FÖRSÄTTSBLAD TENTAMEN/ EXAMINATION COVER

Jag intygar att mobiltelefon och annan otillåten elektronisk utrustning är avstängd och förvaras på anvisad plats. / I hereby confirm that mobile phones and other unauthorized electronic equipment is shut off and placed according to instructions

MARKERA MED "X"/
MARK WITH "X" ☐

IFYLLES AV STUDENT OCH TENTAMENSVAKT/
TO BE FILLED IN BY THE STUDENT AND THE INVIGILATOR:

KURSKOD / COURSE CODE: DD2423	EFTERNAMN / FAMILY NAME: BANSAL
KURSNAMN / COURSE NAME: IMAGE ANALYSIS AND COMPUTER VISION	FÖRNAMN / FIRST NAME: HRITIK
PROVKOD / TEST CODE:	NAMNTECKNING / YOUR SIGNATURE: HBansal
TENTAMENSDATUM / EXAMINATION DATE Y/Y/Y/Y - M/M - D/D: 2018 - 12 - 18	PERSONNUMMER / PERSONAL NUMBER Y/Y/M/M/D/D: 981110 - T014

PROGRAMKOD / PROGRAM CODE:	INLÄMNINGSTID / TIME SUBMITTED:	SIGNATUR TENTAMENSVAKT / SIGNATURE INVIGILATOR:	ANTAL SIDOR / NO OF PAGES:
EXCHANGE IIT DELHI INDIA	5:15 pm		17

MARKERA BEHANDLADE UPPGIFTER MED "X "OCH EJ BEHANDLADE UPPGIFTER MED "-" /
MARK WITH "X" PROBLEMS SOLVED. MARK WITH "-" PROBLEMS NOT ATTEMPTED

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

IFYLLES AV INSTITUTIONEN / TO BE FILLED IN BY THE DEPARTMENT:

BEDÖMNING / ASSESSMENT

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
8.5	7.5	5.5	6	6	10														

BONUSPOÄNG/ BONUS POINTS: 3,0

SLUTSUMMA / FINAL POINTS: 48,5

BETYG/ GRADE: A

0406186511

Godkänns av examinator / approved by Examiner: Mårten Björkman

BANSAL, HRITIK | 981110-T014 | | 1 | 1

Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

Question 1

(a) For a camera centered ~~by~~ at $c = (0, 0, 0)^T$. The projection matrix is written as :- $\boxed{\text{In Terms of Homogeneous Coordinates}}$

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

$$\begin{pmatrix} cx \\ cy \\ c \end{pmatrix} = \underbrace{\begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{P:\ \text{Projection Matrix}} \begin{pmatrix} KX \\ KY \\ KZ \\ K \end{pmatrix} \qquad f: \text{focal length}$$

(x, y, z) are image coordinates, (X, Y, Z) are world coordinates.
If the camera is now centered at $C = (1, 0, 1)^T$, this is equivalent to translation.

$$x = \frac{f(X-1)}{Z-1} \qquad y = \frac{fY}{Z-1}$$

$$\begin{pmatrix} cx \\ cy \\ c \end{pmatrix} = \underbrace{\begin{pmatrix} 5 & 0 & 0 & -1\times5 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}}_{\text{Required Projection Matrix}} \begin{pmatrix} KX \\ KY \\ KZ \\ K \end{pmatrix}$$

$$\boxed{P' = \begin{pmatrix} 5 & 0 & 0 & -5 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}}$$

(b) $T_1 = \begin{pmatrix} 1 & -1 & 3 \\ 1 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} \implies$ Similarity Transformation

$T_2 = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} \implies$ Projective Transformation

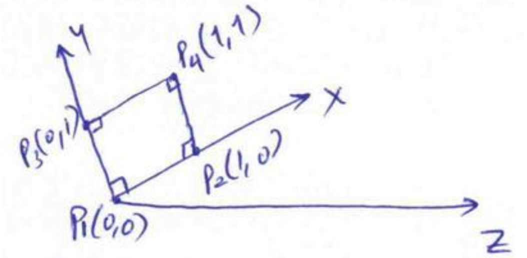
$T_3 = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix} \Rightarrow$ Affine Transformation

In case of affine transformations, parallel lines in real world remain parallel in image space.

BANSAL, HRITIK | 981110-T014 | | 2 | 1

Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

(c)



$$T_3 = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\begin{pmatrix} kx \\ ky \\ k \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & -1 \\ \boxed{0 \quad 0 \quad 2} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

As we can see above that it is better to transform the last row from (0 0 2) to (0 0 1) for easy computations.

This requires each element of T_3 be divided by 2.

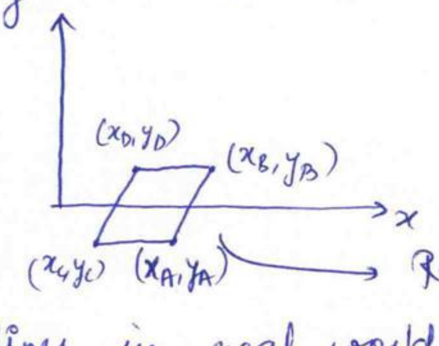
$$\boxed{\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 & -2/2 & 3/2 \\ 2/2 & 0 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}} \quad \text{———} \quad (1)$$

Now applying Equation (1) to all the four points of the given square.

$(X_A, Y_A) \equiv (0,0)$ $\quad (X_B, Y_B) = (1,0)$ $\quad (X_C, Y_C) = (0,1)$ $\quad (X_D, Y_D) = (1,1)$

$x_A = 1.5$ $\quad x_B = 2$ $\quad x_C = 0.5$ $\quad x_D = 1$

$y_A = -0.5$ $\quad y_B = 0.5$ $\quad y_C = -0.5$ $\quad y_D = 0.5$



Resulting Shape.

Parallel lines in real world are mapped to parallel lines in image coordinates under affine transformation.

BANSAL, HRITIK	981110-T014		3	1
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

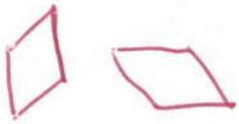
(a) Projective Transformation in this case has 8 dofs. [Parallel lines do not remain parallel in image]

Affine Transformation has 6 dofs. [includes shear]

Similarity Transformation has ~~5~~ 4 dofs. [Same as affine but does not allow change in angles (shear)]

Change of angles, shape, scale occurs in case of projective transformation.

-0.5p



sheer 2 dof.

8.5/9

BANSAL, HRITIK	981110-T014		4	2
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

Question 2

(a) $g_1 = [1, 0, -1] \Rightarrow$ Edge Detection ($\propto L_x$)

$g_2 = [1, -2, 1] \Rightarrow$ Blob Detection ($\propto L_{xx}$)

$g_3 = [1, 2, 1] \Rightarrow$ Noise Removal ($\propto$ Low Pass Filter)

(b) For me g_2 is 2^{nd} order differential filter.

$$g_2 = [\underset{x=-1}{1}, \underset{x=0}{-2}, \underset{x=1}{1}]$$ [Defining the Reference]

$$g_2(x) = 1\delta(x+1) - 2\delta(x) + 1\delta(x-1)$$

$$G_2(\omega) = \int_{-\infty}^{\infty} g_2(x) e^{-i\omega x} dx$$

$$= \int_{-\infty}^{\infty} \delta(x+1) e^{-i\omega x} dx - 2\int_{-\infty}^{\infty} \delta(x) e^{-i\omega x} + \int_{-\infty}^{\infty} \delta(x-1) e^{-i\omega x} dx$$

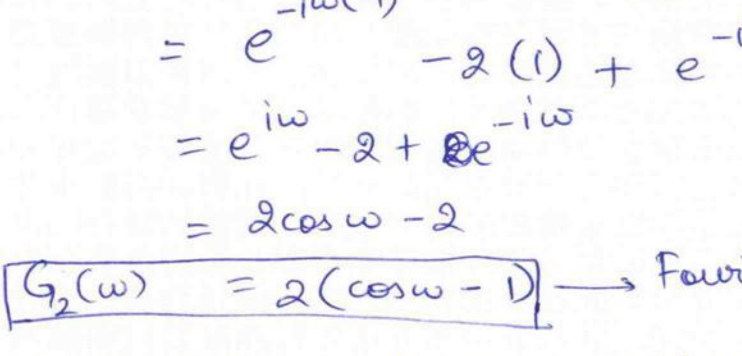
$$= e^{-i\omega(-1)} - 2(1) + e^{-i\omega(1)}$$

$$= e^{i\omega} - 2 + e^{-i\omega}$$

$$= 2\cos\omega - 2$$

$$\boxed{G_2(\omega) = 2(\cos\omega - 1)} \longrightarrow \text{Fourier transform}$$

Magnitude of fourier transform ; $\|G_2(\omega)\| = 2\|\cos\omega - 1\|$



Required Diagram

It is a High Pass Filter.

BANSAL, HRITIK | 981110-T014 | | 5 | 2
Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

(c) $g = [b, a, 0, -a, -b]$ with $x=-2$, $x=-1$, $x=0$, $x=1$, $x=2$ [Defining the reference]

$$g(x) = b\delta(x+2) + a\delta(x+1) - a\delta(x-1) - b\delta(x-2)$$

$$G(\omega) = b\int_{-\infty}^{\infty}\delta(x+2)e^{-i\omega x}dx + a\int_{-\infty}^{\infty}\delta(x+1)e^{-i\omega x}dx - a\int_{-\infty}^{\infty}\delta(x-1)e^{-i\omega x}dx - b\int_{-\infty}^{\infty}\delta(x-2)e^{-i\omega x}dx$$

$$= be^{-i\omega(-2)} + ae^{-i\omega(-1)} - ae^{-i\omega(+1)} - be^{-i\omega(2)}$$

$$= be^{i2\omega} + ae^{i\omega} - ae^{-i\omega} - be^{-i2\omega}$$

$$= b(e^{i2\omega} - e^{-i2\omega}) + a(e^{i\omega} - e^{-i\omega})$$

$$= ib2\sin(2\omega) + i2a\sin\omega = \boxed{2i(b\sin 2\omega + a\sin\omega)}$$

Using the given Taylor Expansion of $\underline{\sin\omega}$

$$G(\omega) = 2i\left(b\left[(2\omega) - \frac{(2\omega)^3}{3!} + \frac{(2\omega)^5}{5!} + O(\omega^7)\right] + a\left[\omega - \frac{\omega^3}{3!} + \frac{\omega^5}{5!} + O(\omega^7)\right]\right)$$

$$= 2i\left(b\left[2\omega - \frac{(2\omega)^3}{3!} + O(\omega^5)\right] + a\left[\omega - \frac{\omega^3}{3!} + O(\omega^5)\right]\right)$$

for lower frequencies, $\underline{G(\omega) \approx i\omega}$ if coefficient of $\underline{\omega^3 = 0}$.

$$-\frac{b8}{6} + a\left(-\frac{1}{6}\right) = 0$$

$$a + 8b = 0$$

$$a = -8b$$

$$\Rightarrow G(\omega) = i\omega' + O(\omega^5)$$

for small frequencies

$$G(\omega) = i\omega'$$

-0.5p

Taking value of $(b = 1, a=-8)$ one of the possible filter values would be

7.5/8

$$\boxed{[1, -8, 0, 8, -1]} \times -\frac{1}{12}$$

BANSAL, HRITIK | 981110-T014[redacted] | | 6 | 3

Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

Question 3

(a) Prior probabilities are $p_A = 0.4$ and $p_B = 0.6$

Let Z denote the random variable associated with the size of the segment.
Let C denote the random variable associated with the colour of the segment.

$p(Z=S, C=R|A) = 0.3$ $\quad$ $p(Z=S, C=R|B) = 0.0$

$p(Z=S, C=G|A) = 0.2$ $\quad$ $p(Z=S, C=G|B) = 0.1$

$p(Z=S, C=B|A) = 0.1$ $\quad$ $p(Z=S, C=B|B) = 0.1$

$p(Z=M, C=R|A) = 0.2$ $\quad$ $p(Z=M, C=R|B) = 0.1$

$p(Z=M, C=G|A) = 0.1$ $\quad$ $p(Z=M, C=G|B) = 0.1$

$P(Z=M, C=B|A) = 0.0$ $\quad$ $p(Z=M, C=B|B) = 0.1$

$P(Z=L, C=R|A) = 0.0$ $\quad$ $p(Z=L, C=R|B) = 0.1$

$p(Z=L, C=G|A) = 0.1$ $\quad$ $p(Z=L, C=G|B) = 0.2$

$p(Z=L, C=B|A) = 0.0$ $\quad$ $p(Z=L, C=B|B) = 0.2$

Using the Bayes' Formula

$$\boxed{P(A|Z=s_i, C=c_i) = \frac{P(Z=s_i, C=c_i|A)P(A)}{P(Z=s_i, C=c_i|A)P(A) + P(Z=s_i, C=c_i|B)P(B)}} \quad (1)$$

$$\boxed{P(B|Z=s_i, C=c_i) = 1 - P(A|Z=s_i, C=c_i)} \quad (2)$$

Applying Eqn (1) and (2) to get:

Remember: if $P(A|Z=s_i, C=c_i) > P(B|Z=s_i, C=c_i)$ then $Z=s_i, C=c_i$ belongs to Class A

$$P(A|Z=S, C=R) = \frac{(0.3)(0.4)}{(0.3)(0.4)+(0.0)(0.6)} = 1 \text{ (Class A)}$$

$$P(A|Z=S, C=G) = \frac{(0.2)(0.4)}{(0.2)(0.4)+(0.1)(0.6)} = \frac{8}{14} \text{ (Class A)}$$

$$P(A|Z=S, C=B) = \frac{(0.1)(0.4)}{(0.1)(0.4)+(0.1)(0.6)} = 0.4 \text{ (Class B)}$$

BANSAL, HRITIK | 981110-T014 | | 7 | 3

Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

(a) contd.

$$P(A \mid Z=M, C=R) = \frac{(0.2)(0.4)}{(0.2)(0.4)+(0.1)(0.6)} = \frac{8}{14} \quad (\text{Class A})$$

$$P(A \mid Z=M, C=G) = \frac{(0.1)(0.4)}{(0.1)(0.4)+(0.1)(0.6)} = 0.4 \quad (\text{Class B})$$

$$P(A \mid Z=M, C=B) = 0.0 \quad (\text{Class B})$$

$$P(A \mid Z=L, C=R) = 0.0 \quad (\text{Class B})$$

$$P(A \mid Z=L, C=G) = \frac{(0.1)(0.4)}{(0.1)(0.4)+(0.2)(0.6)} = \frac{4}{16} = 0.25 \quad (\text{Class B})$$

$$P(A \mid Z=L, C=B) = 0.0 \quad (\text{Class B})$$

Size

Colour	S	M	L
R	A	A	B
G	A	B	B
B	B	B	B

Matrix (1)

(b) Optimal classifier results in an incorrect prediction if it does not assign a new segment having a particular size & colour to the expected class according to the matrix(1) found above.

Probability that the optimal classifier results in an incorrect prediction = 1 − Probability that all predictions are correct

$$= 1 - (1)\left(\frac{8}{14}\right)\left(\frac{6}{10}\right)\left(\frac{8}{14}\right)\left(\frac{6}{10}\right)(1)(1)\left(\frac{3}{4}\right)(1)$$

−1.5p

$$= 1 - \left(\frac{6}{10}\right)^2\left(\frac{3}{4}\right)\left(\frac{64}{196}\right)$$

Not correct!

Sum up all cases that doesn't result in correct result.

$$= 1 - \frac{36}{100} \cdot \frac{3}{4} \cdot \frac{\cancel{64}\,16}{\cancel{196}\,49}$$

$$= 1 - \frac{(36)(12)}{4900} = \boxed{\frac{4468}{4900}}$$

BANSAL, HRITIK | 981110-T014 | | 8 | 3
Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

(c) Assume that training samples (with class) are already present in the feature space. Given the features of a new sample, K-nearest neighbour finds the ~~a~~ K-nearest samples to it based on Euclidean distance ~~btw~~ between their feature vectors. The class associated with the new sample is the majority vote of classes associated with those K-nearest samples.

Algorithm
1. Finding K-nearest samples in feature space.
2. New sample's class is the one which has maximum # of votes in the above selected samples.

K-nearest neighbour cannot be used in the case mentioned above because (1) we need to find the Euclidean distance btw feature vectors in feature space. In the above case, Feature vectors are not numbers.
(2) For the same feature like (Z = Sm, C = R) the sample can belong to any of the class A or B.
(3) Prior probabilities are not taken into account in K-nearest neighbours.

5.5/7

KTH VETENSKAP OCH KONST

BANSAL, HRITIK	981110-T014		9	4
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

Question 4

(a) The fundamental difference between image sharpening and contrast enhancement. ~~is that~~

Image Sharpening : We would like to improve finer details in the image like edges using a suitable filter.

Contrast Enhancement : ~~to~~ the histogram of the image is made uniform so details in the image are clearly visible by eyes.

~~(b) Image blurring caused by lens system are due to intrinsic parameters of the camera. Sharpening of the image to correct for blur uses some kind of filter to do that. No filter can perfectly remove the external noise caused due to in intrinsic parameters of camera.~~

(b) Sharpening of the image is typically performed by some kind of high pass filters. The thing with high pass filter is that ~~tt~~ they make the edges in the image finer as well as increase the noise component of the image. So, this increment in the noise components adds to the imperfections caused by lens system and we won't get a perfect result.

BANSAL, HRITIK	981110-T014		10	4
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

(c) $p_R(r) = \dfrac{\pi}{2+\pi}\left(1+\sin(\pi r)\right)$, $r \in (0,1)$

$p_S(s) = 1$, $s \in (0,1)$

$p_R(r)\,dr = p_S(s)\,ds$

$$\boxed{p_R(r) = \frac{ds}{dr} = T'(r)} \quad — (1)$$

$[T(0)=0]$

$$\int_0^r T'(r)\,dr = \int_0^r p_R(\cancel{t})\,d\cancel{t}$$

$$T(r) = \frac{\pi}{2+\pi}\left[\cancel{t} - \frac{\cos(\pi \cancel{t})}{\pi}\right]_0^r$$

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

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

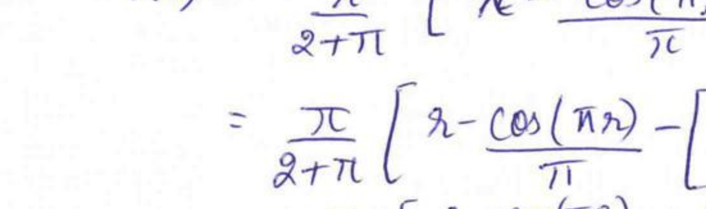
$$\boxed{T(r) = \frac{\pi}{2+\pi}\left[r\right] + \frac{1}{2+\pi}\left[1-\cos(\pi r)\right]} \quad — (2)$$

Stretching occurs when $T'(r) > 1$ (from (1))

$$T'(r) = \frac{\pi}{2+\pi}\left[1+\sin(\pi r)\right] > 1$$

$$1+\sin(\pi r) > \frac{\pi+2}{\pi} \Rightarrow 1+\sin(\pi r) > 1+\frac{2}{\pi}$$

$$\Rightarrow \boxed{\sin(\pi r) > \left(\frac{2}{\pi}\right)}$$



$r \in (0,1)$

So, stretching of grey ~~pixel~~ levels occurs between r_1 and r_2 $(r_1 < r < r_2)$

where $\boxed{r_1 = \frac{1}{\pi}\sin^{-1}\left(\frac{2}{\pi}\right)}$

$$\boxed{r_2 = 1 - \frac{1}{\pi}\sin^{-1}\left(\frac{2}{\pi}\right)}$$

BANSAL, HRITIK | 981110-T014 | | 11 | 4
Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

(a) 1) Sharpening is done using high pass filters on the image. Gradient derivatives such as ∇f or $|\nabla f|$ can be used ~~by~~ as high pass filters.

Not enough!

2) A 3-point filter that can be used to sharpen the image is

$g = [1, 0, -1]$ ex. $[-1, 3, -1]$

not symmetric   symmetric

3) $f * g$

```
              0   0   0   1   3   6   8   9   9   9
                                          1   0  -1
          ------------------------------------------
          0   0   0  -1  -3  -6  -8  -9  -9  -9
      0   0   0   0   0   0   0   0   0   0
  0   0   0   1   3   6   8   9   9   9
  ------------------------------------------------
  0   0   0   1   3   5   5   3   1   0  -9  -9
```

-2p

$l = f * g = [0, 1, 3, 5, 5, 3, 1, 0]$

4) The image l indeed gets sharpened as the frequency of intensity changes in l are more than frequency of intensity ~~of~~ ~~at~~ image f.

6/8

It should leave uniform areas untouched, but it doesn't. It also treats black→white and white→black differently, which it shouldn't.

KTH VETENSKAP OCH KONST

BANSAL, HRITIK	981110-T014		12	5
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

Questions

(a) Corresponding essential matrix E is given by $\boxed{E = RT_t}$

where R is the rotation matrix and T_t is the skew symmetric Translation Matrix.

$$T_t = \begin{pmatrix} 0 & t_x & 0 \\ -t_x & 0 & t_z \\ 0 & -t_z & 0 \end{pmatrix} \qquad t_y = 0$$

$$E = RT_t$$

$$= \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} 0 & t_x & 0 \\ -t_x & 0 & t_z \\ 0 & -t_z & 0 \end{pmatrix}$$

$$\boxed{E = \begin{pmatrix} 0 & t_x\cos\theta + \sin\theta t_z & 0 \\ -t_x & 0 & t_z \\ 0 & t_x\sin\theta - t_z\cos\theta & 0 \end{pmatrix}}$$

which is of form

$$E = \begin{pmatrix} 0 & e_{12} & 0 \\ e_{21} & 0 & e_{23} \\ 0 & e_{32} & 0 \end{pmatrix}$$

(b) $P_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ $\qquad P_2 = \begin{pmatrix} \cos\theta & 0 & -\sin\theta & t_x \\ 0 & 1 & 0 & 0 \\ \sin\theta & 0 & \cos\theta & t_z \end{pmatrix}$ Not P_1 and P_2!

Epipoles Projection of $O_1 = (0,0,0)$ in second camera.

$Eq = 0 \Rightarrow \ldots$ $\begin{pmatrix} \cos\theta & 0 & -\sin\theta & t_x \\ 0 & 1 & 0 & 0 \\ \sin\theta & 0 & \cos\theta & t_z \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} t_x \\ 0 \\ t_z \end{pmatrix} \equiv \underline{\left(\frac{t_x}{t_z}, 0\right)}$

True!

Projection of $O_2 = (t_x \; 0 \; t_z)$ in first camera.

$q^T E = 0 \Rightarrow \ldots$ $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} t_x \\ 0 \\ t_z \\ 1 \end{pmatrix} = \begin{pmatrix} t_x \\ 0 \\ t_z \end{pmatrix} \equiv \left(\frac{t_x}{t_z}, 0\right)$

not True!

Please Turn Over

Equation of Epipolar line $\boxed{q'^T E q = 0}$ where left image point is q and q' is the right image epipolar <u>Line</u>.

$$E = \begin{pmatrix} 0 & e_{12} & 0 \\ e_{21} & 0 & e_{23} \\ 0 & e_{32} & 0 \end{pmatrix} \qquad \cancel{q} = ($$

$$q'^T E q = 0 \Rightarrow q^T E^T q' = 0$$

$$(x_l \; y_l \; 1) \begin{pmatrix} 0 & e_{21} & 0 \\ e_{12} & 0 & e_{32} \\ 0 & e_{23} & 0 \end{pmatrix} \cancel{\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}} = 0$$

→ ~~$y_l = 0$ $x_l =$~~

$$(x_l \; y_l \; 1) \begin{pmatrix} e_{21} y \\ e_{12} x + e_{32} \\ e_{23} y \end{pmatrix} = 0$$

ok epipolar line!

~~It should be independent on (x, y) !!~~

$$x_l e_{21} y + (e_{12} x + e_{32}) y_l + e_{23} y = 0$$

$$\boxed{(e_{21} y) x_l + (e_{12} x + e_{32}) y_l + e_{23} y = 0} \longrightarrow \text{Equation of epipolar line in the left image.}$$

-TP

BANSAL, HRITIK	981110-T014		13	5
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

(c) If we don't really know R_θ and t. Then we have ~~4~~ 3 unknown parameters aka $\theta, t_x, t_y^{=0}, t_z$. In this case.

~~One~~ One correspondence matching gives us ~~two~~ one equations, ~~one~~ in x-coordinates and the other in y-coordinates.

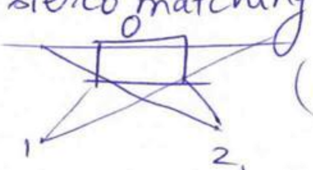
-5p

Hence 2 image correspondences between the two cameras will result into ~~4~~ 2 equations. ~~4~~ 2 equations & ~~4~~ 3 unknowns -5 scale can be solved.

scale is irrelevant!

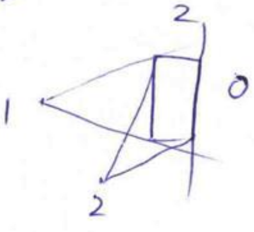
(d) Other reasons of why stereo matching is difficult:

(1) Occlusion



(Hidden regions seen by just one of the cameras)

(2) Foreshortening

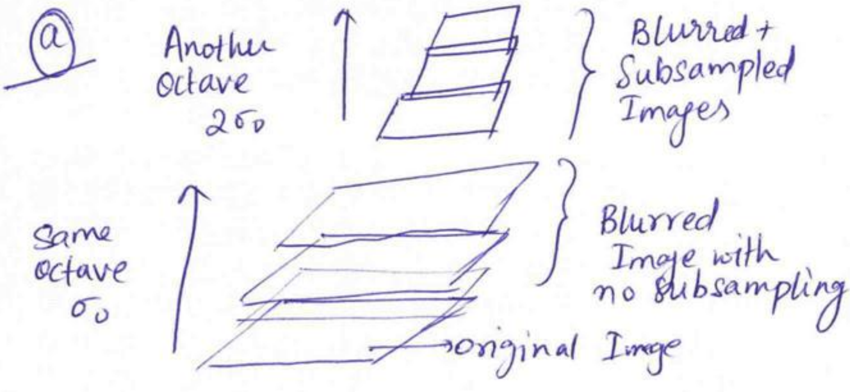


(Due to perspective projection, same object is smaller in one of the cameras)

6/8.

BANSAL, HRITIK	981110-T014		14	6
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

(a)



The blurring in a particular octave is done at different scales. From σ_0 to $\sigma_0 2^{\frac{1}{s+1}}$ if s is the # of images in that octave.

Gaussian Pyramid representation requires to extend the original image in different scales or octaves. We move from finer scales to coarser scales in gaussian pyramid. At coarser scales, finest image structures get blurred so subsampling doesnot lead to loss of information.

Finding local extrema in space as well as scale representation helps us to ~~get~~ find the size of an image structure in original image. These representations are used for feature detection because they can give an idea about the size of image structures as well as features detector through this representation are scale invariant and many more invariances ~~etc~~ can be easily introduced.

(b) Convolution of two functions $f(x;t)$ and $g(x;s)$ is given by

$$h = \boxed{f * g = \int_{\alpha} f(\alpha)\, g(x-\alpha)\, d\alpha}$$

$$\cancel{h = \int_{\alpha} \frac{1}{2\pi t} \exp\left(\frac{-\alpha^2}{2t}\right) \frac{1}{2\pi s} \exp\left(\frac{(x-\alpha)^2}{2s}\right) d\alpha}$$

$$\cancel{h = \int_{-\infty}^{\infty} \frac{1}{2\pi t} e^{-\frac{\alpha^2}{2t}} \cdot \frac{1}{2\pi s}}$$

~~We also know that convolution of two functions in spatial domain is same as multiplication in fourier domain.~~

KTH VETENSKAP OCH KONST

BANSAL, HRITIK | 981110-7014 | | 15 | 6
Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

contd.

(b) $$h = \int_{-\infty}^{\infty} \frac{1}{2\pi t} e^{-\frac{\alpha^2}{2t}} \cdot \frac{1}{2\pi s} e^{-\frac{(x-\alpha)^2}{2s}} \, d\alpha$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi t} \cdot \frac{1}{2\pi s} e^{-\frac{\alpha^2}{2t} - \frac{x^2}{2s} - \frac{\alpha^2}{2s} + \frac{2\alpha x}{2s}} \, d\alpha$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi t} \frac{1}{2\pi s} e^{-\left(\alpha^2\left(\frac{1}{2t}+\frac{1}{2s}\right) - \alpha\left(\frac{x}{s}\right) + \frac{x^2}{2s}\right)} d\alpha \quad \text{——— (1)}$$

we will try to ~~com~~ make it in square form.

forming square

$$\cancel{\alpha^2\left(\frac{1}{2t}+\frac{1}{2s}\right) - \alpha\left(\frac{x}{s}\right) + \left(\frac{x^2}{2s}\right)}$$

$$\cancel{(\alpha-k)^2 = \alpha^2 - 2\alpha K + K^2}$$

$$\cancel{\frac{1}{\frac{1}{2t}+\frac{1}{2s}} = \frac{-2k}{-\frac{x}{s}} = \frac{k^2}{\frac{x^2}{2s}}}.$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi t} \cdot \frac{1}{2\pi s} e^{-\left(\frac{t+s}{2st}\right)\left[\left(\alpha - \frac{xt}{t+s}\right)^2 + \frac{x^2}{(t+s)}\left(\frac{st}{(t+s)}\right)\right]} d\alpha$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi t} \frac{1}{2\pi s} e^{-\frac{(t+s)}{2st}\left(\alpha - \frac{xt}{t+s}\right)^2} \cdot e^{-\frac{x^2}{2(t+s)}} d\alpha$$

Substitute $\alpha' = \left(\alpha - \frac{xt}{t+s}\right) \Rightarrow d\alpha' = d\alpha$ [Limits won't change]

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi t} \cdot \frac{1}{2\pi s} e^{-\frac{(t+s)}{2st}\alpha'^2} e^{-\frac{x^2}{2(t+s)}} d\alpha'$$

Much easier with Fourier!

after substitution again

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi(t+s)2\pi} e^{-\frac{m^2}{2}} e^{-\frac{x^2}{2(t+s)}} \, dm$$

$$\boxed{\because \int_{-\infty}^{\infty} e^{-\frac{m^2}{2}} dm = 2\pi \ (\text{for 2D})}$$

$$= \frac{1}{2\pi(t+s)2\pi} e^{-\frac{x^2}{2(t+s)}} \cdot 2\pi = \boxed{\frac{1}{2\pi(t+s)} e^{-\frac{x^2}{2(t+s)}} = g(x; t+s)} \text{ Proved}$$

BANSAL, HRITIK | 981110-T014 | | 16 | 6

Family name, first name | Personal Registration Number | Programme | Sheet no. | Problem no.

(c) $f(x;t) \rightarrow$ shape of the Gaussian Blob.

$h(x;s) = g(x;s) - g(x;2s)$ [filter to detect the blob]

Applying $h(x;s)$ filter to $f(x;t)$ gaussian blob to get:

$$L(x;\cdot) = f(x;t) * h(x;s)$$

$$= f(x;t) * [g(x;s) - g(x;2s)]$$

$$\boxed{L(x) = g(x;t+s) - g(x;2s+t)}$$

$$L(x) = \frac{1}{2\pi(t+s)} e^{-\frac{x^2}{2(t+s)}} - \frac{1}{2\pi(t+2s)} e^{-\frac{x^2}{2(t+2s)}}$$

for the output to be maximum at center of the blob $\boxed{\left.\frac{\partial L}{\partial s}\right|_{x=0} = 0}$

$$\frac{\partial L}{\partial s} = \frac{1}{2\pi(t+s)} e^{-\frac{x^2}{2(t+s)}} \left(\frac{x^2}{2(t+s)^2}\right) + e^{-\frac{x^2}{2(t+s)}} \frac{-1}{2\pi(t+s)^2}$$

$$- \left[\frac{1}{2\pi(t+2s)} e^{-\frac{x^2}{2(t+2s)}} \frac{2x^2}{2(t+2s)} + e^{-\frac{x^2}{2(t+2s)}} \left(-\frac{2}{2\pi(t+2s)^2}\right)\right]$$

at $\boxed{\left.\frac{\partial L}{\partial s}\right|_{x=0} = 0}$ gives.

$$-\frac{1}{2\pi(t+s)^2} + \frac{2}{2\pi(t+2s)^2} = 0$$

$$\frac{2}{2\pi(t+2s)^2} = \frac{1}{2\pi(t+s)^2}$$

Nice!

$$\frac{(t+2s)^2}{2} = (t+s)^2$$

$$\frac{t+2s}{\sqrt{2}} = t+s$$

$$t + 2s = \sqrt{2}t + \sqrt{2}s \Rightarrow$$

$$s(2-\sqrt{2}) = t(\sqrt{2}-1)$$

$$s = \frac{t(\sqrt{2}-1)}{\sqrt{2}(\sqrt{2}-1)} = \frac{t}{\sqrt{2}}$$

$\Rightarrow \boxed{s = t/\sqrt{2}}$ Required answer

BANSAL, HRITIK	981110-T014		17	6
Family name, first name	Personal Registration Number	Programme	Sheet no.	Problem no.

(d) Three kinds of invariances that SIFT descriptor has:

① Rotational Invariance

② Scale Invariance

10/10 ③ Illuminance Invariance