2007 IIB 4F12 Prof R Cipolla
cipolla@eng.cam.ac.uk

4F12 Engineering Tripos Part II Computer Vision

Q1

(a) – Remove effect of image noise before differentiation in edge detection
– blur image \ remove high frequency spatial details
– scale-space representation of image
– smoothing before sampling to avoid aliasing

(b) $$S(x,y) = \sum_{i=-n}^{n} \sum_{j=-n}^{n} I(x-i,\, y-j)\, G_\sigma(i,j)$$

(c) $$S(x,y) = I(x,y) * G_\sigma(x,y) \qquad (A)$$

$$S(x,y) = I(x,y) * g_\sigma(x) * g_\sigma(y) \quad (B) \quad \text{where } g_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\cdots}$$

since $G_\sigma(x,y)$ is seperable.

$G_\sigma(x,y)$ kernel is $(2n+1)^2$ in size

$g_\sigma(x)$ kernel is $(2n+1)$ in size.

Scheme (A) requires $(2n+1)^2$ multiplications per single pixel

Scheme (B) requires $2(2n+1)$ multiplications per single pixel

Saving $\frac{(2n+1)}{2}$ per pixel.

1(d)

Sample $g_\sigma(x)$ at $x = \{-n \ldots -1, 0, 1, \ldots n\}$
and ignore values below $\frac{1}{1000}$ peak value to determine n

Filter coefficients are values $g_\sigma(x)$ since area $\simeq g_\sigma(x) \times 1$

Find n by $e^{-\frac{(n+1)^2}{2\sigma^2}} < \frac{1}{1000}$

$\therefore n > 3.7\sigma - 1$ (find nearest integer to)

For $\sigma = 1$, $(2n+1) = 7$

	-3	-2	-1	n=0	1	2	3
$g_\sigma(x)$	0.004	0.054	0.242	0.399	0.242	0.054	0.004

(e) Isotropic band-pass filter for finding features at a certain scale, such as blobs

Difference of gaussians of diff. scale or a Laplacian of Gaussian produces a ~~band~~ of spatial frequencies (band) only.

$\nabla G_\sigma^{(x,y)} \simeq G_{k\sigma}^{(x,y)} - G_\sigma^{(x,y)}$

2)

(a) Valid if – no non-linear lens distortion
– "pin-hole" camera model for central perspective projection (rays pass through a single pt)

(b)

$$P = K[R \mid t]$$

where $K = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$

$\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$ = principal pt

f = focal length in mm

$\left.\begin{matrix} k_u \\ k_v \end{matrix}\right\}$ scale factors, pixels per mm

$R = [\text{orientation}]$

$\underline{t}$ = position of camera rel. to world co-ordinate system

– Decompose 3×3 matrix KR by QR decomposition
– recover $\underline{t}$ from $[K\underline{t}]$
– cannot decompose fk_u, fk_v etc.

(c) Bookwork (lecture notes)

$$u_K = \frac{p_{11}X_K + p_{12}Y_K + p_{13}Z_K + p_{14}}{p_{31}X_K + p_{32}Y_K + p_{33}Z_K + p_{34}}$$

$$v_K = \frac{p_{21}X_K + p_{22}Y_K + p_{23}Z_K + p_{24}}{p_{31}X_K + p_{32}Y_K + p_{33}Z_K + p_{34}}$$

2 linear equations in unknown p_{ij}

(a) $(p_{11}X_K + p_{12}Y_K + p_{13}Z_K + p_{14}) - u_K(p_{31}X_K + p_{32}Y_K + p_{33}Z_K + p_{34}) =$

(b) and $(p_{21}X_K + p_{22}Y_K + p_{23}Z_K + p_{24}) - v_K(p_{31}X_K + p_{32}Y_K + p_{33}Z_K + p_{34})$

We can rewrite as

$$\underbrace{\underline{A}}_{2n\times 12 \text{ elements of coefficients in (a) and (b)}}\ \underbrace{\underline{p}}_{12\times 1} = \underline{0}$$

Recover approx $\underline{p}$ by SVD [~~largest svector~~ singular vector corresponding to smallest singular value] or eigenvector corresponding to smallest eigenvalue of $A^T A$.

Need $n \geq 6$ pts with independent $\{X_K, Y_K, Z_K\}$ i.e. not co-planar

Perform nonlinear optimization to reduce non-linear effects and noise. Min reprojection

$$\text{Find } \min_{\underline{p}} \sum_{k=1}^{k=n} \left[(u_K - \hat{u}_K)^2 + (v_K - \hat{v}_K)^2 \right]$$

(d)

Consider pts // to X-axis

$$VP_x \simeq \begin{pmatrix} p_{11} \\ p_{21} \\ p_{31} \end{pmatrix} = \begin{pmatrix} p_{11}/p_{31} \\ p_{21}/p_{31} \end{pmatrix}$$

cartesian pixel-coordinates

Consider pts // to Y-axis

$$VP_y \simeq \begin{pmatrix} p_{12} \\ p_{22} \\ p_{32} \end{pmatrix} = \begin{pmatrix} p_{12}/p_{32} \\ p_{22}/p_{32} \end{pmatrix}$$

Consider pts // to Z-axis

$$VP_z \simeq \begin{pmatrix} p_{13} \\ p_{23} \\ p_{33} \end{pmatrix} = \begin{pmatrix} p_{13}/p_{33} \\ p_{23}/p_{33} \end{pmatrix}$$

$$\therefore KR = \begin{bmatrix} VP_x & VP_y & VP_z \end{bmatrix}$$

$$R_1 \simeq K^{-1} VP_x$$
$$R_2 \simeq K^{-1} VP_y$$
$$R_3 \simeq K^{-1} VP_z$$

where each column is normalized to $\underline{1}$

Q3.

(a) In view 1

$$\underset{3\times1}{\underline{w}} = \underset{3\times3}{K}[\underset{3\times3}{I}|\underset{3\times1}{0}]\underset{4\times1}{\underline{X}}$$

and in view 2 $\underline{w}' = K[R|0]\underline{X}$ since rotation only

$$\underline{\underline{w}' = KRK^{-1}\underline{w}} = T\underline{w}$$ where $T = 3\times3$ transformation matrix

$$\begin{bmatrix} u'_k \\ v'_k \\ w'_k \end{bmatrix} = \begin{bmatrix} & t_{ij} & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} u_k \\ v_k \\ w_k \end{bmatrix}$$

where (u'_k, v'_k, w'_k) is correspondence of (u_k, v_k, w_k)

Given a set of correspondences

(b) (i) $$u'_k = \frac{t_{11}u_k + t_{12}v_k + t_{13}w_k}{t_{31}u_k + t_{32}v_k + t_{33}w_k}$$

$$v'_k = \frac{t_{21}u_k + t_{22}v_k + t_{23}w_k}{t_{31}u_k + t_{32}v_k + t_{33}w_k}.$$

Rewrite as $\underset{2n\times9}{A}\,\underset{9\times1}{\underline{t}} = \underline{0}$ and solve by SVD / least-squares

Q4

(a) In stereo vision

$$\underline{X}_c' = R\underline{X}_c + \underline{t}$$

and $\underline{p}' = K'^{-1}\underline{w}'$ and $\underline{p} = K^{-1}\underline{w}$

ray

homogeneous vector of pixels

Epipolar constraint is that

(A) $\underline{X}_c'$, $\underline{X}_c$ and $\underline{t}$ are coplanar

$$\therefore (\underline{X}_c' \times t) . \underline{X}_c' = 0 \quad \text{or} \quad X_c' . (t_\times X_c') = 0$$

or

$$X_c' . (t_\times R X_c) = 0$$

$$\underline{X}_c'^T \; TR \; \underline{X}_c = 0$$

(B) In terms of rays, p and p'

$$\underline{p}' . (TR)\underline{p} = 0$$

or $\underline{p}'^T \; TR \; \underline{p} = 0$

(c) and pixels $\underline{w}$ and $\underline{w}'$

$$w'^T K'^{-T} \; TR \; K^{-1} w = 0$$

$$\underline{w}'^T F \underline{w} = 0$$

3x3

where $F = K^{-T} T . R . K^{-1}$

4(b). A pt on an epipolar line in right view corresponding to $\underline{\omega}$ in left lies on a line $\underline{L}'$ such that

$$\underline{\omega}' \cdot \underline{L}' = 0$$

where line is represented by $\begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$ and $\underline{\omega}' =$

But $\underline{\omega}' \cdot F\underline{\omega} = 0$ and hence $\underline{L}' = F\underline{\omega}$.

(c). Each pair of correspondences gives 1 equation in unknown F element f_{ij}

$\therefore$ $A\underline{f} = 0$ and solve by least-squares

Matrix F has max rank 2 $(\det F = 0)$. Enforce by finding nearest matrix F^* to estimate $\hat{\underline{E}}$ by minimizing Frobenius norm.

(d) $\underline{F}$ can be decomposed into $\underline{e}'_x [M]_x$ where $\underline{e}'$ is an epipole in right view

or if internal cameras are known it can be decomposed into $\underline{t}$ and $\underline{R}$

$P = K[I|0]$ and $\underline{P}' = K'[R|t]$

$\therefore$ need to know $\underline{K}$ and $\underline{K}'$ of cameras.

Q5.(a) Stereo (book work)

- rectification
- cross-correlation
- DP algorithms
- noise figural continuity; uniqueness + ordering

5(b). SIFT (book work).

- interest pts + scale by Laplacian of gaussian over scale
- gradients
- histograms
- vectors + Kd matching.

6