Q-2)

Part a

Eq 6 can be written as

$$E(I_1) = \rho((f_k * I_1)_i) + \rho((f_k * (I - I_1))_i) +$$

$$\lambda \sum_{i \in S_{1,k}} \rho((f_k * I - f_k * I_1)_i) +$$

$$\lambda \sum_{i \in S_{2,k}} \rho((f_k * I_1)_i)$$

We changed the eq. so that we can write the expression

acc. to eq 7

eq 7 can be written as

$$J_3(v) = \sum_{j=1}^{4} \sum_{k} \rho(A_{j,k} v - b_{j,k})$$

Defining matrix $A_{j,k}$

Let $I = n_1 \times n_2$ dimension

0 = zero matrix of dim $n_2 \times n_2$

$f_k = k^{th}$ filter

f_k^i = ith row of filter f_k

$T_{f_k^i}$ = Toeplitz matrix of dimension $n_2 \times n_2$

$$A_{j,k} = \begin{array}{c} \\ n_1 \text{ blocks} \end{array} \overbrace{\begin{bmatrix} T_{f_k^1} & 0 & 0 & \cdots & 0 \\ T_{f_k^2} & T_{f_k^1} & 0 & \cdots & 0 \\ & & & 0 & 0 \\ T_{f_k^3} & \cdots & \cdots & \cdots & \\ \vdots & \vdots & \vdots & T_{f_k^2} & T_{f_k^1} \\ \vdots & \vdots & & & \end{bmatrix}}^{n_1 \text{ blocks}}_{n \times n}$$

$v = \mathrm{vec}(I_1)$ of size $n \times 1$ $\quad$ b_i = is a vect. of size $n \times 1$

Now we define the corresponding filter to construct matrix $A_{j,k}$ and also mention values of b_j for each of the 4 terms in eq 5

	$A_{j,k} \rightarrow$	b_j
1 $\sum_{j,k} \rho((f_k \cdot I_1)_j)$	f_k	$0 \rightarrow$ vector
2 $\sum_{i,k} \rho((f_k \cdot (I - I_1))_i)$	$-f_k$	$\boxed{-f_k I}$ *
3 $\lambda \sum_{i \in S_1, k} \rho((f_k \cdot I - f_k \cdot I_1)_i)$	$-f_k$	$\boxed{-f_k I}$ *
4 $\lambda \sum_{i \in S_2, k} \rho((f_k \cdot I_1)_i)$	f_k	0

* Is constructed by making a matrix like $A_{j,k}$ (mentioned above) corresponding to filter f_k & I is converted to vector of $n \times 1$ dimension to given b_j of dimension $n \times 1$.

Point b)

$$J_2(I_1) = \overbrace{\sum_{i,k} \rho(f_{i,k} \cdot I_1) + \rho(f_{i,k} \cdot (I - I_1))}^{\text{likelihood}}$$

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

likelihood used in the paper

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

prior used in the paper

$$P_r(x) = \frac{\pi}{2s_1} e^{-|x|/s_1} + \frac{\pi}{2s_2} e^{-|x|/s_2}$$

Part c)

Reason

1. The histogram of derivative filters are peaked at zero and fall of much faster than a Gaussian.

2. Papper investigate the importance of the sparse likelihood model. for first & second derivative filters they compared sparse likelihood that was fitted to the distribution of edges in natural images with the simpler laplacian & Gaussian prior.

They found that highly non-sparse nature of the Gaussian prior result in a very bad decomposition.

The laplacian prior behaves much better than Gaussian prior, but the actual sparse prior that was fitted to the distribution of filters in real images outperforms the laplacian prior.

Therefore likelihood is given by sparse prior

eq. $Pr(x) = \frac{\pi_1}{2s_1} e^{-|x|/s_1} + \frac{\pi_2}{2s_2} e^{-|x|/s_2}$