PROF R CIPOLLA

PART IIB 2011 4F12: COMPUTER VISION AND ROBOTICS

4F12 solutions (2011)

Q1. (a) - Remove high frequency addition noise which is amplified by differentiation. Reduce high-spatial frequencies to select image targets of scale of interest.

- Gaussian kernel $G_\sigma(x,y) = \frac{1}{(\sqrt{2\pi}\sigma)^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$

- $S(x,y) = I(x,y) * G_\sigma(x,y) = I(x,y) * G_\sigma(x) * G_\sigma(y)$

where $g_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$

(b) $$S(x,y) = \sum_{-n}^{-n} \sum_{-n}^{n} I(x-u, y-v)\, g_\sigma(u)\, g_\sigma(v)$$

for $\sigma = 1$, $n = 3$ (Kernel is of size 7, $2n+1$).

(c) Image pyramid: $L(x,y,t) = G(x,y,t) * I(x,y)$

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

Sample such: $G(\sigma_{i+1}) = G(\sigma_i) * G(\sigma_K)$

$$\sigma_{i+1} = 2^{\frac{i}{s}} \sigma_0$$

In each octave s smoothed images with last image blurred by $2\sigma_0$ and then subsampled to generate new octave

1 (d) Band-pass filtering = Laplacian of gaussian $= \nabla^2 I * G(x,y)$

$$\approx [G(k\sigma) - G(\sigma)] * I(x,y)$$

Approximated by differences of gaussian blurred images in pyramid.

Search for max/min in scale and image position for blob-like features

(e) Find, position + scale. Gradient histogram. Normalize to SIFT descriptor
(bookwork)

Good understanding and well-attempted by most candidates. Localisation of blob in scale-space was only component poorly understood.

2(a) Pin-hole camera, <u>central perspective</u>, no non-linear distortion

$$\begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

perspective projection

Allow $\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \longrightarrow \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$ where $W=0$ for pts at ∞.

(b)

$$u = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

$$v = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

Ray is rep by 2 equations of planes which are NOT parallel

$$0 = X(p_{11} - p_{31}u) + Y(p_{12} - p_{32}u) + Z(p_{13} - p_{33}u) + (p_{14} - p_{34})$$

$$0 = X(p_{21} - p_{31}v) + Y(p_{22} - p_{31}v) + Z(p_{23} - p_{33}v) + p_{24} - p_{34}$$

2(c).

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 000 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = K[R|T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

(d). $N > 6$, non-coplanar pts, span ΔZ and field of view

Features are easy to localize + extract + match.

Good attempts by most candidates. Many struggled with the algebraic equations of a ray.

Q3(a) (i) View plane:

(ii) Rotation about axis

$$\underset{z=0 \text{ WLOG.}}{\underline{u} = K[I|0]\underline{X}} \qquad \underline{u}' = K'[R|T]\begin{bmatrix} \cancel{\underline{u}} \\ 0 \\ 1 \end{bmatrix} \quad z=0$$

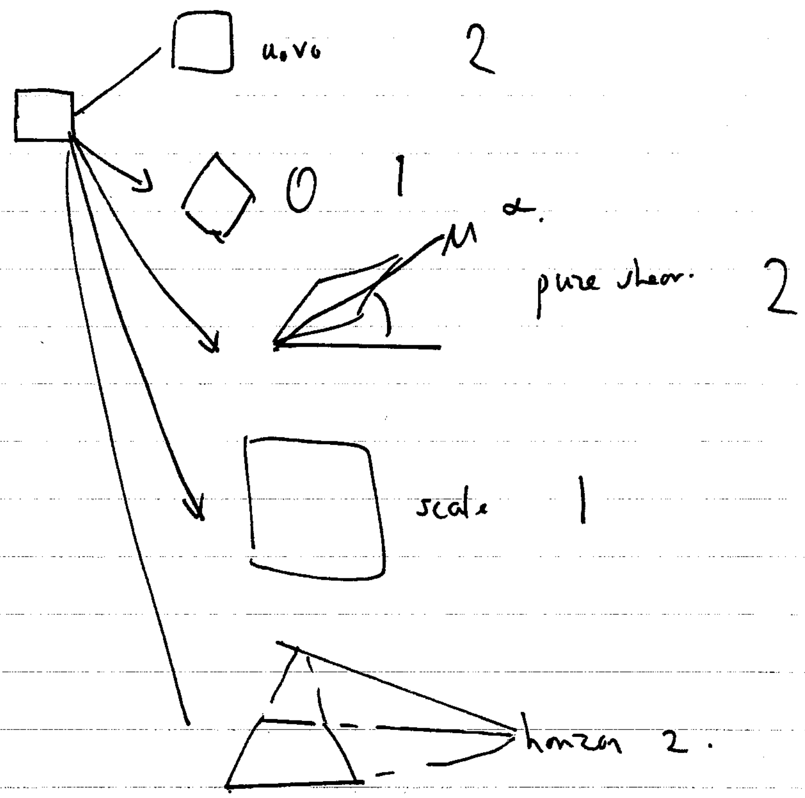
$$\underline{u} = K[I|0]\underline{X} \qquad \underline{u}' = K'[R|0]\underline{X}$$

$$\underline{u}' = \underbrace{K'[R]K^{-1}}_{\underline{T}}\underline{u} \quad \text{rotation}$$

$$\text{or} \quad \underline{u}' = K'[r_{11}\ r_{12}\ \underline{t}][\ \underline{u}\ ]$$

plane – plane

(b) 8 dof



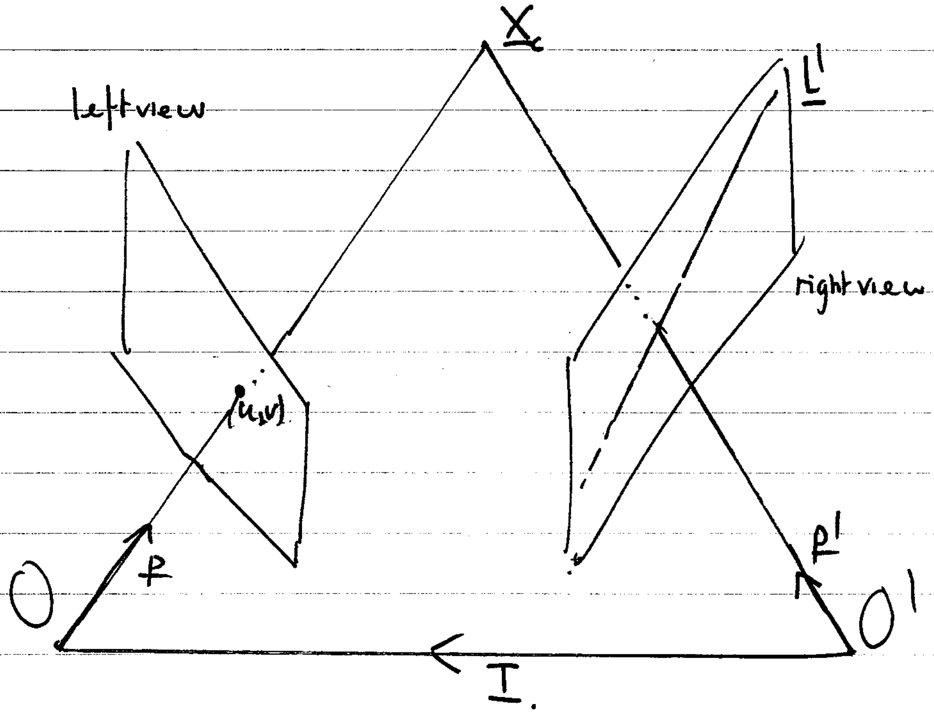
Under weak-perspective

$$\underline{T} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \longrightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \quad 6 \text{ dof.}$$

Part (a) - viewing conditions for homography - was poorly answered. Rest was well-attempted.

4(a)

$$\underline{X}'_{c} = R\underline{X}_c + \underline{T}$$



$\underline{\omega} = K\underline{p}$ and $\underline{\omega}' = K'\underline{p}'$

$\underline{X}' = R\underline{X} + T$ $\quad \underline{X}' \cdot (\underline{T}_x R \underline{X}) = 0$ by coplanarity of $[\underline{X}', \underline{X}, \underline{T}]$

$$\underline{p}'^T \; \underline{T}_x R \; \underline{p} = 0.$$

$$\text{or } \underline{\omega}'^T K'^{-T} T_x R K^{-1} \underline{\omega} = 0$$

$$\underline{\omega}'^T F \underline{\omega} = 0 \text{ where } \underline{F = K'^{-T} T_x R K^{-1}}$$

$$\therefore \underline{\underline{\underline{L}' = K'^{-T} \underline{T}_x \underline{R} K^{-1} \underline{\omega}}}$$

4(b). Reduce 2D search to 1D. Use appearance/feature descriptors

(c) Triangulate rays by solving least-squares
4 equations in 3 unknowns.

$$\underline{u} = K[I|0]\,\underline{X}$$

$$\underline{u}' = K'[R|T]\,\underline{X}$$

(d) SFM (book work).

This question tested the candidates understanding of stereo vision. Very good answers to most components except the equation of an epipolar line.