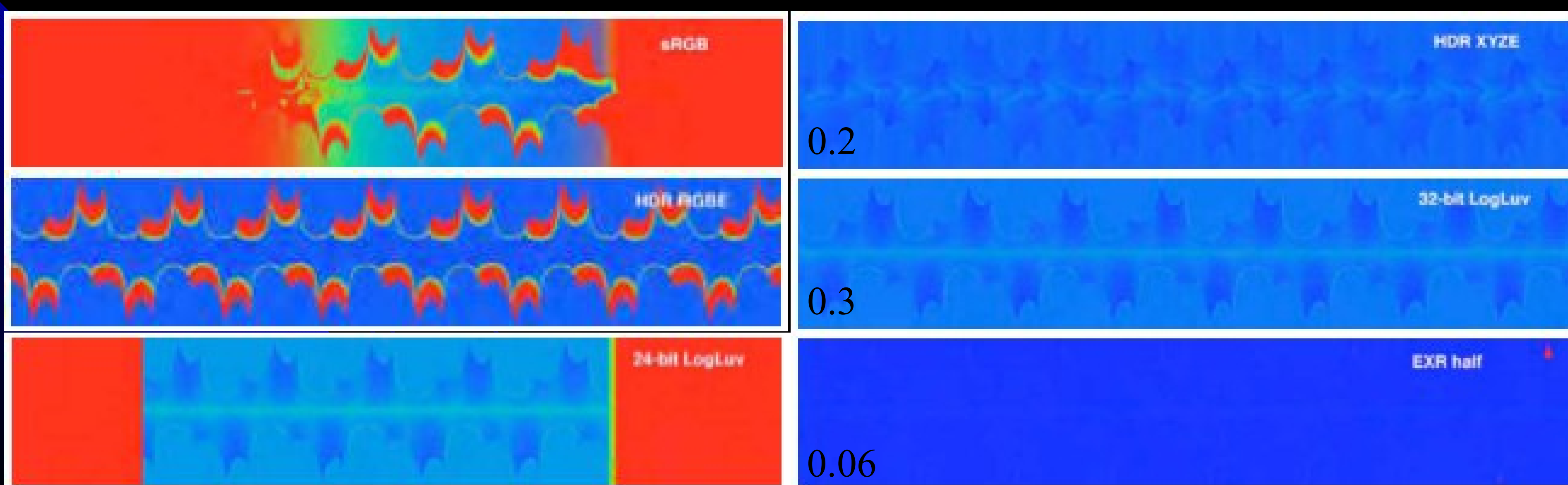
TABLE 3.2 HDR pixel encodings in the order of introduction.

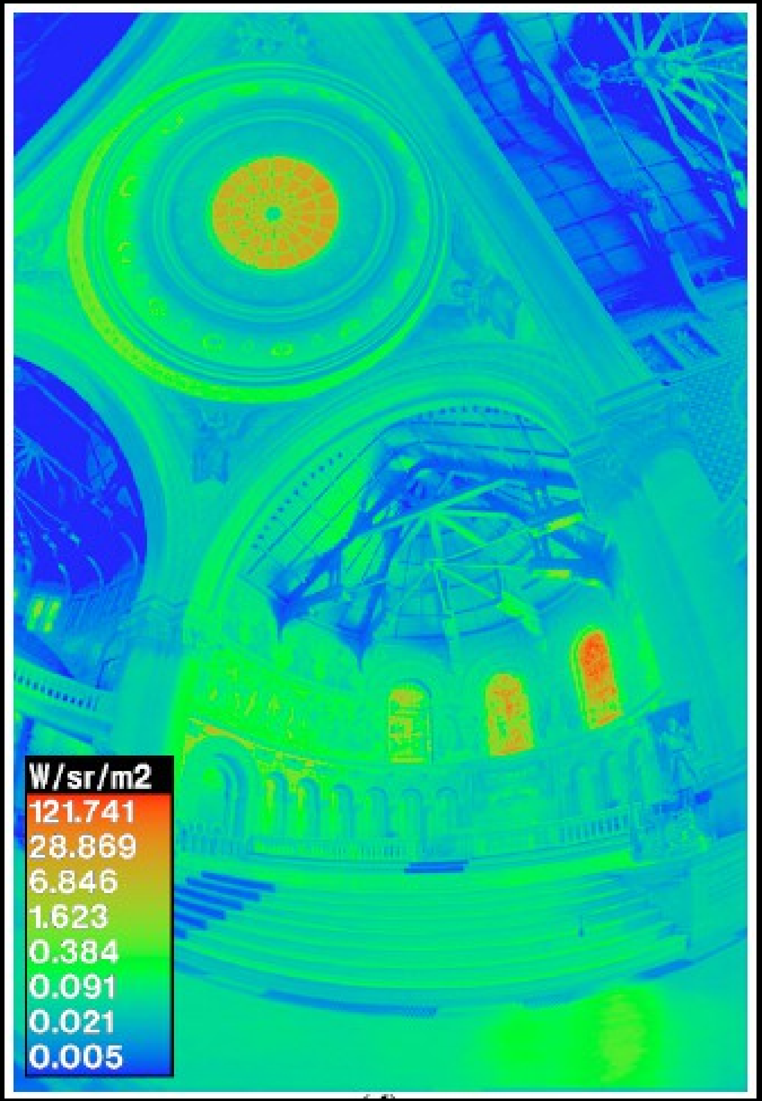


8 rádov


0.5
1.5
2.5
3.5
4.5
5.5
6.5
7.5
CIE ΔE*


I/O Performance
EXR Half
LogLuv 32
Radiance XYZE
Read
Write
0.0 1.0 2.0 3.0 4.0 5.0 6.0
Mpixels/s
Size efficiency
EXR Half
LogLuv 32
Radiance XYZE
0.0 1.0 2.0 3.0 4.0
Mbytes/Mpixel


sRGB
HDR RGBE
24-bit LogLuv
HDR XYZE
0.2
32-bit LogLuv
0.3
EXR half
0.06


W/sr/m2
121.741
28.869
6.846
1.623
0.384
0.091
0.021
0.005

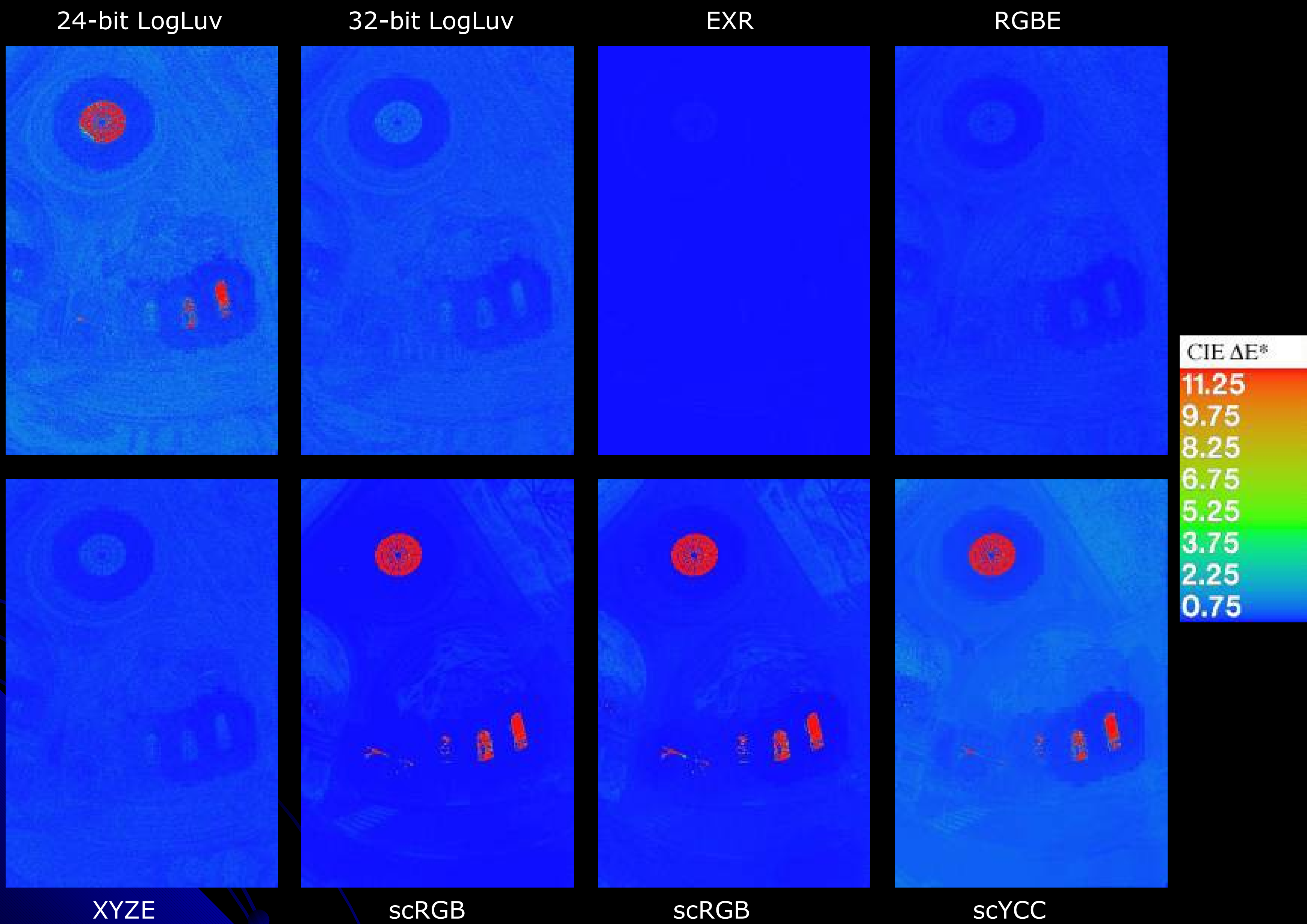
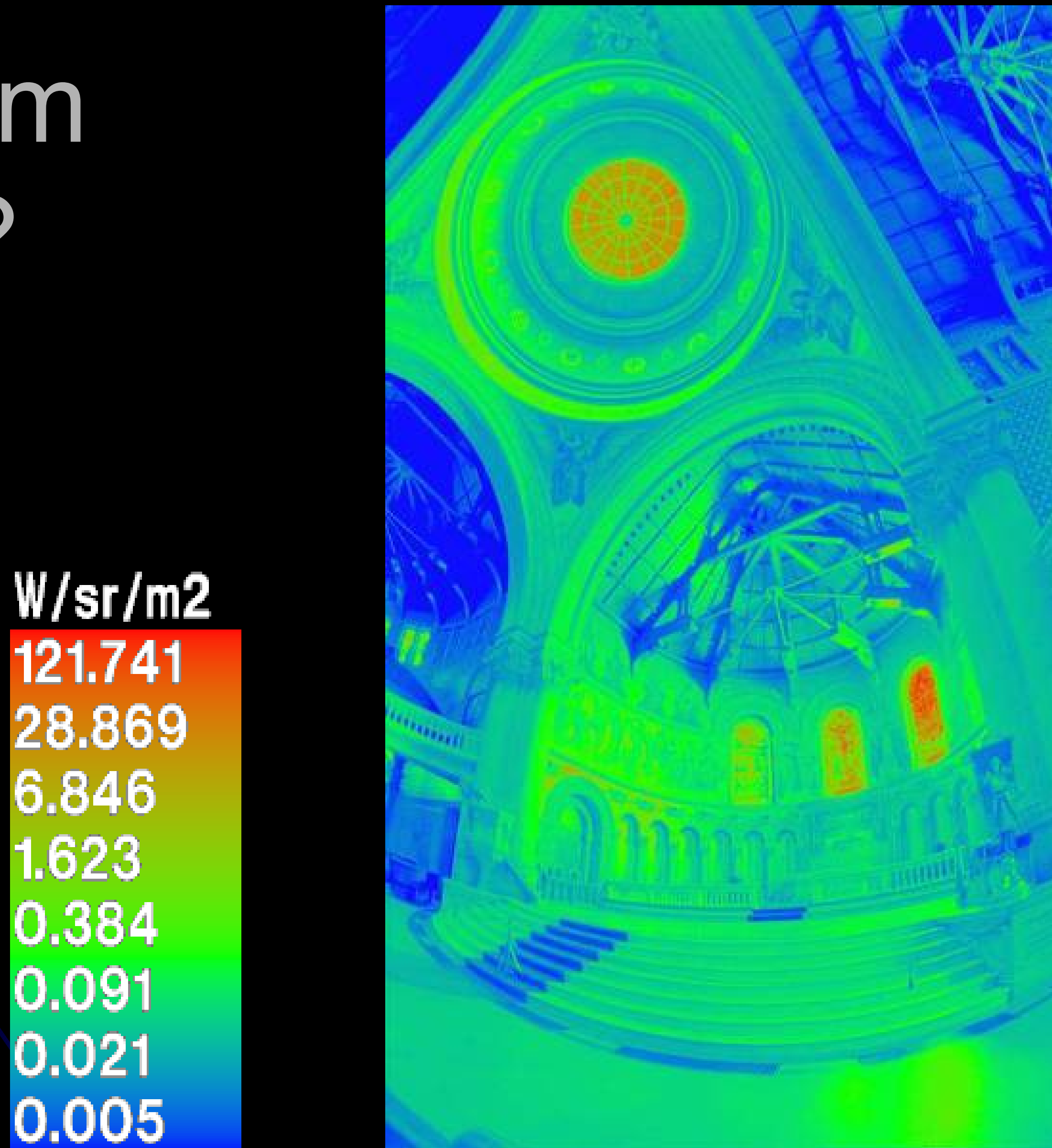

24-bit LogLuv
32-bit LogLuv
EXR
RGBE
CIE ΔE*
11.25
9.75
8.25
6.75
5.25
3.75
2.25
0.75
XYZE
scRGB
scRGB
scYCC

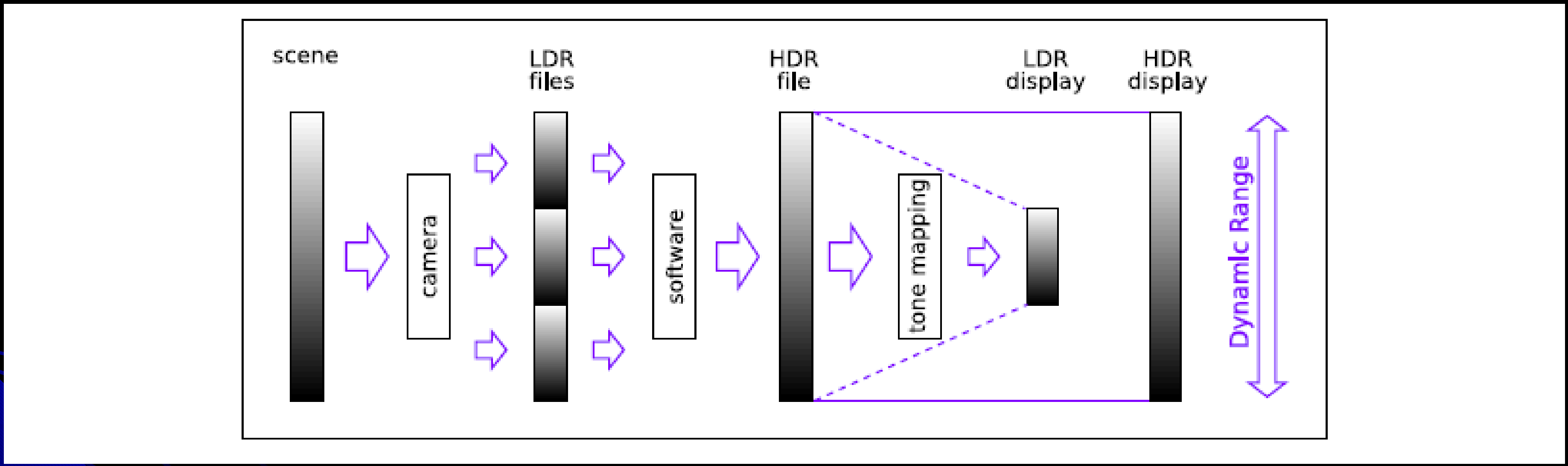
Figure 8: *A spherical HDR capture on top of Mount Vesuvius.*

Sony Pi
Spheron C
Productions

http://www.youtube.co
v=Luprwm9JOi4

Čo s tým ďalej?



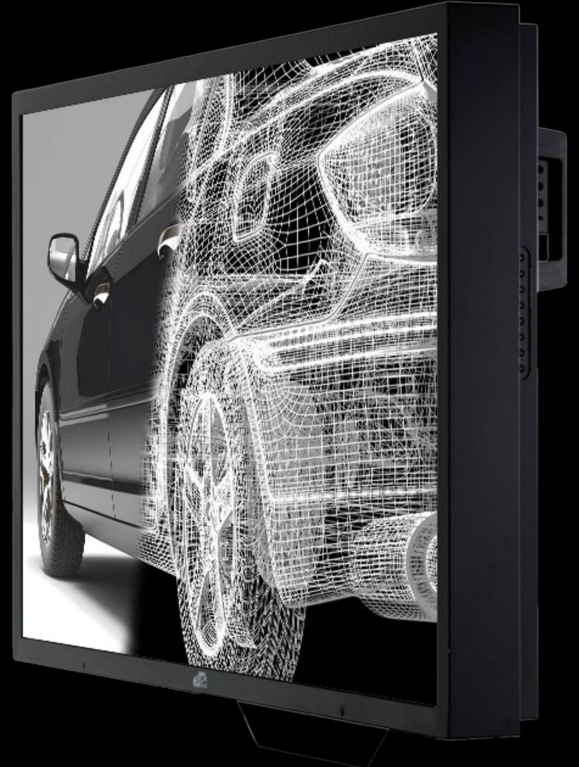

scene
camera
LDR
files
software
HDR
file
tone mapping
LDR
display
HDR
display
Dynamic Range

Brightside DR37-P teraz Dolby prototyp

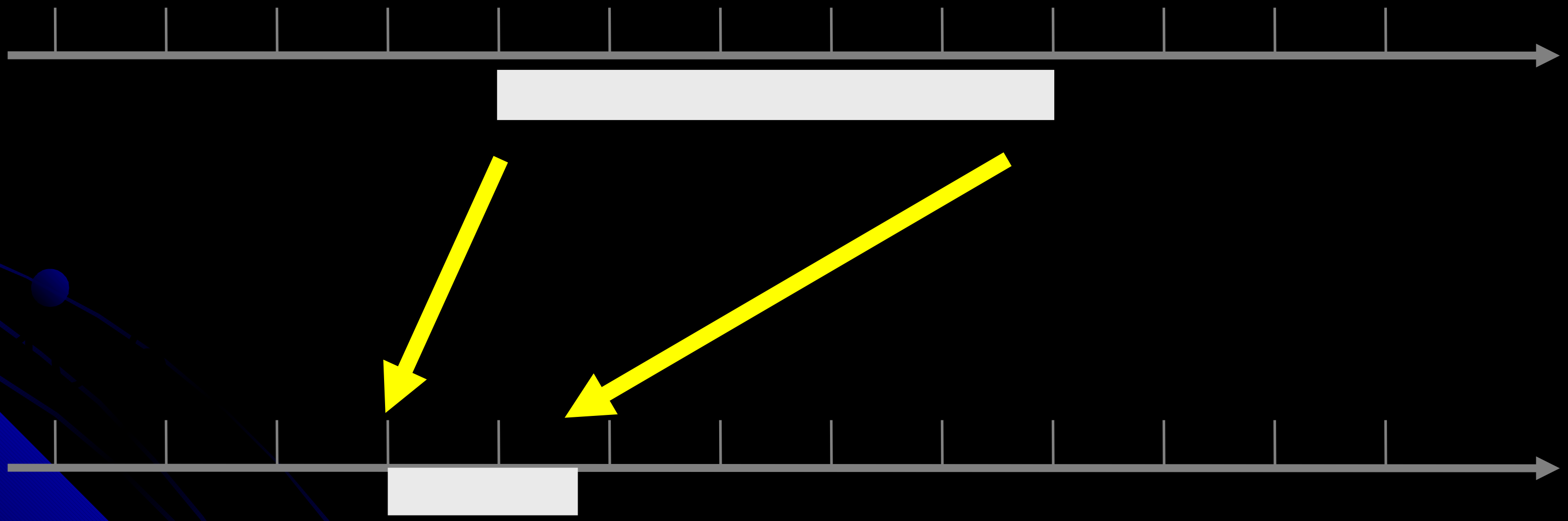
- Over 3,000 cd/m2 brightness
- 0.015 cd/m2 black level
- Contrast ratio > 200,000:1
- High-definition 1,920 x 1,080
- 37-inch screen
- 16 bits per color

SIM2 HDR47E S 4K

Technology	HDR LCD Display with individually controlled LED backlight modulation
Resolution	1920 x 1080 pixels
Display size	47”
Panel aspect ratio	16:9
Number of real colours	16 bit per channel
Number of LED	2202
Brightness	4000 cd/m2
ANSI Contrast	>20.000:1
FULL ON/OFF Contrast	virtually infinite (>1.000.000:1)



Tone Mapping



Tone Mapping

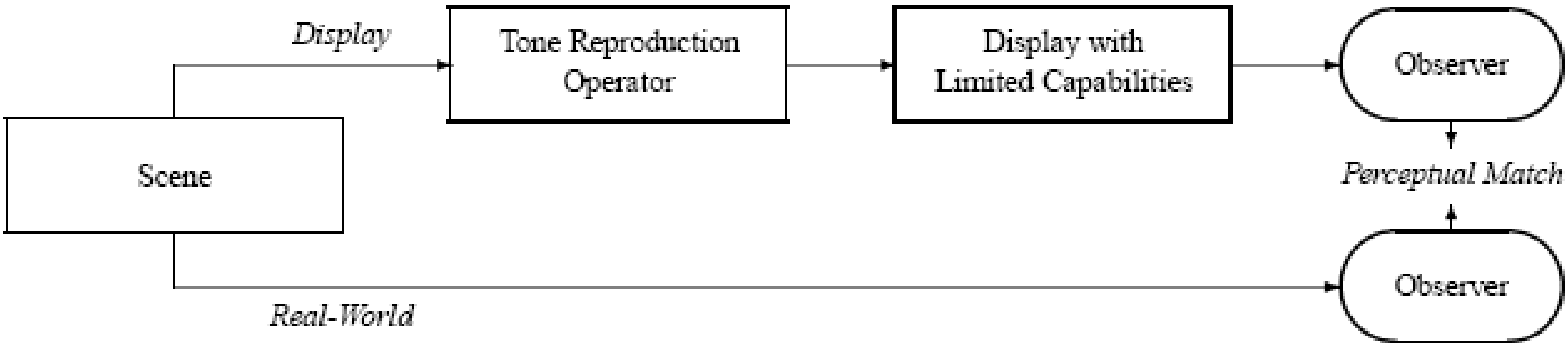


Figure 2: Ideal tone reproduction process

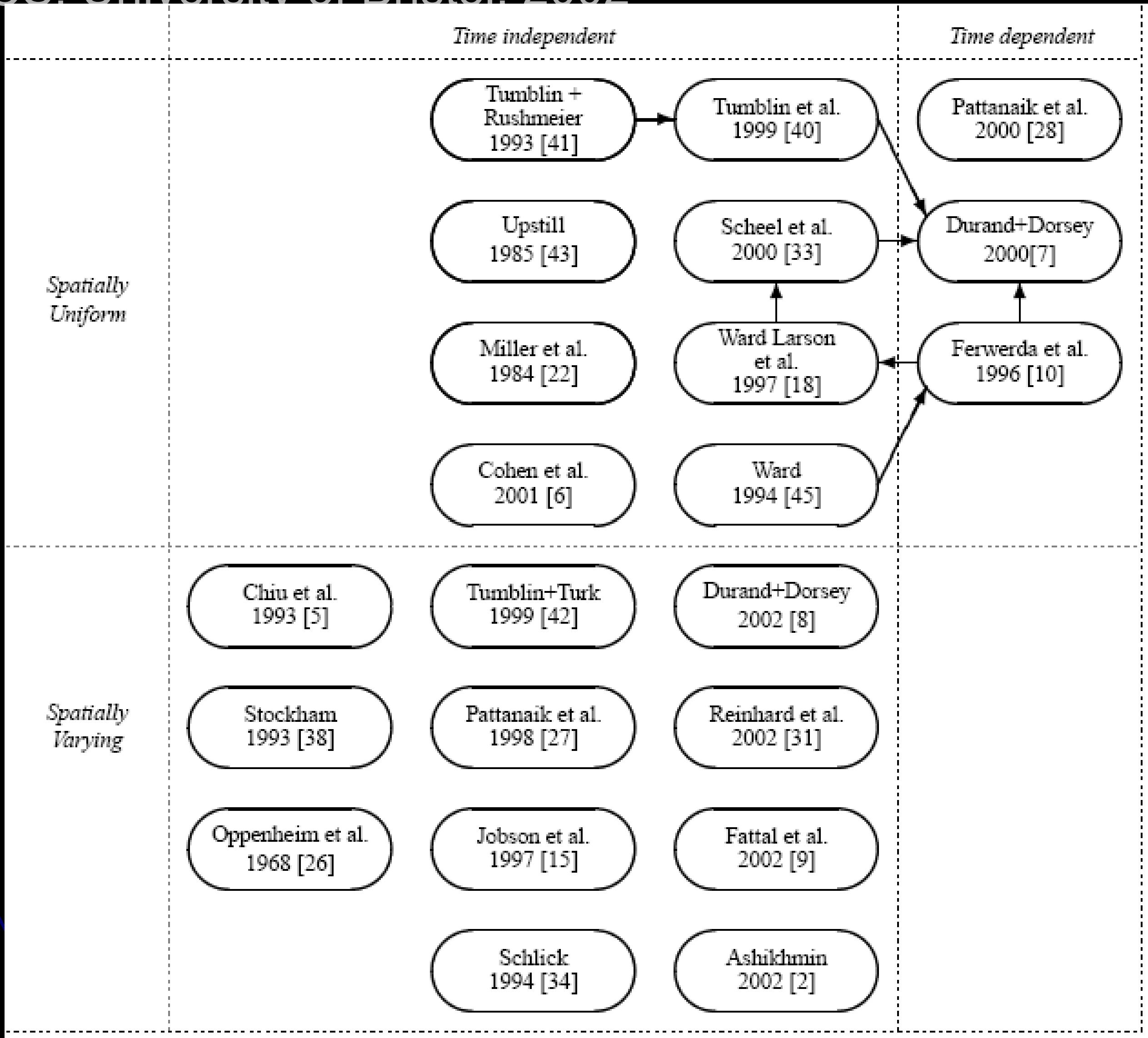
Typy operátorov

- Globálne
  - Rovnaká nelineárna krivka aplikovaná na všetky body obrazu
- Lokálne
  - Adaptačná krivka je rôzna pre každý bod podľa jeho okolia
- Frekvenčné
  - DR je menený podľa frekvencií obsiahnutých v obraze
- Gradientné
  - Mení sa derivácia obrazu

A review of tone reproduction techniques
Kate Devlin
DCS, University of Bristol, 2002

$$L_d(x, y) = mL_w(x, y)$$

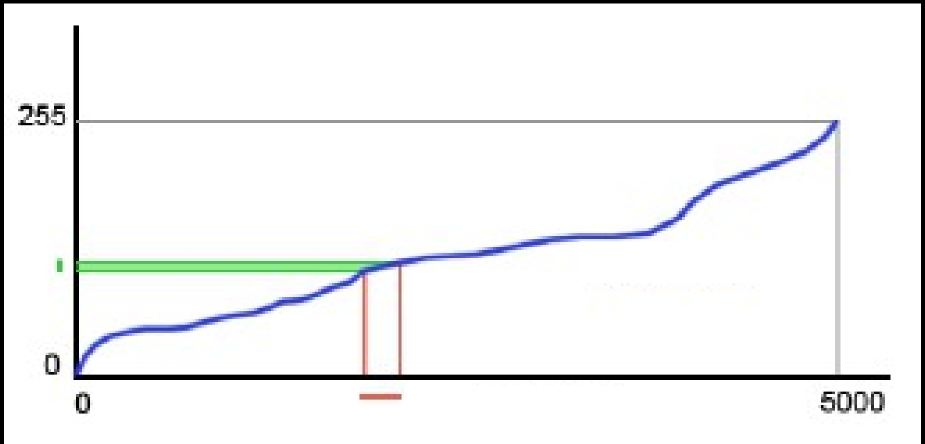
$$L_d(x, y) = m(x, y)L_w(x, y)$$

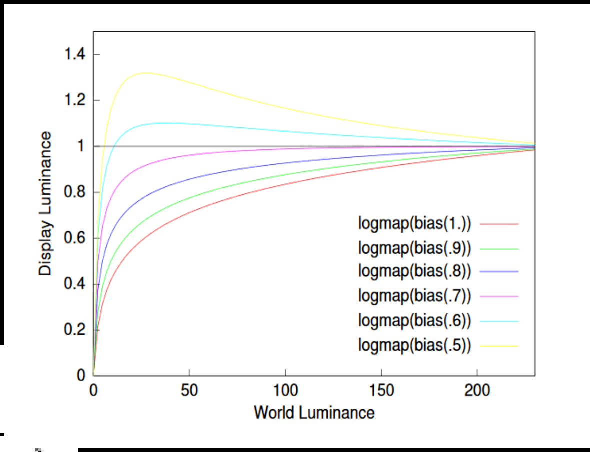


Globálny Operátor

A Contrast-based Scalefactor for
Luminance Display
Ward, 1994

$$I_D = \frac{I_W}{I_{d\,max}}\left[\frac{1,219 + (I_{d\,max}/2)^{0,4}}{1,219 + I_{wa}^{0,4}}\right]^{2,5}$$

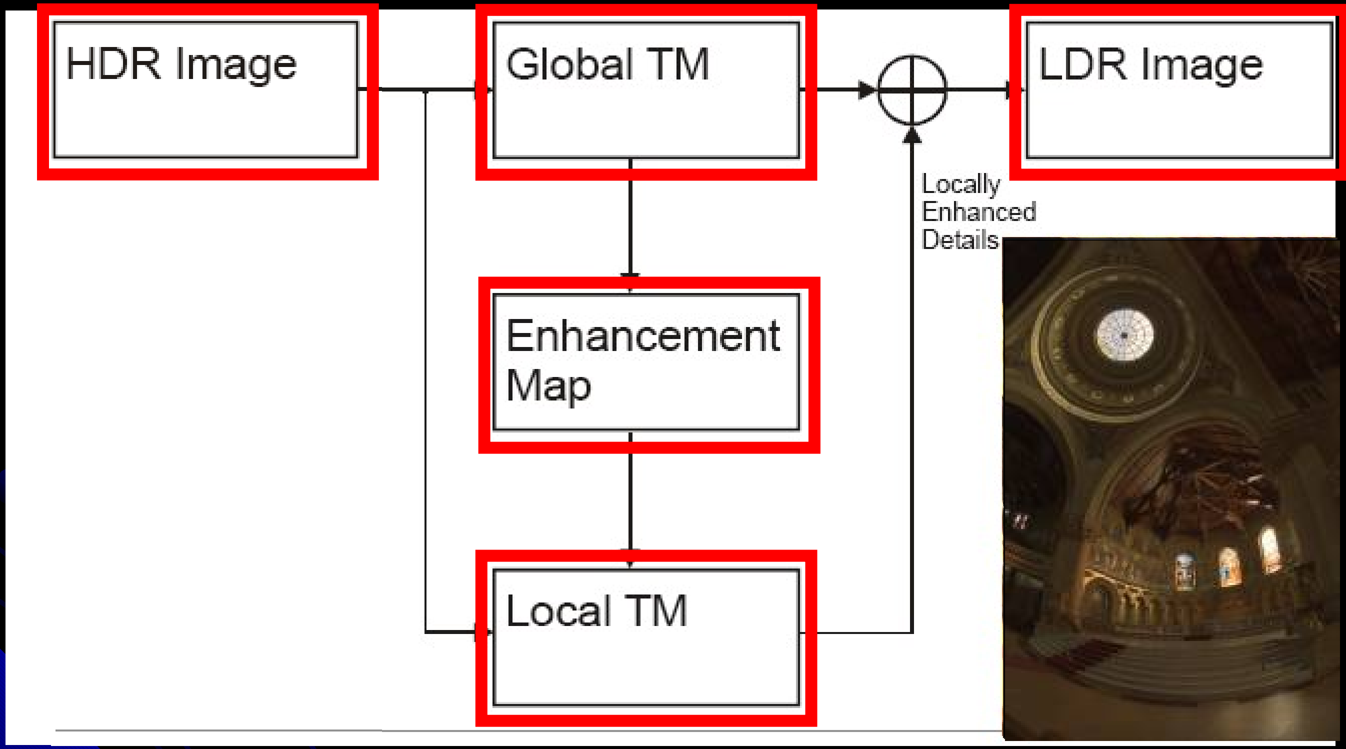


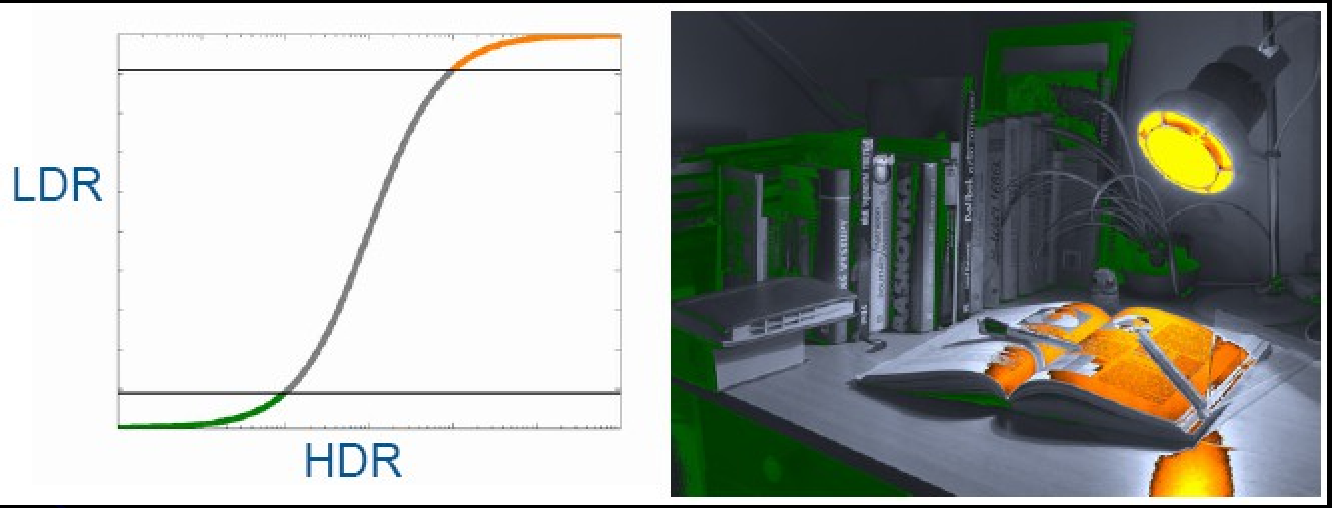


Adaptive Logaritmic Mapping For
Displaying High Contrast Scenes.
Drago, Myszkowski, Annen, Chiba, 2003

$$L_d = \frac{L_{dmax} \cdot 0,01}{\log_{10}(L_{wmax}+1)} \cdot \frac{\log(L_w + 1)}{\log\left(2 + \left(\left(\frac{L_w}{L_{wmax}}\right)^{\frac{\log(b)}{\log(0,5)}}\right) \cdot 8\right)}$$

Perception Motivated Hybrid Approach to Tone Mapping
Martin Čadík, 2007

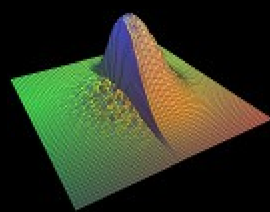




Enhancement map

Fast Bilateral Filtering for the Display of HDR Images
Frédo Durand & Julie Dorsey

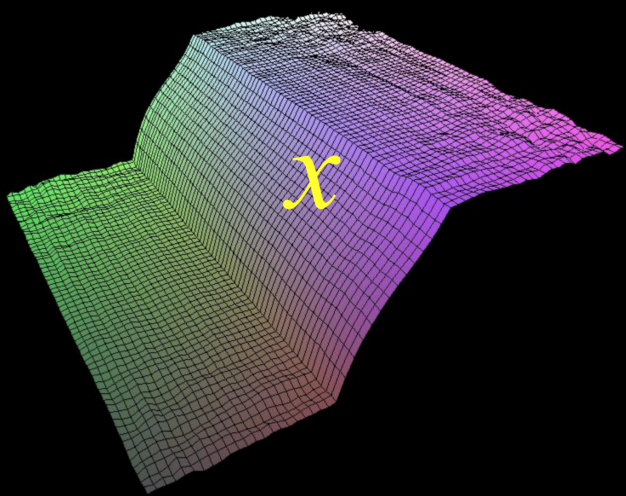
Fast Bilateral Filtering
[Tomasi a Manduchi 1998]

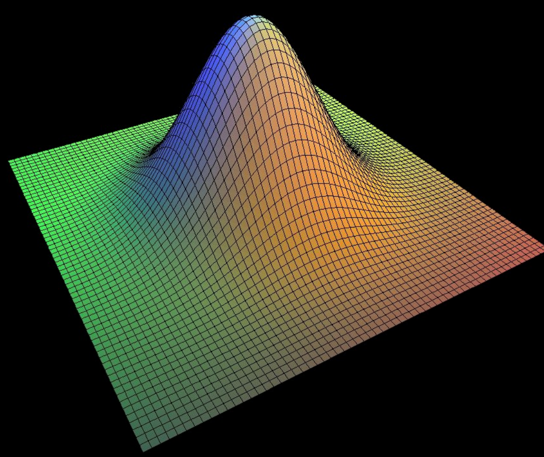


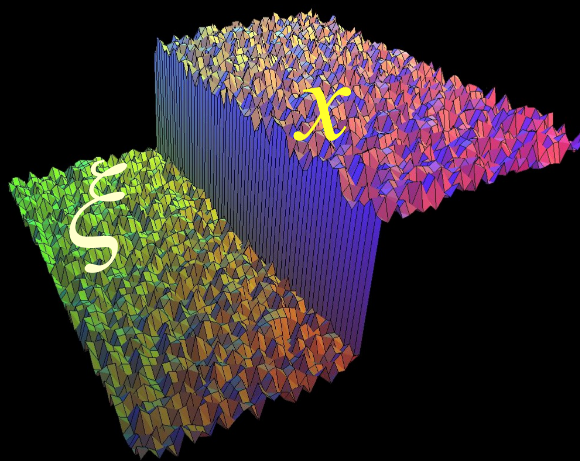
Detail = Intenzita – báza

Bilaterálny filter

$$J(x) = \sum_{\xi} f(x,\xi)\, I(\xi)$$

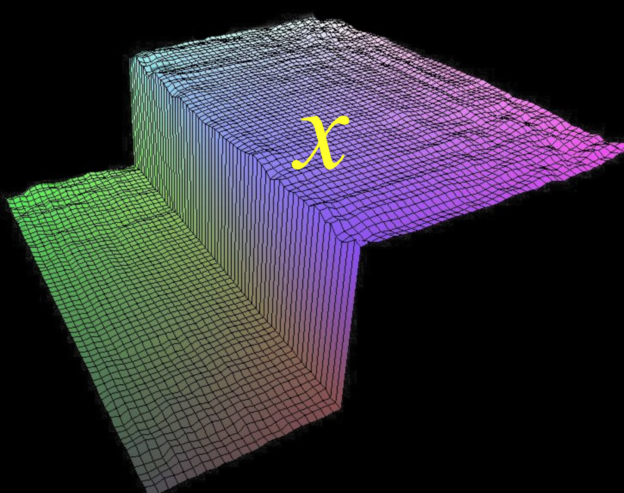


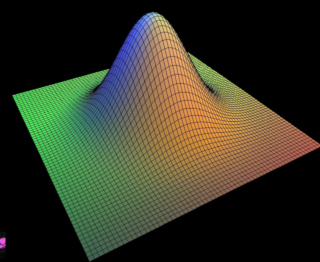


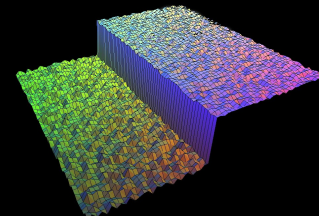


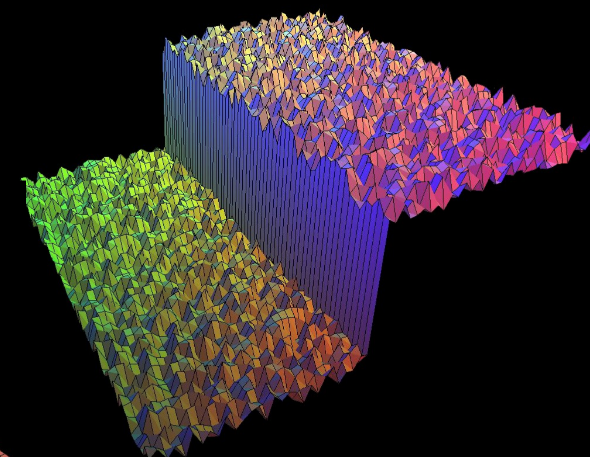
output

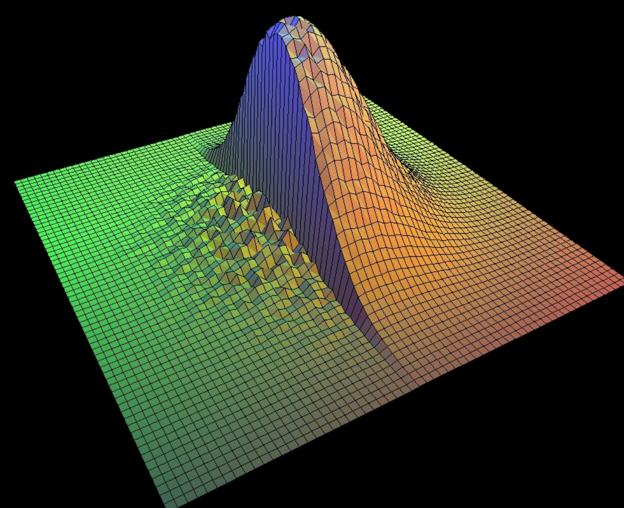
$$J(x)=\frac{1}{k(x)}\sum_{\xi} f(x,\xi) \quad g(I(\xi)-I(x)) \quad I(\xi)$$











output