Image formation and color representation

Javier Vazquez-Corral
Universitat Autònoma de Barcelona

javier.vazquez@cvc.uab.cat

Slides credits: Marcelo Bertalmío

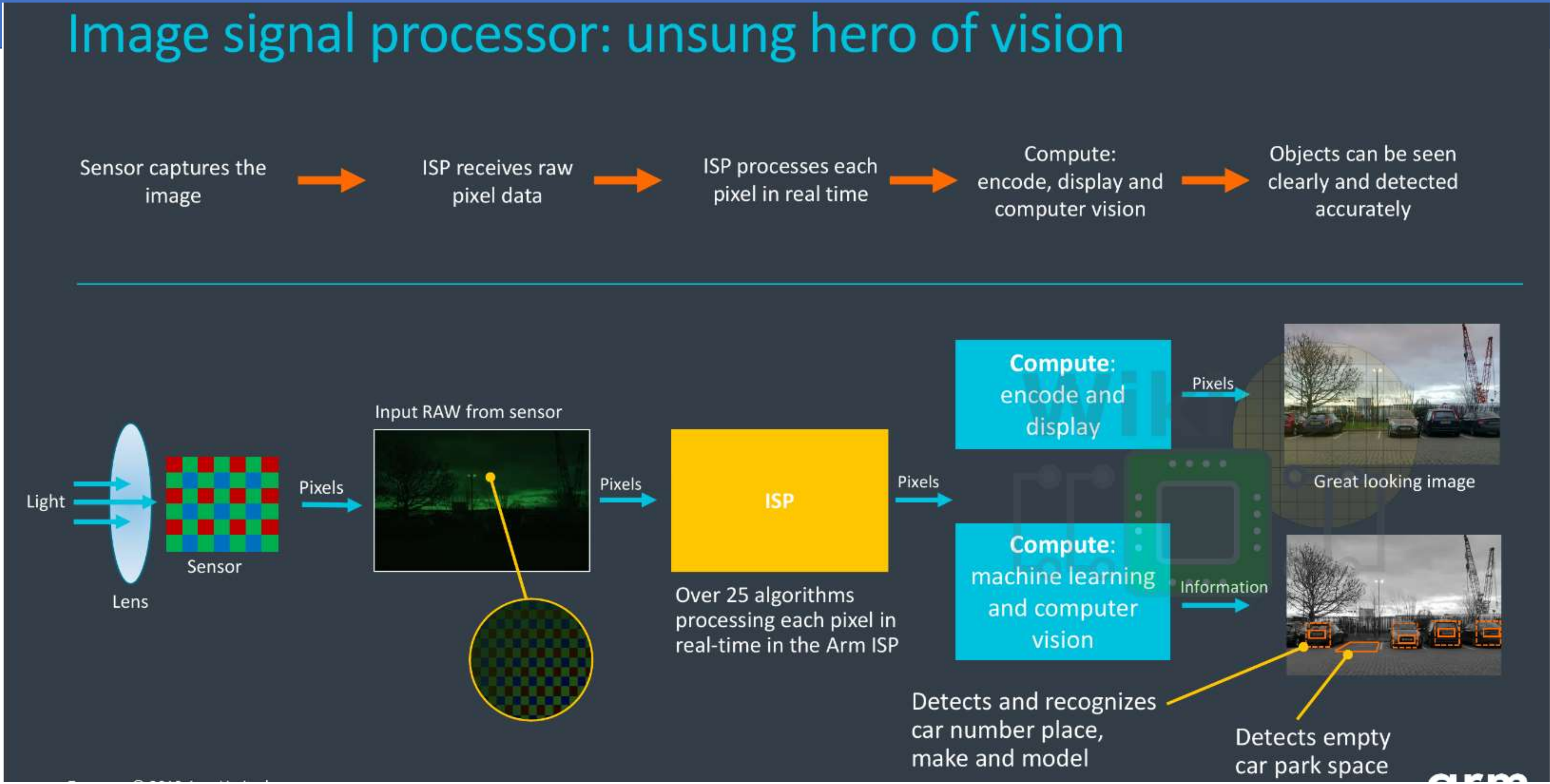
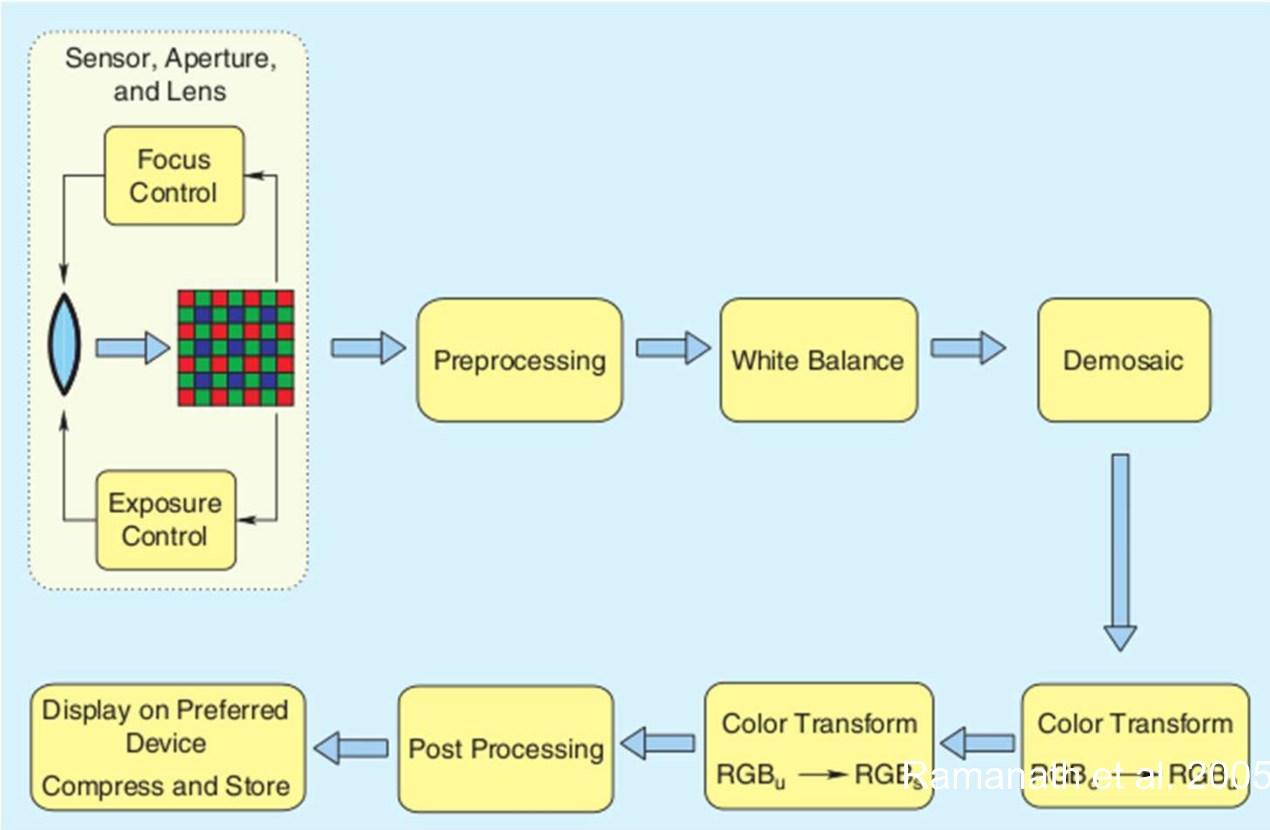


Image signal processor: unsung hero of vision
Sensor captures the image
ISP receives raw pixel data
ISP processes each pixel in real time
Compute: encode, display and computer vision
Objects can be seen clearly and detected accurately
Light
Lens
Sensor
Pixels
Input RAW from sensor
Pixels
ISP
Over 25 algorithms processing each pixel in real-time in the Arm ISP
Pixels
Compute: encode and display
Pixels
Great looking image
Compute: machine learning and computer vision
Information
Detects and recognizes car number place, make and model
Detects empty car park space

The camera color imaging pipeline

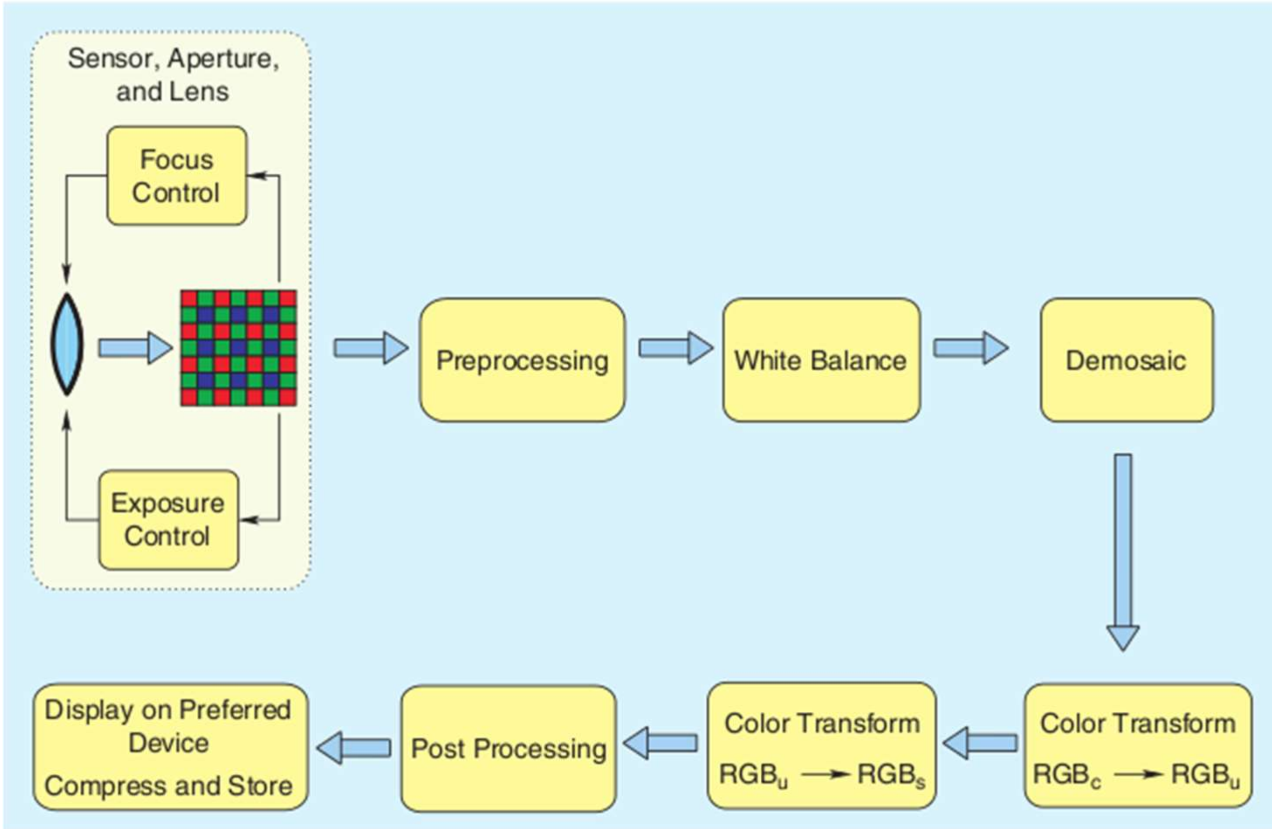


Ramanath et al. 2005

The camera color imaging pipeline

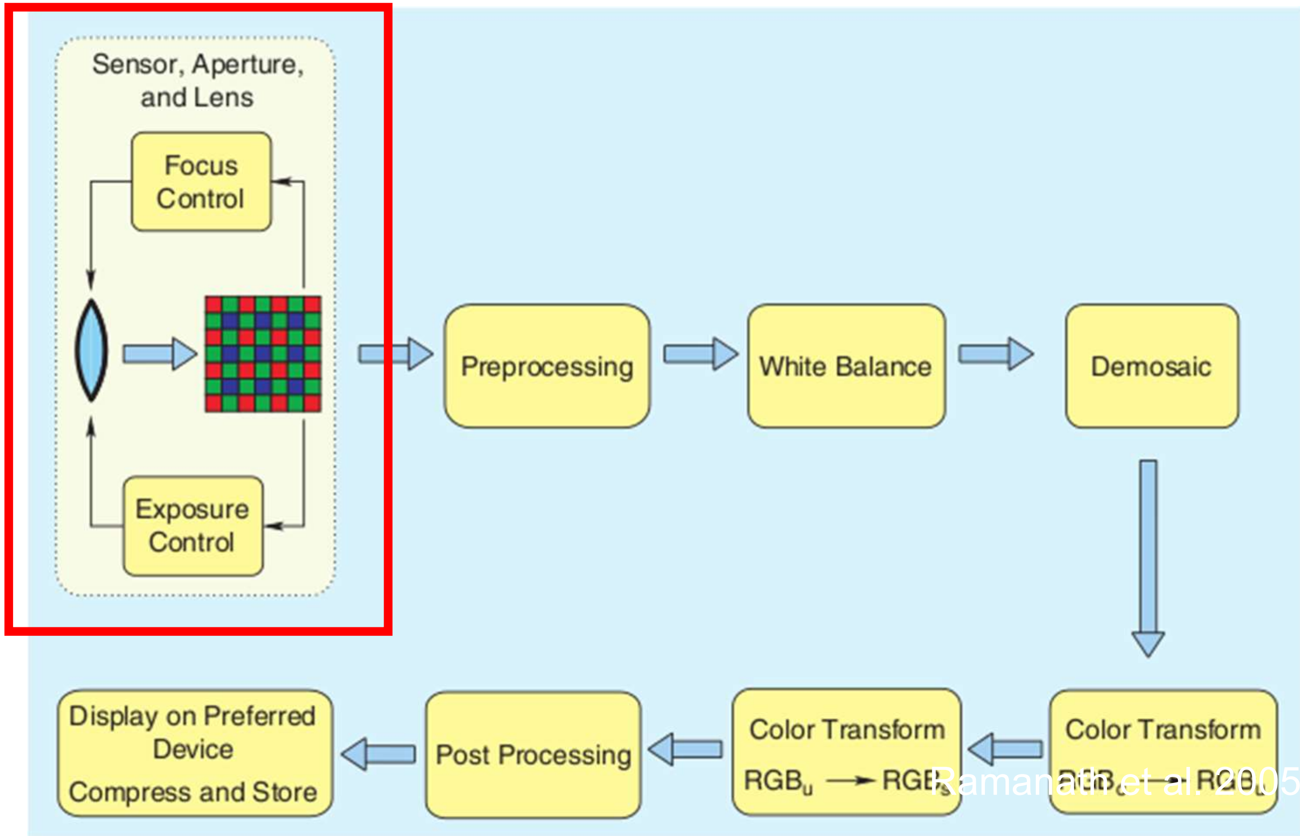
Camera makers usually don't make public this information, but digital cameras commonly perform the following operations.

Each camera maker may change the order/add some extra operations



Ramanath et al. 2005

The camera color imaging pipeline

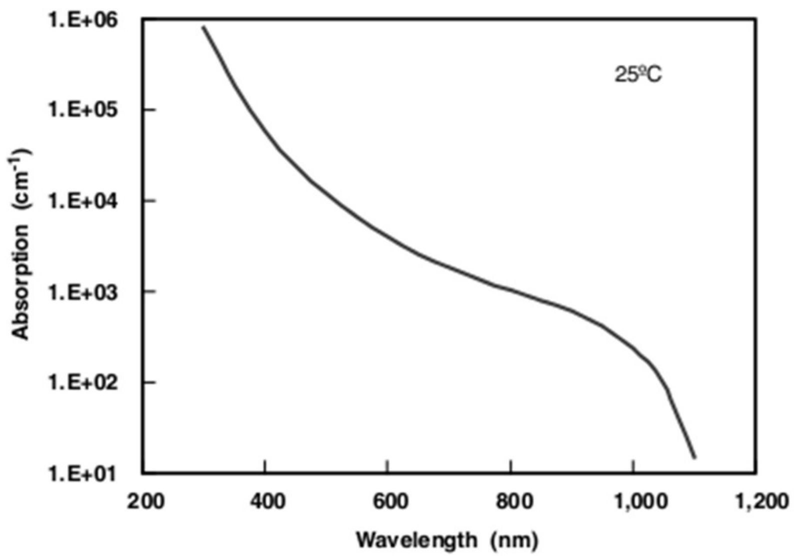


Ramanath et al. 2005

Silicon absortion

Image sensors are semiconductor devices that use the photoelectric effect to convert photons into electrical signals.

They are mainly made of silicon.



Silicon absortion

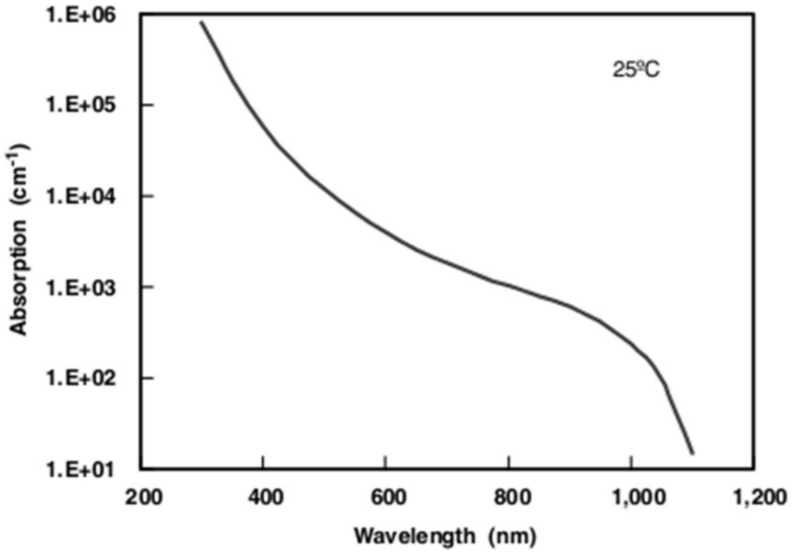
Silicon absortion

Image sensors are semiconductor devices that use the photoelectric effect to convert photons into electrical signals.

They are mainly made of silicon.

Problem?

Silicon does not absorve photons at large wavelengths



Silicon absortion

Silicon absortion

Image sensors are semiconductor devices that use the photoelectric effect to convert photons into electrical signals.

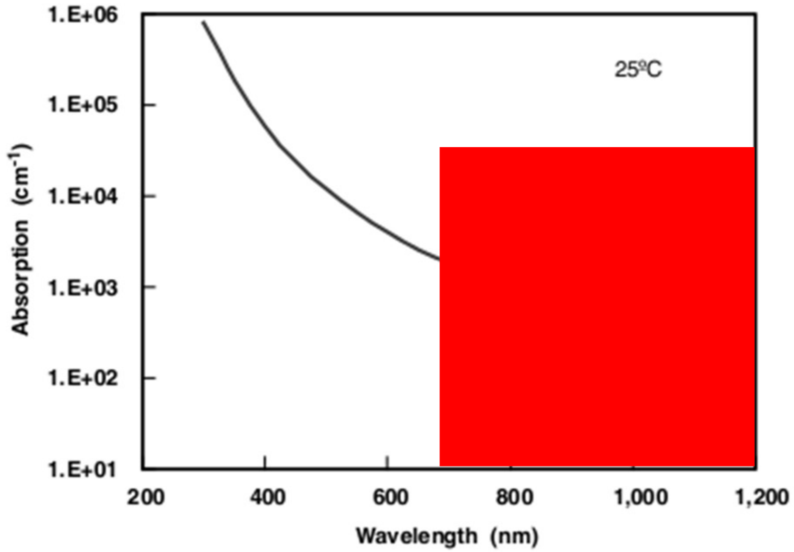
They are mainly made of silicon.

Problem?

Silicon does not absorve photons at large wavelengths

Solution:

Use hot mirror (i.e. a coated optical surface) to block wavelengths >700

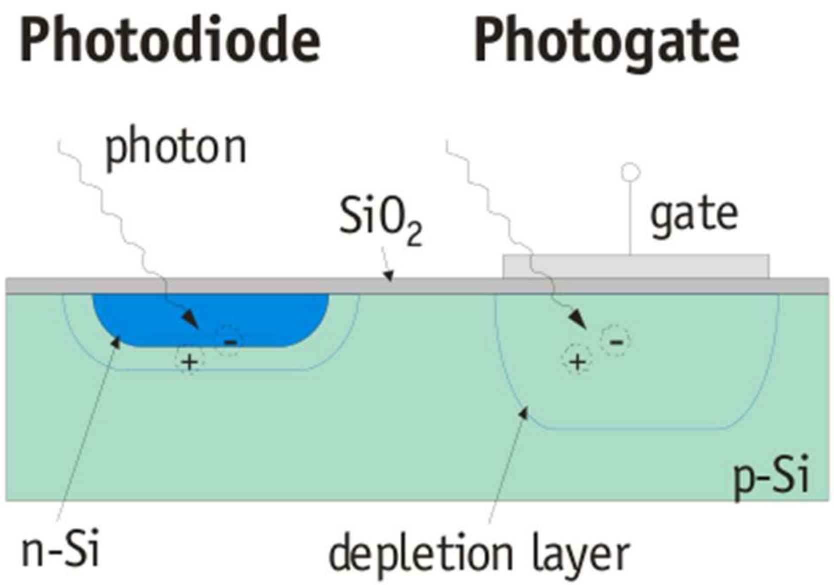


Silicon absortion

Pixel classes

- Photodiodes store the electrical charge around metal junctions created by implanting ions into the silicon.

– Photogates store the electrical charge in “potential walls” created by capacitors.



Pixel characteristics

Fill factor: the portion of the pixel area that is photosensitive.

- Photogates have a very high fill factor, of almost 100%.
- Photodiodes have a fill factor of 30-50%.

Pixel characteristics

Fill factor: the portion of the pixel area that is photosensitive.

- Photogates have a very high fill factor, of almost 100%.
- Photodiodes have a fill factor of 30-50%.

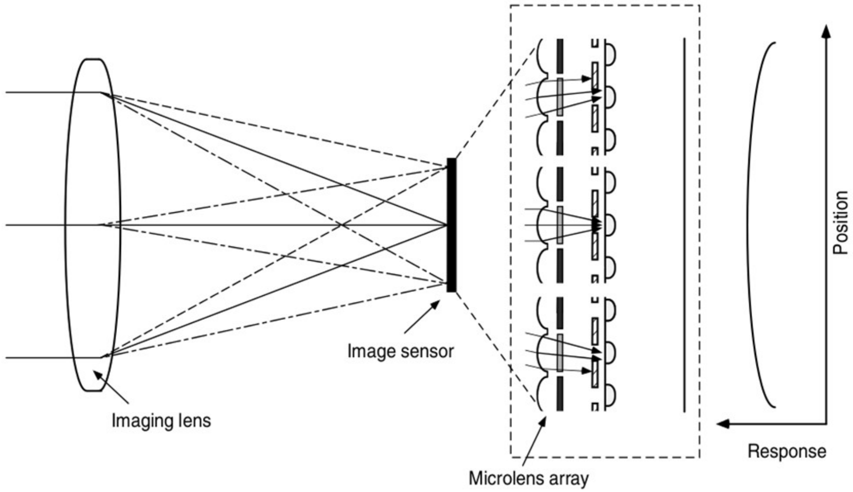
Sensitivity: Can be mesured by "quantum efficiency" (QE): the ratio of incident photons to converted charge carriers which are read out as a signal from the device

- Photogates have a reduced sensitivity in the blue end of the spectrum.
- Photodiodes have better sensitivity than photogates.

Pixel characteristics

Photogates can be improved by using very thin polysilicon gates so as not to compromise sensitivity.

Photodiodes may use microlenses to increase fill factor. But microlenses produces shading in the image, because now the percentage of photons that are focused onto the photosensitive part of the pixel depends on the angle of incidence of the light.



Sensor classes

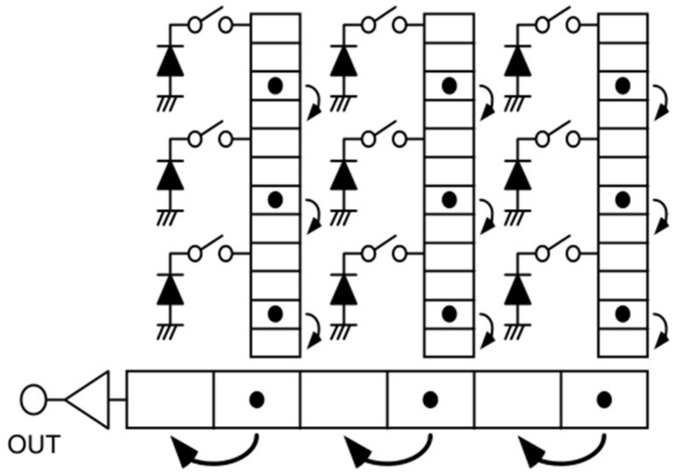
The electrical charge accumulated while the sensor is being exposed to light must be converted into voltage or current through the scanning of the image array.

Sensor classes

The electrical charge accumulated while the sensor is being exposed to light must be converted into voltage or current through the scanning of the image array.

Two classes of sensors depending on how the scanning is performed:

- Charged Coupled Device (CCD): transfers the charge vertically from pixel to pixel over the entire array, then it is transferred horizontally and converted into voltage at just one output amplifier.



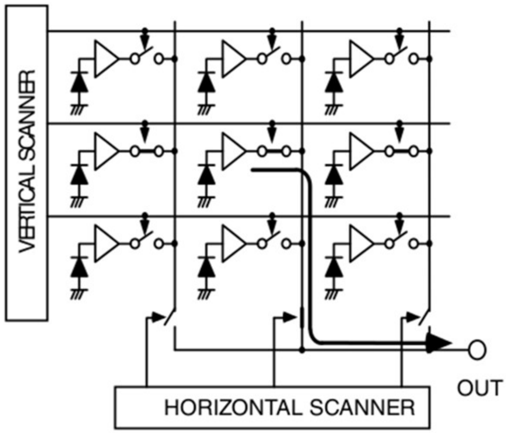
Sensor classes

The electrical charge accumulated while the sensor is being exposed to light must be converted into voltage or current through the scanning of the image array.

Two classes of sensors depending on how the scanning is performed:

- Charged Coupled Device (CCD): transfers the charge vertically from pixel to pixel over the entire array, then it is transferred horizontally and converted into voltage at just one output amplifier.

- CMOS (Complementary Metal Oxide Semiconductor): performs the charge-to-voltage conversion at each pixel location.



Sensor classes

The electrical charge accumulated while the sensor is being exposed to light must be converted into voltage or current through the scanning of the image array.

Two classes of sensors depending on how the scanning is performed:

- Charged Coupled Device (CCD): transfers the charge vertically from pixel to pixel over the entire array, then it is transferred horizontally and converted into voltage at just one output amplifier.

- CMOS (Complementary Metal Oxide Semiconductor): performs the charge-to-voltage conversion at each pixel location.

- CCDs use global shutter, i.e. it exposes the entire image simultaneously.

- CMOSs usually use Rolling shutter, i.e. exposes different portions of the frame at different points in time, “rolling” through the frame.

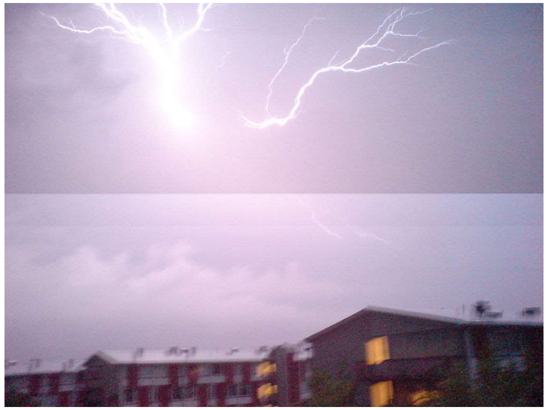
Issues of CMOSs vs CCDs

The vertical charge transfer of CCD can cause smear (blooming)

Issues of CMOSs vs CCDs

The vertical charge transfer of CCD can cause smear (blooming)

Rolling shutter (CMOS) can provoke distortions of fast-moving objects or rapid flashes of light.

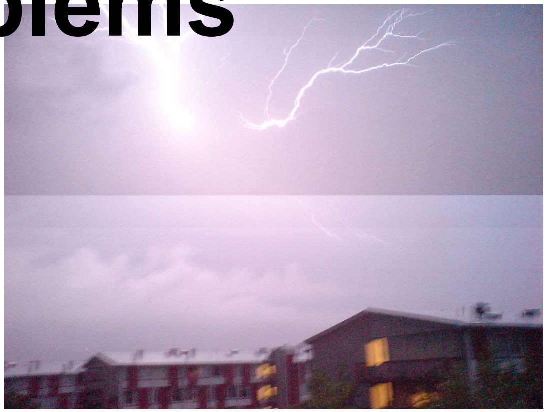


Issues of CMOSs vs CCDs

The vertical charge transfer of CCD can cause smear (blooming)

Camera makers add techniques to alleviate these problems

Rolling shutter (CMOS) can provoke distortions of fast-moving objects or rapid flashes of light.



Back in 2009 CCDs and CMOSs were used in high-end cinema cameras

Updated: 9/13/2009 9:45 PM

FLETCHER CAMERA - CHICAGO 312-932-2700 - MICHIGAN 248-471-9906 - IOWA 800-635-3824

Sorted by: Price

Camera	Imager (Actual Size)	Lens Mount	ISO Base	Latitude	Frame Rates	Digital Sampling on Recorded Media	Recorded Bit Depth Format & Time	Weight	Power Draw	Highlighted Positives	Notable Credits	Average NATIONAL Daily Rental
 Phantom HD Gold	CMOS 36 Ø 25.6 x 25.6mm	PL PV Nikon Canon	250	10 Stops	5-1052 fps @ 1920x1080	2048 x 2048 Uncompressed RAW 1920 x 1080	14 Bit Uncompressed RAW 32G Internal RAM 8.9s @ 1000fps 512G CineMags 3m14s @ 1000fps[2]	12 lbs. 14 lbs. w/ CineMag	80w	Compact Camera Size Uncompressed RAW CineStation Workflow Anamorphic Lenses 14 Bit RAW	*"Hurt Locker"* [1] *"Max Payne"* [1] *"The Spirit"* [1] Numerous National Spots	$5000 w/ 2 CineMag (1+TB) $2500 Camera Only (32GB)
 Weisscam HS-2	CMOS 31.4Ø 22.2 x 22.2mm	PL PV Nikon Canon	320	9 Stops	1-2000 fps @ 1920x1080	2048 x 1536 Uncompressed RAW 1920 x 1080	12 Bit Uncompressed RAW 32G Internal RAM 12s @ 1000fps DigiMag (2 TB) 12m30s @ 1000fps	13.2 lbs 29 lbs w/ DigiMag	70w Camera 85w DigiMag	Simultaneous HD-SDI & RAW Wireless Interface Anamorphic Lenses	Numerous National Spots	$3750 w/ 1 Digi-Mag (2 TB) $3000 Camera Only (16GB)
 Sony F35	CCD 27.1 Ø 23.6 x 13.3mm	PL	320	12 Stops	1-50 fps @ 1920x1090	1920 x 1080	10 Bit HDCAM SR 50 Min 10 Bit DPX s.two OB-1 30 Min	24 lbs. w/ SRW1 18lbs. w/OB-1	106w w/SRW1 75w w/OB-1	4:4:4 Color Sampling 14 bit A/ D Low Light Performance Easy and Established HDCAM SR Workflow	*"CSI-NY"* *"Cold Case"* *"Weeds"* Numerous National Spot	$3500 with SRW-1 & SRPC-1 or OB-1
 Arri D-21 w/OB-1	CMOS 29.6Ø 23.7x17.8m	PL	200	11 Stops	1-60 fps 4:2:2 1-30 fps 4:4:4 1-25 fps RAW	2880 x 2160 Uncompressed ARRIRAW 1920 x 1080	12 Bit ARRIRAW OB-1 30 min 10 Bit HDCAM SR 50 min	23 lbs. w/OB-1 16 lbs. Camera Only	60w Camera 56w SRW-1	4:4:4 Color Sampling Optical Viewfinder Anamorphic lenses	*"Five Killers"* *"The Company"* *"The Bank Job"*	$2750 with SRW-1 & SRPC-1 or OB-1
 Sony F23	2/3" 2/3" 2/3" CCD 11 Ø 8.8 x 6.6mm	B4	320	11 Stops	1-60 fps @ 1920x1080	1920 x 1080	10 Bit HDCAM SR 50 Min	11 lbs. Camera Only 24 lbs. w/SRW-1	106w w/SRW1 75w w/OB-1	4:4:4 Color Sampling Established Workflow Wide Latitude Wide Gamut	*"Speed Racer"* *"Public Enemies"*	$2750 with SRW-1 & SRPC-1
 Red One	CMOS 27.9Ø 24.4 x 13.7mm	PL PV Nikon Canon	250 to 320	10 Stops	1-30p fps @4K 1-60 fps @3K 1-120fps @2K	2K 3K 4K	12 Bit Card (16GB) 9 mins Ram (120GB) 60 min Drive (320GB) 120 min	9 lbs. Camera Only	65w	Cost vs Performance Compact Size	*"District 9"* *"The Informant"* *"Book of Eli"* Numerous National Spots	$1200 with 16 GB Cards & RAM Drive

[1] Portions of the Project [2] Top speed direct record is 450fps

Camera Must Be Available For Sale

Back in 2009 CCDs and CMOSs were used in high-end cinema cameras

Updated: 9/13/2009 9:45 PM

FLETCHER CAMERA - CHICAGO 312-932-2700 - MICHIGAN 248-471-9906 - IOWA 800-635-3824

Sorted by: Price

FLETCHER	Imager (Actual Size)	Lens Mount	ISO Base	Latitude	Frame Rates	Digital Sampling on Recorded Media	Recorded Bit Depth Format & Time	Weight	Power Draw	Highlighted Positives	Notable Credits	Average NATIONAL Daily Rental
 SI-2K	CMOS CMOS 11.7 Ø 10.3x5.76mm	PL B4 C Nikon	250	11 Stops	1-25 fps @ 2K 1-30 fps @ 1080p 1-85 fps @ 720p	2K 1920 x 1080	12 Bit Uncompressed Raw 10 Bit CineForm 40GB/Hour	16 lbs. Standard 1.2 lbs. Mini	60w	Mini Remote Head Live 3D-LUT/Keyer Direct to Edit AVI/MOV Compact Small Size Popular for 3D	*"Slumdog Millionaire"*[1] *"My Bloody Valentine 3D"*[1] Numerous National Spots	$1200 with 128G Flash RAM
 Panasonic HPX-3000/3700	2/3" 2/3" 2/3" CCD 11 Ø 8.8 x 6.6mm	B4	400	10 Stops	**3000** 24, 25p, 30p **3700** 1- 30 fps	1920 x 1080	10 Bit 5 - 32GB P2 Cards 200 min	10.5 lbs	43w	"HD D5" Quality Cost vs Performance FCP 7 Supports Native AVC-I 3700 Offers 4:4:4 Out	*"Party Down"* *"Doll House"* Numerous National Spots	$1200 with 2 16GB Cards
 Sony PDW-F800	2/3" 2/3" 2/3" CCD 11 Ø 8.8 x 6.6mm	B4	500	10 Stops	1080 1-30fps 720 1-48 fps @ 23.98 1-60 fps @ 29.97	1920x1080 1280x720	8 Bit XDCAM Optical Disks 95 Min	9 1/2 lbs.	44w	Combines the best of a tape and tapeless world. Low Cost Media Native support by Major NLE	*"Parks & Recreation"* Mark Burnett Productions	$1,000
 Panasonic HPX-2000/2700	2/3" 2/3" 2/3" CCD 11 Ø 8.8 x 6.6mm	B4	640	10 Stops	**2000** 24, 25 30p, 50p, 60P **2700** 1-60fps	1280x720	10 Bit 5 - 32GB P2 Cards 200 min	10 lbs.	43w	HD D-5 quality @ Relatively low data & Rental rates Also Record 1080i	Numberous Regional Spots	$800 w/ five 16GB Cards
 Sony PMW-EX3	1/2" 1/2" 1/2" CMOS 8 Ø 6.9 x 3.2mm	EX Mount	400	7 Stops	1-30fps @1080p 1-60fps @ 720p	1920 x 1080	8 Bit 2 - 32GB SxS Cards 200 min	7.9 lbs.	13.5w	Cost vs Performance Size Flexibility to mount Professional 2/3" Lenses	*"Melrose Place"*[1] Numerous Regional Spots	$500 w/ two 16GB Cards
 Sony PMW-EX1	1/2" 1/2" 1/2" CMOS 8 Ø 6.9 x 3.2mm	FIXED	400	7 Stops	1-30 fps @ 1080p 1-60 fps @ 720p	1920 x 1080	8 Bit 2 - 32GB SxS Cards 200 min	6.2 lbs.	12w	Cost 1/2" Imager Full 1920x 1080 imager and recording	*"Dexter"*[1] Numerous Regional Spots	$375 w/ two 16GB cards
 for comparison Arricam ST	Full Aperture 31.1Ø 24.9x18.1mm Kodak FILM	PL	500	14 Stops	1-60fps	1920x1080 2K 4K 6K Uncompressed	16 bit (Linear) 10 bit (log) 3 Perforation 14m48s 1000' 4 Perforation 11m 06s 1000'	25 lbs. 400' Load 28 lbs. 1000' Load	55w	4:4:4 Color Sampling Established Workflow Excessive Latitude Proven Archival Value	*"The Black Dahlia"* *"Chicago"* *"Lord of the Rings"* *"Slumdog Millionaire"*1	$2,000 w/2 Mags

[1] Portions of the Project [2] Top speed direct record is 450fps

Camera Must Be Available For Sale

By 2014, CMOSs dominated the market

Los Angeles - Chicago - New Orleans - Miami

2014 CAMERA COMPARISON CHART

www.cineverse.net

CINEVERSE

Camera	Imager (Actual Size)	ISO	Latitude	Frame Rates (Progessive Only)	Digital Sampling on Recorded Media	Recorded Bit Depth Format & Time	Data	Weight Body Only	Power	Highlighted Positives	Notable Credits	Average National Daily Rental BODY ONLY
 ALEXA XT Plus	CMOS 33.5mm Ø 28.2 × 18.13mm 4:3	800	14+ Stops	**Open Gate** .75-75 **4:3 Mode** .75-90fps **16:9 Mode** .75-120 fps	ARRIRAW Open Gate 3414 x 2198 4:3 2880 x 2160 16:9 2880 x 1620	12 Bit ARRIRAW Uncompressed 48 min - 512GB in 16:9 36 min - 512GB in 4:3 Optional (1) SxS or CFast 2.0 Slot Records all **Classic ALEXA** formats (See Below)	10.6 GB per Min 16:9 14.2 GB per Min 4:3 @ 24fps	16.1 lbs Plus 18.3 lbs Studio	103w Arriraw 100w ProRes	True Anamorphic 4:3 Imager Wide Latitude Gently Rolls Off Highlights Proven Reliable, Post-Efficient ARRIRAW, ProRes or DNxHD Optical Viewfinder & Mirror Shutter (Studio Version)	*Fault In Our Stars Birdman Ruth & Alex Unbroken Guardians of the Galaxy*	$3,000 XT Studio $2,700 XT Plus $2,500 XT
 Sony F65 & SR-R4	CMOS 27.9mm Ø 24.7 x 13.1mm 17:9	800	14 Stops	1 to 120 fps Focal Plane Shutter OFF 1 to 60 fps Shutter ON	8K 8192 x 2160 2K 1920x1080	16 Bit F65RAW - SQ @ 3:1 30 min- 512GB 60 min - 1TB (Also Records RAW-Lite SR-File Codec @ 10/12bit)	17 GB per Minute @ 24fps	12 lbs 16 lbs w/ SR-R4	62w Body Only 105w w/SR-R4	16 Bit 8K RAW to 120fps Rotary Shutter 2x Anamorphic De-Squeeze OLED Viewfinder Independent 1D and 3D LUT Live Streaming & Control w/ WiFi	*Oblivion After Earth Tomorrowland Lucy Masters of Sex Monkey Kingdom*	$2,500
 Arri ALEXA XT M	CMOS 29.7mm Ø 23.8 × 17.8mm 4:3	800	14+ Stops	**4:3 Mode** .75-120fps **16:9 Mode** .75-120 fps	ARRIRAW 4:3 2880 x 2160 16:9 2880 x 1620	12 Bit ARRIRAW Uncompressed 48 min - 512GB in 16:9 36 min - 512GB in 4:3 Optional (1) SxS or CFast 2.0 Slot Records all **Classic ALEXA** formats (See Below)	10 GB per Minute @ 24fps	6.4 lbs Head 12.8 lbs XT 12.1 lbs SxS	35w Head 65w Body 131w XT System	Small Ideal for 3D & Cars True Anamorphic 4:3 Imager Low Light Performance Wide Latitude Pleasing Skin Tones Gently Rolls Off Highlights	*Dawn of the Planet of the Apes*[3] *Into the Storm X-Men: Days of Future Past*[3]	$2,500
 ALEXA Classic EV/Plus	CMOS 27.2mm Ø 23.8 × 13.4mm 16:9	800	14+ Stops	.75-120 fps @ ProRes 4:2:2 HQ .75-60 fps @ 2K ProRes or 4:4:4	2K 1920 x 1080	12 Bit ProRes 4:4:4 32GB SxS - 14 min 64GB SxS - 28 min	2 GB per Minute @ 24fps	13.7 lbs EV 15.4 lbs Plus	85w	Low Light Performance Wide Latitude Cost Effective Workflow Pleasing Skin Tones Gently Rolls Off Highlights	*Game of Thrones Homeland Downton Abbey Random*	$1,800 Plus $1,600 EV w/120fps $1,400 EV
 Arri AMIRA	CMOS 27.2mm Ø 23.8 × 13.4mm 16:9	800	14+ Stops	.75-200fps Advanced & Premium .75-100fps Standard	2K Premium 1920 x 1080 Standard & Advanced	12 Bit ProRes 4:4:4 Premium 10 Bit ProRes 4:2:2 HQ Adv. 10 Bit ProRes 4:2:2 Stnd. 120GB CFast - 48 min	1.9 GB per Minute @ 24fps	9.2 lbs	50w	Lightweight 200 fps Slow Motion WiFi *(Advance and Premium)* Single Use Operator Faster Media for Off Loading Established ALEXA Image Quality	Camera Newly Released Credits Coming Soon	$1,500
 Sony F55 w/R5	CMOS 27.1mmØ 24 x 12.7mm 17:9	1250	14 Stops	24, 25, 30, 50, 60 1-60fps @ 4K 1-180fps @ 2K 1-180fps @ HD 1-240fps @ 2K Raw	4K 4096 x 2160 3840 x 2160 2K HD	16 Bit F55 RAW @ 3.6:1 512GB AXSM - 66 min XAVC HD 2K 4K - 10 Bit SR-SQ & SR-Lite 64GB SxS - 30 min	7.7 GB per Minute @ 24fps	6.4 lbs F55 & R5 5 lbs F55	48w System 25w Body Only	Light Weight and Small Profile 1250 ISO Sensor w/ Global Shutter 2x Anamorphic De-Squeeze 2K Mode Uses Full Image Sensor 4K XAVC is Small File Size Wide Latitude & Color w/ S-Log3	*Blacklist Big Bang Theory Annie Alpha House Deliver us from Evil*	$1,250

By 2014, CMOSs dominated the market

Los Angeles - Chicago - New Orleans - Miami

2014 CAMERA COMPARISON CHART

www.cineverse.net

CINEVERSE	Imager (Actual Size)	ISO	Latitude	Frame Rates (Progessive Only)	Digital Sampling on Recorded Media	Recorded Bit Depth Format & Time	Data	Weight Body Only	Power	Highlighted Positives	Notable Credits	Average National Daily Rental BODY ONLY
 Red Epic Mysterium-X	CMOS 31.4mm Ø 27.7 × 14.6mm 1.90:1	800	12+ Stops 15-18+ w/ HDRx	1-120 fps @5K 1-120 fps @4.5K 1-150 fps @4K 1-300fps @2K	5K FF, A2:1, 2:1 2.4:1 4.5K 2.4:1 4K 16:9, HD, A2:1, 2.1 2K 16:9, A2:1, 2.1 1080 & 720 16:9	12 Bit REDCODE - 5K FF @ 7:1 128GB SSD - 40 min 256GB SSD - 80 min (HDRx cuts time in half)	3.2 GB per Minute @ 24fps	5 lbs	60w	Self Contained - Ideal for 3D HDRx High Dynamic Range Established R3D workflow Modular Design High Frames per Second	*Great Gatsby* [3] *Hobbit* [3] *House of Cards* *Justified*	$1,000
 Canon EOS C500	CMOS 27.3mm Ø 24.6 x 13.8mm 16:9	850	12 Stops	24p, 25p, 30p, 50i, 60i @ HD 24, 25, 30, 50, 60, 120 @ 4K to External Recorder	1920 x 1080 Outputs 2K & 4K to External Recorder	8 Bit MPEG2-4:2:2 MXF 64GB CF - 160 min 10 Bit 4K RAW to External Recorder	0.4 GB per Minute @ 24fps	4 lbs	11.4w	4K Output w/ External Recorder PL or EF mount High Dynamic Range Small Self Contained Ideal for 3D	*Need for Speed* *Amityville Horror: The Lost Tapes* *Fathers & Daughters*	$900
 Sony F5	CMOS 27.1mm Ø 24 x 12.7mm 17:9	2000	14 Stops	24, 25, 30, 50, 60 1-180fps @ 2K & HD 1-240fps @ 2K Raw 1-60fps @ 4KRaw	2K 1920 x 1080 Outputs 4K to External Recorder	10 Bit XAVC & SR-SQ & Lite 128GB SxS - 71 min 4K F5RAW to External Recorder	1.8 GB per Minute @ 24fps	5 lbs	25w	Light Weight & Small Profile Good Low Light Performance 2K Mode Uses Full Image Sensor Wide Latitude & Color w/ S-Log3 180fps Recording with SxS media	*Man from Reno* *Parts UnKnown* *The Voice* *Amazing Race*	$450
 Canon EOS C300	CMOS 27.3mm Ø 24.6 x 13.8mm 16:9	850	12 Stops	24, 25, 30 fps @ 1080	1920 x 1080	8 Bit MPEG2-4:2:2 MXF 64GB CF - 160 min	0.4 GB per Minute @ 24fps	3.2 lbs	11.4w	Incredible Low Light Performance Small Size C-Log Workflow Dual Pixel Auto Focus (EF Version Only)	*The Green Inferno* *Blue Ruin* *La vie d'Adèle*	$450
 ARRICAM ST - 35mm Film	Full Aperture 31.1mm Ø 24.9x18.7mm 4:3	500	15-16 Stops	1-60fps	6K - 4K - 2K Uncompressed *(via Scanner)*	16 Bit (Linear) 2P 22m12s 1000' 3P 14m48s 1000' 4P 11m06s 1000'	N/A	25 lbs 400' Load 28 lbs 1000' Load	55w	4:4:4 Color Sampling Established Workflow Widest Available Latitude Proven Archival Value	*12 Years a Slave* *King's Speech* *Cloud Atlas* *Silver Linings* *Playbook*	$1,000 w/Mags
 Phantom Flex 4K	CMOS 31.7mm Ø 27.6 x 15.5mm	800	12 Stops	1000fps @ 4K 2000fps @2K 2000fps @ 1080	CineMag IV 4096 x 2304	12 Bit - RAW 64GB Internal RAM 4.7s @1000fps - 4K	750GB per Minute @ 1000fps	13 lbs	110w	High Resolution for Repositioning Sync Sound also works as 24fps Familiar User Interface Attachable On-Board Battery	Camera Newly Released Credits Coming Soon	$5,500 w/ (2) 2TB CineMags $3,500 Camera Only (64GB)
 Phantom Flex 2K	CMOS 30.1mm Ø 25.6 x 16mm	1250	10 Stops	1-1617 fps @ 2560x1440 1-2564 fps @ 1920x1080	2560 x 1440 Uncompressed RAW 1920 x 1080	12 Bit - RAW 32GB Internal RAM 4.5s @ 2500fps 512G CineMags 1m14s@2500fps	170 GB per Minute @ 1000fps	12 lbs	100w	Efficient Professional Workflow Industry Standard Low Light Performance Uncompressed RAW 4:4:4 Output	*Godzilla (2014)* [2] *Gangster Squad* [2] *Iron Man 3* [2] *Rush* [2]	$5,000 w/ CineMags (512GB) $3,000 Camera Only (32GB)

CMOSs vs CCDs today

	CCD	CMOS
Resolution	Up to 100+ Megapixels	Up to 100+ Megapixels
Frame rate	Best for lower frame rates	Best for higher frame rates
Noise figure	Lower noise floor → Higher image quality	Higher noise floor → Lower image quality
Responsivity and linearity	Lower responsivity, broader linear range	Higher responsivity, lower linear range (saturates early)
Limit of detection	Low (more sensitive at low intensity)	High (less sensitive at low intensity)
Color depth	Higher (16+ bits is typical for expensive CCDs)	Lower, although becoming comparable to CCDs (12-16 bits is typical)

https://octopart.com/blog/archives/2020/05/ccd-vs-cmos-sensor-comparison-which-is-best-for-imaging

Other capture problems: Aliasing

Artifacts appear due to limited spatial resolution

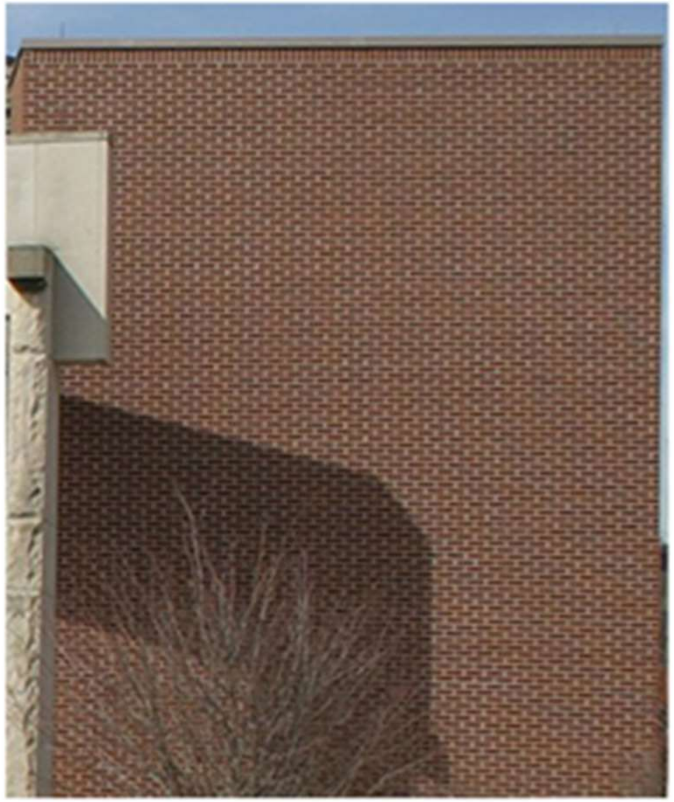
Other capture problems: Aliasing

Artifacts appear due to limited spatial resolution

Solution?

Use an Optical Low Pass Filter (OLPF) -Anti-Aliasing filter.-

The OLPF stop high-frequency image information reducing the effects caused by high-frequency waves in images.

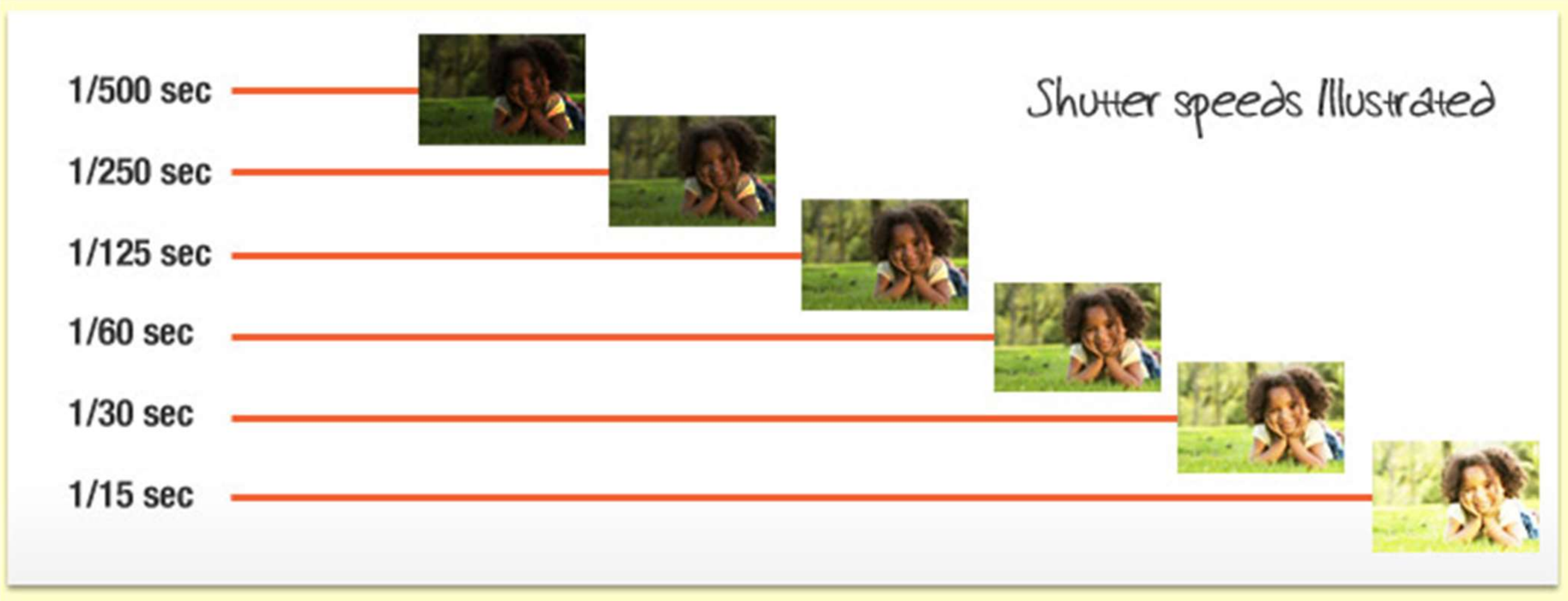


The aperture of the sensor

It is expressed as the ratio of the lens opening

It is used to control the amount of light reaching the image sensor.

The shutter speed



ww2.lifepics.com

It denotes the amount of time the lens is open to capture light

Exposure control

Exposure is the amount of light that is allowed to reach the sensor while capturing an image.

It depends in both the aperture and the shutter speed.

Exposure control

Exposure is the amount of light that is allowed to reach the sensor while capturing an image.

It depends in both the aperture and the shutter speed.

An exposure value (EV) denotes all combinations of relative aperture and shutter speed that provide the same exposure:

$$EV = \log_2 (F^2/t)$$

F is the f-number
t is the exposure time (in seconds)

Auto-Exposure

Most auto-exposure algorithms work as follows:

1) Take a (temporary, not to be recorded) picture with a predetermined exposure value EV_{pre}.
2) Compute a single brightness value B_{pre} from that picture.
3) Select an optimal brightness B_{opt}

Auto-Exposure

Most auto-exposure algorithms work as follows:

1) Take a (temporary, not to be recorded) picture with a predetermined exposure value EV_{pre}.
2) Compute a single brightness value B_{pre} from that picture.
3) Select an optimal brightness B_{opt}

Then, they compute the optimal exposure is:

$$EV_{opt} = EV_{pre} + \log_2(B_{pre} / B_{opt})$$

Auto-Exposure

Most auto-exposure algorithms work as follows:

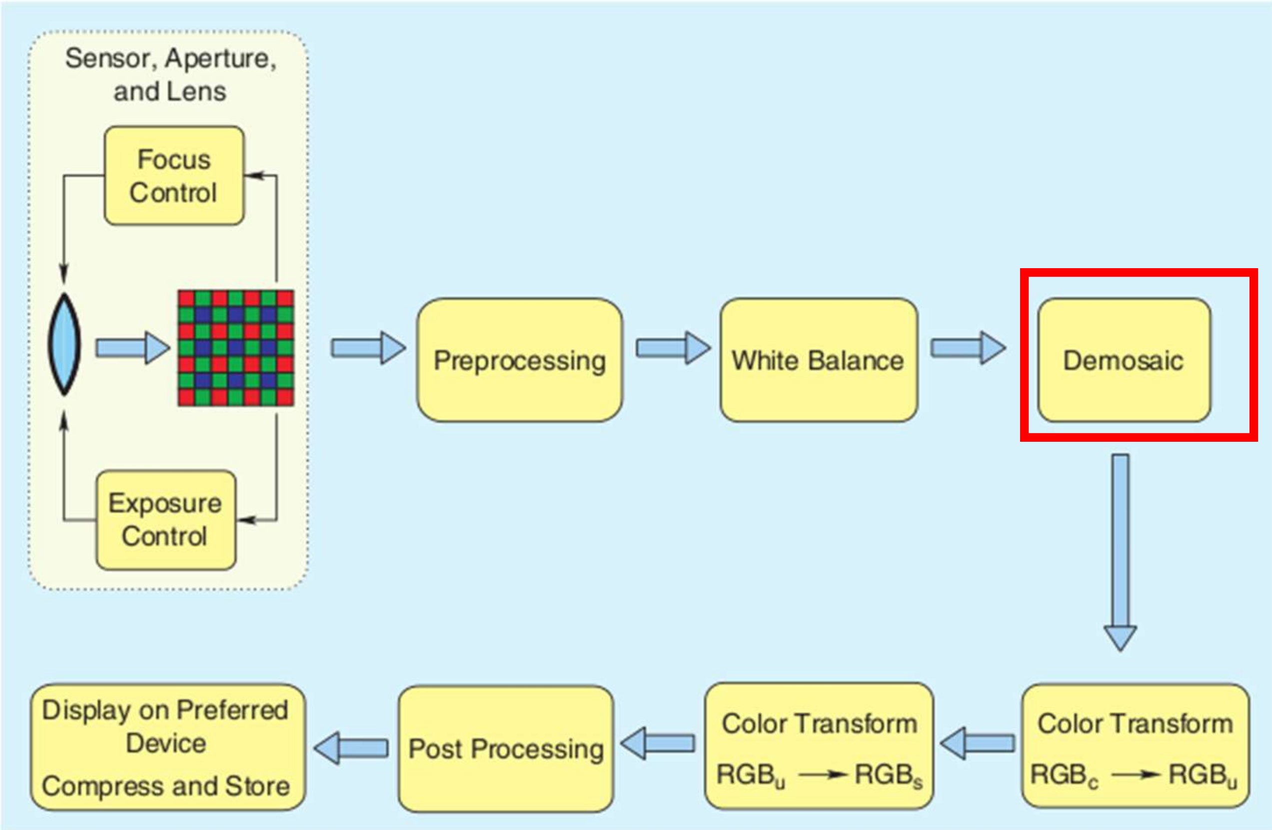
1) Take a (temporary, not to be recorded) picture with a predetermined exposure value EV_{pre}.
2) Compute a single brightness value B_{pre} from that picture.
3) Select an optimal brightness B_{opt}

Then, they compute the optimal exposure is:

$$EVopt = EV_{pre} + \log_2(B_{pre} / B_{opt})$$

Finally, they modify the aperture and shutter speed (and gain) to take the exposure-corrected image.

The camera color imaging pipeline



Ramanath et al. 2005

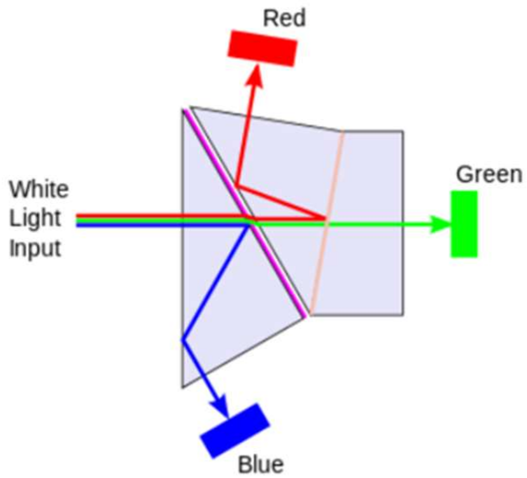
Why demosaicing?

To convert the light reaching the sensor into colors there are two posibilities:

Why demosaicing?

To convert the light reaching the sensor into colors there are two posibilities:

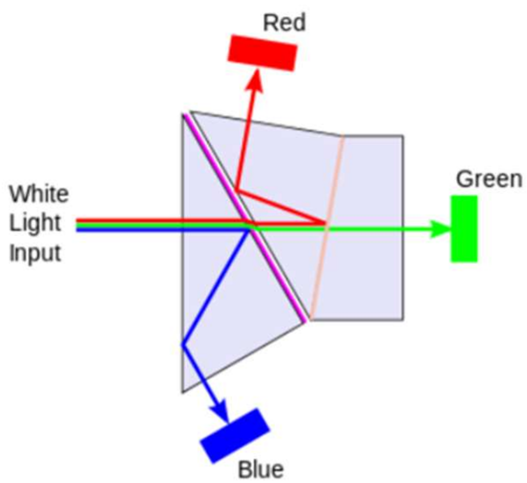
a) A three-sensor systems. Incoming light is separated into three different color channels using a beam splitter (a type of prism) and there is one sensor devoted to each.



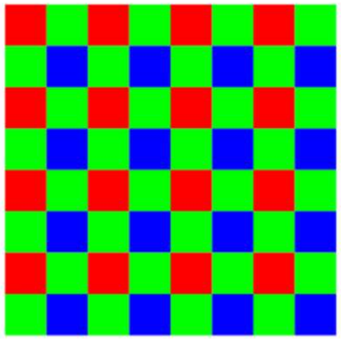
Why demosaicing?

To convert the light reaching the sensor into colors there are two posibilities:

a) A three-sensor systems. Incoming light is separated into three different color channels using a beam splitter (a type of prism) and there is one sensor devoted to each.



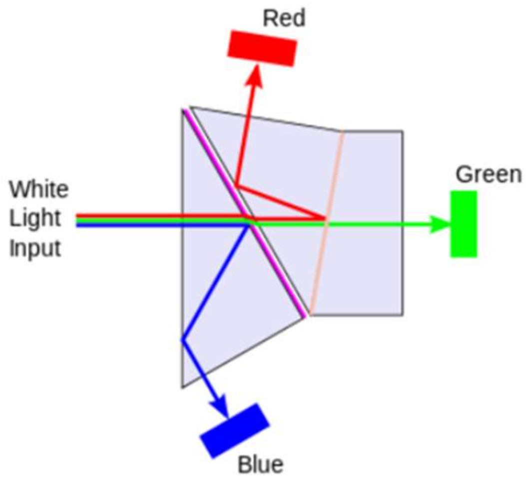
b) A Color Filter Array (CFA). In this one-sensor approach, the image array is covered by a mosaic color filter of three colors, making each pixel in the array capture one color channel only.



Why demosaicing?

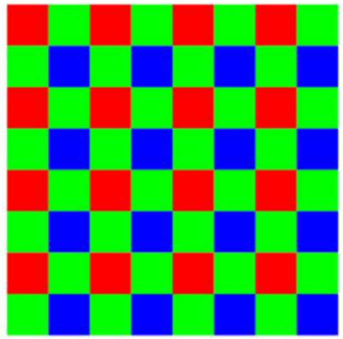
To convert the light reaching the sensor into colors there are two posibilities:

a) A three-sensor systems. Incoming light is separated into three different color channels using a beam splitter (a type of prism) and there is one sensor devoted to each.



Most common approach:

b) A Color Filter Array (CFA). In this one-sensor approach, the image array is covered by a mosaic color filter of three colors, making each pixel in the array capture one color channel only.



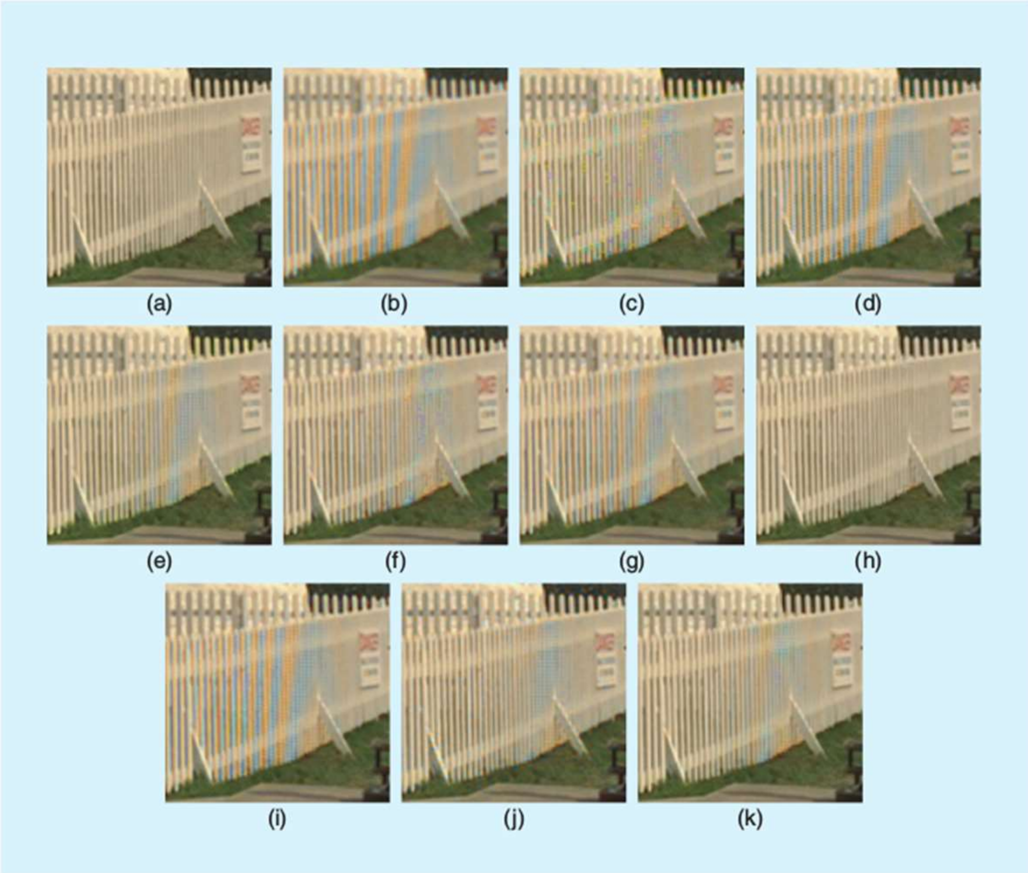
Mosaicked image

The goal of demosaicking is to obtain for each pixel the three colour values: R, G, and B.

Demosaicking is just an interpolation problem.

Main demosaicking problem

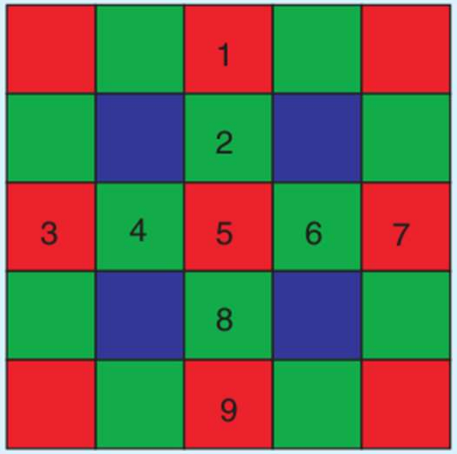
Artifacts in the edges (specially on diagonal ones).



An example of a demosaicing algorithm (Gunturk et al 2005)

First: interpolate the green channel looking at the existence of edges.

(Interpolate first the green channel it is a standard procedure)



1. Calculate horizontal gradient $\Delta H = |(R3 + R7)/2 - R5|$
2. Calculate vertical gradient $\Delta V = |(R1 + R9)/2 - R5|$
3. If $\Delta H > \Delta V$,

   $G5 = (G2 + G8)/2$

   Else if $\Delta H < \Delta V$,

   $G5 = (G4 + G6)/2$

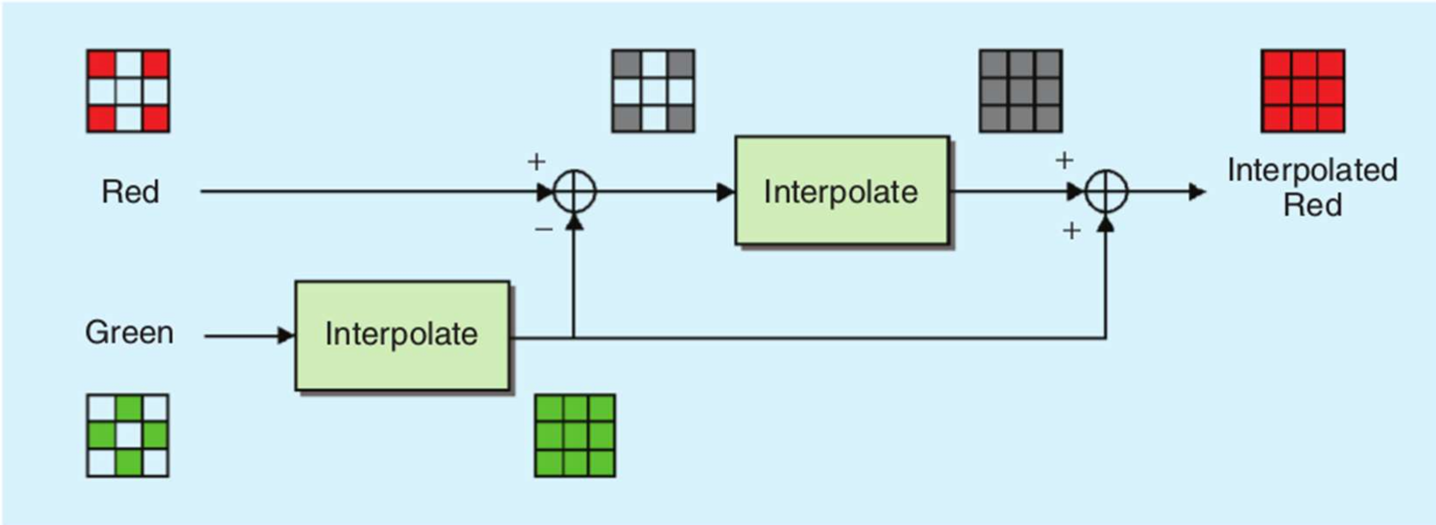
   Else

   $G5 = (G2 + G8 + G4 + G6)/4$

[FIG4] Edge-directed interpolation in [23] is illustrated for estimating the *G* value at pixel 5. The R values are used to determine the edge direction. When the missing *G* value is at a B pixel, the B values are used to determine the edge direction.

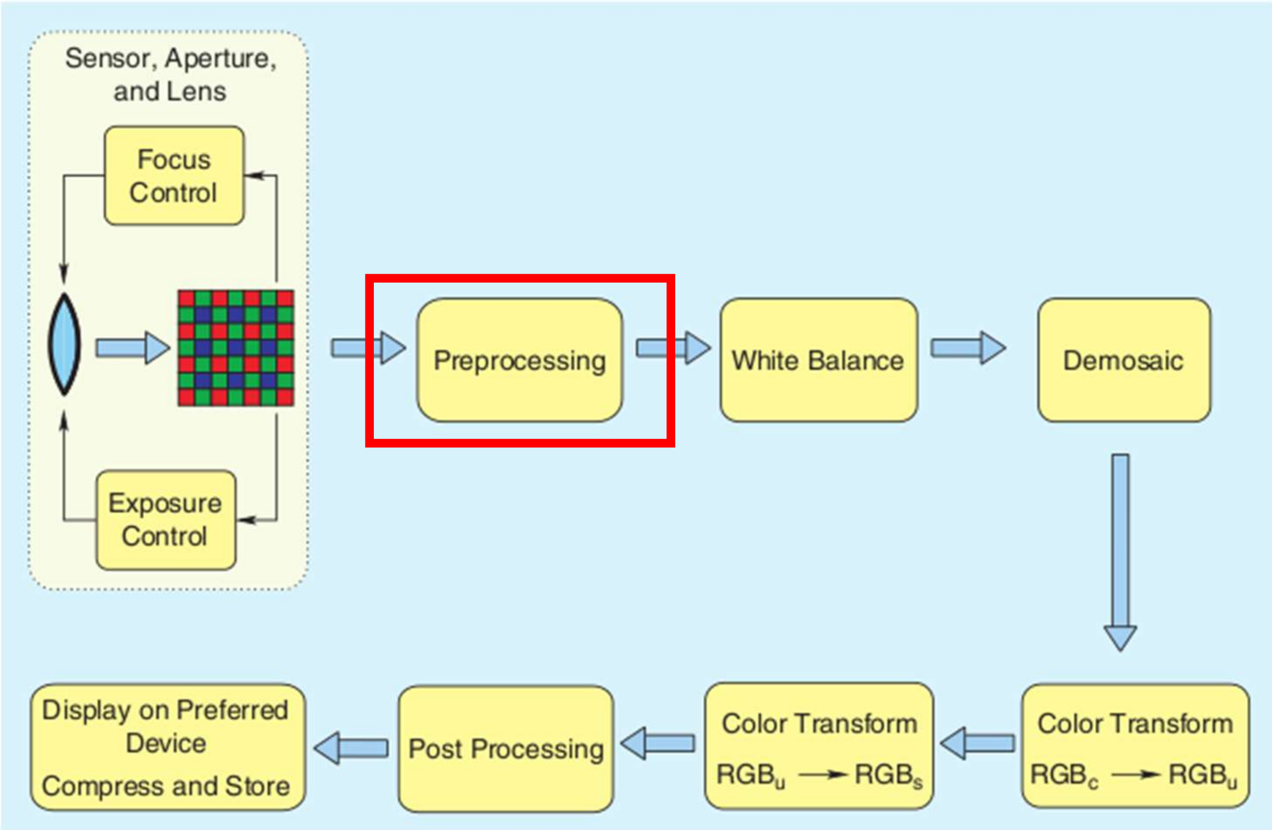
An example of a demosaicing algorithm (Gunturk et al 2005)

Second: Use the information from the green interpolation to help the blue and red ones.



[FIG5] Constant-difference-based interpolation is illustrated for the R channel. The B channel is interpolated similarly.

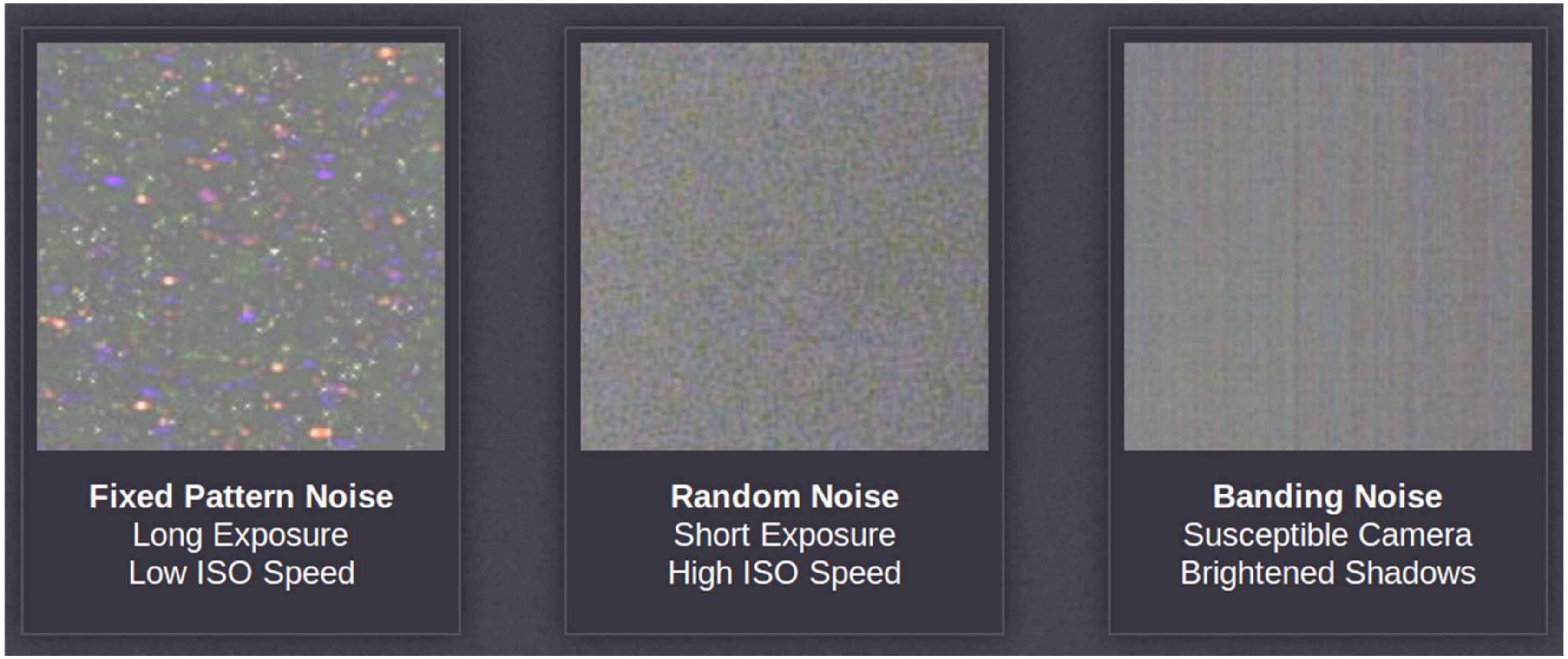
The camera color imaging pipeline



Ramanath et al. 2005

Denoising

Depending on the capture conditions, the obtained image can have different types of noise.



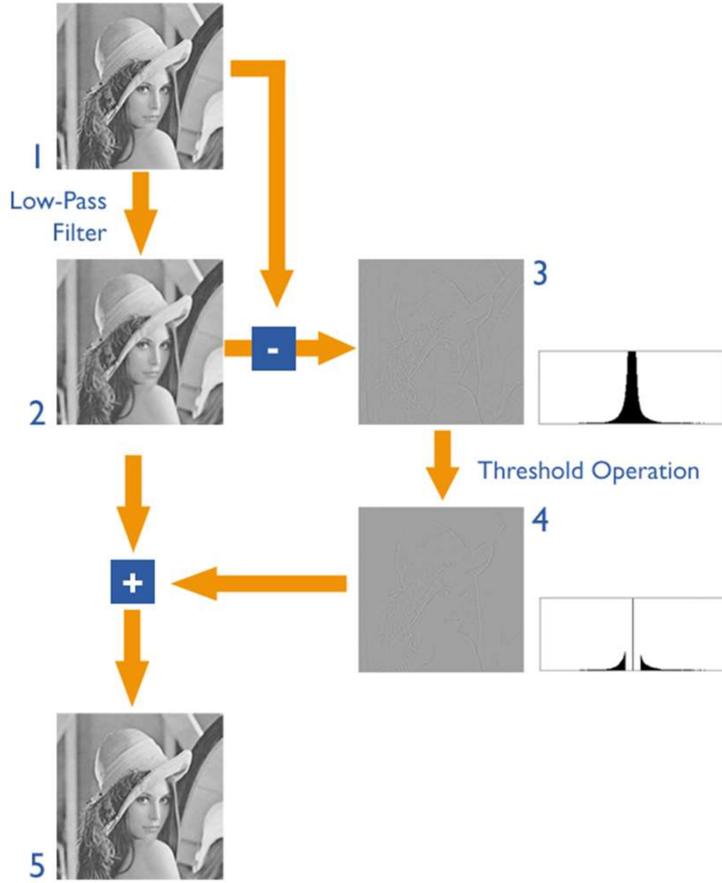
http://www.cambridgeincolour.com

A very simple denoising technique

1) Split the images into a low-pass and a high-pass one.

2) Denoise the high-pass image by thresholding

3) Joing back the low-pass and high-pass images.



http://www.image-engineering.de

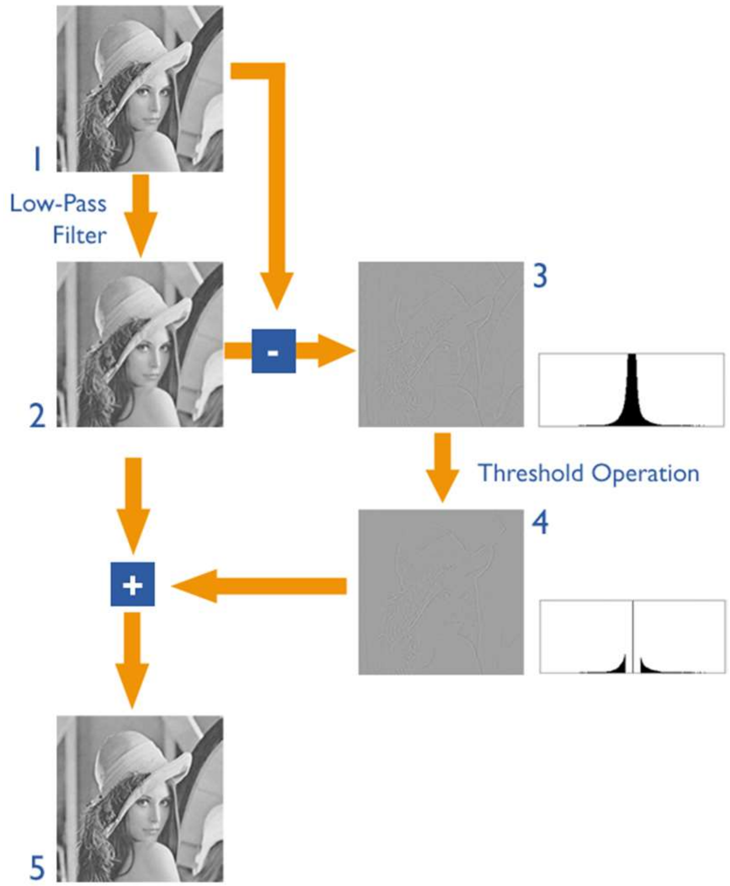
A very simple denoising technique

1) Split the images into a low-pass and a high-pass one.

2) Denoise the high-pass image by thresholding

3) Joing back the low-pass and high-pass images.

This technique is related to the behaviour of our Human Visual System (HVS), as we are less sensitive to changes in higher frequencies (Weber's law).



http://www.image-engineering.de

A very simple denoising technique

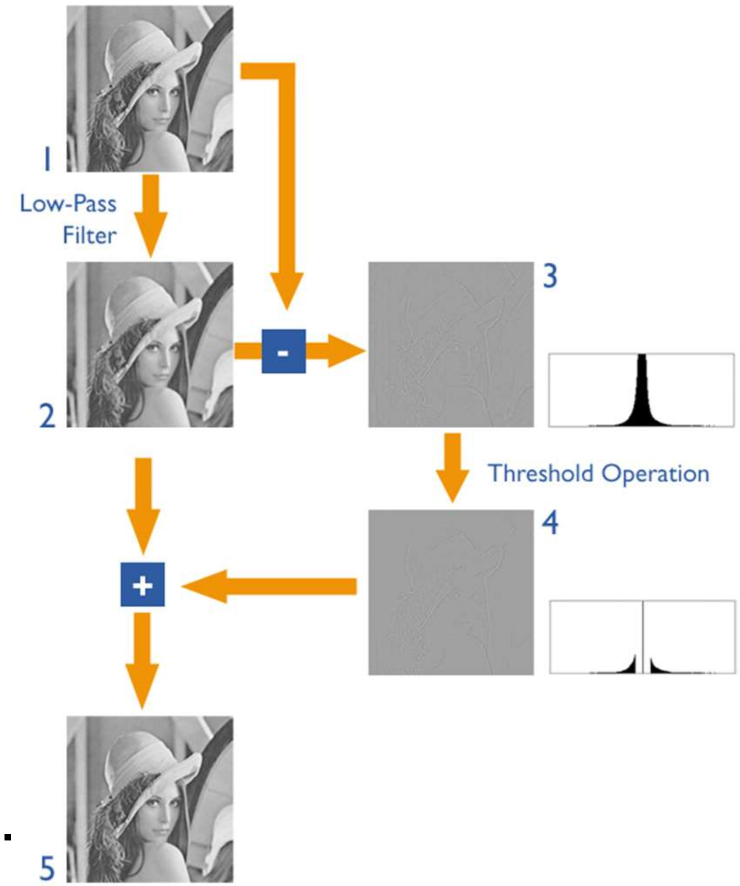
1) Split the images into a low-pass and a high-pass one.

2) Denoise the high-pass image by thresholding

3) Joing back the low-pass and high-pass images.

This technique is related to the behaviour of our Human Visual System (HVS), as we are less sensitive to changes in higher frequencies (Weber's law).

Very similar technique to what is used for Image Compression.



http://www.image-engineering.de

Denoising methods

There exist many denoising methods currently. A non-exhaustive list of breakthroughs can be:

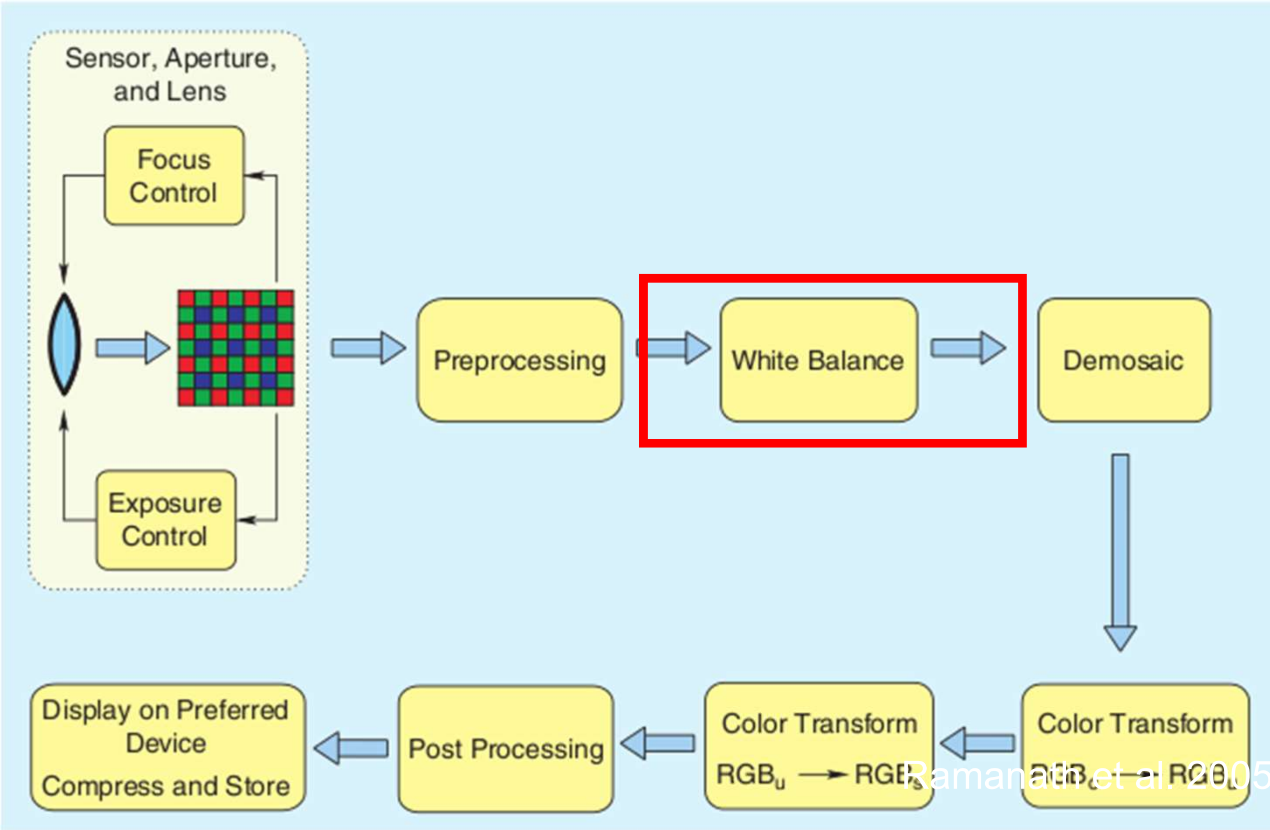
Variational methods:
ROF (Rudin-Osher-Fatemi , 1992)

Non-local methods:
Non-local means (Buades et al, 2005)

Non-local + frequency domain:
BM3D (Foi et al, 2007)

Deep-learning methods:
DnCNN -Residual Learning for Denoising- (Zhang et al, 2017)

The camera color imaging pipeline



Ramanath et al. 2005

White balance

The goal of White Balance is to emulate the human color constancy ability we explained on the previous class.

I.E. We want to record an image as if it was recorded under a canonical –usually White- illuminant.

White balance

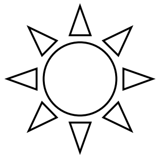
The goal of White Balance is to emulate the human color constancy ability we explained on the previous class.

I.E. We want to record an image as if it was recorded under a canonical –usually White- illuminant.

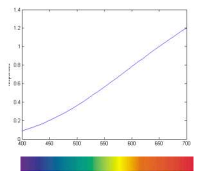
Colour formation

Colour is a function of three physical properties

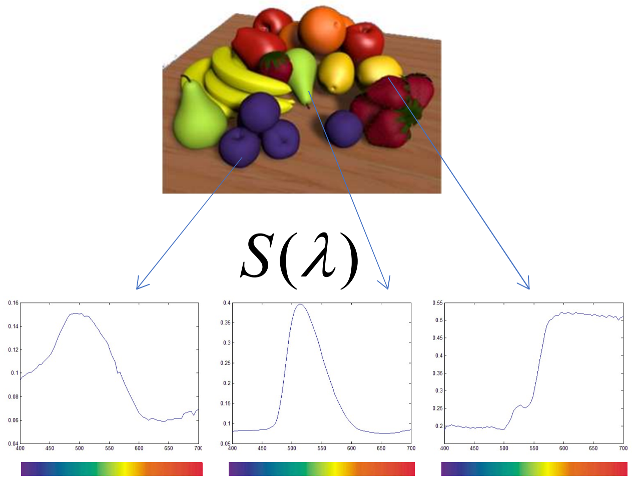
Light source



$$E(\lambda)$$



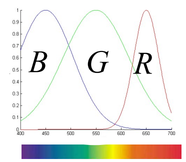
Object reflectance



$$S(\lambda)$$

Camera sensitivities

$$R_i(\lambda)$$

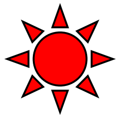


$$I_{R,G,B} = \int_{\omega} R_i(\lambda) E(\lambda) S(\lambda) d\lambda$$

White balance

Modelling illuminant change

Common assumption: Image formation is linear, then

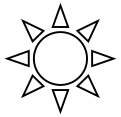


$E_2(\lambda)$

$$I^2_{R,G,B} = \int_{\omega} R_i(\lambda) E_2(\lambda) S(\lambda) d\lambda$$

3 by N matrix

N: Number of image pixels



$E_1(\lambda)$

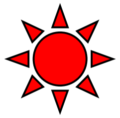
Canonical Image

$$I^1_{R,G,B} = \int_{\omega} R_i(\lambda) E_1(\lambda) S(\lambda) d\lambda$$

White balance

Modelling illuminant change

Common assumption: Image formation is linear, then



$E_2(\lambda)$

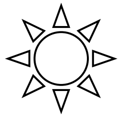
$$I^2_{R,G,B} = \int_\omega R_i(\lambda) E_2(\lambda) S(\lambda)\, d\lambda$$

3 by N matrix
N: Number of image pixels

Illuminant change model

$$I^1_{R,G,B} = \begin{pmatrix} \alpha & \delta & \varepsilon \\ \mu & \beta & \eta \\ \varsigma & \tau & \gamma \end{pmatrix} \cdot I^2_{R,G,B}$$

Canonical Image



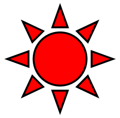
$E_1(\lambda)$

$$I^1_{R,G,B} = \int_\omega R_i(\lambda) E_1(\lambda) S(\lambda)\, d\lambda$$

White balance

Modelling illuminant change

Common assumption: Image formation is linear, then



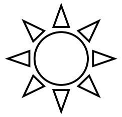
$E_2(\lambda)$

$$I^2_{R,G,B} = \int_\omega R_i(\lambda) E_2(\lambda) S(\lambda)\, d\lambda$$

3 by N matrix
N: Number of image pixels

Illuminant change model

$$I^1_{R,G,B} = \begin{pmatrix} \alpha & \delta & \varepsilon \\ \mu & \beta & \eta \\ \varsigma & \tau & \gamma \end{pmatrix} \cdot I^2_{R,G,B}$$



$E_1(\lambda)$

Canonical Image

$$I^1_{R,G,B} = \int_\omega R_i(\lambda) E_1(\lambda) S(\lambda)\, d\lambda$$

Solving computational colour constancy: Finding the illuminant change matrix *(9 unknowns)* -> ill-posed problem (multiple solutions).

White balance

The problema is usually simplified following the Von Kries law (see last class)

$$R' = k_r R, \quad G' = k_g G, \quad B' = k_b B,$$

Therefore, we just have 3 gains: One for each color channel

$$\begin{pmatrix} \alpha & \delta & \varepsilon \\ \mu & \beta & \eta \\ \varsigma & \tau & \gamma \end{pmatrix} \longrightarrow \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$$

Simple methods

Grey-world (Bunchsbaum 1980): Assumes that the average scene is grey

$$R' = \alpha R$$
$$G' = \beta G$$
$$B' = \gamma B$$

Where:

$$\alpha = 127/R_{average}$$
$$\beta = 127/G_{average}$$
$$\gamma = 127/B_{average}$$

Gray World

Gray World

Simple methods

Max-RGB (Bunchsbaum 1980): Assumes there is either a White object, or red, green, and blue objects

$$R' = \alpha R$$
$$G' = \beta G$$
$$B' = \gamma B$$

Where:

$$\alpha = 255/R_{max}$$
$$\beta = 255/G_{max}$$
$$\gamma = 255/B_{max}$$

White Patch

White Patch

Other methods

Statistical methods
- Shades of Grey (Finlayson and Trezzi, 2005)
- Grey-Edge (Van de Weijer et al, 2007)
- Bag of Pixels (Chakrabarty et al, 2008)
- Corrected moments (Finlayson, 2015)

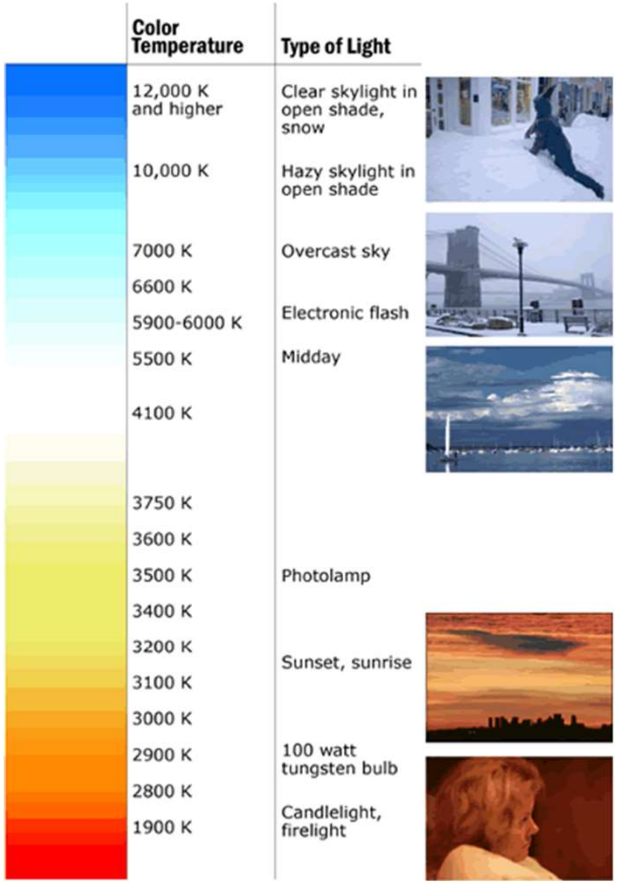
Voting methods
- Color by Correlarion (Finlayson et al, 2001)
- Category hypothesis (Vazquez-Corral et al, 2012)

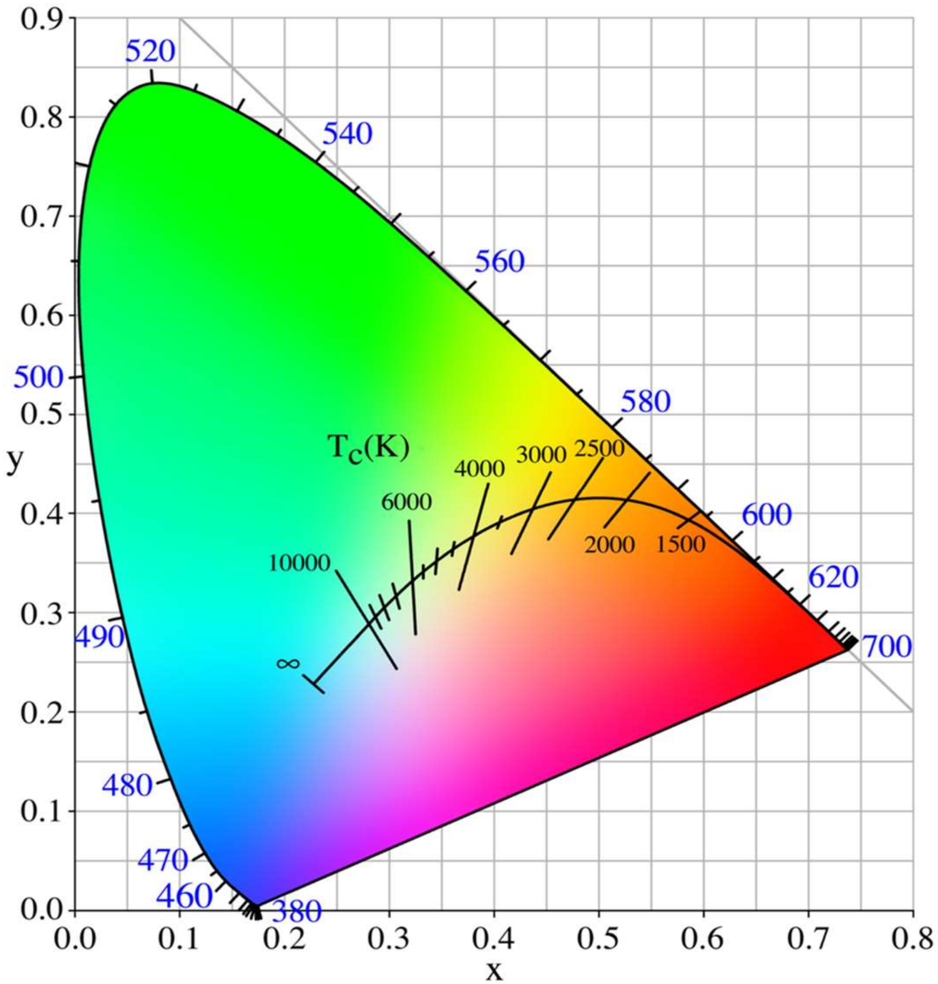
Physics-based methods:
- Specularities (Barnard and Funt, 1999)

Learning methods:
- FFCC (Barron and Tsai, 2017)

Deep learning methods (current state of the art):
- FC4 (Hu et al, 2017)

Planckian illuminants





Planckian illuminants

WB Settings	Color Temperature	Light Sources
	10000 - 15000 K	Clear Blue Sky
	6500 - 8000 K	Cloudy Sky / Shade
	6000 - 7000 K	Noon Sunlight
	5500 - 6500 K	Average Daylight
	5000 - 5500 K	Electronic Flash
	4000 - 5000 K	Fluorescent Light
	3000 - 4000 K	Early AM / Late PM
	2500 - 3000 K	Domestic Lightning
	1000 - 2000 K	Candle Flame

Manual illuminant selection

Tungsten (2850K)

Manual illuminant selection

Fluorescent (3800K)

Manual illuminant selection

Daylight/Flash
(5500K)

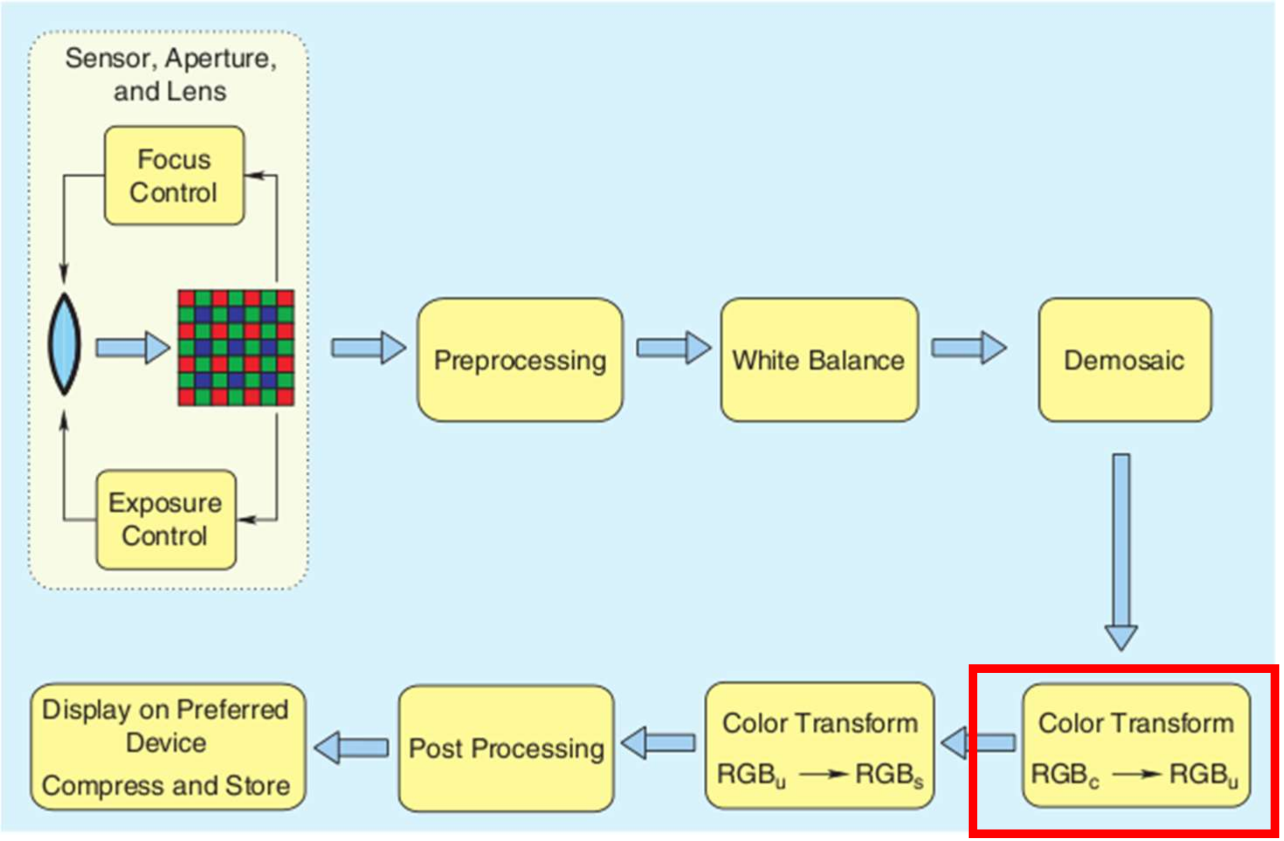
Manual illuminant selection

Overcast (6500K)

Manual illuminant selection

Shade (7500K)

The camera color imaging pipeline



Ramanath et al. 2005

Colour transform to a canonical space

The goal is to convert the camera RGBs values –which are inherently dependent on the sensors- to a canonical colour space that does not depend on the camera.

The typical canonical colour space is XYZ (or sRGB).

Colour transform to a canonical space

The goal is to convert the camera RGBs values –which are inherently dependent on the sensors- to a canonical colour space that does not depend on the camera.

The typical canonical colour space is XYZ (or sRGB).

This transformation is learnt offline by the camera manufacturer. It looks at the best linear 3 x 3 matrix that either:

1) Convert the camera sensors to the XYZ color matching functions

$$\begin{bmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} r(\lambda) \\ g(\lambda) \\ b(\lambda) \end{bmatrix}$$

Colour transform to a canonical space

The goal is to convert the camera RGBs values –which are inherently dependent on the sensors- to a canonical colour space that does not depend on the camera.

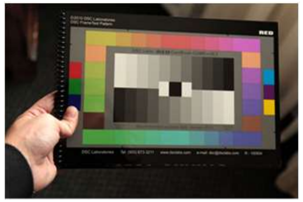
The typical canonical colour space is XYZ (or sRGB).

This transformation is learnt offline by the camera manufacturer. It looks at the best linear 3 x 3 matrix that either:

1) Convert the camera sensors to the XYZ color matching functions

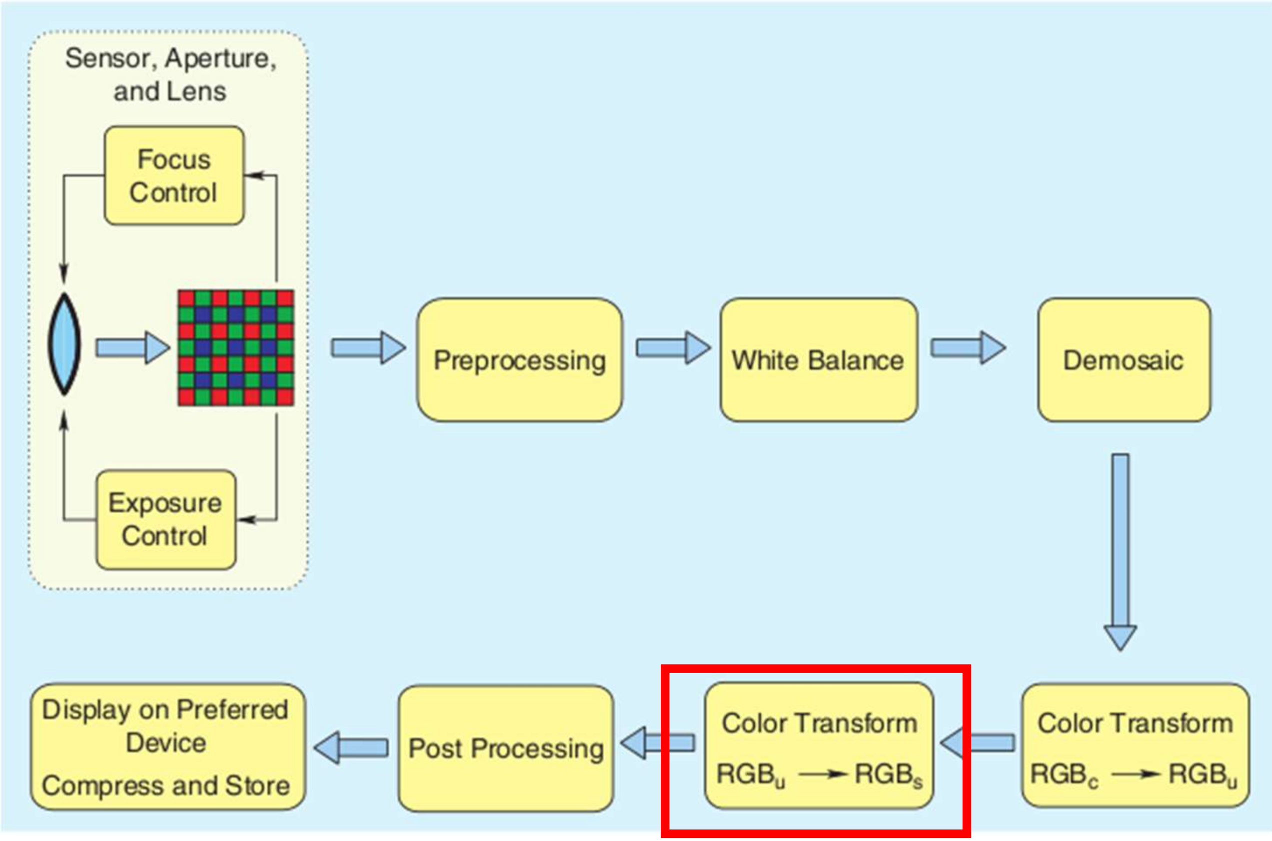
$$\begin{bmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} r(\lambda) \\ g(\lambda) \\ b(\lambda) \end{bmatrix}$$

2) Convert some known colour patches.



$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = A \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The camera color imaging pipeline



Ramanath et al. 2005

Colour transform to the output space

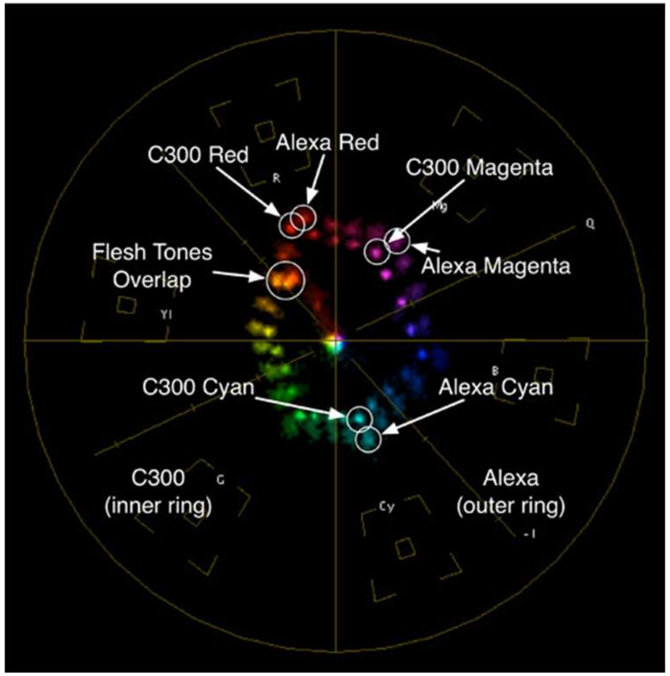
At this stage, each camera manufacturer gives their trade look to the camera (for example, Canon images are usually more saturated than Nikon’s one)

It it usually also a 3x3 linear matrix multiplication, that is embedded with the previous one in a single multiplication.

Colour transform to the output space

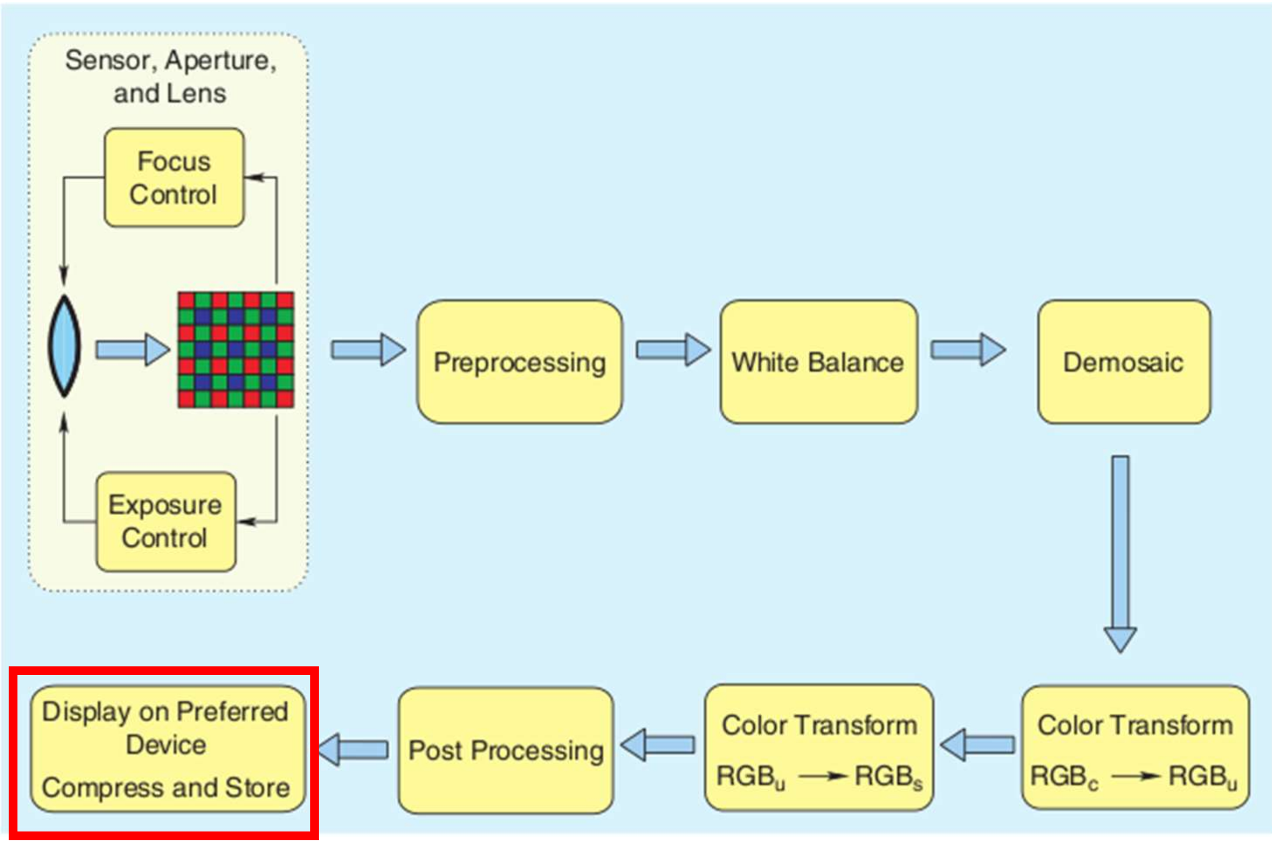
At this stage, each camera manufacturer gives their trade look to the camera (for example, Canon images are usually more saturated than Nikon’s one).

It it usually also a 3x3 linear matrix multiplication, that is embedded with the previous one in a single multiplication.



Comparison between the Canon C-300 and the Alexa colours.

The camera color imaging pipeline



Ramanath et al. 2005

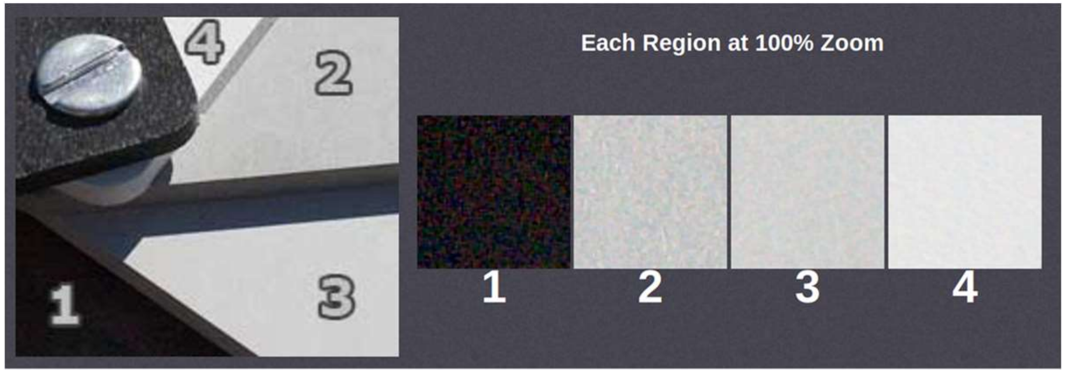
Why gamma correction?

Originally the CRT monitors had a non-linear relation between the voltage applied and the intensity shown.

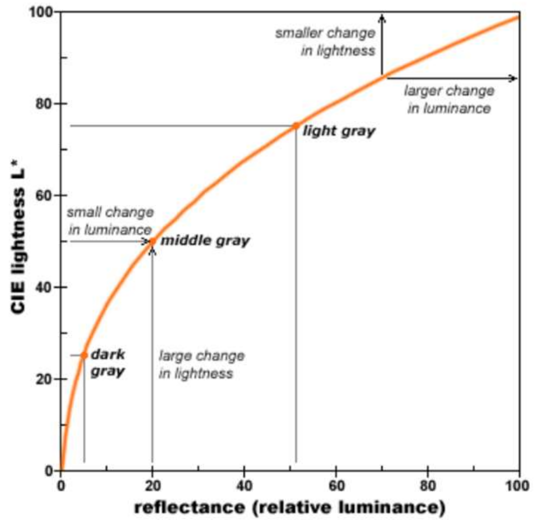
Why gamma correction?

Originally the CRT monitors had a non-linear relation between the voltage applied and the intensity shown.

But also, our perception of intensity is not linear (Weber's law).



http://www.cambridgeincolour.com



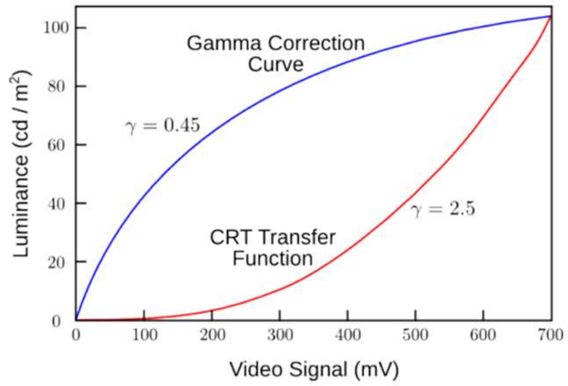
Why gamma correction?

Originally the CRT monitors had a non-linear relation between the voltage applied and the intensity shown.

But also, our perception of intensity is not linear (Weber's law).

Originally for both of these reasons –compensating the display non-linearity and improving our perception of it-, and currently just for the second reason, the camera applies a non-linear transformation to the output image in the form of a power-law

$$I_{out}=I_{in}^{\gamma}$$



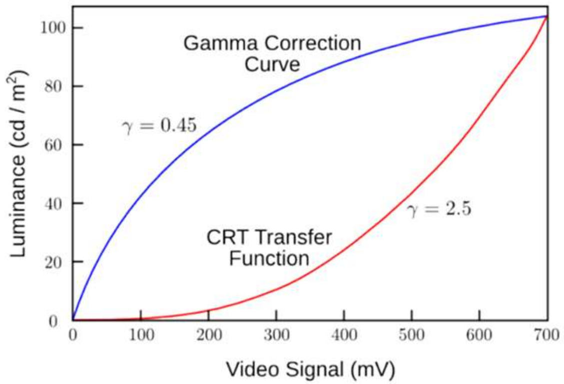
Why gamma correction?

Originally the CRT monitors had a non-linear relation between the voltage applied and the intensity shown.

But also, our perception of intensity is not linear (Weber's law).

Originally for both of these reasons –compensating the display non-linearity and improving our perception of it-, and currently just for the second reason, the camera applies a non-linear transformation to the output image in the form of a power-law

$$I_{out}=I_{in}^{\gamma}$$



There are not perfectly inverse to improve the aesthetics of the image

Logarithmic correction

For High-Dynamic range images, gamma-correction has shown to not be good enough –in HDR images it may wash-out important details on the bright areas, and create artifacts in them-.

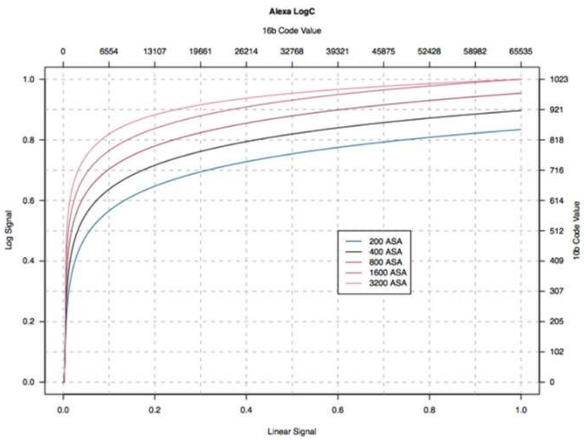
Logarithmic correction

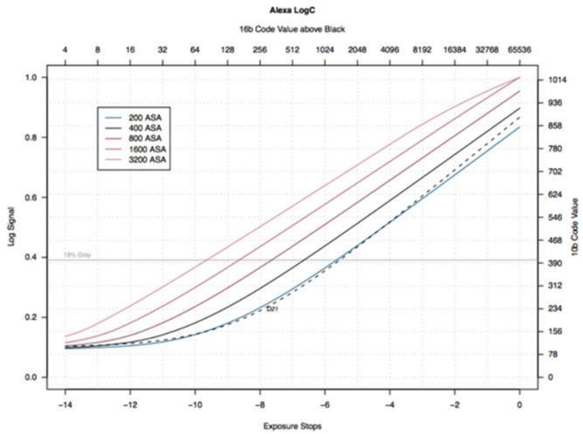
For High-Dynamic range images, gamma-correction has shown to not be good enough –in HDR images it may wash-out important details on the bright areas, and create artifacts in them-.

For HDR image logarithmic correction is used:

$$I_{out}=c\cdot\log(a*I_{in}+b)+d$$

Where a,b,c,d are parameters





Gamma and log correction (example)

Gamma-encoded

Log-encoded
(the monitor is not aimed at displaying it)

Image compression

Cameras need to write the images into output files, which are delimited by the number of bits that can be written for each pixel

Image compression

Cameras need to write the images into output files, which are delimited by the number of bits that can be written for each pixel

Usually, cameras use 12 or 14 bits (which means integer values in the ranges 0-4048 or 0-16192) for the RAW files –which are the unprocessed reads from the sensor-.

Image compression

Cameras need to write the images into output files, which are delimited by the number of bits that can be written for each pixel

Usually, cameras use 12 or 14 bits (which means integer values in the ranges 0-4048 or 0-16192) for the RAW files –which are the unprocessed reads from the sensor-.

However, the output after all the camera imaging pipeline is usually a 8-bit image (integers in the range 0-255).

Image compression

Cameras need to write the images into output files, which are delimited by the number of bits that can be written for each pixel

Usually, cameras use 12 or 14 bits (which means integer values in the ranges 0-4048 or 0-16192) for the RAW files –which are the unprocessed reads from the sensor-.

However, the output after all the camera imaging pipeline is usually a 8-bit image (integers in the range 0-255).

Compressing an image means to go from the 12 or 14 bits image to the 8 bits one.

Image compression: how to perform it

The main idea in image compression is to compress more the information to which we, as humans, are less sensitive to.

Image compression: how to perform it

The main idea in image compression is to compress more the information to which we, as humans, are less sensitive to.

In particular:

1) We are more sensitive to difference in lightness more than to differences in chroma (so we can downsample the chroma part of the image).

Image compression: how to perform it

The main idea in image compression is to compress more the information to which we, as humans, are less sensitive to.

In particular:

1) We are more sensitive to difference in lightness more than to differences in chroma (so we can downsample the chroma part of the image).

2) We are more sensitive to the low frequencies than to the high frequencies.

Image compression: how to perform it

The main idea in image compression is to compress more the information to which we, as humans, are less sensitive to.

In particular:

1) We are more sensitive to difference in lightness more than to differences in chroma (so we can downsample the chroma part of the image).

2) We are more sensitive to the low frequencies than to the high frequencies.

From 2, we can get the idea of decomposing the image into different frequencies → Make use of the DCT decomposition!

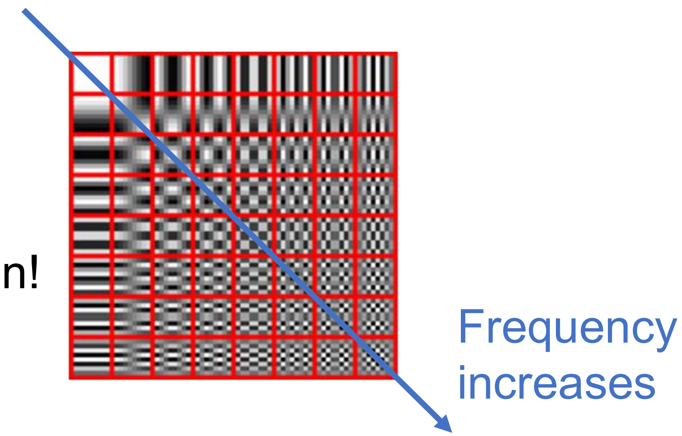
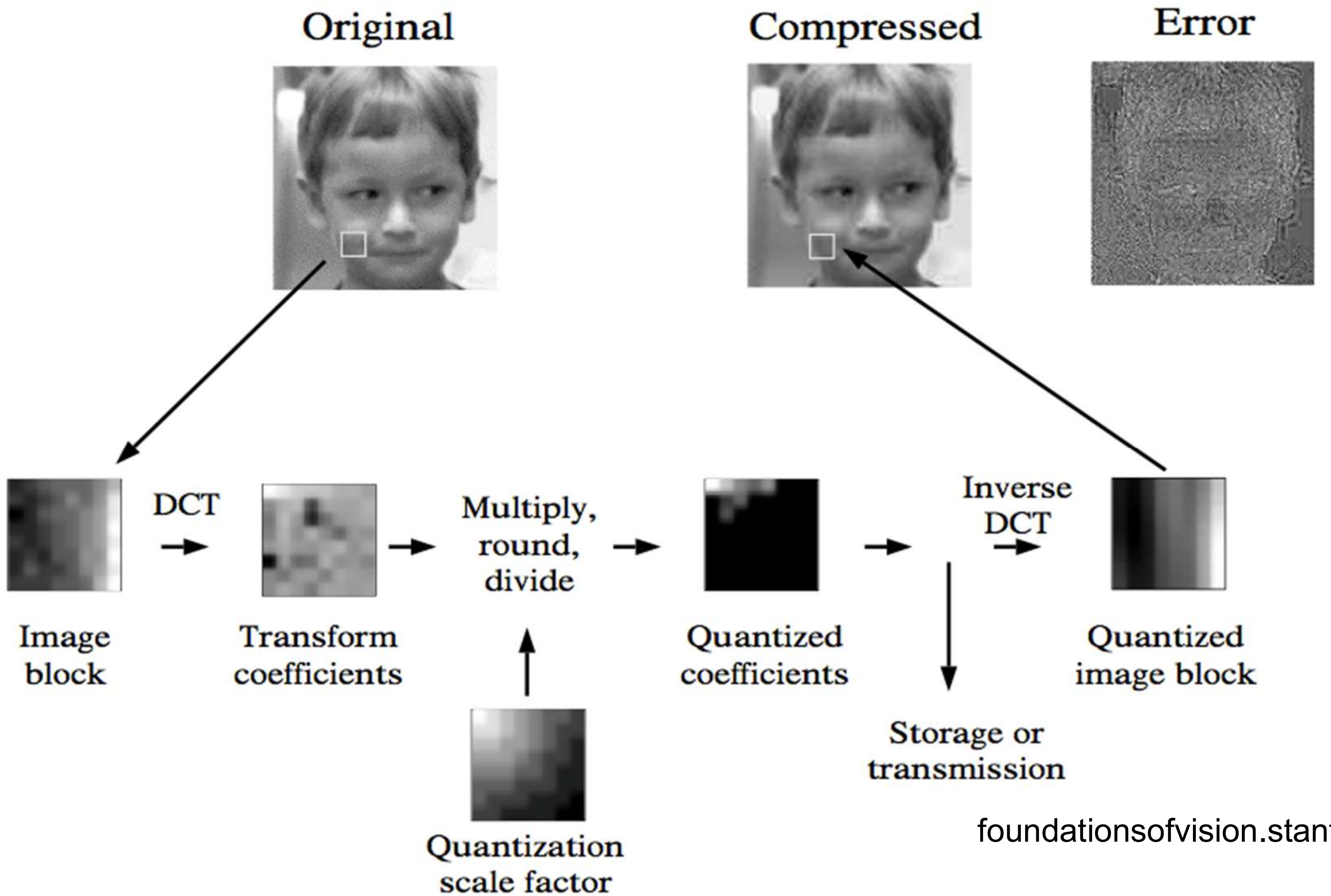


Image compression (the general pipeline)



foundationsofvision.stanford.edu

Image compression (example results)

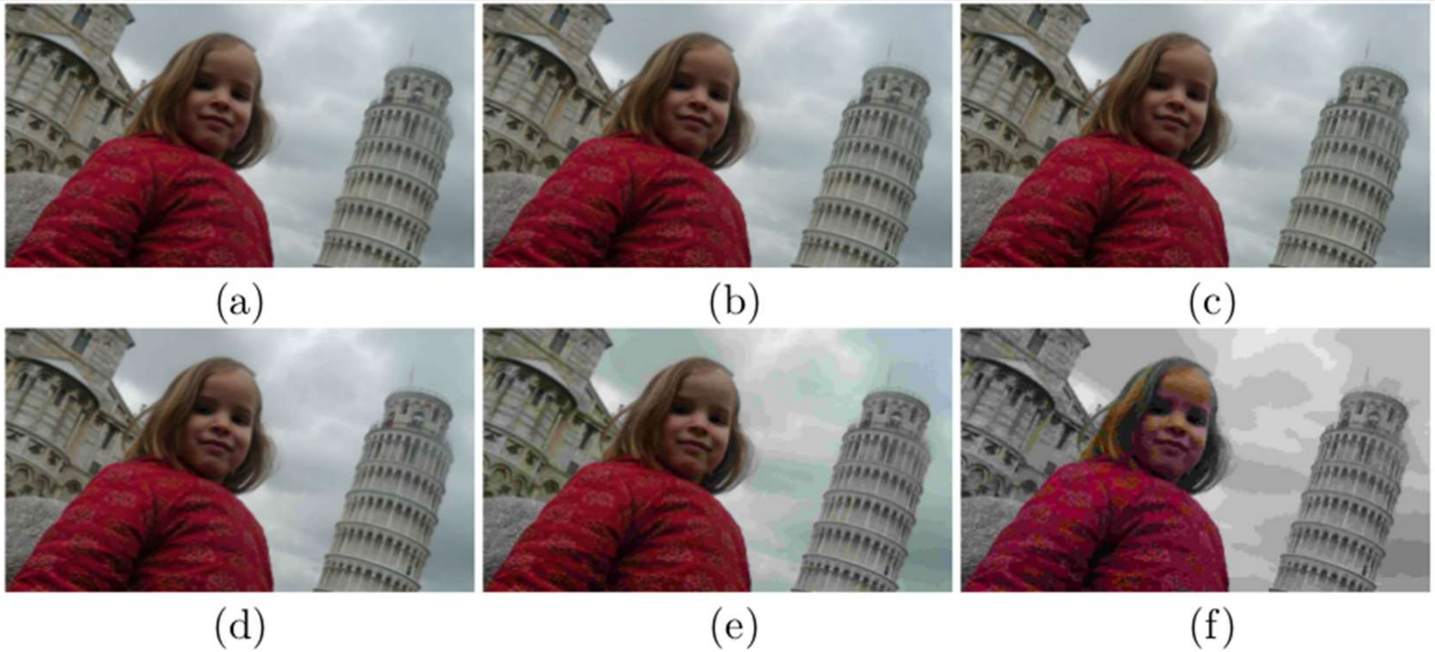
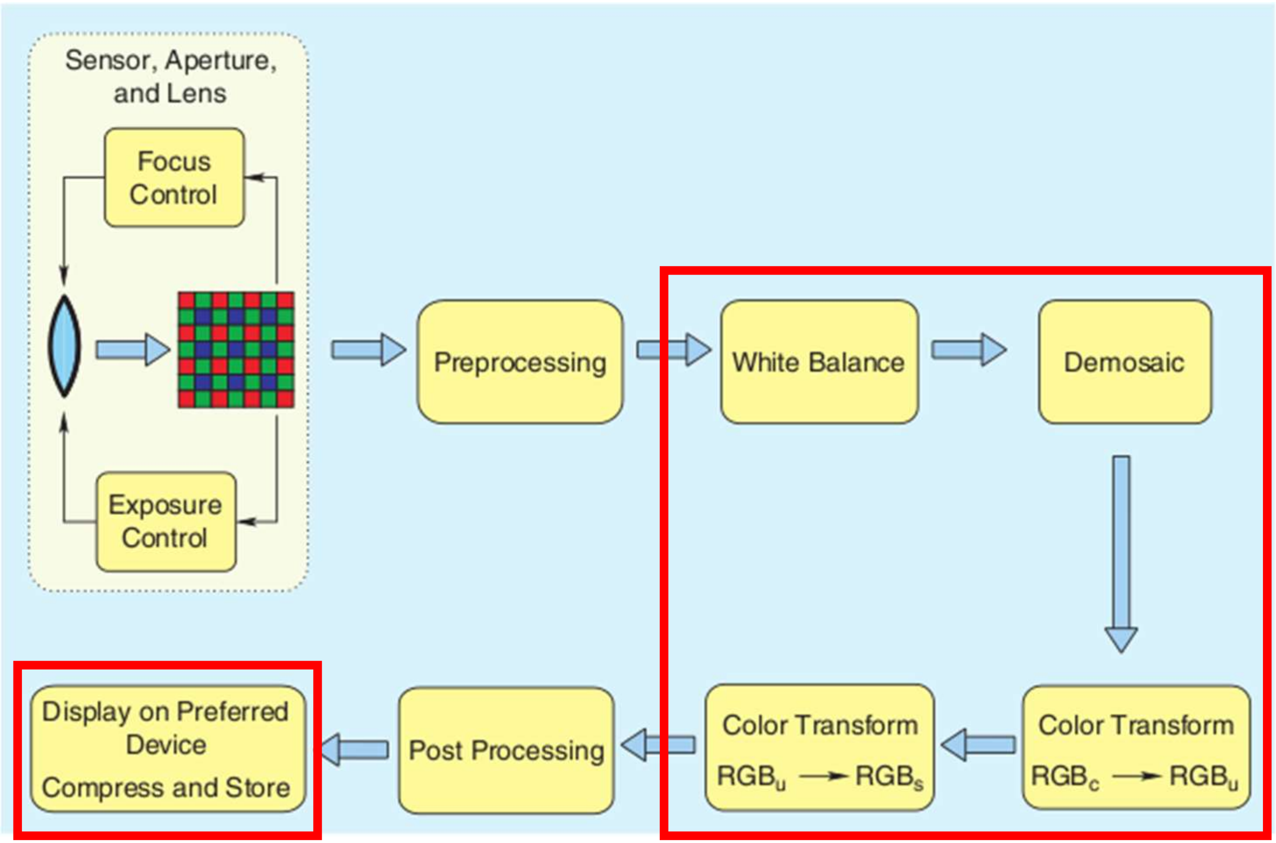


FIGURE 4.5: Different JPEG outputs varying the quality of the compression. The uncompressed original has a size of $3.9MB$. (a) 95% quality, $507KB$ size. (b) 75%, $138KB$. (c) 50%, $90KB$. (d) 25%, $58KB$ size. (e) 10%, $33KB$. (f) 5%, $22KB$.

The camera color imaging pipeline



Ramanath et al. 2005

The color imaging pipeline from input to output

We can approximate the color image pipeline with the next formula:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{out} = \left(\alpha \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \underbrace{\boxed{\begin{bmatrix} r_{AWB} & 0 & 0 \\ 0 & g_{AWB} & 0 \\ 0 & 0 & b_{AWB} \end{bmatrix}}}_{\text{White balance}} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{in} \right)^{\gamma}$$

The color imaging pipeline from input to output

We can approximate the color image pipeline with the next formula:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{out} = \left(\alpha \underbrace{\boxed{\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}}}_{\text{Color correction}} \begin{bmatrix} r_{AWB} & 0 & 0 \\ 0 & g_{AWB} & 0 \\ 0 & 0 & b_{AWB} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{in} \right)^{\gamma}$$

The color imaging pipeline from input to output

We can approximate the color image pipeline with the next formula:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{out} = \left(\boxed{\alpha} \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} r_{AWB} & 0 & 0 \\ 0 & g_{AWB} & 0 \\ 0 & 0 & b_{AWB} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{in} \right)^{\gamma}$$

Gain

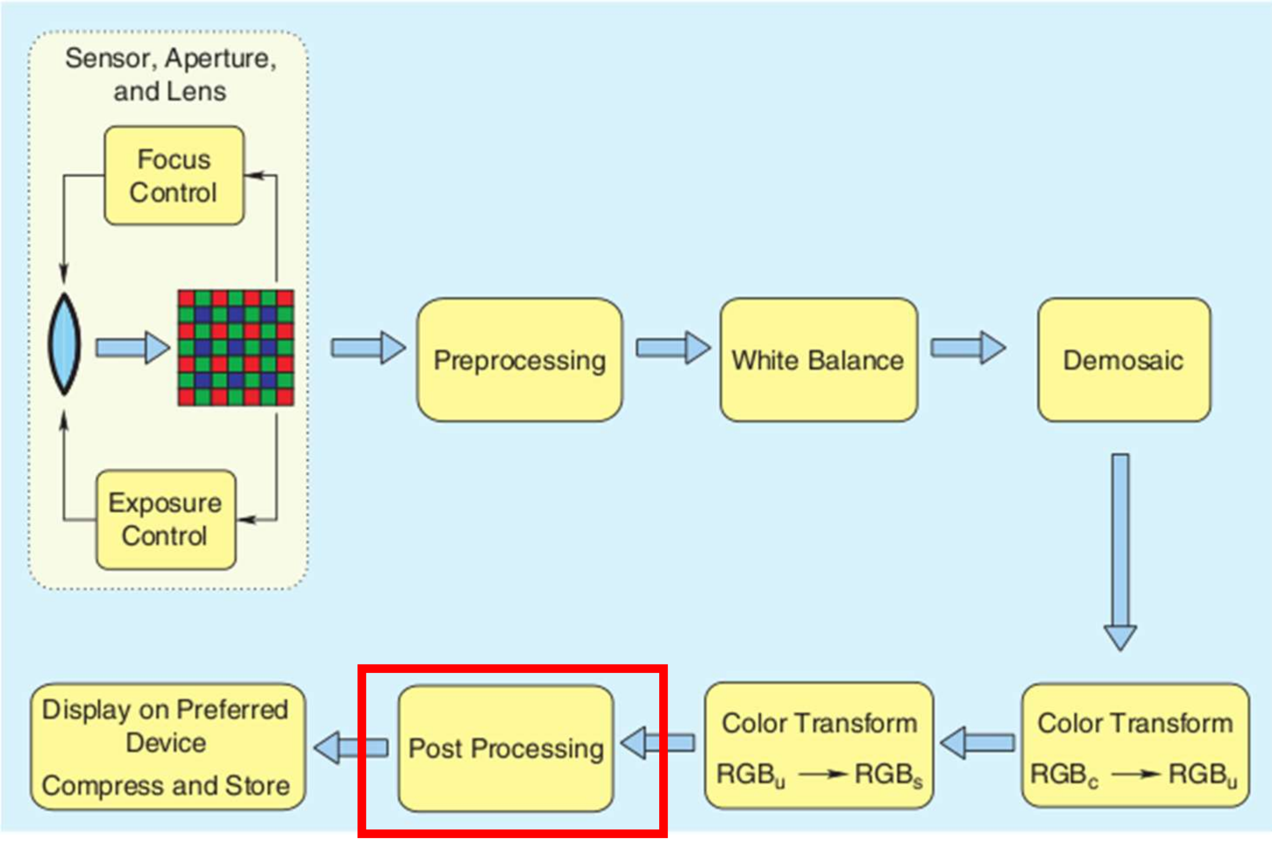
The color imaging pipeline from input to output

We can approximate the color image pipeline with the next formula:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{out} = \left(\alpha \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} r_{AWB} & 0 & 0 \\ 0 & g_{AWB} & 0 \\ 0 & 0 & b_{AWB} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{in} \right)^{\boxed{\gamma}}$$

Display non-linearity

The camera color imaging pipeline



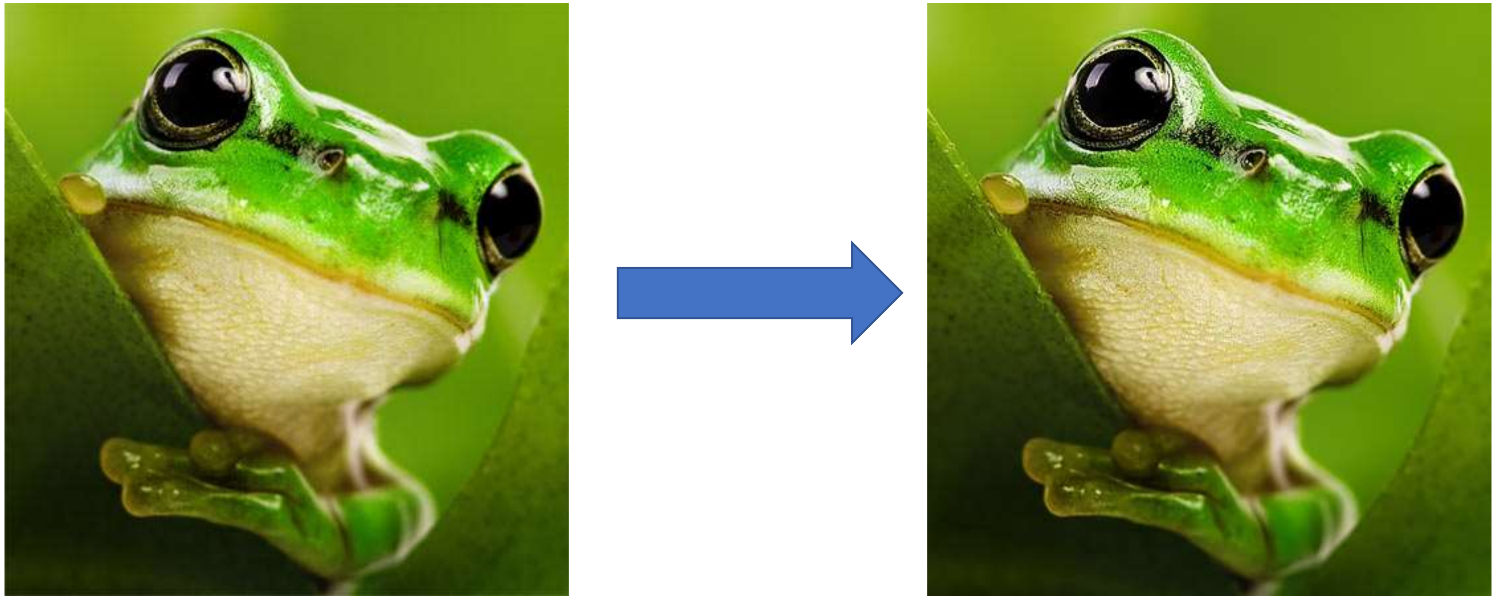
Ramanath et al. 2005

Edge enhancement

The lens system, the optimal aliasing filter and the sensor aperture produce image blurring.

This results in a reduction of the sharpness of the image.

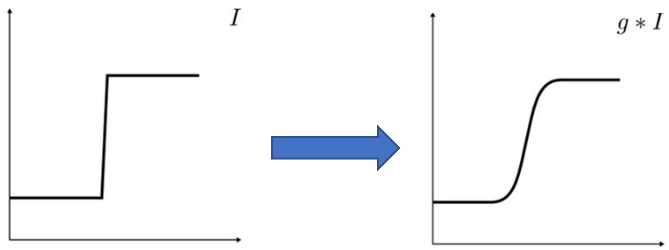
Therefore, camera makers aim at increasing the sharpening of the image.



Unsharp masking: a basic sharpening method

1) Blur the original image (I) by convolving with a Gaussian Kernel (g)

$$I_{blur}=I*g$$



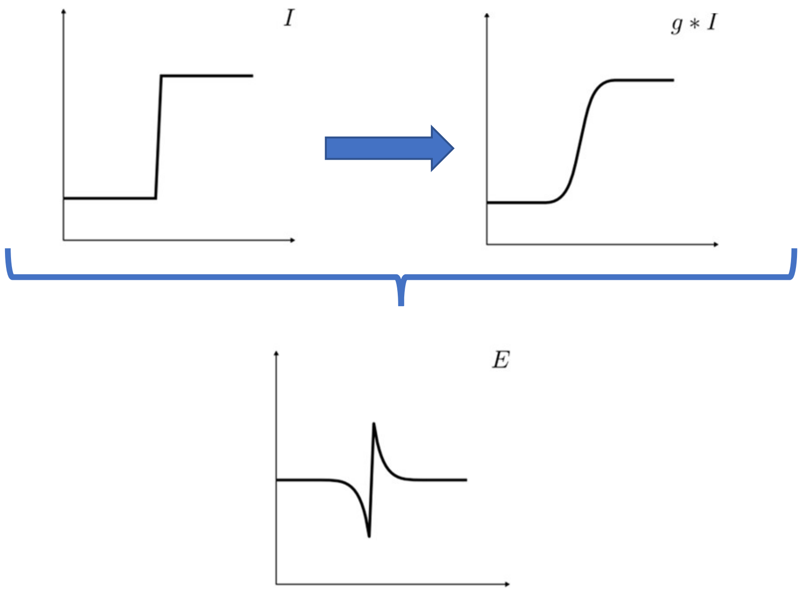
Unsharp masking: a basic sharpening method

1) Blur the original image (I) by convolving with a Gaussian Kernel (g)

$$I_{blur}=I*g$$

2) Substract the blur image from the original image, and obtain E (called E because it represents the edges of the image)

$$E=I-I_{blur}$$



Unsharp masking: a basic sharpening method

1) Blur the original image (I) by convolving with a Gaussian Kernel (g)

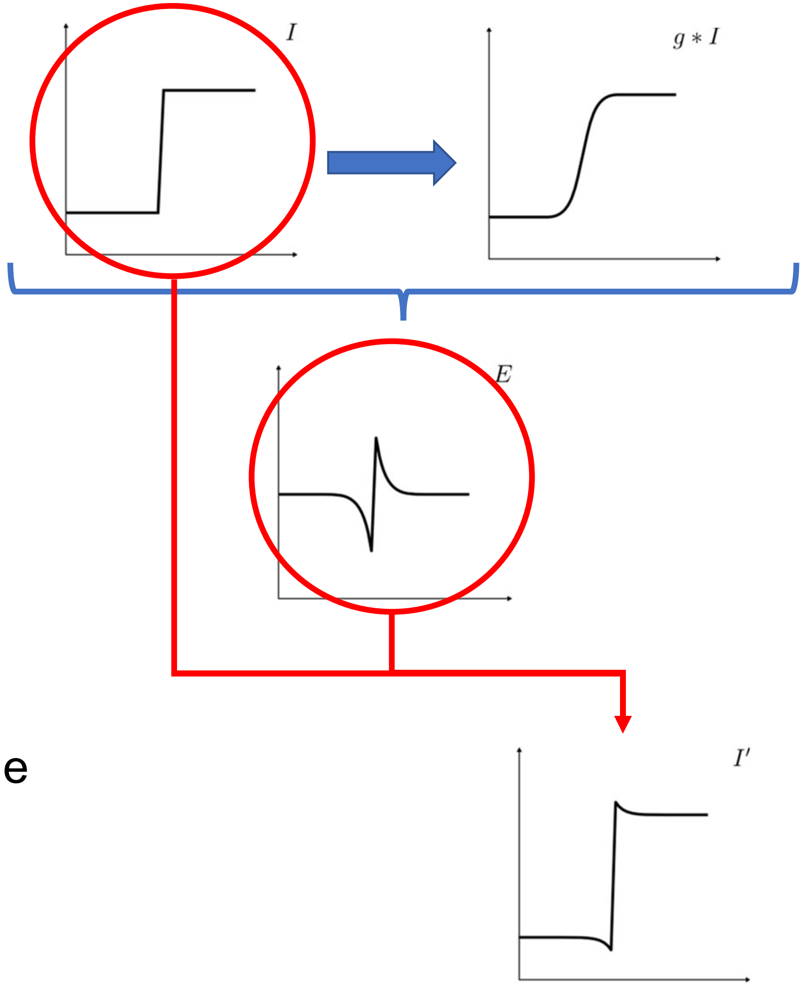
$$I_{blur}=I*g$$

2) Substract the blur image from the original image, and obtain E (called E because it represents the edges of the image)

$$E=I-I_{blur}$$

3) Obtain the sharpened image I' by adding to the original image the image E (weighted by a constant k)

$$I'=I+kE$$



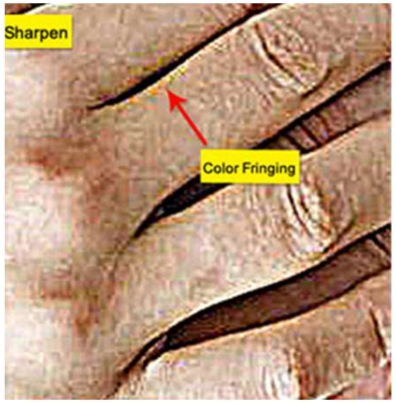
Sharpening problems

An excessive sharpening can lead to different problems:

Sharpening problems

An excessive sharpening can lead to different problems:

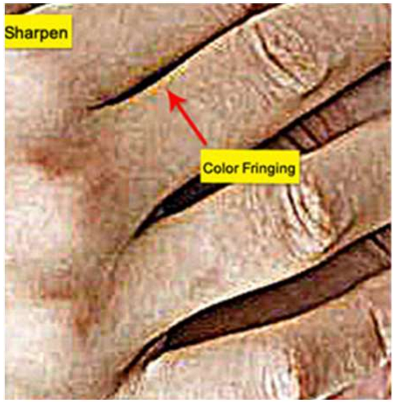
Color fringing:



Sharpening problems

An excessive sharpening can lead to different problems:

Color fringing:



Artifacts:

And now?

Neural ISPs are starting to emerge

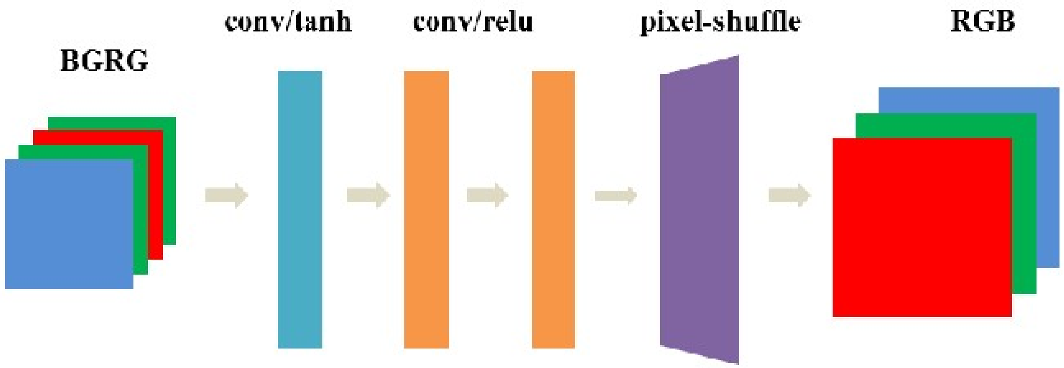


Figure 3. Smallnet architecture proposed by team dh_isp.

Image from: Ignatov, Andrey D. et al. “Learned Smartphone ISP on Mobile NPUs with Deep Learning, Mobile AI 2021 Challenge: Report.” *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2021): 2503-2514.

Image formation and color representation

Javier Vazquez-Corral
Universitat Autònoma de Barcelona

javier.vazquez@cvc.uab.cat

Slides credits: Marcelo Bertalmío