Sampling

Continuous Domain → Discrete

JPEG

CO DEC — ENCODER — DECODER

20x compression

BLOCKING ARTIFACTS

RGB

VISUALIZATION

Human eye cones have peak response at 3 wavelenght

R G B

G ≅ light human eye more sensitive to luminance than color

YCbCr

is a color space

Y = luminance

Cb Cr crominance components

HSV

image analysis

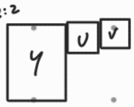
HUE
SATURATION
VALUE

YUV

compression

Down sampled version

4:2:2

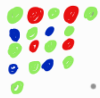


CONVERSION MATRIX

BAYER PATTERN

CCD capture image (array of cells)

50% GREEN

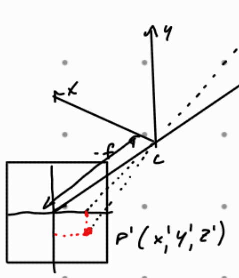


QUANTIZATION

8 bpp is fine

↳ 6 or lower CONTOURING

PIN HOLE CAMERA MODEL

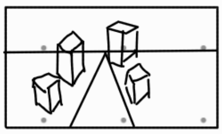


$z' = -f$

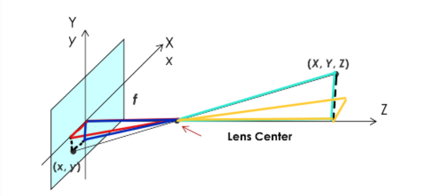
$x' = -f\frac{x}{z}$

$y' = -f\frac{y}{z}$

- FAR OBJECTS APPEAR SMALLER
- PARALLEL lines converge to
  ↳ HORIZON OF THE PLANE
- vertical lines are ⊥ to horizon



Prospective Projection



ORTOGRAPHIC Projection

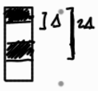
- PARALLEL RAYS

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Distance does not affect dimension

- GOOD Approximation

RESOLVING POWER



Capacity to resolve 2 points

ex.

CCD = 10 Mpixel

img = 4000 (col) x 2500 (row)

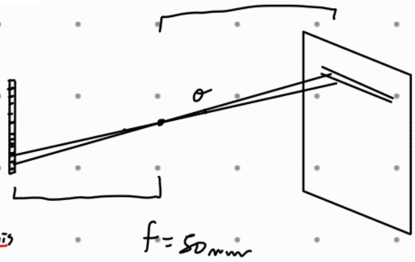
Sensor Dim = 36 mm x 24 mm

SPACING BETWEEN ELEMENTS

$\Delta = \frac{24}{2500} = 0{,}01$ mm VERTICAL

RESOLVING POWER

$R_p = \frac{1}{2\Delta} = 50$ lines/mm



$f = 50$ mm

! assuming

CCD ≫ θ

f ≪ L

TAYLOR

$2\Delta = f \sin(\theta)$

$\simeq f\theta$

$\theta = \frac{\Delta}{25} \simeq 4 \cdot 10^{-4}$ rad

$M = L\theta = 4 \cdot 10^{-4} \cdot 5000 = 2$ mm

OPTICAL FLOW

- ASSUMPTION
  - STATIC OBJECT illumination
  - DISTANCE DON'T CHANGE MUCH
  - EACH POINT shifted in $[x+dx, y+dy]$

} NEVER TRUE BUT GOOD APPROXimation

For hypothesis

APPEARANCE — OBSERVING

$$\psi(x+dx, y+dy, t+dt) = \psi(x,y,t)$$

displacement — time pass — Simplifico

↳ taylor ⇒ $\frac{\partial\psi}{\partial x}dx + \frac{\partial\psi}{\partial y}dy + \frac{\partial\psi}{\partial t}\cancel{dt} = 0$

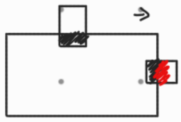
$$\frac{\partial\psi}{\partial x}v_x + \frac{\partial\psi}{\partial y}v_y + \frac{\partial\psi}{\partial t} = 0 \quad \text{OR} \quad \Delta\psi^T\vec{v} + \frac{\partial\psi}{\partial t} = 0$$

$\frac{d}{t} = v$

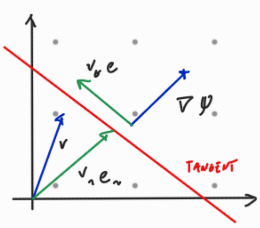
SE QUESTA VARIA → NO VELOCITÀ

APERTURE PROBLEM

- NON SAI SE UN OGGETTO SI MUOVE SE LO GUARDI DALLA PROSPETTIVA SBAGLIATA



EDGES AND OPTICAL FLOW

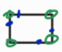


- MOTION ON EDGES DECOMPOSING V
  ↳ $V_n e_n + V_t e_t$

V_t UNDETERMINED

Motion DETECTION

- INTENSITY BASED
- FEATURE BASED



CHANGE DETECTION

BG - SUBTRACTION

Frame differencing

if frame difference > THRESHOLD movement

- illumination is a problem (cumulative difference)

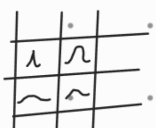
$FD_{k,k-1}(x_1,x_2) = s(x_1,x_2,k) - s(x_1,x_2,k-1)$

apply treshold

↳ k can be different time intervals
⇒ DIFFERENT RESULTS

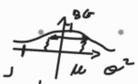
GOOD
- Quick
- Find CONTOURS OF OBJ

BAD
- ONCE IT STOPS NOT ANYMORE DETECTED
- ONLY FEW PIXEL DETECTED
- PROBLEM WITH BG ACQUISITION
- CAMERA MOTION



GAUSSIAN AVERAGE

+ ADAPTIVE
+ FLEXIBLE



current pixel

$\mu_t = \alpha I_t + (1-\alpha)\mu_{t-1}$

MEAN

EACH pixel has 2 values, MODELING A GAUSSIAN to compute if next pixel is inside the distribution

BG - SUBTRACTION

MIXTURE OF GAUSSIAN

Tracking
- Region based
- Contour based
- Feature based
- Template based

REGION BASED TRACKING
Uniform appearence regions
fast (> 30 fps)
use HSV should prevent light changes
- OK indoor
- NOT OUTDOOR

ADAPTIVE BG SUBTRACTION
α PARAM. To weight FD AND BS
"Learning rate" — has no memory
$$B_t = \alpha I_t + (1-\alpha) B_{t-1}$$
(20 frames)
$\alpha = 0{,}05$
GOOD FOR PEDESTRIANS

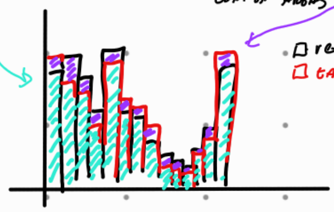
- OTHER METHODS ABSORBS NEW OBJ TO BG
- if you enlarge time does not work well
like periodic movement
PENDULUM CLOCK

=> Solution mix of gaussian
$$P(x_t) = \sum_{i=1}^{k} w_{i,t}\, \eta(x_t, \mu_{i,t}, \Sigma_{i,t})$$
weight — each pixel
Treshold
$$\sum_{i=1}^{B} w_i > T$$

Histogram based approach
- Bg subtraction
- for each moving obj compute histogram
- evaluate histogram O^t and compare to O^r
[QUANTIZED]

□ intersection
$$\cap(O^r_i, O^t_i) = \sum_{m=1}^{U} \min\{O^r_{i,m}, O^t_{i,m}\}$$
"SCORE" HIGHER = BETTER, COMMON AREAS

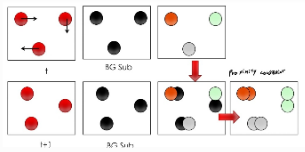
□ $SSD(O^t_i, O^r_i) = \sum_{1}^{U} (O^r_{i,n} - O^t_{i,n})^2$
"ERROR" LOWER = BETTER



□ reference t
□ TARGET t+1

TARGET ASSOCIATION
- ONCE YOU FIND MOVING OBJECT HOW DO YOU KEEP TRACK?

Proximity basis
How?
- OVERLAPPING BLOBS — SCALE
- CENTROID DISTANCE — SCALE
- BBOX OVERLAP — Prospective
- MIN DISTANCE BBOX
- BBOX TO CENTROID DISTANCE

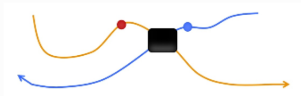


How ABOUT SPLITTING? MERGING?

COMPLEXITY -> DISCRETIZING NEIGHBOURS COLORS
- BETTER FOR LOCAL VARIATION OF LUMINATION
- BETTER IN CHANGE OF DIMENSION

↳ if pdf shifts slower than quantization
PERFECT MATCH

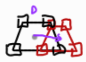
OCCLUSION



ONE OF THE BIGGEST Problem in tracking
HOW TO DEAL WITH THAT?

FEATURE BASED TRACKING
$A = \{A(0), A(1) \dots A(J-1)\}$ images
$m_i(x_i, y_i)$ Position of feature

OBJECTIVE
DETERMINE Displacement vector
$d_i = (dx_i, dy_i)$
↳ $m_{i+1}(x_{i+1}, y_{i+1}) = m_i + d_i$



- OCCHIO ALLE OMBRE

BLOBS EXTRACTION

HOW TO TRACK FEATURES
$$A_i(D_m - d) = A_{i+1}(m)$$
Deformation matrix — DISPLACEMENT — Position of feature Point

Window

CENTRAL HIGHEST Priority

EQUALITY
DOES NOT HOLD (noise)
- ASSUMPTION -> SMALL TRANSITION

DISSIMILARITY MEASURE — Sum if discrete
$$\varepsilon = \iint_W [A_i(m-d) - A_{i+1}(m)]^2 \omega(m)\, dm$$
WEIGHTED FUNCTION
WHEN IT GETS TOO HIGH ABBANDON POINT
WHAT IF feature point disappear

Minimize residual
$$e = 2\iint_W [A_i(m-d) - A_{i+1}(m)]\, g(m)\, w(m)\, dm$$
$$g(m) = \begin{bmatrix} \dfrac{\partial (A_i(m) - A_{i+1}(m))}{\partial x} \\ \dfrac{\partial (A_i(m) - A_{i+1}(m))}{\partial y} \end{bmatrix}$$
solution for displacement vector
=> $Zd = e$

GOOD FEATURES
GOOD FEATURES -> GOOD TRACKING
$$Z = \begin{bmatrix} \sum_w J_x^2 & \sum_w J_x J_y \\ \sum_w J_y J_x & \sum_w J_y^2 \end{bmatrix} > \text{Treshold}$$
CLASSIFIED AS FEATURES
J_x J_y are gradients evaluated on x and y within W window ($n \times n$)

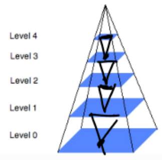
GOOD if smallest eigenvalue of Z > treshold
if ONLY 1 λ > treshold
NOT GOOD

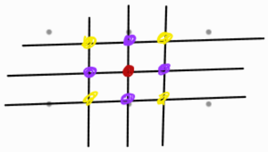
LUCAS-KANADE
OPTIMIZATION FOR TRACKING
$u = [u_x, u_y]$ $v = [v_x, v_y]$
Frame I — J
GOAL ② $v = u + d$

d is the vector that minimizes
$$\varepsilon(d) = \varepsilon(d_x, d_y) = \sum_{x=u_x-w_x}^{u_x+w_x} \sum_{y=u_y-w_y}^{u_y+w_y} (I(x,y) - J(x+d_x, y+d_y))^2$$
2D because of aperture problem
d — Displacement vector

- ACCURACY — Lowest LEVEL
- ROBUSTNESS — Highest level

PYRAMIDAL IMPLEMENTATION
- 0 is original image
- h image at lowest resolution
- L-th level is a linear comb. of elements in previous level





$$I^L(x,y) = \frac{1}{4} I^{L-1}\boxed{(2x, 2y)} +$$
$$\frac{1}{8}\left(I^{L-1}(2x-1, 2y) + I^{L-1}(2x+1, 2y) + I^{L-1}(2x, 2y-1) + I^{L-1}(2x, 2y+1)\right) +$$
$$\frac{1}{16}\left(I^{L-1}(2x-1, 2y-1) + I^{L-1}(2x+1, 2y+1) + I^{L-1}(2x-1, 2y+1) + I^{L-1}(2x+1, 2y+1)\right)$$

PYRAMID
$$\varepsilon^L(d^L) = \varepsilon^L(d_x^L, d_y^L) = \sum_{x=u_x^L-w_x}^{u_x^L+w_x} \sum_{y=u_y^L-w_y}^{u_y^L+w_y} \left(I^L(x,y) - J^L(x+g_x^L+d_x, y+g_y^L+d_y^L)\right)^2$$
GUESS — Second image — Small

Propagated at higher level
$$g^{L-1} = 2(g^L + d^L)$$
$$\hookrightarrow d = \sum_{L=0}^{L_m} 2^L d^L$$

BAYESIAN TRACKING
Previous state — ex rumore
$$x_k = f_k(x_{k-1}, w_{k-1})$$
$Sk \sim 1_k$

ONLINE METHOD
initial PDF $P(x_0|z_0)$ Given
GOAL $P(x_k|z_k)$

Kalman filter
- MEASUREMENT
- ERROR
- Derive state of the

START
$z_1, \sigma^2_{z_1}$
$\hat{x}_1 = z_1$
$\sigma_1^2 = \sigma^2_{z_1}$

STEPS
ADD MEASUREMENT
COMBINE
ITERATE

NEXT STATE

PROCESS NOISE

- Prediction
- UPDATE

System from the measurement

MEASUREMENT → STATE VECTOR

$$z_k = h_k(x_k, v_k)$$

MEASURE NOISE

$$x_k = f_x(x_{k-1}, w_{k-1}) \to P(x_k \mid x_{k-1})$$

$$z_k = h_k(x_k, v_k) \to P(z_k \mid x_k)$$

$$z_1, \sigma_{z1}^2 \qquad \hat{x}_1 = z_1 \qquad \sigma_1^2 = \sigma_{z1}^2$$

$$\to \hat{x}_2 = \hat{x}_1 + K_2(z_2 - \hat{x}_1) \qquad K = \frac{\sigma^2}{\sigma_{z1}^2 + \sigma_{z2}^2}$$

$$\to \frac{1}{\sigma_2^2} = \frac{1}{\sigma_{z1}^2} + \frac{1}{\sigma_{z2}^2}$$

WEIGHTED AVERAGE

UPDATE

$$P(x_k \mid z_k) = \frac{P(z_k \mid x_k)\, P(x_k \mid z_{k-1})}{P(z_k \mid z_{k-1})}$$

$$P(z_k \mid z_{k-1}) = \int P(z_k \mid x_k)\, P(x_k \mid z_{k-1})\, dx_k$$

Propagated forward

$$\underbrace{P(x_k \mid z_{k-1})}_{\text{Prior}} = \int \underbrace{P(x_k \mid x_{k-1})}_{\text{MODEL}}\, \underbrace{P(x_{k-1} \mid z_{k-1})}_{\text{Posterior}}\, dx_{k-1}$$

NUTSHELL

- Difference not all noise
- MOTION OCCURS
- MOTION MODEL (pos, velocity)
  while
  - Predict state and uncertainty
  - correct with new measurement

PRACTICE KF

Efficient solution to least squares method

ASSUMPTION

w_k v_k normal distributions

$$\underset{\text{PROCESS}}{P(w)} = N(0, Q) \quad \text{noises}$$

$$\underset{\text{MEASUREMENT}}{P(v)} = N(0, R)$$

STATE

$$\underbrace{x_k}_{\text{CURRENT STATE}} = \underbrace{A_k}_{\text{STATE TRANSITION MATRIX}} \underbrace{x_{k-1}}_{\text{PREVIOUS}} + \underbrace{B_k u}_{\text{CONTROL INPUT}} + \underbrace{w_{k-1}}_{\text{NOISE}}$$

$$\underbrace{z_k}_{\text{MEASUREMENT}} = \underbrace{H_k}_{\text{MEASURE MATRIX}} x + \underbrace{v_k}_{\text{NOISE}}$$

GEOMETRY

AFFINE TRANSFORMATION

$$\begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{SHIFT}} \underbrace{\begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{SCALE}} \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{ROTATION}} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

EPIPOLAR CONSTRAINT

if 2 cameras are parallel and allied
a match will happen on an horizontal
line

EPIPOLAR PLANE

Plane that connects the focus of both cameras

3D POSITION ESTIMATION

(parallel and aligned system)

$${}^wP_z = \frac{\overset{\text{FOCUS}}{f}\, \overset{\text{TRANSLATION BETWEEN 2 CAMERAS}}{b}}{{}^Lp_x - {}^Rp_x}$$

LEAST SQUARE METHOD

- To solve over determined systems

Minimises

$$y = f(x) = Ax + B \quad \to \quad \varepsilon = \sum_{j=i}^{N} (Ax_j + B - y_j)^2$$

- Partial derivatives for each unknown

$$\frac{\partial \varepsilon}{\partial A}, \quad \frac{\partial \varepsilon}{\partial B}$$

CAMERA MODEL

$$[p_r, p_c]^T = {}^I p = \underbrace{{}^I_F S}_{\text{mm} \to \text{pixel}}\, \underbrace{{}^F_C \Pi(f)}_{\text{CALIBRATION}}\, \underbrace{{}^C_W TR(\alpha, \beta, \gamma, t_x, t_y, t_z)}_{\text{ROTOTRANSLATION}}\, {}^w P$$

CAPYBARA!!!

!!!!!!!!!!!