

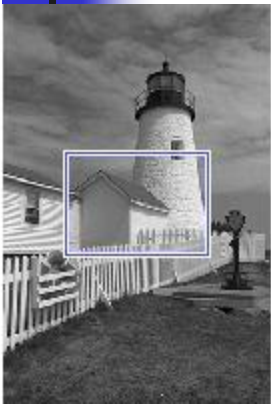


# Fundamentals of Image Processing and Computer Vision

---

**Jayanta Mukhopadhyay**  
**Dept. of Computer Science and Engg.**  
**Indian Institute of Technology, Kharagpur**

# How images are represented in computer?

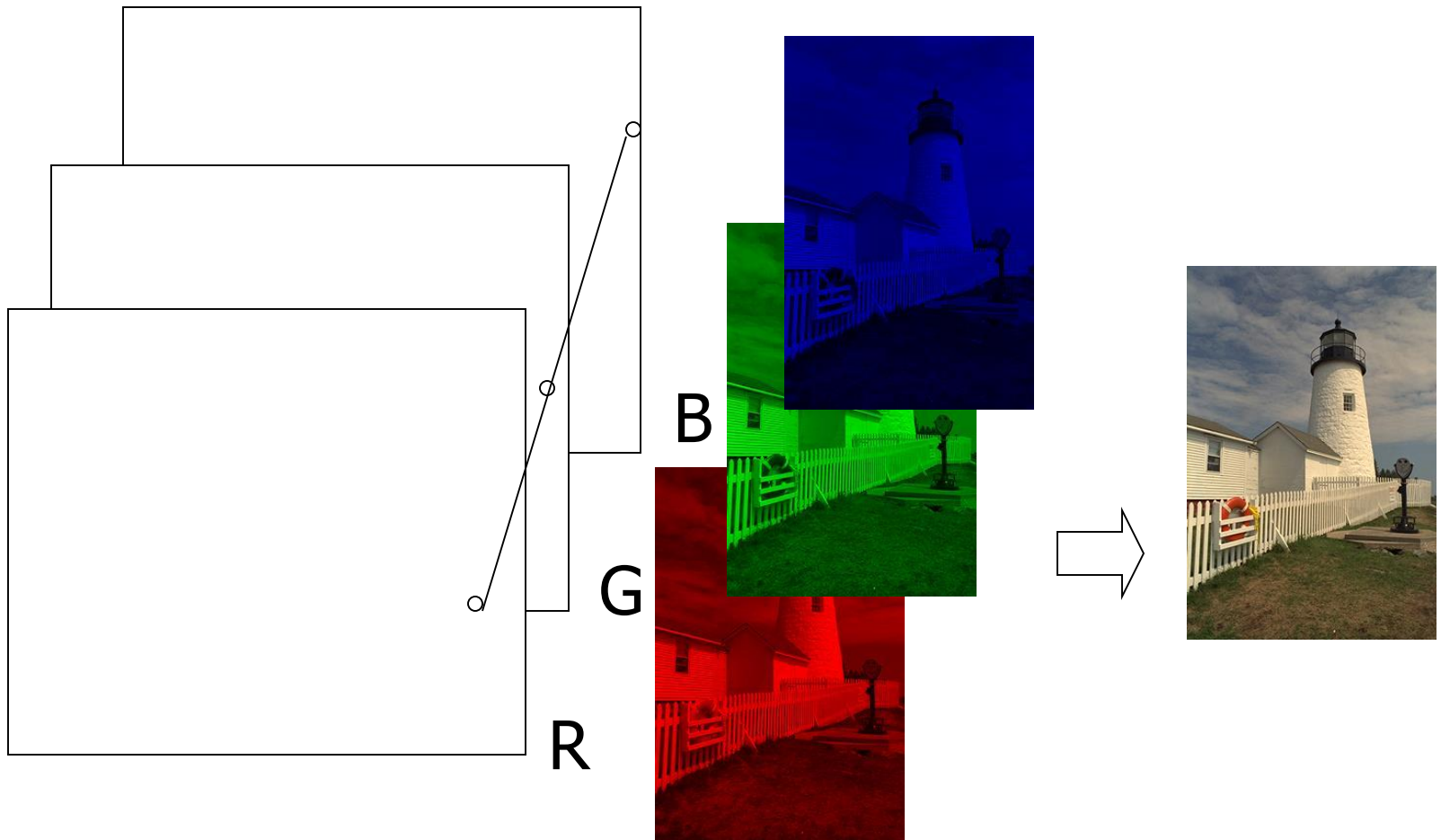


Width: 256  
Height: 384

9	95	95	103	88	73	74	89	103	101	97	96	96	99	98	9
1	107	90	103	118	122	115	99	85	78	81	91	97	100	97	9
4	73	88	196	175	129	118	120	123	116	103	85	75	73	81	9
3	81	67	157	220	226	217	184	152	130	128	125	125	111	95	8
9	92	88	71	75	104	147	187	219	227	209	170	138	129	133	1
0	102	104	100	93	84	72	69	87	122	165	208	225	213	185	1
0	110	122	114	111	108	103	101	95	81	73	79	105	147	190	2
3	119	137	127	125	121	117	114	111	108	107	103	92	83	83	9
3	123	147	141	137	134	130	127	122	121	121	119	114	113	110	9
7	126	150	148	148	143	143	141	135	135	134	127	126	123	123	1

A 2D Array of integers

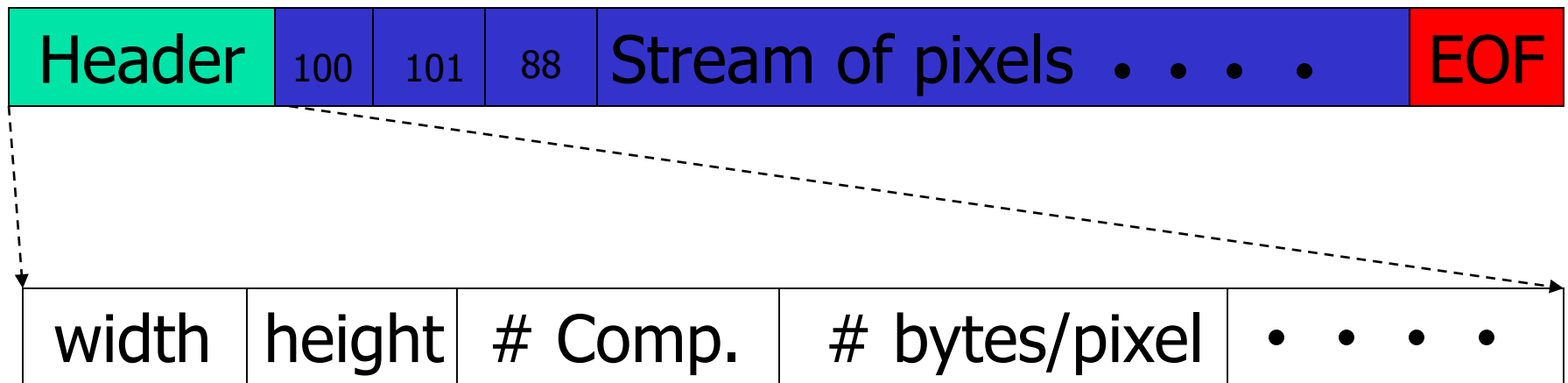
# Color images





# Image File

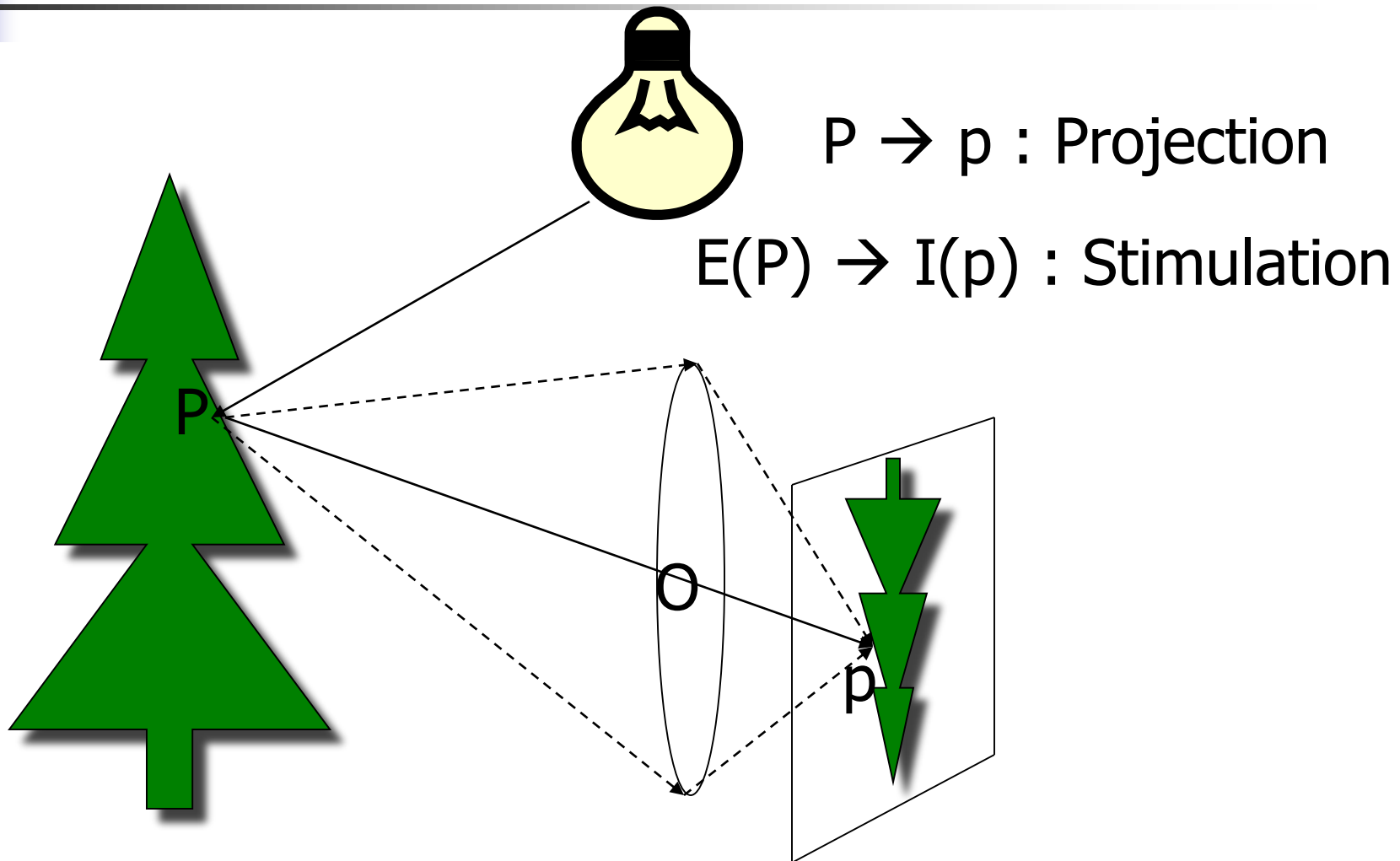
---



Standard File Formats:

TIFF, BMP, GIF, PGM, PPM, JPEG, DICOM, ....

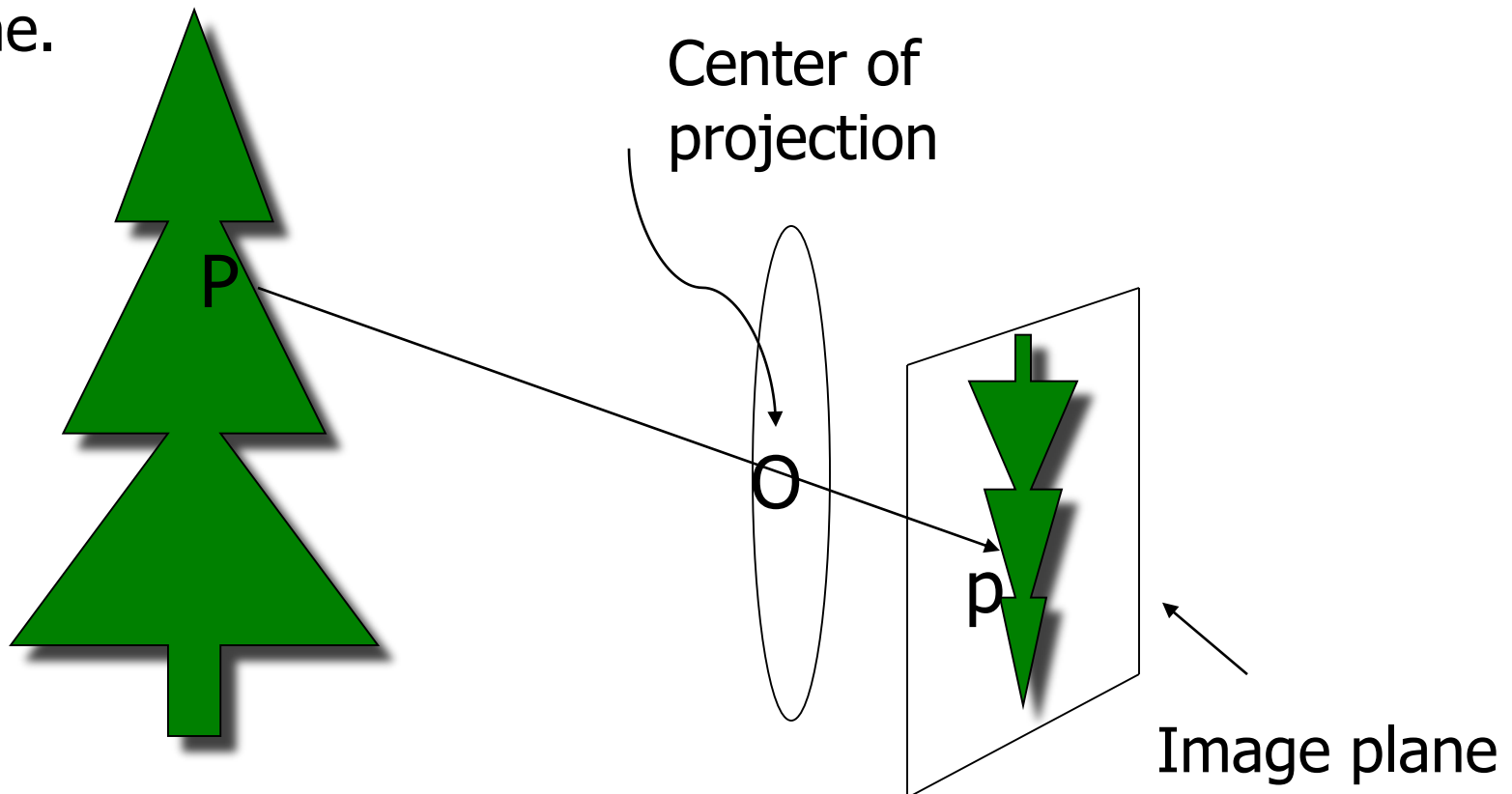
# Image formation in optical camera



# Rules of projection

## Perspective projection

Image points formed by intersection of the ray from a point  $P$  passing through the center of projection  $O$  with the image plane.



Parallel  
projection

# Other imaging principle

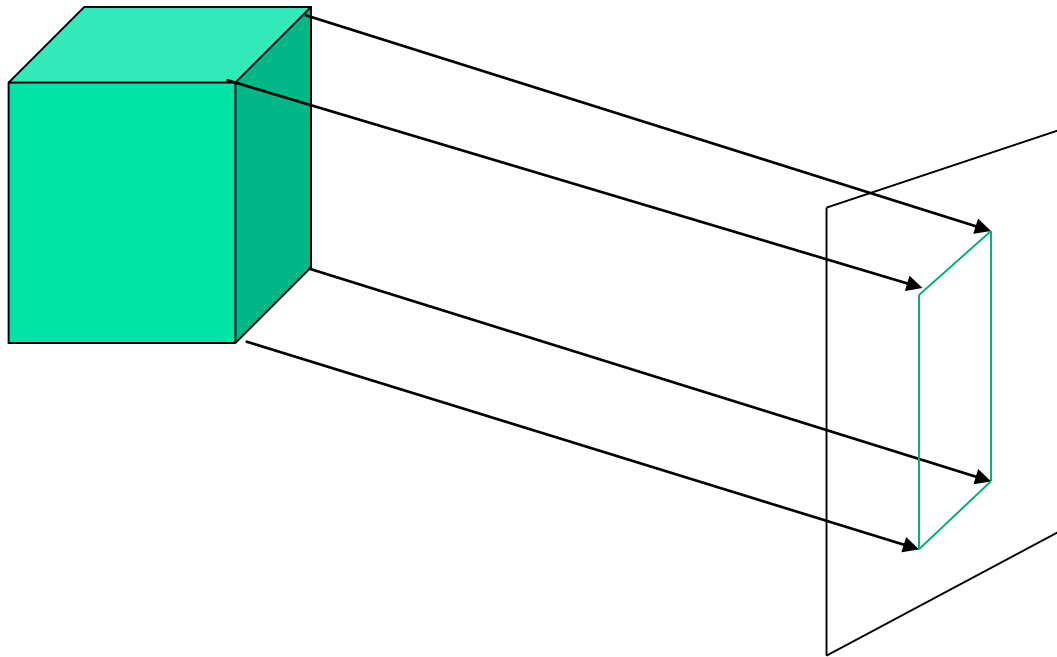
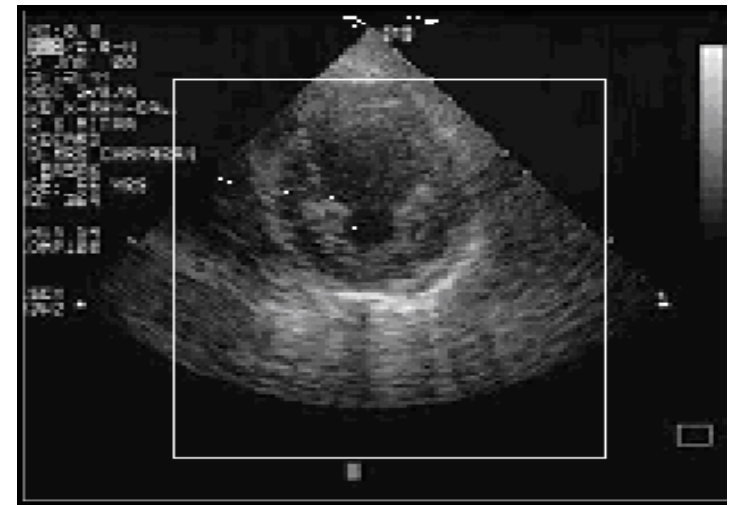
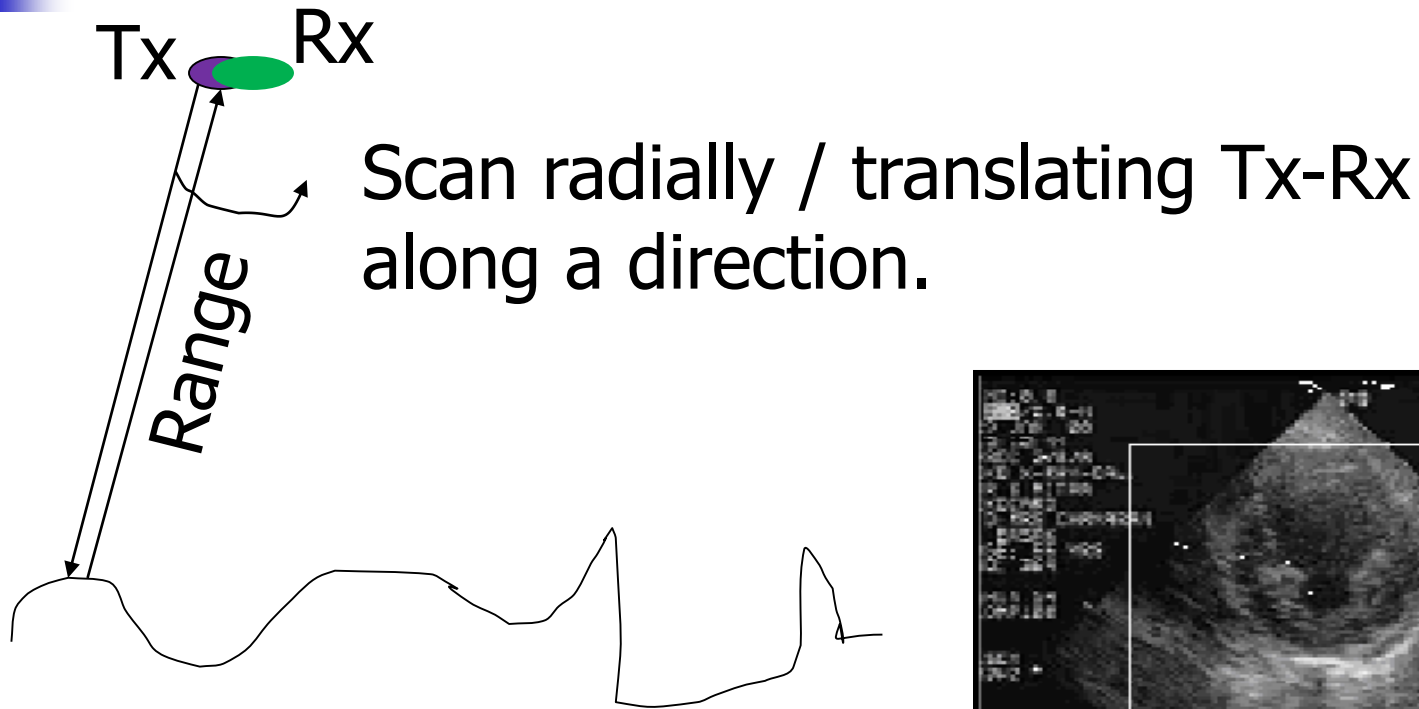


Image points formed by intersection of parallel rays with the image plane.

# Other imaging principle





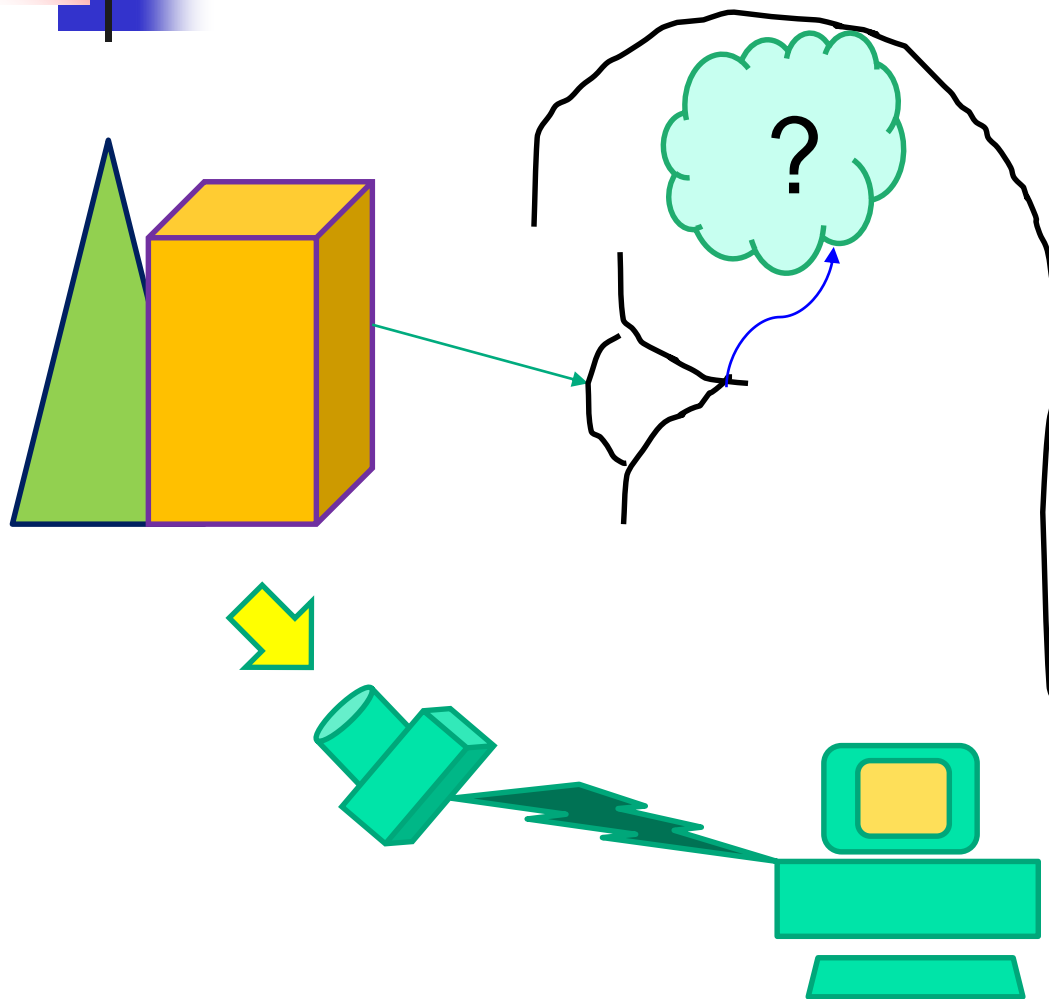


# What is an image?

---

- Impression of the physical world.
- Spatial distribution of a measurable quantity, encoding the geometry and material properties of objects.

# What is Computer Vision?



*The science of facilitating human-like capability of seeing and understanding the environment to a machine or computer.*



# What is Computer Vision?

---

- Concerned with the automatic extraction, analysis and understanding of useful information from a single image, a set of images, or a sequence of images.
- Involving the development of a theoretical and algorithmic basis to achieve automatic visual understanding.



# A few typical problems

---

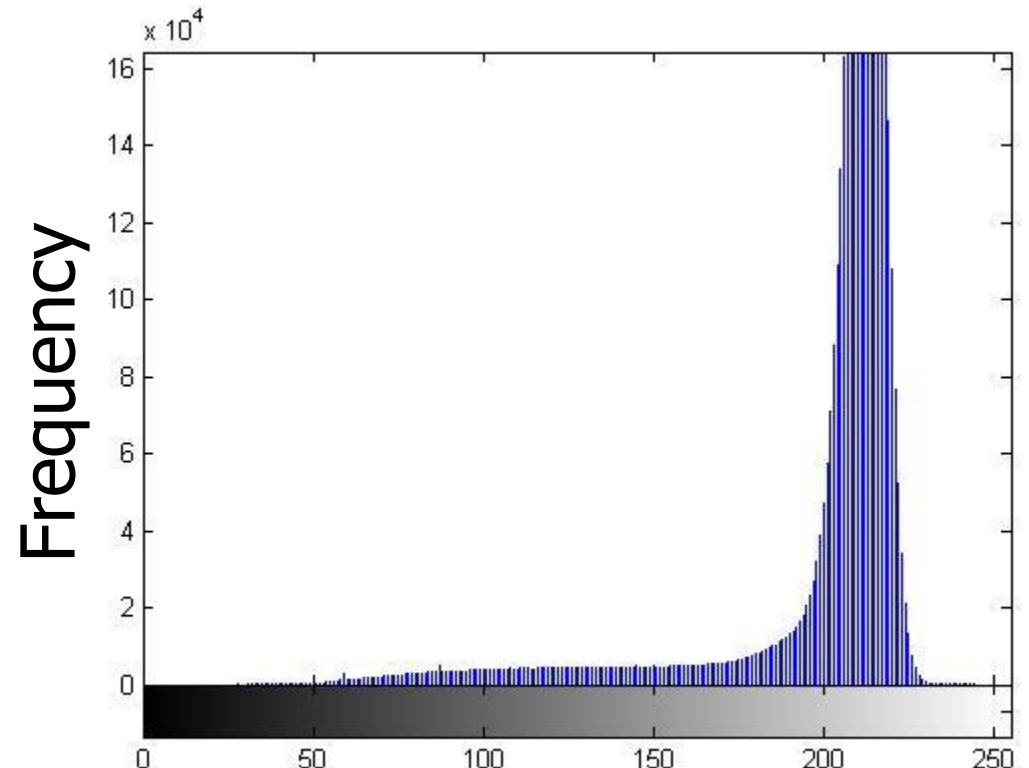
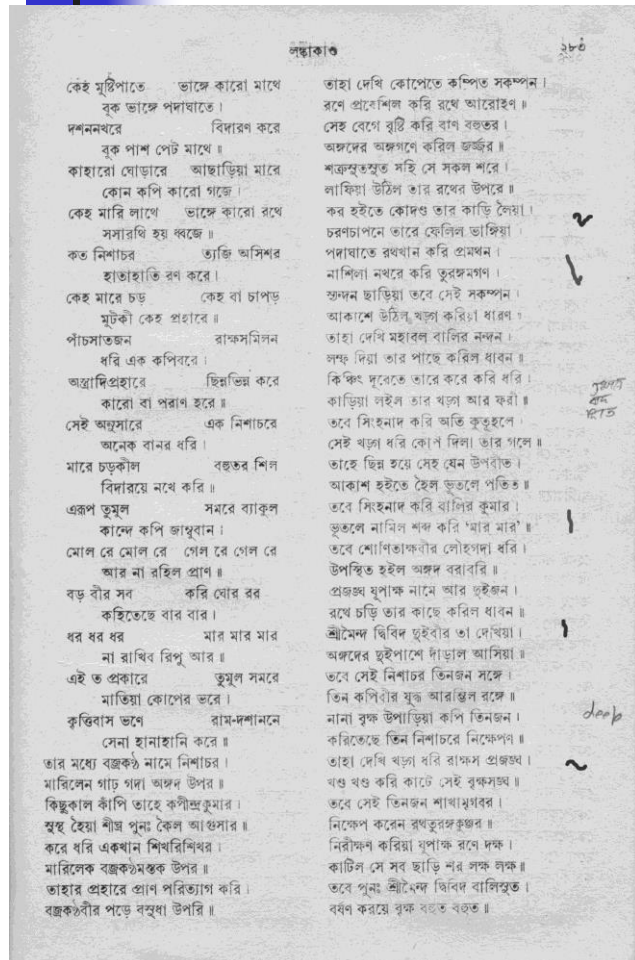
- To determine structure / geometry of objects.
- To determine material properties of objects.
- To determine object's position, category, and its role / interaction with other objects.
- To aid / enhance our perception of the physical world.



---

# A few concepts and operations on digital images

# Image histogram



Normalized Histogram  $\rightarrow p(x)$



# Binarization (Thresholding)

---

- Separation of foreground from background.
  - Foreground: Dark pixels (text)
  - Background: Bright pixels (white region)
- Pixels in a binarized image set to one of the two values (e.g. 255 or 0).
  - 255 for background.
  - 0 for foreground.



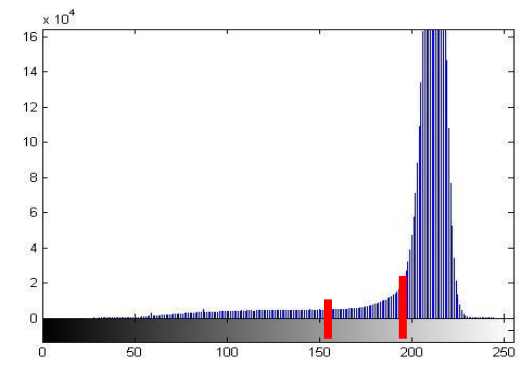
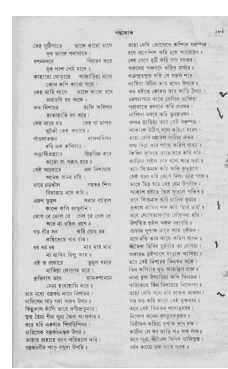
# Binarization (Thresholding)

---

- A simple algorithm
  - Choose a threshold value  $T$ .
  - A pixel greater than  $T$  is set to 255 (background), otherwise to 0 (foreground).



# Thresholding



Can you automate this operation?

লঙ্কাভাণ্ড

২৮৬

কেই মুষ্টিপাতে ভাসে কারো মাথে  
বুক ভাসে পদাধাতে।  
দর্শনমথরে বিদারণ করে  
বুক পাশ পেট মাথে।  
কাহারো যোড়ারে আছাড়িয়া মারে  
কোন কপি কারো গাড়ে।  
কেহ মারি লাগে ভাসে কারো রথে  
সদারণি হয় ধ্বজে।  
কত নিশাচর তাজি অসিপর  
হাভাহাতি রণ করে।  
কেহ মারে চড় কেহ বা চাপড়  
মুটকী কেহ প্রহারে।  
পাঁচাত্তজন রাক্ষসমিলন  
ধরি এক কপিবারে।  
অগ্নিপ্রহারে ছিন্নভিন্ন করে  
কারো বা পরাণ হরে।  
সেই অহুসারে এক নিশাচরে  
অনেক বানর ধরি।  
মারে চড়কিল বহুতর শিল  
বিদারয়ে নখে করি।  
এরূপ তুমুল সমরে ব্যাকুল  
কাপে কপি জাম্বুবান।  
মোল রে মোল রে গেল রে গেল রে  
আর না রহিল প্রাণ।  
বড় বীর সব করি ঘোর রর  
কহিতেছে বার বার।  
ধর ধর ধর মার মার মার  
না রাখিব রিপু আর।  
এই ত প্রকারে তুমুল সমরে  
মতিয়া কোপের ভরে।  
কৃতিবাস ভণে রাম-দশাননে  
সেনা হানাহানি করে।  
তার মধ্যে বজ্রকণ্ঠ নামে নিশাচর।  
মারিলেন গাট গরা অঙ্গন উপর।  
কিছুকাল কাঁপি তাহে কপীপ্রহুমার।  
হুহু হৈয়া শীত পুনঃ কৈল আগুনার।  
করে ধরি একখান শিখরিশিখর।  
মারিলেক বজ্রকণ্ঠমস্তক উপর।  
ভাহার প্রহারে প্রাণ পরিত্যাগ করি।  
বজ্রকণ্ঠবীর পড়ে লম্বা উপরি।

তাহা দেখি কোপেতে কণ্ঠিত সঙ্কম্পন।  
রণে প্রবেশিল করি রথে আরোহণ।  
সেহ বেগে বুরি করি বাণ বহুতর।  
অঙ্গদের অঙ্গগণে করিল ছাড়র।  
শক্রসুতমুত সহি সে সকল শরে।  
লাফিয়া উঠিল তার রথের উপরে।  
কর হইতে কোদণ্ড তার কাড়ি লৈয়া।  
চরণচাপনে তারে ফেলিল ভাঙ্গিয়া।  
পদাধাতে রথখান করি প্রমথন।  
নাশিলা নথরে করি তুরঙ্গমণ।  
স্বদন ছাড়িয়া তবে সেই সঙ্কম্পন।  
আকাশে উঠিল খড়্গা করিয়া ধারণ।  
তাহা দেখি মহাবল বাসির নন্দন।  
লক্ষ দিয়া তার পাছে করিল ধাবন।  
কিঞ্চিৎ দূরেতে তারে করে করি ধরি।  
কাড়িয়া লইল তার খড়া আর ফরা।  
তবে সিংহনাদ করি অতি ক্রুদ্ধহলে।  
সেই খড়া ধরি কোণ দিলা তার গলে।  
তাহে ছিন্ন হয়ে সেই বেন উপবীত।  
আকাশ হইতে হৈল ভূতলে পতিত।  
তবে সিংহনাদ করি বাসির কুমার।  
ভূতলে নামিল শব্দ করি 'মার মার'।  
তবে শোণিতাকবীর লৌহগদা ধরি।  
উপস্থিত হইল অঙ্গদ বরাবরি।  
প্রজন্ম যুগাৎ নামে আর দুইজন।  
রথে চড়ি তার কাছে করিল ধাবন।  
ঐন্মৈন্দ খিবিদ দুইবার তা দেখিয়া।  
অঙ্গদের দুইপাশে পাড়াল আসিয়া।  
তবে সেই নিশাচর ভিনজন সঙ্গে।  
তিন কপিবীর যুদ্ধ আরম্ভিল সঙ্গে।  
নানা বৃক্ষ উপাড়িয়া কপি ভিনজন।  
করিতেছে তিন নিশাচরে নিক্ষেপণ।  
তাহা দেখি খড়া ধরি রাক্ষস প্রজন্ম।  
খণ্ড খণ্ড করি কাটে সেই বৃক্ষসজ্জ।  
তবে সেই ভিনজন শাখাশূন্যবর।  
নিক্ষেপ করেন রথতরঙ্গকুঞ্জর।  
নিরীক্ষণ করিয়া যুগাৎ রণে দক্ষ।  
কাটিল সে সব ছাড়ি শর লক্ষ লক্ষ।  
তবে পুনঃ ঐন্মৈন্দ খিবিদ বাসিমুত।  
বর্ধন করয়ে বৃক্ষ বহুত বহুত।

লঙ্কাভাণ্ড

২৮৬

কেই মুষ্টিপাতে ভাসে কারো মাথে  
বুক ভাসে পদাধাতে।  
দর্শনমথরে বিদারণ করে  
বুক পাশ পেট মাথে।  
কাহারো যোড়ারে আছাড়িয়া মারে  
কোন কপি কারো গাড়ে।  
কেহ মারি লাগে ভাসে কারো রথে  
সদারণি হয় ধ্বজে।  
কত নিশাচর তাজি অসিপর  
হাভাহাতি রণ করে।  
কেহ মারে চড় কেহ বা চাপড়  
মুটকী কেহ প্রহারে।  
পাঁচাত্তজন রাক্ষসমিলন  
ধরি এক কপিবারে।  
অগ্নিপ্রহারে ছিন্নভিন্ন করে  
কারো বা পরাণ হরে।  
সেই অহুসারে এক নিশাচরে  
অনেক বানর ধরি।  
মারে চড়কিল বহুতর শিল  
বিদারয়ে নখে করি।  
এরূপ তুমুল সমরে ব্যাকুল  
কাপে কপি জাম্বুবান।  
মোল রে মোল রে গেল রে গেল রে  
আর না রহিল প্রাণ।  
বড় বীর সব করি ঘোর রর  
কহিতেছে বার বার।  
ধর ধর ধর মার মার মার  
না রাখিব রিপু আর।  
এই ত প্রকারে তুমুল সমরে  
মতিয়া কোপের ভরে।  
কৃতিবাস ভণে রাম-দশাননে  
সেনা হানাহানি করে।  
তার মধ্যে বজ্রকণ্ঠ নামে নিশাচর।  
মারিলেন গাট গরা অঙ্গন উপর।  
কিছুকাল কাঁপি তাহে কপীপ্রহুমার।  
হুহু হৈয়া শীত পুনঃ কৈল আগুনার।  
করে ধরি একখান শিখরিশিখর।  
মারিলেক বজ্রকণ্ঠমস্তক উপর।  
ভাহার প্রহারে প্রাণ পরিত্যাগ করি।  
বজ্রকণ্ঠবীর পড়ে লম্বা উপরি।

তাহা দেখি কোপেতে কণ্ঠিত সঙ্কম্পন।  
রণে প্রবেশিল করি রথে আরোহণ।  
সেহ বেগে বুরি করি বাণ বহুতর।  
অঙ্গদের অঙ্গগণে করিল ছাড়র।  
শক্রসুতমুত সহি সে সকল শরে।  
লাফিয়া উঠিল তার রথের উপরে।  
কর হইতে কোদণ্ড তার কাড়ি লৈয়া।  
চরণচাপনে তারে ফেলিল ভাঙ্গিয়া।  
পদাধাতে রথখান করি প্রমথন।  
নাশিলা নথরে করি তুরঙ্গমণ।  
স্বদন ছাড়িয়া তবে সেই সঙ্কম্পন।  
আকাশে উঠিল খড়্গা করিয়া ধারণ।  
তাহা দেখি মহাবল বাসির নন্দন।  
লক্ষ দিয়া তার পাছে করিল ধাবন।  
কিঞ্চিৎ দূরেতে তারে করে করি ধরি।  
কাড়িয়া লইল তার খড়া আর ফরা।  
তবে সিংহনাদ করি অতি ক্রুদ্ধহলে।  
সেই খড়া ধরি কোণ দিলা তার গলে।  
তাহে ছিন্ন হয়ে সেই বেন উপবীত।  
আকাশ হইতে হৈল ভূতলে পতিত।  
তবে সিংহনাদ করি বাসির কুমার।  
ভূতলে নামিল শব্দ করি 'মার মার'।  
তবে শোণিতাকবীর লৌহগদা ধরি।  
উপস্থিত হইল অঙ্গদ বরাবরি।  
প্রজন্ম যুগাৎ নামে আর দুইজন।  
রথে চড়ি তার কাছে করিল ধাবন।  
ঐন্মৈন্দ খিবিদ দুইবার তা দেখিয়া।  
অঙ্গদের দুইপাশে পাড়াল আসিয়া।  
তবে সেই নিশাচর ভিনজন সঙ্গে।  
তিন কপিবীর যুদ্ধ আরম্ভিল সঙ্গে।  
নানা বৃক্ষ উপাড়িয়া কপি ভিনজন।  
করিতেছে তিন নিশাচরে নিক্ষেপণ।  
তাহা দেখি খড়া ধরি রাক্ষস প্রজন্ম।  
খণ্ড খণ্ড করি কাটে সেই বৃক্ষসজ্জ।  
তবে সেই ভিনজন শাখাশূন্যবর।  
নিক্ষেপ করেন রথতরঙ্গকুঞ্জর।  
নিরীক্ষণ করিয়া যুগাৎ রণে দক্ষ।  
কাটিল সে সব ছাড়ি শর লক্ষ লক্ষ।  
তবে পুনঃ ঐন্মৈন্দ খিবিদ বাসিমুত।  
বর্ধন করয়ে বৃক্ষ বহুত বহুত।

Th=156

Th=192

# Bayesian Classification of foreground and background pixels

Pixels belonging to two classes:

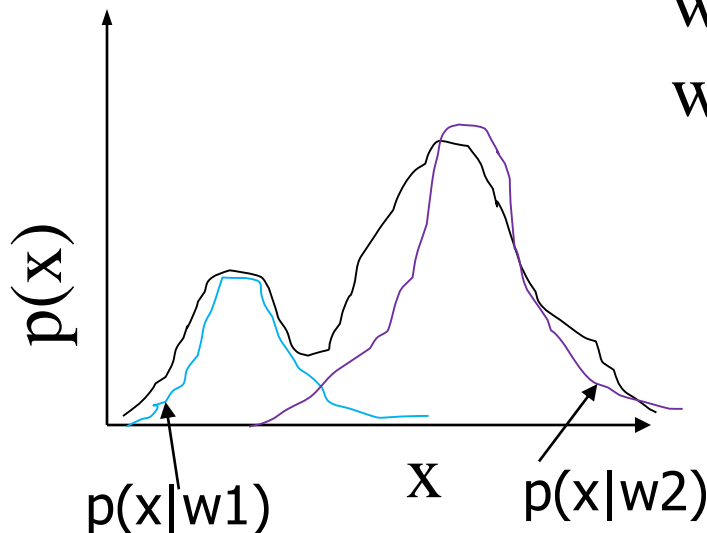
$w_1$ : Foreground

$w_2$ : Background

Compute  $p(w_1|x)$  and  $p(w_2|x)$ .

Bayes' theorem:

$$p(\omega | x) = \frac{p(\omega) p(x | \omega)}{p(x)}$$



Bayes' classification rule:

Assign  $x$  to  $w_1$  if  $p(w_1|x) > p(w_2|x)$  else to  $w_2$ .



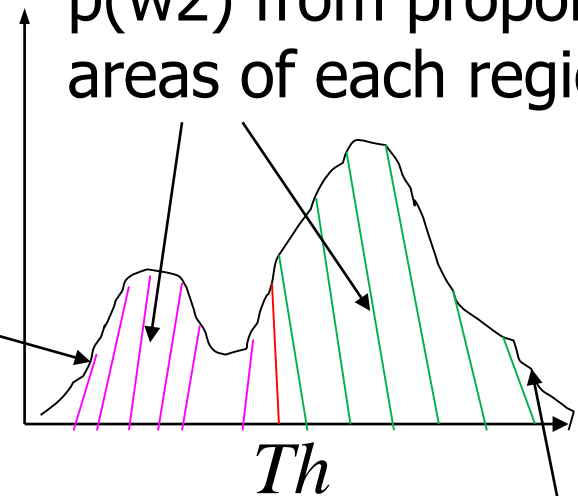
To check whether  $p(w_1)p(x|w_1) > p(w_2)p(x|w_2)$

# Expectation-Maximization Algorithm

2. Compute parameters of  $p(x|w_1)$  by assuming it Gaussian.

$\mu_1, \sigma_1$

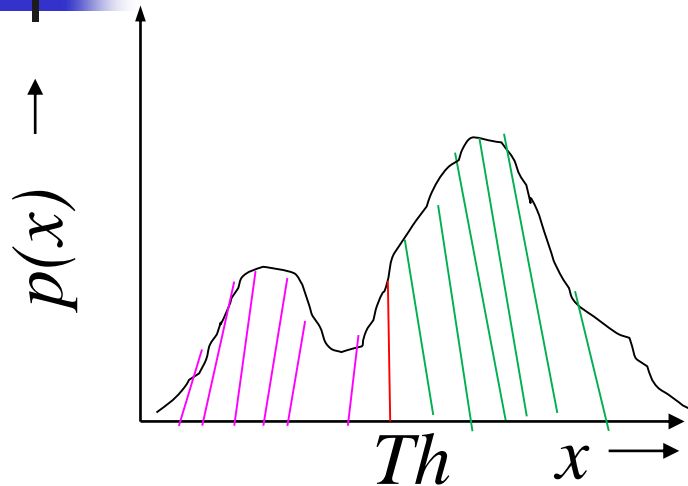
1. Compute  $p(w_1)$  and  $p(w_2)$  from proportional areas of each region.



3. Compute parameters of  $p(x|w_2)$  by assuming it Gaussian.  $\mu_2, \sigma_2$

$$N(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

# Expectation-Maximization Algorithm



4. Compute new threshold value so that for  $x < Th$ ,  $p(w1/x) > p(w2/x)$ , and vice versa.
5. Iterate till the value of  $Th$  converges.

$$p(w1) = \sum_{x=0}^{Th} p(x)$$

$$p(w2) = 1 - p(w1)$$

$$\mu_1 = \sum_{x=0}^{Th} x \cdot p(x)$$

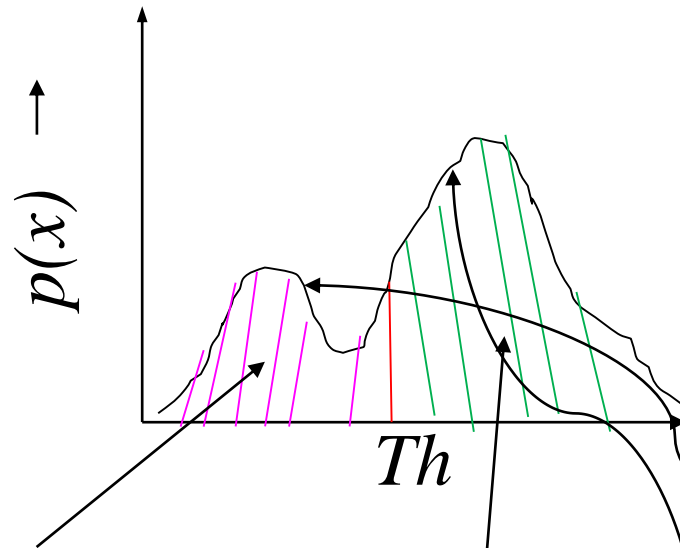
$$\sigma_1^2 = \sum_{x=0}^{Th} x^2 \cdot p(x) - \mu_1^2$$

$$\mu_2 = \sum_{x=Th+1}^{255} x \cdot p(x)$$

$$\sigma_2^2 = \sum_{x=Th+1}^{255} x^2 \cdot p(x) - \mu_2^2$$

# Otsu Thresholding

- Choose a threshold value, which maximizes between class variance ( $\sigma_B^2$ ).



$$\sigma_B^2 = p(w1)p(w2)(\mu_2 - \mu_1)^2$$

# An example

Otsu:157

লঙ্কা৩

২৮৬

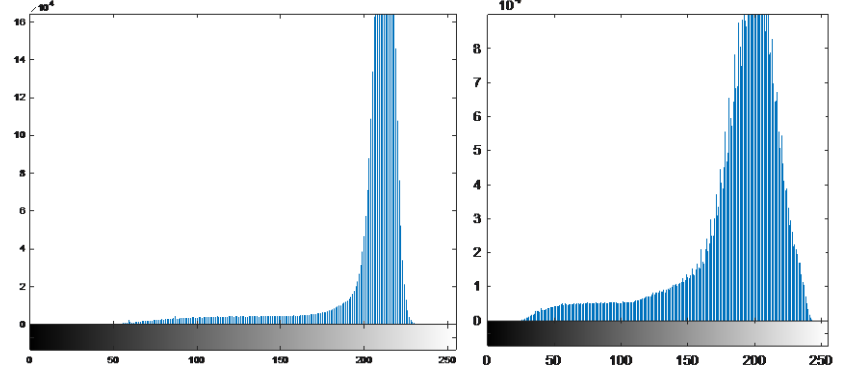
কেই মুষ্টিপাতে ভাঙ্গে কারো মাথে  
বুক ভাঙ্গে পদাঘাতে ।  
দশননখরে বিদারণ করে  
বুক পাশ পেট মাথে ॥  
কাহারো ঘোড়ারে আছাড়িয়া মারে  
কোন কপি কারো গজে ।  
কেহ মারি লাখে ভাঙ্গে কারো মাথে  
সমারথি হয় ধ্বজে ॥  
কত নিশাচর তাজি অসিশর  
হাতাহাতি রণ করে ।  
কেহ মারে চড় কেহ বা চাপড়  
মুটকী কেহ প্রহারে ॥  
পাঁচসাতজন রাক্ষসমিলন  
ধরি এক কপিবারে ।  
অগ্নাদিপ্রহারে ছিন্নভিন্ন করে  
কারো বা পরাণ হরে ॥  
সেই অমুসারে এক নিশাচরে  
অনেক বানর ধরি ।  
মারে চড়কীল বজ্রতর শিল  
বিদারয়ে নখে করি ॥  
এরূপ তুমুল সমরে ব্যাকুল  
কান্দে কপি জাণুবান ।  
মোল রে মোল রে গেল রে গেল রে  
আর না রহিল প্রাণ ॥  
বড় বীর সব করি ঘোর রর  
কহিতেছে বার বার ।  
ধর ধর ধর মার মার মার  
না রাখিব রিপু আর ॥  
এই ত প্রকারে তুমুল সমরে  
মাতিয়া কোপের ভরে ।  
কৃতিবাস ভণে রামদশনানে  
সেনা হানাহানি করে ॥  
তার মধ্যে বজ্রকণ্ট নামে নিশাচর ।  
মারিলেন গাঢ় গাণ্ডা অঙ্গদ উপর ॥  
কিছুকাল কাঁপি তাহে কপীন্দ্রকুমার ।  
বৃহ হৈয়া শীঘ্র পুনঃ কৈল আগ্রসার ॥  
করে ধরি একখান শিখরিশিখর ।  
মারিলেক বজ্রকণ্টমস্তক উপর ॥  
তাহার প্রহারে প্রাণ পরিত্যাগ করি ।  
বজ্রকণ্টবীর পড়ে বনুধা উপরি ॥

তাহা দেখি কোপেতে কল্লিত সঙ্কম্পন ।  
রণে প্রবেশিল করি রথে আরোহণ ॥  
সেহ বেগে বৃষ্টি করি বাণ বজ্রতর ।  
অঙ্গদের অঙ্গগণে করিল জর্জর ॥  
শক্রসুতসুত সহি সে সকল শরে ।  
লাফিয়া উঠিল তার রথের উপরে ॥  
কব হইতে কোদণ্ড তার কাড়ি লৈয়া ।  
চরণচাপনে তারে ফেলিল ভাঙ্গিয়া ॥  
পদাঘাতে রথখান করি প্রমথন ।  
নাশিলা নখরে করি তুরঙ্গমগণ ॥  
স্বন্দন ছাড়িয়া তবে সেই সঙ্কম্পন ।  
আকাশে উঠিল খজা করিধা ধারণ ॥  
তাহা দেখি মহাবল বালির নন্দন ।  
লক্ষ দিয়া তার পাছে করিল ধাবন ॥  
কিঞ্চিৎ দূরেতে তারে করে করি ধরি ।  
কাড়িয়া লইল তার খজা আর ফরা ॥  
তবে সিংহনাদ করি অতি কুতূহলে ।  
সেই খজা ধরি কোণ দিলা তার গলে ॥  
তাহে ছিন্ন হয়ে সেই যেন উপবীত ।  
আকাশ হইতে হৈল ভূতলে পতিত ॥  
তবে সিংহনাদ করি বালির কুমার ।  
ভূতলে নামিল শব্দ করি 'মার মার' ॥  
তবে শোণিতাকবীর লৌহগদা ধরি ।  
উপস্থিত হইল অঙ্গদ বরাবরি ॥  
প্রজন্ম যুগাৎ নামে আর দুইজন ।  
রথে চড়ি তার কাছে করিল ধাবন ॥  
শ্রীমৈন্দ দ্বিবিদ দুইবার তা দেখিয়া ।  
অঙ্গদের দুইপাশে দাঁড়াল আসিয়া ॥  
তবে সেই নিশাচর তিনজন সঙ্গে ।  
তিন কপিবার যুদ্ধ আরম্ভিল রঙ্গে ॥  
নানা বৃক্ষ উপাড়িয়া কপি তিনজন ।  
করিতেছে তিন নিশাচরে নিক্ষেপণ ॥  
তাহা দেখি খজা ধরি রাক্ষস প্রজন্ম ।  
বগু বগু করি কাটে সেই বৃক্ষসম্ম ॥  
তবে সেই তিনজন শাখাযুগবর ।  
নিক্ষেপ করেন রথতুরঙ্গর ॥  
নিরীক্ষণ করিয়া যুগাৎ রণে দক্ষ ।  
কাটিল সে সব ছাড়ি শব্দ লক্ষ লক্ষ ॥  
তবে পুনঃ শ্রীমৈন্দ দ্বিবিদ বালিব্রত ।  
বর্ধন করয়ে বৃক্ষ বজ্রত বজ্রত ॥

১৫৮৫  
১৫৮৫

১৫৮৫

Bayesian  
:157



লঙ্কা৩

২৮৬

কেই মুষ্টিপাতে ভাঙ্গে কারো মাথে  
বুক ভাঙ্গে পদাঘাতে ।  
দশননখরে বিদারণ করে  
বুক পাশ পেট মাথে ॥  
কাহারো ঘোড়ারে আছাড়িয়া মারে  
কোন কপি কারো গজে ।  
কেহ মারি লাখে ভাঙ্গে কারো মাথে  
সমারথি হয় ধ্বজে ॥  
কত নিশাচর তাজি অসিশর  
হাতাহাতি রণ করে ।  
কেহ মারে চড় কেহ বা চাপড়  
মুটকী কেহ প্রহারে ॥  
পাঁচসাতজন রাক্ষসমিলন  
ধরি এক কপিবারে ।  
অগ্নাদিপ্রহারে ছিন্নভিন্ন করে  
কারো বা পরাণ হরে ॥  
সেই অমুসারে এক নিশাচরে  
অনেক বানর ধরি ।  
মারে চড়কীল বজ্রতর শিল  
বিদারয়ে নখে করি ॥  
এরূপ তুমুল সমরে ব্যাকুল  
কান্দে কপি জাণুবান ।  
মোল রে মোল রে গেল রে গেল রে  
আর না রহিল প্রাণ ॥  
বড় বীর সব করি ঘোর রর  
কহিতেছে বার বার ।  
ধর ধর ধর মার মার মার  
না রাখিব রিপু আর ॥  
এই ত প্রকারে তুমুল সমরে  
মাতিয়া কোপের ভরে ।  
কৃতিবাস ভণে রামদশনানে  
সেনা হানাহানি করে ॥  
তার মধ্যে বজ্রকণ্ট নামে নিশাচর ।  
মারিলেন গাঢ় গাণ্ডা অঙ্গদ উপর ॥  
কিছুকাল কাঁপি তাহে কপীন্দ্রকুমার ।  
বৃহ হৈয়া শীঘ্র পুনঃ কৈল আগ্রসার ॥  
করে ধরি একখান শিখরিশিখর ।  
মারিলেক বজ্রকণ্টমস্তক উপর ॥  
তাহার প্রহারে প্রাণ পরিত্যাগ করি ।  
বজ্রকণ্টবীর পড়ে বনুধা উপরি ॥

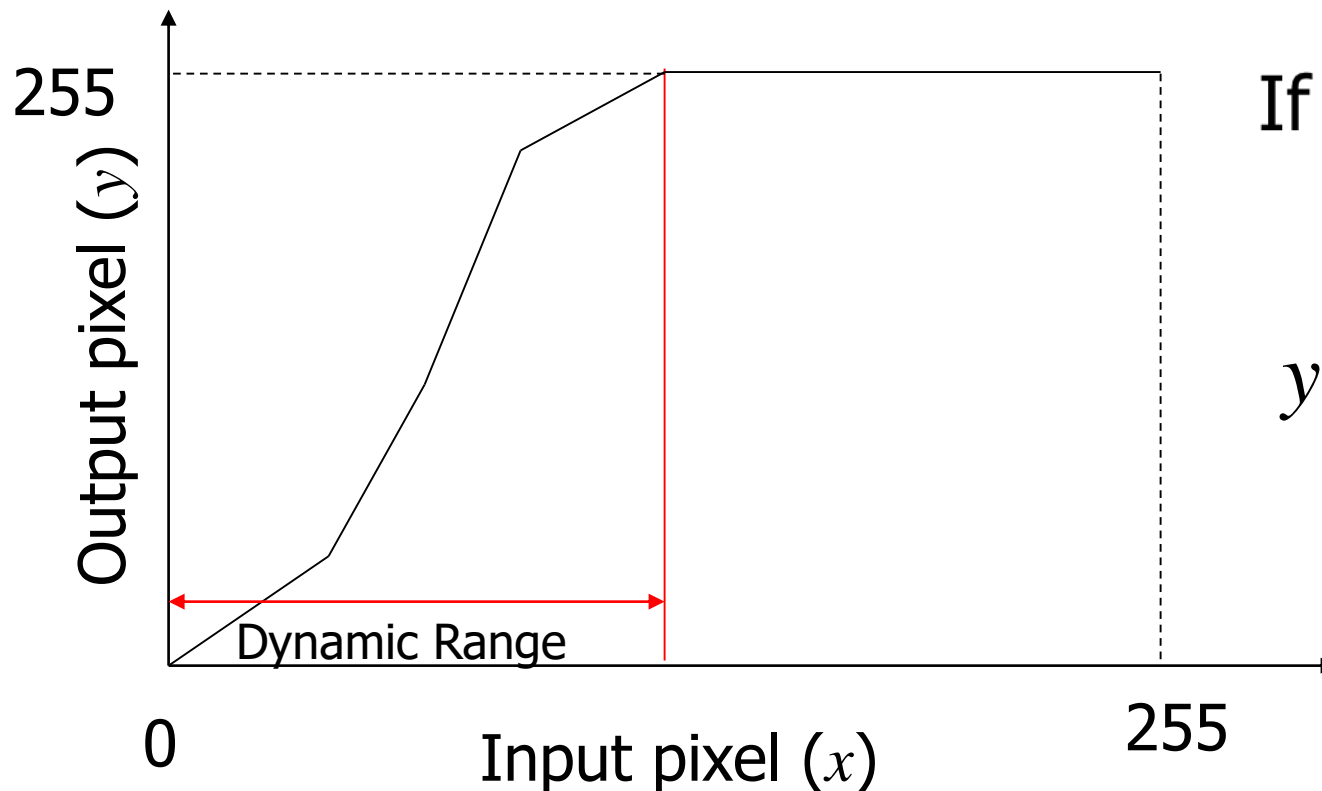
তাহা দেখি কোপেতে কল্লিত সঙ্কম্পন ।  
রণে প্রবেশিল করি রথে আরোহণ ॥  
সেহ বেগে বৃষ্টি করি বাণ বজ্রতর ।  
অঙ্গদের অঙ্গগণে করিল জর্জর ॥  
শক্রসুতসুত সহি সে সকল শরে ।  
লাফিয়া উঠিল তার রথের উপরে ॥  
কব হইতে কোদণ্ড তার কাড়ি লৈয়া ।  
চরণচাপনে তারে ফেলিল ভাঙ্গিয়া ॥  
পদাঘাতে রথখান করি প্রমথন ।  
নাশিলা নখরে করি তুরঙ্গমগণ ॥  
স্বন্দন ছাড়িয়া তবে সেই সঙ্কম্পন ।  
আকাশে উঠিল খজা করিধা ধারণ ॥  
তাহা দেখি মহাবল বালির নন্দন ।  
লক্ষ দিয়া তার পাছে করিল ধাবন ॥  
কিঞ্চিৎ দূরেতে তারে করে করি ধরি ।  
কাড়িয়া লইল তার খজা আর ফরা ॥  
তবে সিংহনাদ করি অতি কুতূহলে ।  
সেই খজা ধরি কোণ দিলা তার গলে ॥  
তাহে ছিন্ন হয়ে সেই যেন উপবীত ।  
আকাশ হইতে হৈল ভূতলে পতিত ॥  
তবে সিংহনাদ করি বালির কুমার ।  
ভূতলে নামিল শব্দ করি 'মার মার' ॥  
তবে শোণিতাকবীর লৌহগদা ধরি ।  
উপস্থিত হইল অঙ্গদ বরাবরি ॥  
প্রজন্ম যুগাৎ নামে আর দুইজন ।  
রথে চড়ি তার কাছে করিল ধাবন ॥  
শ্রীমৈন্দ দ্বিবিদ দুইবার তা দেখিয়া ।  
অঙ্গদের দুইপাশে দাঁড়াল আসিয়া ॥  
তবে সেই নিশাচর তিনজন সঙ্গে ।  
তিন কপিবার যুদ্ধ আরম্ভিল রঙ্গে ॥  
নানা বৃক্ষ উপাড়িয়া কপি তিনজন ।  
করিতেছে তিন নিশাচরে নিক্ষেপণ ॥  
তাহা দেখি খজা ধরি রাক্ষস প্রজন্ম ।  
বগু বগু করি কাটে সেই বৃক্ষসম্ম ॥  
তবে সেই তিনজন শাখাযুগবর ।  
নিক্ষেপ করেন রথতুরঙ্গর ॥  
নিরীক্ষণ করিয়া যুগাৎ রণে দক্ষ ।  
কাটিল সে সব ছাড়ি শব্দ লক্ষ লক্ষ ॥  
তবে পুনঃ শ্রীমৈন্দ দ্বিবিদ বালিব্রত ।  
বর্ধন করয়ে বৃক্ষ বজ্রত বজ্রত ॥

১৫৮৫  
১৫৮৫

১৫৮৫

# Contrast Enhancement: Pixel mapping

Function is monotonically increasing.

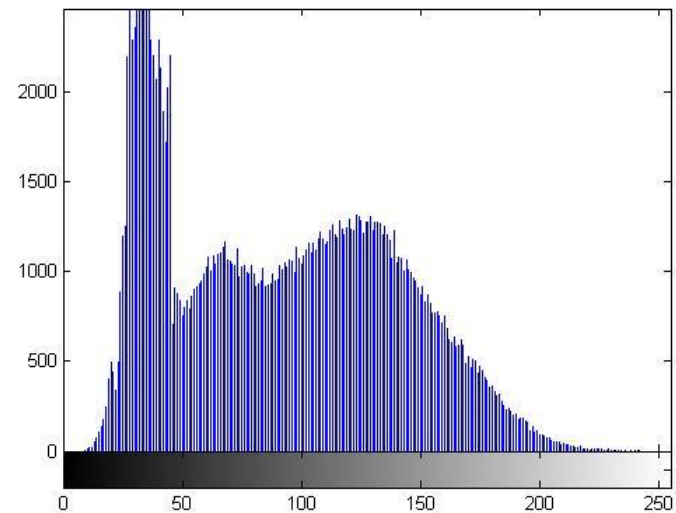
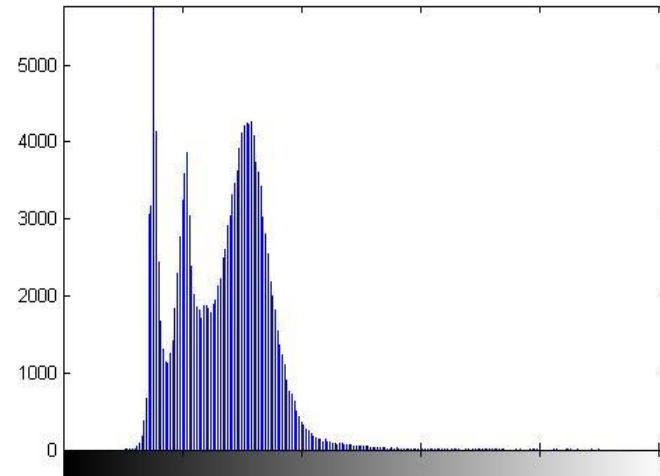


If  $x_1 > x_2$ ,  $y_1 \geq y_2$

$$y = 255 \cdot \sum_{t=0}^x p(t)$$



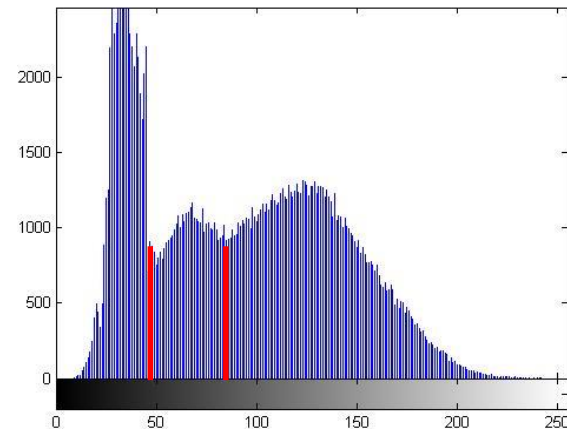
# Histogram Equalization





# Segmentation

- Partitioning image pixels into meaningful non-overlapping sets.
  - Binarization: A special case.
  - Can be extended to group pixels in more than two labels.
    - Expectation-Maximization technique following Bayesian approach.



# Probabilistic modeling with mixture of densities

Number of components

$$P(\mathbf{x}) = \sum_{i=1}^K P(\mathbf{x}|G_i)P(G_i)$$

Component density

Component proportion

- $G_i$  defines the  $i$ th segment or cluster.
- $K$  is a hyper-parameter and should be known.
- For multivariate Gaussian distribution:

- $P(\mathbf{x}/G_i) \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$

$$P(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

- To estimate  $\boldsymbol{\mu}_i$ ,  $\boldsymbol{\Sigma}_i$ , and  $P(G_i)$  for all  $i$ . from the set of *iid.* input samples:  $X = \{\mathbf{x}^t\}, t=1, 2, \dots, N$



# Mixture of Gaussians: Probabilistic inference

---

- Technique could be refined by computing probabilities of belongingness to a segments.

Parametric PDF:

$$p(x|\{\pi_k, \mu_k, \Sigma_k\}) = \sum_k \pi_k N(x|\mu_k, \Sigma_k)$$

Mixing coefficients

Mean vector

Covariance matrix



# Expectation Maximization (EM) Algorithm

$$z_{ik} = \frac{1}{Z_i} \pi_k N(x_i | \mu_k, \Sigma_k)$$

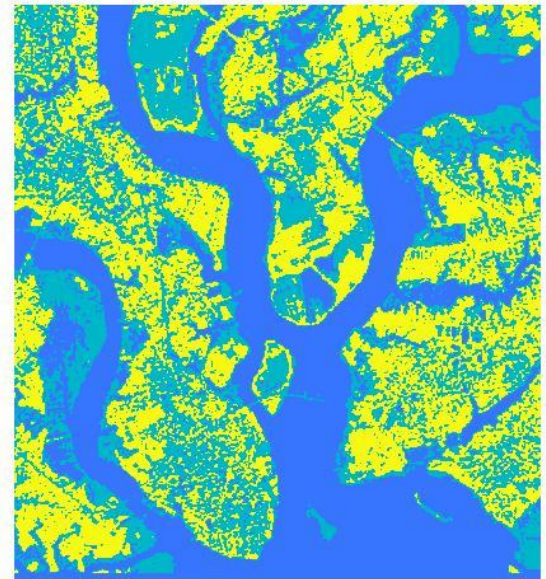
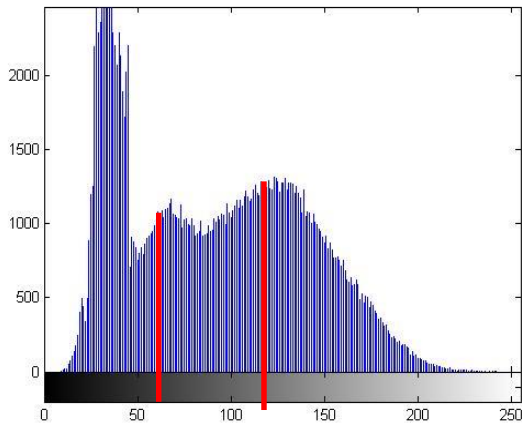
Normalizing  
factor

$$Z_i = \sum_k \pi_k N(x_i | \mu_k, \Sigma_k)$$

- Start with initial set :  $\{\pi_k, \mu_k, \Sigma_k\}$ .
- E-Step (Expectation stage)
  - Compute **probability** ( $z_{ik}$ ) of  $x$  belonging to  $k$ th Gaussian cluster.
  - Assign  $x$  to the  $m$ th cluster whose **probability** is maximum.
- M-Step (Maximization Stage)
  - Re-estimate parameters ( $\{\pi_k, \mu_k, \Sigma_k\}$ ) from class distribution
- Iterate above two steps till it converges.

Optional step.  
Decision to be  
taken at the end. →

# EM thresholding: An example



Intervals:  $[0, 60]$ ,  $[61, 119]$ ,  $[120, 255]$

Blue

Green

Yellow



# K-means clustering: Segmentation through estimating means

---

- Given  $N$   $d$ -dimensional data points,
  - compute  $K$  partitions (clusters) in them
    - so that it minimizes the sum of square of distances between a data point and the center of its respective partition (cluster).

## Optimization problem

Minimization of Sum of Squared Errors (SSE)

$$E = \sum_k \sum_{\forall x \in C_k} \|x - c_k\|^2$$

where

$$c_k = \frac{1}{|C_k|} \sum_{\forall x \in C_k} x$$



# The Lloyd algorithm (1957) (Batch K-Means)

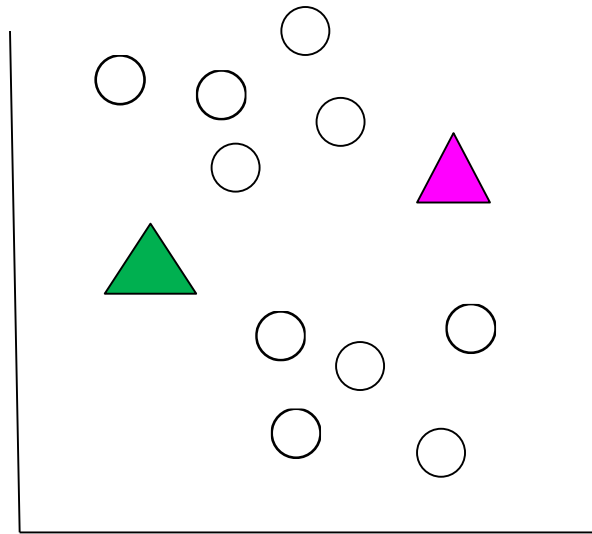
---

- Given  $K$  initial centers, assign a point to the cluster represented by its center, if it is the closest among them.
- Update the centers.
- Iterate above two steps, till the centers do not change their positions.



# K-means: example ( $k=2$ )

---



Choose initial centers.

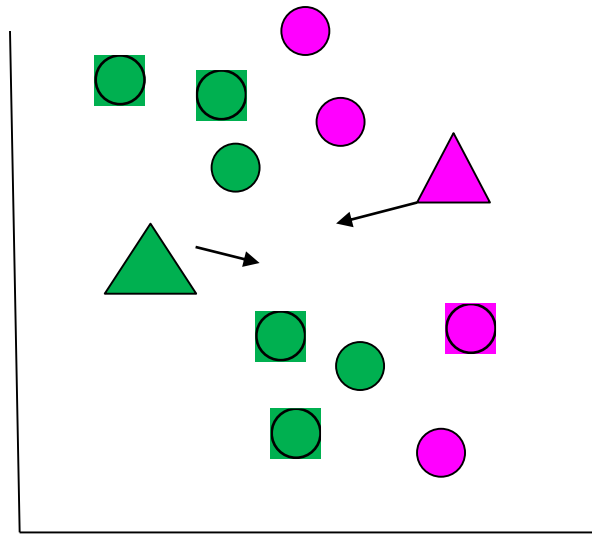
Compute partitions.





# K-means: example ( $k=2$ )

---



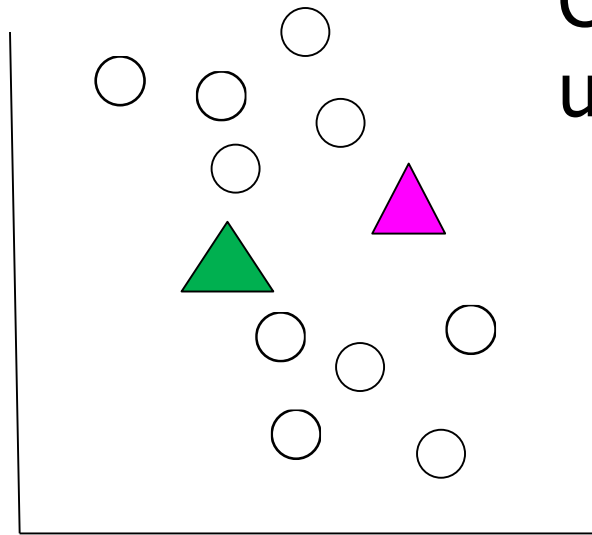
Compute partitions.

Update centers.



# K-means: example ( $k=2$ )

---

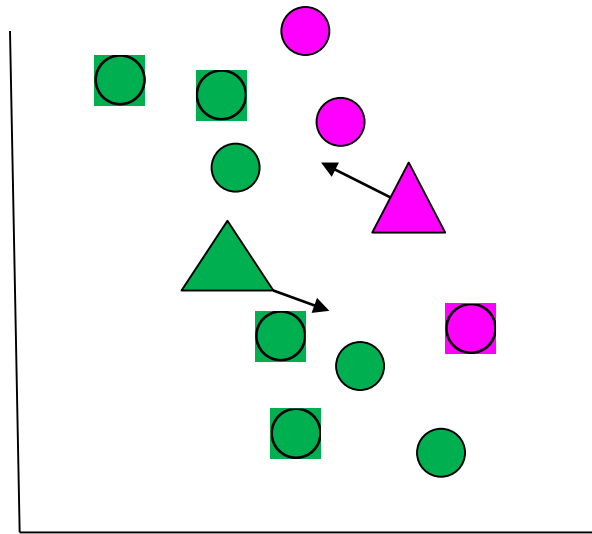


Compute new partitions with updated centers.



# K-means: example ( $k=2$ )

---

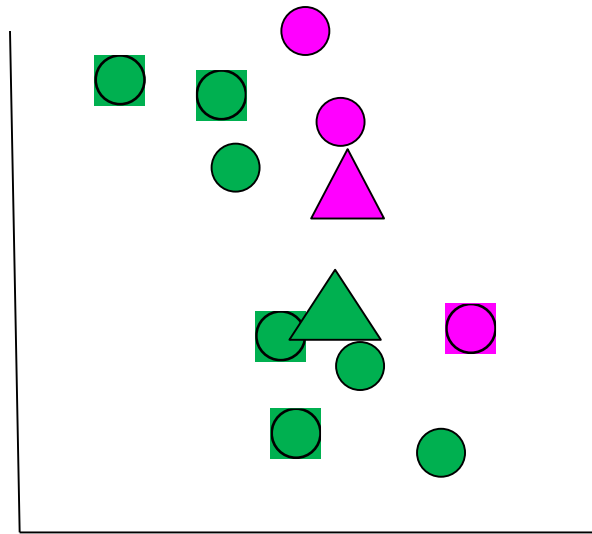


Update centers.



# K-means: example (k=2)

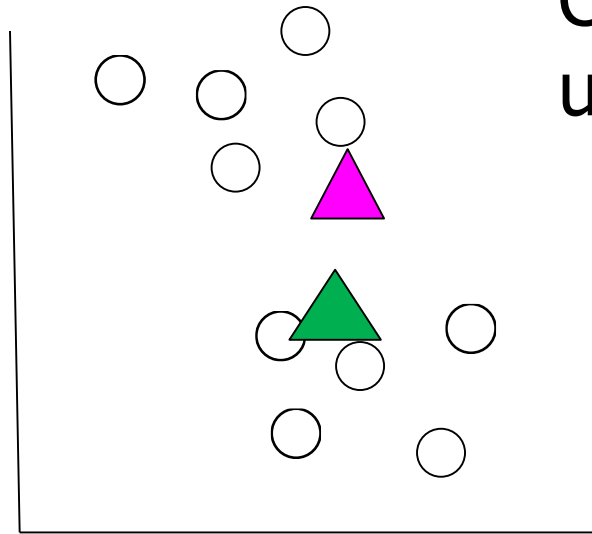
---





# K-means: example ( $k=2$ )

---

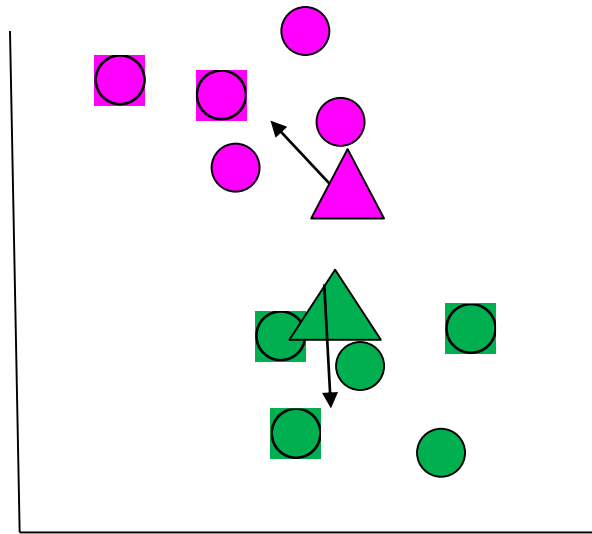


Compute new partitions with updated centers.



# K-means: example ( $k=2$ )

---

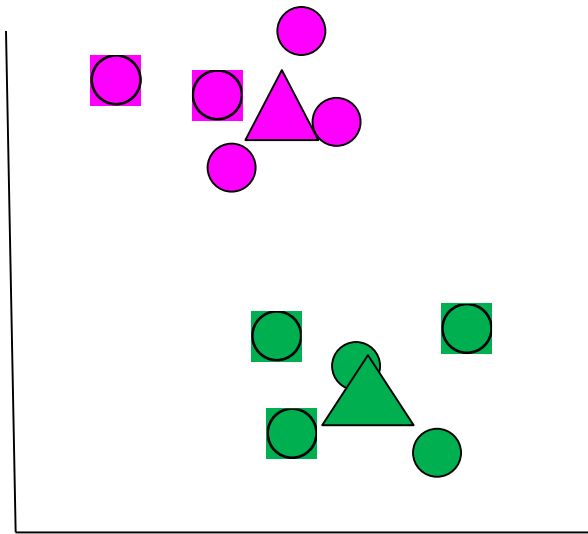


Update centers.



# K-means: example ( $k=2$ )

---



Stop at no change  
(or a very little  
change in cluster  
centers).



# Histogram analysis using K-Means clustering

---

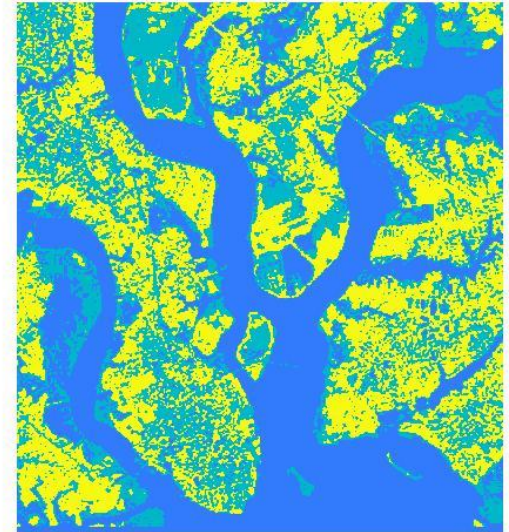
- Compute the histogram  $h(x)$ ,  $x=0,1,2,...N$
- Choose initial  $k$  brightness levels for the set of means, say  $m_1, m_2, ..., m_k$  such that  $0 < m_1 < m_2 < ... < m_k < N$
- Update the  $i$ th means, for all  $i$

$$m_i = \frac{\sum_{(m_{i-1}+m_i)/2}^{(m_i+m_{i+1})/2} xh(x)}{\sum_{(m_{i-1}+m_i)/2}^{(m_i+m_{i+1})/2} h(x)}$$

- Iterate till convergence



# K-Means Segmentation: An example

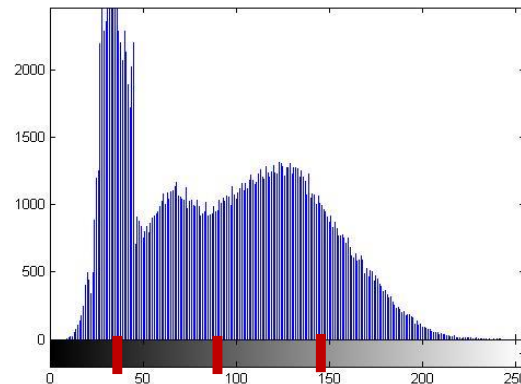


k-means for  $k=3$

$m_1=39$ ,

$m_2=90$ ,

$m_3=147$



Blue: 0-65

green: 66-119

yellow: 120-255



# Mean Shift Algorithm: Segmentation through estimating modes (peaks)

---

- A probability density estimation method
  - Non-parametric
  - Does not require the number of components ( $K$ ) as in K-Means.
- Identify peaks of distribution.
- Each peak defines a cluster / segment.
  - Pixels / elements having shortest path to a peak defines a segment

# Probability density estimation

- Probability distribution using Parzen Windows

- A function of some finite number of data points  $x_1 \dots x_n$

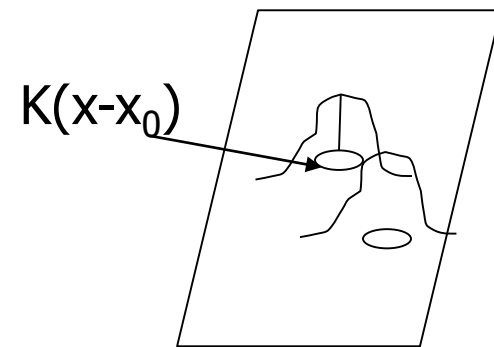
$$P(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K(\mathbf{x} - \mathbf{x}_i)$$

- Kernel Properties:

- Normalized
- Symmetric

$$\int_{R^d} K(\mathbf{x}) d\mathbf{x} = 1$$
$$\int_{R^d} \mathbf{x} K(\mathbf{x}) d\mathbf{x} = 0$$

- Exponential weight Decay  $\lim_{\|\mathbf{x}\| \rightarrow \infty} \|\mathbf{x}\|^d K(\mathbf{x}) = 0$





# Examples of kernels

---

- Epanechnikov

$$K_e(\mathbf{x}) = \begin{cases} c(1 - \|\mathbf{x}\|^2) & \|\mathbf{x}\| \leq 1 \\ 0 & \textit{otherwise} \end{cases}$$

- Uniform

$$K_u(\mathbf{x}) = \begin{cases} c & \|\mathbf{x}\| \leq 1 \\ 0 & \textit{otherwise} \end{cases}$$

- Gaussian

$$K_g(\mathbf{x}) = ce^{-\frac{1}{2}\|\mathbf{x}\|^2}$$

# Mode selection

- Compute the gradient of distribution

$$\nabla P(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \nabla K(\mathbf{x} - \mathbf{x}_i) \quad K(\mathbf{x} - \mathbf{x}_i) = ck \left( \left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)$$

$$\nabla P(\mathbf{x}) = \frac{c}{n} \sum_{i=1}^n \nabla k_i = \frac{c}{n} \left[ \sum_{i=1}^n g_i \right] \square \left[ \frac{\sum_{i=1}^n \mathbf{x}_i g_i}{\sum_{i=1}^n g_i} - \mathbf{x} \right]$$

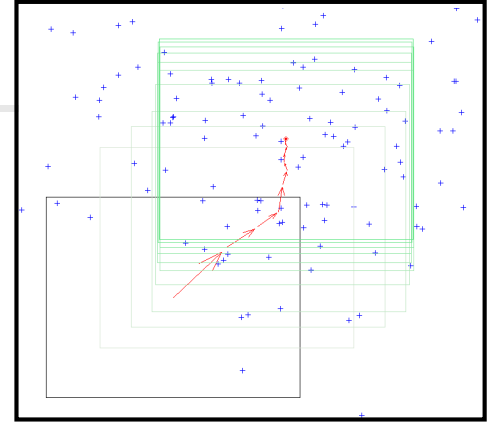
Size of window  $h$

$g(\mathbf{x}) = -k'(\mathbf{x})$

$$\nabla P(\mathbf{x}) = 0 \quad \Rightarrow \quad \mathbf{m}(\mathbf{x}) = \frac{\sum_{i=1}^n \mathbf{x}_i g \left( \frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{h} \right)}{\sum_{i=1}^n g \left( \frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{h} \right)} - \mathbf{x} = 0 \quad \Rightarrow \quad \mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{m}(\mathbf{x}^{(i)})$$

# Mean shift algorithm

- Searches a *mode* or local maximum of density of a given distribution from a point
- Choose a search window (width and location)
- Compute the mean of the data in the search window
- Center the search window at the new mean location
- Repeat until convergence
  - Converged point is a mode.
- From every point do the same.
- Set of points arriving at the same mode forms a segment



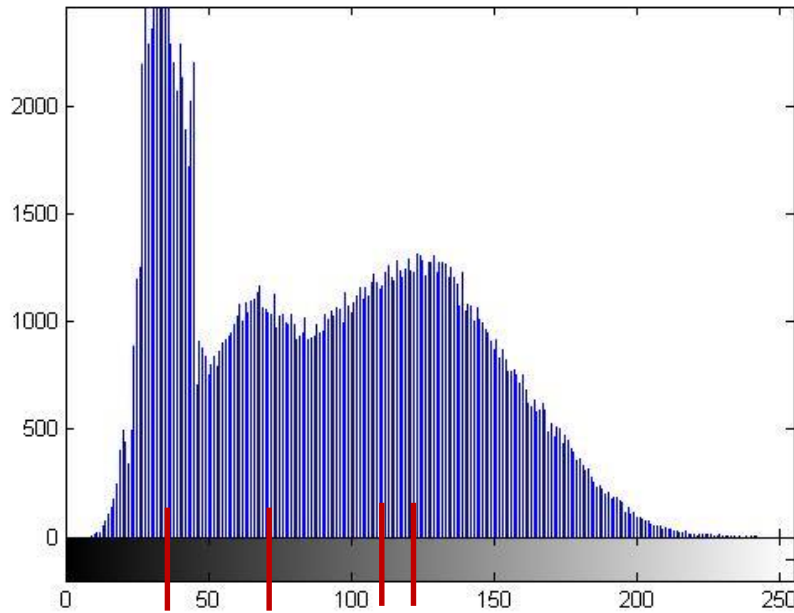


# Mean shift analysis of histogram

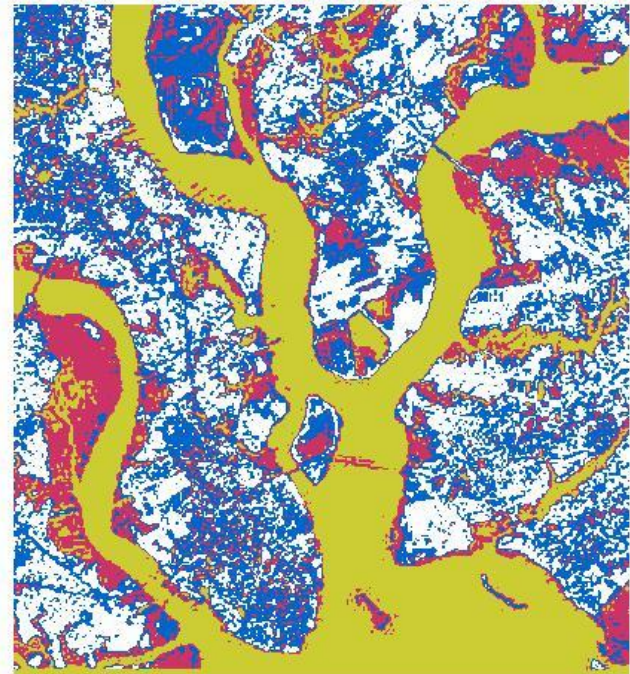
---

- Compute the histogram  $h(x)$ ,  $x=0,1,2,..N$
- For each  $x$  determine the mode  $m(x)$  in the histogram.
- Ensure monotonicity in  $m(x)=0,1,2,...,N$ 
  - $m(x_1) \leq m(x_2)$ , for  $x_1 < x_2$
- Prune spurious modes
  - sufficient support (brightness interval),
  - sufficient gap between adjacent distinct pair, and
  - sufficient strength (number of pixels).
- Get brightness interval for each mode

# Mean shift analysis of histogram and segmentation



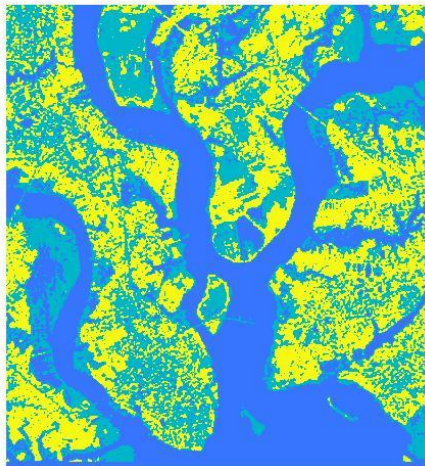
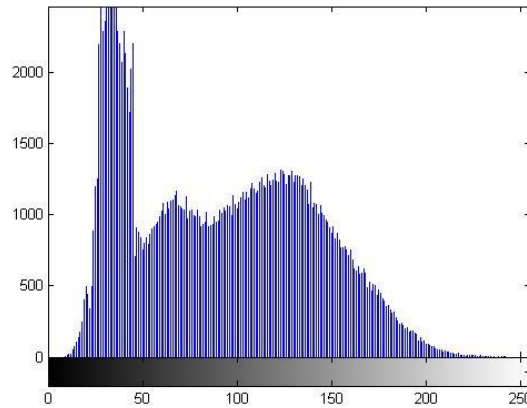
modes:  
32,67,117,130



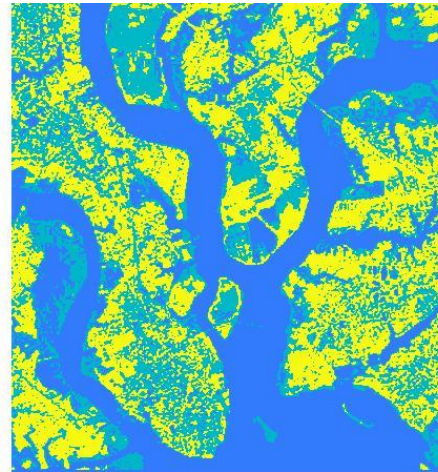
yellow:[1,51]  
red:[52,82]  
blue:[83,128]  
white:[129,256]



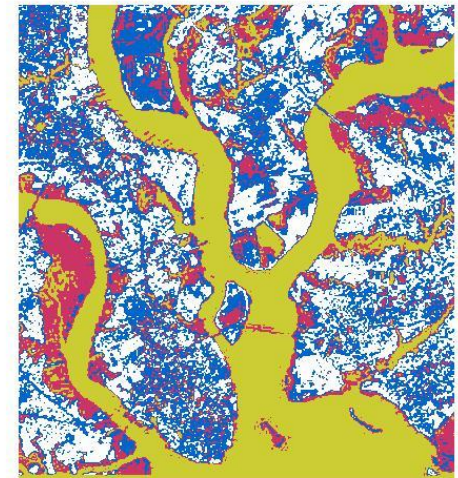
# Segmentation results from various approaches



GMM



K-Means



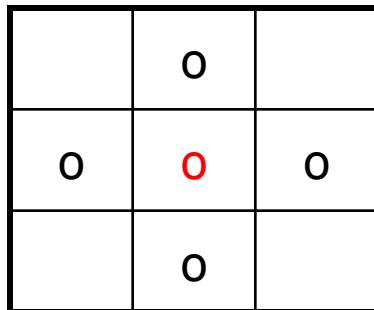
Mean Shift



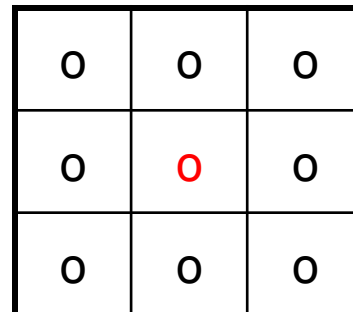
# Component Labeling

---

- Partitioning **connected** image pixels into meaningful non-overlapping sets.
  - Neighborhood definition.
  - 4-neighbor, 8-neighbor



4-neighbors



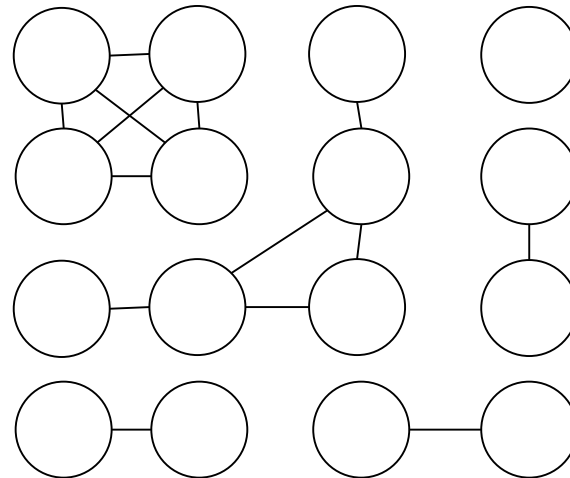
8-neighbors

0	0	0
0	0	0
0	0	0

# Component Labeling

- Form a graph with edges between neighboring pixels having same labels.
- Compute connected components.
  - Graph traversal algorithms

20	20	50	20
20	20	50	100
50	50	50	100
100	100	20	20

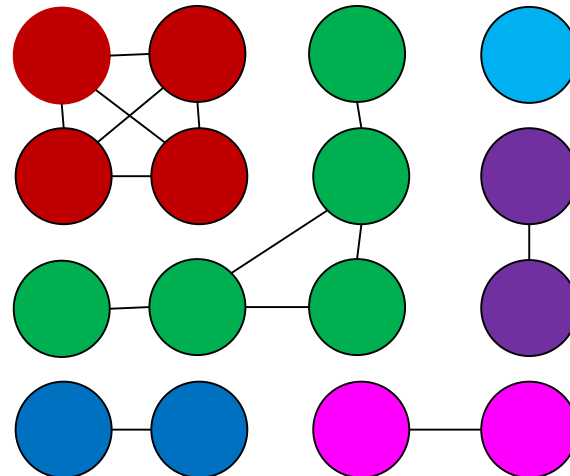


0	0	0
0	0	0
0	0	0

# Component Labeling

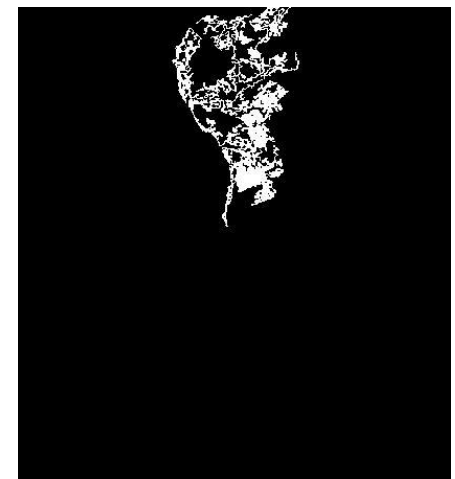
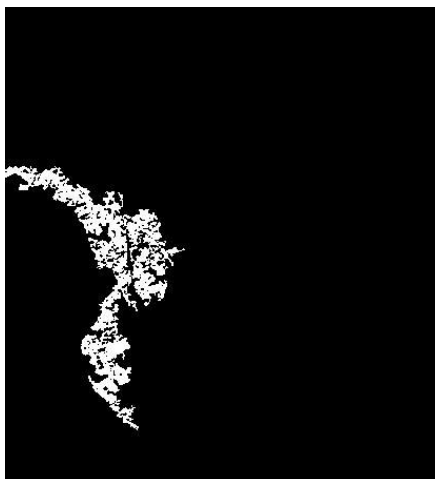
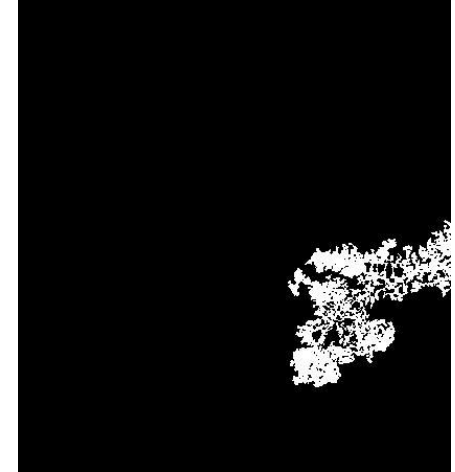
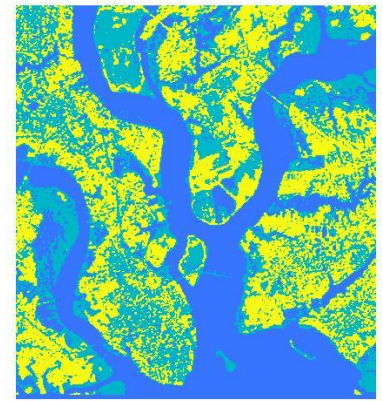
- Form graph with edges between neighboring pixels having same labels.
- Compute connected components.
  - Graph traversal algorithms

20	20	50	20
20	20	50	100
50	50	50	100
100	100	20	20



Do you require  
an explicit  
graph  
representation?  
Can you  
compute using  
only the image  
array?

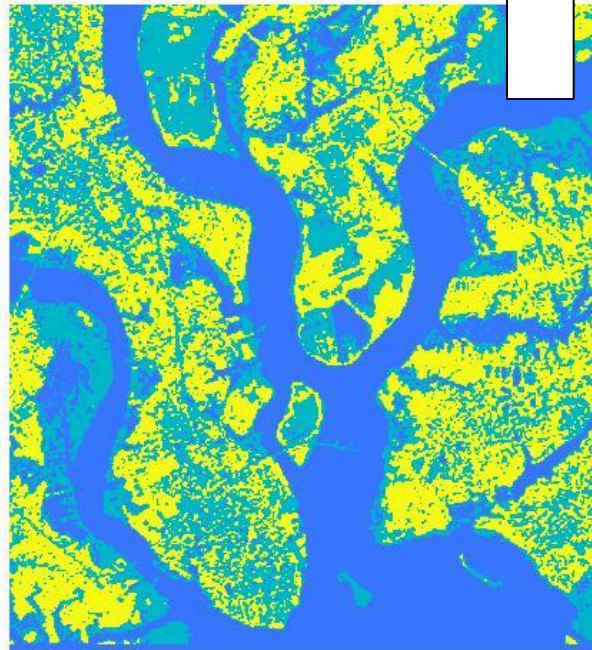
# Examples of components







?



Why a part of river channel is missing?

# Gradient Operations

Consider the image as a 2D function:  $f(x, y)$



$$\frac{\partial f(x, y)}{\partial x} = f(x+1, y) - f(x, y)$$

$$\frac{\partial f(x, y)}{\partial y} = f(x, y+1) - f(x, y)$$

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \hat{i} + \frac{\partial f(x, y)}{\partial y} \hat{j}$$

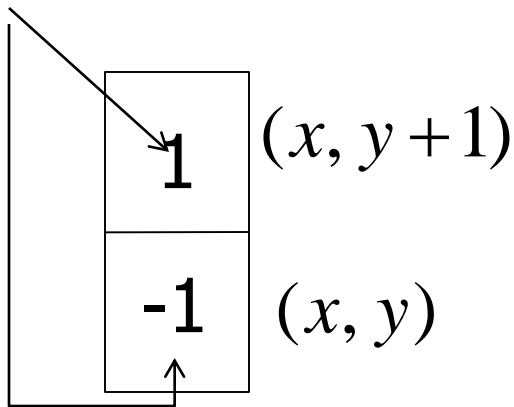


# Computation with mask

-1	1
----	---

$(x, y)$   $(x+1, y)$

Weights



1. Scan the image top to bottom and left to right.
2. At every point  $(x, y)$  place the mask and compute the weighted sum.

$$g(x, y) = (1) \cdot f(x, y+1) + (-1) \cdot f(x, y)$$

3. Write the value  $g(x, y)$  at  $(x, y)$  pixel position of the processed image.



# Robust gradient computation

Averaging neighboring gradient values

-1	-1	1
-1	1	1
-1	-1	1



-1	0	1
-1	0	1
-1	0	1

1	1	1
0	0	0
-1	-1	-1

Prewitt operator

(6 times of the gradient value in any direction)



# Robust gradient computation

Weighted average of neighboring gradient values

1x	-1	$\begin{matrix} -1 \\ 1 \end{matrix}$	1	⇒	-1	0	1	1	2	1
2x	-1	$\begin{matrix} -1 \\ 1 \end{matrix}$	1		-2	0	2	0	0	0
1x	-1	$\begin{matrix} -1 \\ 1 \end{matrix}$	1		-1	0	1	-1	-2	-1

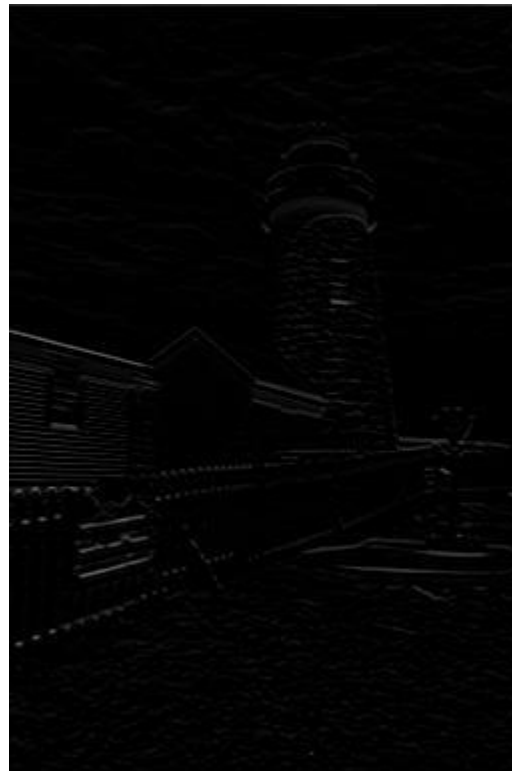
Sobel operator

(8 times of the gradient value in any direction)

# Results of gradient operations



Vertical

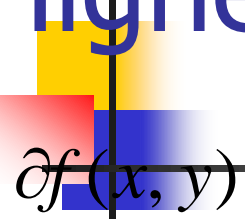


Horizontal



Resultant

# Higher order gradients!



$$\frac{\partial f(x, y)}{\partial x} = f(x+1, y) - f(x, y)$$

1	-2	1
(x-1,y)	(x,y)	(x+1,y)

$$\frac{\partial^2 f(x, y)}{\partial x^2} = \frac{\partial f(x, y)}{\partial x} - \frac{\partial f(x-1, y)}{\partial x}$$

$$\approx f(x+1, y) - f(x, y) - (f(x, y) - f(x-1, y))$$

$$= f(x+1, y) - 2f(x, y) + f(x-1, y)$$

$$\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}$$



0	1	0
1	-4	1
0	1	0

**OR**

1	0	1
0	-4	0
1	0	1

Laplacian Operator

# Weighted combination!

Laplacian Operator

$$\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}$$



1 x

0	1	0
1	-4	1
0	1	0

+

4 x

1	0	1
0	-4	0
1	0	1



$\frac{1}{5}$  x

4	1	4
1	-20	1
4	1	4



# Image sharpening

How do you  
sharpen a color  
image?

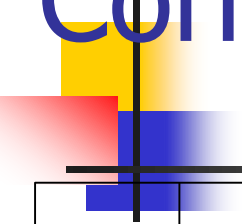
- Add a factor proportional to the gradient values to the pixel values.
  - Sharpens edges as gradients are high there.

$$g(x, y) = f(x, y) + c \cdot ||\nabla f(x, y)|| \quad c > 0$$

- To make it insensitive to direction Laplacian operator may be used.
  - The center value is subtracted. Hence the contribution to be subtracted.

$$g(x, y) = f(x, y) - c \cdot ||\nabla^2 f(x, y)|| \quad c > 0$$

# More on computation with mask: Correlation Filter

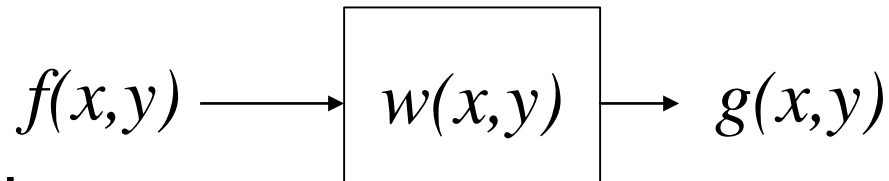


$w_1$	$w_2$	$w_3$
$w_4$	$w_c$	$w_5$
$w_6$	$w_7$	$w_8$

Cross Correlation:

$$g(x, y) = w(x, y) \bullet f(x, y) = \sum_{s=-K/2}^{K/2} \sum_{t=-K/2}^{K/2} w(s, t) f(x+s, y+t)$$

$$g(x, y) = w_1 f(x-1, y+1) + w_2 f(x, y+1) + w_3 f(x+1, y+1) + w_4 f(x-1, y) + w_c f(x, y) + w_5 f(x+1, y) + w_6 f(x-1, y-1) + w_7 f(x, y-1) + w_8 f(x+1, y-1)$$

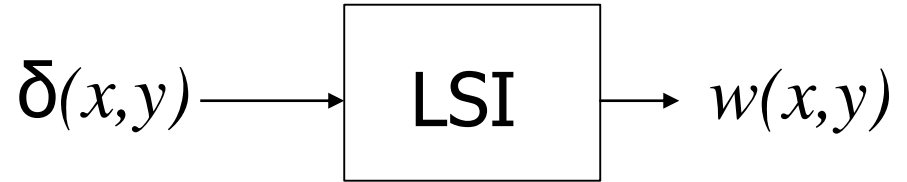


Template Matching:

Mask  $\rightarrow$  Template weights ( $w(x, y)$ )

# Convolution

Linear Shift Invariant System

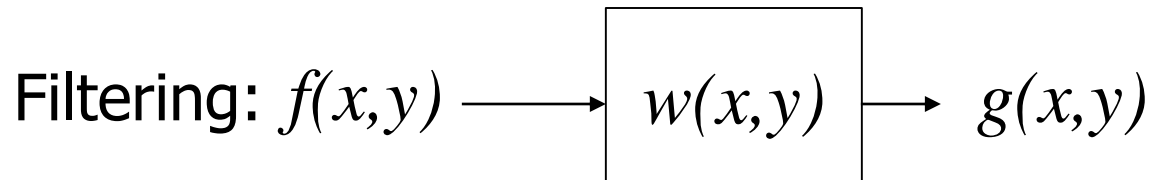


$w_1$	$w_2$	$w_3$
$w_4$	$w_c$	$w_5$
$w_6$	$w_7$	$w_8$

$$g(x, y) = w(x, y) * f(x, y) = \sum_{s=-K/2}^{K/2} \sum_{t=-K/2}^{K/2} w(s, t) f(x-s, y-t)$$

Same as correlation except that the mask is flipped both horizontally and vertically.

$$g(x, y) = w_8 f(x-1, y+1) + w_7 f(x, y+1) + w_6 f(x+1, y+1) + w_5 f(x-1, y) + w_c f(x, y) + w_4 f(x+1, y) + w_3 f(x-1, y-1) + w_2 f(x, y-1) + w_1 f(x+1, y-1)$$



If  $w(x, y)$  is symmetric, that is  $w(x, y) = w(-x, -y)$ , then convolution is equivalent to correlation!

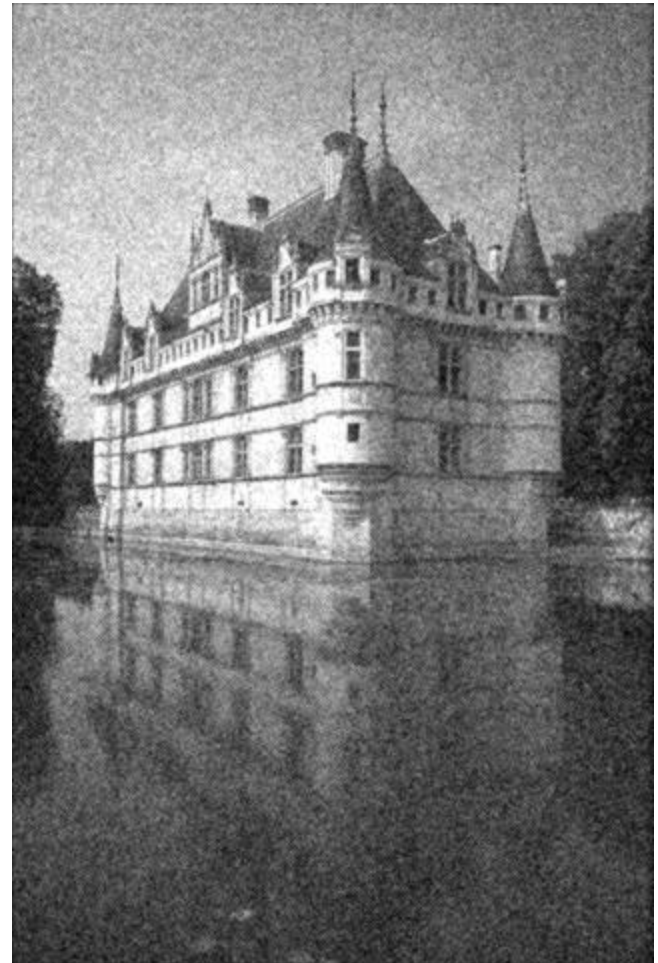
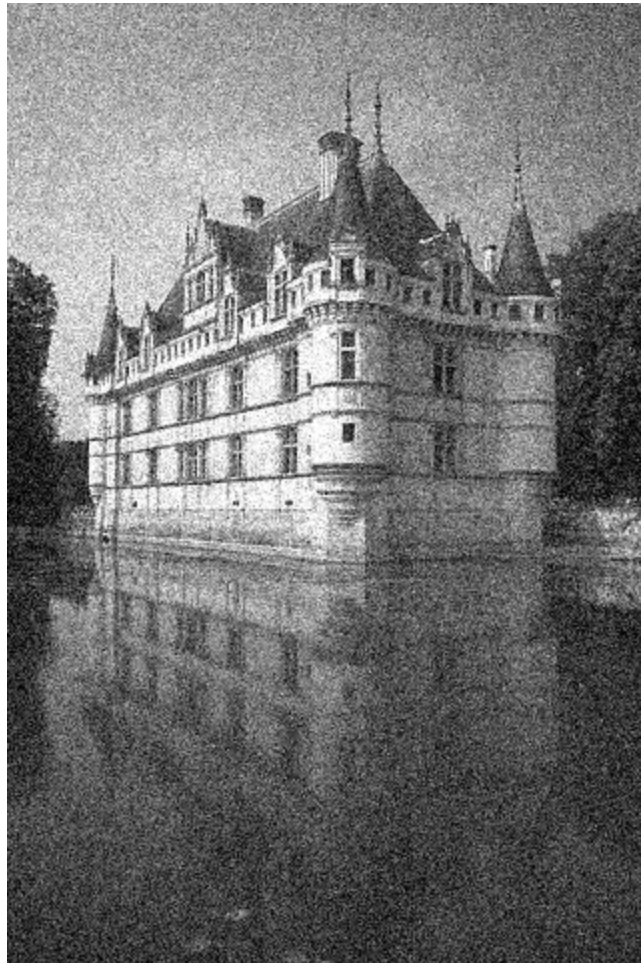


# Additive Noise Filtering

Weighted Mean  
filter: sum of the  
weights= 1, and  
all positive.

$c$	$b$	$c$
$b$	$a$	$b$
$c$	$b$	$c$

$a=0.5$ ,  $b=0.3/4$ ,  
and,  $c=0.2/4$



# Gaussian Smoothing

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-((x-x_c)^2 + (y-y_c)^2)}{2\sigma^2}}$$

$$g(x, y) = f(x, y) * G(x, y)$$

$$\sigma=2$$

Mask size: 9x9





# Median Filtering

---

$g(x,y)$  = the median value among  
the neighbors.

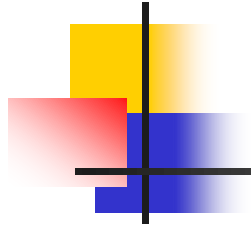




# Summary

---

1. Images formed by optical camera through perspective projection.
2. Discussed different operations in image processing.
  - Binarization
    - Thresholding
  - Contrast enhancement
    - Pixel mapping
  - Segmentation
    - Histogram analysis
    - Multilevel thresholding
  - Component Labeling
  - Gradient computation.
    - Computing edges
  - Correlation and Convolution
    - Gaussian filtering
  - Median filtering



Thank you!