Qn.3 (1) Eq.6: $J_2(I_1) = \sum_{i,k} \rho(f_{i,k} \cdot I_1) + \rho(f_{ik} \cdot (I - I_1))$

$$+ \lambda \sum_{i \in S_1,k} \rho(f_{i,k} \cdot I - f_{ik} \cdot I_1)$$

$$+ \lambda \sum_{i \in S_2,k} \rho(f_{i,k} \cdot I_1)$$

This is rewritten as:

Eq. 7: $J_2(v) = \sum_j \rho_j (A_{j \rightarrow} v - b_j)$

Here, v is vectorized form of Image I_1
(columns placed one below each other)

b_j is either vectorized form of derivative of input image I or 0.
$\therefore$ b_j is either vectorized $(-f_{ik} \cdot I)$ or 0

A_j represents matrix which when multiplied with v gives same result as convolving with filter f_{ik}.
~~It will have~~ Each row of A_j will have values of f_{ik} in row form and shifted.
it is basically implementation of convolution as matrix multiplication.
when we want to consider only $i \in S_1$ or $i \in S_2$ then the corresponding rows in A_j are set to zero too.

Scanned by CamScanner

Let A be matrix form of convolution with f_{ik} filter. and b be vectorized form of input image derivative

$\therefore$ We have following relations.

j	Term	A_j	b_j
0	$\rho(f_{ik} \cdot I_1)$	A	0
1	$\rho(f_{ik}(I - I_1))$	$-A$	$-b$
2	$\rho(f_{ik} I - f_{ik} I_1)_{i \in S_1}$	$-A$ (only rows $\in i$ are non zero)	$-b$
3	$\rho(f_{ik} I_1)_{i \in S_2}$	A (only rows $\in i$ are non zero)	0

(2)

$$\left.\begin{aligned} J_2(I_1) &= \sum_{i,k} \rho(f_{ik} \cdot I_1) \\ &+ \sum_{i,k} \rho(f_{ik} \cdot (I - I_1)) \end{aligned}\right\} \text{prior}$$

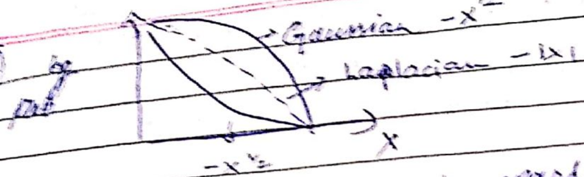
$$\left.\begin{aligned} &+ \lambda \sum_{i \in S_1, k} \rho(f_{ik} I - f_{ik} \cdot I_1) \\ &+ \lambda \sum_{i \in S_2, k} \rho(f_{ik} I_1) \end{aligned}\right\} \text{likelihood}$$

where $\rho(x) = \log\left(\frac{\pi_1}{2s_1} e^{\frac{-|x|}{s_1}} + \frac{\pi_2}{2s_2} e^{-|x|/s_2}\right)$

Mixture of laplacian is used as ~~[illegible]~~ ~~[illegible]~~ prior.

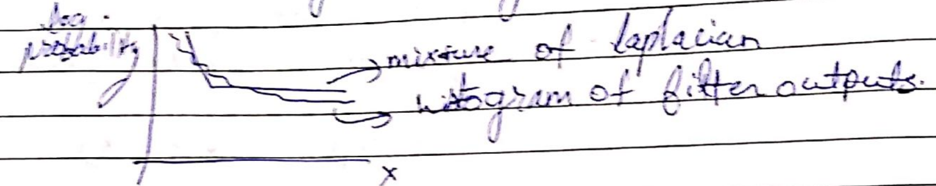
Likelihood is term is value ~~[illegible]~~ of derivative in I_1 should be same as derivatives in I.

3



Gaussian distribution is not sparse as the log probability for higher values is high as compared to other distributions.

$-x^{1/2}$ is sparse distribution.
~~else~~ Laplacian distribution lies on the border however mixture of laplacian is sparse and it represents the filter outer histogram more ~~[illegible]~~ accurately.



Here the paper uses a likelihood term that is different from Gaussian distribution prior

Page: Date: / /

Q.4) $y = \phi x + \eta$ $\quad y \in \mathbb{R}^m, \phi \in \mathbb{R}^{m \times n}, m \ll n$

$x \in \mathbb{R}^n, \eta \in N(0, \sigma^2 I_{m \times m})$

Given: $x \sim N(0, \Sigma_x)$ $\quad \Sigma_x \in \mathbb{R}^{n \times n}$

$$\therefore p(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_x|}} \exp\left(\frac{-1}{2} x^T \Sigma_x^{-1} x\right)$$

$$p(y|x) = \frac{1}{\sqrt{(2\pi)^m |\sigma^2 I_{m \times m}|}} \exp\left(\frac{-1}{2}(y - \phi x)^T \frac{I_{m \times m}}{\sigma^2} (y - \phi x)\right)$$

$$\hat{x}_{MAP}(y) = \underset{x}{\arg\max}\ p(x|y)$$

$$= \underset{x}{\arg\max}\ p(y|x)\, p(x)$$

($p(y)$ doesnot depend on x)

$$= \underset{x}{\arg\max}\ \exp\left(\frac{-1}{2}(y-\phi x)^T \frac{I_{m \times m}}{\sigma^2}(y - \phi x)\right)$$

$$\times \exp\left(\frac{-1}{2} x^T \Sigma_x^{-1} x\right)$$

$$\times \left(1 / (2\pi)^{\frac{m+n}{2}} \sqrt{|\Sigma_x| |\sigma^2 I_{m \times m}|}\right)$$

$$= \underset{x}{\arg\min}\ (y - \phi x)^T \frac{I_{m \times m}}{\sigma^2} (y - \phi x)$$

$$+ x^T \Sigma_x^{-1} x$$

putting $\frac{d}{dx}\left((y - \phi x)^T \frac{I_{m \times m}}{\sigma^2}(y - \phi x) + x^T \Sigma_x^{-1} x\right) = 0$

$$\Rightarrow \therefore 2(y - \phi x)^T \frac{I_{m \times m}}{\sigma^2}(-\phi) + 2x^T \Sigma_x^{-1} = 0$$

$$\therefore x^T \Sigma_x^{-1} = \frac{(y - \phi x)^T \phi}{\sigma^2}$$

$$\therefore \Sigma_x^{-T} x = \frac{\phi^T (y - \phi x)}{\sigma^2}$$

$$\therefore \sigma^2 \Sigma_x^{-T} x = \phi^T y - \phi^T \phi x$$

$$\therefore (\sigma^2 \Sigma_x^{-T} + \phi^T \phi) x = \phi^T y$$

$$\therefore \boxed{\hat{x}_{MAP} = (\phi^T \phi + \sigma^2 \Sigma_x^{-T})^{-1} \phi^T y}$$