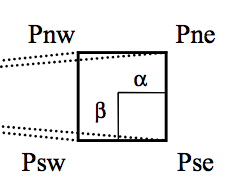
IMG PROCESSING

**. From Signal to Digital Image**

* 2d signal i=f(x,y) => image f(N,N) of pixel
* Sampling: digitalization of spatial coordinate (continuous in origin)
* Quantization: digitalization of pixel amplitude (intensity in defined coordinate spatial at time t)
* Bad sampling => aliasing
* Rotation sometimes is **not revertible** because of interpolation
* Interpolation
  + Sub-pixel precision: I(P) = I(Pse) + alpha (P(Psw-Pse)) + beta (I(Pne) – I(Pse)) + alpha. Beta (Pse + Pnw – Psw – Pne)
  + Linear (Zooming)

|  |  |  |
| --- | --- | --- |
| a | (a+b)/2 | b |
| (a+c)/2 | (a+b+c+d)/4 | (b+d)/2 |
| c | (c+d)/2 | d |

* Shannon’s theorem: me
* In practice, low pass filter -> sampling
* Quantization:
  + Scalar: if (d\_i <= f(x,y) < d\_i+1) => fq(x,y)= r\_i (value between a range)
  + \*\* Granular noise: inherently exist in emited wave
* Histogram: no localization information (better to histogram background, frontground)
  + Equalization: Khoros routine, N pixels, K gray level, calculate cumulative hist CH(k). k <~ INT(CH(k) \* K-1/N)
  + Thresholding: change to binary image (2 levels)

**. Filtering**

* Fourier Transform (use log version to get better result)
  + Spatial -> Spectral, rotate -> rotate, transition -> not change, scale -> scale inverse
* Low frequency: uniform areas, High: edge or noise
* Maxtrix multiplication in Fourier Frequency (2 matrix have same size) or convolution in Spatial
* Convolution:
* Average: Low-pass filter
* Median: non-linear(sorting), no create new value, remove extreme (reserve edge, remove noise)

**. Edge detection**

* f(x,y); Gradient ; Laplacian:
* => **edge “point”**
* **Gradient based: SURF, FAST, DAISY, SIFT**
* **Prewitt** filter: computing the gradient for the image intensity function. Makes use of the maximum directional gradient. very sensitive to noise.
* **Sobel** filter: Detects edges are where the gradient magnitude is high. This makes the Sobel edge detector more sensitive to diagonal edge than horizontal and vertical edges.
* , **edge point:** zero crossing-point
* Local binary pattern => 00110011. LBP =
* Hough Transform: designed line detection.
  + 1-to-n or ax+by=0 for line, for circle ( 1 point -> 1 curve in new coordinate)
  + n-to-1

**. Segmentation**

detect and merge homogenous zone -> split until reaching a given uniformity criterion and merge neighbouring areas to a similarity criteria

* Split:
* Merge: start from the lowest.
  + Check criteria after merge: SED(L)

**. Mathematical Morphology**

Pre-processing for binary edge map to remove outliner: We have object A, Structure element B\_p

* erosion: the whole form of matrix fit in figure
* dilation: at least of 1 px in the form of matrix must belong to figure.
* erosion is not an inverse of dilation. Dilation -> erosion -> same origina => inverse
* opening: erosion then dilation closing: dilation then erosion
* duality:

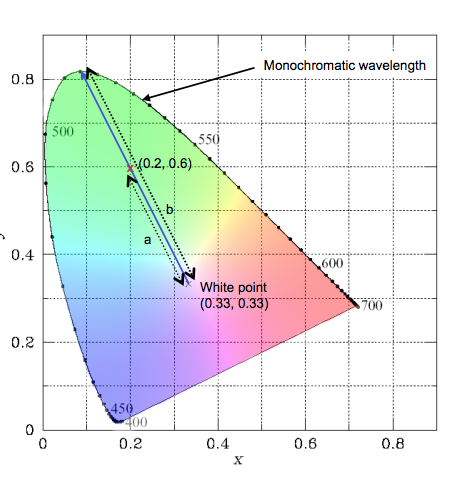
**. Motion estimation (optical flow)**

* assume that the intensity of pixel still remains constant over time
* +

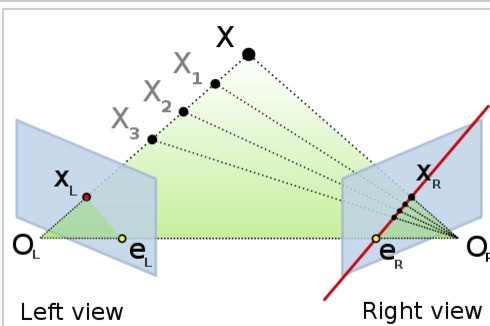
**. Color**

* Perception: source, charasteristic of object, human eyes and human brain
* Visible spectrum: <1% sun spectrum , color by weightlength and brightness by amplitude
* Light source defined by temperature and spectral power distribution.
* Source temperature: at specificed T, source emit spectrum rays with different frequency. The most frequency emitted by source based on absolute object’s temperature
* Retina have 2 cell typess: RODs (120 millions – around vision area) brightness, illumination. CONES (5 million, distribute most in fovea) sensitive to color/ frequency. In CONE, 64% RED, 32% GREEN, 4% BLUE (no blue at centre)
* **Metamer:** different light specture produce same color ( eye perception)
* 11 culture color: black, white, red, green, yellow, blue, brown, purple, pink, orange, gray
* **Grassman’s law: (additive color mixing)**
  + Law 1: Any color C can matched by linear combination of 3 others color.
  + Law2: Mixture of 2 color can matched by linearly adding their component
  + Law3: Proportionality
  + Not work in practice because negative weight
* Subtractive color mixing (yellow, cyan, megenta): start with white and substract color. Used in printing and photography
* white object will have equal values of reflectivity for each primary color: R = G = B

**. Colour Mode**

* 3 manner to classify:
  + Visual: Munshell, Chevreul, Ostwald
  + Physical: RGB, HSV, CIE XYZ,…
  + Physical and psychometrical: CIE Lab, CIE Luv,..
* RGB color space: cube => RGB triangle (**Maxwell)** along gray-scale axis
  + Some color is combination of negative contribution
  + Some color is not represented in RBG
  + Can’t not use in printing (not get right color)
* **Munsell** based on value (brightness/illumination), hue (colour: red, green, yellow, blue, purple with 2nd dimension between colour), Chroma/saturation (purity of the hue).
* **HSV** (Hue-Saturation-Value)/ **HLS** (Hue-Lightness-Saturation) Value = square root of Lightness
* **CIE XYZ:** linear combination of RGB (some colour doesn’t exist in visible spectrum) => more standardized
* **CIE xyY chromatically diagram:** idea from divided into 2 part: chromaticity and brightness
  + xyY is a color space in a surface that is away from origin by 1
  + we can derive X,Z from xyY:
  + White point (0.33,0.33). Draw a line from white through your color point
    - Boundary: weightlength
    - Percentage of distance from point to edge: purity(saturation)
    - P\_1 and P\_2 complement when line cross white point
  + Additive color mixing
    - 2 color:
    - 3 color:
    - in case negative distribution, P point will be outside triangle so distance from P to specific edge will <0
* **Mac Adam’s ellipses:** Ellipse = minimum perceivable color difference (color in ellipse is not different in human perception)
* **USC diagram:**
* **Lab color space:** L is luminosity, a\* is red/green axis, b\* is yellow/blue axis. Metrics is representative of mathematically distance are perceptual distance
* **YIQ color:** commercial color TV. **Y: illumination** and I,Q: chromaticity.More bit to encode Y (people are more sensitive with illumination), less bit to encode I,Q
* **YCrCb (YUV)** : PAL analog video, CCIR 601 standard for digital video
  + Eye is not sensitive to green => can accept some error in calculation
  + Scale and filtered version: U=0.5(B-Y),V=0.877(R-Y)

**. 3D**

* The goal: extract 3D information of a scene from multi-views
* Active:
  + Time of flight: 1 IR emitter + 1 detector, indoor & outdoor, fast & precise but low-resol and expensive
    - Continuous wave modulation: frequency vary by time, estimation based on phase shift, measured between receiver and emitter
    - Pulsed Wave modulation: estimate distance by computing time needed for light to travel
  + Structured light: Kinect, indoor only, measure distort of IR, fast and precise
* Passive, common and popular: Stereo setup: 2 calibrated camera, triangle like human eyes, not very precise
* Depth map from **stereo vision system** if 2 video camera are parallels:
* Epipolar geometry: in case your camera are not parallel, and you know your projection of X on left view. So projection of X on right view is only on a line e\_r x\_r