INDEX

IT 524 - Autumn 2012-13

NAME: Milind Padalkar STD.: PhD SEC.: M1 ROLL NO.: 201121015 SUB.: Computer Vision

S. No.	Date	Title	Page No.	Teache Sign / Remark
27	15/10/2012	Lecture 27	122	
28	16/10/2012	Lecture 28	124	
29.	19/10/2012	Lecture 29	126	

classmate
Date 23/07/2012
Page 1

Lecture 1.

2011034918

Books :-
① Introductory techniques for 3D Computer Vision.
 – Emanuele Trucco et al. Prentice Hall (Xerox with Sir)
② Robot Vision – B.K.P. Horn (MIT Press)
③ Image processing, analysis & machine vision – Sonka et al.
④ Computer Vision – A modern Approach – Forsyth & Ponce.

Computer Vision :- (2D to 3D)
A set of computational techniques for estimating the 3D properties. These may be geometric, dynamic etc.
Geometric – Finding the 3D shape / depth
Dynamic – Finding velocity of a 3D objects.

IP : 2D to 2D
CV : Inverse Problem 2D to 3D
CG : Forward Problem 3D to 2D.

Conferences :- ICCV, CVPR, ECCV, ACCV, ICIP, ICPR, ICVGIP, etc.
Journals : IJCV, PAMI, IVC, CVIU, IEEE trans. on IP, PR, MI, PR Letters, SPM (Good for literature survey), special issues

People.
① Pushmit Kohli
② Pawan Kumar
③ P. Anandan
④ Manik Varma
} Microsoft Research
⑤ Andrew Zisserman, (U.K.)
⑥ Jitendra Malik. (Berkely)
⑦ P. Jawahar (NIT Warangal)

Websites for material
① Computer Vision Homepage on Google search.
② Keith Bibliography
③ cvonline
④ vislist.

Industries:
Sarnoff, HCL, MS Research, Etica Solutions, Online Solutions, etc

Lecture 2

24/07/2012

Software for CV programs.
1. Matlab
2. Mathematica
3. Scilab
4. Numerical Recipies in C (Book + s/w)

→ Simulated Anneling.

Course content:
→ To get 3D information from 2D images.
→ Study computational techniques to perform above task.

Whenever a scene (3D) is imaged (2D), one dimension lost. How can it be recovered?
→ An illposed problem ... to be solved during the course.

Optical Camera v/s Range Camera.

→ Optical Camera capture info. of a 3D object on a 2D image plane.
→ Range Camera captures depth info. of a 3D object.

Course Outline.
→ 2D to 2D transformations (Geometric transformations).
→ 3D to 2D transformations (Perspective Projections)
→ ~~Thin~~ Pin-hole camera (No lens).
→ Real Aparture imaging (thin lens model).
→ Photometric transformation
→ Bidirectional Reflectance Distribution Function (BRDF) - Scen
→ Lambertian surface.
→ Specular surface.

→ 1st Approach for 3D estimation
  → Shape from shading (SFS).
  → Variational approach.
→ 2nd Approach
  → Photometric Stereo
  → Least Squares Approach (LS)
  → Constrained LS
  → Total LS
  → SVD
  → PCA.
  → EM algorithm.
→ 3rd Approach.
  → Steropsis or stereo.
  → Optical Flow.
→ ~~4~~th Approach
  → Depth from focus.
→ 5th Approach
  → Depth from defocus / blur.

→ Markov Random Field (MRF)
→ Super resolution imaging
→ Image Inpainting
→ Conditional Random Field.

$$Y = AX$$

We have A & Y & aim to find X. If A is square then A^{-1} may exist & then $A^{-1}Y = A^{-1}AX = IX = X$ ... — (1)

But if A is not a square matrix then

$$Y = AX \Rightarrow A^TY = A^TAX$$ ... A^TA is now square

$$(A^TA)^{-1}A^TY = (A^TA)^{-1}(A^TA)\hat{X}$$ ... X may not be exactly obtained. Yet

$$\therefore \boxed{\hat{X} = (A^TA)^{-1}A^TY}$$

$\hat{X}$ gives $\hat{Y} = A\hat{X}$ such that $(Y-\hat{Y})^2$ is

– gives the solution in the LS sense.

Thus $\Sigma(y_i - \hat{y}_i)^2 \neq 0$. If $=0$ then $X = \hat{X}$.

Symmetry of a matrix. – $(A^TA)^T = (A^T)(A^T)^T = A^TA$

Singular matrix has determinant $= 0$.

Lecture 3

31/07/2012.

* 2D to 2D transformations.

Recap:-

Q. What is an ill posed problem?

Ans. A problem is said to be ill posed if it has
① no solution.
② multiple solutions (no unique solution)
③ a small perturbation in the data changes the solution drastically.

Eg. ⓐ $x+y=1$
ⓑ $x-y=1$

The above two equations have a unique solution $(x,y)=(1,0)$. However if

① ⓐ $x+y=1$
ⓑ $x+y=2$ } has no solution

Whereas

② ⓐ $x+y=1$
$2x+2y=2$ } are linearly dependant equations & has multiple/infinite solutions.

Further,

③ ⓐ $x_1+x_2=2$
ⓑ $x_1+1.0001\,x_2=2.0001$ } has solution $x_1=1, x_2=1$

While,

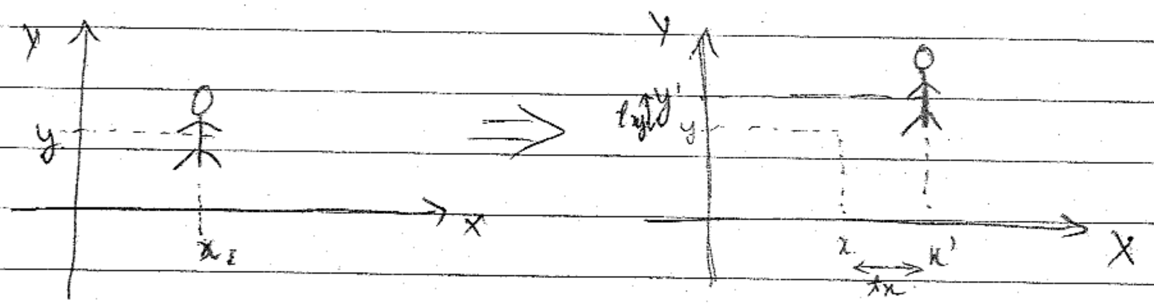
ⓐ $x_1+x_2=2$
ⓑ $x_1+1.0001\,x_2=2.0002$ } gives $x_1=0$ & $x_2=2$. Thus, has large change in solution for small change in data. Gives an ill-conditioned matrix.

Vision problems are ill posed. To make them better-posed we use constraints.

2D Geometric Transformations

① Rigid body / Euclidean Transformation
Properties of the objects are unchanged with transformations (lengths, angle, etc).
Eg:-
2D: ~~Tot~~ Translation & Rotation (2D To 2D or image-to-imag

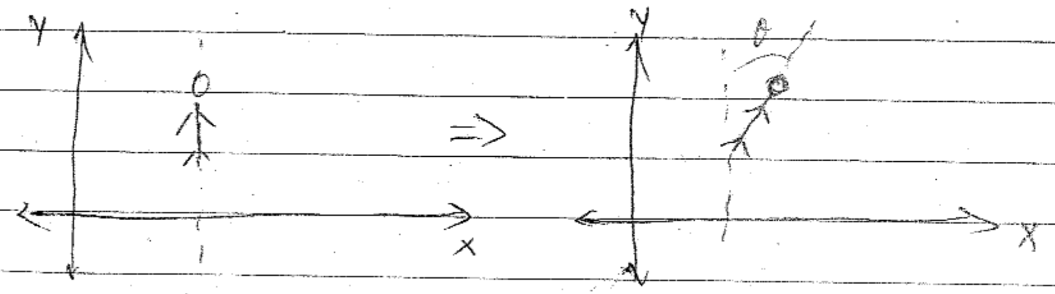
Translation :-



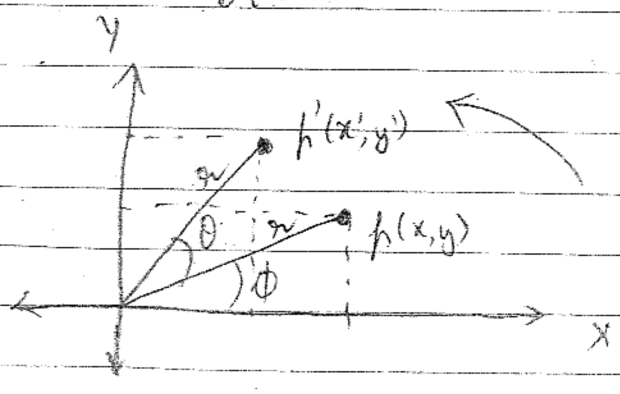
$$\therefore \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \equiv \begin{array}{l} x' = x + t_x \\ y' = y + t_y \end{array}$$

or $\underline{p}' = \underline{p} + \underline{t}$

Rotation :-



or.



Now,

$$\frac{\cancel{y}}{r} = \sin\phi \quad \& \quad \frac{\cancel{x}}{r} = \cos\phi .$$

$$\therefore \cancel{x}y = r\sin\phi \quad \& \quad \cancel{y}x = r\cos\phi$$

similarly,

$$y' = r\sin(\theta+\phi) \quad \& \quad x' = r\cos(\theta+\phi)$$

$$y' = r[\sin\theta.\cos\phi + \cos\theta\sin\phi] \qquad x' = r[\cos\theta.\cos\phi - \sin\theta\sin\phi].$$

$$\therefore y' = r\cos\phi.\sin\theta + r\sin\phi.\cos\theta \qquad x' = r\cos\phi.\cos\theta - r\sin\phi.\sin\theta .$$

$$\therefore y' = x\sin\theta + y\cos\theta \qquad x' = x\cos\theta - y\sin\theta .$$

$$\therefore y' = y\cos\theta + x\sin\theta \qquad x' = -y\sin\theta + x\cos\theta .$$

$$\therefore \begin{bmatrix} x' \\ y' \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{\text{Rotation matrix}} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\therefore p' = R.p$$

Homework: ① Program for rotation and translation of grey-level images. -done.
② Bilinear & Linear interpolation. - done.

Rotation + translation : $p' = R.p + t$

Now, $(p')^T(p') = ?$ $\qquad p' = (x', y')'$

$$\therefore (x', y')\begin{pmatrix} x' \\ y' \end{pmatrix} = (x')^2 + (y')^2 = r^2 \qquad — ①.$$

Also, $p^T p = (x, y)\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 = r^2$ — (2).

If we consider only rotation, then $p' = Rp^*$

$\therefore (Rp^*)^T (Rp^*) = p^T R^T R p = r^2 = p^T p$ ... from (1) & (2)

$\therefore R^T R = I$.

$\therefore$ R has to be orthonormal matrix.

$\therefore R^T = R^{-1}$ ... orthogonal matrix in linear algebra

Orthonormal = Orthogonal + Unit norm.

(2) Similarity transformation

Includes, uniform scaling.

$$p' = sRp + \underline{t}$$

$s \downarrow$ Scaling factor (a scalar)

(3) Affine transformations.

Non-uniform scaling.

$$p' = Ap + \underline{t} \qquad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

A & $\underline{t}$ are affine parameters.

→ W.a.p. for affine transform. ... similar to rotation.

Eg:- $(x, y) \rightarrow (ax, by)$ ... preserves lines & parallelism

In our case,

$x' = ax + by$

& $y' = cx + dy$

<u>Homogeneous coordinates.</u>

Normally it is difficult to represent $\underline{X}_1 = \underline{X} + \underline{t}$ as $\underline{X}_1 = A\underline{X}$. For this purpose, homogeneous coordinates are used.

Eg:- $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$.

Now, $\underset{3\times1}{\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}} = \underset{3\times3}{\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}} \underset{3\times1}{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}$

$x' = x + t_x$
$y' = y + t_y$
$\& \ 1 = 1$

→ H.W. represent rotation, scaling using homogeneous coordinates

① <u>Rotation:-</u> ~~Carten~~ Euclidean coordinates.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Homogeneous coordinates are.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

② <u>Affine transformation (Non-uniform scaling)</u>

Euclidean coordinates.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

Homogenous Coordinates are.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$\therefore$ $x' = ax + by + t_x$
$y' = cx + dy + t_y$.

Most General form to get homogeneous coordinates using transformation is

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

Such that.

$$\frac{x'}{w'} = a.\frac{x}{w} + b\frac{y}{w} + c.\frac{w}{w},$$

$$\frac{y'}{w'} = d.\frac{x}{w} + e.\frac{y}{w} + f.\frac{w}{w} \text{ and}$$

$$1 = g.\frac{x}{w} + h.\frac{y}{w} + i.$$

Lecture 4

03/08/2012

3D to 3D transformations

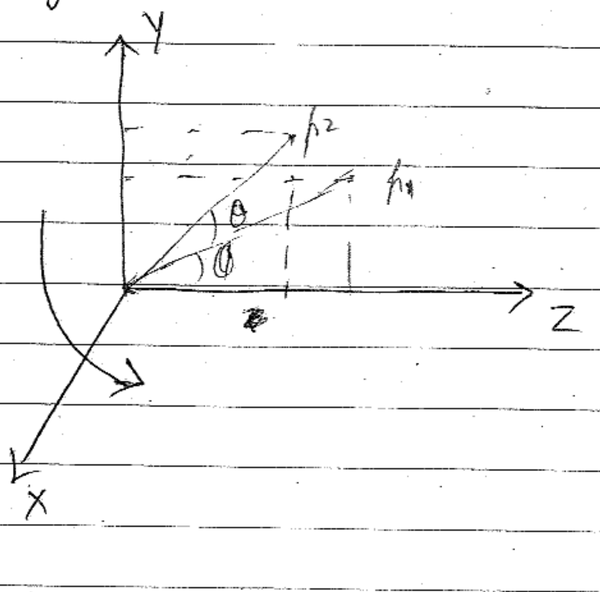
$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ z + t_z \end{bmatrix}$$

... translation in 3D.
... Euclidean coordinates.

∴ Homogeneous coordinates are

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Similarly in 3D we perform rotation & denote using homogeneous coordinates as follows.

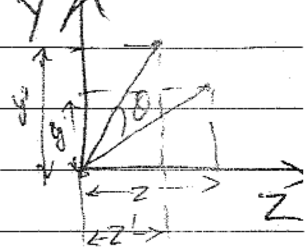


counter clockwise rotation

When we rotate about X axis, the x coordinate remains constant & the coordinates in Y-Z plane are transformed. Thus in Euclidean coordinates we have

$$\cancel{\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \cos\theta & -\cos\theta & \\ & & \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}}$$

Rotation about X axis.



$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sin\theta & \cos\theta \\ 0 & \cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

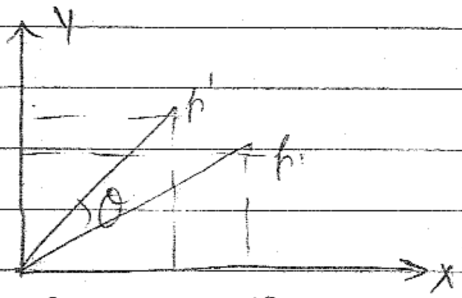
$z' = z\cos\theta - y\sin\theta$

$\therefore \cancel{x' = x}, \cancel{y' = z\cos\theta - y\sin\theta}, \cancel{z' = z\sin\theta}$ $\quad y' = z\sin\theta + y\cos\theta$

$$\therefore \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Rotation about Z axis : z is constant.

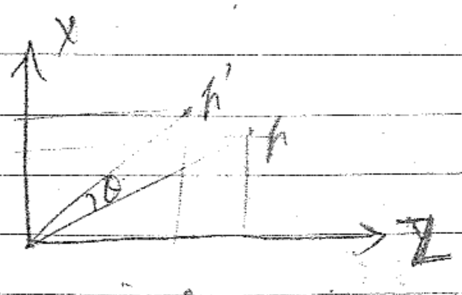
$\therefore$



$x' = x\cos\theta - y\sin\theta$
$y' = x\sin\theta + y\cos\theta$

$$\therefore \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Rotation about Y axis, we have y constant.



$z' = z\cos\theta - x\sin\theta$, $x' = z\sin\theta + x\cos\theta$.

$$\therefore \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Homogeneous coordinates:

① Rotation about X axis.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

② Rotation about Z axis.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

③ Rotation about Y axis.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

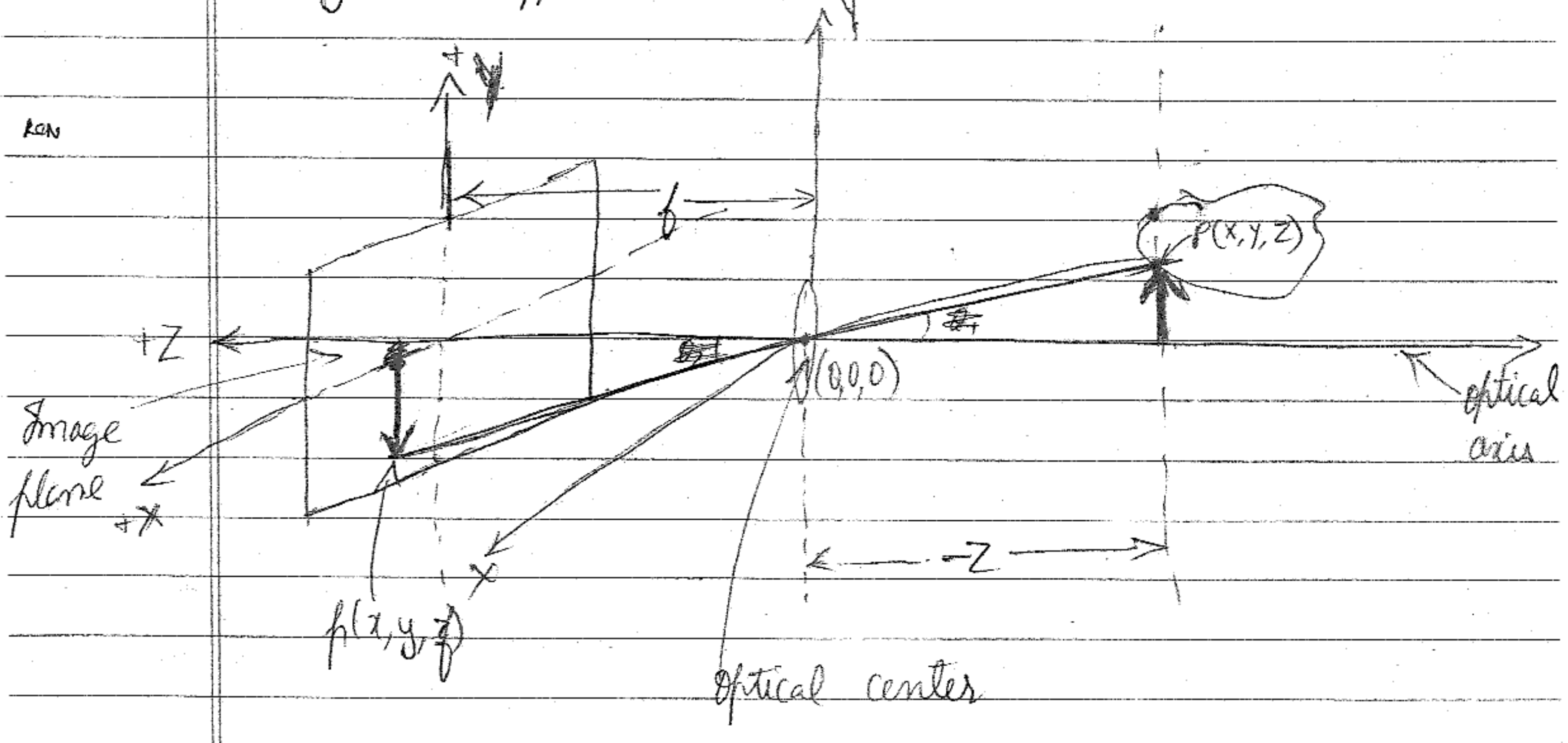
Image formation.

Perspective Projection (Central Projection) (3D to 2D)

Shapes are described as they are & not as they appear.

Refer Forsyth & Ponce

→ Pinhole camera :- A model, which has source light emerging from a hole having negligable diameter. No lens is used here. This model is used in many CV applications.



Now, ~~$\theta_1 = \theta_2$~~

~~$\therefore \tan\theta_1 = \tan\theta_2$~~

$\therefore$

$$\cancel{\frac{f}{-Z} = \frac{x}{-X}}$$

$$\cancel{\therefore x = f\left(\frac{-X}{-Z}\right) = f\frac{X}{Z} = X\left(\frac{f}{Z}\right)}$$

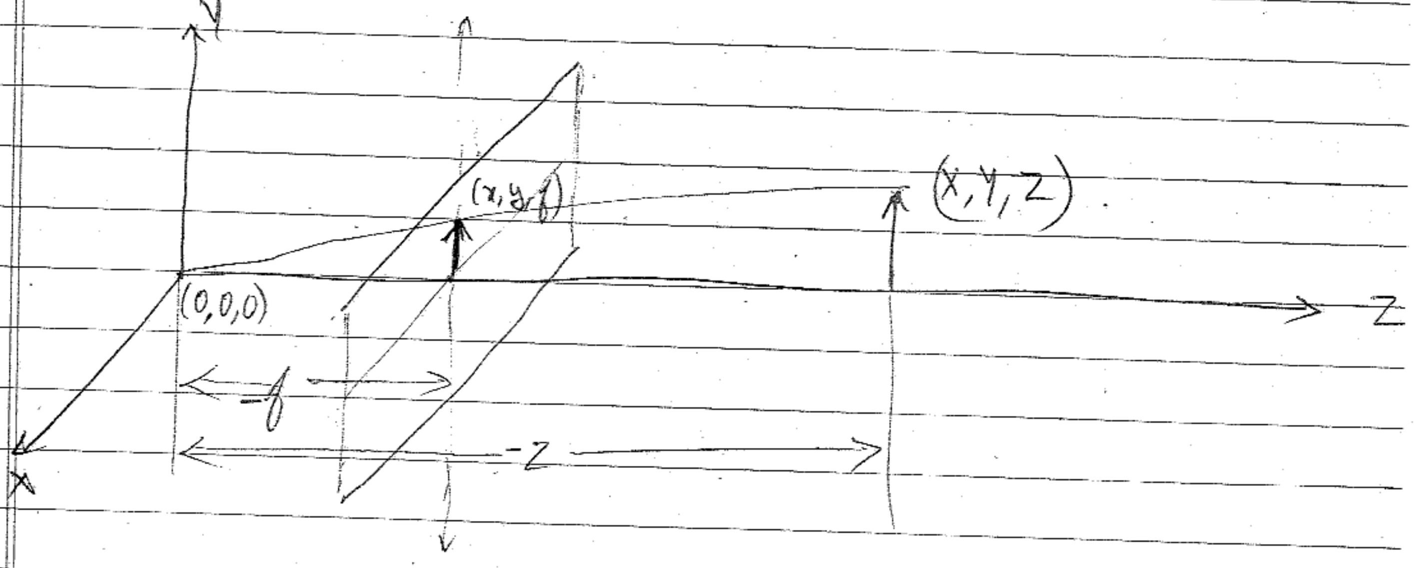
~~Similarly:~~ $\frac{f}{-Z}$

$$\frac{f}{x} = \frac{-Z}{-X} \quad \therefore x = X \cdot \left(\frac{f}{Z}\right),$$

$$\frac{f}{y} = \frac{-Z}{-Y} \quad \therefore y = f \cdot \frac{Y}{Z} = Y \cdot \left(\frac{f}{Z}\right).$$

$z = f$.

In order to get a non-inverted image, the image plane is now kept between the object and the optical center.



Again $x = X \cdot \frac{f}{Z}$, $y = Y \cdot \frac{f}{Z}$

The perspective projection is a non-linear transformation

Why?

Condition for linearity is $T(\alpha x_1(t) + \beta x_2(t)) = \alpha T[x_1(t)] + \beta T[x_2(t)]$

P.T.O.

Here, let us take 2 points $P_1(X_1, Y_1, Z_1)$ & $P_2(X_2, Y_2, Z_2)$

Now, for P_1,

$$x_1 = f\frac{X_1}{Z_1}, \quad y_1 = f\frac{Y_1}{Z_1}$$

for P_2,

$$x_2 = f\frac{X_2}{Z_2}, \quad y_2 = f\frac{Y_2}{Z_2}.$$

Let $P_3 = P_1 + P_2$

$$\therefore P_3 = (X_1 + X_2,\ Y_1 + Y_2,\ Z_1 + Z_2).$$

$$\therefore x_3 = f\left(\frac{X_1 + X_2}{Z_1 + Z_2}\right) \ \& \ y_3 = f\left(\frac{Y_1 + Y_2}{Z_1 + Z_2}\right).$$

But,

$$x_1 + x_2 = f\left(\frac{X_1}{Z_1}\right) + f\left(\frac{X_2}{Z_2}\right) = f\left(\frac{X_1}{Z_1} + \frac{X_2}{Z_2}\right) \neq x_3$$

$$\& \ y_1 + y_2 = f\left(\frac{Y_1}{Z_1}\right) + f\left(\frac{Y_2}{Z_2}\right) = f\left[\frac{Y_1}{Z_1} + \frac{Y_2}{Z_2}\right] \neq y_3.$$

... for $\alpha = \beta = 1$.

Thus,

$$T[P_{1x} + P_{2x}] = f\left[\frac{X_1 + X_2}{Z_1 + Z_2}\right] \neq f\left[\frac{X_1}{Z_1} + \frac{X_2}{Z_2}\right] = T(P_{1x}) + T(P_{2x}),$$

$$\& \ T[P_{1y} + P_{2y}] \neq T[P_{1y}] + T[P_{2y}]$$

Therefore the perspective projection is non-linear.

To represent (3D-2D) transformation using homogeneous coordinates we have,

why?

$$\Longrightarrow \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$\therefore \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\therefore \lambda x = fX$$
$$\lambda y = fY$$
$$\& \ \lambda = Z$$

$$\therefore \boxed{x = f\frac{X}{Z}, \quad y = f\frac{Y}{Z}}$$

From Euclidean coordinates $(x, y)'$ is transformed to $(x, y, 1)'$ or $(\lambda x, \lambda y, \lambda)'$

i.e. $\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ or $\begin{pmatrix} \lambda x \\ \lambda y \\ \lambda \end{pmatrix}$

similarly, $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \rightarrow \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$

Lecture 5

06/08/2012

Most general form of 3D to 2D projective geometry is.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ T_{21} & T_{22} & T_{23} & T_{24} \\ T_{31} & T_{32} & T_{33} & T_{34} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}.$$

Now; $x = \frac{x_1}{x_3}$, $y = \frac{x_2}{x_3}$.

& $X = \frac{X_1}{X_4}$, $Y = \frac{X_2}{X_4}$, $Z = \frac{X_3}{X_4}$.

The above equation is used for projective camera.
→ To preserve parallelism an affine camera is used. Projective camera can only preserve lengths but not parallelism.
→ Affine camera can preserve both lengths & parallelism

The homogeneous equation for an affine camara is,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ T_{21} & T_{22} & T_{23} & T_{24} \\ 0 & 0 & 0 & T_{34} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

Now,

$$x_1 = T_{11}X_1 + T_{12}X_2 + T_{13}X_3 + T_{14}X_4.$$
$$x_2 = T_{21}X_1 + T_{22}X_2 + T_{23}X_3 + T_{24}X_4.$$
$$x_3 = T_{34}X_4.$$

And.

$$x = \frac{x_1}{x_3} = \left(\frac{T_{11}}{T_{34}}\right)\frac{X_1}{X_4} + \left(\frac{T_{12}}{T_{34}}\right)\frac{X_2}{X_4} + \left(\frac{T_{13}}{T_{34}}\right)\frac{X_3}{X_4} + \left(\frac{T_{14}}{T_{34}}\right)$$

Similarly,

$$y = \frac{x_2}{x_3} = \left(\frac{T_{21}}{T_{34}}\right)\frac{X_1}{X_4} + \left(\frac{T_{22}}{T_{34}}\right)\frac{X_2}{X_4} + \left(\frac{T_{23}}{T_{34}}\right)\frac{X_3}{X_4} + \left(\frac{T_{24}}{T_{34}}\right)$$

Let $\frac{T_{14}}{T_{34}} = t_x$ & $\frac{T_{24}}{T_{34}} = t_y$.

& $\left(\frac{T_{11}}{T_{34}}\right) = a$, $\left(\frac{T_{12}}{T_{34}}\right) = b$, $\left(\frac{T_{13}}{T_{34}}\right) = c$.

$\left(\frac{T_{21}}{T_{34}}\right) = d$, $\left(\frac{T_{22}}{T_{34}}\right) = e$, $\left(\frac{T_{23}}{T_{34}}\right) = f$.

$$\therefore\ x = a\left(\frac{X_1}{X_4}\right) + b\left(\frac{X_2}{X_4}\right) + c\left(\frac{X_3}{X_4}\right) + t_x$$
$$= aX + bY + cZ + t_x$$

$$\&\ y = d\left(\frac{X_1}{X_4}\right) + e\left(\frac{X_2}{X_4}\right) + f\left(\frac{X_3}{X_4}\right) + t_y$$
$$= dX + eY + fZ + t_y$$

$$\therefore \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

which is nothing but equation of an affine transformation i.e. non-uniform scaling & translation

P.T.O.

Weak perspective (wp) projection :-

Variation in the depth of an object is very less compared to the average depth of that object.
we have,

$$x = f\frac{X}{Z} \quad , \quad y = f\frac{Y}{Z}$$

Let Z_{avg} be the average depth & Δz be ~~the~~ a small incremental depth value.

$$\therefore x = f\frac{X}{Z_{avg} + \Delta Z} = f\frac{X}{Z_{avg}\left(1 + \frac{\Delta Z}{Z_{avg}}\right)}$$

But $\frac{\Delta Z}{Z_{avg}} \approx$ small.

$$\therefore 1 + \frac{\Delta z}{Z_{avg}} \approx 1.$$

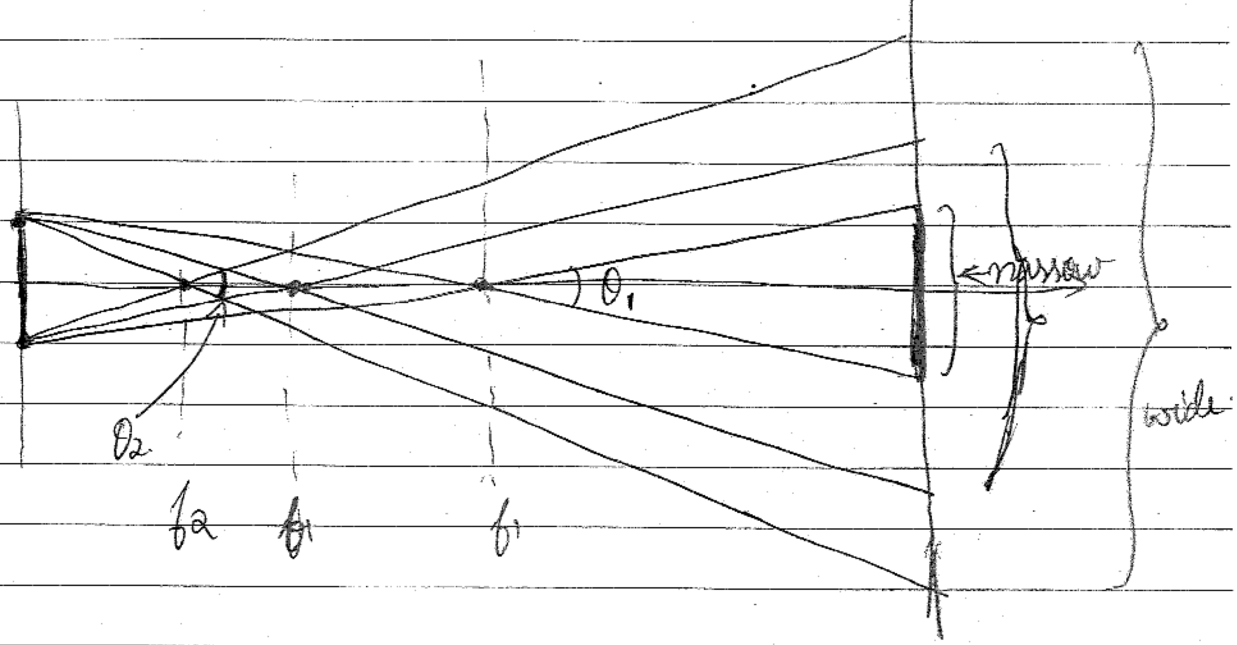
$$\therefore x = f\frac{X}{Z_{avg}}$$

Similarly $y = f\frac{Y}{Z_{avg}}$.

Let $\frac{f}{Z_{avg}} = m$.

$\therefore x = mX$ & $y = mY$ ... scaled orthographic projectio

Zoom Lens / ~~Telescophic~~ Telephoto lenses can be accurately modelled by using weak-perspective projection.

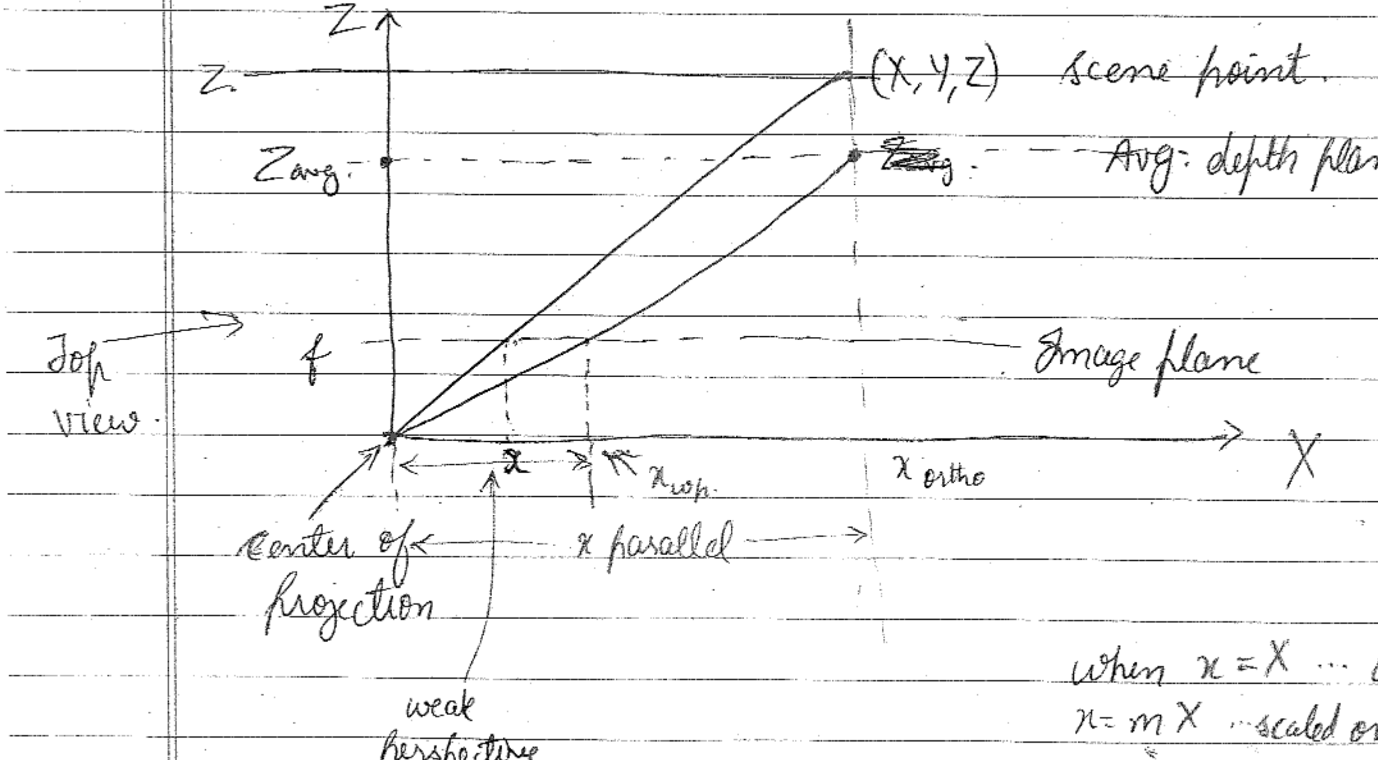


Parallel projection.

If $m=1$ then the weak perspective projection becomes parallel projection.

$\therefore x = X$ & $y = Y$

Thus, the rays are parallel to the optical axis.



when $x = X$ ... orthographic

$x = mX$ ... scaled orthographic

Lecture 6

07/08/2012.

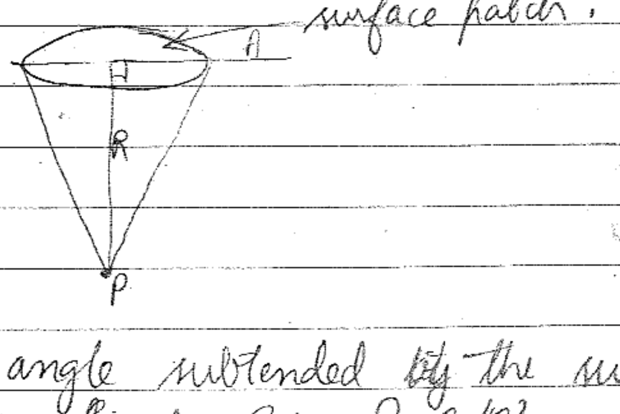
Physics of image formation.

→ Photometric transformations.
→ What is the value of the image formed at a point?
Terms:-
ⓐ Solid angle, ⓑ Irradiance ⓒ Radiance.

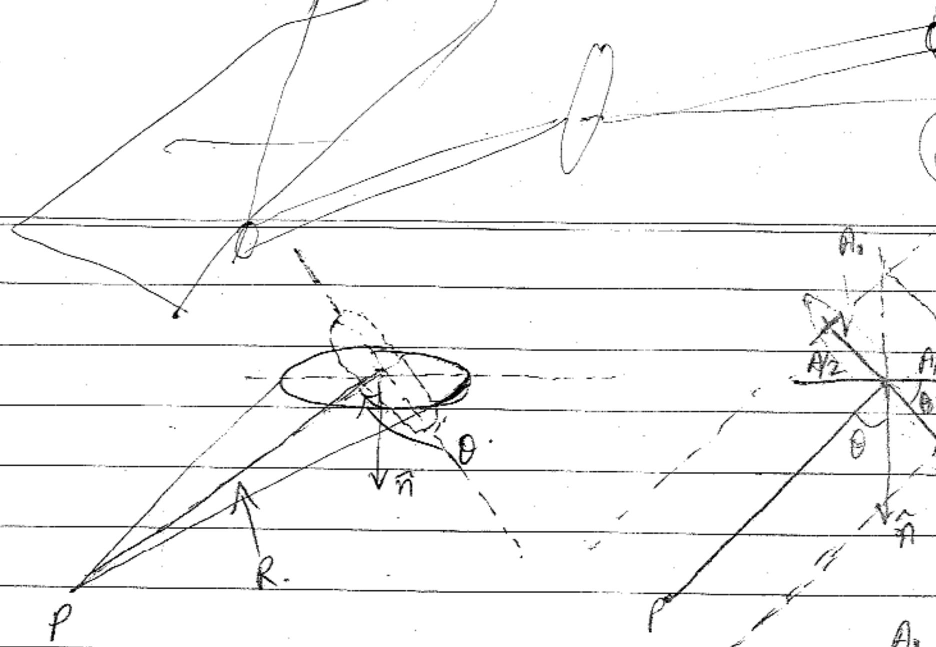
Irradiance:-
→ Amount of light falling on an object.
→ Measured in Watts/m^2
→ Denoted by (E).

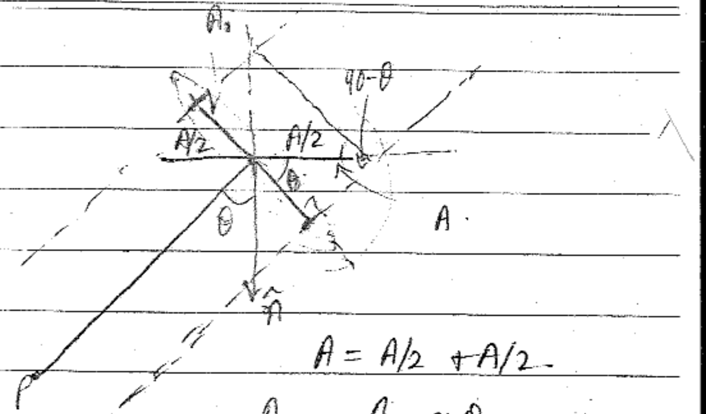
Radiance:-
→ Amount of light emitted / reflected from a surface.
→ Measured in Watts/m^2/Sr.
→ Denoted by (L)

Solid Angle:-
→ Denoted by Ω.
→ Measures the proportion of space contained between a point & a surface.



∴ Solid angle subtended by the surface patch as seen from the point P is $\Omega = A/R^2$
If the patch is tilted or not completely visible, then the solid angle by the patch as seen from P is given by $\Omega = \frac{\text{foreshortened area}}{R^2}$





$$A = A/2 + A/2$$

$$\frac{A_1}{2} = \frac{A}{2}\cos\theta$$

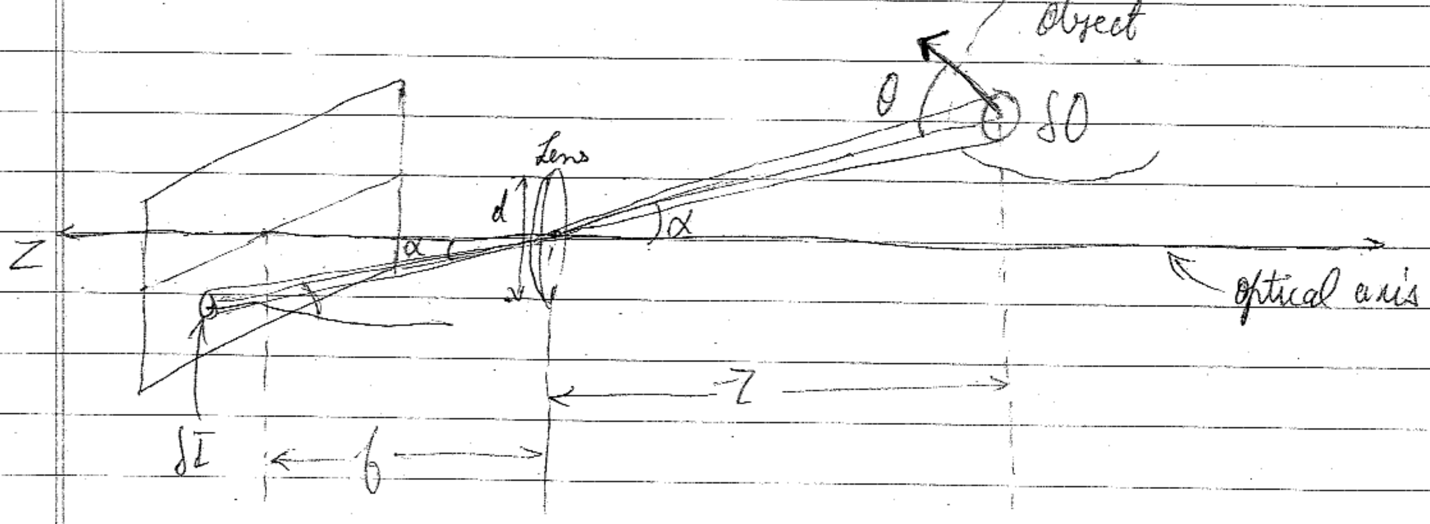
$$\therefore A_1 = A\cos\theta$$

$$\cos\theta = \frac{A_1/2}{A/2}$$

$$\therefore \Omega = \frac{A\cos\theta}{R^2}$$

Irradiance - Radiance Relation.

~~Image Formation~~: Pinhole model.



Derivation:

Lens diameter : d

Object surface area : δO

Image patch area : δI

Focal length : f

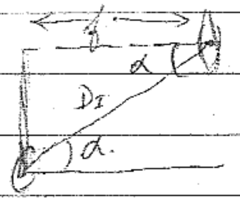
Distance of object pt. : -Z

Angle of foreshortening : θ.

Direction of the object w.r.t. the optical axis : α

∴ Solid angle subtended by image patch as seen from optical center is

$$\Omega_I = \frac{\delta I \cos\alpha}{(D_I)^2}$$

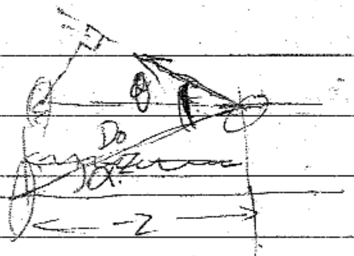


But $\cos\alpha = \frac{f}{D_I}$.

$$\therefore D_I = f/\cos\alpha.$$

$$\therefore \Omega_I = \frac{\delta I \cos\alpha}{(f/\cos\alpha)^2} = \frac{\delta I \cos^3\alpha}{f^2}$$

Similarly,



$$\Omega_o = \frac{\delta O \cos\theta}{(D_o)^2}$$

But $\cos\alpha = \frac{-Z}{D_o}$.

$$\therefore D_o = \frac{-Z}{\cos\alpha}$$

$$\therefore \Omega_o = \frac{\delta O \cos\theta}{(Z/\cos\alpha)^2} = \frac{\delta O \cos\theta \cdot \cos^2\alpha}{+Z^2}$$

For a pinhole model, $\Omega_I = \Omega_o$.

$$\therefore \frac{\delta I}{f^2}\cos^3\alpha = \frac{\delta O}{Z^2}\cos\theta \cos^2\alpha$$

$$\therefore \boxed{\frac{\delta I}{\delta O} = \left(\frac{f}{Z}\right)^2 \frac{\cos\theta}{\cos\alpha}} \quad \text{— (1)}$$

Solid angle subtended by the lense as seen from the object patch is,

$$\Omega = \frac{\pi (d/2)^2 \cos\alpha}{(z/\cos\alpha)^2} = \frac{\pi}{4}\frac{d^2}{} \frac{\cos^3\alpha}{z^2}.$$

$$\therefore \boxed{\Omega = \frac{\pi}{4}\left(\frac{d}{z}\right)^2 \cos^3\alpha} \qquad \text{—(2)}$$

Scene Radiance L i.e. light emitted from the object patch is,

$$L = \frac{\delta P}{[\delta O \cos\theta]\cdot \Omega} = \frac{\delta P}{\delta O \cos\theta \cdot \left(\frac{\pi}{4}\right)\left(\frac{d}{z}\right)^2 \cos^3\alpha}.$$

forshortening

$$\therefore \boxed{\delta P = L \cdot \delta O \cos\theta \left(\frac{\pi}{4}\right)\left(\frac{d}{z}\right)^2 \cos^3\alpha} \qquad \text{—(3)}.$$

This power is completely absorbed by the image patch i.e. image irradiance E.

$$\therefore E = \frac{\delta P}{\delta I}$$

... here no foreshortening is considered as the power is absorbed completely by the image patch.

$$\therefore E = L \cdot \frac{\delta O}{\delta I} \cos\theta \left(\frac{\pi}{4}\right)\left(\frac{d}{z}\right)^2 \cos^3\alpha.$$

From (1) we have

$$\boxed{E = L\left(\frac{z}{f}\right)^2 \frac{\cos\alpha}{\cos\theta} \cdot \cos\theta \cdot \left(\frac{\pi}{4}\right)\left(\frac{d}{z}\right)^2 \cos^3\alpha = L \cdot \left(\frac{\pi}{4}\right)\left(\frac{d}{f}\right)^2 \cos^4\alpha}$$

~~Here~~, for a camera, both 'd' and 'f' are fixed. Also $(\pi/4)$ is a constant. If α is constant, then

$$\boxed{E \propto L}$$

The term $\left(\frac{f}{d}\right)$ is called as the "f" number. It gives quantitive measure of lens speed & is a dimensionless quantity.

Lecture 7

13/08/2012.

Scene radiance & image irradiance.

$E \propto L$ & $E \propto (\text{angle})^4$

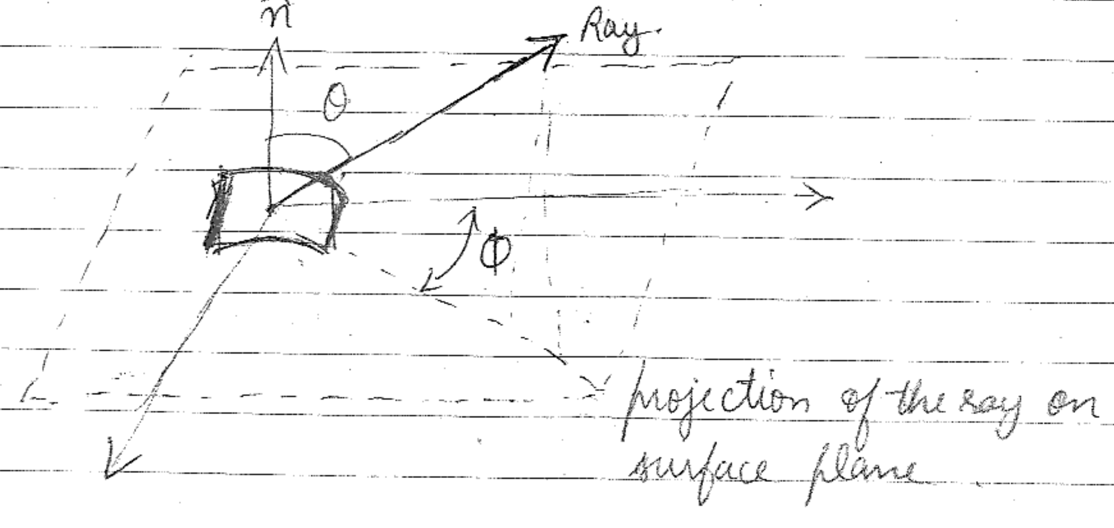
→ E is decreasing as we move away from the image plane.
→ Not a problem for zoom-lens (narrow field of view).

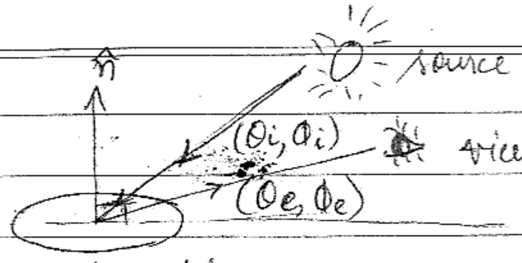
Bidirectional Reflectance Distribution Function (BRDF)

→ Defined for the object.
→ Tells us about how bright a surface patch looks when the light falls on the surface from some direction & gets reflected in the viewing direction.

→ What factors are going to determine L?
→ ① L depends on amount of light falling on surface.
→ ② depends on the amount of light getting reflected from the surface.
→ ③ Property of the object / material / scene / surface. viz. matte / specular / mixture of matte & specular.

L → Radiance (outgoing) θ → polar angle
E → Irradiance (in-coming) ϕ → azimuth angle.





BRDF which is a function of $\theta_i, \phi_i, \theta_e, \phi_e$ is defined as the ratio of radiance of the the scene as viewed from the direction (θ_e, ϕ_e) to the irradiance, resulting due to illumination in directio (θ_i, ϕ_i).

$$\therefore \boxed{f(\theta_i \phi_i, \theta_e, \phi_e) = \frac{\delta L(\theta_e, \phi_e)}{\delta E(\theta_i, \phi_i)}} \quad \text{—} (1)$$

Helmholtz Reciprocity Theorem:-

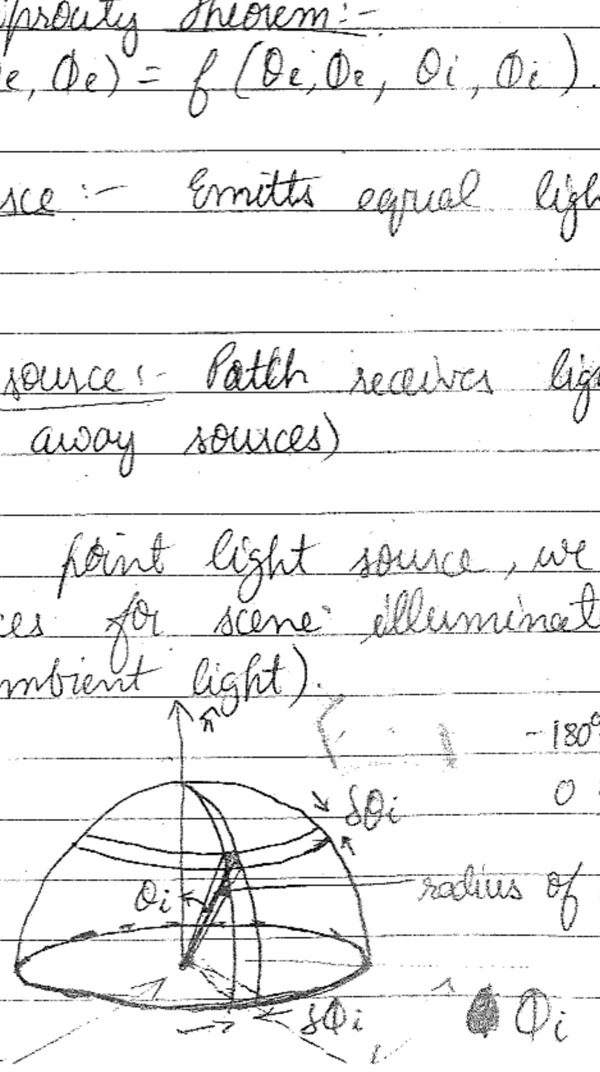
$$f(\theta_i, \phi_i, \theta_e, \phi_e) = f(\theta_e, \phi_e, \theta_i, \phi_i).$$

Point light source :- Emitts equal light / energy in al directions.

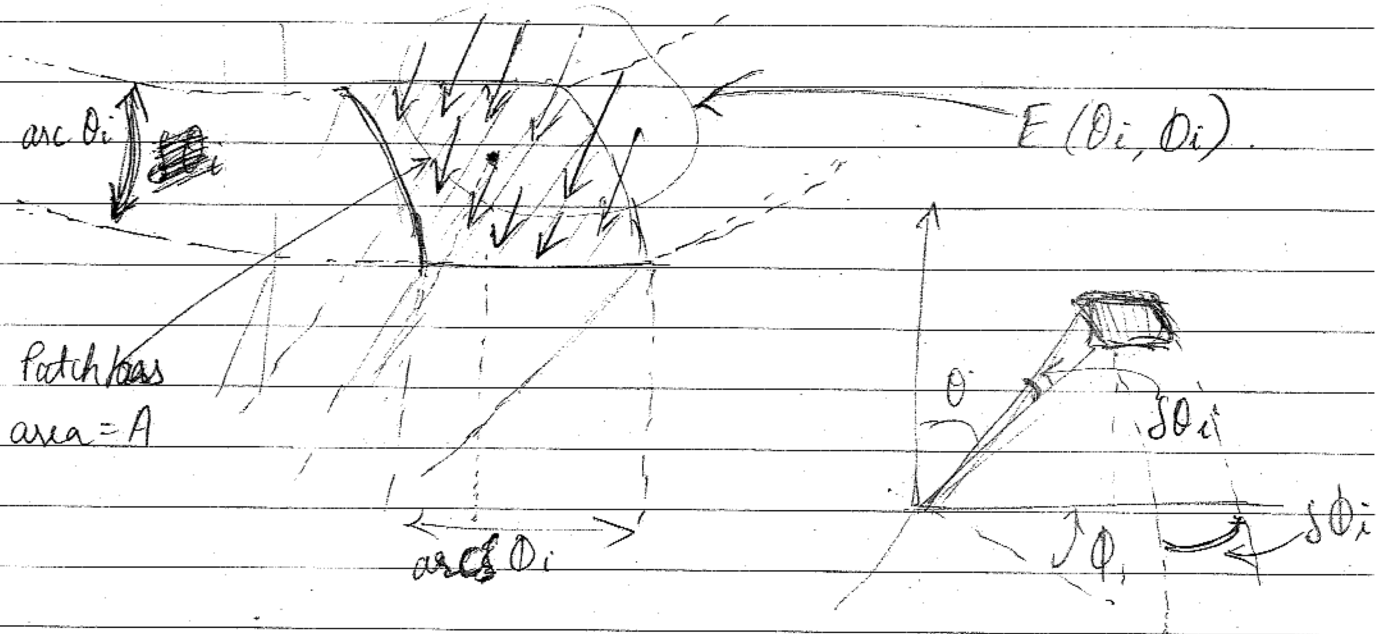
Extended light source :- Patch receives light from many sources (far away sources)

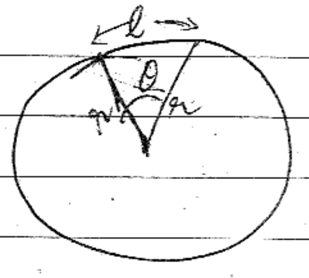
Instead of a point light source, we have a number of light sources for scene illumination.
Eg:- sky (ambient light).

$-180° \leq \phi \leq 180°$
$0 \leq \theta \leq 90°$.



Let $E(\theta_i, \phi_i)$ be radiance per unit solid angle from the direction (θ_i, ϕ_i)





$$\frac{l}{2\pi r} = \frac{\theta}{2\pi} \quad \therefore \quad l = \theta r.$$

$$\therefore \boxed{\text{arc } \theta_i = \delta\theta_i \cdot r}$$

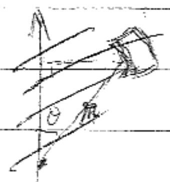
~~$E(\theta_i, \phi_i) = \frac{A \sin\theta_i}{r^2}$~~ Similarly $\boxed{\text{arc } \phi_i = \delta\phi_i \cdot r}$

$\therefore$ Area of patch under consideration is
$A = (r\,\delta\theta_i)(r\,\delta\phi_i)$.
However, as $\theta_i \to 0$, we observe that this area shrinks; while it is maximum when $\theta = \pi/2$ or $90°$.

$$\therefore \boxed{A = (r\,\delta\theta_i)(r\cdot\delta\phi_i)\cdot\sin\theta_i}$$ ~~foreshortened area as seen from the center of hemisphere~~

$\therefore$ The solid angle subtended by the patch as seen from the point of convergence is A/r^2 i.e.

$$\delta\omega_i = \frac{r\,\delta\theta_i\, r\,\delta\phi_i \sin\theta_i}{r^2} \qquad \delta\theta_i\,\delta\phi_i \sin\theta_i$$



~~Since we assume that the light falling on the patch is same as the light emitted by the source, we have the patch radiance equal to $E(\theta_i, \phi_i)$ per unit solid angle from direction (θ_i, ϕ_i) equal to $E(\theta_i, \phi_i)$.~~

~~Then, the~~

Let, $E(\theta_i, \phi_i)$ be the radiance per unit solid angle from the direction (θ_i, ϕ_i). Then, the radiance from the patch under consideration is

$$E(\theta_i \phi_i) \cdot \delta\omega$$

$$= E(\theta_i \phi_i) \cdot \delta\theta_i \, \delta\phi_i \sin\theta_i$$

which is nothing but the light coming inside the sphere from the patch under consideration. i.e. light over the total solid angle of that patch.

Now if we integrate the radiance from all patches we get the total incoming light from all direction or the total irradiance of the surface.

$$\therefore E_o = \int_{\theta=0}^{\pi/2} \int_{\phi=-\pi}^{\pi} E(\theta_i, \phi_i) \sin\theta_i \cos\theta_i \, d\theta_i \, d\phi_i \, , \qquad \text{--- (2)}$$

where $\cos\theta_i$ accounts for the foreshortening of the surface as seen from the direction (θ_i, ϕ_i).

... see fig. on p. 23

Now, $BRDF = \frac{Radiance}{Irradiance}$.

$\therefore$ Radiance = product of BRDF & irradiance.

$$\therefore L(\theta_e, \phi_e) = \int_{\theta=0}^{\pi/2}\int_{\phi=-\pi}^{\pi} f(\theta_i, \phi_i, \theta_e, \phi_e) \cdot E(\theta_i, \phi_i) \cdot \sin\theta_i \cdot \cos\theta_i \, d\theta_i \, d\phi_i$$

where again $\cos\theta_i$ accounts for foreshortening of the patch as seen from the direction (θ_i, ϕ_i).

We have considered this scenario for an extended light source and a Lambertian surface.

Lambertian surface:-

(a) Appears equally bright from all viewing directions independent of (θ_e, ϕ_e)

(b) Reflects all incident light & absorbs none.

(c) $f(\theta_i, \phi_i, \theta_e, \phi_e)$ = constant.

Eg: wood, paper.

For a Lambertian surface, BRDF is constant = f.

$$\therefore \quad L(\theta_e, \phi_e) = f \cdot E_o.$$

Now, $L(\theta_e, \phi_e)$ represents the light emitted in the direction (θ_e, ϕ_e). Then the total light emitted in the complete ~~hemisphere~~ patch hemisphere is given by the total radiance L as follows.

$$L = \int_{\theta=0}^{\pi/2}\int_{\phi=-\pi}^{\pi} L(\theta_e, \phi_e) \cdot \underbrace{\sin\theta_e}_{\text{for area depending on } \theta} \underbrace{\cos\theta_e}_{\text{for foreshortening}} \, d\theta_e \, d\phi_e$$

Since all the incident light is completely reflected, $L = E_0$.

$$\therefore E_0 = \int_{\theta=0}^{\pi/2}\int_{\phi=-\pi}^{\pi} f.E_0 \sin\theta_e \cos\theta_e \, d\theta_e \, d\phi_e$$

$$\therefore 1 = f \cdot \int_{\theta=0}^{\pi/2}\int_{\phi=-\pi}^{\pi} \sin\theta_e \cos\theta_e \, d\theta_e \, d\phi_e$$

$$= f \cdot [\phi]_{-\pi}^{\pi} \int_0^{\pi/2} \sin\theta_e \cos\theta_e \, d\theta_e$$

$$= f \cdot [\pi - (-\pi)] \int_0^{\pi/2} \sin\theta_e \cos\theta_e \, d\theta_e$$

$$= 2\pi f \int_0^{\pi/2} \sin\theta_e \cos\theta_e \, d\theta_e$$

$$= 2\pi f \int_0^{\pi/2} \frac{2 \sin\theta_e \cos\theta_e}{2} \, d\theta_e$$

$$= \frac{2\pi f}{2} \int_0^{\pi/2} \sin(2\theta_e) \, d\theta_e$$

$$= \pi f \cdot \left[\frac{-\cos(2\theta_e)}{2}\right]_0^{\pi/2}$$

$$= \pi f \left[\frac{-\cos(\pi) - (-\cos 0)}{2}\right]$$

$$= \frac{\pi f}{2}\left[\cos 0 - \cos\pi\right] = \frac{\pi f}{2}\left[1 - (-1)\right] = \frac{2\pi f}{2} = \pi f$$

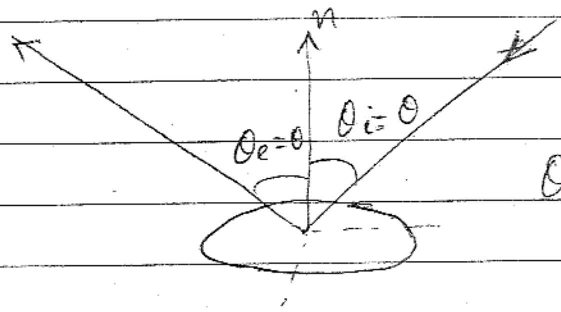
$\therefore \boxed{f = \frac{1}{\pi}}$

$\therefore \quad f(\theta_i \, \phi_i \, \theta_e \, \phi_e) = \frac{1}{\pi}$

$$\boxed{L(\theta_e \phi_e) = \frac{1}{\pi} E_0}$$

Lecture 8

Ideal specular surface / Mirror-like surface



$\theta_e = \theta_i = \theta$ for a specular surface.

Also, $\phi_e = \phi_i + \pi$

Specular surface reflects light arriving from direction (θ_i, ϕ_i) to direction (θ_e, ϕ_e) i.e. $(\theta_i, \phi_i + \pi)$.

Here, the BRDF is proportional to $\delta(\theta_e - \theta_i)$ & $\delta(\phi_e - (\phi_i + \pi))$.

$$\therefore \text{BRDF} \propto \delta(\theta_e - \theta_i)\, \delta(\phi_e - \phi_i - \pi)$$
$$= k \cdot \delta(\theta_e - \theta_i)\, \delta(\phi_e - \phi_i - \pi), \quad \int \delta(\theta_e - \theta_i)\, d\theta_i = 1$$

But, what is k?

→ Total radiance = $\int$ emittance in all directions &

→ Total radiance = irradiance.

$$\therefore \int_{\theta=0}^{\pi/2} \int_{\phi=-\pi}^{\pi} k \cdot \delta(\theta_e - \theta_i)\, \delta(\phi_e - \phi_i - \pi) \sin\theta_e \cdot \cos\theta_e \, d\theta_e \, d\phi_e = 1$$

R.H.S. = 1 because, total irradiance = $\int \delta(\theta_e - \theta_i)\, d\theta_i = 1$.

$$\cancel{\therefore \int_{\theta=0}^{\pi/2} \int_{\phi=-\pi}^{\pi} k \cdot 1 \cdot 1 \cdot \sin\theta_e \cos\theta_e \, d\theta_e \, d\phi_e = 1}$$

$$\therefore \frac{k}{2}\left[\phi_e\right]_{-\pi}^{\pi} \int \frac{2\sin\theta_e \cos\theta_e \, d\theta_e}{2} = 1.$$

$$\therefore \pi k \int \sin(2\theta_e)\, d\theta_e = 1 \qquad \text{Here, } k = \frac{1}{\sin\theta_i \cos\theta_i}$$

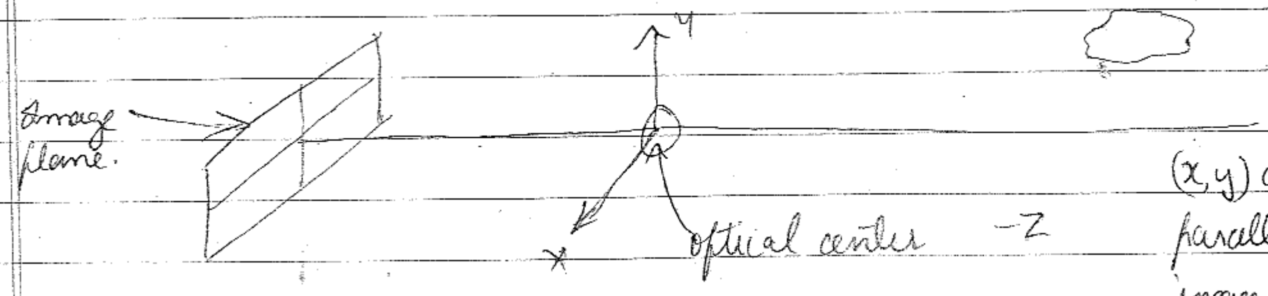
or $\quad k \cdot \sin\theta_i \cos\theta_i = 1.$

$$\therefore \boxed{f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{\delta(\theta_e - \theta_i)\, \delta(\phi_e - \phi_i - \pi)}{\sin\theta_i \cos\theta_i}}$$

In some of the methods used for finding the depth from images (eg: shape from shading), we make an assumption about the light source. We take the source to be a point light source.
Eg:- Stars, etc. ∴ pt. light source in astronomy.

Surface Orientation. / Normal to a patch.

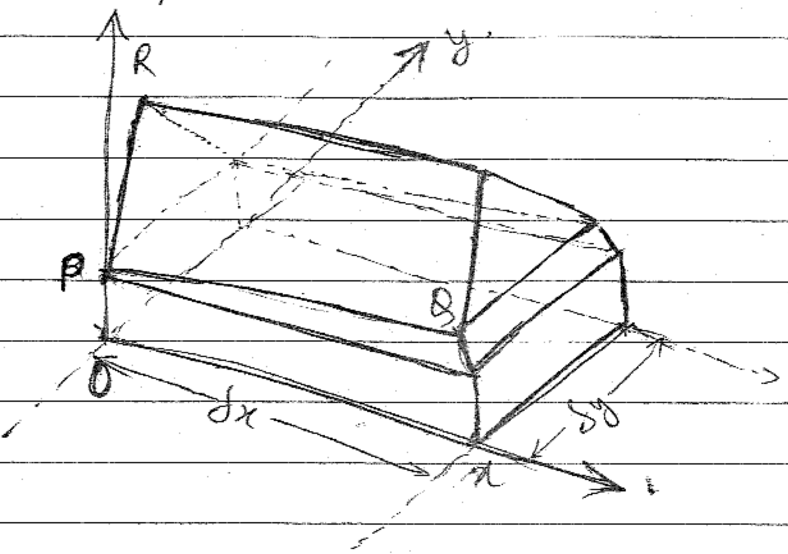
Assume the surface to be smooth i.e. has tangent plane at every point. The surface normal, ~~which~~ is a unit vector perpendicular to the tangent plane.



(x, y) axis is parallel to the image plane

Consider depth as a 2D function given by $Z(x, y)$.
This is also known as Monge representation of the surface

Consider an infinitesimily small patch of the surface (planar patch).



$$Z(x,y) = Z(x_0,y_0) + \frac{\partial Z}{\partial x}(x,y)\Big|_{x_0,y_0}\delta x + \frac{\partial Z}{\partial y}(x,y)\Big|_{x_0,y_0}\delta y + \underbrace{\cdots\cdots}_{\text{neglected}}$$

... By using Taylor's series.

Now, $\delta Z = \frac{\partial Z}{\partial x}\delta x + \frac{\partial Z}{\partial y}\cdot\delta y = p\delta x + q\delta y$.

$$\left.\begin{aligned}\overrightarrow{PQ} &= (\delta x, 0, p\,\delta x) = (1, 0, p)\\ \overrightarrow{PR} &= (0, \delta y, q\,\delta y) = (0, 1, q)\end{aligned}\right\}$$ why? ... take $\overrightarrow{PQ}\times\overrightarrow{PR}$

divide by increments.

$\overrightarrow{PQ}\times\overrightarrow{PR}$ gives normal at P i.e. $\hat{n}$.

$$\hat{n} = \begin{vmatrix} i & j & k \\ 1 & 0 & p \\ 0 & 1 & q \end{vmatrix} = i(-p) + j(-q) + k.$$

$$= \begin{pmatrix} -p \\ -q \\ 1 \end{pmatrix}$$ ... since $p = \frac{\partial z}{\partial x}$ & $q = \frac{\partial z}{\partial y}$,

p & q are surface gradients.

Lecture 9

17/08/2012.

Surface Expression:

$$\hat{n} = \begin{pmatrix} -p \\ -q \\ 1 \end{pmatrix}$$

Since the normal is only a direction, it has a unit magnitude.

$$\therefore \hat{n} = \begin{bmatrix} \dfrac{-p}{\sqrt{p^2+q^2+1}} \\[2ex] \dfrac{-q}{\sqrt{p^2+q^2+1}} \\[2ex] \dfrac{1}{\sqrt{p^2+q^2+1}} \end{bmatrix}$$

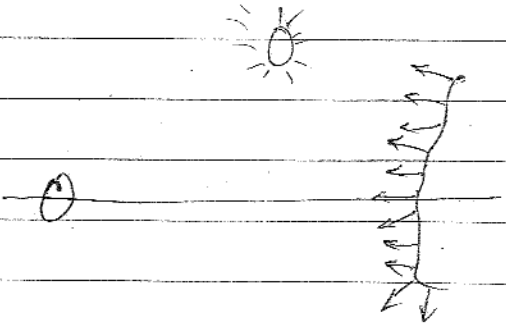
$$p = \frac{\partial z(x,y)}{\partial x}, \quad q = \frac{\partial z(x,y)}{\partial y}$$

p, q are surface gradients in x & y directions, respectively.

(p, q) exist for ~~ev~~ every point on the scene/surface.

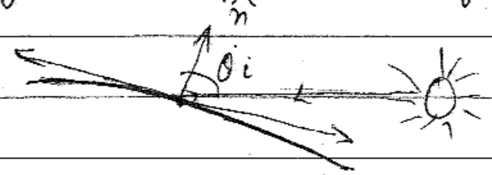
Consider a point light source emitting light on an object

$$\hat{s} = \begin{bmatrix} \dfrac{-p_s}{\sqrt{p_s^2+q_s^2+1}} \\[2ex] \dfrac{-q_s}{\sqrt{p_s^2+q_s^2+1}} \\[2ex] \dfrac{1}{\sqrt{p_s^2+q_s^2+1}} \end{bmatrix} \Rightarrow \text{direction of the point light source.}$$



Reflectance Map:-

Gives relationship between surface orientation & the brightness of the surface.



Consider a point light source and a Lambertian surface

$$\therefore L = \frac{E}{\pi} \cos\theta_i$$

scene radiance

source radiance

angle between source direction & the normal to the patch.

$\theta_i \geq 0$.

Now, $\hat{n}\cdot\hat{s} = |\hat{n}||\hat{s}| \cos\theta_i$

But $|\hat{n}| = |\hat{s}| = 1$.

$$\therefore \hat{n}\cdot\hat{s} = \cos\theta_i$$

$$\therefore \cos\theta_i = \frac{(-p, -q, 1)\cdot(-p_s, -q_s, 1)^T}{\sqrt{p^2+q^2+1}\ \sqrt{p_s^2+q_s^2+1}}$$

$$= \frac{pp_s + qq_s + 1}{\left(\sqrt{p^2+q^2+1}\right)\left(\sqrt{p_s^2+q_s^2+1}\right)}$$

The reflectance map of a Lambertian surface illuminated by a pt. light source is denoted by

$R(p,q) = \cos\theta_i$

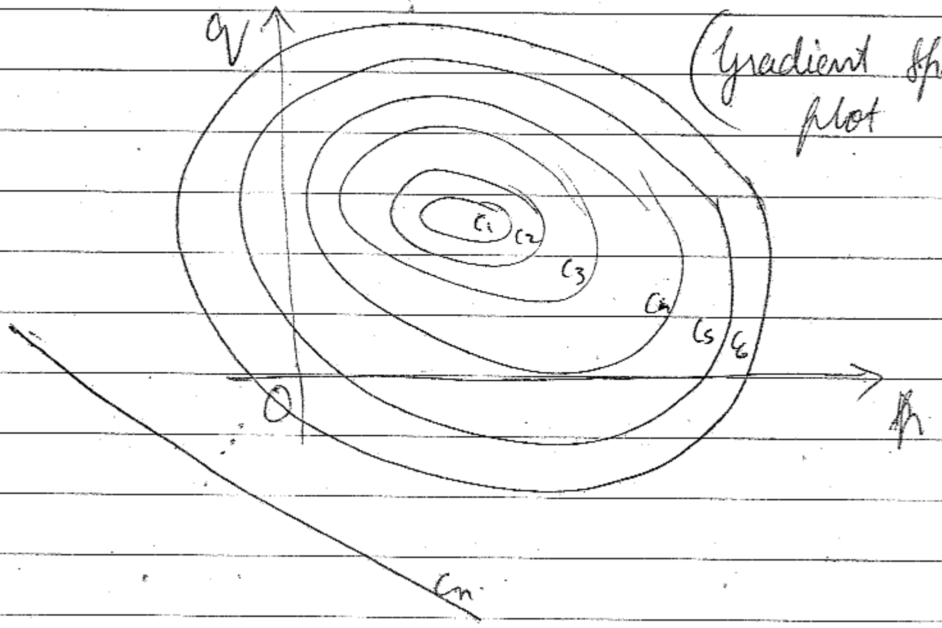
$$\therefore \boxed{R(p,q) = \cos\theta_i = \frac{p p_s + q q_s + 1}{\sqrt{p_s^2+q_s^2+1}\,\sqrt{p^2+q^2+1}}}$$

& since $\theta_i \geq 0$ we have $0 \leq \cos\theta_i \leq 1$

$\therefore$ $0 \leq R(p,q) \leq 1$

Let $R(p,q) = C$

If p_s, q_s are known then we can have many (p,q) such that $R(p,q) = C$; giving a plot like a circle. If we have $C_1, C_2, C_3, \ldots C_n$ as shown in the following plot, then for C_n, we would have an arc such that it is a straight line.



(Gradient Space plot) or (plot of $R(p,q)$ in p,q.)

When $R(p,q) = 0$ we have,

$$\frac{p p_s + q q_s + 1}{\sqrt{p_s^2+q_s^2+1}\,\sqrt{p^2+q^2+1}} = 0.$$

$$\therefore \; p p_s + q q_s + 1 = 0.$$

$$q = \left(-\frac{p_s}{q_s}\right) p + \left(\frac{-1}{q_s}\right) \quad \ldots \; mx + C \text{ form}$$

Shape from Shading:

Use shading cue (hint) to solve for 'z' depth.

Shading:-

We know that $R(p,q)$ captures the dependance of brightness on surface orientation. The image irradiance $E(x,y)$ at a point is proportional to the radiance at a corresponding point on the surface. If (p,q) is generated at a point on the surface and a constant of proportionality is considered as 1, then we have.

$$L = \frac{E}{\pi} \cos\theta_i = \frac{E}{\pi} R(p,q).$$

$$L \propto R(p,q).$$

But $E \propto L.$

$$\therefore E \propto R(p,q)$$

$$\therefore E(x,y) = t \cdot R(p,q) \quad \& \quad t = 1$$

$$\therefore E(x,y) = R(p,q)$$

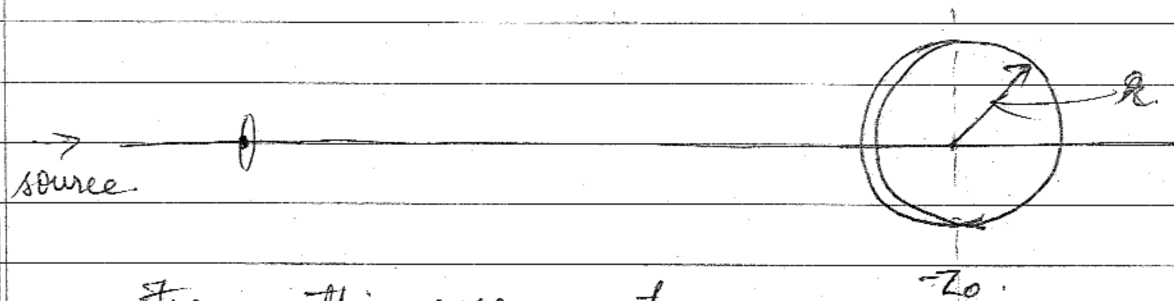
i.e. $\boxed{E(x,y) = R\left(p(x,y), q(x,y)\right)}$

Lecture 10

$E(x,y) = R(p,q)$.

If (p,q) varies smoothly (gradual depth variation) as a smooth surface, the image brightness pattern will change smoothly, giving rise to what is known as shading.

Eg: Consider a sphere on the optical axis.



From this arrangement,
For the source, the source position becomes $(0,0,f)$.

$\therefore \frac{\partial f}{\partial x} = p_s = 0$ & $\frac{\partial f}{\partial y} = q_s = 0$.

$$\therefore \hat{s} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\therefore R(p,q) = \frac{p(0)+q(0)+1}{\sqrt{0^2+0^2+1}\sqrt{p^2+q^2+1}} = \frac{1}{\sqrt{p^2+q^2+1}}$$

Now, $Z = Z_0 + \sqrt{r^2-(x^2+y^2)}$ for $\forall\ x^2+y^2 \geq r$.

$$\therefore p = \frac{\partial z}{\partial x} = \frac{1}{2\sqrt{r^2-(x^2+y^2)}}(-2x) = \frac{-x}{\sqrt{r^2-(x^2+y^2)}} = \frac{-x}{Z-Z_0}$$

Similarly $q = \frac{\partial z}{\partial y} = \frac{-y}{\sqrt{r^2-(x^2+y^2)}} = \frac{-y}{Z-Z_0}$

Now, $R(p,q) = \frac{1}{\sqrt{p^2+q^2+1}}$ for $\hat{s} = (0,0,1)^T$

$$= \frac{1}{\sqrt{\left(\frac{-x}{z-z_0}\right)^2 + \left(\frac{-y}{z-z_0}\right)^2 + 1}}$$

$$= \frac{1}{\sqrt{\frac{x^2+y^2+(z-z_0)^2}{(z-z_0)^2}}}$$

$$= \frac{z-z_0}{\sqrt{x^2+y^2+\left[r^2-(x^2+y^2)\right]}}$$

$$= \frac{z-z_0}{\sqrt{r^2}} = \frac{z-z_0}{r}.$$

$$= \frac{\sqrt{r^2-(x^2+y^2)}}{r}.$$

$$\therefore \boxed{R(p,q) = \sqrt{1-\frac{x^2+y^2}{r^2}}}$$

i.e. $R\left(p(x,y), q(x,y)\right) = \sqrt{1-\frac{(x^2+y^2)}{r^2}}$

∴ function of function = functional & can be solved using calculus of variation.

Exercise 1 :- Image Rotation.
Exercise 2 :- (a) For source at (0,0,1), get images from $R(p,q)$.
(2) For source at (0.5,0.5,1) get $R(p,q)$ & $E(x,y)$ & show images. ... 2 weeks deadline.

For $R(p,q)$ we need to find p, q, given 1 equation i.e. 2 degrees of freedom & 1 constraint.

Shape From Shading :-

Given (p,q), unique mapping to $E(x,y)$. But given $E(x,y)$, many solutions to (p, q) as.

$$\underset{\text{given}}{E(x,y)} = \frac{p p_s + q q_s + 1}{\sqrt{p_s^2 + q_s^2 + 1}\ \underbrace{\sqrt{p^2+q^2+1}}_{\text{many}}}$$

SFS :- A single image is given & object is imaged under controlled lighting conditions (not ambient light). Assume source position is known for a pt. light source & find the shape / depth.

To solve, we make assumptions,
1. Object is Lambertian
2. Source is pt. light source.
3. 3D-2D mapping, we will assume parallel projection i.e. $x = X$, $y = Y$ i.e. $Z(X,Y) = z(x,y)$.

solution: We have $E(x,y) = R(p,q) = \dfrac{p p_s + q q_s + 1}{\sqrt{p^2+q^2+1}\ \sqrt{p_s^2+q_s^2+1}}$

Given image $\Rightarrow$ get p-map + q-map & later Z map.

Lecture 11

since we have 1 equation & 2 unknowns (p, q) at every (x, y) there are infinite solutions for (p, q).

The problem can be solved (or a better solution) can be obtained by considering the problem as follows.

Considering the continuous case, a smoothness (contraint) can be imposed on the solution by considering / formulating the problem as

minimize for all p, q. $\displaystyle\iint_{x\,y} (p_x^2 + p_y^2 + q_x^2 + q_y^2)\, dx\, dy$.

$p_y = \dfrac{\partial p}{\partial y}$ $\quad p_x = \dfrac{\partial p}{\partial x}$, $\quad q_y = \dfrac{\partial q}{\partial y}$, $\quad q_x = \dfrac{\partial q}{\partial x}$.

& $p = \dfrac{\partial z}{\partial x}$ & $q = \dfrac{\partial z}{\partial y}$

→ subject to the condition, $E(x, y) = R(p, q)$.

Constriained minimization can be solved using Lagrange Multiplier method.

<u>Unconstrained Minimization (Regularization)</u>

$E(x, y) = R(p, q)$ ... not always true. In practice, some noise due to sensor exists.

$\therefore E(x, y) = R(p, q) + n(x, y)$

With this hint, the problem can now be formulated as follows:-

Consider

$$E_1 = \iint_{x\,y} \left[E(x,y) - R(p(x,y), q(x,y))\right]^2 dx\,dy.$$

$$\& \; E_s = \iint_{x\,y} \left(p_x^2 + q_x^2 + p_y^2 + q_y^2\right) dx\,dy. \Rightarrow \text{Regularization Term.}$$

Minimize

$$\epsilon = \iint_{x\,y} \left[\left(p_x^2 + q_x^2 + p_y^2 + q_y^2\right) + \lambda\left(E(x,y) - R(p(x,y), q(x,y))\right)^2\right] dx\,dy.$$

∴ λ is taken to give weightage.

i.e.

minimize $\forall\, p, q$ $\quad \epsilon = \iint_{x\,y} \left[\left(p_x^2 + p_y^2 + q_x^2 + q_y^2\right) + \lambda(E-R)^2\right] dx\,dy.$

If we take $\lambda = 0$, then $\min_{\forall (p,q)} \epsilon = 0$

i.e. $p = 0$, $q = 0$ i.e. a flat surface & if we keep increasing λ, it will emphasize $(E-R)^2$ term.

$\underset{\text{p-map}}{\begin{bmatrix} 0 & 0 & 0 \\ \cdots & & 0 \end{bmatrix}} = \underset{\text{q-map}}{\begin{bmatrix} 0 & 0 & 0 \\ \cdots & & 0 \end{bmatrix}}$

λ depends on the chosen model.

(However λ can be obtained, given the data. Out of scope for this course.)

P.T.O.

Let $I = \iint\limits_{x\,y} F(p, q, p_x, q_x, p_y, q_y)\,dx\,dy$.

Then, $\underset{V(p,q)}{\text{minimize}}$ I for a functional.

... (Review the course on differential geometry in CV)

The solution can be obtained by using Calculus of Variations. There are 2 unknown functions $p(x,y)$ & $q(x,y)$ $\therefore$ 2 Euler equations.

$$F_p - \frac{\partial F_{p_x}}{\partial x} - \frac{\partial F_{p_y}}{\partial y} = 0, \quad \text{— (1)}$$

$$F_q - \frac{\partial F_{q_x}}{\partial x} - \frac{\partial F_{q_y}}{\partial y} = 0. \quad \text{— (2)}$$

& $F_p = \frac{\partial F}{\partial p}$, $F_q = \frac{\partial F}{\partial q}$.

Now,

$$F = p_x^2 + q_x^2 + p_y^2 + q_y^2 + \lambda\left[E(x,y) - R\left(p(x,y), q(x,y)\right)\right]^2$$

$$\therefore \frac{\partial F}{\partial p} = 2\left[\lambda(E-R)\cdot(-\partial R/\partial p)\right] \quad \& \quad \frac{\partial F}{\partial p_x} = 2p_x \quad \frac{\partial F}{\partial p_y} = 2p_y$$

$$\therefore 2\left[\lambda(E-R)(-\partial R/\partial p)\right] - \frac{\partial 2p_x}{\partial x} - \frac{\partial 2p_y}{\partial y} = 0.$$

$$\therefore p_{xx} + p_{yy} = -\lambda(E-R)\frac{\partial R}{\partial p} = \nabla^2 p.$$

$\therefore p_{xx} = \dfrac{\partial p_x}{\partial x}$, $p_{yy} = \dfrac{\partial p_y}{\partial y}$

$\therefore \nabla^2 p = -\lambda\,(E-R)\cdot R_p$.

Similarly $\nabla^2 q = -\lambda\,(E-R)\cdot R_q$.

Lecture 12

28/08/2012

Shape from Shading :-

Trying to solve by posing it as a minimizati problem.

i.e.

minimize $\forall p(x,y), q(x,y)$: $\epsilon = \iint \left[\left(p_x^2 + q_x^2 + p_y^2 + q_y^2 \right) + \lambda (E - R)^2 \right] dx\,dy.$

Considering continuous case i.e. x & y are real variables, we solve the problem using calculus of variations

Solution :-

$$p_{xx} + p_{yy} = \nabla^2 p = -\lambda (E-R) R_p \quad —①.$$

$$\& \quad q_{xx} + q_{yy} = \nabla^2 q = -\lambda (E-R) R_q \quad —②$$

Solve ① ② to get $p(x,y)$ & $q(x,y)$ for all (x,y).

Discrete case :-

When the intensities are available at discrete (x,y) i.e. say (i,j) $\forall i,j$

∇^2 i.e. Laplacian Operator can be represented by $\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}$ or $\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}$, we use the later.

$$\therefore \nabla^2 p = -4 p(i,j) + \left[p(i+1,j) + p(i-1,j) + p(i,j+1) + p(i,j-1) \right].$$

$$= -\lambda \left[E(i,j) - R\left(p(i,j), q(i,j) \right) \right] \frac{\partial R}{\partial p}$$

Let $\bar{p}(i,j) = \dfrac{p(i,j-1) + p(i,j+1) + p(i-1,j) + p(i+1,j)}{4}$

$$4p(i,j) = 4\bar{p}(i,j) + \lambda\left[E(i,j) - R(p(i,j), q(i,j))\right]\frac{\partial R}{\partial p}.$$

$$\therefore p(i,j) = \bar{p}(i,j) + \beta\left[E(i,j) - R(p(i,j), q(i,j))\right]\left.\frac{\partial R}{\partial p}\right|_{p(i,j)}, \quad \beta = \frac{\lambda}{4} \qquad (1)$$

Similarly,

$$\nabla^2 q = -4q(i,j) + \left[q(i,j-1) + q(i,j+1) + q(i+1,j) + q(i-1,j)\right].$$

$$\& \ \nabla^2 q = -\lambda(E-R)\frac{\partial R}{\partial q}.$$

Let $\bar{q}(i,j) = \dfrac{q(i,j+1) + q(i,j-1) + q(i+1,j) + q(i-1,j)}{4}.$

$$\therefore 4q(i,j) = 4\bar{q}(i,j) + \lambda\left[E(i,j) - R(p(i,j), q(i,j))\right]\left.\frac{\partial R}{\partial q}\right|_{q(i,j)}$$

$$\therefore q(i,j) = \bar{q}(i,j) + \beta\left[E(i,j) - R(p(i,j), q(i,j))\right]\left.\frac{\partial R}{\partial q}\right|_{q(i,j)}, \quad \beta = \lambda/4 \qquad (2)$$

Equations (1) & (2) can be solved iteratively as given every n^{th} iteration,

$$p^{n+1}(i,j) = \bar{p}^{n}(i,j) + \beta\left[E(i,j) + R(p^{n}(i,j), q^{n}(i,j))\right] \cdot \left.\frac{\partial R}{\partial p^{n}}\right|_{\forall i,j.}$$

$$\& \ q^{n+1}(i,j) = \bar{q}^{n}(i,j) + \beta\left(E(i,j) + R(p^{n}(i,j), q^{n}(i,j))\right)\left.\frac{\partial R}{\partial q^{n}}\right|_{\forall i,j.}$$

Directly minimizing the discrete case

Objective
Minimize
$\forall p(x,y)$ & $q(x,y)$

$$\iint_{x\,y} \left(\left(p_x^2 + q_x^2 + p_y^2 + q_y^2 \right) + \lambda \left(E - R \right)^2 \right) dx\,dy.$$

Directly minimizing the integral for discrete version, smoothness at a point can be given/measured by

$$s(i,j) \; (E_s) = \overset{\text{constant}}{\frac{1}{4}} \left[\left(p(i,j) - p(i-1,j) \right)^2 + \left(p(i,j) - p(i,j-1) \right)^2 + \left(q(i,j) - q(i-1,j) \right)^2 + \left(q(i,j) - q(i,j-1) \right)^2 \right]$$

The term corresponding to $(E-R)^2$ is.

$$r(i,j) = \left[E(i,j) - R\left(p(i,j), q(i,j) \right) \right]^2$$

$\therefore$ The discrete cost function to be minimized is now.

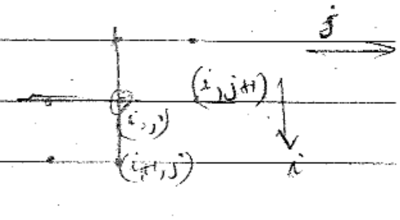
$$e = \min_{\substack{\forall p(i,j) \\ \& q(i,j)}} \sum_i \sum_j s(i,j) + \lambda r(i,j)$$

The minimization can be achieved by differentiating w.r.t $p(i,j)$ & $q(i,j)$ & equating it to 0.

$$\therefore \frac{\partial s(i,j)}{\partial p(i,j)} = \frac{1}{4} \left[\begin{array}{l} 2\left(p(i,j) - p(i-1,j) \right)(-1) \\ + 2\left(p(i,j \right. \end{array} \right.$$

Now, $\epsilon = \epsilon_s + \epsilon_r$.

$$= \sum_{i,j} s(i,j) + \sum_{i,j} \lambda r(i,j).$$



$$\epsilon_s = s(0,0) \cdots + s(i,j) + s(i+1,j) + s(i,j+1) + \cdots$$

$$= \cdots + \frac{1}{4}\Big[\big(p(i,j) - p(i-1,j)\big)^2 + \big(p(i,j) - p(i,j-1)\big)^2 + \big(q(i,j) - q(i-1,j)\big)^2 + \big(q(i,j) - q(i,j-1)\big)^2\Big]$$

$$+ \frac{1}{4}\Big[\big(p(i+1,j) - p(i,j)\big)^2 + \big(p(i+1,j) - p(i+1,j-1)\big)^2 + \big(q(i+1,j) - q(i,j)\big)^2 + \big(q(i+1,j) - q(i+1,j-1)\big)^2\Big]$$

$$+ \frac{1}{4}\Big[\big(p(i,j+1) - p(i-1,j+1)\big)^2 + \big(p(i,j+1) - p(i,j)\big)^2 + \big(q(i,j+1) - q(i-1,j+1)\big)^2 + \big(q(i,j+1) - q(i,j)\big)^2\Big]$$

$$+ \cdots$$

$$\frac{\partial \epsilon}{\partial p(i,j)} = \frac{\partial \epsilon_s}{\partial p(i,j)} + \frac{\partial \epsilon_r}{\partial p(i,j)}$$

$$\frac{\partial \epsilon_s}{\partial p(i,j)} = \frac{1}{4}\Big[2\big(p(i,j) - p(i-1,j)\big)(1) + 2\big(p(i,j) - p(i,j-1)\big)(1) + 2\big(p(i+1,j) - p(i,j)\big)(-1) + 2\big(p(i,j+1) - p(i,j)\big)(-1)\Big]$$

$$= \frac{1}{4}\Big[8\,p(i,j) - 2p(i-1,j) - 2p(i,j-1) - 2p(i+1,j) - 2p(i,j+1)\Big]$$

$$= 2p(i,j) - \frac{1}{2}\Big[p(i-1,j) + p(i,j-1) + p(i+1,j) + p(i,j+1)\Big]$$

Similarly,

$$e_r = \sum_{i,j} \lambda \cdot r_{(i,j)} = \sum_{i,j} \lambda \left[E(i,j) - R\left(p(i,j), q(i,j)\right)\right]^2$$

$$= \lambda\left[E(\cdot,\cdot) - R\left(p(\cdot,\cdot), q(\cdot,\cdot)\right)\right]^2$$

$$+ \vdots$$

$$+ \lambda\left[E(i,j) - R\left(p(i,j), q(i,j)\right)\right]^2$$

$$+ \vdots$$

$$\therefore \frac{\partial e_r}{\partial p(i,j)} = 2\lambda\left[E(i,j) - R\left(p(i,j), q(i,j)\right)\right] \cdot \frac{\partial R}{\partial p(i,j)} (-1).$$

$$\& \frac{\partial\left(R\left[p(i,j), q(i,j)\right]\right)}{\partial p(i,j)} = R_p.$$

$$\therefore \frac{\partial e_r}{\partial p(i,j)} = \cancel{R_p} - 2\lambda\left[E(i,j) - R\left(p(i,j), q(i,j)\right)\right] R_p.$$

Now, $\frac{\partial e}{\partial p(i,j)} = 0.$

$$\therefore 2\left\{p(i,j) - \frac{1}{2}\left(p(i-1,j) + p(i,j-1) + p(i+1,j) + p(i,j+1)\right)\right.$$

$$= 2\lambda\left[E(i,j) - R\left(p(i,j), q(i,j)\right)\right] R_p.$$

Let $p(i-1,j) + p(i,j-1) + p(i+1,j) + p(i,j+1) = 4\bar{p}(i,j).$

$$\therefore \; 2p(i,j) - \frac{1}{2}\, 4\bar{p}(i,j) = 2\lambda\left[E(i,j) - R(p(i,j), q(i,j))\right] R_p$$

$$\therefore \; p(i,j) = \bar{p}(i,j) + \lambda\left[E(i,j) - R(p(i,j), q(i,j))\right] R_p.$$

i.e. Let $\lambda = \beta$.

$$\boxed{p^{n+1}(i,j) = \bar{p}^{n}(i,j) + \beta\left[E(i,j) - R(p^{n}(i,j), q^{n}(i,j))\right] \cdot \frac{\partial R}{\partial p^{n}_{(i,j)}}}$$

similar $\dfrac{\partial \epsilon}{\partial q(i,j)} = 0$.

$$\therefore \; \boxed{q^{n+1}(i,j) = \bar{q}^{n}(i,j) + \beta\left[E(i,j) - R(p^{n}(i,j), q^{n}(i,j))\right] \cdot \frac{\partial R}{\partial q^{n}_{(i,j)}}}$$

Lecture 13

$$p^{n+1}(i,j) = p^{n}(i,j) + \lambda\left[E(i,j) - R\left(p^{n}(i,j), q^{n}(i,j)\right)\right]\frac{\partial R^{n}}{\partial p}\Big|_{p^{n}(i,j)}.$$

Similar equation for q^{n+1}.

$$f = \frac{2p}{1+\sqrt{p^2+q^2+1}}, \quad g = \frac{2q}{1+\sqrt{p^2+q^2+1}}, \quad f(i,j), g(i,j)$$

$$p = \frac{4f}{4-f^2-g^2}, \quad q = \frac{4g}{4-f^2-g^2}.$$

H.W. How do we go from p, q to f, g?
What are the drawbacks of working in p, q domain?
Check - Horn - Ikeuchi Approach.

Assignment 3:- (Check Lecture folder).

Smith's Approach:-
Use of fact that $p_y = q_x$ for any valid surface.

i.e. $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$

or $\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial^2 z}{\partial x \partial y}$.

Use this as the regularization prior information.

& $\min_{\forall p, q} \iint \left[(E-R)^2 + \lambda (p_y - q_x)^2\right] dx\, dy$.

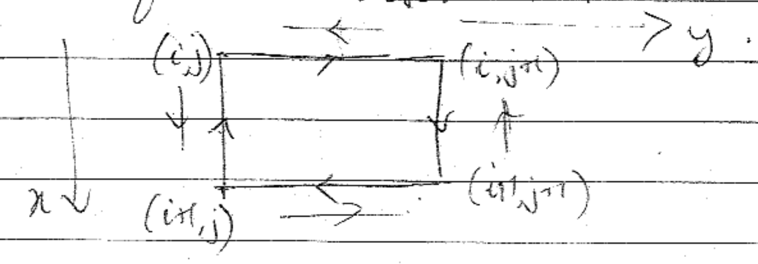
Use of Euler equations (i.e. calculus of variations) to arrive at the solution for the continuous case.

Strat's approach:-

$$p\,dx + q\,dy = \delta z.$$

$$\oint_C \left(p(x,y)\,dx + q(x,y)\,dy \right) = 0$$ for a closed contour, true $\forall$ closed contours.

In discrete case, we ~~minimize~~ may use $\left(\oint_C p\,dx + q\,dy \right)^2$ as minimum $\forall C$ & gain minimize the resulting equation for the discrete case.



i.e. minimize$_{\forall p,q}$ $\sum\sum \left((E\{-R)^2 + \lambda \epsilon_{ij}^2 \right)$

$$\epsilon_{ij} = \int p\,dx + q\,dy = \frac{1}{2}\left[p(i,j) + p(i+1,j) \right] - \frac{1}{2}\left[p(i,j+1) + p(i+1,j+1) \right] + \frac{1}{2}\left[q(i,j) + q(i+1,j) \right] - \frac{1}{2}\left[q(i,j+1) + q(i+1,j+1) \right]$$

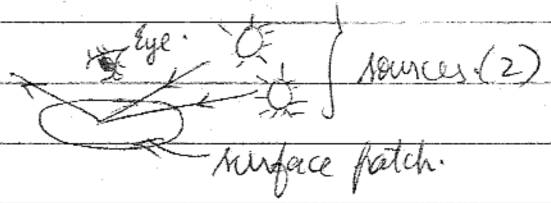
Try solving for iterative solution.

C.V. Lecture 14

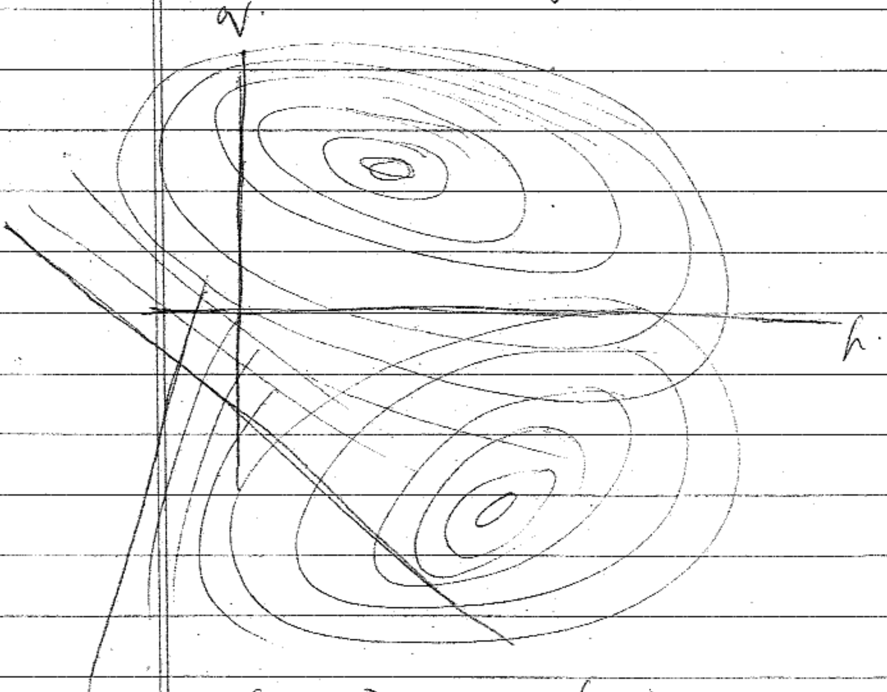
Recap:

Problems with SFS
① Surely contrained
② Single image

Using Photometric stereo, we can get better estimates of p & q since it uses more than 1 image & in turn has more information.



p_{s_1} q_{s_1}
p_{s_2} q_{s_2}

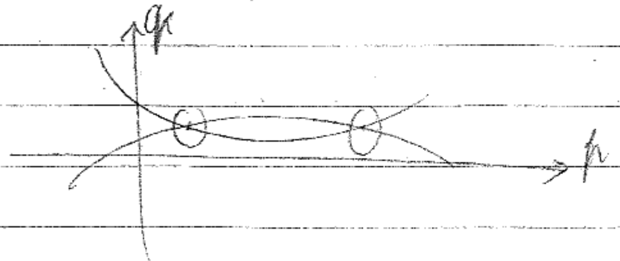


$(p, q) \leftarrow (p_{s_1}, q_{s_1})$

$(p, q) \leftarrow (p_{s_2}, q_{s_2})$

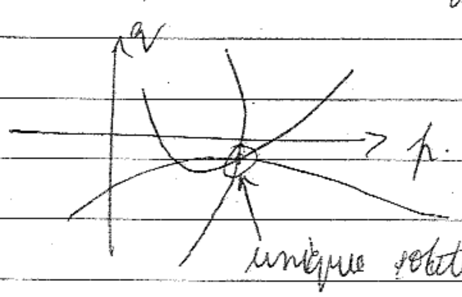
$R_1(p, q) = E_1(x, y)$
$R_2(p, q) = E_2(x, y)$

Either no solution or 1 solution or max. 2 solutions.



single camera (fixed) + multiple source positions — Photometric Stereo
single source + multiple camera pos. } Structure from
+ — " — object pos. } motion.

We can get a unique solution if we use 3 source positions



Albedo $\rightarrow \rho$.

$R(p,q) \propto \cos\theta_i$, $R(p,q) = \rho.\cos\theta_i$

ρ is a function of (x,y) , $0 \leq \rho \leq 1$

Actually for $E(x,y) \propto R(p,q)$
$\therefore R(p,q) = \cos\theta_i$

$\because R(p,q) = E(x,y)$, we have , $E(x,y) = \rho\cos\theta_i$
$= \rho.\hat{n}\cdot\hat{s}$.

Consider 3 measurements

$\therefore E_1(x,y) = \rho\hat{n}\hat{s}_1 = \rho(x,y)\hat{n}\hat{s}_1$
$E_2(x,y) = \rho\hat{n}\hat{s}_2 = \rho(x,y)\hat{n}\hat{s}_2$
$E_3(x,y) = \rho\hat{n}\hat{s}_3 = \rho(x,y)\hat{n}\hat{s}_3$

$$\therefore E_1 = \frac{\rho\,(pp_{s_1} + qq_{s_1} + 1)}{\sqrt{p_{s_1}^2+q_{s_1}^2+1}\,\sqrt{p^2+q^2+1}} \quad \& \text{ similarly } E_2, E_3$$

$$\therefore \underbrace{\begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}}_{\text{known}} = \underbrace{\begin{bmatrix} -p_{s_1}/\sqrt{p_{s_1}^2+q_{s_1}^2+1} & -q_{s_1}/\sqrt{p_{s_1}^2+q_{s_1}^2+1} & 1/\sqrt{p_{s_1}^2+q_{s_1}^2+1} \\ \vdots & \vdots & \vdots \\ & & \end{bmatrix}}_{\text{known}} \underbrace{\begin{bmatrix} \frac{-\rho p}{\sqrt{p^2+q^2+1}} \\ \vdots \\ \frac{\rho}{\sqrt{p^2+q^2+1}} \end{bmatrix}}_{\text{unknown}}$$

$\therefore E = S(\rho \underline{n})$

$\therefore (\rho \underline{n}) = S^{+}E$. $\therefore |\rho \underline{n}| = |S^{+}E|$

But $|\rho \underline{n}| = \rho|\underline{n}| = \rho$. $\because |\underline{n}| = 1$.

$\therefore \rho = |S^{+}E|$

Now,

$$\underline{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad \ldots -p = a/c, \quad -q = b/c$$

$$a = \frac{-p\rho}{\sqrt{p^2+q^2+1}}, \quad b = \frac{-q\rho}{\sqrt{p^2+q^2+1}}, \quad c = \frac{\rho}{\sqrt{p^2+q^2+1}}$$

Till now, we considered noise only in $E(x,y)$ i.e. noise due to sensor. However, there can also be error in reading the source directions i.e. error in S.

Therefore, 3 equations cannot be consistent or linearly independent. This makes us increase the number of readings or measurements, so as to increase the accuracy of determining the unknowns.

Now,

$$E = S\rho\underline{n} \equiv Y = AX.$$

If $\underline{Y} = \underline{A}X$, then LS solution for X can be obtained a

$$\underset{\forall X}{\text{minimize}} \quad \epsilon = \|\underline{Y} - A\underline{X}\|^2.$$

This can be solved by equating the derivative taken w.r.t. the unknown $(\underline{X})$ to zero.

$$\therefore \frac{\partial \epsilon}{\partial \underline{X}} = \underline{0}$$

$$\therefore \frac{\partial \|Y - AX\|^2}{\partial \underline{X}} = \underline{0}$$

$$\therefore -2A^T(Y - AX)^* = \underline{0}.$$

$$\therefore A^T(Y - AX) = \underline{0}.$$

$$\therefore A^TY - A^TAX = \underline{0}.$$

$$\therefore A^TY = A^TAX.$$

$$\therefore \underline{\hat{X} = (A^TA)^{-1}(A^TY)}.$$

Why $\dfrac{\partial A\underline{X}}{\partial \underline{X}} = A^T$? Consider $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Let $\underline{Z} = A\underline{X} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1a_{11} + x_2a_{12} + x_3a_{13} \\ x_1a_{21} + x_2a_{22} + x_3a_{23} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$

$$\therefore \frac{\partial \underline{Z}}{\partial \underline{X}} = \begin{bmatrix} \dfrac{\partial z_1}{\partial \underline{X}} & \dfrac{\partial z_2}{\partial \underline{X}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial z_1}{\partial x_1} & \dfrac{\partial z_2}{\partial x_1} \\ \dfrac{\partial z_1}{\partial x_2} & \dfrac{\partial z_2}{\partial x_2} \\ \dfrac{\partial z_1}{\partial x_3} & \dfrac{\partial z_2}{\partial x_3} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = A^T$$

$$\boxed{\therefore \frac{\partial (A\underline{X})}{\partial \underline{X}} = A^T}$$

CV Lecture 15

Photometric Stereo:-

More than 1 pt. sources ~~can~~ are be used to capture images. The camera & object positions are fixed.

$$\underline{E} = S(\rho \underline{n})$$

Solution using LS, here, pseudo inverse.

But sometimes pseudo inverse may not exist. Then SVD is used.

Any matrix has a S.V.D. $A = U \Sigma V^T$

U & V are orthonormal matrices. & Σ has diagonal entries called singular values.

$$\begin{aligned} AA^T &= (U\Sigma V^T)(U\Sigma V^T)^T \\ &= (U\Sigma V^T)(V\Sigma^T U^T) \\ &= U\Sigma V^T V \Sigma^T U^T \end{aligned}$$

But V is orthonormal $\therefore V^T = V^{-1}$ $\therefore V^T V = I$.

$\therefore AA^T = U \Sigma\Sigma^T U^T$

Again, U is orthonormal $\therefore U^T = U^{-1}$.

$\therefore AA^T = U\Sigma_1 U^{-1}$ $\quad \because \Sigma\Sigma^T = \Sigma_1$.

AA^T is a symmetric square matrix, U diagonaliz AA^T.

$\therefore$ U has eigenvectors of AA^T. & Σ_1 has the corresponding eigenvalues.

Symmetry :- Let $B = AA^T$.

$\therefore B^T = (AA^T)^T = (A^T)^T (A^T) = AA^T = B$.

Now, since U^{-1} exists, the eigenvectors of AA^T have to be <u>linearly independent</u>.

Similarly, V diagonalizes A^TA

$\therefore V$ has eigenvectors of A^TA.

Also, $\Sigma_1 = \sqrt{\Sigma\Sigma^T} = \sqrt{\Sigma^T\Sigma}$.

Number of singular values determine the rank of a matrix

Conditional Number (c.n.) : $\dfrac{\sigma_{max}}{\sigma_{min}}$ ... $\sigma_{min} \neq 0$ ... the least eigenvalue.

A matrix is best for inversion when c.n. = 1.
& bad when c.n. is very high.

$$\boxed{}_{200\times200} \xrightarrow{S.V.D.} \boxed{U}_{200\times200}\ \boxed{\Sigma}_{200\times200}\ \boxed{V^T}_{200\times200} \cong \boxed{U}_{200\times k}\ \boxed{\Sigma}_{k\times k}\ \boxed{V^T}_{k\times200}$$

$$\underbrace{\sigma_1 u_1 v_1 + \sigma_2 u_2 v_2 + \cdots}_{k \text{ terms.}}$$

<u>Assignment :-</u>

Image : $\rightarrow$ SVD $\Rightarrow$ Reduction $\rightarrow$ Reconstruction.

Write the compression ratio & try on color, gray & natural images.

Lecture 16

10/09/2012

Photometric Stereo (contd)

~~th~~ The SVD solution gives LS estimate of unknowns

$$\left(\frac{-\rho p}{\sqrt{p^2+q^2+1}}, \frac{-\rho q}{\sqrt{p^2+q^2+1}}, \frac{\rho}{\sqrt{p^2+q^2+1}} \right).$$

This is done at every image pt. (i,j) & we get the p, q & ρ maps.

The solution is in LS sense. The solution can be improved by regularization.

Now, $E_\ell = \rho n S_\ell = \rho R_\ell(p,q)$.

$\therefore$ obj. f^n for LS is $\epsilon_d = \sum_i \sum_j \sum_{\ell=1}^{k} \| E_\ell(i,j) - \rho(i,j) R_\ell(p(i,j), q(i,j)) \|^2$

By adding the regularization term, we have

$\epsilon_1 = \epsilon_d + \lambda \epsilon_s$. ... Regularized Photometric Stereo.

Where, ϵ_s is the regularization term given by,

$$\epsilon_s = \sum_i \sum_j \Big[\big(p(i,j) - p(i-1,j)\big)^2 + \big(p(i,j) - p(i,j-1)\big)^2$$
$$+ \big(q(i,j) - q(i-1,j)\big)^2 + \big(q(i,j) - q(i,j-1)\big)^2$$
$$+ \big(\rho(i,j) - \rho(i-1,j)\big)^2 + \big(\rho(i,j) - \rho(i,j-1)\big)^2$$

$$= \sum_i \sum_j \left(\epsilon_{sp} + \epsilon_{sq} + \epsilon_{s\rho} \right)$$

Similarly in f, g domain

Depth Estimation: from p, q map.

$$p(x,y) = \frac{\partial z(x,y)}{\partial x}, \quad q(x,y) = \frac{\partial z(x,y)}{\partial y}.$$

$\therefore \; p = z_x \; , \; q = z_y.$

& $p_x = z_{xx} \; , \; q_y = z_{yy}.$

$$\text{Min.} \iint_{x\,y} \underbrace{\left((p - z_x)^2 + (q - z_y)^2\right)}_{F} \, dx \, dy.$$

Using calculus of variations,

$$F_z - \frac{\partial F_{z_x}}{\partial x} - \frac{\partial F_{z_y}}{\partial y} = 0. \quad , \; \cancel{F_z = 0} \quad \text{—} (1).$$

$$F = (p - z_x)^2 + (q - z_y)^2.$$

$\therefore F_z = 0.$

$$F_{z_x} = +2(p - z_x)(-1). \; , \; F_{z_y} = +2(q - z_y)(-1).$$

$\therefore$ ~~$+2p - 2z - 2p + 2z_x - 2q + 2z_y = 0$~~.

$\therefore$ ~~z~~

$$\therefore \frac{\partial z_x}{\partial x} = -2p_x + 2z_{xx} \; \& \; \frac{\partial z_y}{\partial y} = -2q_y + 2z_{yy}.$$

$$\therefore -2(p_x - z_{xx}) - 2(q_y - z_{yy}) = 0.$$

$$\therefore \boxed{p_x + q_y = z_{xx} + z_{yy}.}$$

$\therefore \nabla^2 z = p_x + q_y.$

$\therefore$ ~~Z+t~~ The approximation in discrete form is

$$z(i,j) = \bar{z}(i,j) + [p(i-1,j) - p(i,j)] + [q(i,j-1), q(i,j)]$$

$$\therefore z^{n+1}(i,j) = \bar{z}^n(i,j) + [p(i-1,j) - p(i,j)] + [q(i,j-1) - q(i,j)].$$

Assignment 5:- Find z by giving a noisy z.

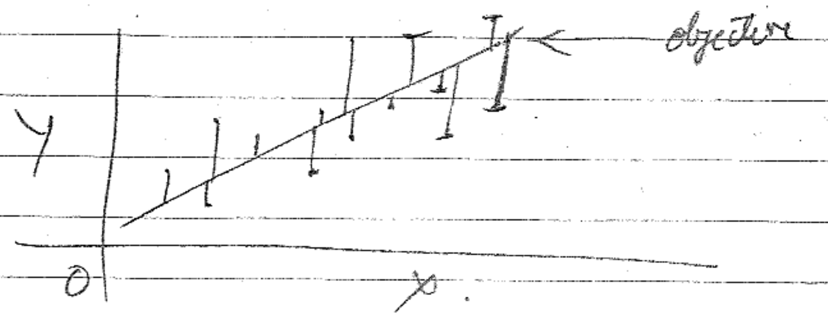
Deterministic Optimization Methods :-

→ Estimating parameters, fitting data to a curve/line, etc. Runner modelling example.

① LS
② WLS
③ CLS
④ TLS
⑤ Lmed (Least Median Squares Solution).

L.S.

Fitting a straight line to a given data (by considering error in Y-direction).



noise

$y_1 = mx_1 + n_1 c$
$y_2 = mx_2 + n_2 c$
$y_3 = mx_3 + n_3 c$
$\vdots$
$y_i = mx_i + n_i c$
$\vdots$
$y_N = mx_N + n_N c$

... line passing through origin.

L.S. sense : Minimize $\sum_{i=1}^{N} (n_i)^2$

$= $ Minimize $\sum_{i=1}^{N} \left(y_i - (mx_i + c)\right)^2$

In general, for any line,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{bmatrix} = \underbrace{\overbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_i & \vdots \\ \vdots & \\ x_N & 1 \end{bmatrix}}^{X}}_{n \times 2} \underbrace{\begin{bmatrix} m \\ c \end{bmatrix}}_{\substack{A \\ 2 \times 1}} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_i \\ \vdots \\ n_N \end{bmatrix}$$

This is of the form $\underline{Y} = A\underline{X}$ + noise.

∴ The problem can be posed as.

Minimize $\| \underline{Y} - A\underline{X} \|^2$

? & we already know that the solution is given by

$\hat{X} = (A^T A)^{-1} (A^T Y)$ ← Pseudo inverse of X.

<u>Estimating the parameters m & c.</u>

Now, we need to minimize (1) Objective function to be minimized

$$\sum_{i=1}^{N} \left(y_i - (mx_i + c)\right)^2.$$

for unknowns m & c.

∴ diff. w.r.t. m & eq. to 0, we have (2) $\frac{\partial}{\partial m}[\,(1)\,] = 0$

$$\frac{\partial}{\partial m}[\;] = \sum_{i=1}^{N} 2\left(y_i - (mx_i + c)\right)(-x_i) = 0.$$

$$\therefore \sum_{i=1}^{N} -x_i y_i + m\sum_{i=1}^{N} x_i^2 + c\sum_{i=1}^{N} x_i = 0.$$

or

$$\boxed{m\sum_{i=1}^{N} x_i^2 + c\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} x_i y_i = 0}$$

$\overline{x_i^2}$, $\overline{x_i}$, $\overline{xy}$, $\overline{x_i}$ (1)

Also, diff. w.r.t. c & eq to 0 we have,

$$\frac{\partial}{\partial c}\{\;\} = \sum_{i=1}^{N} 2\left(y_i - (mx_i + c)\right)(-1) = 0$$

$$= \sum_{i=1}^{N} -y_i + m\sum_{i=1}^{N} x_i + c\sum_{i=1}^{N} 1 = 0.$$

$$\therefore \boxed{m\sum_{i=1}^{N} x_i + c\cdot N - \sum_{i=1}^{N} y_i = 0}$$

$m\overline{x_i^2} + c\overline{x_i} = \overline{xy}$

$m\overline{x_i} + c = \overline{y}$

$\boxed{\overline{x_i}, \overline{y_i}}$

$m(\overline{x_i^2} - (\overline{x_i})^2) = \overline{xy} - \overline{x_i}\,\overline{y}$ (2)

$m = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x_i^2} - (\overline{x_i})^2}$

Eq Multiplying eq ① with N & eq ② with $\sum_{i=1}^{N} x_i$, we have.

$$\boxed{Nm \sum_{i=1}^{N} x_i^2 + Nc \sum_{i=1}^{N} x_i - N \sum_{i=1}^{N} x_i y_i = 0.} \quad -④$$

$$\boxed{m \left(\sum_{i=1}^{N} x_i\right)^2 + Nc \sum_{i=1}^{N} x_i - \left(\sum_{i=1}^{N} y_i\right)\left(\sum_{i=1}^{N} x_i\right) = 0} \quad -⑤$$

Eq ④ − eq ⑤ gives,

$$-m\left[-N \sum_{i=1}^{N} x_i^2 + \left(\sum_{i=1}^{N} x_i\right)^2\right] + \left[\left(\sum_{i=1}^{N} y_i\right)\left(\sum_{i=1}^{N} x_i\right) - N \sum_{i=1}^{N} x_i y_i\right] = 0.$$

$$\therefore m = \frac{\left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right) - N \sum_{i=1}^{N} x_i y_i}{\left(\sum_{i=1}^{N} x_i\right)^2 - N \sum_{i=1}^{N} x_i^2}$$

Multiplying & dividing RHS by N^2 we have

$$m = \frac{\left(\frac{1}{N} \sum_{i=1}^{N} x_i\right)\left(\frac{1}{N} \sum_{i=1}^{N} y_i\right) - \left(\frac{1}{N} \sum_{i=1}^{N} x_i y_i\right)}{\left(\frac{1}{N} \sum_{i=1}^{N} x_i\right)^2 - \frac{1}{N} \sum_{i=1}^{N} x_i^2}$$

$$\therefore \; m = \frac{\bar{x}\cdot\bar{y} - \frac{1}{N}\sum_{i=1}^{N} x_i y_i}{\bar{x}^2 - \frac{1}{N}\sum_{i=1}^{N} x_i^2}$$

$$= \frac{\frac{1}{N}\sum_{i=1}^{N} x_i y_i - \bar{x}\cdot\bar{y}}{\frac{1}{N}\sum_{i=1}^{N} x_i^2 - \bar{x}^2} = \frac{E(X\cdot Y) - E(X)\cdot E(Y)}{E(X^2) - (E(X))^2}$$

$$\therefore \; \boxed{\hat{m} = \frac{cov(x,y)}{var(x)}.}$$

$Cov(XY) = E(XY) - E(X)\cdot E(Y)$
$Var(X) = (E(X^2)) - (E(X))^2$

$$= \frac{s(x,y)}{(s_x)}$$

From eq ② we have,

$$c = \frac{1}{N}\left(\sum_{i=1}^{N} y_i - \hat{m}\sum_{i=1}^{N} x_i\right).$$

$$= \frac{1}{N}\sum_{i=1}^{N} y_i - \hat{m}\,\frac{1}{N}\sum_{i=1}^{N} x_i.$$

$$\therefore \; \boxed{\hat{C} = \bar{y} - \hat{m}\,\bar{x}}.$$

$$cor(x,y) = r_{(x,y)} = \frac{cov(x,y)}{\sqrt{var(x)}\cdot\sqrt{var(y)}} = \frac{cov(x,y)}{var(x)}\cdot\frac{\sqrt{var(x)}}{\sqrt{var(y)}}.$$

$$\therefore cov(x,y) = \hat{m}\cdot\frac{\sqrt{var(x)}}{\sqrt{var(y)}} \quad \text{i.e.} \quad r_{(x,y)} = \hat{m}\cdot\frac{s_x}{s_y}.$$

$$\therefore \; \boxed{\hat{m} = r_{(x,y)}\cdot\frac{s_y}{s_x}}$$

Now, $y_i = \hat{m} x_i + \hat{c}$

$$\therefore \; y_i = r_{(x,y)} \cdot \frac{s_y}{s_x} \cdot x_i + \bar{y} - r_{(x,y)} \frac{s_y}{s_x} \cdot \bar{x}$$

$$\therefore (y_i - \bar{y}) = r_{(x,y)} \frac{s_y}{s_x} \left(x_i - \bar{x} \right).$$

... of the form $\underline{Y} = \underline{A} X$

//

Lecture 17

11/09/2012

Ex. Fitting a straight line for a given data.

x	0	3	6
y	1	4	5

$$Y = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 \\ 3 & 1 \\ 6 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} m \\ c \end{bmatrix}$$

$$\underline{Y} = A\underline{X}$$

We need to find $\underline{X}$'s estimate
Using L.S. we have $\hat{\underline{X}} = (A^TA)^{-1}A^T\underline{Y}$.

$$\therefore \cancel{A^TA = \begin{bmatrix} 0 & 1 \\ 3 & 1 \\ 6 & \end{bmatrix}\begin{bmatrix} 0 & 3 & 6 \\ 1 & 1 & 1 \end{bmatrix}}$$

$$\cancel{= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 10 & 19 \\ & & \end{bmatrix}}$$

$$A^TA = \begin{bmatrix} 0 & 3 & 6 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 3 & 1 \\ 6 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 45 & 9 \\ 9 & 3 \end{bmatrix}$$

$$|A^TA| = \begin{vmatrix} 45 & 9 \\ 9 & 3 \end{vmatrix} = 135 - 81 \neq 0$$

$\therefore A^TA$ is invertible

For a matrix M, the inverse (if exists) is given by

$$M^{-1} = \frac{Adj(M)}{|M|}.$$

$$\therefore (A^T A)^{-1} = \frac{1}{54}\begin{bmatrix} 3 & -9 \\ -9 & 45 \end{bmatrix} = \begin{bmatrix} 1/18 & -1/6 \\ -1/6 & 5/6 \end{bmatrix}$$

$$\therefore \hat{X} = \frac{1}{54}\begin{bmatrix} 3 & -9 \\ -9 & 45 \end{bmatrix}\begin{bmatrix} 0 & 3 & 6 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}$$

$$= \frac{1}{54}\begin{bmatrix} 3 & -9 \\ -9 & 45 \end{bmatrix}\begin{bmatrix} 42 \\ 10 \end{bmatrix}$$

$$\begin{array}{r} 150 \\ 126 \\ \hline 24 \end{array}$$

$$= \frac{1}{18}\begin{bmatrix} 1 & -3 \\ -3 & 15 \end{bmatrix}\begin{bmatrix} 42 \\ 10 \end{bmatrix}$$

$$= \frac{1}{18}\begin{bmatrix} 12 \\ 24 \end{bmatrix} = \frac{12}{18}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$= \frac{2}{3}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 4/3 \end{bmatrix}.$$

$\therefore \hat{m} = 2/3$, $\hat{c} = 4/3$.

$\therefore \hat{y}_1 = (2/3)(0) + 4/3 = 4/3$

$\hat{y}_2 = (2/3)(3) + 4/3 = 10/3$

$\hat{y}_3 = (2/3)(6) + 4/3 = 16/3$

$\therefore e_1 = y_1 - \hat{y}_1 = 1 - 4/3 = -1/3$, $e_2 = 4 - 10/3 = 2/3$.

$\&\ e_3 = 5 - 16/3 = -1/3$ $\therefore \sum_i e_i^2 = \frac{1}{9} + \frac{4}{9} + \frac{1}{9} = \frac{6}{9} = \frac{2}{3}$

$$m = \frac{\mathrm{Cov}(XY)}{\mathrm{var}(X)} = \frac{E(XY) - E(X)E(Y)}{E(X^2) - [E(X)]^2}$$

$$= \frac{42}{3} - \frac{3 \cdot 16}{3} = \frac{13}{15} \; 4$$

Let $\hat{m} = 1/3$, $\hat{c} = 4/3$.

$\therefore \hat{y}_1 = (1/3)(0) + 4/3 = 4/3$

$\hat{y}_2 = (1/3)(3) + 4/3 = 7/3$

$\hat{y}_3 = (1/3)(6) + 4/3 = 10/3$.

$$\therefore e_1^2 + e_2^2 + e_3^2 = \left(1 - \frac{4}{3}\right)^2 + \left(4 - \frac{7}{3}\right)^2 + \left(5 - \frac{10}{3}\right)^2$$

$$= \frac{1}{9} + \frac{25}{9} + \frac{25}{9} = \frac{51}{9} = \frac{17}{3}$$

Thus, this error is more.

Q. Find the best quadratic LS fit to a given data.

$\therefore y = c_0 + c_1 x + c_2 x^2$.

where, c_0, c_1, c_2 are to be determined.

For LS, we need to minimize Z, such that,

$$Z = \sum_i \left(y_i - (c_0 + c_1 x_i + c_2 x_i^2)\right)^2$$

Now to find optimal c_0, c_1, c_2, $\frac{\partial Z}{\partial c_0} = 0$, $\frac{\partial Z}{\partial c_1} = 0$, $\frac{\partial Z}{\partial c_2}$

$\frac{\partial Z}{\partial c_0} \Rightarrow \therefore 2\sum_{i=1}^{N} \left(y_i - (c_0 + c_1 x_i + c_2 x_i^2)\right) \cdot (-1) = 0$.

$$\therefore \boxed{\sum_{i=1}^{N} y_i - N c_0 - c_1 \sum_{i=1}^{N} x_i - c_2 \sum_{i=1}^{N} x_i^2 = 0} \quad \text{— (1)}$$

$\frac{\partial Z}{\partial c_1} \Rightarrow 2\sum_{i=1}^{N} \left(y_i - (c_0 + c_1 x_i + c_2 x_i^2)\right)(-x_i) = 0$.

$$\therefore \boxed{\sum_{i=1}^{N} x_i y_i - c_0 \sum_{i=1}^{N} x_i - c_1 \sum_{i=1}^{N} x_i^2 - c_2 \sum_{i=1}^{N} x_i^3 = 0}$$

$$\frac{\partial Z}{\partial c_2} = 0 \Rightarrow 2\sum_{i=1}^{N}\left(y_i - (c_0 + c_1 x_i + c_2 x_i^2)\right)(-x_i^2) = 0.$$

$$\therefore \boxed{\sum_{i=1}^{N} x_i^2 y_i - c_0 \sum_{i=1}^{N} x_i^2 - c_1 \sum_{i=1}^{N} x_i^3 - c_2 \sum_{i=1}^{N} x_i^4 = 0.} \quad \text{—(3)}$$

We can then arrive at values of c_0, c_1 & c_2 by solving eq. (1), (2) & (3) simultaneously.

However, if we write

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{bmatrix}, \quad X = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$$

then $\underline{Y} = A\underline{X}$

? & we minimize $\|Y - AX\|^2$
whose LS solution is
$$\boxed{\hat{X} = (A^T A)^{-1} A^T Y}$$

While for L.S. data is assumed to be corrupted at every pt. with noise of same variance, for every pt. the noise has same variance.

If the variances for every pt. are known to be distinct then we may go for weighted least squares (WLS) method

i.e. $\text{Min } Z = \sum w_i (y_i - (mx_i + c))^2$, -- known x_i, y_i

Objective function changes to

i.e. $\boxed{\text{Min } \|W^{1/2}(Y - AX)\|^2}$

P.T.O.

or $\quad \text{Min.} \; \| W Y - W A X \|^2$

$\therefore$ The solution can be given by diff. w.r.t $\underline{X}$ & equating to 0.

$\therefore \quad -2 (WA)^T (W\underline{Y} - WA\underline{X}) = 0 \quad \therefore A^T W^T W \underline{Y} = A^T W^T W A \underline{X}$

$\therefore \; \cancel{2W^2 [A^T (Y - AX)] = 0} \quad \therefore \hat{\underline{X}} = (A^T W^T W A)^{-1} (A^T W^T W \underline{Y})$

Constrained Least Squares :- (CLS)

If $\underline{A X = 0}$ with constraints $\|X\|^2 = 1$,
Homogeneous equation

then we have the CLS solution.

This is used because, when A is full ranked $\underline{X = 0}$ is a trivial solution, which we wish to avoid. Hence $\|\underline{X}\|^2 = 1$ is used.

2

The problem is posed as,

$$\text{Min} \quad Z = \|A\underline{X}\|^2$$
$$\text{s.t.} \quad \|X\|^2 = 1.$$

Using Lagrange multiplier method, we have,

$$F = \|A\underline{X}\|^2 + \lambda(\|\underline{X}\|^2 - 1)$$

or,

$$F = \|A\underline{X}\|^2 + \lambda(1 - \|\underline{X}\|^2)$$

& $\quad \dfrac{\partial F}{\partial X} = 0, \quad \dfrac{\partial F}{\partial \lambda} = 0.$

$$\frac{\partial F}{\partial X} = 0 \Rightarrow 2A^T(AX) \cancel{+ \lambda 2 X^T X} - \lambda(2X) = 0.$$

$$\therefore \quad A^T A X - \lambda X = 0.$$

$$\boxed{A^T A X = \lambda X} \quad \leftarrow (1)$$

classmate
Date
Page 13

This is of the form $B\underline{X} = \lambda \underline{X}$

Thus, B matrix has the eigen vector $\underline{X}$ & λ as its corresponding eigenvalue.

Now, $\|AX\|^2 = \cancel{A^TA}\,(AX)^T(AX) = X^TA^TAX$

But $A^TAX = \lambda X$ from (1).

$\therefore \|AX\|^2 = X^T\lambda X = \lambda X^TX = \lambda\|X\|^2$

But $\|X\|^2 = 1$. ... the constraint.

$\therefore \|AX\|^2 = \lambda$

For $\|AX\|^2$ to be minimum, we need λ to be minimum. In other words $\|AX\|^2$ is equal to the smallest eigen value when $\|AX\|^2$ is minimum.

$\therefore$ The solution X ~~that corresponds to the smallest eigen value λ~~, is that eigen vector of (A^TA), which is associated or corresponds to the smallest eigenvalue of (A^TA).

Q. Why is $\dfrac{\partial\|Y-AX\|^2}{\partial X} = 2A^T(Y-AX)$?

$y_1 = mx_1 + c$

$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_{3\times1}$ $A = \begin{bmatrix} m \\ c \end{bmatrix}_{2\times1}$ $X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{bmatrix}_{3\times2}$

2. or simply, why is $\dfrac{\partial AX}{\partial X} = A^T$?

solution Consider, $\underline{a} = (a_1\ a_2\ a_3\ a_4)^T$, $\underline{x} = (x_1\ x_2\ x_3\ x_4)^T$

$= \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$

P.T.O.

& $z = a'x$
$= a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4$

Now, $\dfrac{\partial(a'x)}{\partial x} = \begin{bmatrix} \dfrac{\partial z}{\partial x_1} \\ \dfrac{\partial z}{\partial x_2} \\ \dfrac{\partial z}{\partial x_3} \\ \dfrac{\partial z}{\partial x_4} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \underline{a}$

Similarly, now consider,

$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$, $\underline{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$

such that $Z = A\underline{x}$

$= \begin{bmatrix} a_{11}x_1 & a_{12}x_2 & a_{13}x_3 \\ a_{21}x_1 & a_{22}x_2 & a_{23}x_3 \end{bmatrix} = \begin{bmatrix} \underline{a}_1^T \\ \underline{a}_2^T \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \underline{a}_1^T \underline{x} \\ \underline{a}_2^T \underline{x} \end{bmatrix}$

then, $\dfrac{\partial z}{\partial \underline{x}} = \begin{bmatrix} \dfrac{\partial a_1^T x}{\partial x} & \dfrac{\partial a_2^T x}{\partial x} \end{bmatrix} = \dfrac{\partial (A\underline{x})}{\partial \underline{x}}$

$= \begin{bmatrix} \dfrac{\partial \underline{a}_1^T \underline{x}}{\partial x_1} & \dfrac{\partial a_2^T \underline{x}}{\partial x_1} \\ \dfrac{\partial \underline{a}_1^T x}{\partial x_2} & \dfrac{\partial a_2^T x}{\partial x_2} \\ \dfrac{\partial \underline{a}_1^T x}{\partial x_3} & \dfrac{\partial \underline{a}_2^T x}{\partial x_3} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = A^T$

$\therefore \boxed{\dfrac{\partial (A\underline{x})}{\partial \underline{x}} = A^T}$ Hence, $\boxed{\dfrac{\partial \|\underline{Y} - A\underline{X}\|^2}{\partial \underline{X}} = 2A^T(Y - AX)}$

$ax + by + c = 0$

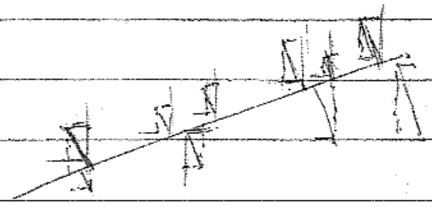
Direction vector = $\begin{bmatrix} -b \\ a \end{bmatrix}$

Normal vector $\begin{bmatrix} a \\ b \end{bmatrix}$

Lecture 18

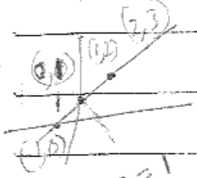
We now use the solution CLS to arrive at the solution of total least squares. (TLS)

In TLS, error is considered to be in both x & y directions



Here, $(y+e_y) = m(x+e_x) + c$

We need to minimize $\sum_{\forall i} (e_{x_i}^2 + e_{y_i}^2)$



$m=1 \quad c=1$

$y = x+1$

$x - y + 1 = 0$

$(-b, a)$

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Normal $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

This is nothing but minimizing the perpendicular distance squares.

Eq. of a straight line : $y = m_1 x + c_1$

or $ax + by + c = 0$. —— ①

s.t. $m_1 = \frac{-a}{b}$, $c_1 = \frac{-c}{b}$

Consider $ax + by = 0$ —— ②

Lines represented by ① & ② are parallel

∴ To find the direction vector of line in ① we can consider the line represented by ②.

(-b, a) satisfies ②

If $P(-b, a)$ is a point, then

$a(-b) + b(a) = -ab + ab = 0$

∴ $ax + by = 0$ is satisfied.

Now consider $a'x + b'y + c' = 0$. —— ③

∴ Direction vector can be given by $P'(-b', a')$

For lines in ① & ③ to be ⊥, we need $P \cdot P' = 0$

For line $ax+by+c=0$, the direction vector is given by $P(-b,a)$ & a normal vector to this line is given by $Q(a,b)$
... this is because $P \cdot Q = (-b,a)\cdot(a,b)^T = -ba+ab = \underline{\underline{0}}$ satisfied.

Suppose, instead of $ax+by+c=0$, we have the line represented by $rax + rby + rc = 0$
$\therefore$ The normal vector is given by $Q(ra, rb)$.
But Q being a direction vector, it is a unit vector

$$\therefore \sqrt{(ra)^2+(rb)^2} = 1.$$
$$\therefore (ra)^2+(rb)^2 = 1.$$
$$\therefore r^2(a^2+b^2) = 1.$$
$$\therefore r^2 = \frac{1}{a^2+b^2}.$$

$$\therefore \boxed{r = \frac{1}{\sqrt{a^2+b^2}}}.$$

Then, the line can be represented as

$$\frac{ax}{\sqrt{a^2+b^2}} + \frac{by}{\sqrt{a^2+b^2}} + \frac{c}{\sqrt{a^2+b^2}} = 0.$$

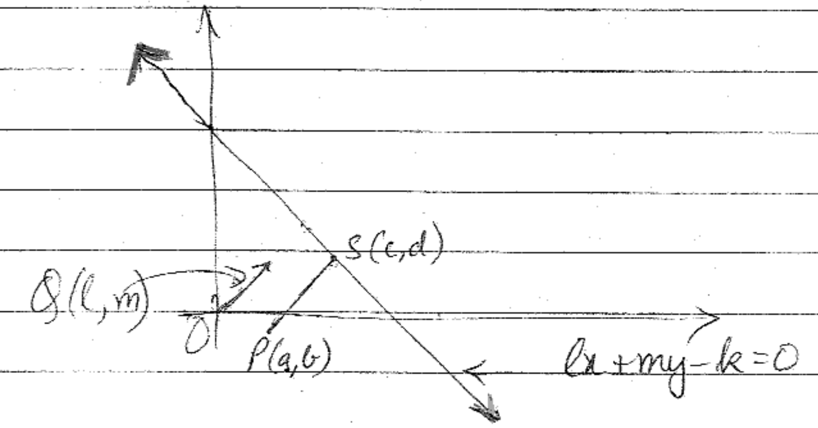
Or, let, $l = \frac{a}{\sqrt{a^2+b^2}}$, $m = \frac{b}{\sqrt{a^2+b^2}}$, $k = \frac{-c}{\sqrt{a^2+b^2}}$.

$\therefore$ The normal equation of a line can be written as,

$$\boxed{lx + my - k = 0}$$

$\therefore$ The unit normal is now $Q(l, m)$ such that $l^2 + m^2 = 1$

Now consider,



Now, OQ & PS are $\parallel$.

Let $PS = r \cdot Q$ where Q or OQ is a unit vector.

$\therefore |r| = (P-S) \cdot Q$ vector $\parallel$ to OQ unit vector. , $r \in \mathbb{R}$.

magnitude   dot product

$\therefore PQ - SQ = |r|$ ... length of vector PS.

Considering the coordinates of P, Q & S we have,

~~al + bm~~ $(a\ b)\begin{pmatrix} l \\ m \end{pmatrix} - (c\ d)\begin{pmatrix} l \\ m \end{pmatrix} = |r|$

$\therefore al + bm - cl - dm = |r|$

But since S lies on the line $lx + my - k = 0$, we have,

$lc + md - k = 0$.

$\therefore lc + md = k$.

$\therefore al + bm - k = |r|$

$\therefore |r| = (a\ b)\begin{pmatrix} l \\ m \end{pmatrix} - k$. or $|r| = \begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} l \\ m \end{pmatrix} - k$

Consider, $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$

$\therefore d_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \cdot \begin{pmatrix} l \\ m \end{pmatrix} - k.$

where d_1 is now the $\perp$ dist. of pt. (x_1, y_1) from line $lx + my - k = 0$ ... similar to $|r_1|$ for pt. $P(a, b)$

$\therefore$ For i^{th} pt., the $\perp$ distance is given by.

$$\boxed{d_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} l \\ m \end{pmatrix} - k}$$

or $\boxed{d_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} - k}$ ... $l = n_x, m = n_y$ & $\sqrt{n_x^2 + n_y^2} = 1$

For TLS, we need to minimize the error which is the sum of all $\perp$ distances.

i.e. $\min Z = \sum_{i=1}^{N} (d_i)^2$

i.e. $\min Z = \sum_{i=1}^{N} \left[\begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} - k \right]^2$

with the constraint $\|\bar{n}\|^2 = 1$ i.e. $n_x^2 + n_y^2 = 1$

Diff. w.r.t. unknown k, we get.

$\frac{\partial Z}{\partial k} = 0 \Rightarrow 2 \sum_{i=1}^{N} \left[\begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} - k \right] (-1) = 0.$

$$\therefore \sum_{i=1}^{N} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} - k \sum_{i=1}^{N} 1 = 0 .$$

$$\therefore \quad kN = \sum_{i=1}^{N} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix}$$

$$\therefore \boxed{k = \frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \begin{pmatrix} n_x \\ n_y \end{pmatrix}} \quad \text{i.e.} \quad \boxed{k = \frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \vec{n}}$$

$$\therefore \quad d_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} - \frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix}$$

$$= \begin{pmatrix} x_i - \frac{1}{N} \sum_{i=1}^{N} x_i \\ y_i - \frac{1}{N} \sum_{i=1}^{N} y_i \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix}$$

$$= \begin{pmatrix} \tilde{x}_i \\ \tilde{y}_i \end{pmatrix} \begin{pmatrix} n_x \\ n_y \end{pmatrix}$$

$$\therefore \boxed{d_i = \begin{pmatrix} \tilde{x}_i \\ \tilde{y}_i \end{pmatrix} \cdot \vec{n}}$$

$$\therefore \quad Z = \sum_{i=1}^{N} \left[\begin{pmatrix} \tilde{x}_i \\ \tilde{y}_i \end{pmatrix} \cdot \vec{n} \right]^2 = \sum_{i=1}^{N} \left(\tilde{x}_i \cdot n_x + \tilde{y}_i \, n_y \right)^2$$

$$\therefore \quad Z = \sum_{i=1}^{N} d_i^2 = \left\| \begin{bmatrix} \tilde{x}_1 & \tilde{y}_1 \\ \tilde{x}_2 & \tilde{y}_2 \\ \vdots & \vdots \\ \tilde{x}_N & \tilde{y}_N \end{bmatrix} \begin{bmatrix} n_x \\ n_y \end{bmatrix} \right\|^2 = \| A \underline{X} \|^2$$

Ankush.

classmate
Date
Page 83

Lecture 19

17/09/2012.

Q. For the following 2D pts. fit a straight line which leads to a solution in the TLS sense. Use SVD.

x	2	0	6	4
y	0	2	4	6

Soln. $N = 4$. $\bar{x} = \frac{1}{4}(2+0+6+4) = \frac{12}{4} = 3$. (1) Find $\bar{x}$

$\bar{y} = \frac{1}{4}(0+2+4+6) = \frac{12}{4} = 3$. (2) Find $\bar{y}$

$$x_i - \bar{x} \qquad \tilde{x} = \begin{array}{c} x_i - \bar{x} \\ \begin{bmatrix} 2-3 \\ 0-3 \\ 6-3 \\ 4-3 \end{bmatrix} \end{array} = \begin{bmatrix} -1 \\ -3 \\ 3 \\ 1 \end{bmatrix}$$ (4) Find $x_i - \bar{x} = \tilde{x}$

$$\tilde{y} = \begin{array}{c} y_i - \bar{y} \\ \begin{bmatrix} 0-3 \\ 2-3 \\ 4-3 \\ 6-3 \end{bmatrix} \end{array} = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$ (5) Find $(y - \bar{y}) = \tilde{y}$

$A = [\tilde{x} \ \tilde{y}]$. (6) Take $A = [\tilde{x} \ \tilde{y}]$

$$\therefore A = \begin{bmatrix} -1 & -3 \\ -3 & -1 \\ 3 & 1 \\ 1 & 3 \end{bmatrix} \quad \& \quad X = \begin{bmatrix} n_x \\ n_y \end{bmatrix}$$

(7) For TLS, $\min \|AX\|^2$ s.t. $\|X\|^2 = 1$.
whose solution is the eigenvector of (A^TA) corresponding to least eigenvalue.

P.T.O

$\therefore A^T A = \begin{bmatrix} -1 & -3 & 3 & 1 \\ -3 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} -1 & -3 \\ -3 & -1 \\ 3 & 1 \\ 1 & 3 \end{bmatrix}$

(7) Find $A^T A = A$

$= \begin{bmatrix} 20 & 12 \\ 12 & 20 \end{bmatrix} = A_1$

Now, $|A_1 - \lambda I| = 0$

(8)(a) Find out eigen values / eigen vectors.

$\therefore \begin{vmatrix} 20-\lambda & 12 \\ 12 & 20-\lambda \end{vmatrix} = 0.$

$\therefore (20-\lambda)^2 - (12)^2 = 0.$

$\therefore (20-\lambda-12)(20-\lambda+12) = 0.$

$\therefore (8-\lambda)(32-\lambda) = 0.$

$\therefore \lambda = 32$ or $\lambda = 8.$

When $\lambda = 32$,
we have

$A_1 v_1 = 32 v_1$

or $\begin{bmatrix} 20 & 12 \\ 12 & 20 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = 32 \begin{bmatrix} V_x \\ V_y \end{bmatrix}$

8(b) Find eigen vectors corresponding to eigen values.

$\therefore 20 v_x + 12 v_y = 32 V_x.$

$\& \; 12 V_x + 20 V_y = 32 V_y$ why

(i) using property of Eigen values +

? (ii) $\sqrt{V_x^2 + V_y^2} = 1$

$\therefore 12 v_x = 12 v_y.$

$\therefore v_x = v_y.$

Also, $\sqrt{v_x^2+v_y^2} = 1.$

$\therefore \sqrt{2v_x^2} = 1.$

$\therefore v_x = \frac{1}{\sqrt{2}} = v_y.$

$\therefore v_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$

When $\lambda = 8$ we have,

$A_1 u_1 = 8 u_1.$

or $\begin{bmatrix}20 & 12\\12 & 20\end{bmatrix}\begin{bmatrix}u_x\\u_y\end{bmatrix} = 8\begin{bmatrix}u_x\\u_y\end{bmatrix}$

$\therefore 20u_x + 12u_y = 8u_x.$

$\& \; 12u_x + 20u_y = 8u_y.$

$\therefore 12u_y = -12u_x.$

$\therefore u_y = -u_x.$

Also, $\sqrt{u_x^2+u_y^2} = 1.$

$\therefore \sqrt{u_x^2+(-u_x)^2} = 1.$

$\therefore \sqrt{2u_x^2} = 1.$

$\therefore u_x = \frac{1}{\sqrt{2}}.$

$\therefore u_y = \frac{-1}{\sqrt{2}}$

$\therefore u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{bmatrix}$

$\therefore$ eigenvector corresponding to the smallest eigen value of A_1 (i.e. of A^TA), which in this is 8), is $u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$.

$\therefore \underline{X} = \begin{matrix}l\\m\end{matrix}\begin{bmatrix}n_x\\n_y\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$ $\quad lx + my - k = 0$

Equation of the best fit line in TLS sense is given by, (Q) Find $k = \frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix}x_i\\y_i\end{pmatrix}\begin{pmatrix}n_x\\n_y\end{pmatrix}$

$lx + my - k = 0$, where $l = n_x$, $m = n_y$.

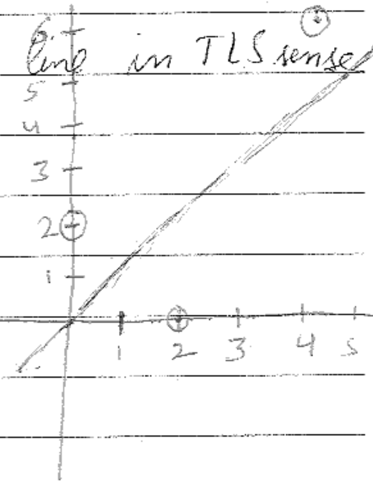
$$k = \frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix}x_i\\y_i\end{pmatrix}\cdot\begin{pmatrix}n_x\\n_y\end{pmatrix}$$

2 $$\therefore k = \frac{1}{4}\left[\begin{array}{l} 2(1/\sqrt{2}) + 0(-1/\sqrt{2}) \\ + 0(1/\sqrt{2}) + 2(-1/\sqrt{2}) \\ + 6(1/\sqrt{2}) + 4(-1/\sqrt{2}) \\ + 4(1/\sqrt{2}) + 6(-1/\sqrt{2}) \end{array}\right] = \frac{1}{4}(0) = 0.$$

$\therefore$ The equation of the best fit line in TLS sense is,

$$lx + my = 0.$$

or $$\frac{x}{\sqrt{2}} - \frac{y}{\sqrt{2}} = 0.$$



or $x - y = 0$ or $\boxed{x = y}$

Special cases:- 4 points such that the adjacent pts are equidistant & so are the diagonals.

$P(A|B) = \frac{P(A\cap B)}{P(B)}$

EM algorithm

Used in case of some missing data

Foundation,

Bayes Rule:- $P(A|B)\cdot P(B) = P(B|A)\cdot P(A)$

Here, Posterior probability, Likelihood function, Prior probability

$P(\text{Parameter}|\text{Data})\cdot P(\text{Data}) = P(\text{Data}|\text{Parameter})\cdot P(\text{Parameter})$

$\therefore P(\theta|Z)\cdot P(Z) = P(Z|\theta)\cdot P(\theta)$

& $\therefore P(\theta|Z) = \dfrac{P(Z|\theta).P(\theta)}{P(Z)}$

But $P(Z)$ is constant

likelihood

$\therefore \underset{\forall \theta}{\text{argmax}}\ P(\theta|Z) = \underset{\forall \theta}{\text{argmax}}\ P(Z|\theta) . P(\theta)$

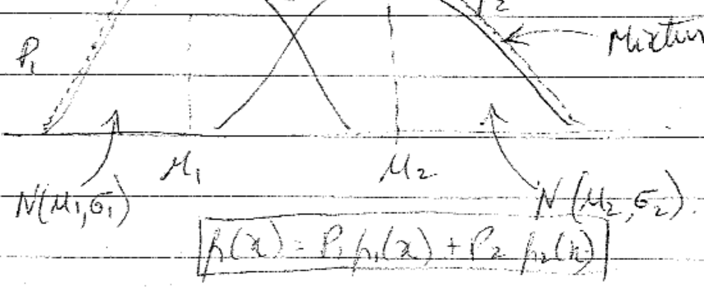
Maximum Aposteriori Estimation (or MAP estimation).
& Posterior probability of θ is $P(\theta|Z)$.
Prior probability of θ is $P(\theta)$.

In case of uniform distribution, we have all unknowns as equiprobable. Hence, if θ is assumed to follow a uniform distribution, then $P(\theta)$ is constant.

$\therefore \underset{\forall \theta}{\text{argmax}}\ P(\theta|Z) = \underset{\forall (\theta)}{\text{argmax}}\ P(Z|\theta)$ ... Maximum likelihood estimate

Check Matlab function 'randn'.

EM is MLE for incomplete data (used to get Gaussian Mixture Model's parameters).
Eg:- GMM.



$\boxed{p(x) = P_1 p_1(x) + P_2 p_2(x)}$

Q. Given a set of 1000 pts, fit a Gaussian curve to the data using MLE.

P.T.O.

Problem Formulation:-

$$\underset{\forall \underline{\theta}}{\text{argmax}}\ P(\underline{\theta} \mid z_1 \ldots z_n) \text{ i.e. } \underset{\forall \mu,\sigma}{\text{argmax}}\ P(\mu\sigma \mid z_1 \ldots z_n)$$

in this case
$\underline{\theta} = (\mu, \sigma)$

Assumption is that the pts. are i.i.d. i.e. all have same μ & σ^2

(For simplicity $\sigma^2 = 1$, $\mu = 0$ ... which is not given, but will be estimated).

Log Likelihood:-

$$\hat{\mu}, \hat{\sigma} = \underset{\forall \mu\sigma}{\text{argmax}}\ P(\mu, \sigma \mid z_1, \ldots, z_n)$$

But due to MLE,

$$\underset{\forall \mu,\sigma}{\text{argmax}}\ P(\mu,\sigma \mid z_1 \ldots z_n) = \underset{\forall \mu,\sigma}{\text{argmax}}\ P(z_1 \ldots z_n \mid \mu, \sigma)$$

$$\therefore \hat{\mu}, \hat{\sigma} = \underset{\forall \mu\sigma}{\text{argmax}}\ P(z_1, \ldots, z_n \mid \mu, \sigma).$$ ① maximize $P(\text{data})$

For a Gaussian curve, $P(z_i \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{z_i - \mu}{2\sigma}\right)^2}$

But all z_i are i.i.d. ② Since z_i are independent $P(z_1, z_2, \ldots z_n) = P(z_1) \cdot P_{z_2} \cdot P_{z_3}$

$\therefore P(z_1 \ldots z_n) = P(z_1) \cdot P(z_2) \ldots P(z_n)$

* Since ln of is monotonically increasing, ~~function~~ the max. of ln is at the same pt as that of the function itself. Since the derivatives are involved in min/max, it is convenient to use ln instead of directly using the function.

$$\therefore P(z_1 \dots z_n | \mu, \sigma) = P(z_1|\mu,\sigma) \cdot P(z_2|\mu\sigma) \cdot \dots \cdot P(z_n|\mu\sigma)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}\left(\frac{z_1-\mu}{\sigma}\right)^2} \dots \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{z_n-\mu}{\sigma}\right)^2}$$

~~(3) Total probability P(data | parameters) =~~

$$= \prod_{i=1}^{n} P(z_i|\mu,\sigma) = \left(\frac{1}{\sqrt{2\pi}\cdot\sigma}\right)^n \left[e^{-\frac{1}{2}\left[\left(\frac{z_1-\mu}{\sigma}\right)^2 + \dots + \left(\frac{z_n-\mu}{\sigma}\right)^2\right]} \right]$$

Now, $\hat{\mu}, \hat{\sigma} = \underset{\forall \mu, \sigma}{\text{argmax}}\ P(z_1 \dots z_n|\mu,\sigma) = \underset{\forall \mu,\sigma}{\text{argmax}}\ \ln\left[P(z_1 \dots z_n|\mu,\sigma)\right]$.

* $\boxed{\text{Note that } \cancel{\text{argmax}}\ \max P(z_1 \dots z_n|\mu\sigma) \neq \max \ln\left[P(z_1 \dots z_n|\mu,\sigma)\right] \text{ However both have max. value at same } \mu, \sigma.}$

(4) Take log likelihood as it is easy to deal with ~~products~~ sum rather than products

$$\therefore \ln\left[P(z_1 \dots z_n|\mu,\sigma)\right] = \ln\left[\prod_{i=1}^{n} P(z_i|\mu\sigma)\right]$$

i.e. $\hat{\mu}, \hat{\sigma} = \underset{\forall \mu\sigma}{\text{argmax}}\ \ln\left[\prod_{i=1}^{n} P(z_i|\mu\sigma)\right] = \underset{\forall \mu,\sigma}{\text{argmax}} \sum_{i=1}^{n} \ln\left[P(z_i|\mu,\sigma)\right]$ ✓

Where $\ln\left[P(z_i|\mu,\sigma)\right] = \ln\left[\frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}\left(\frac{z_i-\mu}{\sigma}\right)^2}\right]$

(5) (a) Calculate log equation.

$$= \ln\left(\frac{1}{\sqrt{2\pi}}\right) + \ln\left(\frac{1}{\sigma}\right) + \left(-\frac{1}{2}\left(\frac{z_i-\mu}{\sigma}\right)^2\right) \cdot \ln e$$

But $\ln e = 1$

$$\therefore \ln\left[P(z_i|\mu,\sigma)\right] = -\ln(\sqrt{2\pi}) - \ln(\sigma) - \frac{1}{2}\left(\frac{z_i-\mu}{\sigma}\right)^2$$

$$\therefore \sum_{i=1}^{n} \ln\left[P(z_i|\mu,\sigma)\right] = \sum_{i=1}^{n}\left[-\ln(\sqrt{2\pi}) - \ln(\sigma) - \frac{1}{2}\left(\frac{z_i-\mu}{\sigma}\right)^2\right]$$

$$\therefore \sum_{i=1}^{n} \ln\left[P(z_i|\mu,\sigma)\right] = -\ln(\sqrt{2\pi})\cdot\sum_{i=1}^{n} i - \ln(\sigma)\sum_{i=1}^{n} i - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{z_i-\mu}{\sigma}\right)^2$$

$$= -n\ln\left(\sqrt{2\pi}\right) - n\ln(\sigma) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{z_i-\mu}{\sigma}\right)^2$$

Here, the unknowns are μ, σ

Let $T = \underset{\forall \mu,\sigma}{\text{argmax}}\left[-n\ln(\sqrt{2\pi}) - n\ln(\sigma) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{z_i-\mu}{\sigma}\right)^2\right]$

(6) Diff equation wrt μ and σ and equate to zero

$\therefore$ (a) $\frac{\partial T}{\partial \mu} = 0$ & (b) $\frac{\partial T}{\partial \sigma} = 0$.

(a) mean variance

Using $\frac{\partial T}{\partial \mu} = 0$ we get,

$$-0 - 0 - \frac{1}{2}\sum_{i=1}^{n} 2\left(\frac{z_i-\mu}{\sigma}\right)\left(\frac{-1}{\sigma}\right) = 0$$

$$\therefore \sum_{i=1}^{n}\left(\frac{z_i-\mu}{\sigma^2}\right) = 0.$$

$$\text{or.} \quad \sum_{i=1}^{n}(z_i-\mu) = 0.$$

$$\therefore \sum_{i=1}^{n} z_i - n\mu = 0.$$

$$\therefore \boxed{\hat{\mu} = \frac{1}{N}\sum_{i=1}^{n} z_i}$$

Using $\frac{\partial T}{\partial \sigma} = 0$, we get, Page 91

$$\cancel{0 - n\left(\frac{1}{\sigma}\right) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{z_i-\mu}{}\right)^2}$$

$$0 - n\left(\frac{1}{\sigma}\right) - \frac{1}{2}\sum_{i=1}^{n}(z_i-\mu)^2\cdot(-2)\cdot\sigma^{-3} = 0.$$

$$\therefore \sum_{i=1}^{n}\frac{(z_i-\mu)^2}{\sigma^3} - \frac{n}{\sigma} = 0.$$

$$\therefore \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-\mu)^2 = n.$$

$$\therefore \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(z_i-\mu)^2}$$

where $\hat{\mu}$, $\hat{\sigma}^2$ are the M.L estimates of Mean & variance, respectively, of a Gaussian curve.

Lecture 20

Now, Z is a r.v.

then

$$\sigma^2 = \int_{-\infty}^{\infty} (z-\mu)^2 f_z(z)\, dz = E\left[(Z-\mu)^2\right]$$

$$\& \ \mu = \int_{-\infty}^{\infty} z\, f_z(z)\, dz = E[Z]$$

If we apply the ML estimation to solve the straight line fitting problem, then we need to find m, c such that

$$\underset{\forall m,c}{\operatorname{argmax}}\ P(y_1, \ldots y_n \mid m, c)$$

Given, $y_i = mx_i + c + \underbrace{n_i}_{\text{noise}}$

Assumption:- $n_i\ \forall i = 1{:}N$ are noise values from Gaussian p.d.f. which is i.i.d. Let mean 0 & variance then

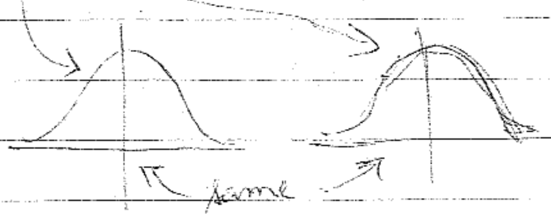
For N_1,

$$P_{N_1}(n) = \frac{1}{\sqrt{2\pi}\,(1)} \cdot e^{-\frac{1}{2}\left(\frac{n-0}{1}\right)^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{n^2}{2}}$$

... $-10 \ldots + 10$. Realization

N_1 : $10, 7, 6, 4, 3, \ldots$

N_2 : $-10, 6, 2, 3, -5, \ldots$



However, the order is differen

$\therefore \; P(y_i \mid m, c) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(y_i - (mx_i + c))^2}{2}}$ ... Modelling the noise.

Since $(Y_i \mid m, c) \; \forall i$ are independent r.v. we have

$$P(y_1 \dots y_N \mid m, c) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(y_i - (mx_i + c))^2}{2}}$$

Now, $\underset{m,c}{\operatorname{argmax}} \; P(y_1 \dots y_N \mid m, c)$

$$= \underset{m,c}{\operatorname{argmax}} \left(\frac{1}{\sqrt{2\pi}}\right)^N \cdot e^{-\frac{1}{2}\sum_{i=1}^{N} (y_i - (mx_i + c))^2}$$

$$= \underset{m,c}{\operatorname{argmax}} \; \ln \left[\left(\frac{1}{\sqrt{2\pi}}\right)^N \cdot e^{-\frac{1}{2}\sum_{i=1}^{N} (y_i - (mx_i + c))^2} \right]$$

$$= \underset{m,c}{\operatorname{argmax}} \left[\underbrace{N \cdot \left(-\ln(\sqrt{2\pi})\right)}_{\text{constant}} - \frac{1}{2} \sum_{i=1}^{N} (y_i - (mx_i + c))^2 \cdot \ln e \right]$$

constant

$$= \underset{m,c}{\operatorname{argmax}} \left[\cancel{\tfrac{1}{2}} - \sum_i (\cancel{x}y_i - (mx_i + c))^2 \right]$$

$$= \underset{m,c}{\operatorname{argmin}} \left[\sum_i (y_i - (mx_i + c))^2 \right] = \text{L.S. estimate of } m, c$$

Thus, M.L.E. with $\mu = 0$ & var $= \cancel{1}^2$ for an i.i.d. Gaussian is the same as L.S. estimate.

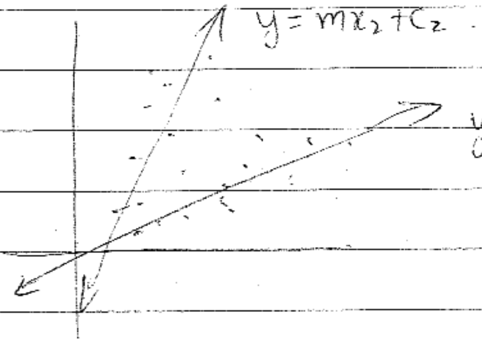
When each r.v. noise has different var., then it becomes WLS

MAP estimate :-

EM is the ML estimate of parameters when there is incomplete data or hidden data.

Take the case of fitting a line. Let the data be such that it fits well if we use more than 1 straight lines.

Eg:



We need to estimate (m_1, c_1) & (m_2, c_2)

Q. Which pt. belongs to which line? But if pts. are known, then we can decide the probability of assigning a pt. to a line.

EM :- Involves 2 steps
1. Expectation E
2. Maximization M.

Let $\underline{\theta}$ be the param. to be estimated.
Let $\underline{y}$ be the observed data (vector)
Let z be the realization of hidden data, a r.v. (Z)

E Step :-
The hidden data is estimated using the current estimate of

M Step :-
Obtain the likelihood function assuming observations & the hidden data. Find ML estimate of θ using the observ & hidden data found in the E step.

Repeat until θ converges.

Lecture 21 $\boxed{???}$

24/09/2012

For M.L. estimate of parameters θ given the observations Y, we have to maximize $L = \ln[P(Y|\theta)] = L(\theta)$.
... this is because, we have already proved that $P(\theta|Y) \equiv P(Y|\theta)$ in this case.

If θ_n corresponds to the estimate of θ after n^{th} iteration, then one can write,

$$L(\theta_n) = \ln[P(Y|\theta_n)]$$

Since the objective is to find that θ which maximizes $L(\theta)$, we look for updating θ such that $\underset{\text{new}}{L(\theta)} > \underset{\text{old}}{L(\theta_n)}$

By this, we are actually looking for maximizing $L(\theta) - L(\theta_n)$

Let,
Z = the hidden/missing data & let z be its realization

$\therefore$ The MLE can be formulated as follows,

$$\underset{\theta}{\operatorname{argmax}} \overset{MLE}{P(\underset{\text{obs.}}{Y}, \underset{\text{hidden}}{Z} | \underset{\text{param.}}{\theta})} = \sum_z P(Y, z|\theta)$$

$$= \sum_z P(Y|z\theta) \cdot P(z|\theta) \quad \text{... chain rule}$$

$\therefore L(\theta) = \ln[P(YZ|\theta)]$ & $L(\theta_n) = \ln[P(Y|\theta_n)]$

$$\therefore L(\theta) = \ln \sum_z P(Y|z\theta) \cdot P(z|\theta)$$

$$L(\theta) - L(\theta_n) = \ln \sum_z P(Y|z\theta) \cdot P(z|\theta) - \ln P(Y|\theta_n)$$

... which is what we actually need to maximize

$$\therefore L(\theta) - L(\theta_n) = \ln\left(\sum_{z} P(Y|z\theta)\cdot P(z|\theta)\,\frac{P(z|Y\theta_n)}{P(z|Y\theta_n)}\right) - \sum_{z} P(z|Y\theta_n)\,\ln[P(Y|\theta_n)]$$

$\because$ as $\sum_{z} P(z|Y\theta_n) = 1$.

$$\therefore L(\theta) - L(\theta_n) = \ln\left(\sum_{z} \frac{P(z|Y\theta_n)\cdot P(Y|z\theta)\cdot P(z|\theta)}{P(z|Y\theta_n)}\right) - \ln[P(Y|\theta_n)]\cdot\sum_{z} P(z|Y\theta_n)$$

Jensen's inequality,

$$\ln \sum_{i} \lambda_i y_i \geq \sum_{i} \lambda_i \ln y_i$$

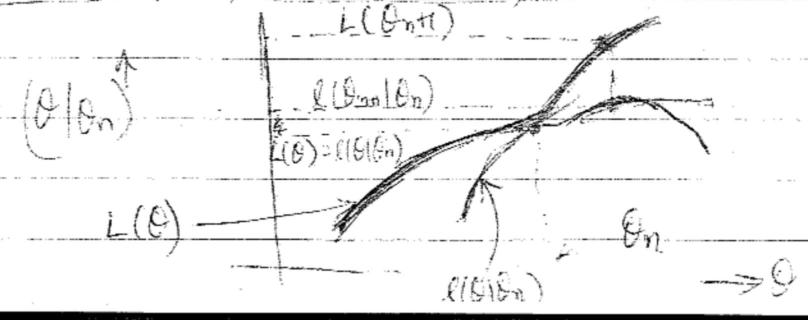
$$\sum_{i} \lambda_i = 1 \quad \& \quad \lambda_i \geq 0.$$

$$\therefore L(\theta) - L(\theta_n) \geq \sum_{z} P(z|Y\theta_n)\left[\ln\left(\frac{P(Y|z\theta)\cdot P(z|\theta)}{P(z|Y\theta_n)}\right) - \ln(P(Y|\theta_n))\right]$$

$$\therefore L(\theta) - L(\theta_n) \geq \underbrace{\sum_{i} P(z_i|Y\theta_n)\cdot \ln\left[\frac{P(Y|z\theta)\cdot P(z|\theta)}{P(Y|\theta_n)\cdot P(z|Y\theta_n)}\right]}_{\Delta(\theta|\theta_n)}$$

$$\therefore L(\theta) \geq \underbrace{L(\theta_n) + \Delta(\theta|\theta_n)}_{l(\theta|\theta_n)}$$

$$\therefore L(\theta) \geq l(\theta|\theta_n)$$



Find $l(\theta|\theta_n)$.

Now,

$$L(\theta_n) + \Delta(\theta|\theta_n) = l(\theta|\theta_n)$$

$$l(\theta_n|\theta_n) = L(\theta_n) + \Delta(\theta_n|\theta_n)$$

$$= L(\theta_n) + \ln(1)$$

$$= L(\theta_n) + 0$$

$$= L(\theta_n)$$

Now, therefor θ which ~~maximizes~~ increases $l(\theta|\theta_n)$ also increases $L(\theta)$ & hence max. can be converted to $\underset{\theta}{\operatorname{argmax}} \{l(\theta|\theta_n)\} = \theta_{n+1}$

$$l(\theta|\theta_n) = L(\theta) + \Delta(\theta|\theta_n)$$

$$\theta_{n+1} = \underset{\theta}{\operatorname{argmax}} \{l(\theta|\theta_n)\}$$

$$= \underset{\theta}{\operatorname{argmax}} \left\{ L(\theta) + \sum_i P(z_i|Y\theta_n) \cdot \ln\left[\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(Y|\theta_n)\, P(z|Y\theta_n)}\right] \right\}$$

$$= \underset{\theta}{\operatorname{argmax}} \left\{ \sum_i P(z_i|Y\theta_n) \ln\left[P(Y|z\theta) \cdot P(z|\theta)\right] \right\}$$

$$= \underset{\theta}{\operatorname{argmax}} \left\{ \sum_i P(z_i|Y\theta_n) \cdot \ln\left[P(Yz|\theta)\right] \right\}$$

$$= \underset{\theta}{\operatorname{argmax}}\ E_{z|Y\theta_n} \cdot \ln\left[P(Yz|\theta)\right]$$

Attempt GMM using EM.

classmate

Date

Page (98)

~~17~~18/10/2012.

Lecture 2~~1~~ Recap.

Steropsis / Stereo Vision.   Book by Trucco

EM Algorithm :-

For a random variable Y, the ML estimation technique tries to identify the parameters(θ) of the distribution / curve(s), which generated Y. This is achieved by maximizing the likelihood $P(Y|\theta)$.

i.e. to find $P(\theta|Y)$, we use Bayes' rule, which says $P(\theta|Y) = \dfrac{P(Y|\theta) \cdot P(\theta)}{P(Y)}$

But here $P(Y)$ is constant & also $P(\theta)$ can be assumed to be constant for a uniform distribution (which we can easily assume as the probability of having the parameter θ_i out of i available parameters is equally likely & is equal to $1/i$). Hence ~~$P(\theta)$~~ the max. value of $P(\theta|Y)$ simply depends on max. value of $P(Y|\theta)$. Therefore, we try to maximize $P(Y|\theta)$

However, sometimes we have some missing data or hidden parameters. For example, say the data (observed) belongs to two different distributions & we are unaware as to which pt. belongs to which distribution. Let this hidden parameter or latent data be Z whose realization is z. Then, we intend to find θ ~~$P(Y,z|\theta)$~~ which maximizes $P(Y,z|\theta)$

Again using Bayes rule

$P(Y,z,\theta) = P(Y|z,\theta) \cdot P(z,\theta)$

~~& $P(Y|Z\theta) = P(YZ|\theta) \cdot P(Z|\theta)$~~

& $P(YZ|\theta) = P(Y|Z\theta) \cdot P(Z|\theta)$.

This comes from the fact that,

$$P(Y,Z,\theta) = P(Y|Z\theta) \cdot P(Z\theta).$$
$$\& \; P(Z\theta) = P(Z|\theta) \cdot P(\theta)$$
$$\therefore P(Y,Z,\theta) = P(Y|Z\theta) \cdot P(Z|\theta) \cdot P(\theta)$$

$$\& \; P(Y,Z,\theta) = P(YZ|\theta) \cdot P(\theta).$$
$$= P(Y|Z\theta) \cdot P(Z|\theta) \cdot P(\theta)$$ ... Chain rule.

$$\therefore \underset{\theta}{\operatorname{argmax}} P(YZ|\theta) = \underset{\theta}{\operatorname{argmax}} \sum_{z} P(Y_z|\theta) \quad \ldots \text{ i.i.d.}$$

$$= \underset{\theta}{\operatorname{argmax}} \sum_{z} P(Y|z\theta) \cdot P(z|\theta)$$

~~Now, for ML, maximizing $P(Y|\theta)$ is same as~~

Now, for ML, the parameter(θ) which maximizes $P(Y|\theta)$ also maximizes $\ln[P(Y|\theta)]$, as ln is a monotonically increasing function.

Let $L(\theta) = \ln[P(Y|\theta)]$, Then after n^{th} iteration we have

$$L(\theta_n) = \ln[P(Y|\theta_n)]$$

$\therefore$ In case of hidden parameters, the parameter θ which maximizes $P(YZ|\theta)$ also maximizes $\ln[P(YZ|\theta)]$.

Let $L(\theta) = \ln[P(YZ|\theta)]$

$$\therefore L(\theta) = \ln\left[\sum_{z} P(Y_z|\theta) \; P(z|\theta)\right]$$

If in the n^{th} iteration we know the hidden param or the hidden param does not exist, then in we can estimate the parameters in the next iteration using the ML estimate from the n^{th} iteration. Since we want ~~the~~ to find ~~$L(\theta)$~~ θ which maximize $L(\theta)$, we look to update θ such that we have ~~max.~~ $L(\theta) > L(\theta_n)$.

In other words, we intend to maximize $L(\theta) - L(\theta$

$$\therefore L(\theta) - L(\theta_n) = \ln\left[\sum_z P(Y|z\theta) \cdot P(z|\theta)\right] - \ln\left[P(Y|\theta_n)\right] \quad —(1)$$

Now, $\sum_z P(z|Y\theta_n) = 1$.

$$\therefore \ln\left[P(Y|\theta_n)\right] = \sum_z P(z|Y\theta_n) \cdot \ln\left[P(Y|\theta_n)\right]$$

$\therefore$ Eq (1) becomes,

$$L(\theta) - L(\theta_n) = \ln\left[\sum_z P(Y|z\theta) \, P(z|\theta) \cdot \frac{P(z|Y\theta_n)}{P(z|Y\theta_n)}\right] - \sum_z P(z|Y\theta_n) \ln\left[P(Y$$

$$= \underbrace{\ln\left[\sum_z P(z|Y\theta_n) \cdot \left(\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n)}\right)\right]}_{A} - \sum_z P(z|Y\theta_n) \ln\left[P(Y|\theta_n\right.$$

Consider the term $A = \ln\left[\sum_z P(z|Y\theta_n) \cdot \frac{P(Y|z\theta) \, P(z|\theta)}{P(z|Y\theta_n)}\right]$

This is of the form $\ln\left(\sum_i \lambda_i X_i\right)$ or $f(\lambda_i X_i$

For a convex function, $f(\lambda x_i + (1-\lambda) y_i) \leq \lambda f(x_i) + (1-\lambda$

$0 \leq \lambda \leq 1$

In general, ~~$\sum f(\lambda_i x_i) \leq$~~

$$\boxed{f\left(\sum_i \lambda_i x_i\right) \leq \sum_i \lambda_i f(x_i)} \quad \ldots \sum \lambda_i = 1, \; 0 \leq \lambda_i \leq 1$$

Now $-\ln$ is a convex function

$$\therefore -\ln\left(\sum_i \lambda_i x_i\right) \leq -\sum \lambda_i \ln(x_i)$$

$$\therefore \boxed{\ln\left(\sum_i \lambda_i x_i\right) \geq \sum \lambda_i \ln(x_i)}$$

... Jensen's inequality

$\therefore$ For the term A, where $P(z|Y\theta)$ can be considered as λ_z, we have,

$$\ln\left[\sum_z P(z|Y\theta_n) \cdot \frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n)}\right] \geq \sum_z P(z|Y\theta_n) \cdot \ln\left[\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n)}\right]$$

$$\therefore L(\theta) - L(\theta_n) \geq \sum_z P(z|Y\theta_n) \cdot \ln\left[\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n)}\right] - \sum_z P(z|Y\theta_n) \cdot \ln\left[P(Y|\theta_n)\right]$$

$$\therefore L(\theta) - L(\theta_n) \geq \sum_z P(z|Y\theta_n) \cdot \ln\left[\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n) \cdot P(Y|\theta_n)}\right] = \Delta(\theta|\theta_n)$$

$$\therefore L(\theta) - L(\theta_n) \geq \Delta(\theta|\theta_n)$$

$$\text{or } L(\theta) \geq L(\theta_n) + \Delta(\theta|\theta_n) = l(\theta|\theta_n)$$

$$\therefore L(\theta) \geq l(\theta|\theta_n)$$

P.T.O.

When $\cancel{k}\theta = \theta_n$, then,

$l(\theta|\theta_n)$ becomes $l(\theta_n|\theta_n) = L(\theta_n) + \Delta(\theta_n|\theta_n)$

$$= P(Y|\theta_n) + \sum_{z} P(z|Y\theta_n) \cdot \ln\left[\frac{P(Y|z\theta_n) \cdot P(z|\theta_n)}{P(z|Y\theta_n) \cdot P(Y|\theta_n)}\right]$$

Now, $P(Y, z, \theta_n) = P(Y|z\theta_n) \cdot P(z\theta_n)$
$= P(Y|z\theta_n) \cdot P(z|\theta_n) \cdot P(\cancel{z}\theta_n)$ — (1)

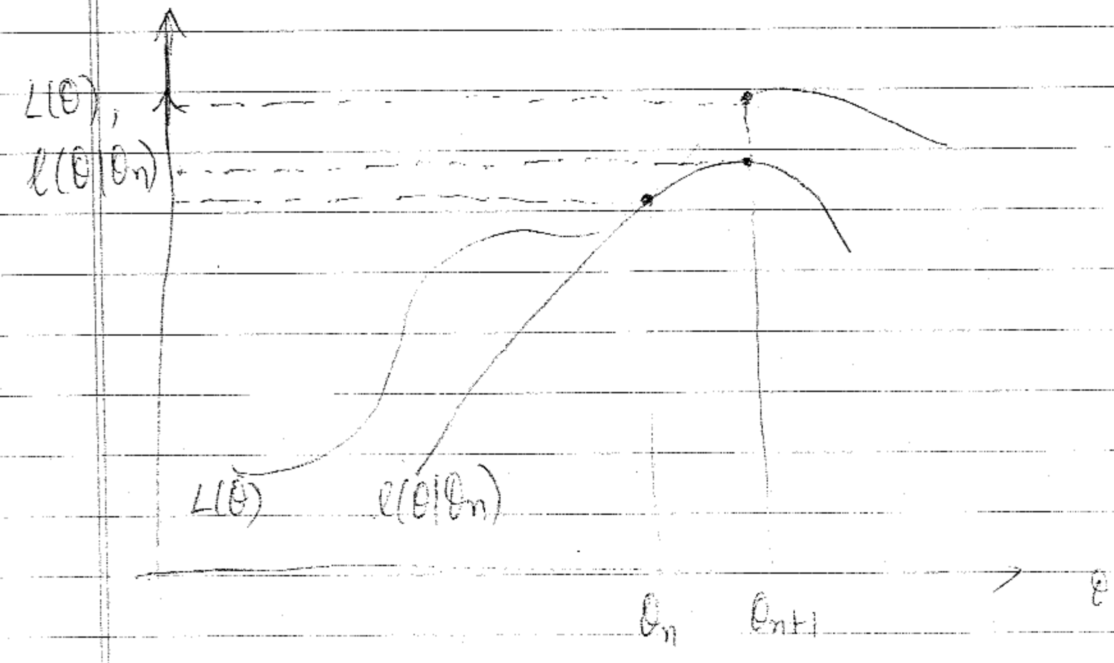
Also, $P(Y, z, \theta_n) = P(z|Y\theta_n) \cdot P(Y\theta_n)$
$= P(z|Y\theta_n) \cdot P(Y|\theta_n) \cdot P(\theta_n)$ — (2)

$\therefore P(Y|z\theta_n) \cdot P(z|\theta_n) = P(z|Y\theta_n) \cdot P(Y|\theta_n)$.

$\therefore l(\theta_n|\theta_n) = P(Y|\theta_n) + \sum_{z} P(z|Y\theta_n) \cdot \ln(1)$.

$= P(Y|\theta_n)$

$= \underline{\underline{L(\theta_n)}}$.



$l(\theta|\theta_n)$ is bounded above by $L(\theta)$ & equals $L(\theta_n)$ at the current estimate $\theta = \theta_n$. Also, this means that any θ which increases $l(\theta|\theta_n)$ also increase $L(\theta)$. Thus maximizing $l(\theta|\theta_n)$ gives the greatest possible increase in $L(\theta)$. Hence, in order to maximize $L(\theta)$, we now maximize $l(\theta|\theta_n)$ for $n+1^{th}$ iteration.

$$\therefore \theta_{n+1} = \underset{\theta}{\operatorname{argmax}} \{l(\theta|\theta_n)\} = \underset{\theta}{\operatorname{argmax}} \{L(\theta_n) + \Delta(\theta|\theta_n)\}$$

$$= \underset{\theta}{\operatorname{argmax}} \left[\sum_z P(z|Y\theta_n) \cdot \ln \left[\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n) \cdot P(Y|\theta_n)} \right] + P(Y|\theta_n) \right]$$

but the terms with θ_n are constants & hence can be neglected in ln.

$$\therefore \theta_{n+1} = \underset{\theta}{\operatorname{argmax}} \left[\sum_z P(z|Y\theta_n) \ln[P(Y|z\theta) \cdot P(z|\theta)] \right]$$

Now, $P(Y|z\theta) = \dfrac{P(Y,z,\theta)}{P(z,\theta)}$ & $P(z|\theta) = \dfrac{P(z\theta)}{P(\theta)}$

$$\therefore \theta_{n+1} = \underset{\theta}{\operatorname{argmax}} \left[\sum_z P(z|Y\theta_n) \cdot \ln \left[\frac{P(Yz\theta)}{\cancel{P(z\theta)}} \cdot \frac{\cancel{P(z\theta)}}{P(\theta)} \right] \right]$$

Also, $P(Yz\theta) = \cancel{P(Y|z\theta)}\ P(Yz|\theta) \cdot P(\theta)$

$\therefore P(Yz|\theta) = \dfrac{P(Yz\theta)}{P(\theta)}$

$$\therefore \theta_{n+1} = \underset{\theta}{\operatorname{argmax}} \left[\sum_z P(z|Y\theta_n) \cdot \ln(P(Yz|\theta)) \right]$$

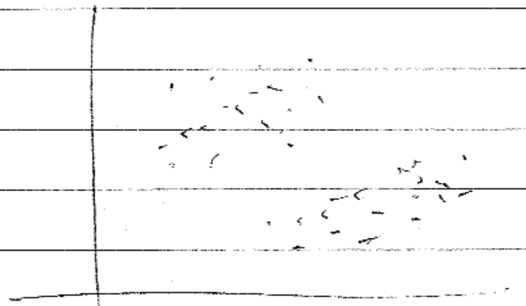
$$= \underset{\theta}{\operatorname{argmax}} \left[E_{z|Y\theta_n} \{\ln[P(Yz|\theta)]\} \right]$$

P.T.O.

This leads to the following.
1^{st} estimate the expectation of log-likelihood of $(Y,Z|\theta)$ i.e. estimate the expectation of $\ln[P(YZ|\theta)]$.
And use this conditional expectation (i.e. $E_{Z|Y\theta}$) of the hidden param/data to find the parameter θ that maximizes this estimate.
i.e.
1. Estimate $E_{Z|Y\theta_n}\{\ln(P(Y|Z\theta)\}$ ... by assuming some θ initially.
2. Maximize :- this expression w.r.t. θ.

Example :-

Estimate two best fit lines for the following shown data, which is corrupted by noise comming from a mixture of Gaussians with mean 0 & variance σ^2.



i.e. Apply EM to GMM.

Assume any two lines initially. say
$y = m_1 x + c_1$ — Line 1
& $y = m_2 x + c_2$ — Line 2.

Here we do not know as to which pt. belongs to which line. Hence, the latent data 'Z' is the assignment of pts to a line.

Given the equation of 2 lines, we can decide which pts are closer to which line & hence make this assignment.
The error between a pt. & the two lines can be used to decide the closeness of lines to the pt.

We assume this error to have come from a mixture of Gaussians with mean 0 & var. σ^2.

$\therefore$ ML estimate that pt. has been corrupted due to noise comming from distribution 1 can be used as the weight to estimate the best fit line.

Now for line 1, we have,

$$r_{1(i)} = (y_i - (m_1 x_i + c_1))^2$$

& for line 2, $\quad r_{2(i)} = (y_i - (m_2 x_i + c_2))^2$

Since the error is gaussian with mean 0, we have,

$$w_{1(i)} = \frac{\frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{1}{2}\frac{(r_{1(i)})^2}{\sigma^2}}}{\frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{1}{2}\left(\frac{r_{1(i)}^2}{\sigma^2}\right)} + \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{r_{2(i)}^2}{\sigma^2}\right)}}$$

$$= \frac{e^{-\frac{1}{2}\left(\frac{r_{1(i)}}{\sigma}\right)^2}}{e^{-\frac{1}{2}\left(\frac{r_{1(i)}}{\sigma}\right)^2} + e^{-\frac{1}{2}\left(\frac{r_{2(i)}}{\sigma}\right)^2}}$$

$$\& \; w_{2(i)} = \frac{e^{-\frac{1}{2}\left(\frac{r_{2(i)}}{\sigma}\right)^2}}{e^{-\frac{1}{2}\left(\frac{r_{1(i)}}{\sigma}\right)^2} + e^{-\frac{1}{2}\left(\frac{r_{2(i)}}{\sigma}\right)^2}}$$

such that $w_{1(i)} + w_{2(i)} = 1$

This completes the E step.

P.T.O.

In the maximization (M) step, we try to find the parameters of the best fitting lines knowing the probable assignment of the data points to the lines. The best fitting lines are the ones having minimum error.

For line 1,

$$\text{Min. } Z = \sum_{i=1}^{n} w_i \left(y_i - (m_1 x_i + c_1) \right)^2$$

$$\frac{\partial z}{\partial m_1} = 0 \Rightarrow \sum_{i=1}^{n} 2w_i \left(y_i - (m_1 x_i + c_1) \right)(-x_i) = 0.$$

$$\therefore \sum_{i=1}^{n} w_i x_i \left(m_1 x_i + c_1 - y_i \right) = 0.$$

$$\therefore m_1 \sum_{i=1}^{n} x_i^2 w_i + c_1 \sum_{i=1}^{n} x_i w_i = \sum_{i=1}^{n} w_i x_i y_i \quad \text{--- (i)}$$

$$\frac{\partial Z}{\partial c_1} = 0 \Rightarrow \sum_{i=1}^{n} 2w_i \left(y_i - (m_1 x_i + c_1) \right)(-1) = 0.$$

$$\therefore \sum_{i=1}^{n} w_i \left(m_1 x_i + c_1 - y_i \right) = 0.$$

$$\therefore m_1 \sum_{i=1}^{n} x_i w_i + c \sum_{i=1}^{n} w_i = \sum_{i=1}^{n} w_i y_i \quad \text{--- (2)}$$

From (1) & (2) we have,

$$\underbrace{\begin{bmatrix} \sum x_i^2 w_i & \sum x_i w_i \\ \sum x_i w_i & \sum w_i \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m_1 \\ c_1 \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} \sum x_i y_i w_i \\ \sum y_i w_i \end{bmatrix}}_{B}$$

$$X = B$$

$\therefore \begin{bmatrix} \cancel{m_1} \\ \cancel{c_1} \end{bmatrix} = \boxed{\hat{X} = A^{-1}B}$ or $\boxed{\hat{X} = (A^TA)^{-1}A^TB}$

Similar estimate of parameters of line 2.

Alternately, since the problem boils down to WLS, we have

$$Z = \text{Min.} (Y - AX)^T W (Y - AX) \text{, where,}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \quad X = \begin{bmatrix} m_1 \\ c_1 \end{bmatrix}, \quad W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & w_n \end{bmatrix}$$

Then the solution is given as

$$\frac{\partial Z}{\partial X} = 0 \Rightarrow -2A^TW(Y - AX) = \underline{0}$$

$$\text{or} \quad A^TWY - A^TWAX = \underline{0}$$

$$\therefore A^TWY = A^TWAX$$

$$\therefore A^TWAX = A^TWY$$

$$\therefore \boxed{\hat{X} = (A^TWA)^{-1}(A^TWY)}$$

Similar expression for line 2.

Alternately repeat these steps till the parameter converge.

$$\text{or} \quad Z = \text{Min} \|W(Y - A\underline{X})\|^2 = \text{Min} \|(W\underline{Y} - WA\underline{X})\|^2$$

$$\therefore \frac{\partial Z}{\partial \underline{X}} = \underline{0} \Rightarrow (WA)^T(W\underline{Y} - WA\underline{X}) = \underline{0}$$

$$\therefore A^TW^TW\underline{Y} - A^TW^TWA\underline{X} = 0$$

$$\therefore \underline{\hat{X}} = (A^TW^TWA)^{-1}(A^TW^TW\underline{Y})$$

Lecture 22

25/09/2012

Stereopsis / Stereo Vision

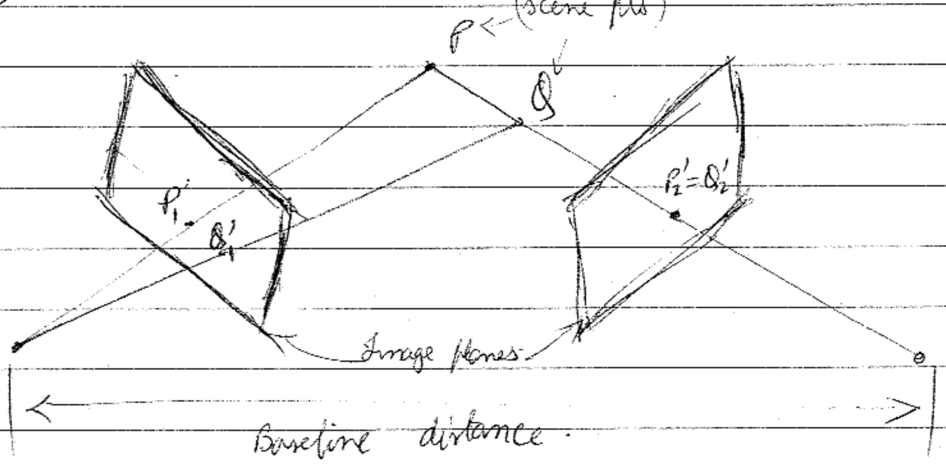
Infer about a scene using 2 images, similar to binocular vision.

Why Stereo vision?

→ ~~Occlusion~~ Occlusion (All pts. ~~on~~ the line of sight get mapped to a single pt in the image plane). Thus there is info. loss when ~~the~~ transforming from 3D to 2D. Here, precisely, the depth info. is lost

Assumptions:-

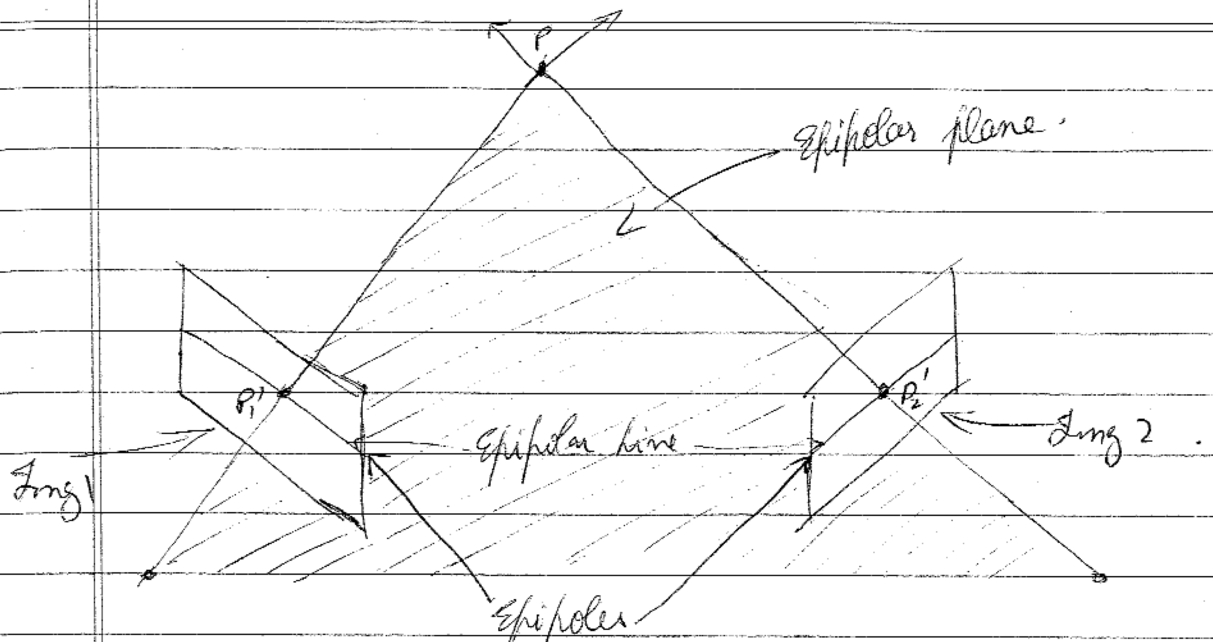
1. We use 2 cameras.
2. Positions of the cameras are known.
3.



Correspondance Problem.-

If we are able to somehow tell as to which pt. in the left image (P_1' which is the mapping of scene pt. P corresponds to a point (P_2', which is also a mapping of P, but in right image plane) in the right image, then we can easily extract depth info. ~~about~~ of the point (scene pt P) under consideration.)

One way to deal with the correspondance problem is to use the epipolar constraint, stated as follows.



(whenever the image plane cuts the shaded triangle).

Corresponding pts lies on the epipolar line in the 2nd image i.e. All pts between P & P_2' can be found on the epipolar line in image 1 & those between P & P_1' can be found on the epipolar line in image 2. Thus, the correspondance problem is reduced to 1D search along the epipolar line.

classmate

Date 28/10/2012

Page 110

Lecture 23

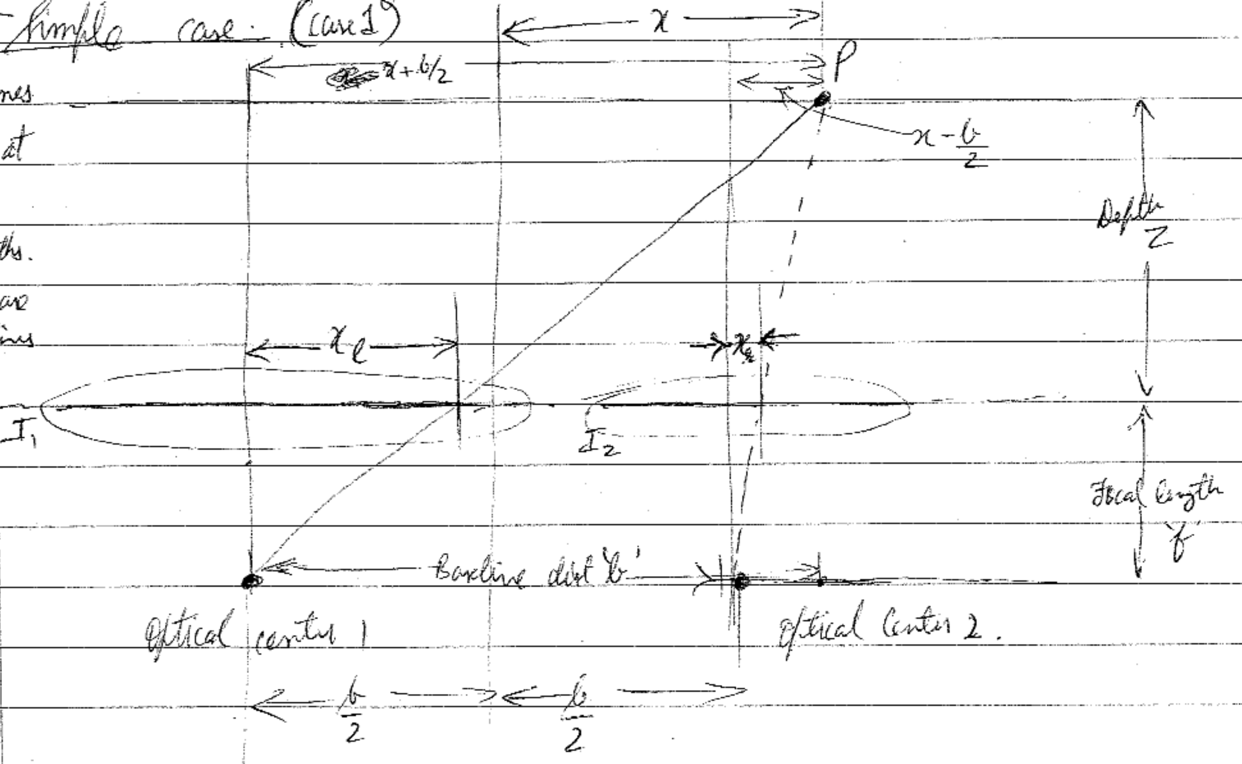
Depth estimation in Stereo -

* Use of epipolar constraint

Simple case (case 1)

* Parallel Img planes
* Focal pts are at same height.
* Same Focal lengths.
* Epipolar lines are horizontal scanlines

I_1, I_2 Image plans



Now, $\dfrac{x_\ell}{f} = \dfrac{x + b/2}{Z}$

Also, $\dfrac{x_r}{f} = \dfrac{x - b/2}{Z}$

$\therefore \dfrac{x_\ell - x_r}{f} = \dfrac{x + b/2 - x + b/2}{Z} = \dfrac{b}{Z}$

$\therefore (x_\ell - x_r) = \dfrac{f \cdot b}{Z}$

or $\boxed{Z = \dfrac{f \cdot b}{(x_\ell - x_r)}}$

Here, $x_\ell - x_r$ is known as the disparity or the displacement

Now, f & b are constants.

$\therefore Z \propto \frac{1}{(x_l - x_r)}$ i.e. depth estimation is reduced to finding $(x_l - x_r)$ - the disparity.

Case 2:- (Same as case 1, except that the image planes are non parallel).
In this case, first rectification is used (which includes interpolation) so as to form the image onto parallel planes, aligned with the optical centers.

For correspondance what can be matched?
1. Objects? (Difficult to identify for a machine).
2. Edges? (Likely, but may fail in case of images like |||| etc.
3. Pixels? (Too lose a criteria, imagine two pictures of a wall).
4. Collection of pixels? (Makes more sense).

Lecture 24 & 25

01/10/2012

The correspondance problem can be solved (attempted) using.

① Local methods:-
Work on (group of pixels) / patches.

(A) correlation - based algorithms :-
Produce a dense set of correspondances. (Sparse set)

(i) SSD (Sum of squared diff.).

$$SSD = \sum_{[i,j] \in R} \left(f(i,j) - g(i,j) \right)^2$$

(ii) Correlation $C_{fg} = \sum_{[i,j] \in R} f(i,j)\, g(i,j)$.

(iii) SAD (Sum of absolute differences).

$$SAD = \sum_{[i,j] \in R} \left| f(i,j) - g(i,j) \right|$$

(iv) Adaptive window, etc.

(v) Now, $SSD = \sum_{(i,j) \in R} \left(f(i,j) - g(i,j) \right)^2$

$$= \sum_{(i,j) \in R} \left[\left(f(i,j) \right)^2 + \left(g(i,j) \right)^2 - 2 f(i,j)\, g(i,j) \right]$$

$$= \sum_{(i,j) \in R} \left(f(i,j) \right)^2 + \sum_{(i,j) \in R} \left(g(i,j) \right)^2 - \underbrace{2 \sum_{(i,j) \in R} f(i,j)\, g(i,j)}_{\text{Cross Correlation as}}$$

Photometric Constraint:
Same world pts has same intensity in both images.
→ This may not always hold because,
(a) Noise, (b) Specularity (c) Foreshortening.

(B) Feature based methods :-
Sparse set of correspondances, using most similar feature pairs
Similarity measure must be adapted to the type of feature
Which features?
(1) Corners — SSD, cross-correlation.
(2) Edges, lines — orientation, contrast, length, coordinates etc.

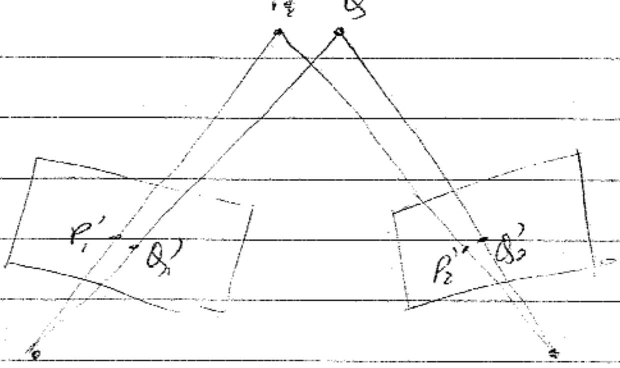
Eg:- Comparing lines.

$$S = \frac{1}{w_l(l_l - l_r)^2 + w_\theta(\theta_l - \theta_r)^2 + w_m\left[(x_l - x_r)^2 + (y_l - y_r)^2\right] + w_c(c_l - c_r)^2}$$

lengths orientation mid-pt contrast

↑ similarity ⟷ ↑ S

Ordering Constraint:-
Usually, order of two pts in image is same.
Eg:-



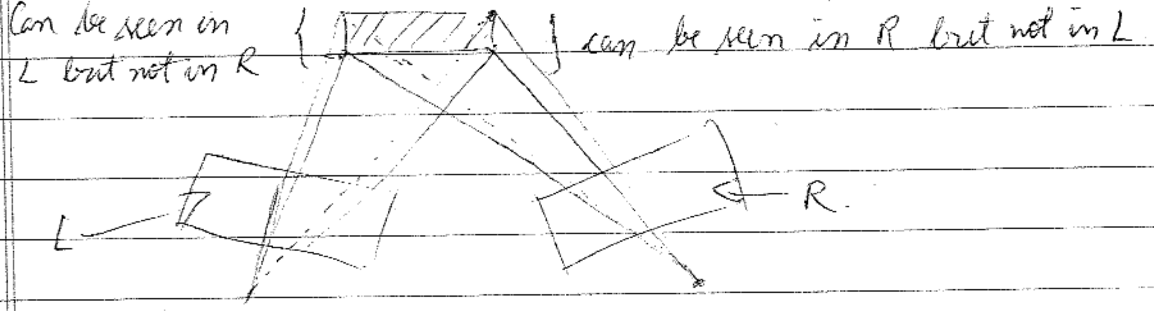
P gets mapped on left of Q.

This constraint may not always be true,
Eg:

P gets mapped on left of Q in L
right in R

Other constraints:-

1. Smoothness: Disparity usually doesn't change too quickly.
2. Uniqueness: Each feature can at most have 1 match.
3. Occlusion & ~~d~~ disparity are connected.



Mean correction, $\overset{\text{for comparing L\&R}}{\text{normalization}}$, soft constraints, hard constraints optimization depending on the chosen constraint.

Normalized pixel:- $\hat{I}(x,y) = \dfrac{I(x,y) - \bar{I}}{\|I - \bar{I}\|_{W_m(x,y)}}$,

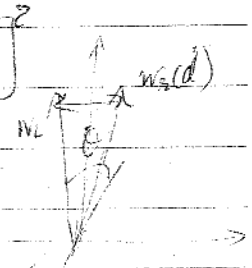
where, $\bar{I} = \dfrac{1}{|W_m(x,y)|} \sum_{(u,v)\in W_m(x,y)} I(u,v)$.

& $\|I\|_{W_m(x,y)} = \sqrt{\sum_{(u,v)\in W_m(x,y)} \left(I(u,v)\right)^2}$

Image metrics:-

Normalized SSD

$C_{SSD}(d) = \sum_{(u,v)\in W_m(x,y)} \left[\hat{I}_L(u,v) - \hat{I}_R(u-d,v)\right]^2$



$= \|\underbrace{w_L}_{\text{vector from left window}} - \underbrace{w_R(d)}_{\text{vector from right window}}\|^2$

Normalized correlation:- $C_{NC}(d) = \sum_{(u,v)\in W_m(x,y)} \hat{I}_L(u,v) \cdot \hat{I}_R(u-d,v) = w_L \cdot w_R(d)$

$$d^* = \arg\min_d \|W_L - W_R(d)\|^2 = \arg\max_d W_L \cdot W_R(d)$$

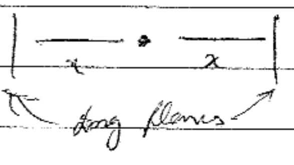
See Middlebury site : stereo images
Taxonomy paper IJCV 2002

Better disparity values → high accuracy → feature based.
* Problem is sparse disparity values — interpolation is needed

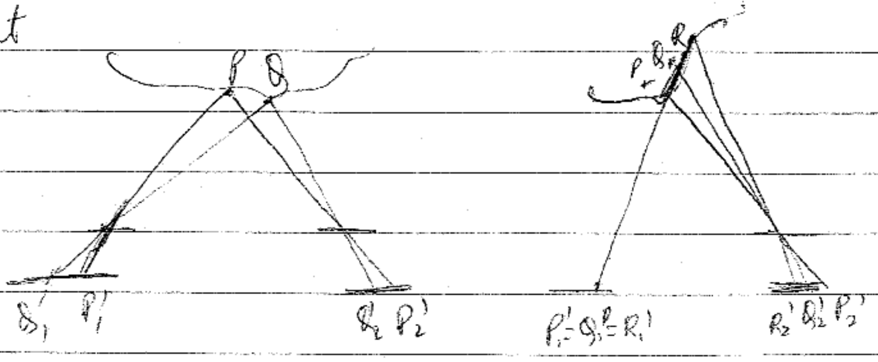
H.W. Get disparity from 2 given images using features.

Global approach:-

1. Epipolar constraint: Common Epipolar line.
2. Photometric compatibility constraint: Intensities on L & R images do not differ much.
3. Geometric similarity constraints:- length, orientation, etc do not differ much in 2 images.
4. Disparity constraint:- $\|(P_L - P_R) - (Q_L - Q_R)\|$ is min. over entire image.
5. Ordering constraint
6. Disparity limit constraint.



6. Uniqueness constraint



Minimizations using graph cuts
Use Gabor filter bank on L & R images & formulate the data term, comp.
Check if the data can be converted to TCS.

Lecture 26

08/10/2012

Global approach:-
Optimization done on entire image & not patchwise.

Cost function = $E_{data} + \lambda E_{prior}$

Assuming the epipolar constraint (remember the simple case, where correspondance search is reduced to search on the epipolar line) we have,

$f_R(x + d(x,y), y)$ for right, given that
$f_L(x, y)$ is the px. val in left image.
Here $d(x,y)$ is the disparity at pt. (x,y).

Hence, cost function is,

$$\text{Cost function} = \sum_x \sum_y \left[f_L(x,y) - f_R(x+d(x,y), y)\right]^2 +$$
$$\lambda \sum_x \sum_y \left[\left(d(x,y) - d(x+1,y)\right)^2 + \left(d(x,y) + d(x,y+1)\right)^2 \right]$$

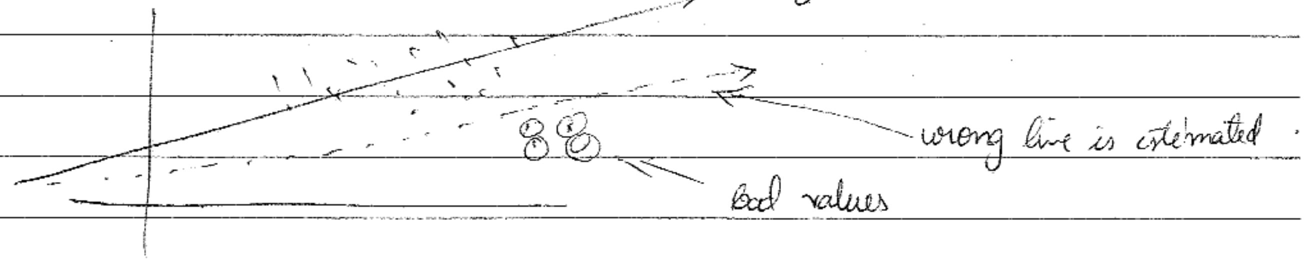
The prior has come from the knowledge that the disparity values do not change much for neighbouring pixels.

Hence, we need to minimize the above cost function & find $d(x,y)$ that gives the minimum cost.

There are many variants to this approach. Graph cuts based approaches are popular in recent times.

Estimators we have studied:
① LS, CLS, WLS, TLS, EM, ML

In these type of estimators, outliers can give a bad estimate.
Breakdown pt. is the smallest proportion of bad data or outliers that can give us a wrong estimate. ~~Eg~~.



Breakdown pt: smallest proportion of bad data contamination that forces the estimate to move to an arbitrary value.

$\therefore$ Breakdown pt $= \dfrac{\text{minimum \# pts to be perturbed}}{\text{Total \# samples}}$

For L.S, W.L.S, C.L.S, T.L.S; this value is $\frac{1}{N}$

Least median squares (Lmeds) has a breakdown pt $= 1/2$.

L_{meds} : $\min\left[\text{median}\left(y_i - (mx_i + c)\right)^2\right]$

This is difficult to differentiate / not differentiable & hence not ~~popu~~ popular.
Here, even if half the data has error, we get a good estimate. Hence, breakdown pt. $= 1/2$.

Now consider the estimate (parameter) itself as a r.v.
Eg:- mean (M) & var. (σ_v^2).
m & σ_v^2 are true values.

$m = \int_{-\infty}^{\infty} x f_x(x)\,dx$, $\sigma_v^2 = \int_{-\infty}^{\infty} (x-m)^2 f_x(x)\,dx$

not practical

$E[X_i] = m.$ $\quad E[(X_i - m)^2] = \sigma_v^2$

Now, $M = \frac{1}{N}\sum_i X_i$ & $\sigma_v^2 = \frac{1}{N}\sum_i (X_i - M)^2$

Unbiased Estimator:-

Expected value of estimate = true value of estimate.

Here, $E[M] = \frac{1}{N} \cdot E[X_1 + X_2 + \dots + X_N]$.

$$= \frac{1}{N} \cdot [E(X_1) + E(X_2) + \dots + E(X_N)]$$

$$= \frac{1}{N} [m + m + \dots + m].$$

$$= \frac{1}{N}(Nm) = m.$$

But m = true value.

$\therefore$ M is an unbiased estimator ~~of M~~.

$$E[\sigma_v^2] = \frac{1}{N} E[(X_1 - m)^2 + (X_2 - m)^2 + \dots + E(X_N - m)^2].$$

$$= \frac{1}{N}[E[(X_1 - m)^2] + E(X_2 - m)^2 + \dots E[(X_N - m)^2]]$$

$$\cancel{= \frac{1}{N}[E[X_1]^2 - [E(X_1)]^2 + \dots E[X_N^2] - E[X_N}$$

$$\text{var}(M) = E[(M - m)^2]$$

$$= E\left[\left(\frac{\sum X_i}{N} - m\right)^2\right]$$

Now considers,

$$E\left[\sigma_v^2\right] = E\left[\frac{1}{N}\sum (X_i - M)^2\right]$$

$$= \frac{1}{N} E\left[\sum \left(X_i^2 + M^2 - 2MX_i\right)\right]$$

$$= \frac{1}{N}\left[\sum E(X_i^2) + E\left(\sum (M^2)\right) - E\left(2M\sum X_i\right)\right]$$

$$\cancel{= \frac{1}{N}\left[\sum\left[(E(X))^2 + \sigma_v^2\right] - E\left[2M(NM)\right] + \sum(M^2)\right]}$$

$$= \frac{1}{N}\left[\sum E(X_i^2) + \sum E(M^2) - 2E\left[M.\sum X_i\right]\right]$$

$$= \frac{1}{N}\left[\sum\left[(E(X_i))^2 + \sigma_v^2\right] + \sum E(M^2) - 2E\left[MNM\right]\right]$$

$$= \frac{1}{N}\left[\sum\left(M^2 + \sigma_v^2\right) + N\left(E[M^2]\right) - 2NE\left[M^2\right]\right]$$

$$= \frac{1}{N}\left[\sum\left(M^2 + \sigma_v^2\right) - N\cdot E\left[M^2\right]\right] \quad — (1)$$

where, $E\left[M^2\right] = E\left[\left(\frac{1}{N}\sum X_i\right)^2\right] = \frac{1}{N^2} E\left[\left(\sum X_i\right)^2\right]$

$$= \frac{1}{N^2} E\left[\left(\sum_i X_i\right)\left(\sum_j X_j\right)\right] = \frac{1}{N^2} E\left[\underbrace{\sum_i X_i^2}_{N \text{ terms}} + \underbrace{\sum_{i\neq j} X_i X_j}_{N^2 - N = N(N-1) \text{ terms}}\right]$$

$$\therefore E\left[M^2\right] = \frac{1}{N^2}\left[E\left[\left(\sum X_i^2\right)\right] + E\left[\sum_{i \neq j} X_i X_j\right]\right]$$

Independe

$$= \frac{1}{N^2}\left[\sum E\left(X_i^2\right) + \sum_{i \neq j} E\left(X_i\right) \cdot E\left(X_j\right)\right]$$

$$= \frac{1}{N^2}\left[\sum\left[\left(E(X_i)\right)^2 + \sigma_v^2\right] + \sum_{i \neq j}\left(E(X_i) \cdot E(X_j)\right)\right]$$

$$\cancel{= \frac{1}{N^2}\left[\sum\left(E(X_i)\right)^2 +\right.}$$

$$= \frac{1}{N^2}\left[\underbrace{\sum\left(M^2 + \sigma_v^2\right)}_{N \text{ terms}} + \underbrace{\sum_{i \neq j} \cancel{\overset{M}{E(X_i)} \cdot \overset{M}{E(X_j)}}}_{N(N-1) \text{ terms}}\right]$$

$$= \frac{1}{N^2}\left[M^2 \sum_{1}^{N} 1 + \sigma_v^2 \sum_{1}^{N} 1 + M^2 \sum_{1}^{N(N-1)} 1\right]$$

$$= \frac{1}{N^2}\left[M^2 N + \sigma_v^2 N + M^2 N(N-1)\right]$$

$$= \frac{1}{N^2}\left[\cancel{M^2 N} + \sigma_v^2 N + M^2 N^2 - \cancel{M^2 N}\right]$$

$$= \left(M^2 + \sigma_v^2 / N\right)$$

$$\therefore E\left[\sigma_v^2\right] = \frac{1}{N}\left[N M^2 + N\sigma_v^2 - N\left(M^2 + \frac{\sigma_v^2}{N}\right)\right]$$

$$= \frac{1}{N}\left[\cancel{N M^2} + N\sigma_v^2 - \cancel{N M^2} - \sigma_v^2\right]$$

$$= \frac{1}{N}\left[N\sigma_v^2 - \sigma_v^2\right]$$

$$= \frac{N-1}{N}\sigma_v^2$$

$$\therefore \boxed{E\left[\sigma_v^2\right] = \frac{N-1}{N}\sigma_v^2}$$

Hence, ~~not~~ a biased estimator.

Instead if we use $\sigma_v^2 = \frac{1}{N-1}\left[\sum (X_i - M)^2\right]$

Then $E[\sigma_v^2] = \frac{N-1}{N-1}\sigma_v^2 = \sigma_v^2$

which becomes an unbiased estimator.

If the estimated parameters have zero variance, then these are consistent estimators.

classmate
Date
Page (122)

Lecture 27

15/10/2012

09/10/2012 → Surprise quiz.

3 questions.

1. 2 ~~they were given~~ transformations were given, for 3D to 2D & was asked whether these were linear or non-lin
2. If non-linear, how to make them linear.
3. How many min. pts. would be required to to determine the unknown parameters.

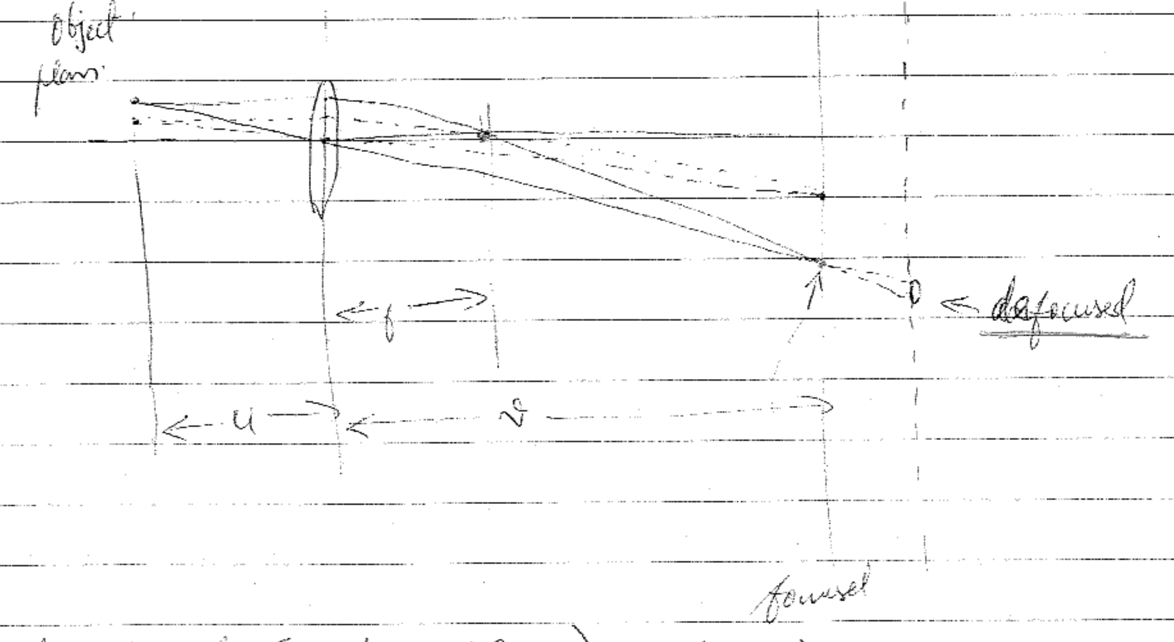
Another approach for depth estimation.. Depth from defocus (DFD)

(1) Uses real aperture or thin lens (not pinhole) model.
.. even this model is not a practical camera model.

(2) DFD uses Len's Law :- $\boxed{\frac{1}{u} + \frac{1}{v} = \frac{1}{f}}$

u → object distance, v → image distance, f → focal length

which holds good only for focused points.



(3) Point Spread Function (PSF) ≡ $h(x,y)$

→ impulse response of the camera
→ Response of the camera to pt. light source

Image formed.

$$g(x,y) = f(x,y) * h(x,y) + n(x,y)$$

noise.

Image ~~formed~~ with thinlens model for ~~an~~ ~~of~~ object pts. in a plane (which are at same dist. from camera)

Original focused image (from pinhole model)

Space invariant blur

Considering camera as a linear space invariant system.
i.e. all object pts. that actually lie on the same obj plane, have same amount of blur.

However, we have space varying blur for a given object having varying depths.

$g(x,y) = f(x,y) * h(x,y) + \cdot$

Lecture 28

16/10/2012

Depth from focus:-

Focus as a cue.

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$ holds for all pts. in focus.

Then, by varying params (v & f) we [can capture multiple] [images (keeping 'u' constant)] so that all pts. are focused in some or the other image.

Problem [Here, the problem is of finding which pts are focused.]

Back to DFD discussion:-

Recall from DFD

(3) $g(t) = \int f(\tau) \cdot h(t-\tau)\, d\tau$

Then $f(x,y)$ can be used as,

Convert (3) into a 2 variable equation

(4) $$g(x,y) = \iint_{l\ m} f(l,m) \cdot h(\underbrace{x-l, y-m}_{\text{Space/shift invariant}})\, dl \cdot dm + n(x,y)$$

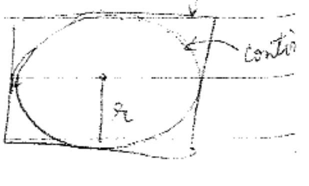
Convert (4) into discrete — integration $\iint$ into summation $\Sigma\Sigma$

(5) In discrete form, we have

$$\sum_l \sum_m f(l,m) \cdot h(x-l, y-m)$$ [space invariant blur

(6) Here $h(x,y)$ is shift invariant & if we consider the blur to have a radius 'r' then

$$\boxed{h(x,y) = \frac{1}{\pi r^2}, \quad x^2+y^2 < r^2}$$



For space variant blur, we have,

$$g(x,y) = \sum_{l}\sum_{m} f(l,m)\cdot h(l,m,x,y) + n(x,y)$$ [space variant blur]

2 Where, $h(l,m,x,y)$ at a location (x,y) is given by

$$h(x,y,l,m) = \frac{1}{\pi r^2(x,y)} \quad \text{for } (l-x)^2 + (y-m)^2 < r^2(x,y)$$

Gaussian : Best for parameterization.

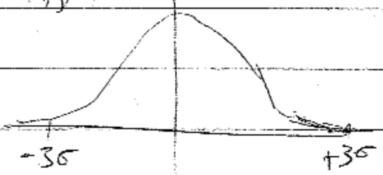
Parameterize the blur by modelling it as Gaussian, with a variance that decides the amt. of blur & hence the depth.

$$\therefore h(x,y,l,m) = \frac{1}{2\pi\sigma^2_{(x,y)}} \cdot e^{-\frac{(l^2+m^2)}{2\sigma^2}}$$

$\therefore$ where $h(x,y,l,m)$ is actually $\dfrac{\left(e^{-\frac{l^2}{2\sigma^2_{(x,y)}}} \cdot e^{-\frac{m^2}{2\sigma^2_{(x,y)}}}\right)}{\left(\sqrt{2\pi}\,\sigma_{(x,y)}\right)\cdot\left(\sqrt{2\pi}\,\sigma_{(x,y)}\right)}$

Assuming the mean $= 0$

$\left.\begin{array}{l} -3 \text{ to } +3 \text{ in } l \text{ direction,} \\ -3 \text{ to } +3 \text{ in } m \text{ direction.} \\ \therefore \text{Mask is of size } 7\times 7 \end{array}\right\}$ if $\sigma(x,y) = 1$



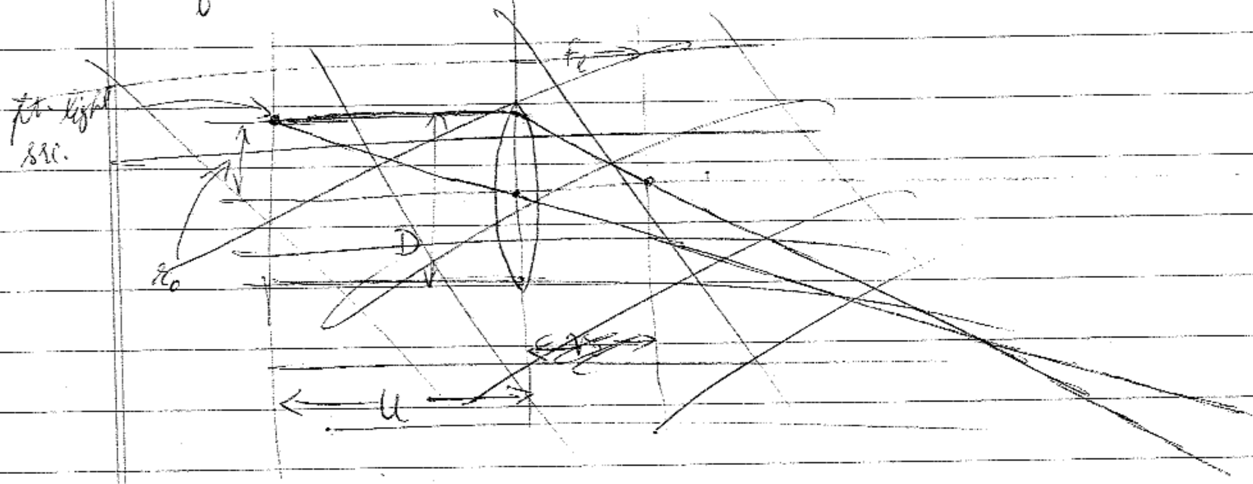
Lecture 29

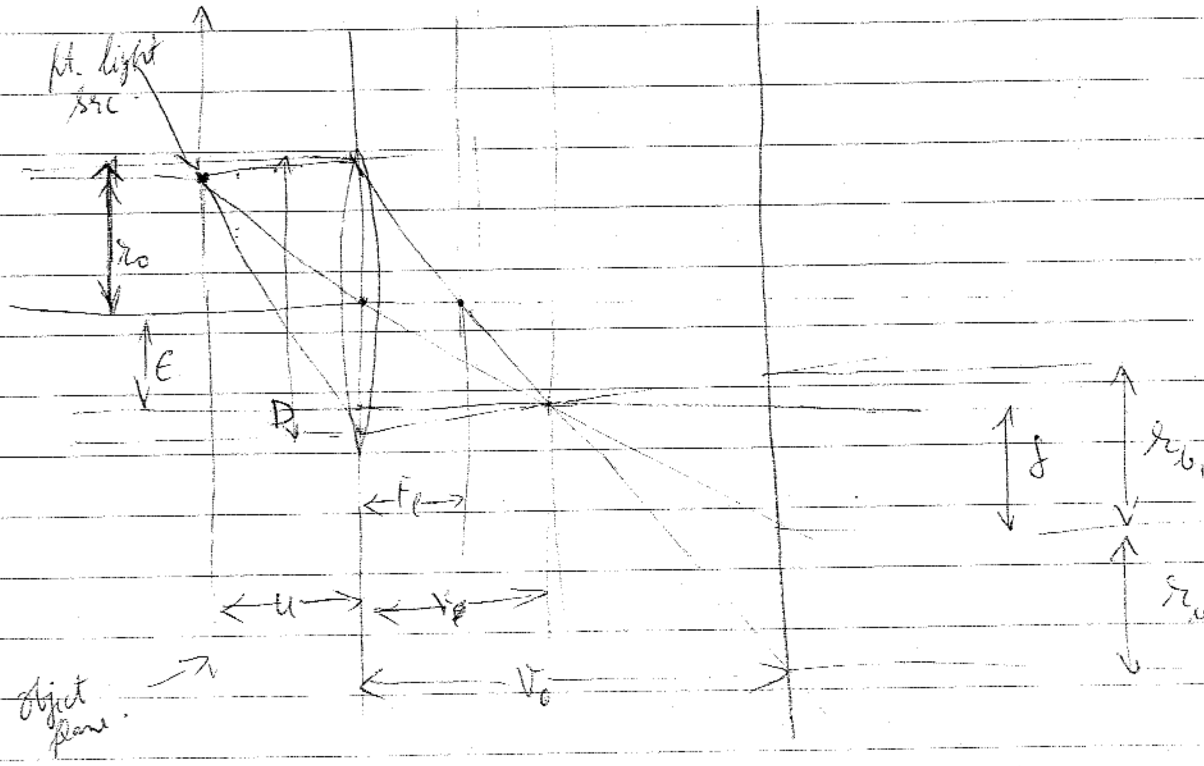
19/10/2012

Depth from Defocus:- Pentland 1982

(1) Blur radius is independent of the location of the pt. src. on the image (object) plane (at a dist. 'u' from lens center).

Proof:-





$$\frac{r_{b_1} - \delta}{v_0 - v} = \frac{r_0 - \epsilon}{v} \quad -(1)$$

$$\& \quad \frac{r_{b_2} + \delta}{v_0 - v} = \frac{r_0 + \epsilon}{v} \quad -(2)$$

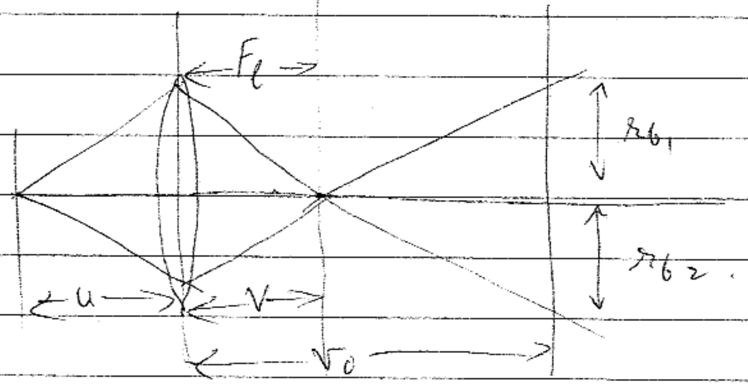
& $\dfrac{\delta}{v_0 - v} = \dfrac{\epsilon}{v}$ —(3)

Using (1) & (3) we get,

$$r_{b_1} = r_0\left(\frac{v_0}{v} - 1\right)$$

similarly, $\boxed{r_{b_2} = r_0\left(\dfrac{v_0}{v} - 1\right)}$

$\therefore\ r_{b_1} = r_{b_2} = r_b$



Again, $r_{b_1} = r_{b_2} = r_0\left(\dfrac{v_0}{v} - 1\right)$ ... $\epsilon = 0, \delta = 0$.

$\therefore$ All pts. at dist 'u' have a blur radius $= r_b = r_0\left(\dfrac{v_0}{v} - 1\right)$

Two images with 2 diff. values for the cam. param (v_1 & v_2). If the blur is modelled as Gaussian, it can be parametrized by variance of the Gaussian blur.

↑ var ⟷ ↑ blur.
↓ var ⟷ ↓ blur.
0 var ⟷ pt. is focused.

$\therefore\ \sigma \propto$ depth $\propto$ blur radius (r_b)

Let $\sigma = \rho r_b$

~~$\sigma = \rho r_b$~~

P.T.O.

$$r_1 = r_0\left(\frac{V_1}{v} - 1\right)$$

$$= r_0 . V_1\left(\frac{1}{v} - \frac{1}{V_1}\right).$$

$$= r_0 V_1\left(\frac{1}{f} - \frac{1}{u} - \frac{1}{V_1}\right).$$

$$\boxed{\therefore \sigma_1 = \rho . r_0 . V_1\left(\frac{1}{f} - \frac{1}{u} - \frac{1}{V_1}\right)}.$$

Also, $\sigma_2 = \rho r_0 V_2\left(\frac{1}{f} - \frac{1}{u} - \frac{1}{V_2}\right).$

Now, $\frac{\sigma_1}{\rho r_0 V_1} = \frac{1}{f} - \frac{1}{u} - \frac{1}{V_1}.$

$$\boxed{\therefore \frac{1}{f} - \frac{1}{u} = \frac{\sigma_1}{\rho r_0 V_1} + \frac{1}{V_1}}$$

Also, $\boxed{\frac{1}{f} - \frac{1}{u} = \frac{\sigma_2}{\rho r_0 V_2} + \frac{1}{V_2}}.$

$$\therefore \frac{\sigma_1}{\rho r_0 V_1} + \frac{1}{V_1} = \frac{\sigma_2}{\rho r_0 V_2} + \frac{1}{V_2}.$$

$$\therefore \left(\frac{1}{V_2} - \frac{1}{V_1}\right) = \frac{1}{\rho r_0}\left(\frac{\sigma_1}{V_1} - \frac{\sigma_2}{V_2}\right)$$

$$\therefore \frac{V_1 - V_2}{V_1 V_2} = \frac{1}{\rho r_0}\left(\frac{V_2\sigma_1 - \sigma_2 V_1}{V_1 V_2}\right).$$

$$\therefore \left(V_1 - V_2\right)\rho r_0 = \left(V_2\sigma_1 - \sigma_2 V_1\right)$$

$$\therefore V_2\sigma_1 = (V_1 - V_2)\, \rho r_0 + \sigma_2 V_1$$

$$\therefore \sigma_1 = \left(\frac{V_1 - V_2}{V_2}\right)\rho r_0 + \sigma_2\left(\frac{V_1}{V_2}\right)$$

$$\therefore \boxed{\sigma_1 = \beta + \alpha'\sigma_2}$$

Now, look for another pt. relating σ_1 & σ_2 to get σ_1 or σ_2 by solving the two equations simultaneously.

We can get 2^{nd} relation between σ_1 & σ_2 using Subbarao's approach. : Parallel depth recovery -- ICCV 1988.

* Qn : Why noise was not considered in this approach?
— see page 130

We know that blur is space varying for solving the depth estimation ~~to~~ problem using defocus as the cue.

Let g_1 & g_2 be the two captured images by varying v_1 & v_2.

$$\therefore g_1(x,y) = \sum_l \sum_m h_1(l,m,x,y)\cdot f(l,m) + n(x,y).$$

$$\& \; g_2(x,y) = \sum_l \sum_m h_2(l,m,x,y)\cdot f(l,m) + n(x,y).$$

$$\text{Where, } h_1(l,m,x,y) = \frac{1}{2\pi\sigma_1^2(x,y)}\cdot e^{-\frac{l^2+m^2}{2\sigma_1^2(x,y)}}$$

spatial frequencies

Using FFT, we have $H_1(u,v) = H_1(u,v,x,y) = e^{-\frac{1}{2}(u^2+v^2)\cdot\sigma_1^2(x,y)}$
similar equation for H_2.

Then, $\dfrac{G_1}{G_2} = \dfrac{H_1 \cancel{F}}{H_2 F}$

P.T.O.

Get equations in terms of σ_1 & σ_2 & then find σ_1, σ_2 using the earlier eq. & this one simultaneously.

Since $\sigma \propto r_0 \propto$ depth, we get depth.

* Why noise was not considered?

Ans: M. Subbarao's approach aimed at recovering depth map along with recovering a focused image. The focused image recovery method involved deconvolution with the pt. spread function. However, deconvolution poses many serious difficulties especially in presence of noise. Therefore the model avoided the noise term.

Miscellaneous 1

Computer Vision

Bayes' Rule :-

$$P(A|B)\cdot P(B) = P(AB) = P(BA) = P(B|A)\cdot P(A)$$

Suppose we have some data and need to estimate the parameters of the distribution to which the data belongs, then

$P(\text{param}\,|\,\text{data})\cdot P(\text{data})$ is to be determined.

But $P(\text{param}\,|\,\text{data})\cdot P(\text{data}) = P(\text{data}\,|\,\text{param})\cdot P(\text{param})$.

or $P(\underline{\theta}|Z)\cdot P(Z) = P(Z|\underline{\theta})\cdot P(\underline{\theta})$

$$\therefore P(\underline{\theta}|Z) = \frac{P(Z|\underline{\theta})\cdot P(\underline{\theta})}{P(Z)}.$$

But $P(Z)$ is a constant

$$\therefore \underset{\forall \underline{\theta}}{\text{argmax}}\, P(\underline{\theta}|Z) = \underset{\forall \underline{\theta}}{\text{argmax}}\, \overset{\text{likelihood}}{P(Z|\underline{\theta})}\cdot P(\underline{\theta}).$$

Here the maximum aposterior (MAP) estimate of parameters is $\underset{\forall \underline{\theta}}{\text{argmax}}\, P(\underline{\theta}|Z)$, i.e. ~~take~~ that value of $\underline{\theta}$ that maximizes $P(\underline{\theta}|Z)$.

The value $P(\underline{\theta}|Z)$ is called the posterior probability while $P(\underline{\theta})$ is called the prior probability of $\underline{\theta}$.

In case of uniform distribution, $P(\underline{\theta})$ is constant

$$\therefore \underset{\forall \underline{\theta}}{\text{argmax}}\, P(\underline{\theta}|Z) = \underset{\forall \underline{\theta}}{\text{argmax}}\, P(Z|\underline{\theta})$$

∴ This is called the maximum likelihood estimate or MLE

The $\text{argmax}\, P(\underline{\theta}|Z) = \text{argmax}\, [\,P(Z|\underline{\theta})\,]$

Q. Given a data set find the parameters of a Gaussian curve to which this data belongs

Soln. Need to estimate mean & variance i.e. $\hat{\mu}, \hat{\sigma}$.
Now,

$$\hat{\mu}, \hat{\sigma} = \underset{\mu, \sigma}{\arg\max}\ P(\mu, \sigma \mid z_1, z_2 \ldots z_n)$$

But due to MLE,

$$\underset{\mu, \sigma}{\arg\max}\ P(\mu, \sigma \mid z_1 \ldots z_n) = \underset{\mu, \sigma}{\arg\max}\ P(z_1 \ldots z_n \mid \mu, \sigma)$$

Assuming that the data has come from i.i.d. Gaussian we have.

$$P(z_1 \ldots z_n) = P(z_1) \cdot P(z_2) \cdots P(z_n)$$

$$\& \ P(z_1 \ldots z_n \mid \mu, \sigma) = P(z_1 \mid \mu, \sigma) \cdot P(z_2 \mid \mu, \sigma) \cdots P(z_n \mid \mu, \sigma).$$

$$\therefore \hat{\mu}, \hat{\sigma} = \underset{\mu, \sigma}{\arg\max} \left[P(z_1 \mid \mu, \sigma) \cdot P(z_2 \mid \mu, \sigma) \cdots P(z_n \mid \mu, \sigma) \right]$$

But $P(z \mid \mu, \sigma) = \dfrac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2}$

$$\therefore \hat{\mu}, \hat{\sigma} = \underset{\mu, \sigma}{\arg\max} \left[\left(\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{z_1-\mu}{\sigma}\right)^2} \right) \cdots \left(\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{z_n-\mu}{\sigma}\right)^2} \right) \right]$$

$$= \underset{\mu, \sigma}{\arg\max} \left[\left(\frac{1}{\sqrt{2\pi}\,\sigma} \right)^n \cdot e^{-\frac{1}{2}\left(\frac{z_1-\mu}{\sigma}\right)^2 - \frac{1}{2}\left(\frac{z_2-\mu}{\sigma}\right)^2 \cdots -\frac{1}{2}\left(\frac{z_n-\mu}{\sigma}\right)^2} \right]$$

$$= \underset{\mu, \sigma}{\arg\max} \left[\left(\frac{1}{\sqrt{2\pi}\,\sigma} \right)^n \cdot e^{-\frac{1}{2} \sum_i \left(\frac{z_i-\mu}{\sigma}\right)^2} \right]$$

the value of μ, σ that maximizes $P(z_1 \dots z_n | \mu, \sigma)$ also maximizes $\ln P(z_1 \dots z_n | \mu, \sigma)$ as $\ln$ is a monotonically increasing function.

$$\therefore \hat{\mu}, \hat{\sigma} = \underset{\mu, \sigma}{\operatorname{argmax}} \left[\ln \left[\left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\frac{1}{2} \sum_i \left(\frac{z_i - \mu}{\sigma} \right)^2} \right] \right]$$

$$= \underset{\mu, \sigma}{\operatorname{argmax}} \left[n \ln \left(\frac{1}{\sqrt{2\pi}\,\sigma} \right) + -\frac{1}{2} \sum_i \left(\frac{z_i - \mu}{\sigma} \right)^2 \ln e \right].$$

$$= \underset{\mu, \sigma}{\operatorname{argmax}} \left[-n \ln(\sqrt{2\pi}) - n \ln(\sigma) - \frac{1}{2} \sum_i \left(\frac{z_i - \mu}{\sigma} \right)^2 \right].$$

But $n \ln(\sqrt{2\pi})$ & ~~$n \ln(\sigma)$~~ ~~are~~ constants.

$$\therefore \hat{\mu}, \hat{\sigma} = \underset{\mu, \sigma}{\operatorname{argmax}} \left[-\frac{1}{2} \sum_i \left(\frac{z_i - \mu}{\sigma} \right)^2 \Big\} - n \ln(\sigma) \right]$$

$$= \underset{\mu, \sigma}{\operatorname{argmax}} \left[\frac{1}{2} \sum_i \left(\frac{z_i - \mu}{\sigma} \right)^2 \Big\} - n \ln(\sigma) \right]$$

$$= \underset{\mu, \sigma}{\operatorname{argmin}} \left[\frac{1}{2} \sum_{i=1}^{n} \left(\frac{z_i - \mu}{\sigma} \right)^2 + n \ln(\sigma) \right]$$

Let $T = \left[\frac{1}{2} \sum_{i=1}^{n} \left(\frac{z_i - \mu}{\sigma} \right)^2 + n \ln(\sigma) \right]$

To find μ & σ that minimize T, we have.

$$\frac{\partial T}{\partial \mu} = 0 \quad \& \quad \frac{\partial T}{\partial \sigma} = 0.$$

classmate

4. Date Page

$\frac{\partial T}{\partial \mu} = 0$ gives $\frac{\partial}{\partial \mu}\left[\left\{\frac{1}{2}\sum_{i=1}^{n}\left(\frac{z_i-\mu}{\sigma}\right)^2\right\} + n\ln(\sigma)\right] = 0.$

$\therefore 0 = \sum_{i=1}^{n} \frac{\partial}{\partial \mu}\left(\left(\frac{z_i-\mu}{\sigma}\right)^2\right) + \frac{\partial\, n\ln(\sigma)}{\partial \mu}$

$= \sum_{i=1}^{n} 2\left(\frac{z_i-\mu}{\sigma}\right)\cdot\left(\frac{1}{\sigma}\right)(-1) + 0$

$= \left(\frac{1}{2}\right)\frac{-2}{\sigma^2}\sum_{i=1}^{n}(z_i-\mu)$

$\therefore = \frac{-1}{\sigma^2}\left\{\sum_{i=1}^{n} z_i - \sum_{i=1}^{n}\mu\right\}$

$= \frac{-1}{\sigma^2}\left\{\sum_{i=1}^{n} z_i - n\cdot\mu\right\}$

But since $\sigma \neq 0$ we have,

$\sum_{i=1}^{n} z_i - n\mu = 0.$

$\therefore n\hat{\mu} = \sum_{i=1}^{n} z_i$

$\therefore \boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} z_i}$

Taking $\frac{\partial T}{\partial \sigma} = 0$ gives,

$$0 = \frac{\partial}{\partial\sigma}\left[\sum_{i=1}^{n}(z_i-\mu)^2\cdot\sigma^{-2}\right] + n\frac{\partial \ln(\sigma)}{\partial\sigma}.$$

$$= (z_i-\mu)^2\cdot\sum_{i=1}^{n}\frac{\partial(\sigma^{-2})}{\partial\sigma} + n\frac{\partial\ln(\sigma)}{\partial\sigma}.$$

$$= (z_i-\mu)^2\cdot\left(\frac{1}{-2}\right)\sigma^{-3} + n\left(\frac{1}{\sigma}\right)$$

$$= -\frac{1}{2}\left(\frac{z_i-\mu}{\sigma}\right)^2\left(\frac{1}{\sigma}\right) + n\left(\frac{1}{\sigma}\right).$$

$$\therefore\ n = \frac{1}{2}\left(\frac{z_i-\mu}{\sigma}\right)^2$$

$$\therefore\ \sigma^2 =$$

$$0 = \frac{\partial}{\partial\sigma}\left\{+\frac{1}{2}\sum_{i=1}^{n}\frac{(z_i-\mu)^2}{\sigma^2} + n\ln(\sigma)\right\}.$$

$$= \left\{\frac{1}{2}\sum_{i=1}^{n}(z_i-\mu)^2\cdot\frac{\partial(\sigma^{-2})}{\partial\sigma}\right\} + \left\{n\frac{\partial\ln(\sigma)}{\partial\sigma}\right\}$$

$$= \frac{1}{2}\sum_{i=1}^{n}(z_i-\mu)^2\left(\frac{\cancel{2}}{\cancel{2}}\right)\sigma^{-3} + n\left(\frac{1}{\sigma}\right)$$

$$= -\frac{1}{\sigma}\sum_{i=1}^{n}\left(\frac{z_i-\mu}{\sigma}\right)^2 + n\left(\frac{1}{\sigma}\right).$$

$$\therefore\ n = \sum_{i=1}^{n}\frac{(z_i-\mu)^2}{\sigma^2} \qquad \therefore\ \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(z_i-\mu)^2}$$

For line fitting of data corrupted by i.i.d. gaussian of $\mu = 0$ & variance σ^2 the MLE is formulated as

$$\underset{\forall m,c}{\operatorname{argmax}} P(m, c \mid z_1 \cdots z_n) = \underset{\forall m c}{\operatorname{argmax}} P(z_1 \cdots z_n \mid m, c).$$

Now $\quad y_i = mx_i + c + n_i$

where n_i is modelled as

$$P_{N_i}(n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{n_i - 0}{\sigma}\right)^2} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{n_i}{\sigma}\right)^2}$$

$$P(y_i \mid m c) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}\left(\frac{y_i - mx_i - c}{\sigma}\right)^2}$$

$$\Rightarrow \underset{\forall m,c}{\operatorname{argmax}} P(z_1 \cdots z_n \mid m c).$$

$$= \underset{\forall m c}{\operatorname{argmax}} \left[\prod_{i=1}^{N} \frac{e^{-\frac{1}{2}\left(\frac{y_i - mx_i - c}{\sigma}\right)^2}}{\left(\sqrt{2\pi}\sigma\right)^N} \right]$$

$$= \underset{\forall m,c}{\operatorname{argmax}} \ln \left[\prod_{i=1}^{N} \frac{e^{-\frac{1}{2}\left(\frac{y_i - mx_i - c}{\sigma}\right)^2}}{\left(\sqrt{2\pi}\sigma\right)^N} \right]$$

$$= \underset{\forall m,c}{\operatorname{argmax}} \left\{ -N \ln\left(\sqrt{2\pi}\right) - N \ln(\sigma) - \frac{1}{2} \sum_{i=1}^{N} \left(\frac{y_i - mx_i - c}{\sigma} \right)^2 \right\}.$$

Let $\sigma^2 = 1$.

$$\therefore \underset{\forall m c}{\operatorname{argmax}} P(m, c \mid y_1 \cdots y_n) = \underset{\forall m,c}{\operatorname{argmax}} \left\{ -\frac{1}{2} \sum_{i=1}^{N} \left(\frac{y_i - mx_i - c}{1} \right)^2 \right\}.$$

$$= \underset{\forall m c}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - (mx_i + c)\right)^2 = \text{L.S. estimate}.$$

classmate

Date
Page

EM algorithm :-

(a) Expectation E

(b) Maximization M.

$$\hat{\theta} = \underset{\forall \theta}{\text{argmax}}\ P(\underline{\theta}|Y)$$

Using MLE we have.

$$\hat{\theta} = \underset{\forall \theta}{\text{argmax}}\ P(Y|\theta)$$

with a hidden variable z, we have.

$$\hat{\theta} = \underset{\forall \theta}{\text{argmax}}\ P(YZ|\theta)$$

$$= \underset{\forall \theta}{\text{argmax}}\ P(Y|Z\theta) \cdot P(Z|\theta)$$

$$= \underset{\theta}{\text{argmax}} \sum_{z} P(Y|z\theta) \cdot P(z|\theta)$$

$$= \underset{\theta}{\text{argmax}}\ \ln \sum_{z} P(Y|z\theta) \cdot P(z|\theta).$$

Let $L(\theta) = \ln P(Y|Z\theta)$

& $L(\theta_n) = \ln P(Y|\cancel{Z}\theta_n)$.

$$\therefore L(\theta) = \ln \sum_{z} P(Y|z\theta) \cdot P(z|\theta)$$

We need to maximize $L(\theta) - L(\theta_n)$.

$$L(\theta) - L(\theta_n) = \ln\left(\sum_{z} P(Y|z\theta) \cdot P(z|\theta)\right) - \ln\left(P(Y|\cancel{Z}\theta_n)\right)$$

8

$$\sum_z P(z|Y\theta_n) = 1.$$

$$\therefore \ln[P(Y|\theta_n)] = \sum_z P(z|Y\theta_n) \cdot \ln[P(Y|\theta_n)].$$

$$\therefore L(\theta) - L(\theta_n) = \ln\left(\sum_z P(Y|z\theta) \cdot P(z|\theta)\right) - \sum_z P(z|Y\theta_n) \cdot \ln\left(P(Y|\theta_n)\right)$$

$$= \ln\left[\sum_z P(Y|z\theta) \cdot P(z|\theta) \cdot \frac{P(z|Y\theta_n)}{P(z|Y\theta_n)}\right] - \sum_z P(z|Y\theta_n) \cdot \ln\left(P(Y|\theta_n)\right)$$

$$= \ln\left[\sum_z P(z|Y\theta_n) \cdot \left(\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n)}\right)\right] - \sum_z P(z|Y\theta_n) \cdot \ln P(Y|\theta_n)$$

Now, $\ln\left[\sum_z P(z|Y\theta_n) \cdot \frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n)}\right]$ is of the form

$$\ln\left[\sum_i \lambda_i \cdot X_i\right] \quad \text{or} \quad f(\lambda_i, X_i)$$

For a convex function,

$$f(\lambda x_i + (1-\lambda)y_i) \leq \lambda f(x_i) + (1-\lambda) f(y_i) \quad \ldots 0 \leq \lambda \leq 1$$

or $$f\left(\sum_i \lambda_i x_i\right) \leq \sum_i \lambda_i f(x_i) \quad \ldots \sum \lambda_i = 1,\ 0 \leq \lambda_i \leq 1$$

$-$ve $\ln$ is convex.

$$\therefore -\ln\left[\sum_i \lambda_i X_i\right] \leq \sum_i \lambda_i \left[-\ln(X_i)\right]$$

$$\therefore \ln\left[\sum_i \lambda_i X_i\right] \geq \sum_i \lambda_i \left[\ln(X_i)\right]$$

$$\therefore \ln\left[\sum_{z} P(z|Y\theta_n) \cdot \frac{P(Y|z\theta)\cdot P(z|\theta)}{P(z|Y\theta_n)}\right] \geq \sum_{z} P(z|Y\theta_n)\cdot \ln\left[\frac{P(Y|z\theta)\cdot P(z|\theta)}{P(z|Y\theta_n)}\right]$$

$$\therefore L(\theta) - L(\theta_n) \geq \sum_{z} P(z|Y\theta_n)\cdot \ln\left[\frac{P(Y|z\theta)\cdot P(z|\theta)}{P(z|Y\theta_n)}\right]$$

$$- \sum_{z} P(z|Y\theta_n)\cdot \ln\left[P(Y|\theta_n)\right]$$

$$\geq \sum_{z} P(z|Y\theta_n)\left\{\ln\left[\frac{P(Y|z\theta)\cdot P(z|\theta)}{P(z|Y\theta_n)}\right] - \ln\left[P(Y|\theta_n)\right]\right\}$$

$$\geq \sum_{z} P(z|Y\theta_n)\left\{\ln\left[\frac{P(Y|z\theta)\cdot P(z|\theta)}{P(z|Y\theta_n)\cdot P(Y|\theta_n)}\right]\right\} = \Delta(\theta|\theta_n)$$

$$\therefore L(\theta) - L(\theta_n) \geq \Delta(\theta|\theta_n)$$

or $L(\theta) \geq \Delta(\theta|\theta_n) + L(\theta_n) = l(\theta|\theta_n)$

i.e. $L(\theta) \geq l(\theta|\theta_n)$

When $\theta = \theta_n$, $l(\theta|\theta_n)$ becomes $l(\theta_n|\theta_n)$

or

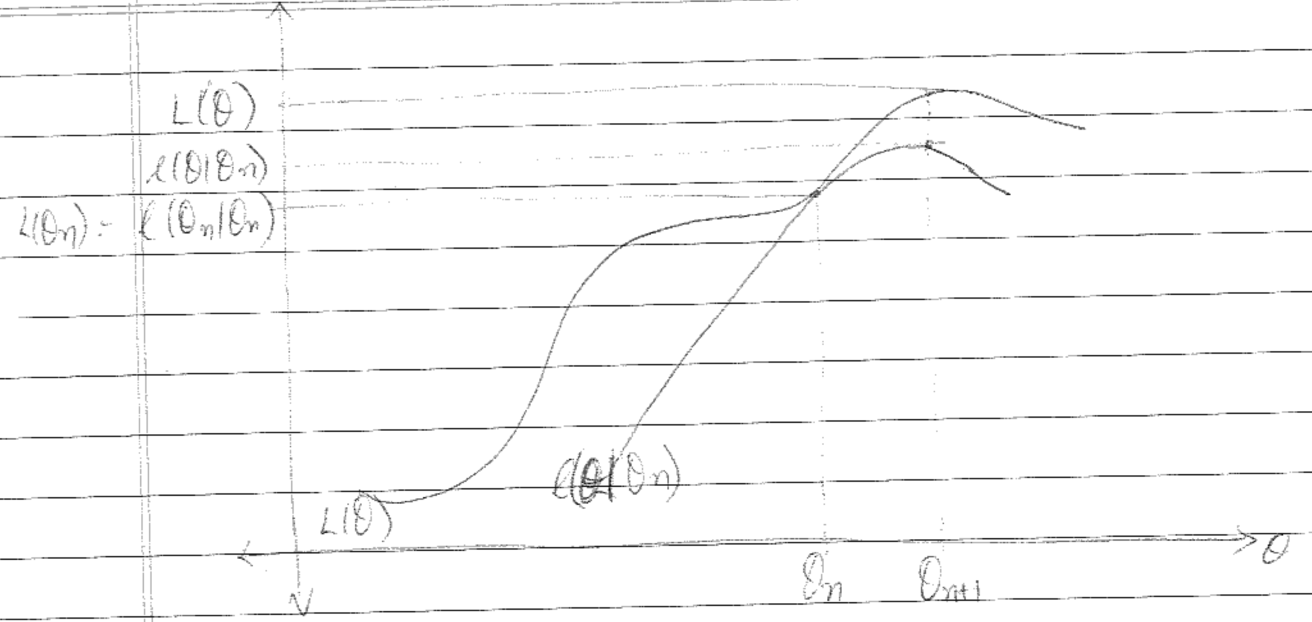
$$l(\theta_n|\theta_n) = \Delta(\theta_n|\theta_n) + L(\theta_n)$$

$$= \sum_{z} P(z|Y\theta_n)\left\{\ln\left[\frac{P(Y|z\theta_n)\cdot P(z|\theta_n)}{P(z|Y\theta_n)\cdot P(Y|\theta_n)}\right]\right\} + L(\theta_n)$$

$$= \sum_{z} P(z|Y\theta_n)\left\{\ln\left[\frac{P(Yz\theta_n)/P(\theta_n)}{P(Yz\theta_n)/P(\theta_n)}\right]\right\} + L(\theta_n)$$

$$= \sum_{z} P(z|Y\theta_n)\cdot \ln(1) + L(\theta_n) = 0 + L(\theta_n)$$

$$\therefore L(\theta_n) = l(\theta_n|\theta_n).$$



$$\theta_{n+1} = \underset{\theta}{\operatorname{argmax}} \{ l(\theta|\theta_n) \}.$$

$$= \underset{\theta}{\operatorname{argmax}} \left\{ \sum_z P(z|Y\theta_n) \cdot \ln \left[\frac{P(Y|z\theta) \cdot P(z|\theta)}{P(z|Y\theta_n) \cdot P(Y|\theta_n)} \right] + \ln P(Y|\theta_n) \right\}$$

Terms with θ_n are constants:

$$\therefore \theta_{n+1} = \underset{\theta}{\operatorname{argmax}} \left\{ \sum_z \left(\ln \left[P(Y|z\theta) \cdot P(z|\theta) \right] \right) \cdot P(z|Y\theta_n) \right\}$$

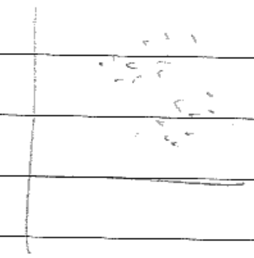
$$= \underset{\theta}{\operatorname{argmax}} \left\{ \sum_z P(z|Y\theta_n) \cdot \ln \left[\left[P(Yz\theta)/P(z\theta) \right] \cdot \left[P(z\theta)/P(\theta) \right] \right] \right.$$

$$= \underset{\theta}{\operatorname{argmax}} \left\{ \sum_z P(z|Y\theta_n) \ln \left(\frac{P(Yz\theta)}{P(\theta)} \right) \right\}$$

$$= \underset{\theta}{\operatorname{argmax}} \left\{ \sum_z P(z|Y\theta_n) \cdot \ln \left(P(Yz|\theta) \right) \right\}$$

$$= \underset{\theta}{\operatorname{argmax}} \left\{ E_{z|Y\theta_n} \cdot \ln \left(P(Yz|\theta) \right) \right\}$$

Eg:



Apply EM to GMM with $\mu = 0$ & var $= \sigma^2$.

$y = m_1 x + c_1$ — line 1

$y = m_2 x + c_2$ — line 2

$r_1(i) = (y_i - (m_1 x_i + c_1))^2$. $r_2(i) = (y_i - (m_2 x_i + c_2))^2$.

$$w_1(i) = \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{r_1(i)-0}{\sigma}\right)^2}}{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{r_1(i)-0}{\sigma}\right)^2} + \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{r_2(i)-0}{\sigma}\right)^2}}$$

Similarly $w_2(i)$. Now $w_1(i) + w_2(i) = 1$.

This is the E step .. by starting with some random values for m_1, c_1 & m_2, c_2.

Now, M step.

$$\hat{m}_1, \hat{c}_1 = \underset{m_1, c_1}{\arg\min} \sum_i w_1(i) (y_i - (m_1 x_i + c_1))^2$$

Similarly $\hat{m}_2, \hat{c}_2$.

Repeat steps E & M till m & c converge.

$M = \frac{1}{N}\sum_i X_i$, $E[X_i] = m$ | $\sigma_v^2 = E[(X_i - m)^2]$
$\sigma_v^2 = \frac{1}{N}\sum_i (X_i - M)^2$

<u>Biased & unbiased estimators.</u>

<u>Unbiased:</u>

Expected value = true value.

$$E[M] = \frac{1}{N} E[X_1 + X_2 + \cdots + X_N]$$

$$= \frac{1}{N}\left(E[X_1] + E[X_2] + \cdots + E[X_N]\right).$$

$$= \frac{1}{N}(m + m + \cdots + m).$$

$$= \frac{Nm}{N} = m = \text{true value.}$$

$\therefore$ M is a unbiased estimator.

Consider

$$E[\sigma_v^2] = E\left[\frac{1}{N}\sum_i (X_i - M)^2\right].$$

$$= \frac{1}{N} E\left[\sum_i (X_i - M)^2\right].$$

$$= \frac{1}{N} E\left[\sum_i (X_i^2 - 2X_iM + M^2)\right]$$

$$= \frac{1}{N}\left[\sum_i E[X_i^2] - 2\sum_i E[X_iM] + \sum_i E[M^2]\right]$$

$$= \frac{1}{N}\left[\sum_i E[X_i^2] - 2E\left[M\sum_i X_i\right] + E\left[M^2\sum_i 1\right]\right]$$

$$= \frac{1}{N}\left[\sum_i \left(E[X_i^2]\right) - 2E[MNM] + E[M^2 \cdot N]\right]$$

$$= \frac{1}{N}\left[\sum_i \left(E[X_i^2]\right) - N E[M^2]\right]$$

Now, $E[(X-\mu)^2] = E[X^2] - (E(X))^2 = \sigma_V^2$.

$$\therefore \quad E[X^2] = \sigma_V^2 + (E[X])^2 .$$

$$\therefore \quad E[\sigma_V^2] = \frac{1}{N}\left[\sum_i \left(\sigma_V^2 + (E(X))^2\right) - N E[M^2]\right].$$

$$= \frac{1}{N}\left[\sum_i \left(\sigma_V^2 + M^2\right) - N E[M^2]\right].$$

Now $M = \frac{1}{N}\sum_i X_i \quad \therefore M^2 = \left(\frac{1}{N}\sum_i X_i\right)^2 = \frac{1}{N^2}\left(\sum_i X_i\right)^2$

$$\therefore E[M^2] = E\left[\frac{1}{N^2}\left(\sum_i X_i\right)^2\right] = \frac{1}{N^2} E\left[\left(\sum_i X_i\right)^2\right].$$

$$\cancel{= \frac{1}{N^2}\left(E\left[\sum\right.\right.} = \frac{1}{N^2} E\left[\left(\sum_i X_i\right)\left(\sum_j X_j\right)\right]$$

~~together~~

$$= \frac{1}{N^2} E\left[\sum_{\substack{i \\ \forall i=j}} X_i X_i + \sum_{\substack{i \\ \forall i \neq j}} X_i X_j\right]$$ independent

$$= \frac{1}{N^2} E\left[\underbrace{\sum_i X_i^2}_{N \text{ terms}} + \underbrace{\sum_{i,j} X_i X_j}_{N^2-N \text{ terms}}\right] = \frac{1}{N^2}\left[\sum_i E[X_i^2] + \sum_i E[X_i]E[X_j]\right]$$

$$= \frac{1}{N^2}\left\{\underbrace{\sum \left(\sigma_V^2 + M^2\right)}_{N \text{ terms}} + \underbrace{\sum M \cdot M}_{N^2-N \text{ term}}\right\}$$

$$\therefore E[M^2] = \frac{1}{N^2}\left\{ N(\sigma_v^2 + M^2) + (N^2 - N)(M^2) \right\}$$

$$= \frac{1}{N^2}\left\{ N\sigma_v^2 + \cancel{NM^2} + N^2M^2 - \cancel{NM^2} \right\}.$$

$$= \frac{1}{N}\sigma_v^2 + M^2$$

$$\therefore E[\sigma_v^2] = \frac{1}{N}\left\{ \sum (M^2 + \sigma_v^2) - N\left(\frac{1}{N}\sigma_v^2 + M^2\right)\right\}.$$

$$= \frac{1}{N}\left\{ N(M^2 + \sigma_v^2) - \sigma_v^2 - NM^2 \right\}.$$

$$= M^2 + \sigma_v^2 - \frac{\sigma_v^2}{N} - M^2$$

$$= \sigma_v^2 - \frac{\sigma_v^2}{N}$$

$$= \frac{N\sigma_v^2 - \sigma_v^2}{N}$$

$$= \frac{N-1}{N}\sigma_v^2 \neq \sigma_v^2$$

$\therefore$ σ_v^2 is an ~~un~~biased estimator.

If we make $\sigma_v^2 = \frac{1}{N-1}\sum_i (X_i - M)^2$ then it is an unbiased estimator.

Miscellaneous 2

Markov Random Fields :-
MRF

Sites S : Let S index discrete set of m sites.

$$S = \{1, 2, \ldots, m\}.$$

Sites on a Lattice are considered spatially regular. A rectangular lattice for a 2D image of size $n \times n$ is denoted by :

$$S = \{(i,j) \mid 1 \le i, j \le n\}.$$

A label is an event that may happen to a site. i.e. a site can take any label from a given set of labels $\mathcal{L}$; $\mathcal{L}_c = [X_l, X_h] \subset R$, $\mathcal{L}_d = \{1, 2, \ldots, M\}$.

$f : S \to \mathcal{L}$ ... mapping site to a label.

Configuration F : $f = \{f_1, f_2, \ldots, f_m\}$ $f_i \in \mathcal{L}$

$\therefore$ # configurations $\mathbb{F} = \underbrace{\mathcal{L} \times \mathcal{L} \times \cdots \times \mathcal{L}}_{m \text{ times}} = \mathcal{L}^m$

Neighbourhood System :-

$$N = \{N_i \mid \forall i \in S\}$$

N_i is the set of sites neighbouring 'i'.

or

$$N_i = \{i' \in S \mid dist(px_i, px_{i'})^2 \le r,\ i \ne i'\}$$

&

$dist(A, B) = \|A, B\|$ ... Euclidean distance

(S, N) is a graph where S are nodes & links between the nodes are determined by N.

P.T.O.

Clique for (S,N) is c such that
c_i = { set of # i sites which are neighbours }

Eg:

$c_1 = \{\{i\} \mid i \in S\}$.

$c_2 = \{\{i, i'\} \mid i' \in N_i,\ i \in S\}$

$c_3 = \{\{i, i', i''\} \mid i' \in N_i, i' \in N_i'', i'' \in N_i, i'' \in N_i', i \in S\}$.

⋮

etc.

& $C = \{C_1 \cup C_2 \cup C_3 \cdots\}$

Random Field :-

$F = \{F_1, F_2, \ldots, F_m\}$ .. a family of random variables defined on S that can take values $f_i \in L$. When F_i takes value f_i, it is denoted by $(F_i = f_i)$ & jointly $(F_1 = f_1, F_2 = f_2, \ldots, F_m = f_m)$.
Probability denoted by $P(F_i = f_i)$ &
$P(F_1 = f_1, F_2 = f_2, \ldots, F_m = f_m)$ ... joint probability.

Markov Random Field :-

F is MRF w.r.t. S if

① $P(f) > 0 \quad \forall f \in \mathbb{F}$

② $P(f_i \mid f_{S-\{i\}}) = P(f_i \mid f_{N_i})$.

where,

$S - \{i\}$ denotes the set S excluding 'i'

& $f_{N_i} = \{f_{i'} \mid i' \in N_i\}$

Gibbs Random Field (GRF) :-

F is said to be GRF w.r.t. S if & only if its configurations obey the Gibbs distribution given by

$$P(f) = Z^{-1} e^{-\frac{1}{T}U(f)}$$

where $Z = \sum_{f \in \mathbb{F}} e^{-\frac{1}{T}U(f)}$ & T is a constant, while

$$U(f) = \sum_{c \in C} V_c(f).$$

$V_c(f)$ is the clique potential of the clique $c \in C$.

GRF is homogeneous if $V_c(f)$ is independent relative to the position of c in S.
GRF is ~~independent~~ isotropic if V_c is independent of the orientation of c.

$$U(f) = \sum_{\{i\} \in C_1} V_1(f_i) + \sum_{\{i,i'\} \in C_2} V_2(f_i, f_{i'}) + \sum_{\{i,i',i''\} \in C_3} V_3(f_i, f_{i'}, f_{i''}) + \cdots$$

2 1 2
1 • 1 V_2
2 1 2

$x(0,0)$	$x(0,1)$	$x(0,2)$	$x(0,3)$
$x(1,0)$	$x(1,1)$	$x(1,2)$	$x(1,3)$
$x(2,0)$	$x(2,1)$	$x(2,2)$	$x(2,3)$
$x(3,0)$	$x(3,1)$	$x(3,2)$	$x(3,3)$

$$U(f) = \sum_{c \in C} V_c(x) = a\left[|x(0,0)| + |x(0,1)| + \cdots + |x(1,0)| + \cdots + |x(3,3)| \right]$$

$$+ \beta\left[|x(1,1) - x(1,2)| + |x(1,1) - x(1,0)| + |x(1,1) - x(2,1)| + |x(1,1) - x(0,1)| \right]$$

$$+ \beta\left[\text{w.r.t. } x(1,2),\ x(2,1),\ x(2,2) \right]$$

$$+ \gamma\left[|x(1,2) - 2x(1,1) + x(0,1)| + |x(0,2) - 2x(1,2) + x(1,1)| + \cdots \right]$$

$\vdots$

when $\alpha_1 = \alpha_2 = \cdots = \alpha_n$ ... homogeneous MRF (i.e. single α, β, γ).
when $\alpha_1 \neq \alpha_2 \neq \cdots \neq \alpha_n$, $\beta_1 \neq \cdots \neq \beta_n$, $\gamma_1 \neq \cdots \neq \gamma_n$ .. non homogeneous MRF

Depth from defocus ... multiple images, space varying blur.
- Raju, Chaudhari.

$g_i = H_i(\sigma_i) \cancel{f} + n_i \qquad , \; i = 1, 2, \dots, p$ # images

MAP-MRF

$$\hat{\sigma}, \hat{f} = \underset{\forall \sigma, f}{\arg\max}\; P(\sigma, f \mid g_1, \dots g_n)$$

$$= \underset{\forall \sigma, f}{\arg\max}\; P(g_1 \cdots g_n \mid \sigma, f) \cdot P(\sigma, f).$$

σ & f are independent

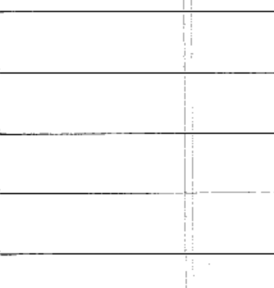
$$\therefore \hat{\sigma}, \hat{f} = \underset{\forall \sigma, f}{\arg\max}\; P(g_1 \cdots g_n \mid \sigma, f) \cdot P(\sigma) \cdot P(f).$$

$$\therefore \underset{\forall \sigma, f}{\arg\max}\; \underbrace{e^{-\sum_{i=1}^{p} \frac{\| g_i - H_i(\sigma_i) f \|^2}{2\sigma_n^2}} \cdot e^{-U(\sigma)} \cdot e^{-U(f)}}$$

$$= \underset{\forall \sigma, f}{\arg\max}\; \ln \Big[\qquad \Big]$$

$$= \underset{\forall \sigma, f}{\arg\max} \left[-\sum_{i=1}^{p} \frac{\| g_i - H_i(\sigma_i) f \|^2}{2\sigma_n^2} - U(\sigma) - U(f) \right]$$

$$= \underset{\forall \sigma, f}{\arg\min} \left[\sum_{i=1}^{p} \| g_i - H_i(\sigma_i) f \|^2 + U(\sigma) + U(f) \right]$$



$$\& \; U(f) = \beta \Big[\sum_i \sum_j \big[f(i,j-1) - f(i,j) \big]^2 + \big[f(i-1,j) - f(i,j) \big]^2 + \big[f(i,j+1) - f(i,j) \big]^2 + \big[f(i+1,j) - f(i,j) \big]^2 \Big]$$

& similarly for $U(\sigma)$ with γ.

$$\arg\max P(\theta|g) = \arg\max\left(\frac{1}{Z^m}\left(\log \prod_i P(g_i|\theta)\right) + P(\theta)\right) \quad \ldots \sigma = 1$$

$$= \arg\max\left(\frac{1}{Z^m}\left(\sum \log P(g_i|\theta) + \log P(\theta)\right)\right)$$

$$\Downarrow \qquad\qquad \Downarrow$$

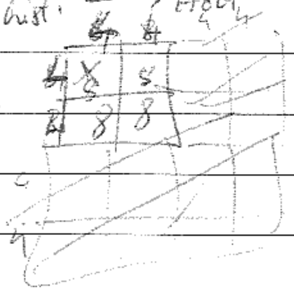
$$\sum_{i \in S} \log P(g_i|\theta_i) + \sum_{i \in S}\sum_{j \in N_i} \phi\, \theta_i \theta_j$$

$$\arg\max\left(\sum_{i \in S} A(\theta_i, g) + \sum\sum I(\theta_i, \theta_j, g)\right) \quad \ldots \text{CRF}$$

SIFT

1. DoG.
2. 26 neighbours. Min-Max detection. (p)
3. Low contrast removal. $0.3 < D(p)$
4. Check edge elimination. $H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix}$

   & $\dfrac{(Tr(H))^2}{\det(H)} > r$
5. ~~Descriptor. feature vector~~ Mag & direction assignment using L. scaling & hist.
6. Keypoint descriptor. ... normalized.
   $4 \times 4 \times 8 = 128$ – descriptor



Conditional Random Field
CRF

→ Discriminative model.
Directly model the conditional distribution over labels.

MRF do not include dependence on the observed data, while CRF captures these dependence between the observed data without resorting to model approximations.

MRF → $P(f)$ → Prior ... $P(f)$
CRF → ~~$P(f|Y)$~~ ~~$P(y|f)$~~ → Posterior ... $P(f|y)$

Let the observed data from the input image be given by $y = \{y_i\}_{i \in S}$ ... y_i is the data from i^{th} site.

Corresponding image labels are given by $f = \{f_i\}_{i \in S}$ or $f_i \in \mathcal{L}$. The random variables f & y are jointly distributed, but in a discriminative framework, then a conditional ~~random field~~ model $P(f|y)$ is constructed from the observations & labels. Here (f, y) is called the CRF ~~of~~ over S if,

$$\boxed{P(f_i | y, f_{S-\{i\}}) = P(f_i | y, f_{N_i})}$$

N_i is the neighbourhood of 'i' just like in MRF.
Ofcourse, $P(f_i|y) > 0 \quad \forall f \in \mathbb{F}$

& $$P(f|y) = \frac{1}{Z} \exp\left(\sum_{i \in S} A_i(f_i, y) + \sum_{i \in S}\sum_{j \in N_i} I_{ij}(f_i, f_j, y)\right).$$

A_i is the association potential
I_{ij} is the interaction potential.
For MRF the same quantity is given by.

$$P(f|y) = \frac{1}{Z_m} \exp\left(\sum_{i \in S} \log p(g_i(y_i)|f_i) + \sum_{i \in S}\sum_{j \in N_i} \beta_m f_i f_j\right).$$

The field for CRF is assumed to be homogeneous as well as isotropic i.e. independent of location & orientation respectively. So, A_i becomes A & I_{ij} becomes I.

$A(f_i, y)$ is a measure of how likely a site 'i' will take the label f_i given the data (here the complete image) y.

If there is a set (φ) of all patches in the image y. & each patch is mapped to a feature vector $\mathbb{R}^l$ then,

$$g : \varphi \rightarrow \mathbb{R}^l$$

If $w_i(y)$ is a patch around location 'i' in image y then $g(w_i(y)) = g_i(y)$ is the ~~feature~~ corresponding feature vector.

$$\text{Now,} \quad A(f_i | y) = \log P'(f_i | g_i(y))$$

i.e.
local class conditional probability at site 'i'.

Log Linear Models & Conditional Random Fields 3

Packages: (1) CRF ++
(2) CRF SGD