Partial Solution of Homework #1

Question 1 (Solution)

The error is given by

\[ E_Q = \int_{L}^{H} \frac{(t - Q(t))^2}{t^2} p(t) dt = \sum_{i=1}^{J} \int_{d_i}^{d_{i+1}} \frac{(t - r_i)^2}{t^2} p(t) dt \]

The optimization problem:

\[ \min_{[d_i, r_i]^{J+1}_{i=1}} E_Q = \min_{[d_i, r_i]^{J+1}_{i=1}} \sum_{i=1}^{J} \int_{d_i}^{d_{i+1}} \frac{(t - r_i)^2}{t^2} p(t) dt \]

a. Recall that \( \frac{\partial}{\partial x} \int_{y}^{x} f(t) dt = f(x) \) and \( \frac{\partial}{\partial x} \int_{x}^{y} f(t) dt = -f(x) \)

Hence, \( \frac{\partial}{\partial d_i} E_Q = \frac{(d_i - r_i)^2}{d_i^2} p(d_i) - \frac{(d_i - r_{i+1})^2}{d_i^2} p(d_i) = 0 \) \( \Rightarrow d_i = \frac{r_i + r_{i+1}}{2} \) \( i = 1, ..., J - 1 \)

b. \( \frac{\partial}{\partial r_j} E_Q = \sum_{i=1}^{J} \int_{d_{i+1}}^{d_i} \frac{t}{t^2} p(t) dt = -2 \int_{d_{i+1}}^{d_i} \frac{(t - r_j)^2}{t^2} p(t) dt = 0 \)

\( \Rightarrow r_j = \frac{\int_{d_{j+1}}^{d_j} p(t) dt}{\int_{d_{j+1}}^{d_j} t^2 p(t) dt} \) \( j = 1, ..., J \)

c. The error metric considers the ratio between the error and the signal-power, i.e., the error for a given x is normalized by its magnitude.
Question 3 (Solution)

(1) a. To find the optimal coefficient for each 3D sampling area (before quantization), we will first calculate the MSE for each area, then calculate its derivative and compare to zero.

\[
MSE_{\Delta_{ijk}}(\varphi_{ijk}) = \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \left[ \varphi(x, y, z) - \varphi_{ijk} \right]^2 \, dx \, dy \, dz =
\]

\[
= \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \left[ \varphi^2(x, y, z) - 2\varphi(x, y, z)\varphi_{ijk} + \varphi_{ijk}^2 \right] \, dx \, dy \, dz =
\]

\[
= \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi^2(x, y, z) \, dx \, dy \, dz - 2\varphi_{ijk} \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi(x, y, z) \, dx \, dy \, dz + \varphi_{ijk}^2
\]

\[
\Rightarrow \frac{\partial MSE_{\Delta_{ijk}}(\varphi_{ijk})}{\partial \varphi_{ijk}} = -2 \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi(x, y, z) \, dx \, dy \, dz + 2\varphi_{ijk} = 0
\]

\[
\Rightarrow \varphi_{ijk}^{\text{opt}} = \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi(x, y, z) \, dx \, dy \, dz
\]

b. The MSE for the optimal coefficient will be:

\[
MSE_{\Delta_{ijk}}(\varphi_{ijk}^{\text{opt}}) = \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi^2(x, y, z) \, dx \, dy \, dz - 2\varphi_{ijk}^{\text{opt}} \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi(x, y, z) \, dx \, dy \, dz + \left(\varphi_{ijk}^{\text{opt}}\right)^2
\]

\[
= \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi^2(x, y, z) \, dx \, dy \, dz - 2\left(\varphi_{ijk}^{\text{opt}}\right)^2 + \left(\varphi_{ijk}^{\text{opt}}\right)^2
\]

\[
= \frac{1}{\text{volume}(\Delta_{ijk})} \iiint \varphi^2(x, y, z) \, dx \, dy \, dz - \left(\varphi_{ijk}^{\text{opt}}\right)^2
\]

(2) Since our continuous signal is a real valued function \( \varphi(x, y, z) \in [\varphi_L, \varphi_H] \), the quantization error for \( b \) bits is the same as in the 1D and 2D cases:

\[
MSE^{\text{quantization}} = \frac{1}{12} \left( \varphi_H - \varphi_L \right)^2 2^{2b}
\]

(3) .

(4) \( B = N_x \cdot N_y \cdot N_z \cdot b \).

\( B \) is the total number of bits that can be used to represent the video.
\( N_x \) is the number of samples in each row for each of the video frames.
\( N_y \) is the number of samples in each column for each of the video frames.
$N_z$ is the number of samples in time (The $z$ axis is the progress in time. If $N_z$ is smaller than the number of frames, it means that each of the sampling 3D areas contains more than one frame).

(5) News presenter - $N_z = 22$, tennis player - $N_z = 88$. The idea of this section is to understand the meaning of $N_x, N_y, N_z$. While $N_x, N_y$ are the vertical and horizontal resolution, and decreasing them will affect the quality of the image (in this case, the quality of each frame), $N_z$ is the "temporal resolution" and decreasing it will affect the flow of the video (if $N_z$ is smaller than the total number of frames, instead of having 88 different frames, the video will have only $N_z$ different frames. The first $\frac{88}{N_z}$ frames will be identical, the next $\frac{88}{N_z}$ frames will be identical and so on). The tennis player video is much more dynamic. From only the eight frames shown in the section one can see that the posture of the player always changes and the position of the tennis racket always changes. In the second video, on the other hand, everything is static, except the head of the news presenter that moves a little bit in each frame. That's why $N_z$ can be smaller for the news presenter video without affecting the MSE too much, while decreasing it in the tennis player video will affect the MSE much more.