최신 Computer Vison 설명

Computer Vision

요약정리

커널연구회 (www.kernel.bz)

정재준 (rgbi3307@nate.com)

목차

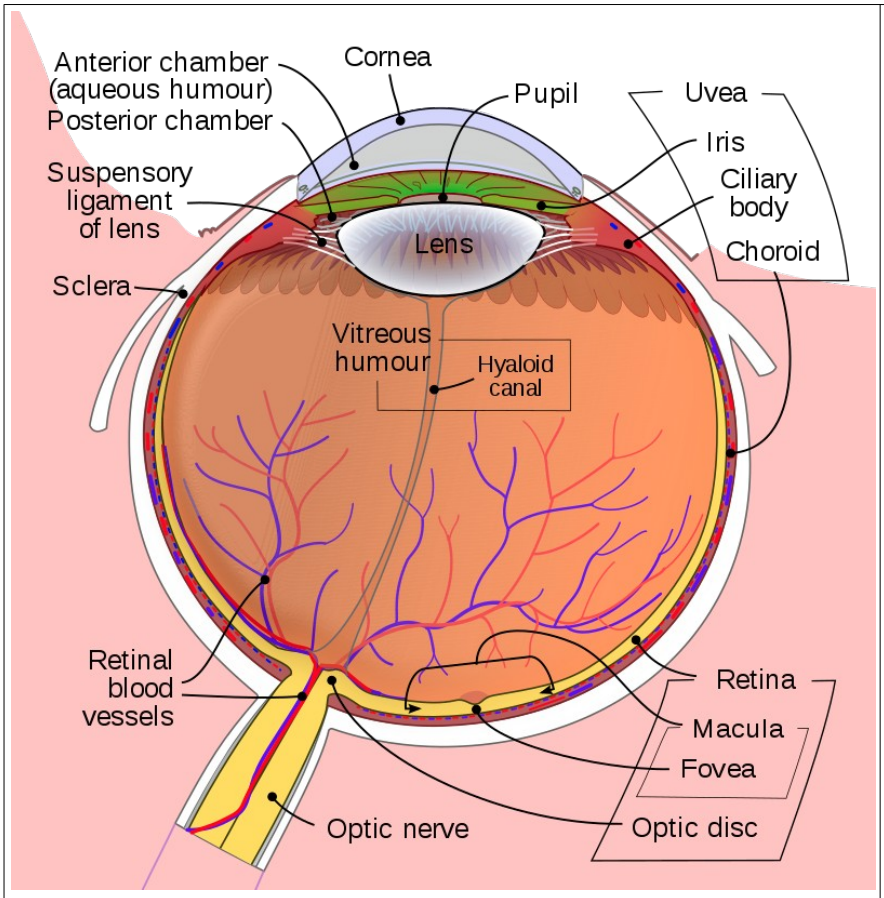
목차

1. human visual system

참조: https://en.wikipedia.org/wiki/Outline_of_computer_vision

참조: https://en.wikipedia.org/wiki/Visual_system

Eye



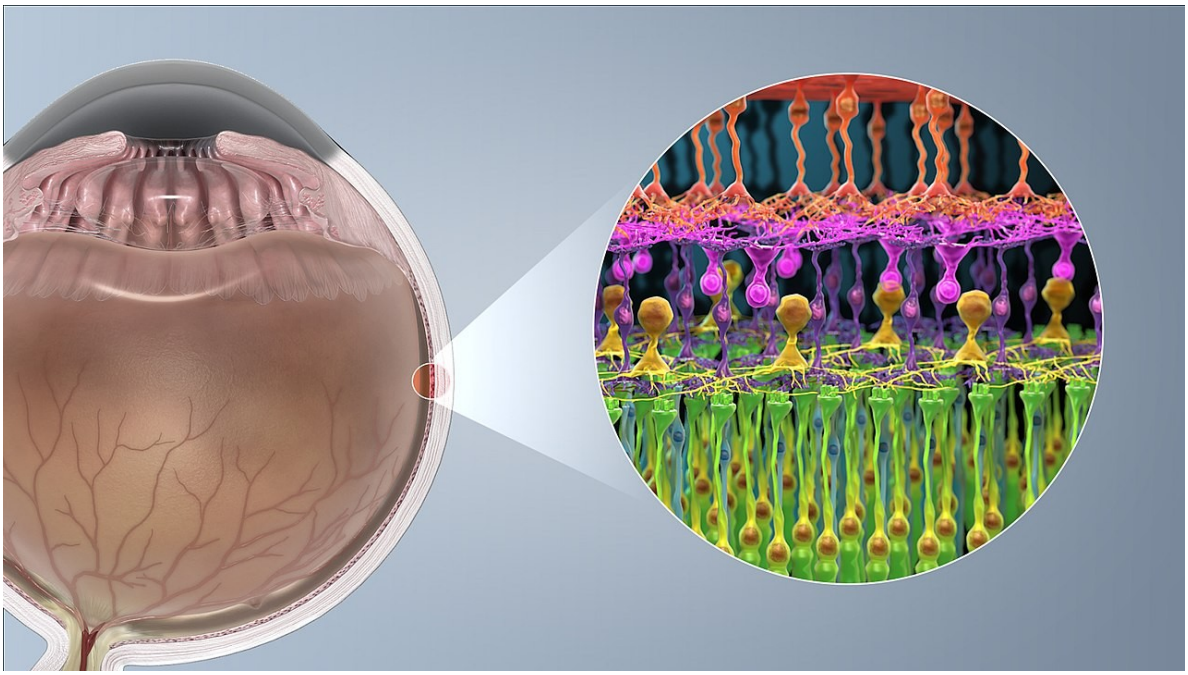
Cornea: 각막
Pupil: 동공, 눈동자
Iris: 홍체
Lens: 수정체
Retina: 망막
Macula: (망막의) 황반
Fovea: (망막의) 중심와
Optic disc:(안구의) 시신경 원판
Optic nerve: 시신경

Vitreous humour: 유리체
Hyaloid canal: 유리체관
Retinal blood vessels: 망막 혈관

그림출처:

https://en.wikipedia.org/wiki/Visual_system#/media/File:Schematic_diagram_of_the_human_eye_en.svg

Photoreceptor cell: 수광 세포



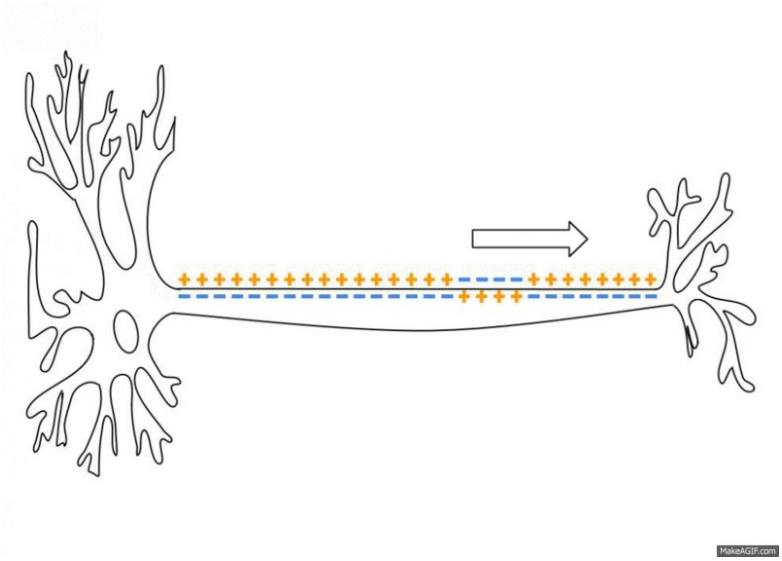
그림출처:

https://en.wikipedia.org/wiki/Photoreceptor_cell#/media/File:Photoreceptor_cell.jpg

그림출처:

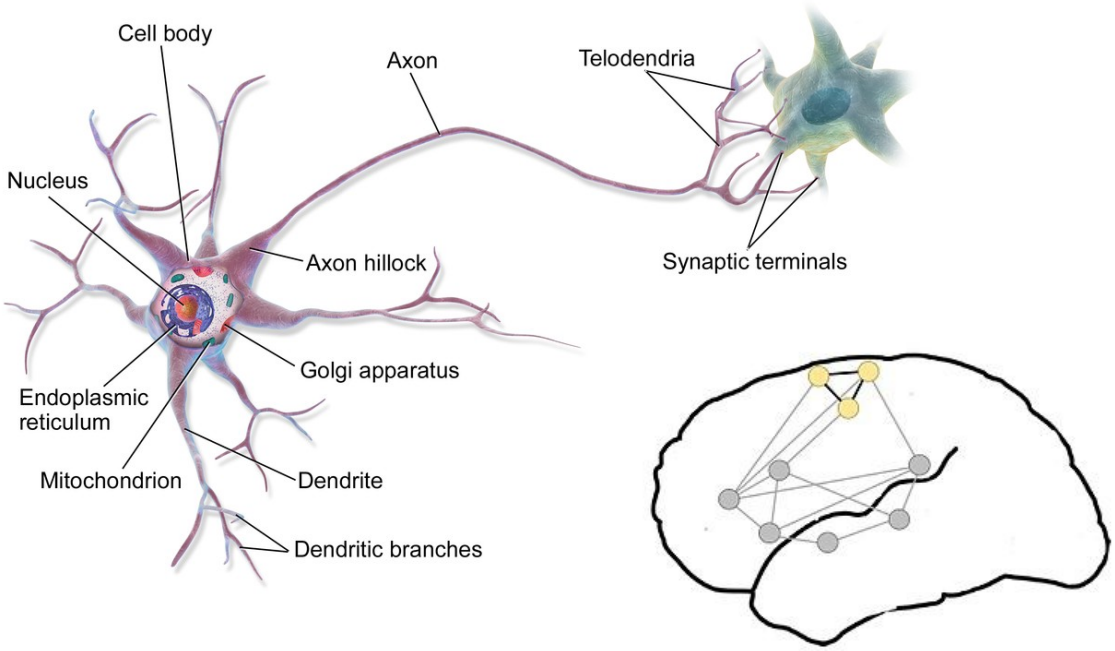
https://en.wikipedia.org/wiki/Photoreceptor_cell#/media/File:1414_Rods_and_Cones.jpg

Action potential

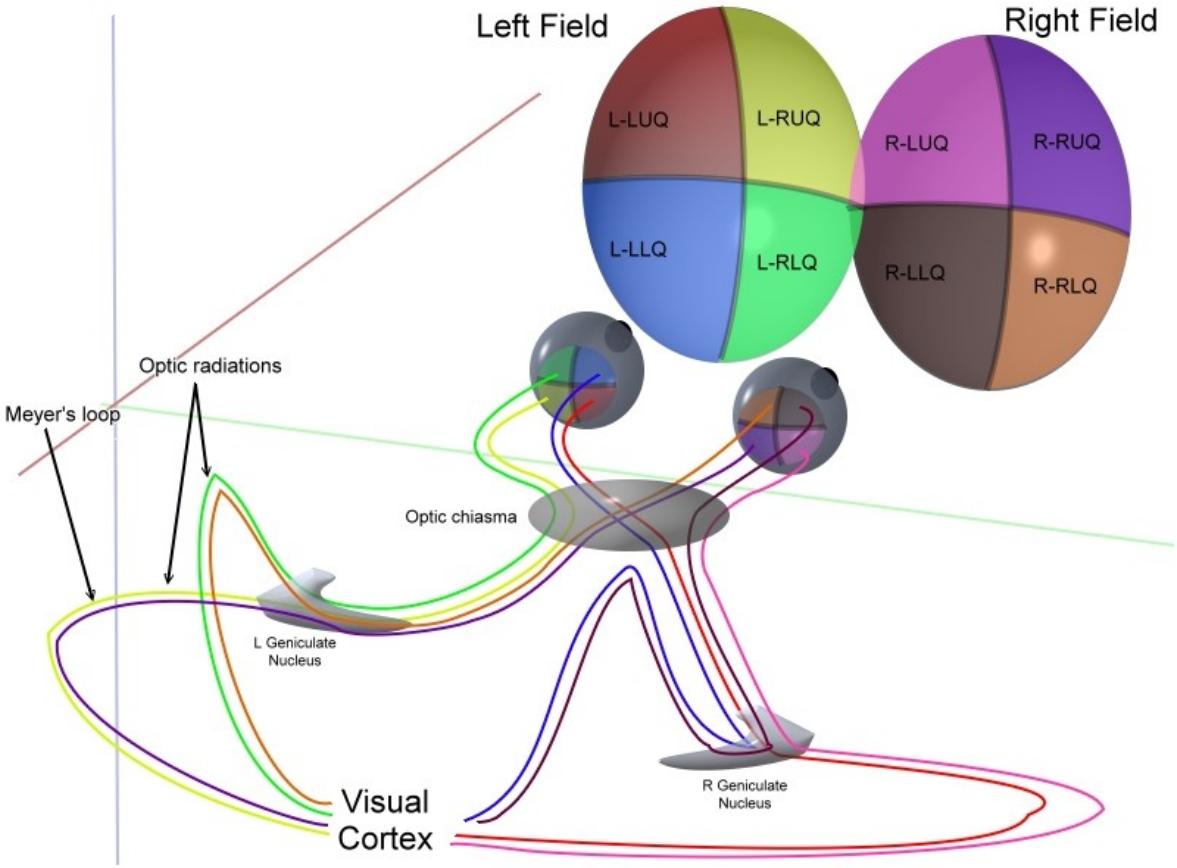


그림출처: https://commons.wikimedia.org/wiki/File:Action_Potential.gif

roughly 1.2 million axons of ganglion cells transmit information **from the retina to the brain** resulting in sensitive to color and indifferent to motion



System overview

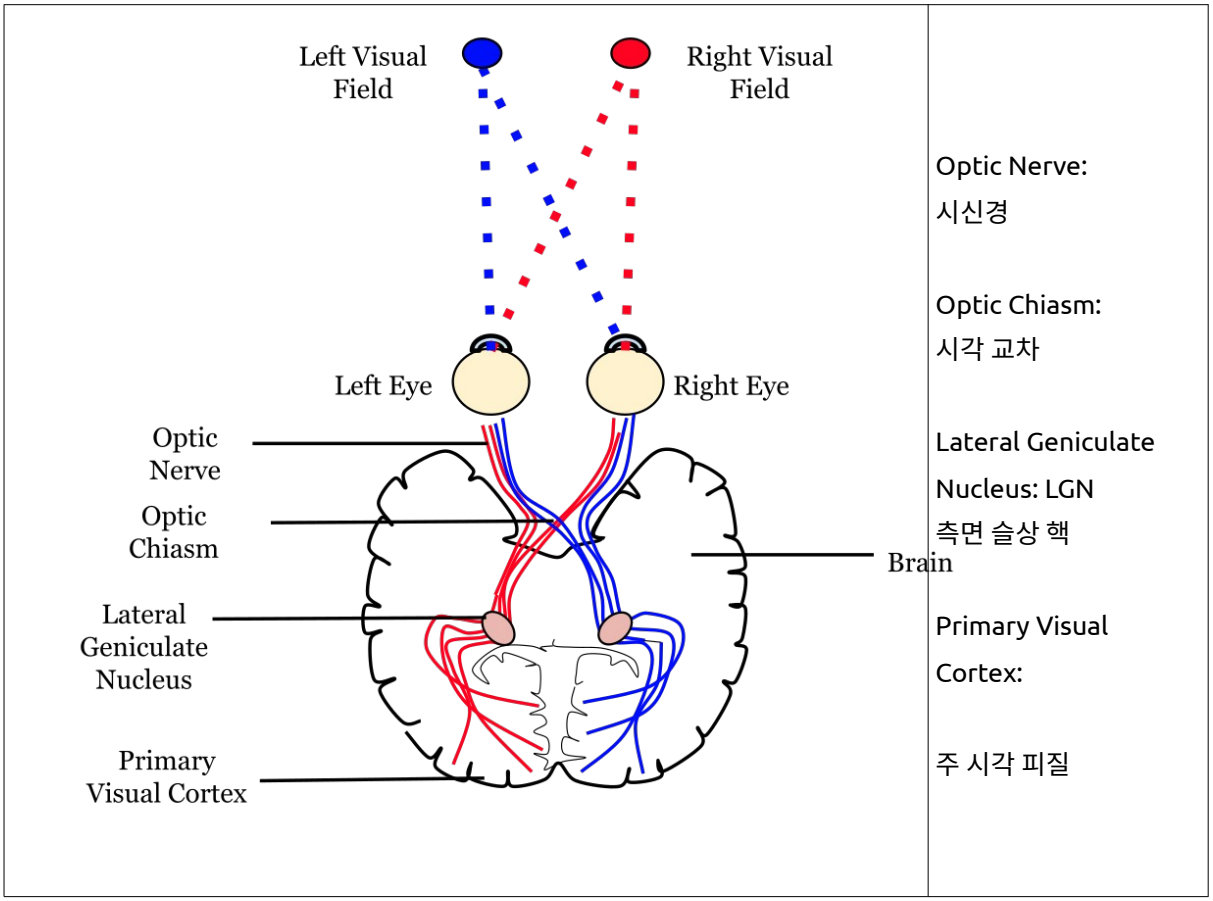


그림출처: https://en.wikipedia.org/wiki/Visual_system#/media/File:ERP_-_optic_cabling.jpg

- V1 performs edge-detection, focusing on even small spatial and color changes
- V2 comparing left and right pulses (2D images, size, color, shape), 40~100ms
- V3 helps process 'global motion' (direction and speed) of objects
- V4 recognizes simple shapes
- V5 integrates local object motion into global motion, eye movement
- V6 analyzes motion of objects relative to the background

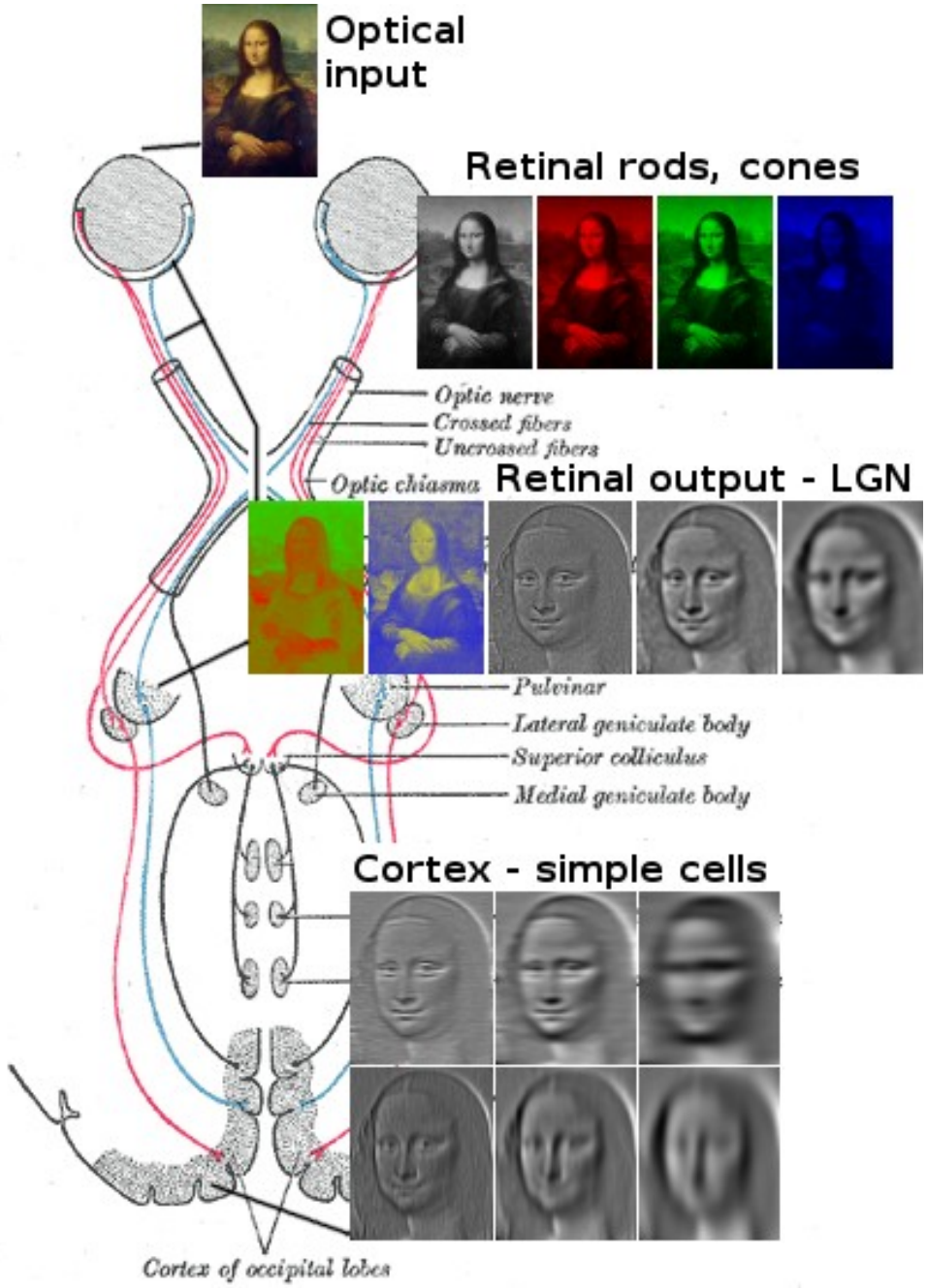
Visual cortex

https://en.wikipedia.org/wiki/Visual_cortex#Primary_visual_cortex_(V1)

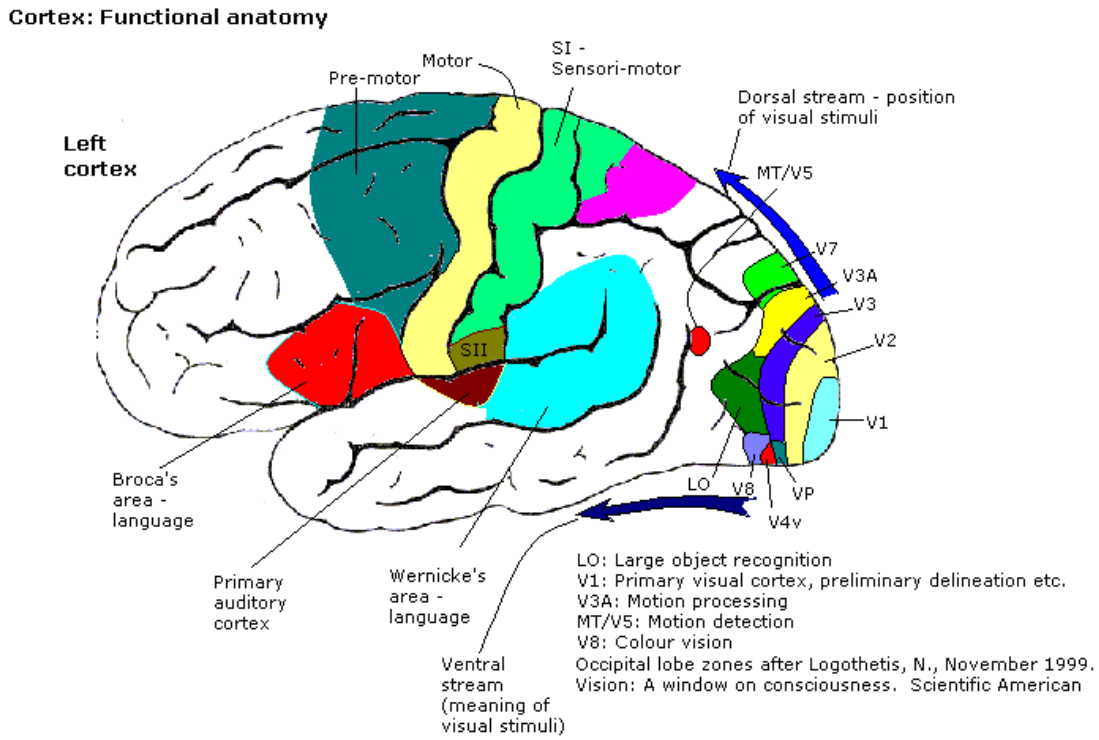


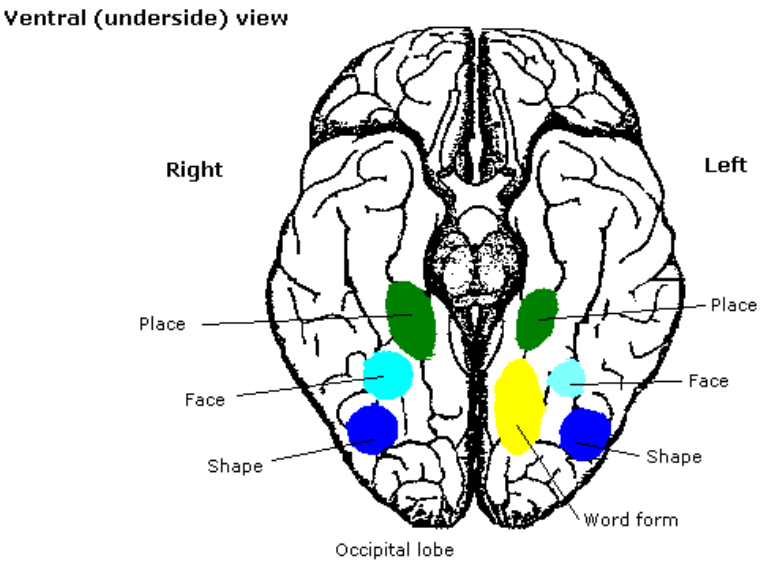
그림출처:

https://en.wikipedia.org/wiki/Visual_cortex#/media/File:Neural_pathway_diagram.svg


Optical input
Retinal rods, cones
Optic nerve
Crossed fibers
Uncrossed fibers
Optic chiasma
Retinal output - LGN
Pulvinar
Lateral geniculate body
Superior colliculus
Medial geniculate body
Cortex - simple cells
Cortex of occipital lobes

functional areas of cortex



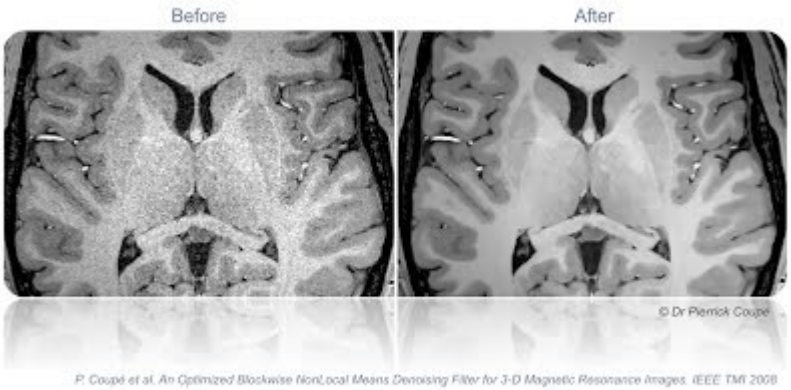


그림참조: https://en.wikipedia.org/wiki/Cerebral_cortex#/media/File:Constudproc.png

2. Image enhancement: 화상 개선

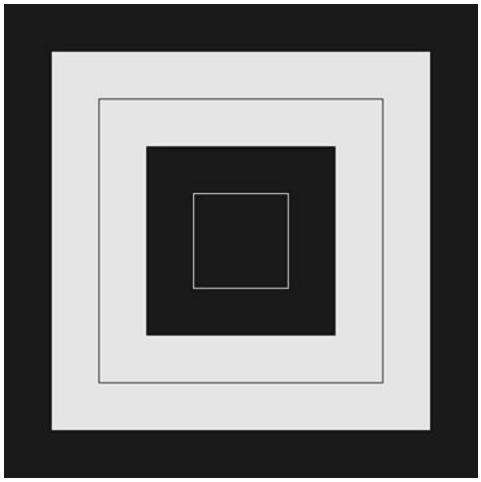
- Image denoising(Noise reduction)
- Image histogram
- Inpainting
- Histogram equalization
- Tone mapping
- Retinex
- Gamma correction
- Anisotropic diffusion (Perona–Malik equation)

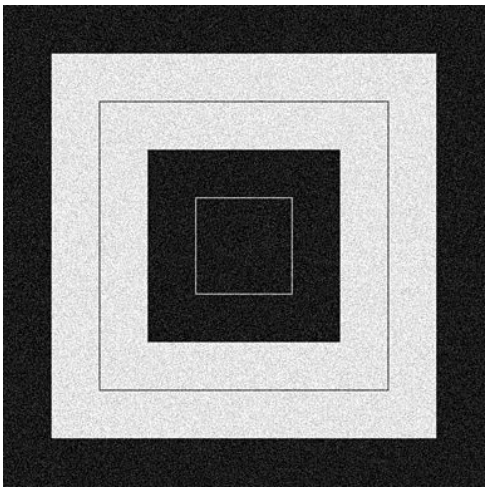
Image denoising(Noise reduction)



그림출처: https://sites.google.com/site/pierrickcoupe/softwares/denoising-for-medical-imaging/mri-denoising

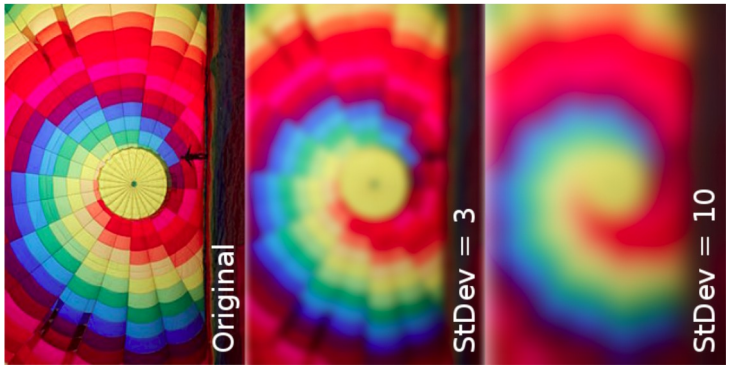
Gaussian noise

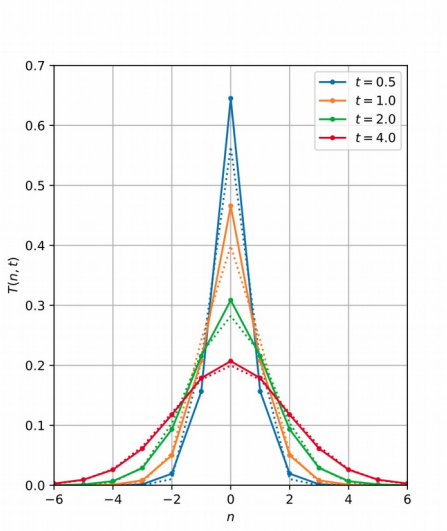




참조: https://en.wikipedia.org/wiki/Gaussian_noise

Linear smoothing filters / Gaussian smoothing / Gaussian function





참조 https://en.wikipedia.org/wiki/Gaussian_blur

출처: https://en.wikipedia.org/wiki/Gaussian_function

Nonlinear filters / median filter

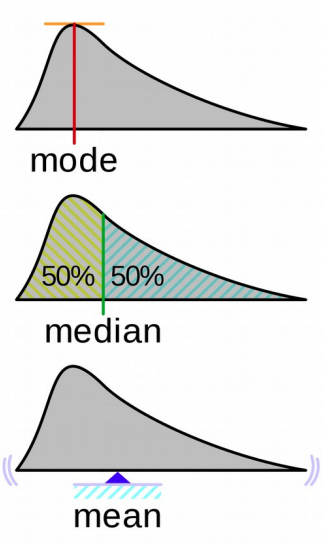
1, 3, 3, **6**, 7, 8, 9

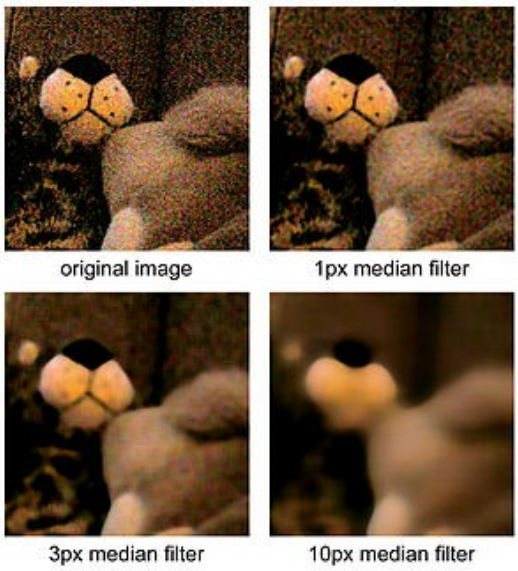
Median = 6

1, 2, 3, **4**, **5**, 6, 8, 9

Median = (4 + 5) ÷ 2

= 4.5

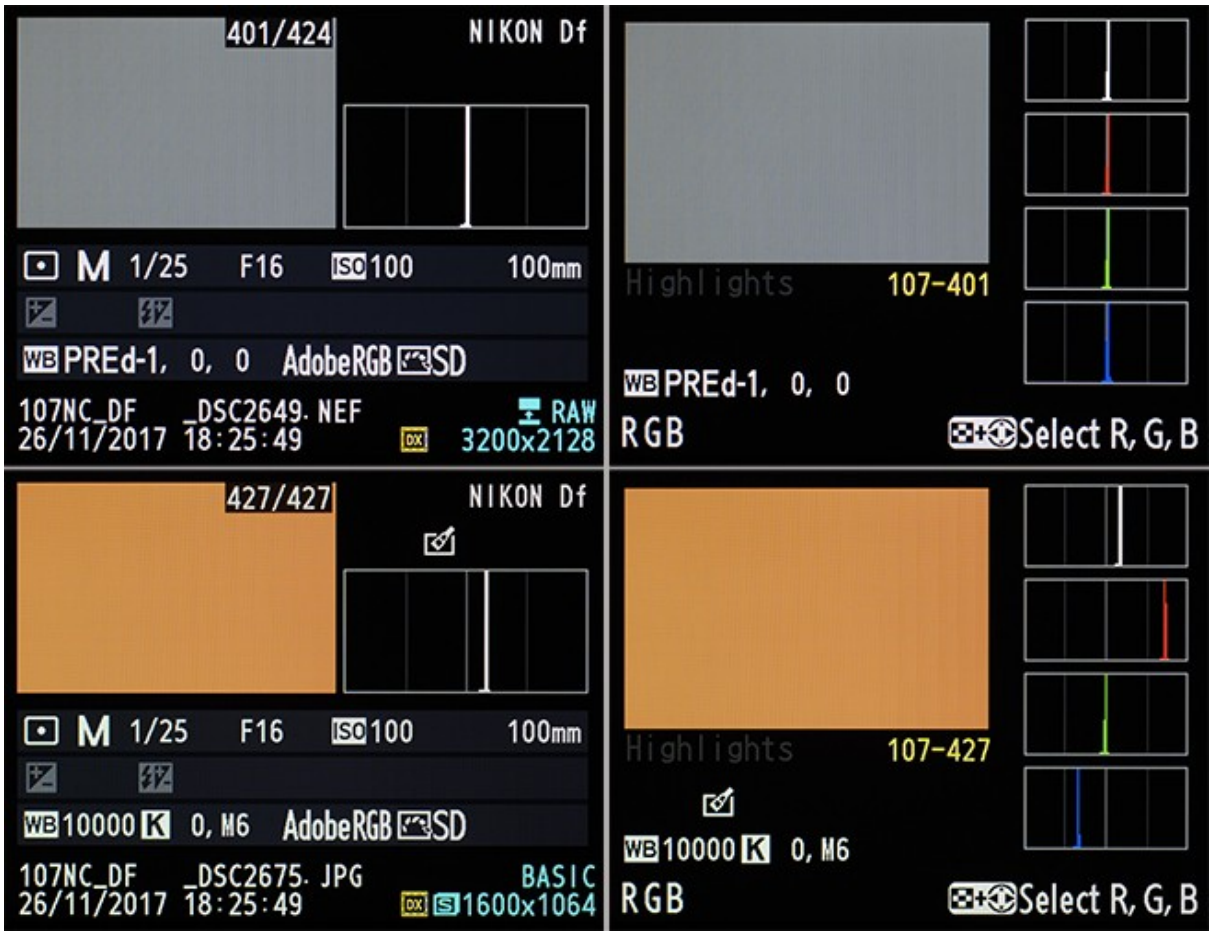




참조: https://en.wikipedia.org/wiki/Median_filter

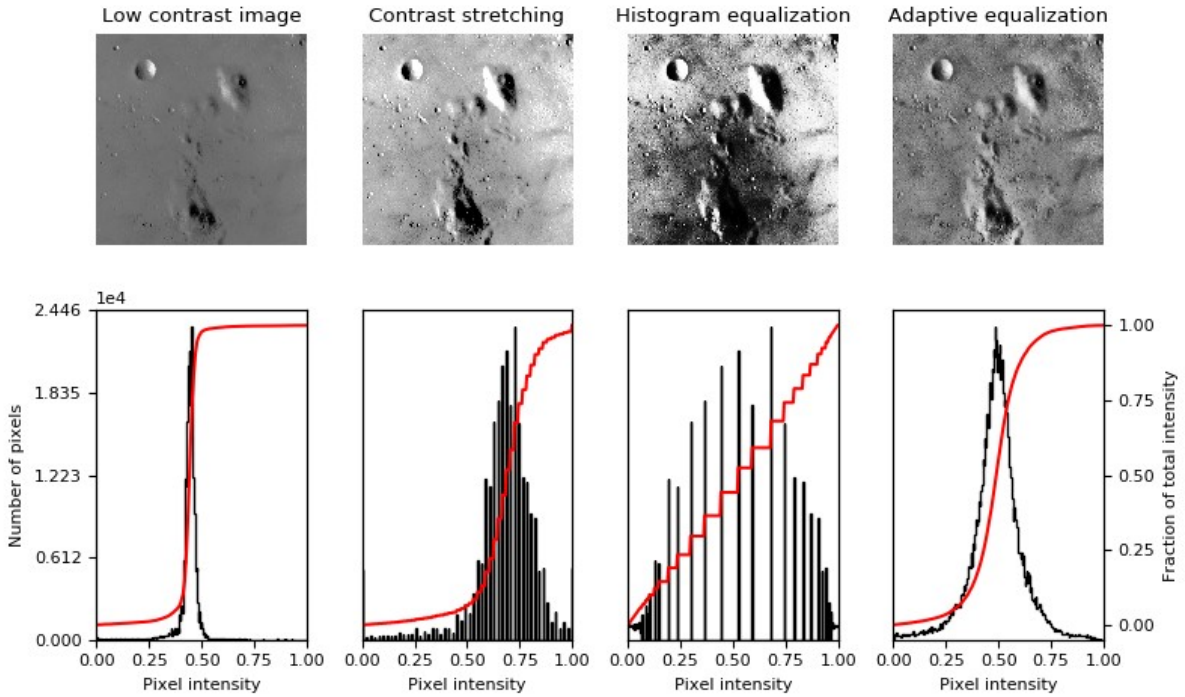
Wavelet transform / Statistical methods / Block-matching algorithms / Random field / Deep learning

Image histogram



그림출처: https://www.fastrawviewer.com/blog/in-camera-histogram-doesn%27t-represent-exposure

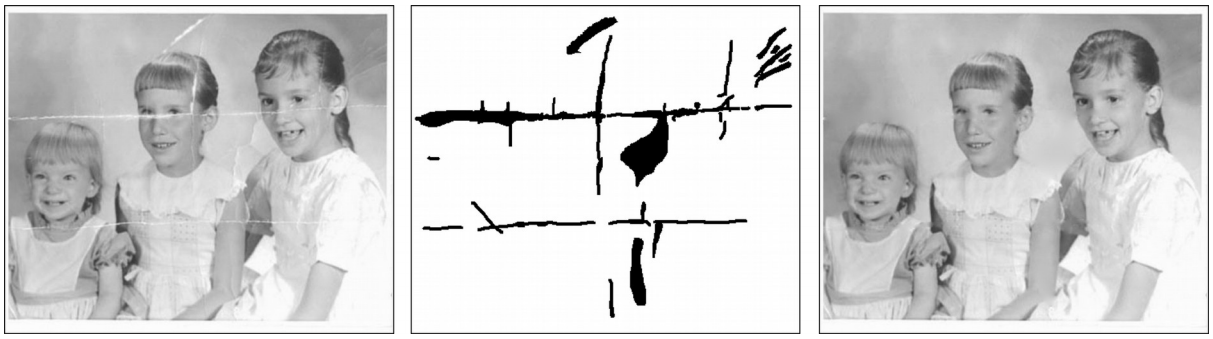
Histogram equalization



그림출처:

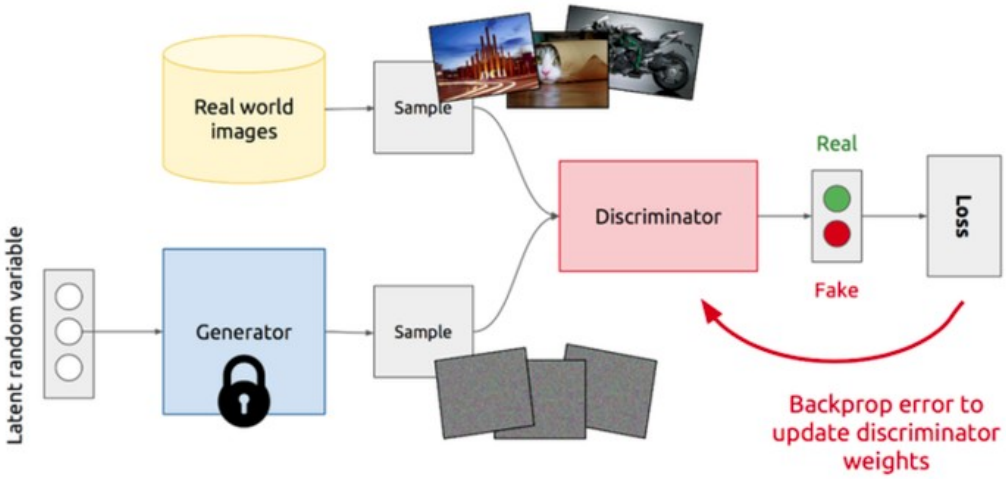
https://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_equalize.html

Inpainting



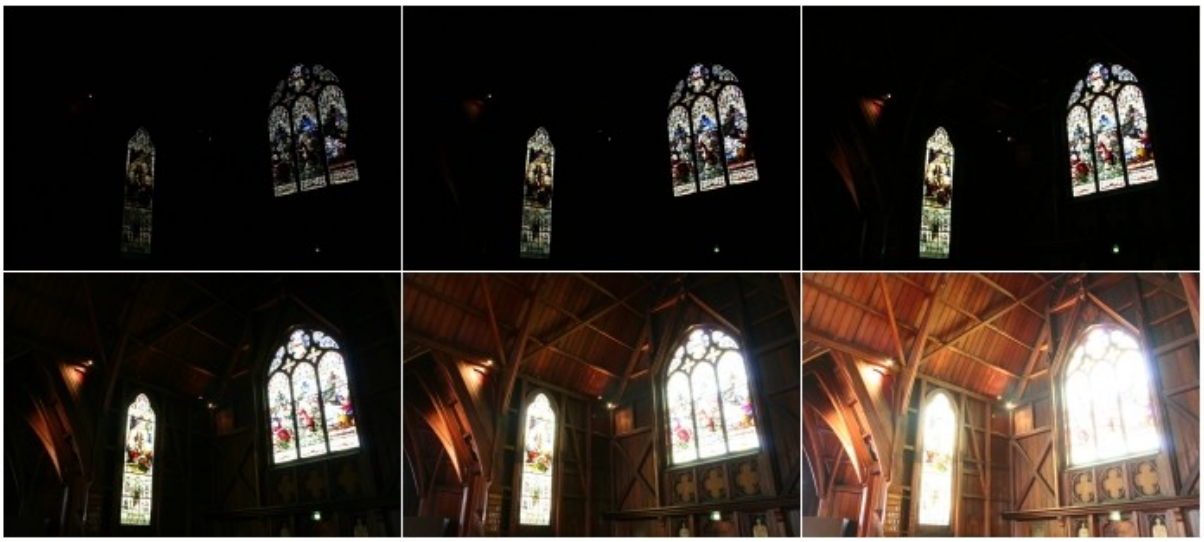
그림출처: https://github.com/Mugichoko445/Fast-Digital-Image-Inpainting

Image Inpainting with GAN



그림출처: https://worthpreading.tistory.com/64

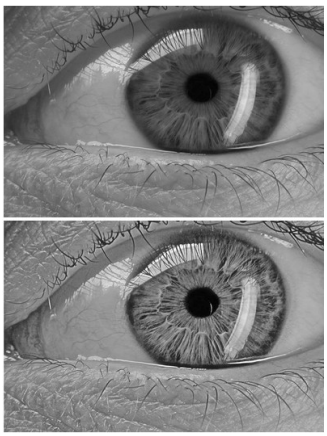
Tone mapping (High-Dynamic-Range imaging)



그림출처: https://en.wikipedia.org/wiki/Tone_mapping

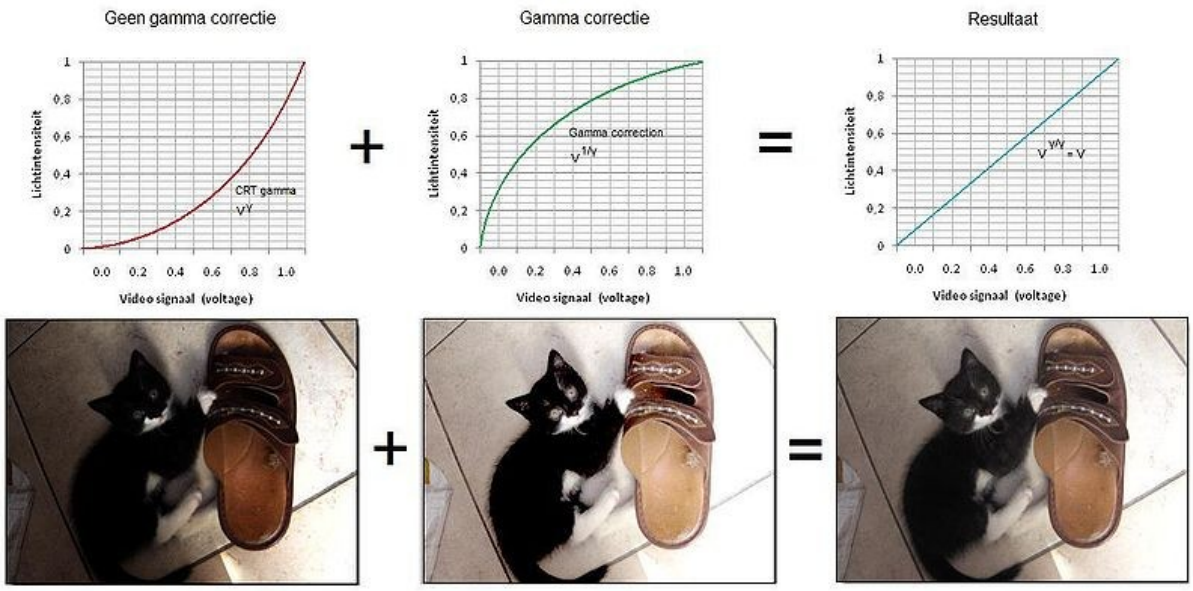
Retinex

입력영상의 배경성분 제거, contrast 향상, sharpness 증진.



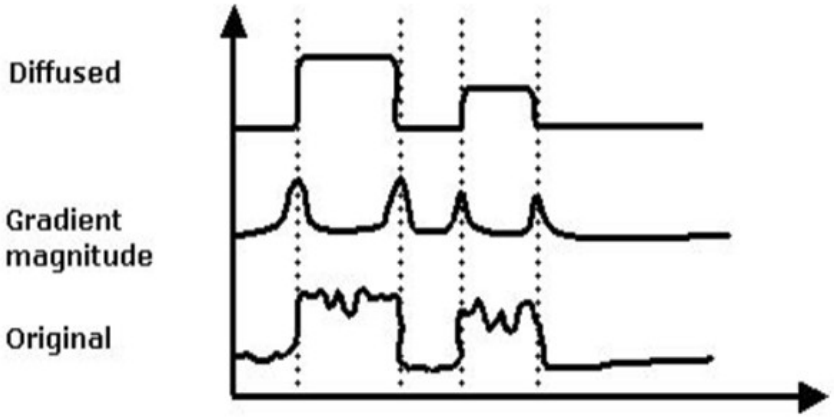
그림출처: https://kipl.tistory.com/65

Gamma correction

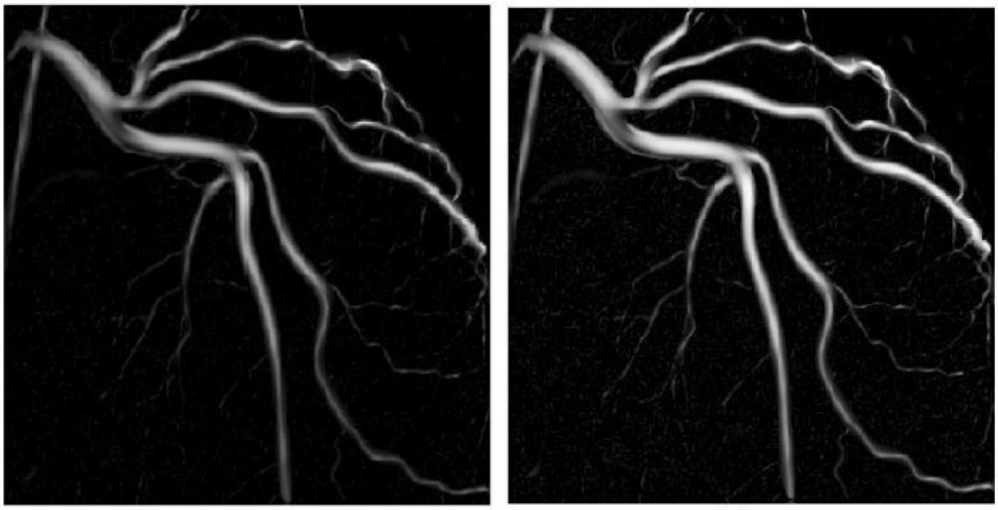


그림출처: https://commons.wikimedia.org/wiki/File:Gamma_correction_brabbit.jpg

Anisotropic diffusion



그림출처: https://dsp.stackexchange.com/questions/14606/anisotropic-diffusion

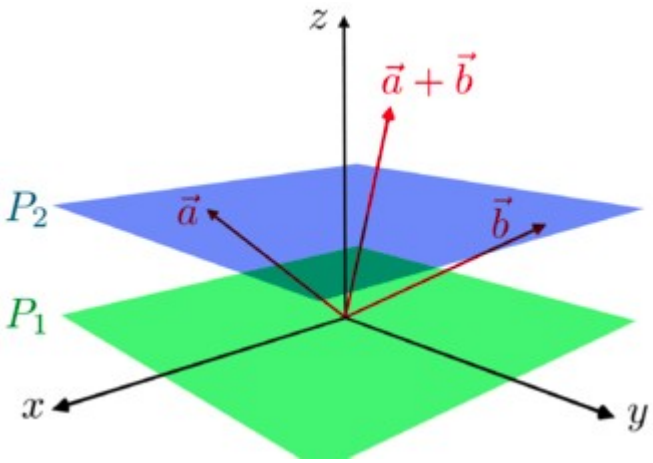


그림출처: https://www.researchgate.net/figure/Segmented-output-After-Anisotropic-Diffusion-Filtering-small-arteries-are-more-prominent_fig3_271483004

3. Transformations: 변형, 변환

- Affine transform
- Homography (computer vision)
- Hough transform
- Radon transform
- Walsh–Hadamard transform

Affine transform



Affine space: https://en.wikipedia.org/wiki/Affine_space

translation, scaling, homothety, similarity transformation, reflection, rotation, shear mapping, compositions

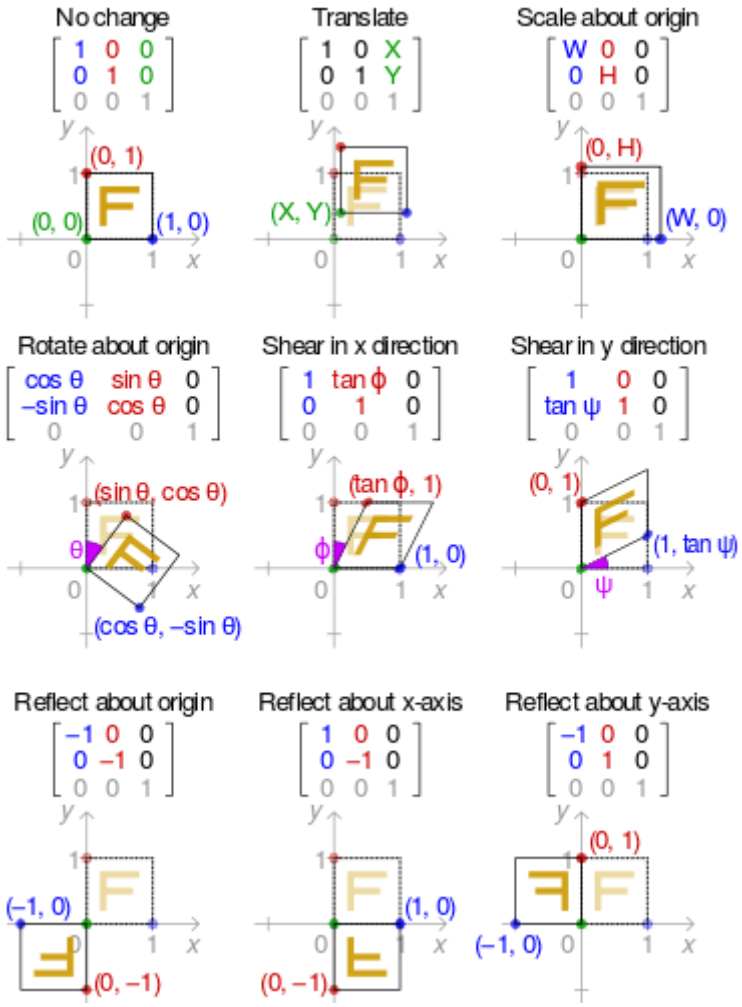
2D affine transformation matrices on a unit square

Homography (computer vision)

Hough transform

Radon transform

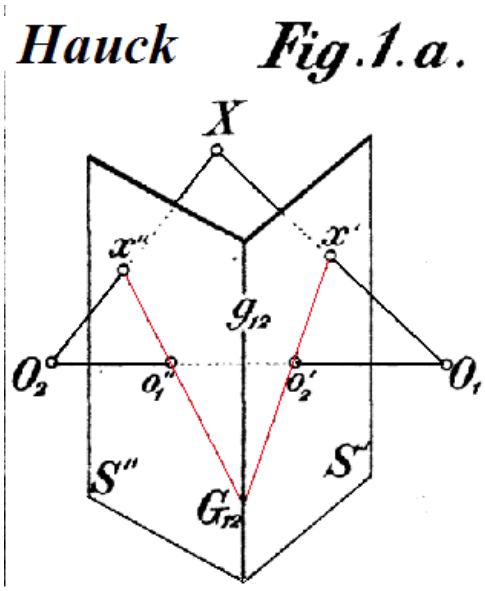
Walsh–Hadamard transform



출처: https://en.wikipedia.org/wiki/Affine_transformation

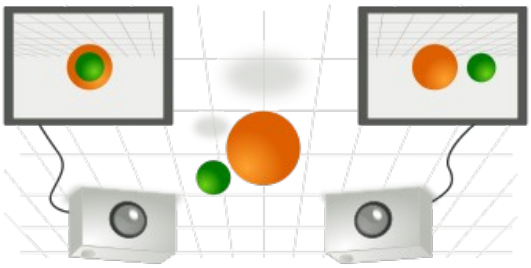
Homography (computer vision)

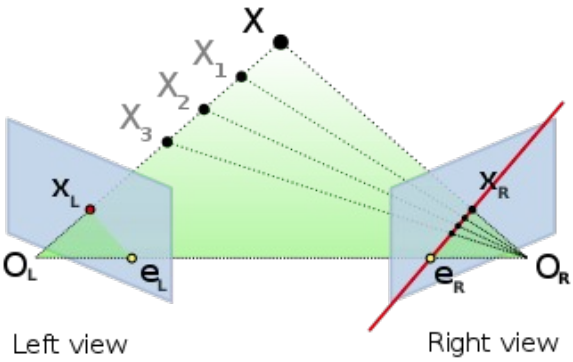
image rectification, image registration, or computation of camera motion—rotation and translation—between two images.



참조: https://en.wikipedia.org/wiki/Homography_(computer_vision)

O1 and O2 both pointed at X in **epipolar geometry**





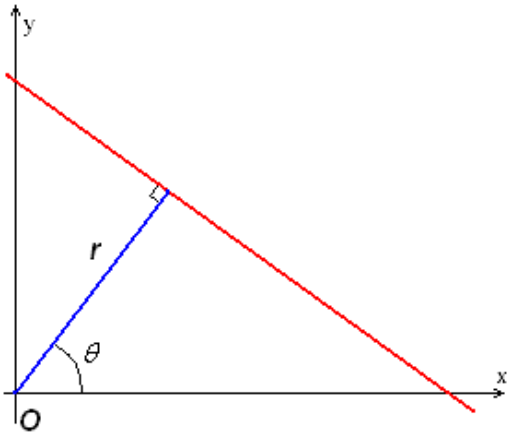
그림출처:

https://en.wikipedia.org/wiki/Epipolar_geometry

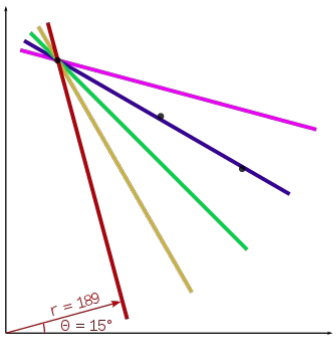
Hough transform

feature extraction, identification of lines in the image, identifying positions of arbitrary shapes

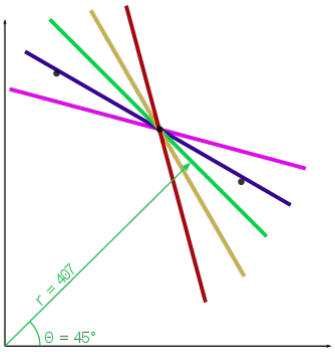
r = x cos θ + y sin θ



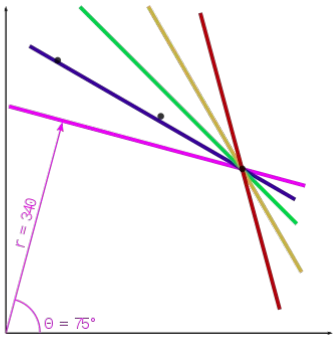
출처: https://en.wikipedia.org/wiki/Hough_transform



Θ	r
15	189.0
30	282.0
45	355.7
60	407.3
75	429.4



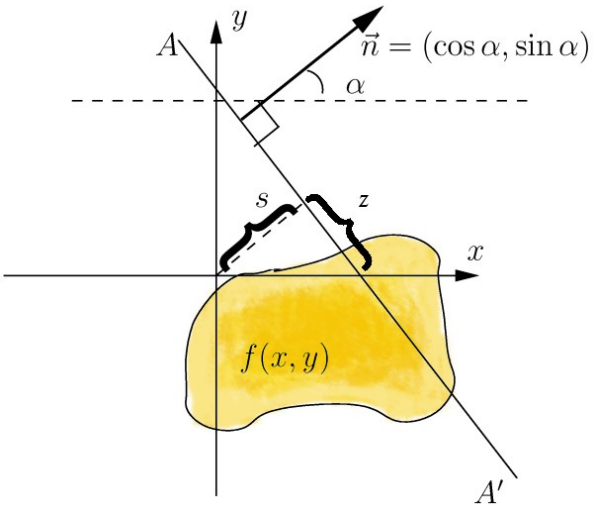
Θ	r
15	318.5
30	376.8
45	407.3
60	409.8
75	385.3



Θ	r
15	419.0
30	443.6
45	438.4
60	402.9
75	340.1

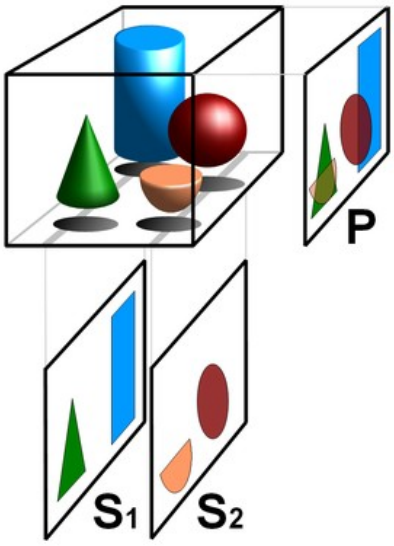
Radon transform

Maps f on the (x, y)-domain to Rf on the (α, s)-domain



출처: https://en.wikipedia.org/wiki/Radon_transform

Radon transform is widely applicable to **tomography**

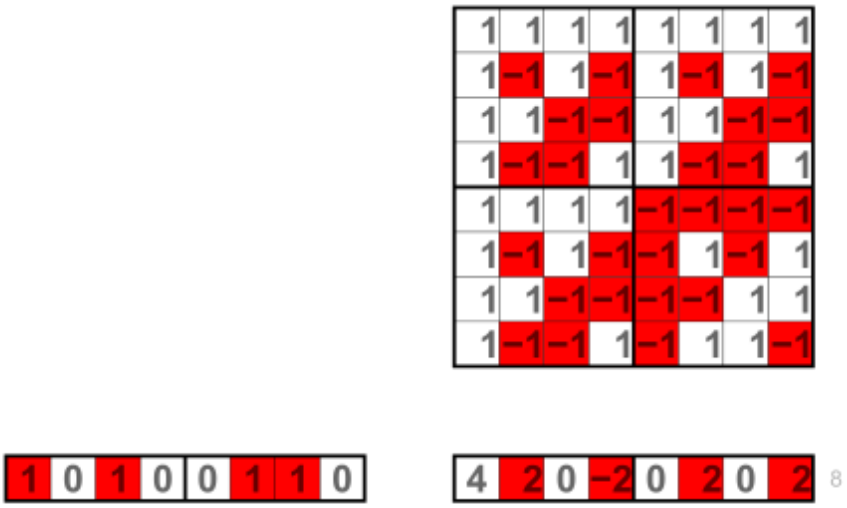


그림출처: https://en.wikipedia.org/wiki/Tomography

Walsh–Hadamard transform

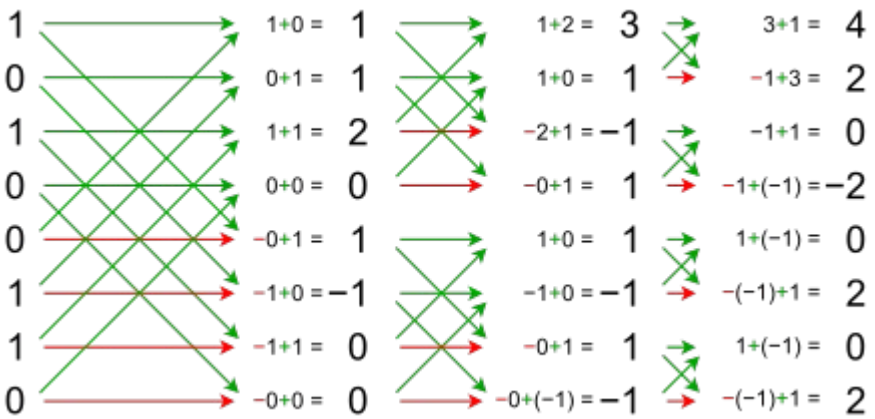
Many **quantum algorithms** use the Hadamard transform as an initial step.

The Hadamard transform is also used in data encryption, video compression,



출처(Walsh matrix): https://en.wikipedia.org/wiki/Hadamard_transform

Fast Walsh–Hadamard transform



4. Filtering: 걸러내기, 선별

- Image compression
- Filter bank
- Gabor filter
- JPEG 2000
- Adaptive filtering

Image compression

참조: https://en.wikipedia.org/wiki/Image_compression

Methods for **lossy** compression: Transform coding

Discrete Cosine Transform (DCT) is JPEG, HEIF

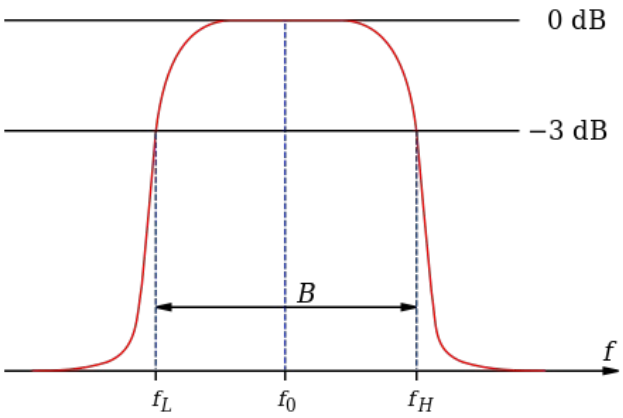
Methods for **lossless** compression:

PCX, BMP, TGA, TIFF, arithmetic coding and Huffman coding

Filter bank

filter bank is an array of **band-pass filters** that separates the input signal into multiple components

참조: https://en.wikipedia.org/wiki/Filter_bank



출처: https://en.wikipedia.org/wiki/Band-pass_filter

Gabor filter

Gabor filter is a linear filter used for texture analysis(질감 분석)

Gabor filters is thought by some to be similar to perception in the human visual system.

참조: https://en.wikipedia.org/wiki/Gabor_filter

외곽선을 검출하는 기능을 하는 필터로, 사람의 시각체계가 반응하는 것과 비슷하다는 이유로 널리 사용되고 있다. Gabor Fiter 는 간단히 말해서 사인 함수로 모듈레이션 된 Gaussian Filter 라고

생각할 수 있다. 파라미터를 조절함에 따라 Edge 의 크기나 방향성을 바꿀 수 있으므로 Bio-inspired 영상처리 알고리즘에서 특징점 추출 알고리즘으로 핵심적인 역할을 하고 있다.

출처: https://thinkpiece.tistory.com/304

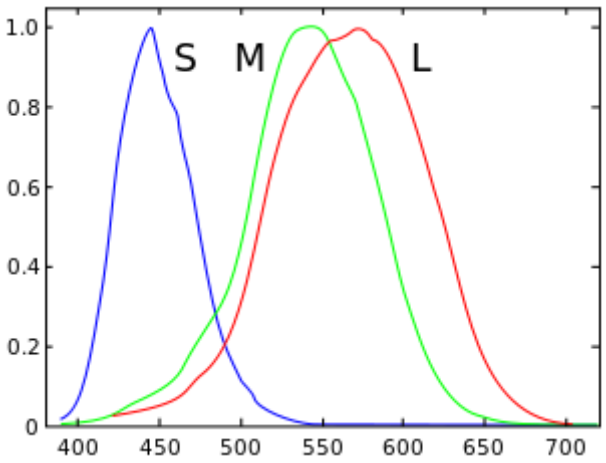
그림출처: https://www.slideshare.net/ankitnayan3/gabor-filtering-for-fingerprint-image-enhancement

5. Color vision: 색상 식별

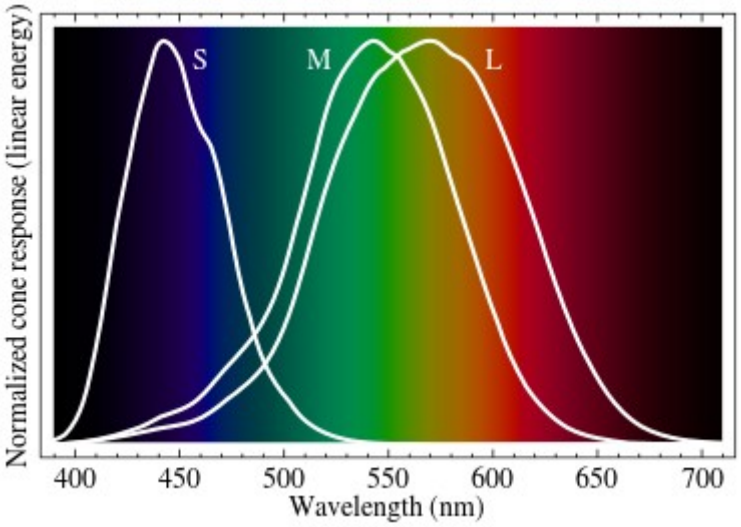
- Visual perception (시각적 인식)
- Human visual system model
- Color matching function
- Color space
- Color appearance model
- Color management system
- Color mapping
- Color model
- Color profile

Color matching function

참고: https://en.wikipedia.org/wiki/CIE_1931_color_space#Color_matching_functions

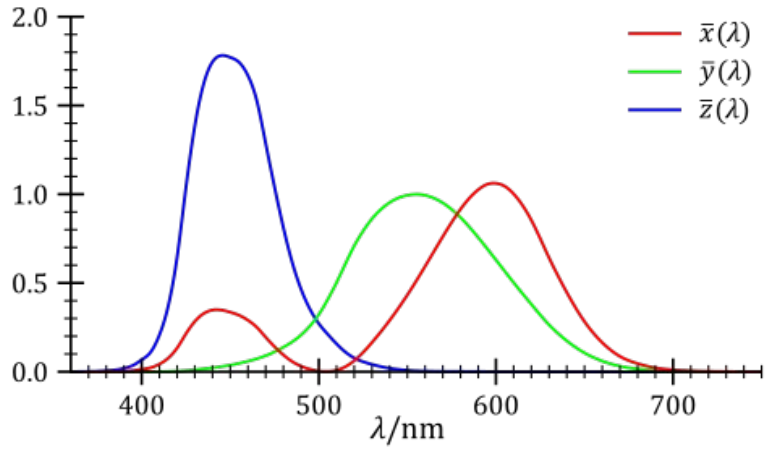


Cone cells, or cones, are photoreceptor cells in the retinas of vertebrate eyes (e.g. the human eye)

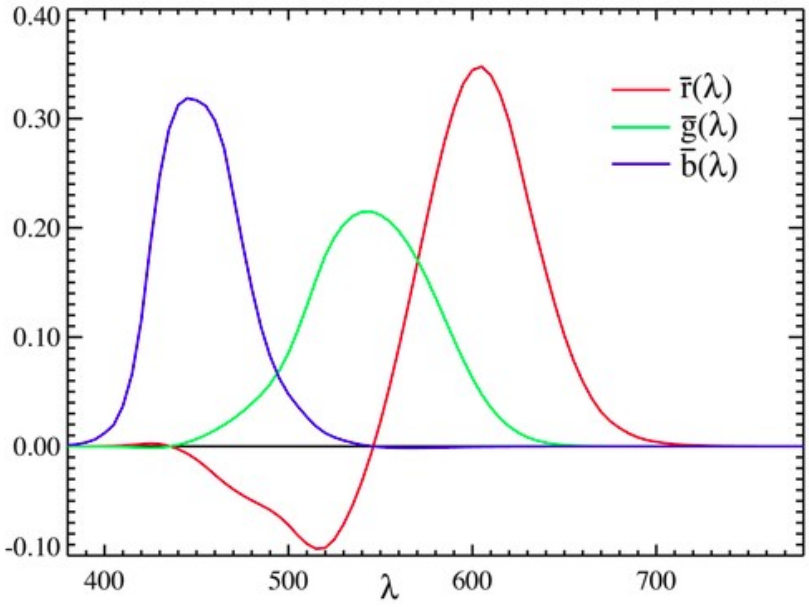


출처(Cone cell): https://en.wikipedia.org/wiki/Cone_cell

The CIE XYZ standard observer **color matching functions**



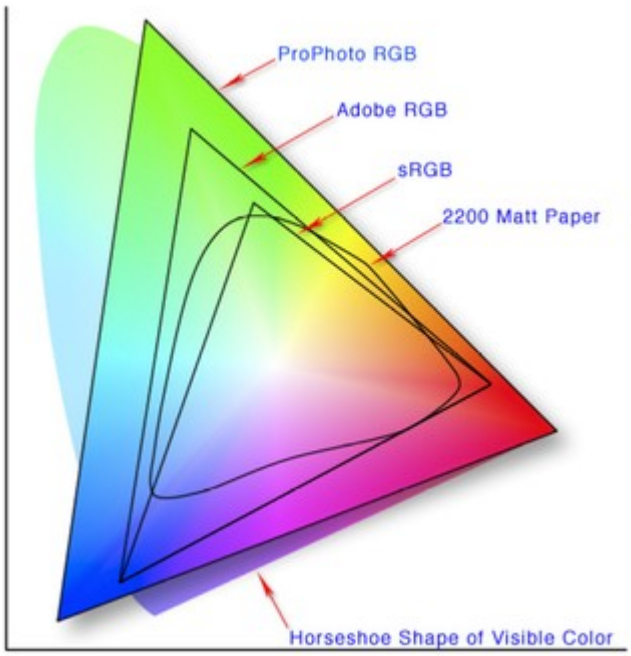
The CIE **RGB color matching functions**

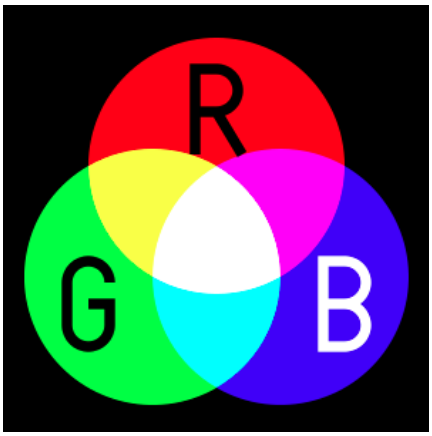


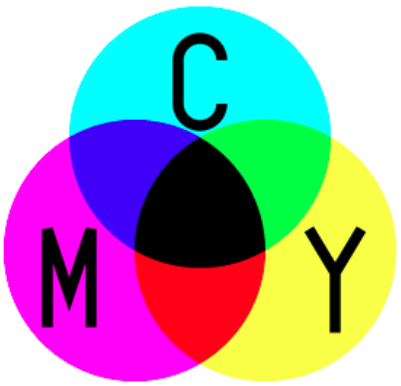
Color space

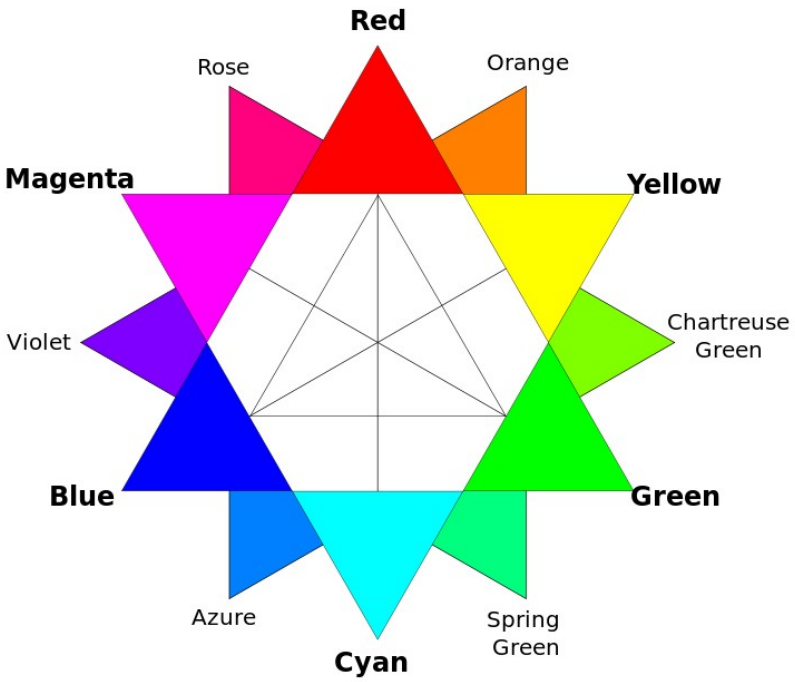
color space is a specific organization of colors

참조: https://en.wikipedia.org/wiki/Color_space









그림출처: https://en.wikipedia.org/wiki/File:Color_star-en.svg

6. Feature extraction: 특징 추출

- Active contour
- Blob detection
- Canny edge detector
- Edge detection
- Harris Corner Detector
- Random sample consensus (RANSAC)

Active contour (snakes)

object tracking, shape recognition, segmentation, edge detection and stereo matching.

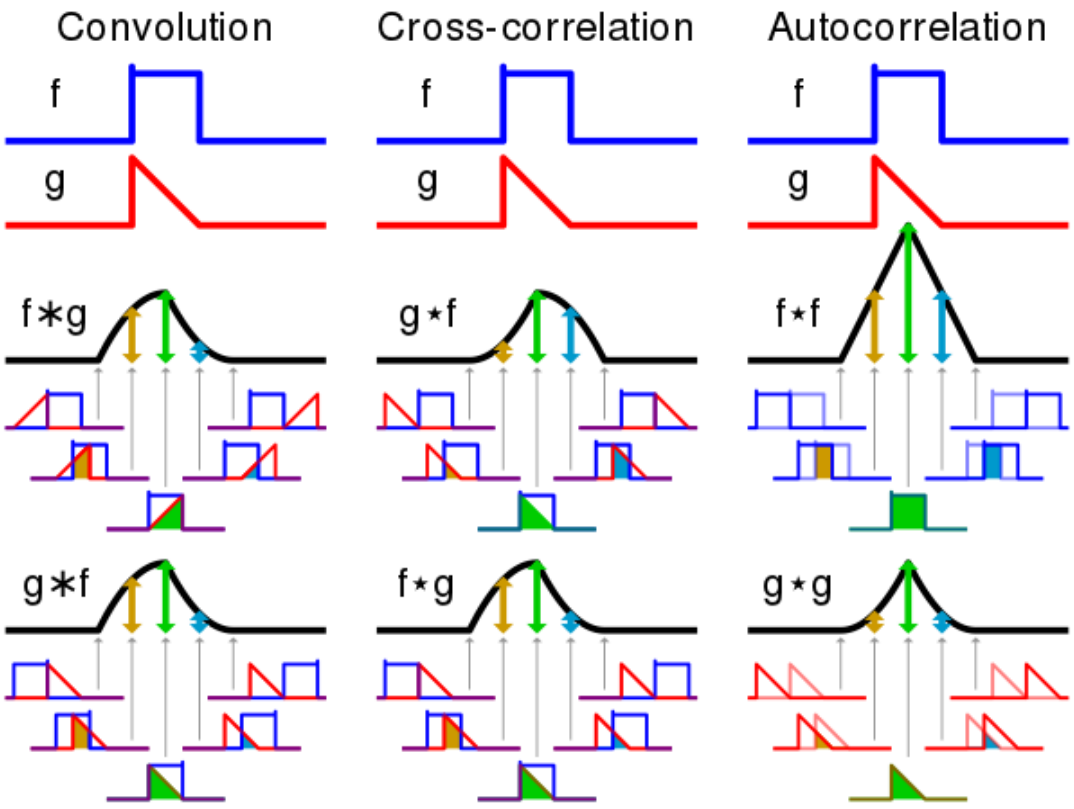
참고: https://en.wikipedia.org/wiki/Active_contour_model

Blob detection

blob detection methods are aimed at detecting regions , brightness or color, compared to surrounding regions.

참고: https://en.wikipedia.org/wiki/Blob_detection

The most common method for blob detection is **convolution**.

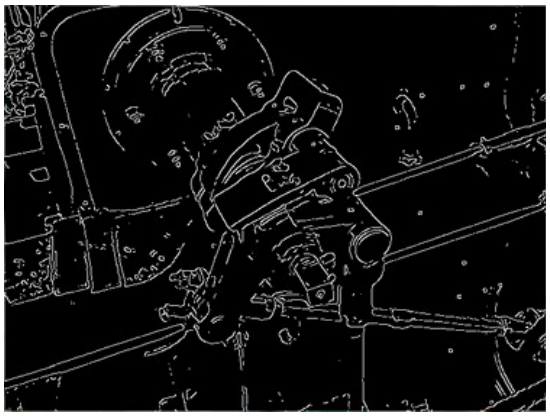


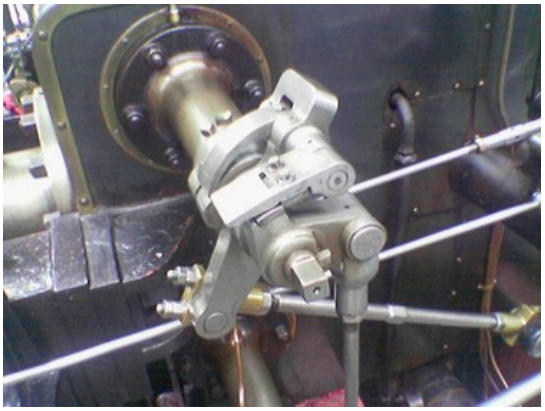
그림참조: https://en.wikipedia.org/wiki/Convolution

Canny edge detector

The Canny edge detector is an **edge detection** operator that uses a multi-stage algorithm to detect a wide range of edges in images.

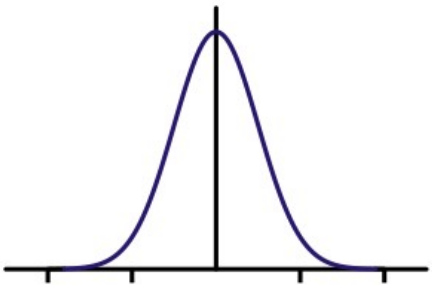
참조: https://en.wikipedia.org/wiki/Canny_edge_detector





Apply Gaussian filter to smooth the image in order to remove the noise

Find the intensity gradients of the image

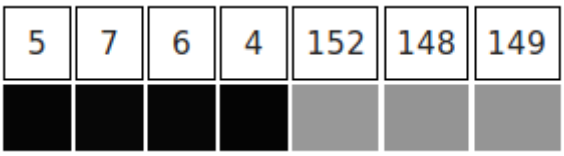


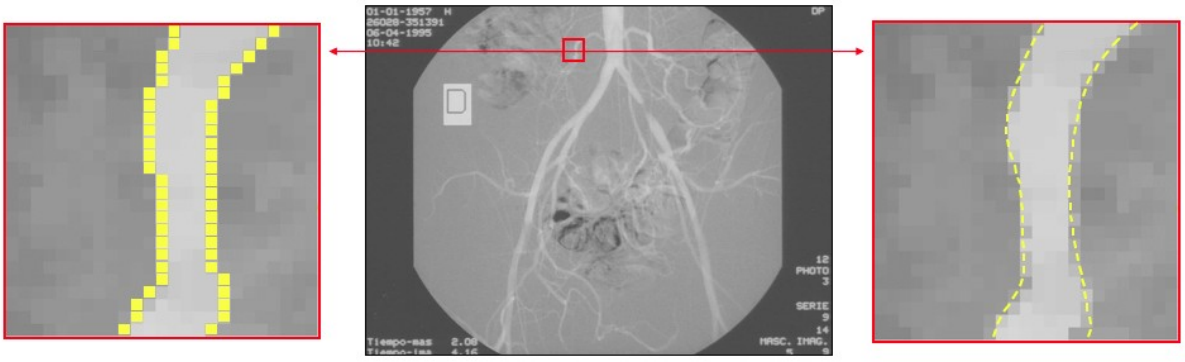
Gaussian Filter: https://en.wikipedia.org/wiki/Gaussian_filter

Edge detection

discontinuities in depth, discontinuities in surface orientation, changes in material properties and variations in scene illumination.

참조: https://en.wikipedia.org/wiki/Edge_detection



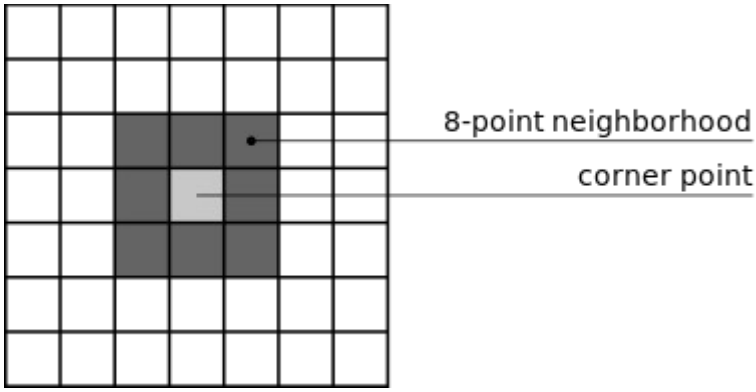


Edge detection on an angiographic image. **On the left**, edge detection is made at a pixel level. **On the right**, **subpixel edge detection** locates the edge inside the pixel precisely

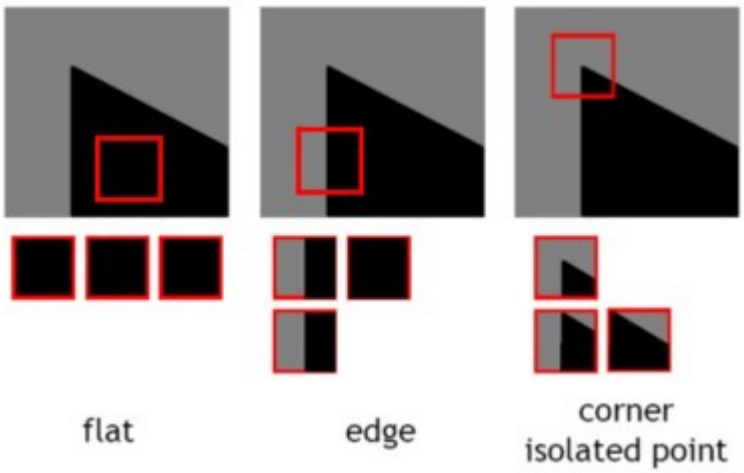
Harris Corner Detector

Harris Corner Detector is a corner detection operator that is commonly used in computer vision algorithms **to extract corners** and **infer features** of an image.

참조: https://en.wikipedia.org/wiki/Harris_Corner_Detector



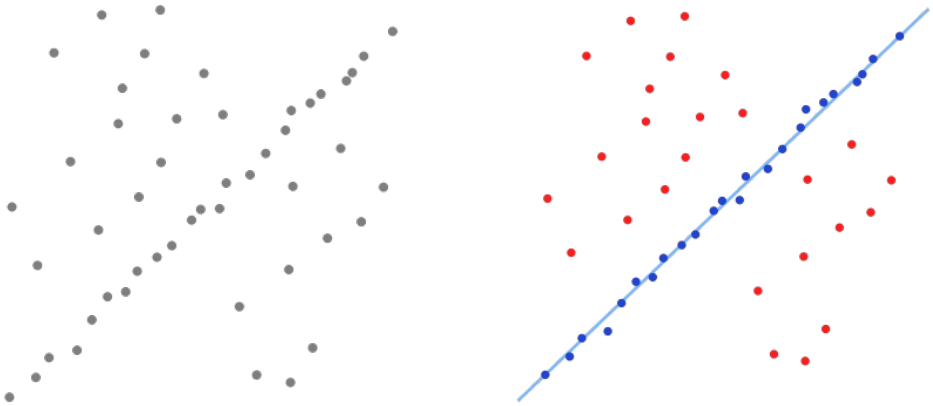
출처: https://en.wikipedia.org/wiki/Harris_affine_region_detector



Random sample consensus (RANSAC)

Random sample consensus (RANSAC) is an iterative method to estimate parameters.

참조: https://en.wikipedia.org/wiki/Random_sample_consensus



The RANSAC algorithm is a **learning technique to estimate parameters of a model** by random sampling of observed data.

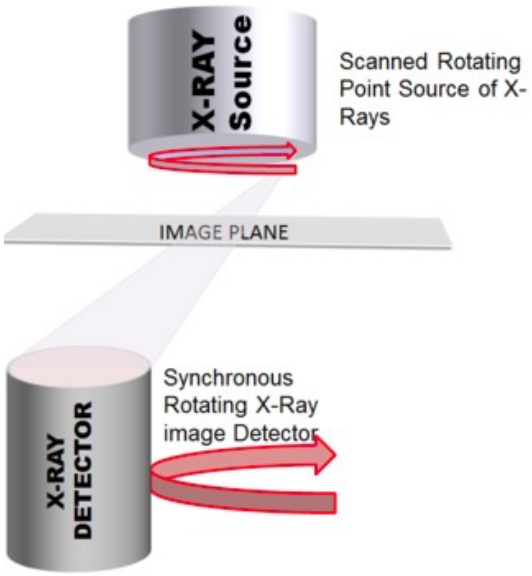
7. Commercial computer vision systems

- 5DX
- Aphelion (software)
- Microsoft PixelSense
- Poseidon drowning detection system
- Visage SDK

5DX

The 5DX was an automated X-ray inspection robot.

참조: https://en.wikipedia.org/wiki/5DX



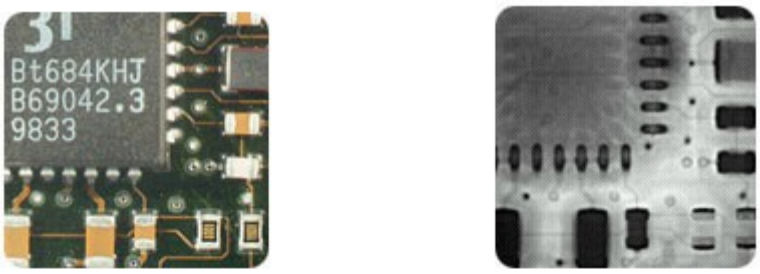


Image of Circuit Board Side A - Optical Image and X-ray Image Using 3D X-ray Inspection (5DX)

Image of Circuit Board Side B - Optical Image and X-ray Image Using 3D X-ray Inspection (5DX)

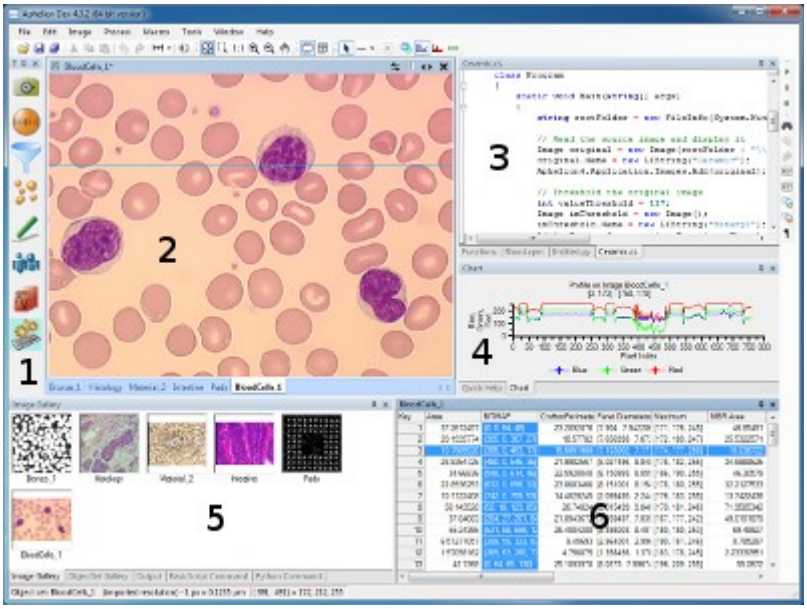
Composite Optical Image and 2D X-ray Image

출처: https://www.keysight.com/main/editorial.jspx?cc=TW&lc=cht&ckey=216395&id=216395

Aphelion (software)

image processing and image analysis applications.

참조: https://en.wikipedia.org/wiki/Aphelion_(software)



Aphelion Dev Graphical User Interface version 4.x: (1) Task Bar, (2) Image Display, (3) Macro editing window/Function window, (4) Charts (a profile is displayed in this example), (5) Image Gallery, (6) Measurement grid.

Microsoft PixelSense

Interactive surface computing platform that allows one or more people to use and touch real-world objects, and share digital content at the same time

참조: https://en.wikipedia.org/wiki/Microsoft_PixelSense

Visage SDK

Visage|SDK allows software programmers to build a wide variety of **face and head tracking** and **eye tracking applications** for various operating systems, mobile and tablet environments, and embedded systems, using computer vision and machine learning algorithms.

참조: https://en.wikipedia.org/wiki/Visage_SDK

http://visagetechnologies.com/

8. Applications

참고(출처): https://en.wikipedia.org/wiki/Outline_of_computer_vision

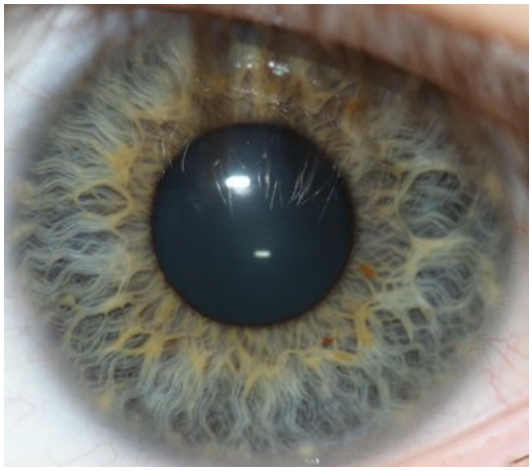
3D reconstruction from multiple images

Augmented reality (AR)

Automatic number plate recognition

Gesture recognition

Iris recognition



Object detection