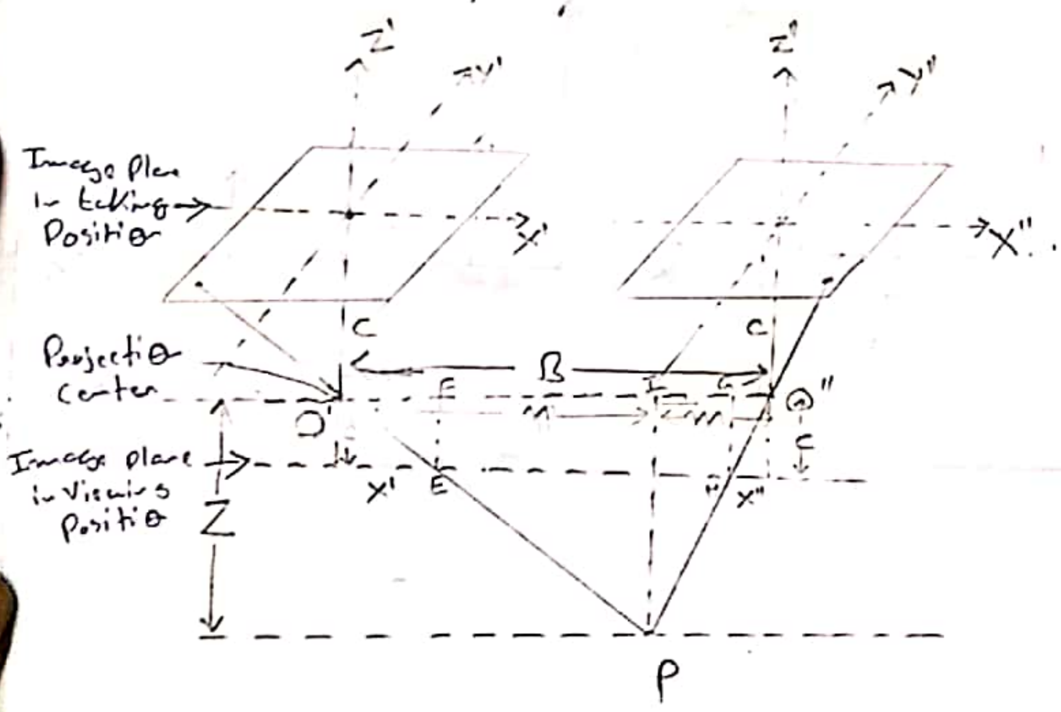
Matching - Part A
(Cross Correlation)

Finding Correspondence   data association

* Stereo Normal Case

→ Special case of Stereo where images have same y-coordinate, same z-coordinate and are shifted in x.



$\Delta O'EF \sim \Delta O'PI$

$$\rightarrow \frac{EF}{O'F} = \frac{PI}{IO'} \Rightarrow \frac{c}{x'} = \frac{Z}{n} \quad —(1)$$

$\Delta O''HG \sim \Delta O''PI$

$$\rightarrow \frac{HG}{GO''} = \frac{PI}{IO''} = \frac{c}{x''} = \frac{Z}{m} \quad —(2)$$

⇒ Rearranging eq ① and ② and adding them we get:-

$$m+n = B = \frac{Z}{c}(x'+x'') \rightarrow P_x$$

(It is also called x disparity or x-parallax)

$$\boxed{\frac{Z}{c} = \frac{B}{P_x}}$$

$$\boxed{Z = c\frac{B}{P_x}}$$

B → (distance between the two camera)

P_x → (x-parallax)

c → (camera constant)

Z → (depth of the point)

⇒ larger the x-parallax closer the point and smaller the x parallax farther is the point.

⇒ Similarly X coordinate of point:

$$m = \frac{Z}{c}x' = x'\frac{B}{P_x}$$

$$\boxed{X = x'\frac{B}{P_x}}$$

⇒ Y-coordinate by mean

$$\frac{Y}{X} = \frac{\frac{y'+y''}{2}}{x'}$$

$$\boxed{Y = \left(\frac{y'+y''}{2}\right)\frac{B}{P_x}}$$

(Average because of error in finding corresponding point)

⇒ If we know the corresponding points, we can perform a 3D reconstruction.

Image matching Problem

→ Identifying and measuring tie points

→ Identifying image regions showing the same object

↑ More often used for automatic procedure

★ Cross correlation

Key assumptions:

→ Image only differ by

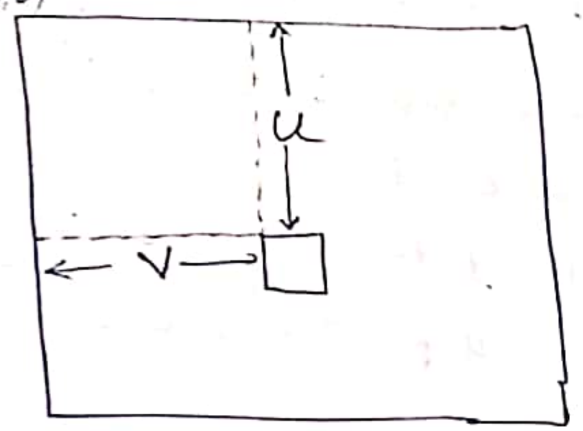
- translation
- brightness
- Contrast

Template matching

→ Find the location of a template within an image

→ Usually size of template << size of image.

(0,0)



⇒ Given an image $g_1(i,j)$ and template $g_2(p,q)$, find offset $[\hat{u}, \hat{v}]$ between g_1 & g_2.

⇒ Geometric transformation

$$\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} i \\ j \end{bmatrix} - \begin{bmatrix} u \\ v \end{bmatrix} \rightarrow \text{Unknown}$$

⇒ Radiometric transformation

$$g_2(p,q) = a + b\, g_1(i,j)$$

→ Unknown

⇒ Typical measures of similarity

① Sum of squared difference

$$SSD = \sum_m \left(g_2(m) - g_1(m) \right)^2$$

② Sum of absolute difference

$$SAD = \sum_m | g_2(m) - g_1(m) |$$

③ Maximum of difference

$$Max = \max_m | g_2(m) - g_1(m) |$$

Scanned by CamScanner

⇒ Problem with the above measure of Similarity is that there is no invariance against changes in brightness and Contrast.

Cross Correlation Function

⇒ Best estimate of the Offset [u, v] is given by maximizing the Cross Correlation coefficient over all possible locations.

$$[\hat{u}, \hat{v}] = \arg\max_{u,v} \rho_{12}(u,v)$$

$$\rho_{12}(u,v) = \frac{\sigma_{g_1 g_2}(u,v)}{\sigma_{g_1}(u,v)\,\sigma_{g_2}}$$

Standard deviation of intensity values of query image g_1 in the area of template g_2 at current offset [u, v]

Standard deviation of intensity value of template g_2

Covariance between the intensity values of g_1 and g_2 in the area of template g_2 at current Position [u, v]..

$\Rightarrow f_{12}(u,v)$ → 1 Maximum Value
→ −1 Minimum Value

Search Strategy

→ Exhaustive Search

→ Coarse to fine Strategy using an Image pyramid.

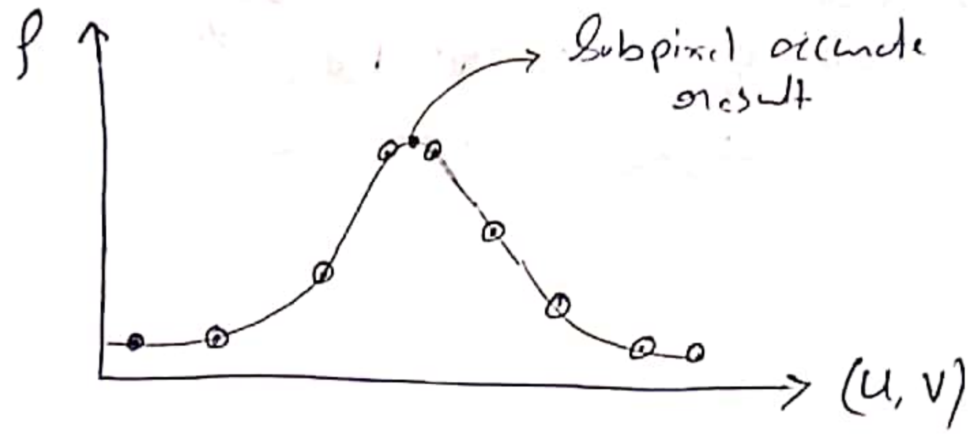
Scanned by CamScanner

(8.b)

Matching - Part B

(Least Square Matching)

★ Subpixel Estimation



Procedure

1) Fit a locally smooth surface through $f_{12}(u,v)$ around the initial position $[\hat{u}, \hat{v}]$.

2) Estimate its local maxima

- Fit a quadratic function around $[\hat{u}, \hat{v}]$

$$f(x) = (x - x^*)^T A (x - x^*) + a$$

$$\left\{ x = \begin{bmatrix} u \\ v \end{bmatrix} \quad x^* \text{ maximum} \right\}$$

- Compute the first derivative

$$\nabla f(x) = \frac{d f(x)}{dx} = 2A(x - x^*)$$

- At maximum: $\nabla f(x^*) = 0$

- Hessian $H_f(x) = 2A$

$\Rightarrow$ So,

$$H_f^{-1} \nabla f(x) = (x - x^*)$$

$$\Rightarrow \boxed{x^* = x - H_f|_x^{-1} \nabla f|_x} \rightarrow \{\text{Current best estimate of } x\}$$

$\Rightarrow$ For an image,

$$x = \begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix}$$

$$\nabla f|_x = \begin{bmatrix} f_i \\ f_j \end{bmatrix}_x \rightarrow \text{Sobel operator}$$

$$H_f|_x = \begin{bmatrix} f_{ii} & f_{ij} \\ f_{ji} & f_{jj} \end{bmatrix}_x \rightarrow \text{Operators for } 2^{nd} \text{ derivative}$$

$$f_i = \frac{\delta f}{\delta u} = \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * f$$

$$f_j = \frac{\delta f}{\delta v} = \frac{1}{8} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * f$$

$$f_{ii} = \frac{\delta^2 f}{\delta u} = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ -2 & -4 & -2 \\ 1 & 2 & 1 \end{bmatrix} * f$$

$$f_{ij} = \frac{\delta^2 f}{\delta u \delta v} = \frac{1}{4} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} * f$$

$$f_{jj} = \frac{\delta^2 f}{\delta v^2} = \frac{1}{4} \begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 2 \\ 1 & -2 & 1 \end{bmatrix} * f$$

★ Least Square Matching (LSM)

⇒ LSM is a generalization of CC.

⇒ LSM supports arbitrary geometric and radiometric transformations.

⇒ LSM requires an initial guess.

LSM in 1D

⇒ Matching two signals, shifted by u

$$y = x - u$$

$$g(x_m) = f(y_m) + n(x_m)$$

(observed signal) (given signal) (noise)

$$m = 1, \dots M \quad \{\text{observation}$$

$$g_m = g(x_m) \quad f_m = f(x_m) \quad n_m = n(x_m)$$

$$g(x_m) - n(x_m) = f(x_m - u)$$

Formulated as a 2D Problem

$$g(x_m, y_m) = f(x_m - u, y_m - v) + n(x_m, y_m)$$

$$m = 1, \dots M \text{ pixel}$$

$$\boxed{g(x_m, y_m) - n(x_m, y_m) = f(x_m - u, y_m - v)}$$

⇒ Objective

→ Find $[\hat{u}, \hat{v}]$, such that $\sum_m n^2(x_m, y_m)$ is minimum

→ Image signal is nonlinear.
→ We need to linearize f

- Linearize f around an initial guess

$$f(x+\Delta u, y+\Delta v) \approx f(x,y) + \left.\frac{\delta f}{\delta u}\right|_{(x,y)} \Delta u + \left.\frac{\delta f}{\delta v}\right|_{(x,y)} \Delta v$$

$$g_m - n_m = f_m + \underbrace{f_{x,m}}_{\text{first derivative of } f_m} \Delta\hat{u} + f_{y,m} \Delta\hat{v}$$

⇒ For all pixels, we can write

$$\underbrace{\begin{bmatrix} g_1 \\ \vdots \\ g_M \end{bmatrix} - \begin{bmatrix} f_1 \\ \vdots \\ f_M \end{bmatrix}}_{\Delta l} - \underbrace{\begin{bmatrix} n_1 \\ \vdots \\ n_M \end{bmatrix}}_{V} = \underbrace{\begin{bmatrix} f_{x,1} & f_{y,1} \\ \vdots & \vdots \\ f_{x,M} & f_{y,M} \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}}_{\Delta x}$$

$$\Delta l - V = A \Delta x$$

⇒ System of Normal Equations

$$(A^T \Sigma_{ll}^{-1} A)\, \Delta\hat{x} = A^T \Sigma_{ll}^{-1} \Delta l$$

$$N \quad \hat{x} = h$$

$$\Sigma_{ll} = \sigma_n^2 I$$

$$(A^T (\sigma_n^2 I)^{-1} A)\, \Delta\hat{x} = A^T (\sigma_n^2 I)^{-1} \Delta l$$

$$(A^T A)\, \Delta\hat{x} = A^T \Delta l$$

$$\boxed{\Delta\hat{x} = (A^T A)^{-1} A^T \Delta l}$$

* Structure of $A^T A$

$$A^T A = \begin{bmatrix} f_{x,1} & \cdots & f_{x,M} \\ f_{y,1} & \cdots & f_{y,M} \end{bmatrix} \begin{bmatrix} f_{x,1} & f_{y,1} \\ \vdots & \vdots \\ f_{x,M} & f_{y,M} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_m f_{x,m}^2 & \sum_m f_{x,m} f_{y,m} \\ \sum_m f_{x,m} f_{y,m} & \sum_m f_{y,m}^2 \end{bmatrix}$$

* Uncertainity of the Shift

⇒ The covariance matrix of the shift is

$$\Sigma_{\hat{x}\hat{x}} = \Sigma_{\Delta\hat{x}\Delta\hat{x}} = \sigma_n^2 N^{-1} = \sigma_n^2 (A^T A)^{-1}$$

$$\Sigma_{\hat{x}\hat{x}} = \sigma_n^2 \begin{bmatrix} \Sigma_m f_{x,m}^2 & \Sigma_m f_{x,m} f_{y,m} \\ \Sigma_m f_{x,m} f_{y,m} & \Sigma_m f_{y,m}^2 \end{bmatrix}^{-1}$$

⇒ Higher the gradient, more precisely we know the shift.

* Covariance of the Gradients

⇒ $$\Sigma_{\nabla f \nabla f} = \begin{bmatrix} \sigma_{f_x}^2 & \sigma_{f_x, f_y} \\ \sigma_{f_x, f_y} & \sigma_{f_y}^2 \end{bmatrix}$$

⇒ Assuming a zero mean $\mu_{f_x} = \mu_{f_y} = 0$ we obtain,

$$\sigma_{f_x}^2 = \frac{1}{M} \sum_{m=1}^{M} f_{x,m}^2 \qquad \sigma_{f_y}^2 = \frac{1}{M} \sum_{m=1}^{M} f_{y,m}^2$$

$$\sigma_{f_x f_y} = \frac{1}{M} \sum_{m=1}^{M} f_{x,m} f_{y,m}$$

* Connection

Covariance matrix of shift and Gradients are related as:

$$\boxed{\Sigma_{\hat{x}\hat{x}} = \frac{\sigma_n^2}{M} \Sigma_{\nabla f \nabla f}^{-1}}$$

⇒ Precision of estimated translation depends on the:

① number M of pixels within used template

② noise variance σ_n^2

③ Covariance matrix of gradients $\Sigma_{\nabla f \nabla f}$

{The steeper the gradient the better the matching}

* Generalization of LSM

⇒ Between two Image I^1 and I^2 :-

- Affine transformation TG
- Linear radiometric transformation.

⇒ Geometric transformation

$$T_G : \begin{bmatrix} p \\ q \end{bmatrix}_m = \begin{bmatrix} a_1 & a_2 \\ a_4 & a_5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_m + \begin{bmatrix} a_3 \\ a_6 \end{bmatrix}$$

⇒ Radiometric transformation

$$T_R : g = a_7 \cdot f + a_8$$

⇒ We may with an identity as initial guess:

$a_{0k} = 1 \quad \forall\, k = 1, 5, 7$

$a_{0k} = 0 \quad \forall\, k = 2, 3, 4, 6, 8$

⇒ After this every thing is similar to only shift case.

* ## Cross Correlation Vs. LSM

→ LSM and Cross Correlation are equivalent, if the
- geometric transformation is a translation.
- radiometric transformation is linear.

→ LSM allows for more general transformations and is more powerful.

→ LSM need an initial guess.

→ CC uses exhaustive search instead of an initial guess (lower dim. space).