五彩斑斓的计算摄影

汪彧之

BigEagle

你是谁

Contact Me

- wangyuzhi[at]megvii.com
- bigeagle
- bigeagle_xd
- bigeagle

I am a research engineer at Megvii Research. Previously I was a postdoctoral researcher at Dept. of CS, Tsinghua University. I received my Ph.D. in Electronic Engineering at NICS lab, Tsinghua University in 2017, supervised by Prof. Huazhong Yang and Prof. Yu Wang. Before that, I received my B.S. degree from Xidian University. Check my google scholar profile for my full publication list.

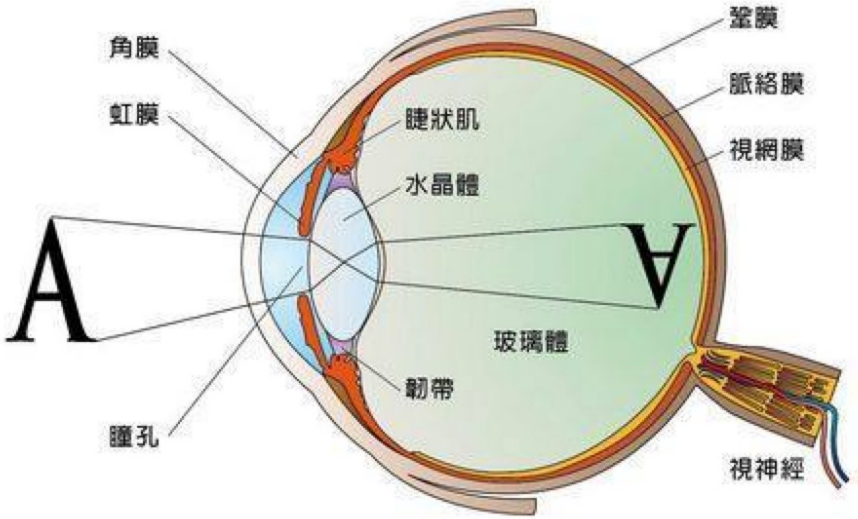
My current research interest includes Low-level Computer Vision, Image Signal Processing, Computational Photography and Software-Hardware Co-Design in these applications. Our solutions on computational photography has been applied to various smartphones since 2019.

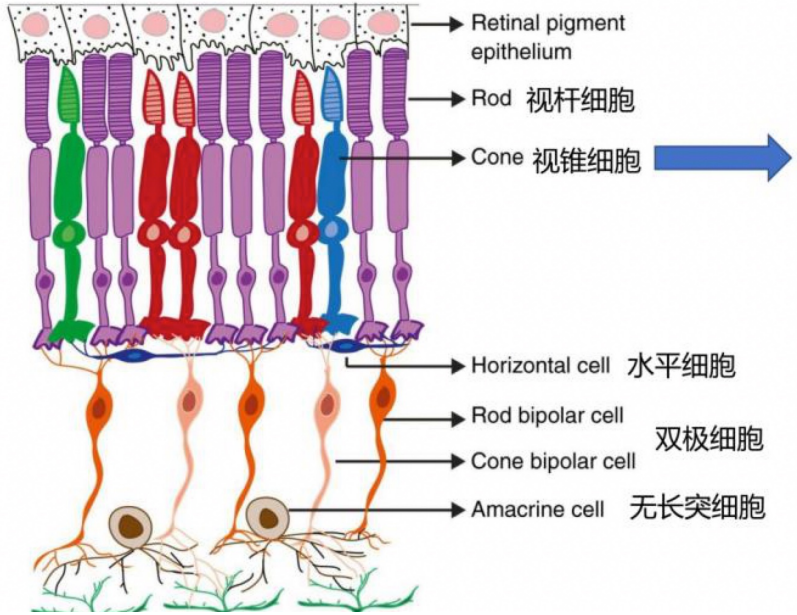
I am also a supporter of Free/Libre Software. I was an experienced user of Arch Linux since 2009, before I betrayed my beliefs and switched to MacBook in 2021. I developed many open source projects when I was a student member of TUNA and a maintainer of TUNA mirror site. Check my Github profile for more details of my contributions to the open source community.

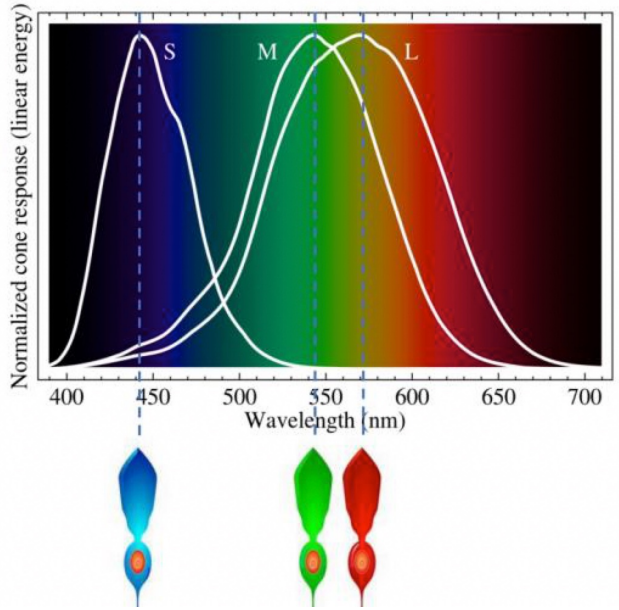
Why

- “你在旷视做什么?”
  - ~~研究非凡科技，为客户和社会创造最大价值~~
  - 亮一点的同时再暗一点
  - 平滑一点的同时，细节再多一点
  - 鲜艳一点的同时，再暗淡一点
- 计算摄影
  - 用算法还原人们心中的美好世界
  - 技术与艺术的结合、科学与玄学的平衡

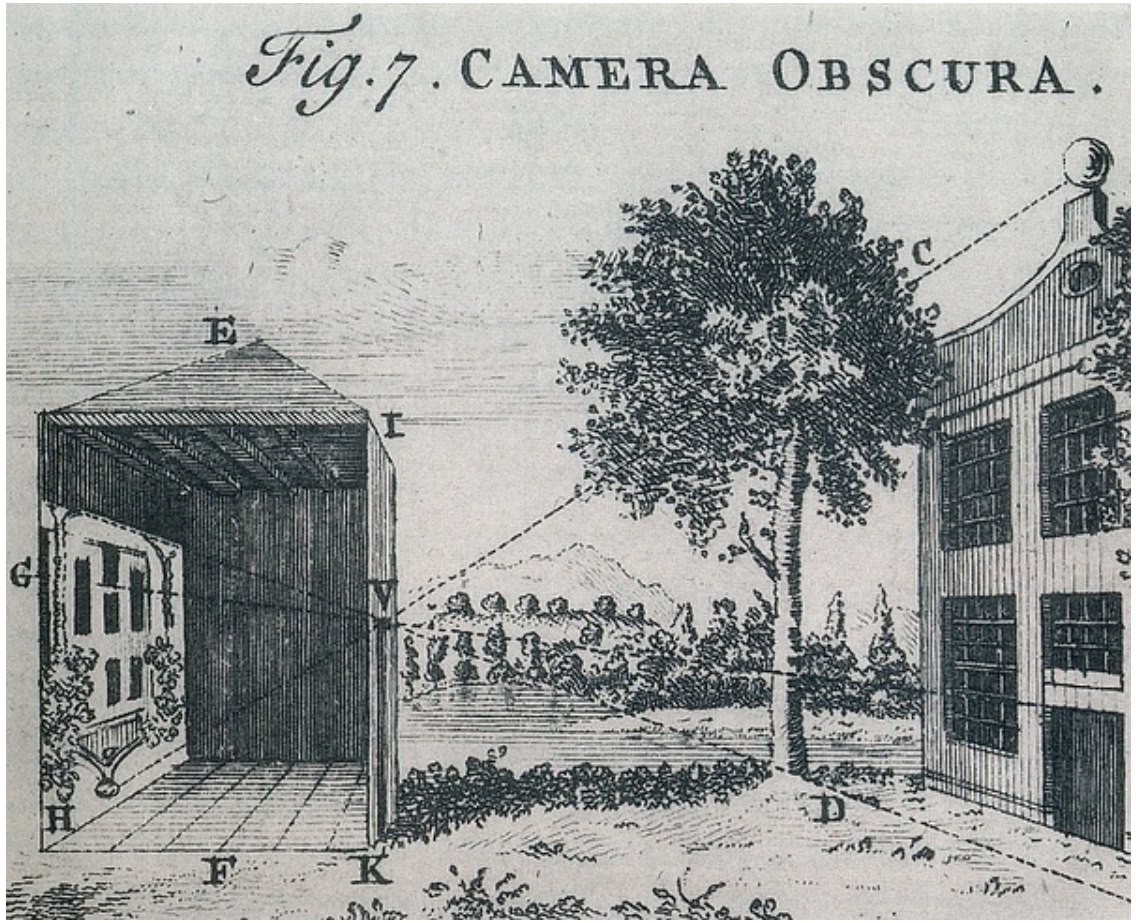
生物摄影

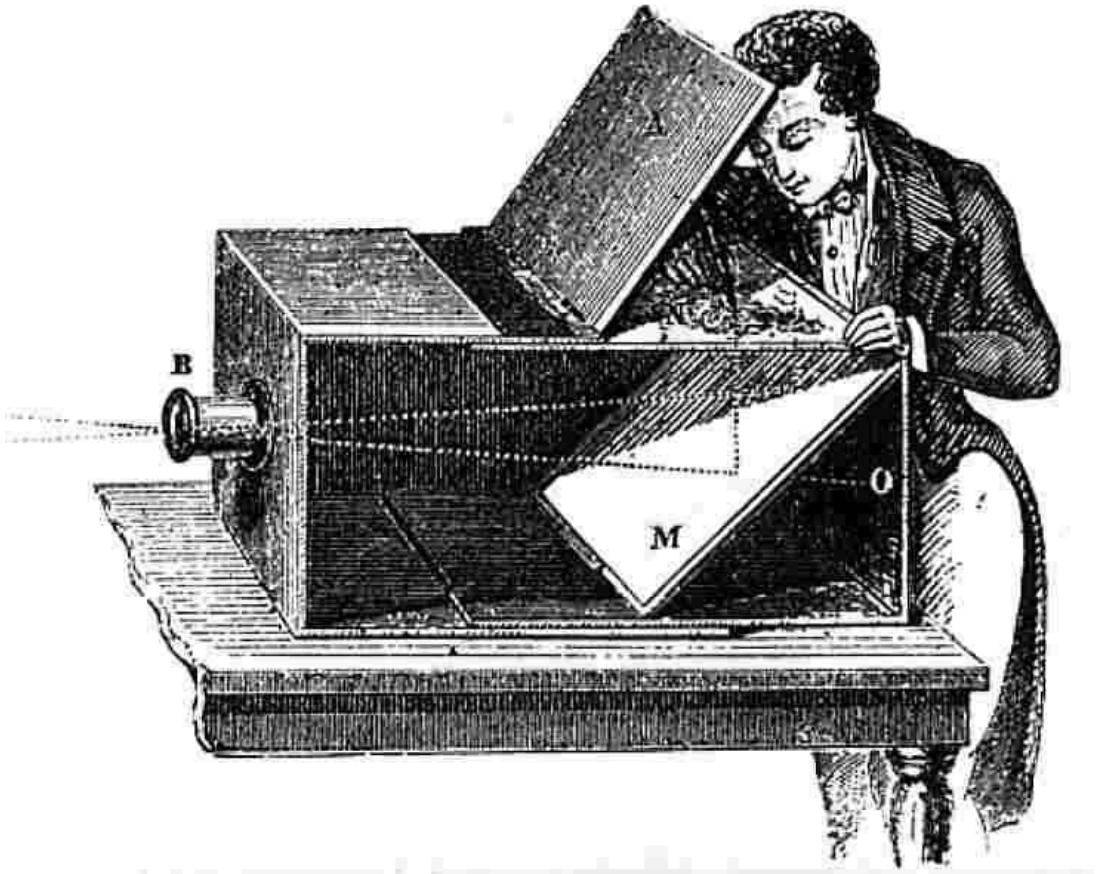

角膜
虹膜
瞳孔
鞏膜
脈絡膜
視網膜
睫狀肌
水晶體
玻璃體
韌帶
視神經


Retinal pigment epithelium
Rod 视杆细胞
Cone 视锥细胞
Horizontal cell 水平细胞
Rod bipolar cell
Cone bipolar cell
双极细胞
Amacrine cell 无长突细胞


S
M
L
Normalized cone response (linear energy)
400
450
500
550
600
650
700
Wavelength (nm)

人工摄影

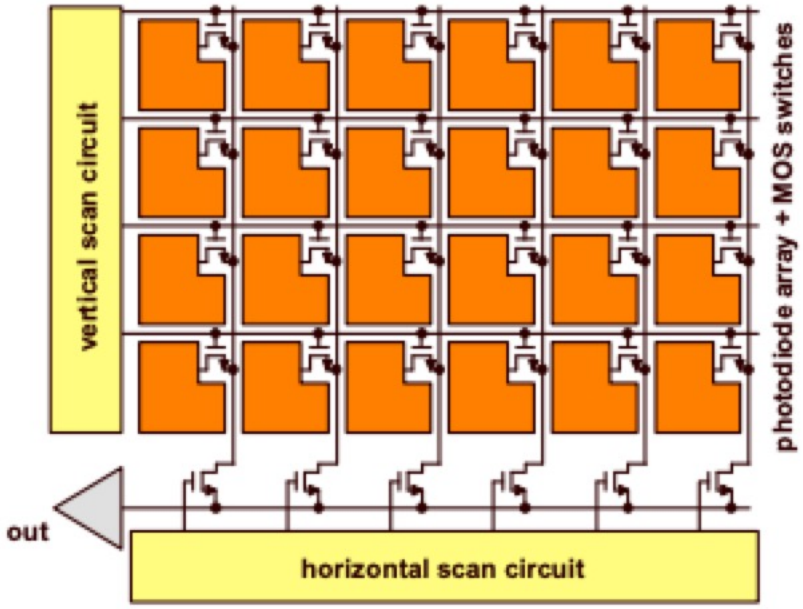


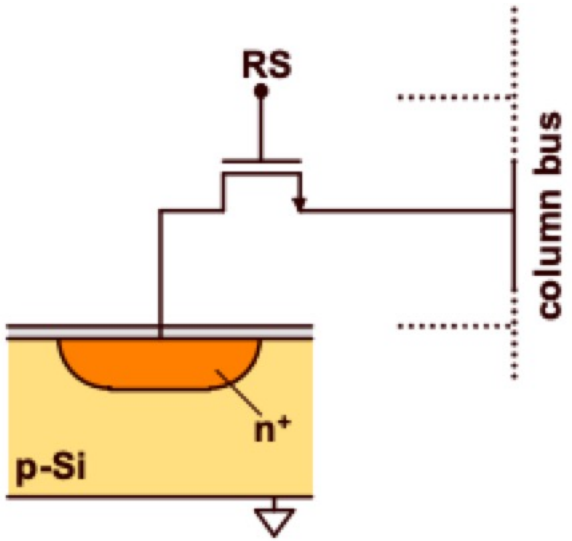


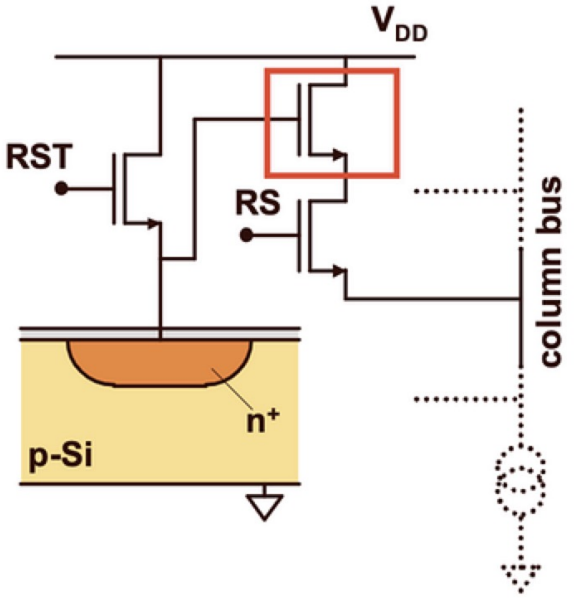
化学摄影

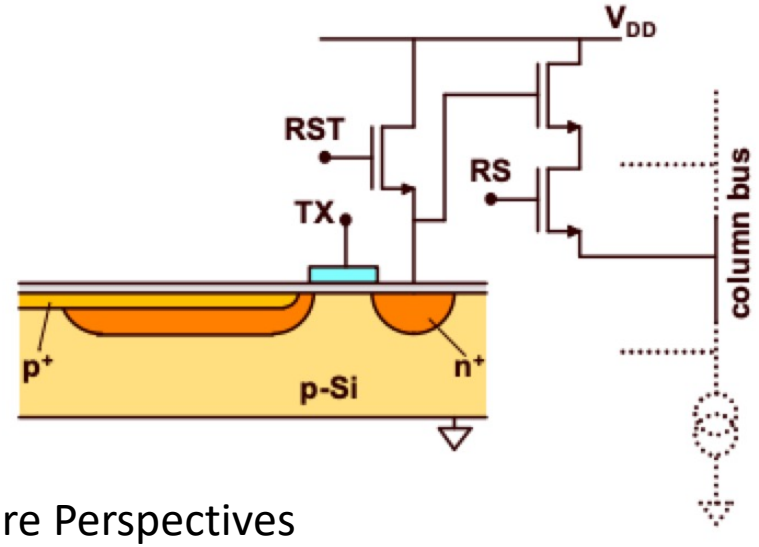
$$2AgBr = 2Ag + Br_2$$

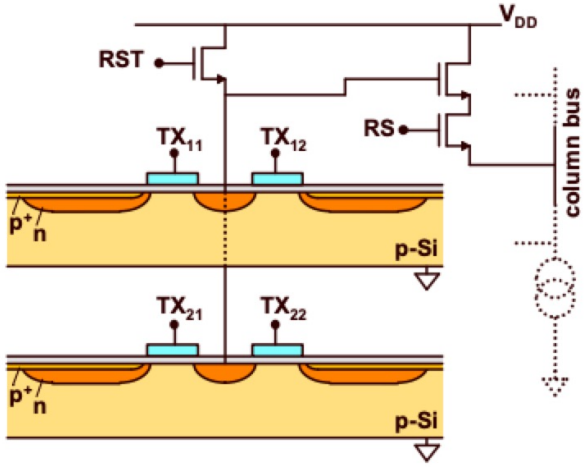
数字摄影











CMOS Image Sensors : State-Of-The-Art and Future Perspectives

4T Active Pixel Sensor

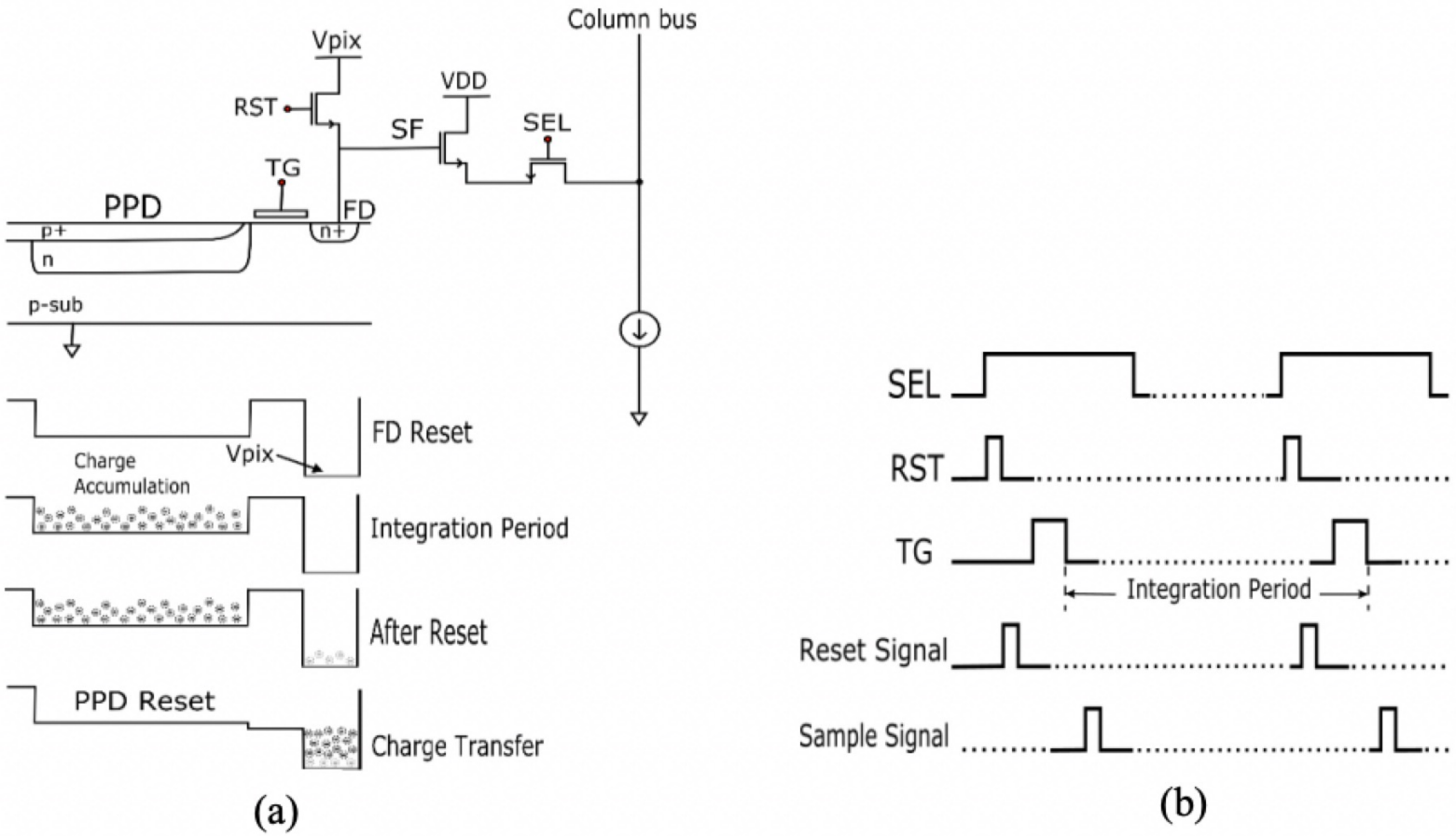
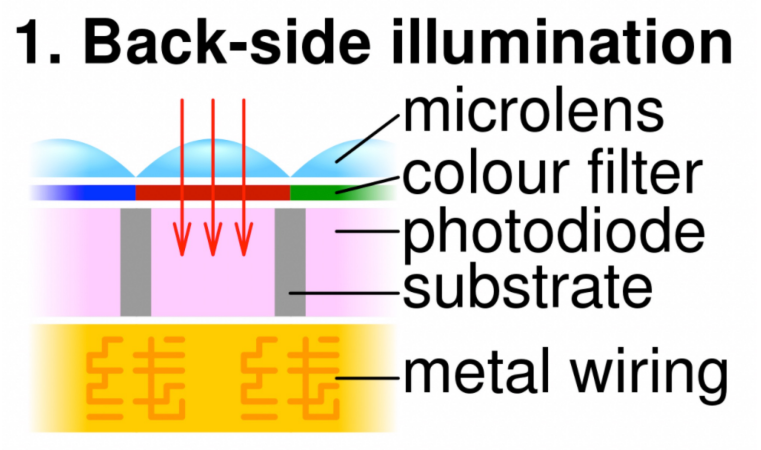
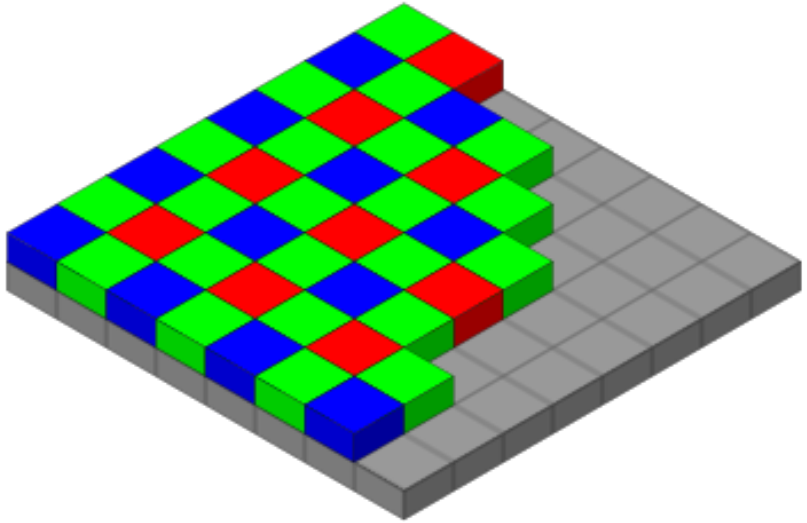
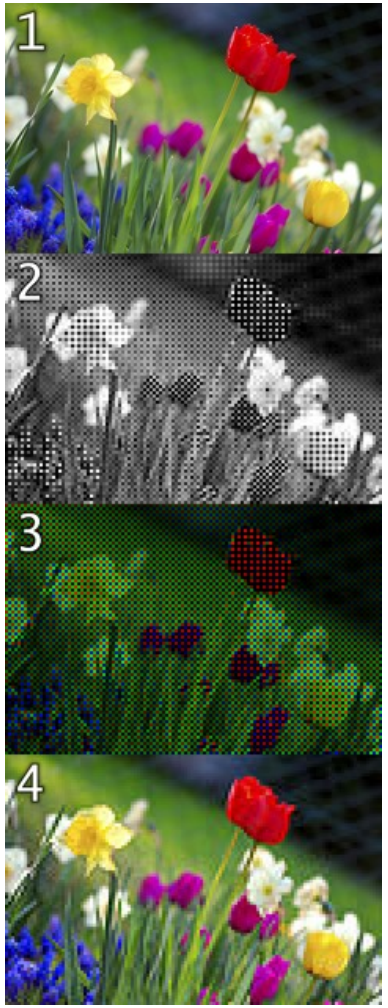


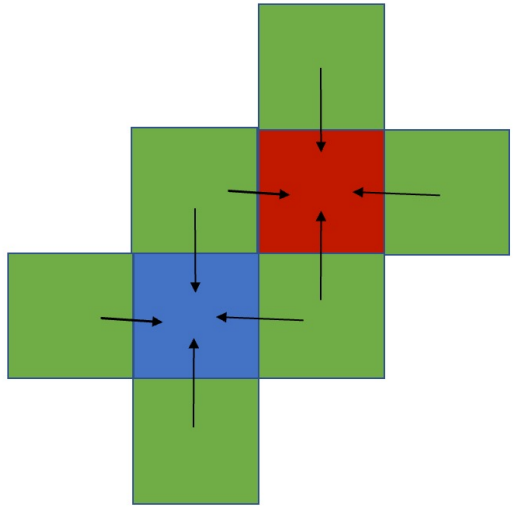
Figure 2.3-2- (a) Operation of 4T APS, (b) Timing diagram of signals.

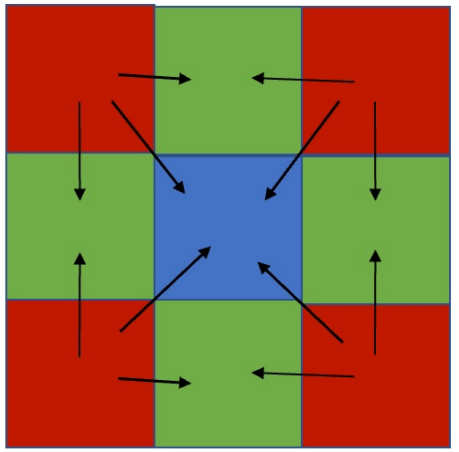
计算摄影


1. Back-side illumination
microlens
colour filter
photodiode
substrate
metal wiring




1
2
3
4





计算摄影

- 在成像过程中使用各种算法，恢复人们心中的世界
- 给硬件修锅

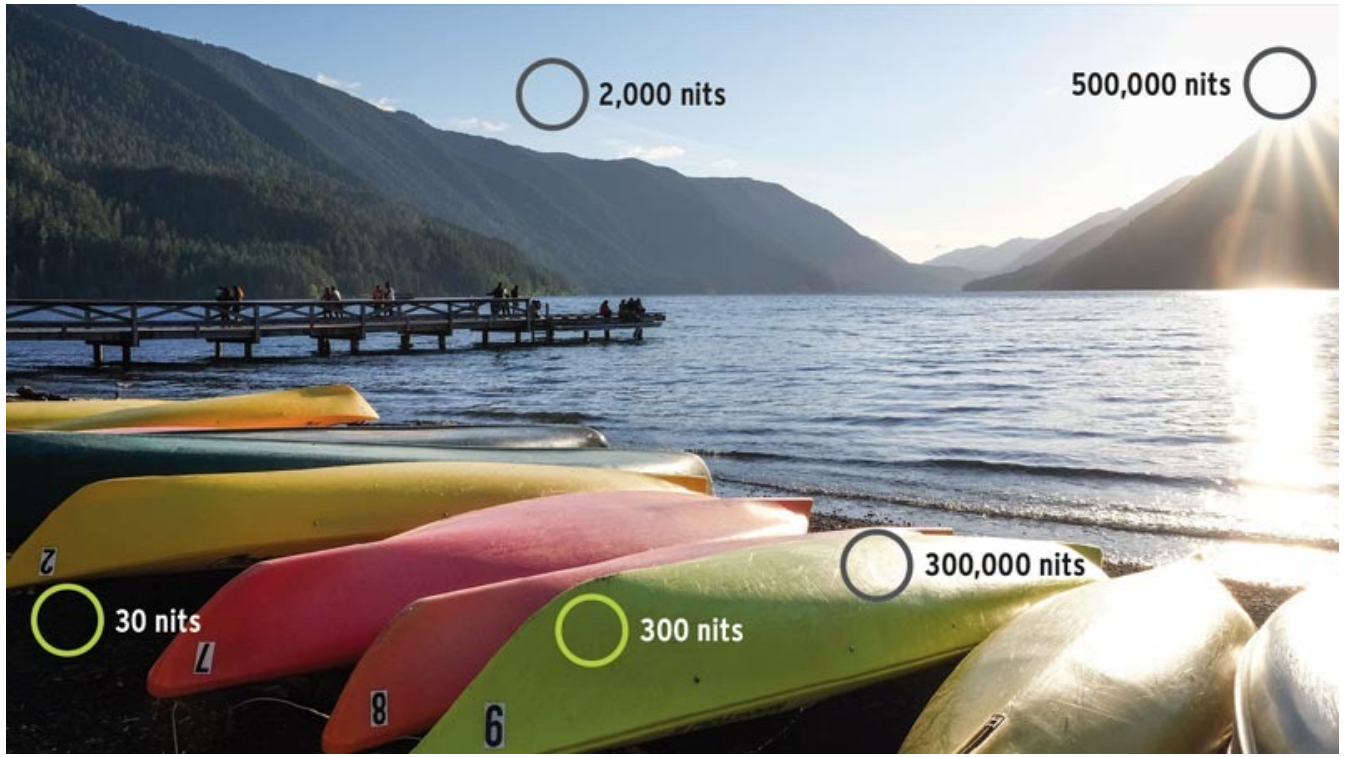
- 亮度与动态范围
- 解析力
- 噪声
- 色彩
- 空间感
- ……

什么是“动态范围”？

能够测量到的最大值与最小值之比

高动态范围：在亮一点的同时暗一点

大自然的动态范围



Illuminance (lux)	Surfaces illuminated by
0.0001	Moonless, overcast night sky (starlight)[4]
0.002	Moonless clear night sky with airglow[4]
0.05–0.3	Full moon on a clear night[5]
3.4	Dark limit of civil twilight under a clear sky[6]
20–50	Public areas with dark surroundings[7]
50	Family living room lights (Australia, 1998)[8]
80	Office building hallway/toilet lighting[9][10]
100	Very dark overcast day[4]
150	Train station platforms[11]
320–500	Office lighting[8][12][13][14]
400	Sunrise or sunset on a clear day.
1000	Overcast day;[4] typical TV studio lighting
10,000–25,000	Full daylight (not direct sun)[4]
32,000–100,000	Direct sunlight

如何估计一个场景的亮度(lux)

Sunny 16 rule & Looney 11 rule

1/100"　ISO 100

满月的“亮度” = 地面大白天的亮度

进光量正比于光圈的平方

Aperture	Lighting conditions	Shadow detail
f/22	Snow/sand	Dark with sharp edges
f/16	Sunny	Distinct
f/11	Slight overcast	Soft around edges
f/8	Overcast	Barely visible
f/5.6	Heavy overcast	No shadows
f/4	Open shade/sunset	No shadows
Add one stop	Backlighting	n/a

1000~2000lux

日常差距1000倍是很正常的

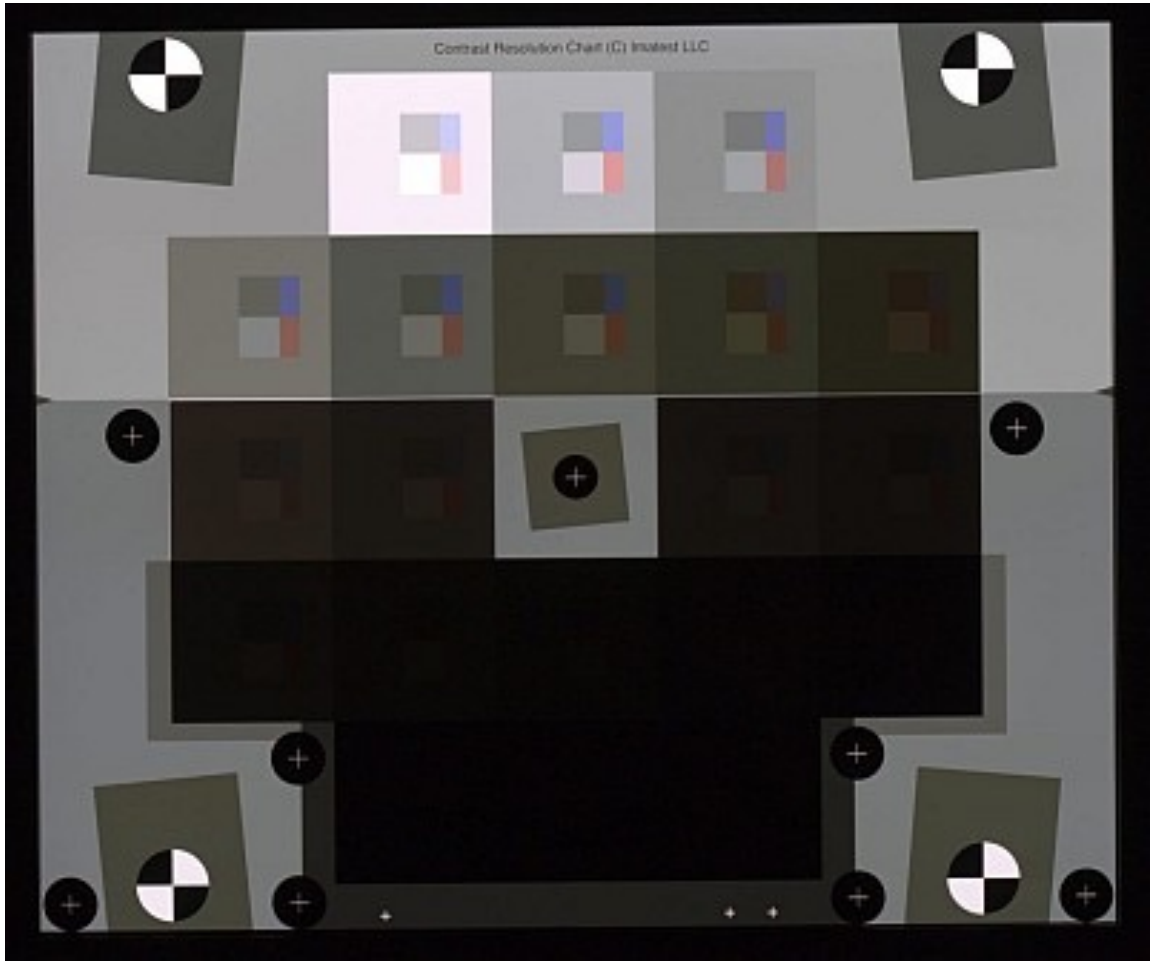
iso 50 1/6655s

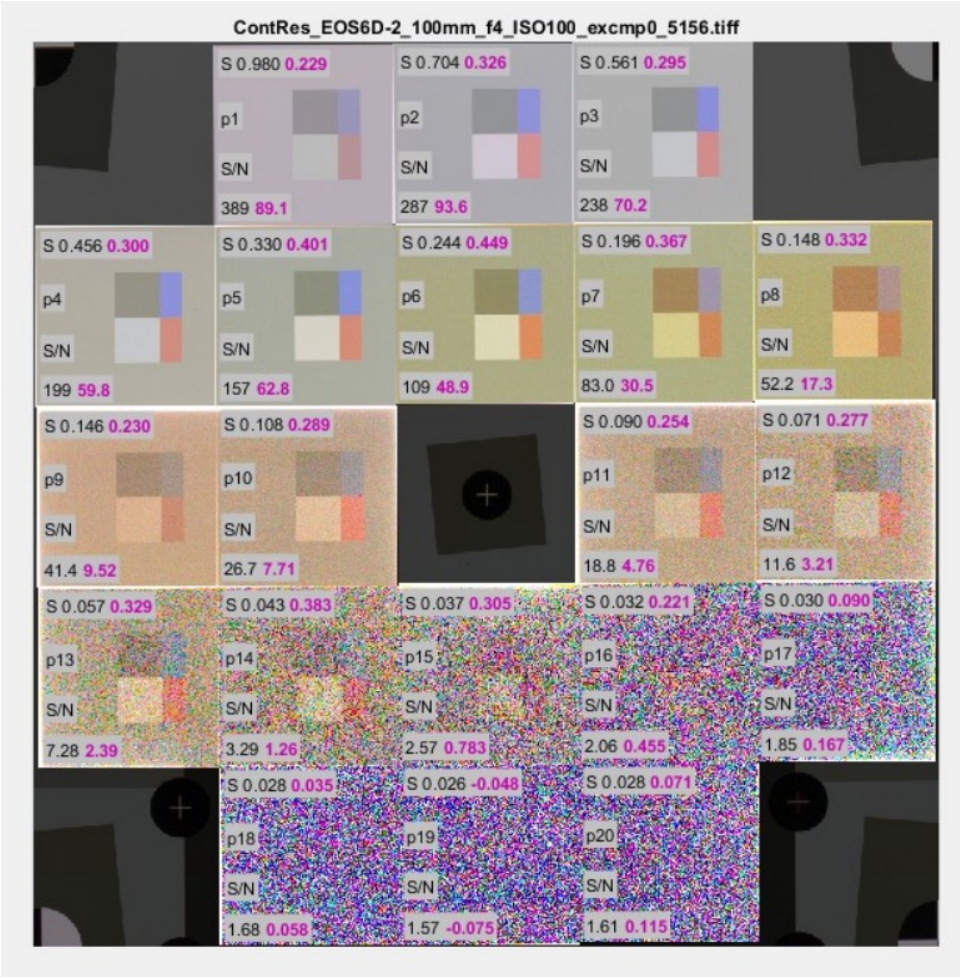
iso 400 1/50s

Sensor的动态范围

- 摄影中一般用 EV 数衡量“曝光量”
  - 进光量每增加一倍，EV 数 +1
- 10bit sensor -> 不会超过10 EV
- 有一些 sensor 用dB来表示。10EV = 60 dB
  - $20 * \log_{10}(2) = 6.02$

动态范围的测量

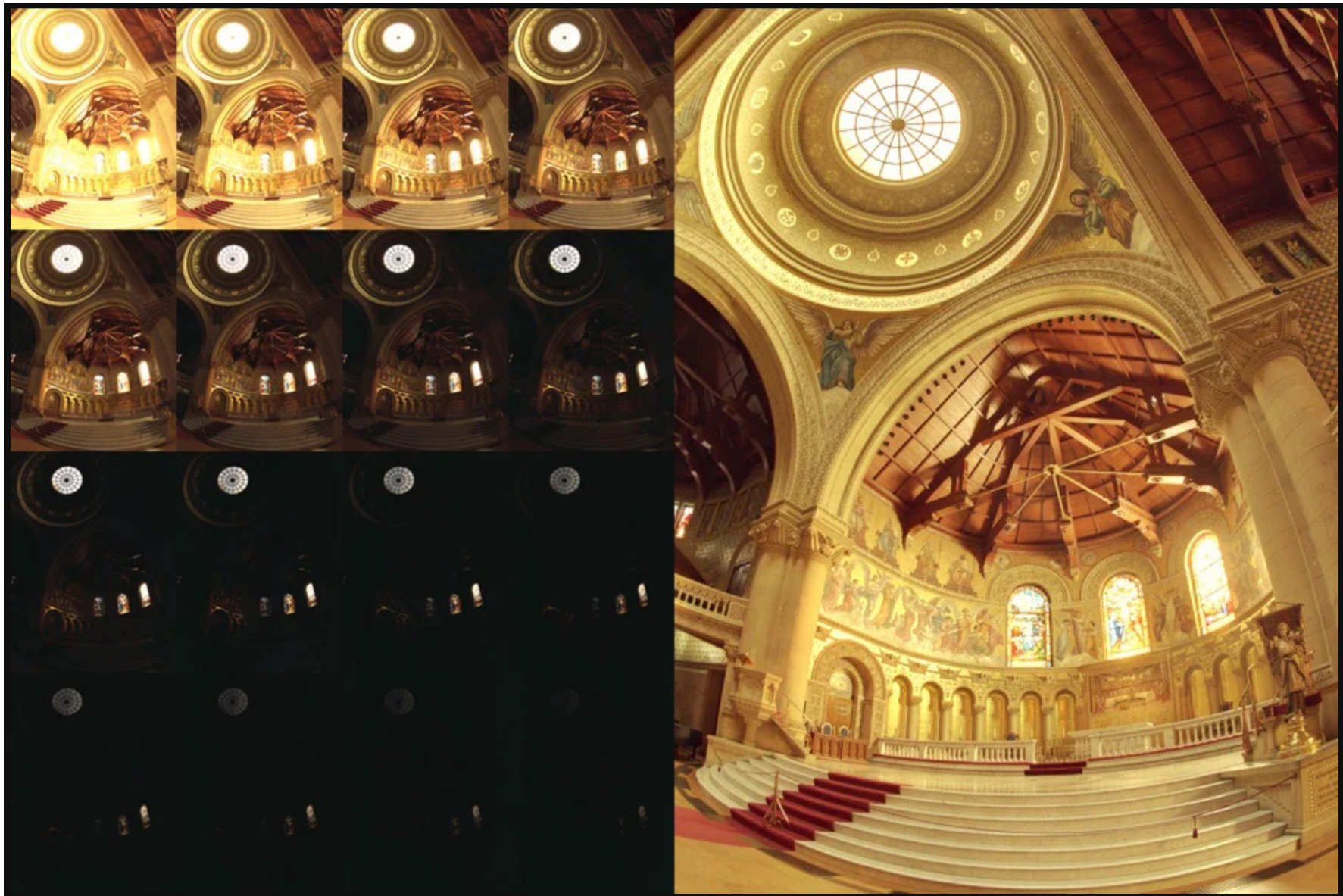




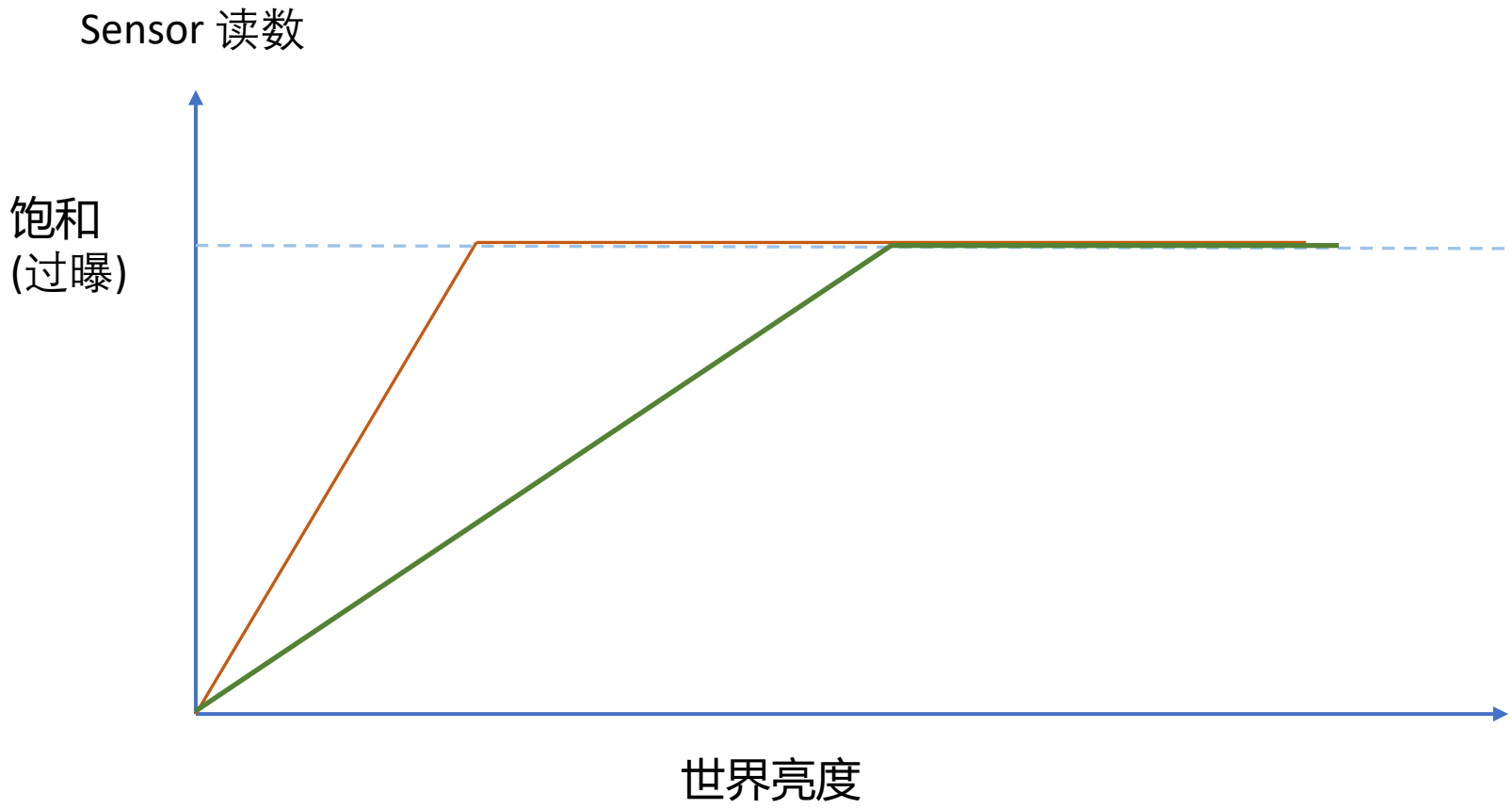
如何提高动态范围

软件方法

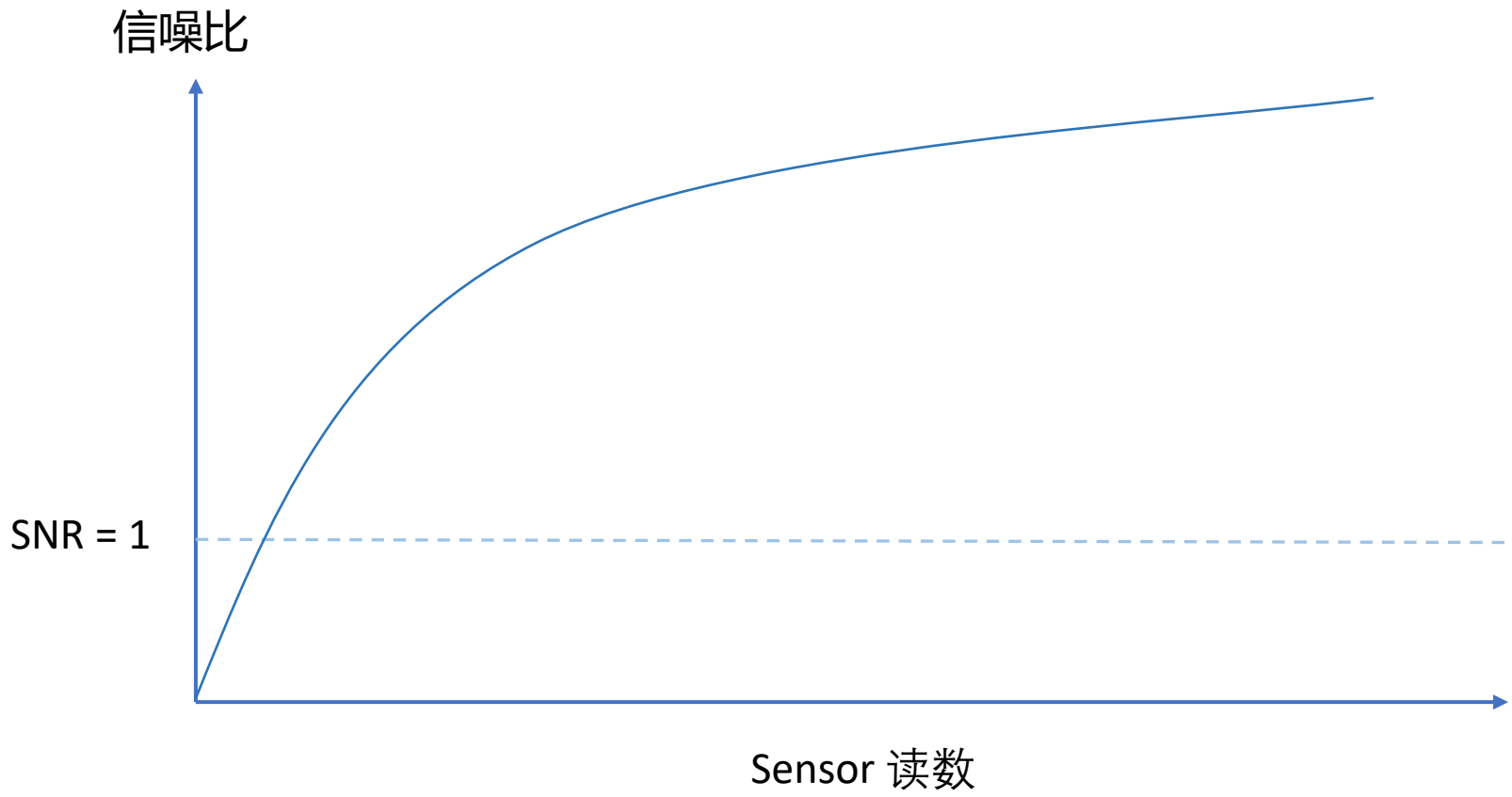
- 多曝光合成
- “包围曝光”



曝光对动态范围的影响



曝光对动态范围的影响



如何选帧

- 对于同一片区域，在某一帧，

- 如果值过大：整体或局部过曝，导致色偏/细节丢失
- 如果值过小：信噪比过低，细节模糊

- 所以我们想要一个又不大又不小的帧

经典的融合算法

Exposure Fusion

Tom Mertens[1] Jan Kautz[2] Frank Van Reeth[1]

[1]Hasselt University — EDM
transationale Universiteit Limburg
Belgium

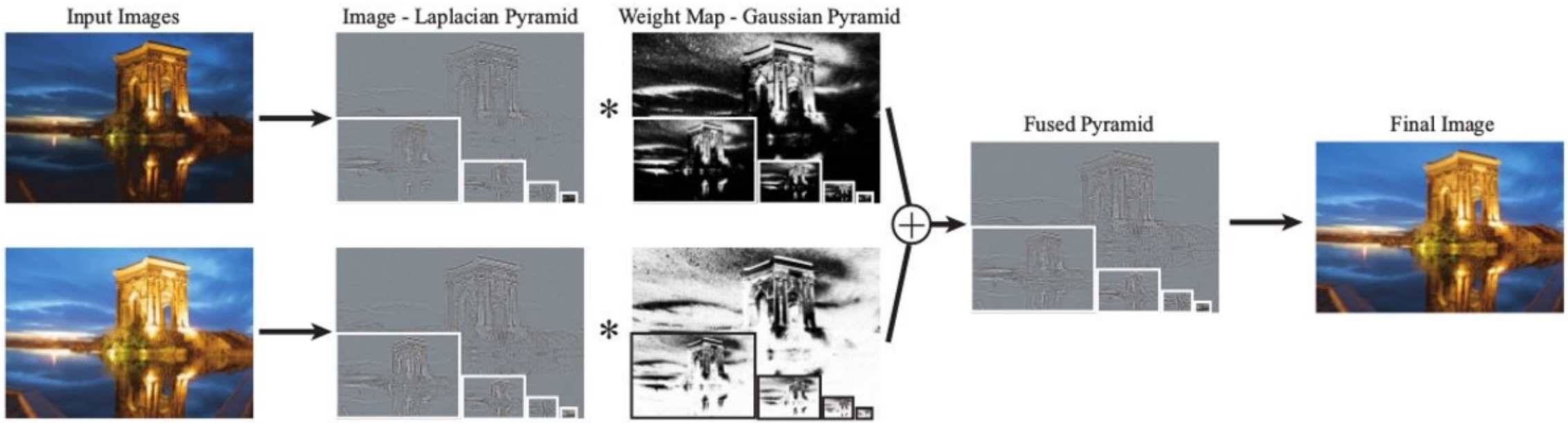
[2]University College London
UK

- Exposure fusion

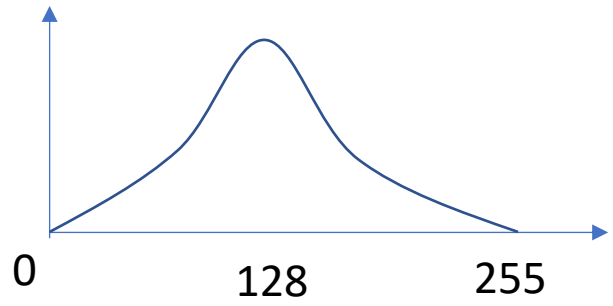
(a) Input images with corresponding weight maps

(b) Fused result

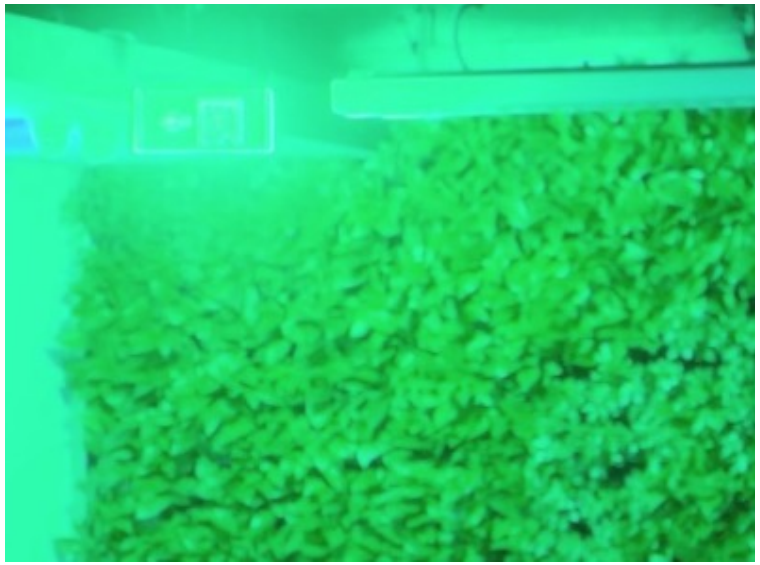
Figure 2. Exposure fusion is guided by weight maps for each input image. A high weight means that a pixel should appear in the final image. These weights reflect desired image qualities, such as high contrast and saturation. Image courtesy of Jacques Joffre.


Input Images
Image - Laplacian Pyramid
Weight Map - Gaussian Pyramid
Fused Pyramid
Final Image

最简易的HDR融合



- 对于一组图像 $X_0, X_1, \ldots, X_N$
- 计算其亮度 L = max(R, G, B)
- 计算其权重 $w_i = K(Li)$, K 为以128为中心点的 Gaussian
- 融合 $I = sum(w_i, X_i)$




野心
敬畏心
同理心


野心
敬畏心
同理心

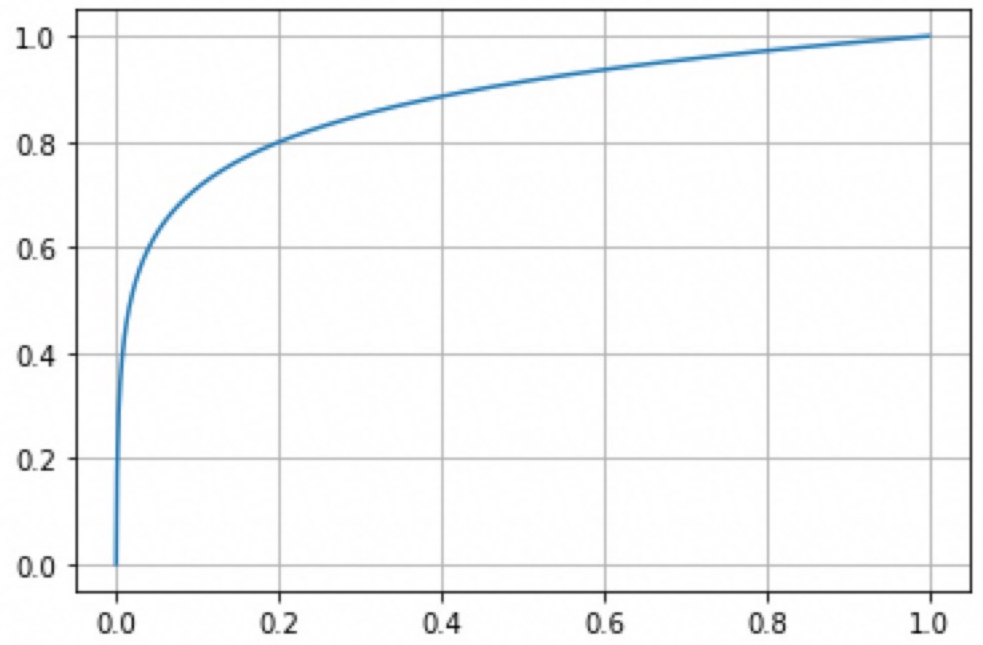

不要和我说什么色相色调明暗饱和度高级灰
老夫要的产品
就是
色彩要丰富
对比要强烈
最重要的是
字要大
不大怎么突出？
我要产品名称占80%
椰树椰汁听说过没？那个就很好嘛
这种也能算渐变？这才叫渐变
又突出又好看！
格我要简约的同时复杂一点
LOGO大一点的同时再缩小一点
微软雅黑怎么不是黑的？
那个黑色你做成五彩斑斓
我要那种世界500强的感觉
太方了
你把它改圆一点
少点感觉你再改一改
再大一点！
算了
还是用初稿吧
你到底会不会啊？还设计师呢

HDR 重建 v.s. HDR 渲染 v.s. HDR 显示

- HDR 重建：获得“真实”的高动态范围图像
  - Ev0 12bit + ev-2 12bit + ev-4 12bit -> 16bit
- HDR 渲染：用低动态范围的图像表示
  - 16bit -> 8bit
- HDR 显示：让显示器可以显示高动态范围的图像
  - 10bit -> 0.0001 ~ 10000 nit

HDR Tone-Mapping

- 假设已经获取了一份高动态范围、线性亮度的图像
- 如何将高动态范围的信息映射到低位宽图像



- 拉个 Curve 怎么样？


MUTATIONS
TINTIN

HDR Tone-Mapping

- 在保持“局部对比度”的同时降低“全局对比度”

LF = GaussianBlur(I, sigma)

HF = I – LF

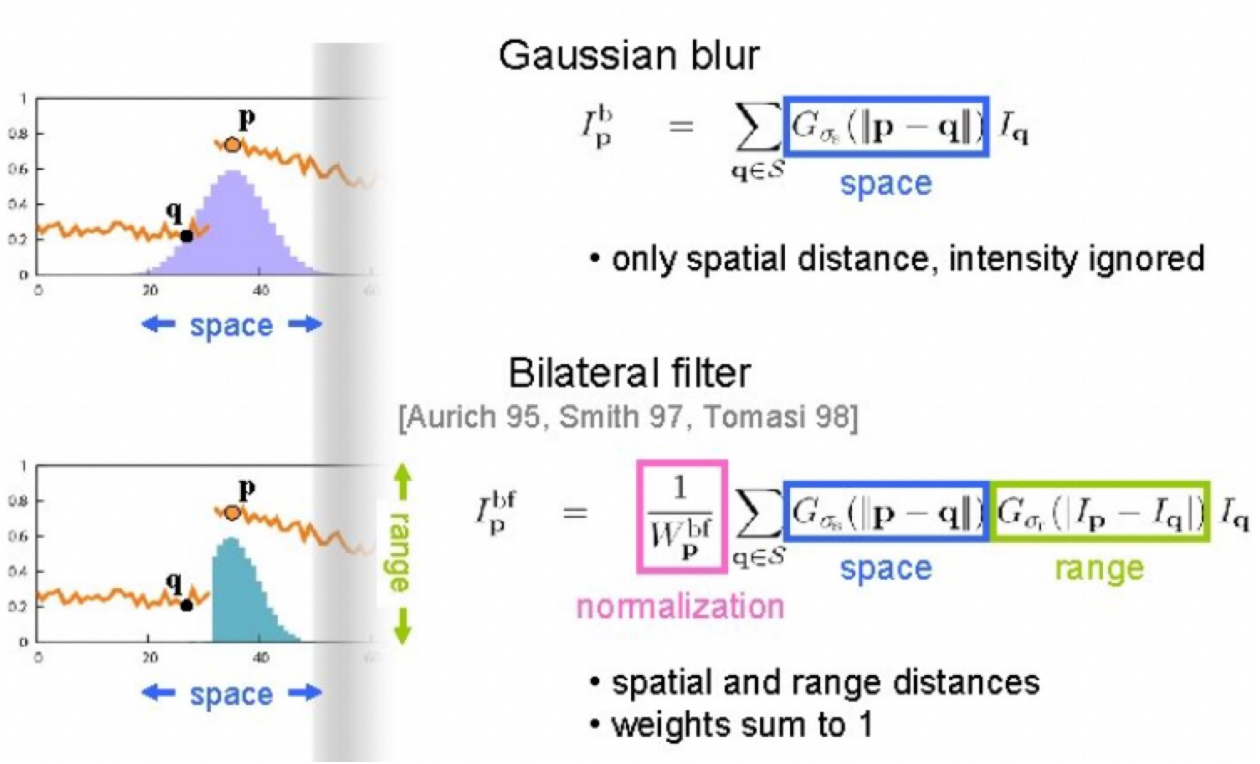
LF_{tmo} = Curve(LF)

$I_{tmo} = LF_{tmo} + HF$


MUTATIONS
TINTIN

HDR Tone-Mapping

- 使用“保边滤波器”获取图像全局分量



HDR Tone-Mapping

- 使用“保边滤波器”获取图像全局分量

$LF = BilateralFilter(I, sigma\_s, sigma\_r)$

$HF = I - LF$

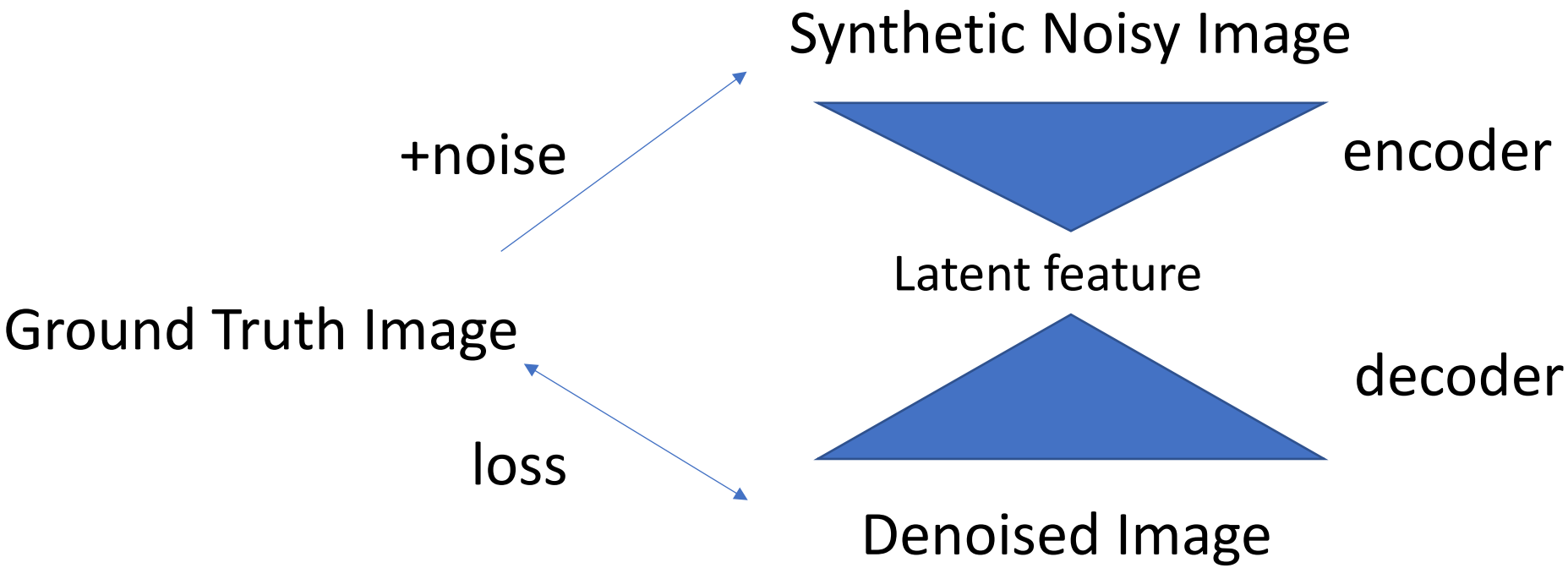
$LF_{tmo} = Curve(LF)$

$I_{tmo} = LF_{tmo} + HF$

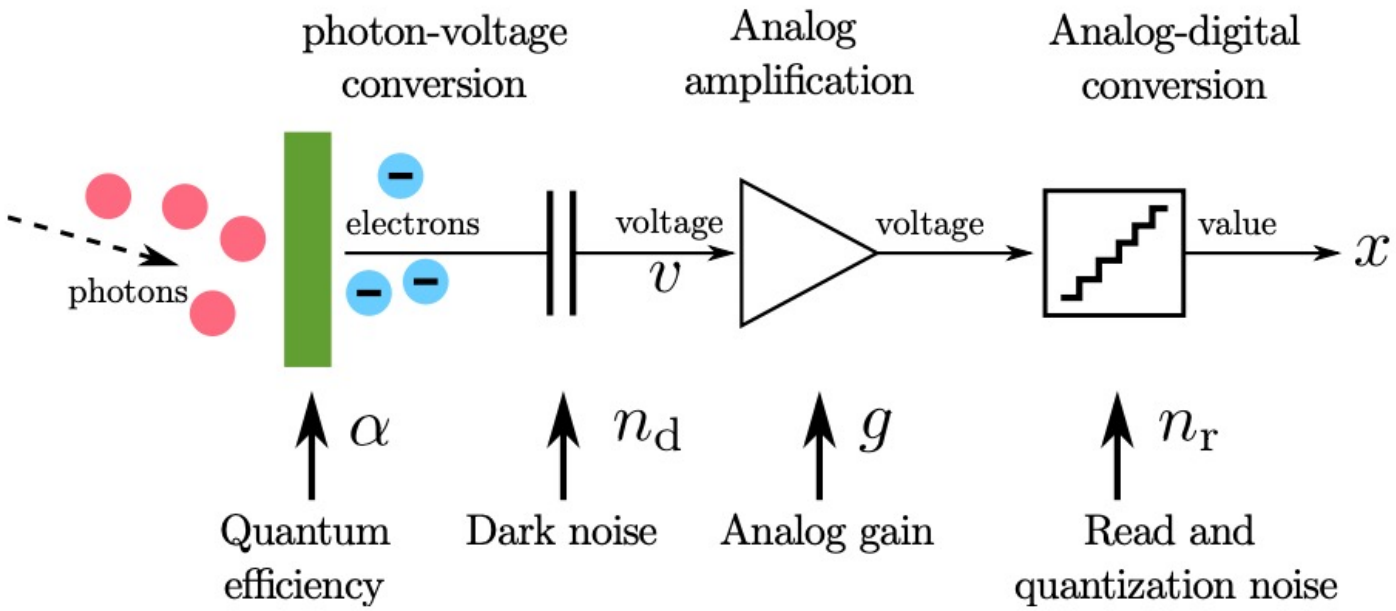

MUTATIONS
TINTIN

- 亮度与动态范围
- 解析力
- 噪声
- 色彩
- 空间感
- ……

Denoising by CNN



泊松-高斯模型



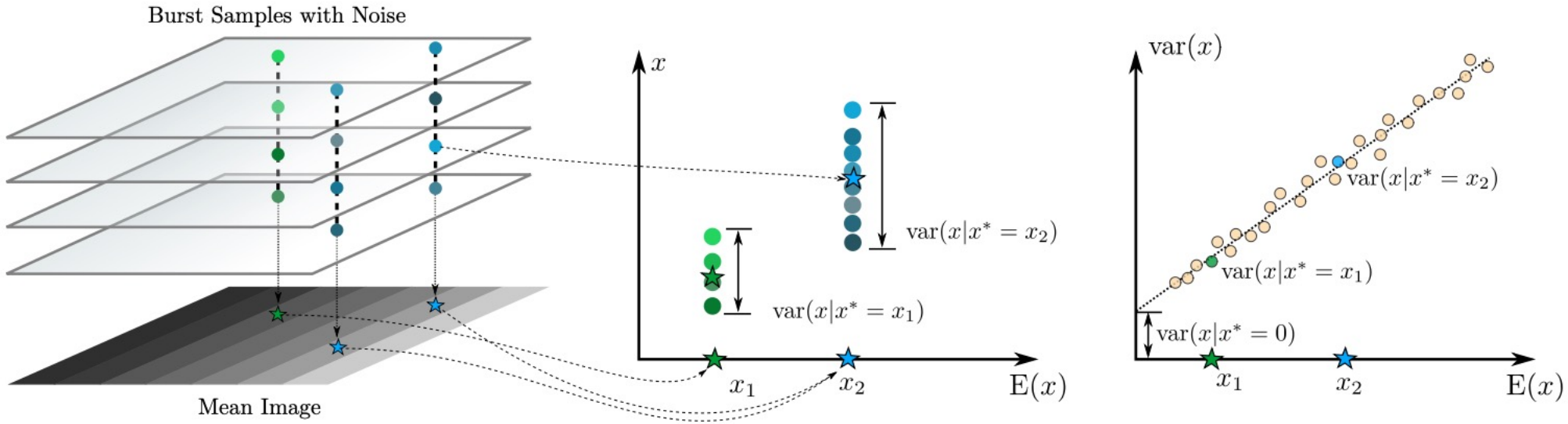
$$x \sim k\mathcal{P}(\frac{x^*}{k}) + \mathcal{N}(0, \sigma^2)$$

$$k = g\alpha$$

$$\sigma^2 = g^2\sigma_d^2 + \sigma_r^2$$

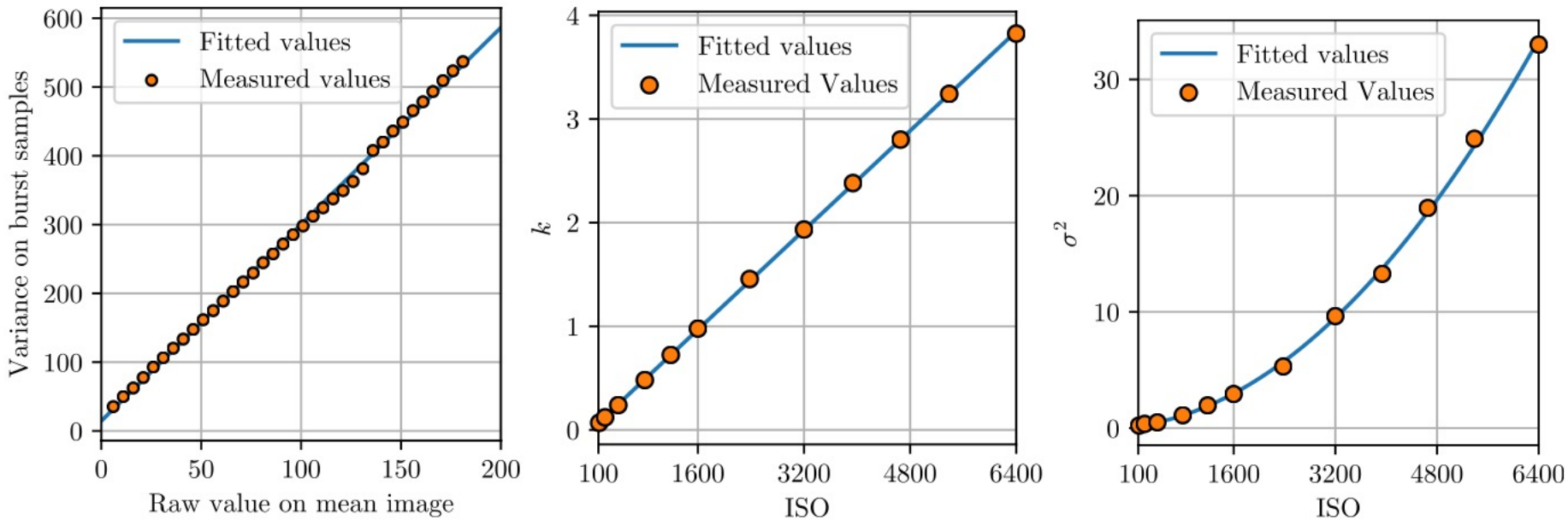
$$\begin{cases} \mathrm{E}(x) & = x^*, \\ \mathrm{Var}(x) & = kx^* + \sigma^2 \end{cases}$$

噪声模型的参数估计

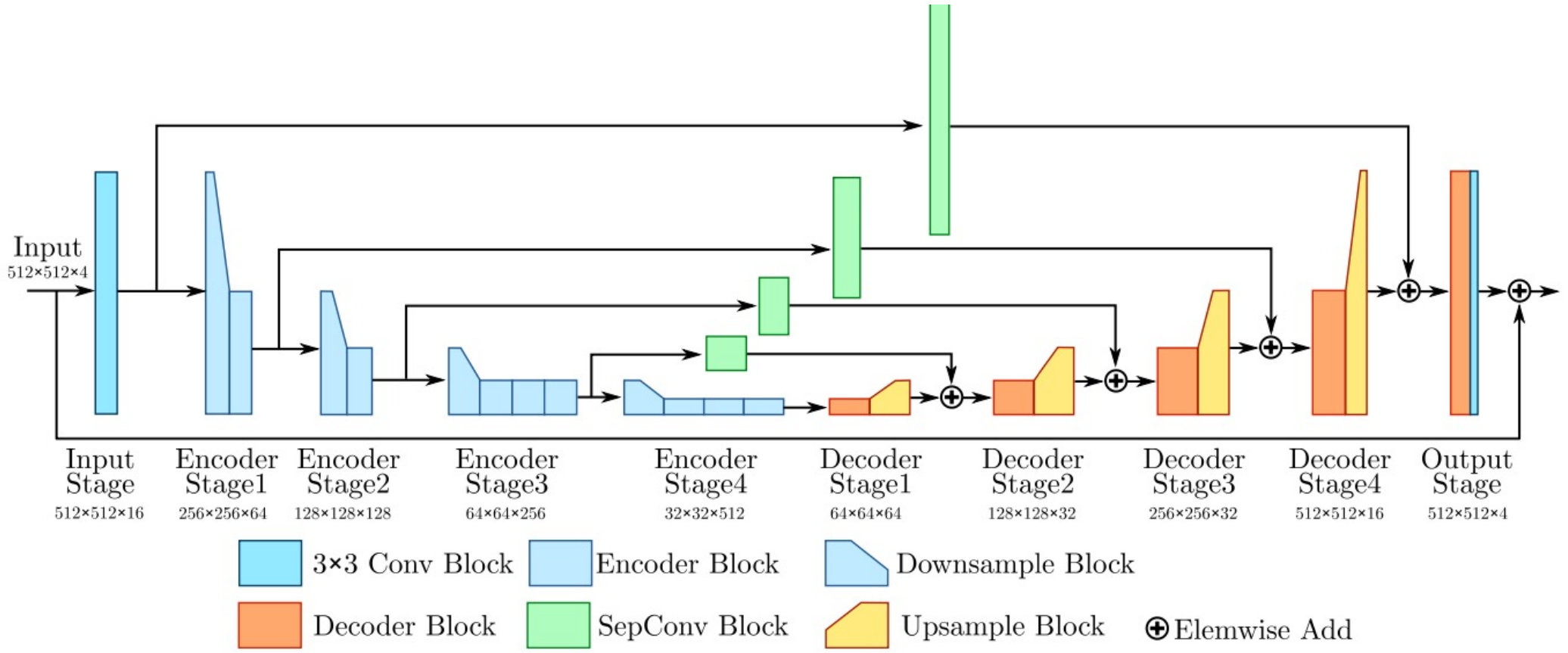


(a) Noise Sampling (b) Noise Statistics (c) Param Estimation

$$\begin{cases} \mathrm{E}(x) & = x^*, \\ \mathrm{Var}(x) & = kx^* + \sigma^2 \end{cases}$$

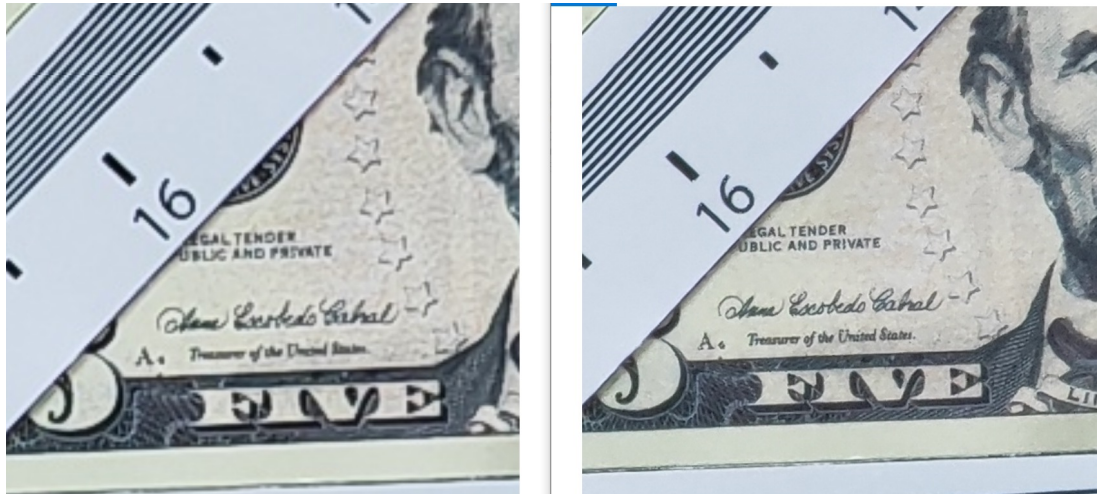

Fitted values
Measured values
Variance on burst samples
Raw value on mean image
0
50
100
150
200
0
100
200
300
400
500
600
Fitted values
Measured Values
k
ISO
100
1600
3200
4800
6400
0
1
2
3
4
Fitted values
Measured Values
σ^2
ISO
100
1600
3200
4800
6400
0
10
20
30

一卷到底


Input
512×512×4
Input Stage
512×512×16
Encoder Stage1
256×256×64
Encoder Stage2
128×128×128
Encoder Stage3
64×64×256
Encoder Stage4
32×32×512
Decoder Stage1
64×64×64
Decoder Stage2
128×128×32
Decoder Stage3
256×256×32
Decoder Stage4
512×512×16
Output Stage
512×512×4
3×3 Conv Block
Encoder Block
Downsample Block
Decoder Block
SepConv Block
Upsample Block
Elemwise Add

What’s More

多摄深度估计模拟大光圈

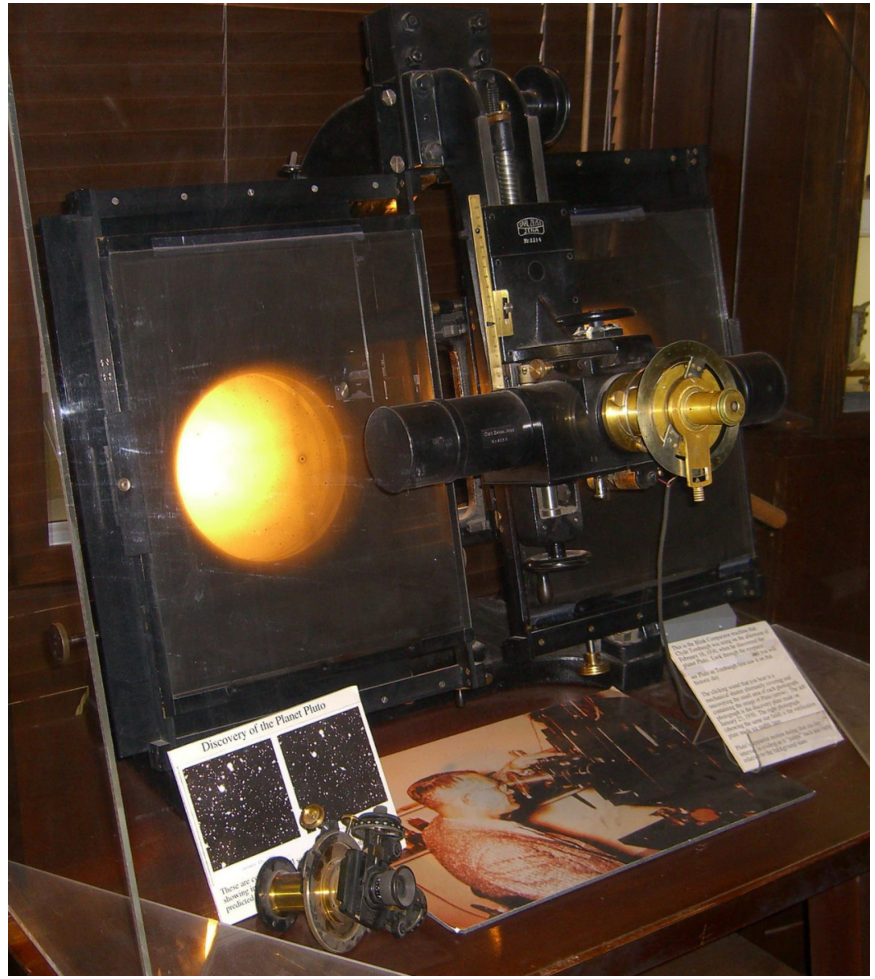


RGB / Mono 融合

128帧超分摄月

MEGSPOT

- 图片&视频浏览、对比软件
- 开源 & 跨平台
- 叠加式 & 拖拽式对比
- 视频动态对比 + 暂停细看
- 友商用了都说好



1930 年，克莱德·汤博使用
“闪烁比对器”发现了冥王星

https://github.com/MegEngine/MegSpot

私货时间