EE5585 Data Compression March 12, 2013

Lecture 14

Instructor: Arya Mazumdar *Scribe: Cheng-Yu Hung*

Scalar Quantization for Nonuniform Distribution

Suppose we have an input modeled by a random variable X with *pdf* $f_X(x)$ as shown in Figure 1 and we wished to quantize this source using a quantizer with M intervals. The endpoints of the intervals are known as *decision boundaries* and denoted as $\{b_i\}_{i=0}^{M}$, while the representative values $\{y_i\}_{i=1}^{M}$ are called *reconstructive levels*. Then, $Q(X) = y_i$ iff $b_{i-1} < X \leq b_i$, where the quantization operation is defined by $Q(\cdot)$.

The mean squared quantization error (*quantizer distortion*) is given by

$$\begin{aligned}\sigma_q^2 &= E[(X - Q(X))^2] && (1)\\ &= \int_{-\infty}^{\infty} (x - Q(x))^2 f_X(x)dx && (2)\\ &= \int_{b_0}^{b_M} (x - Q(x))^2 f_X(x)dx && (3)\end{aligned}$$

$$\Rightarrow \sigma_q^2 = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} (x - y_i)^2 f_X(x)dx \tag{4}$$

Thus, we can pose the optimal quantizer design problem as the followings:

Given an input *pdf* $f_X(x)$ and the number of quantization levels M in the quantizer, find the decision boundaries $\{b_i\}$ and the reconstruction levels $\{y_i\}$ so as to minimize the mean squared quantization error.

If we know the *pdf* of X, a direct approach to find the $\{b_i\}$ and $\{y_i\}$ that minimize the mean squared quantization error is to set the derivative of (4) with respect to b_j and y_j to zero, respectively. Then,

$$\begin{aligned}\frac{\partial \sigma_q^2}{\partial y_j} &= \frac{\partial}{\partial y_j}[\int_{b_{j-1}}^{b_j} (x - y_j)^2 f_X(x)dx] && (5)\\ &= \frac{\partial}{\partial y_j}[\int_{b_{j-1}}^{b_j} x^2 f_X(x)dx - 2y_j \int_{b_{j-1}}^{b_j} x f_X(x)dx + y_j^2 \int_{b_{j-1}}^{b_j} f_X(x)dx] && (6)\\ &= -2\int_{b_{j-1}}^{b_j} x f_X(x)dx + 2y_j \int_{b_{j-1}}^{b_j} f_X(x)dx = 0 && (7)\end{aligned}$$

$$\Rightarrow y_j = \frac{\int_{b_{j-1}}^{b_j} x f_X(x)dx}{\int_{b_{j-1}}^{b_j} f_X(x)dx} \tag{8}$$

$$\begin{aligned}\frac{\partial \sigma_q^2}{\partial b_j} &= \frac{\partial}{\partial b_j}[\int_{b_{j-1}}^{b_j} (x - y_j)^2 f_X(x)dx + \int_{b_j}^{b_{j+1}} (x - y_{j+1})^2 f_X(x)dx] && (9)\\ &= (b_j - y_j)^2 f_X(x)dx - (b_j - y_{j+1})^2 f_X(x)dx = 0 && (10)\end{aligned}$$

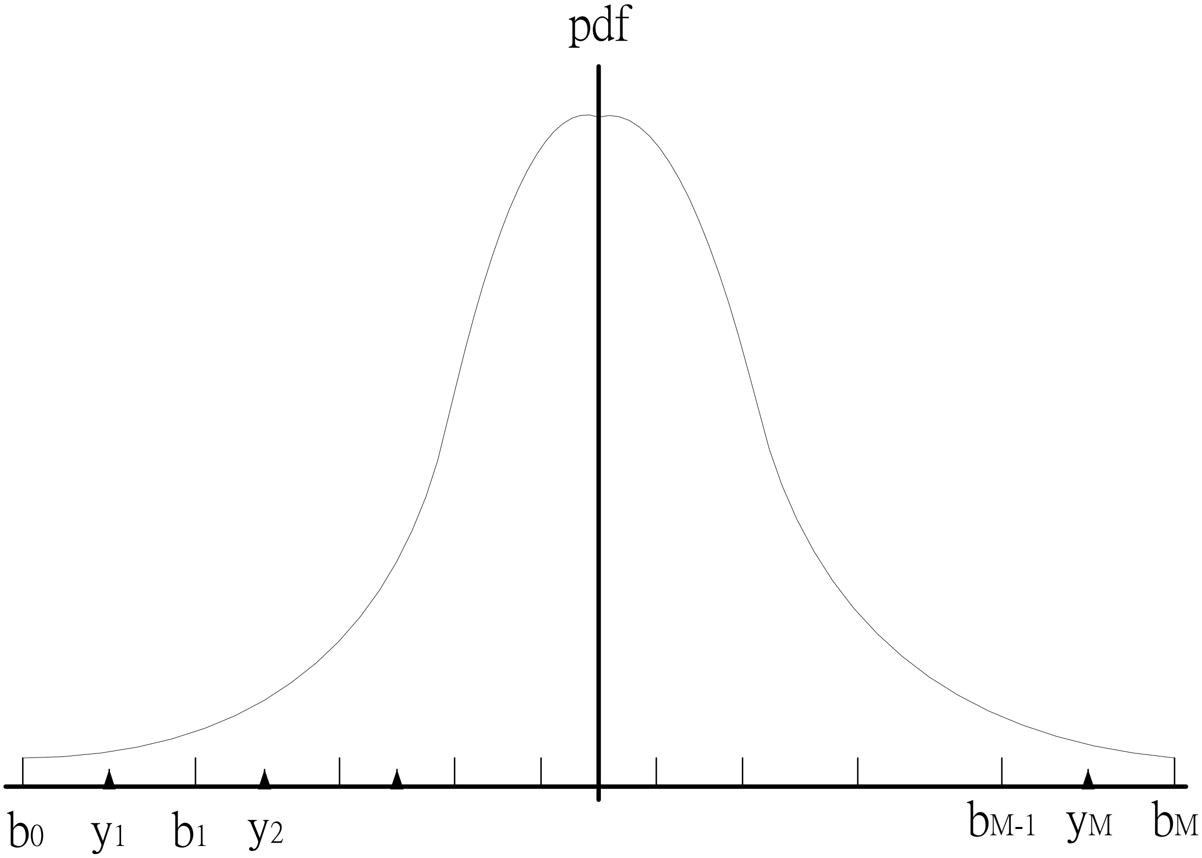

Figure 1: Nonuniform distribution of X.

Then,

$$(b_j - y_j)^2 = (b_j - y_{j+1})^2 \tag{11}$$
$$b_j - y_j = -(b_j - y_{j+1}) \tag{12}$$
$$\Rightarrow b_j = \frac{y_j + y_{j+1}}{2} \tag{13}$$

$$\Rightarrow y_{j+1} = 2b_j - y_j \tag{14}$$

The decision boundary is the midpoint of the two neighboring reconstruction levels. Solving these two equations (8) and (14) listed above will give us the values for the reconstruction levels and decision boundaries that minimize the mean squared quantization error. Unfortunately, to solve for y_j, we need the values of b_j and b_{j-1}, and to solve for b_j, we need the values of y_j and y_{j+1}. Therefore, the Lloyd-Max algorithm is introduced how to solve these two equations (8) and (14) iteratively.

Lloyd-Max Algorithm

Suppose $f_X(x)$ and $b_0 = -\alpha, b_M = +\alpha$ is given, Find $\{b_i\}_{i=0}^{M}$ and $\{y_i\}_{i=1}^{M}$. Assume a value for y_1, then
From (8), find b_1
From (14), find y_2
From (8), find b_2
From (14), find y_3
$\vdots$
From (8), find b_{M-1}
From (14), find y_M. Since we know $b_M = +\alpha$, we can directly compute $y'_M = \frac{\int_{b_{M-1}}^{b_M} x f_X(x)dx}{\int_{b_{M-1}}^{b_M} f_X(x)dx}$ and compare it with the previously computed value of y_M. If the difference is less than some tolerance threshold, we can stop. Otherwise, we adjust the estimate of y_1 in the direction indicated by the sign of the difference and repeat the procedure.

Properties of the Optimal Quantizer

The optimal quantizers have a number of interesting properties. We list these properties as follow:
1. Optimal quantizer must satisfy (8) and (14).
2. $EX = EQ(X)$
proof: Since $Q(X) = y_i$ iff $b_{i-1} < X \leq b_i$ and $Pr(Q(X) = y_i) = Pr(b_{i-1} < X \leq b_i)$, then

$$EQ(X) = \sum_{i=1}^{M} y_i Pr(Q(X) = y_i) \tag{15}$$

$$= \sum_{i=1}^{M} y_i Pr(b_{i-1} < X \leq b_i) \tag{16}$$

$$= \sum_{i=1}^{M} \frac{\int_{b_{i-1}}^{b_i} x f_X(x)dx}{\int_{b_{i-1}}^{b_i} f_X(x)dx} \int_{b_{i-1}}^{b_i} f_X(x)dx \tag{17}$$

$$= \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} x f_X(x)dx \tag{18}$$

$$= \int_{b_0}^{b_M} x f_X(x)dx \tag{19}$$

$$= \int_{-\infty}^{+\infty} x f_X(x)dx \tag{20}$$

$$= EX \tag{21}$$

The reason of (19) to (20) is that the value of $f_X(x)$ beyond b_0 and b_M is zero.

3. $EQ(X)^2 \leq EX^2$
proof: If $g_X(x) = f_X(x)/(\int_{b_{i-1}}^{b_i} f_X(x)dx)$, then $\int_{b_{i-1}}^{b_i} g_X(x)dx = 1$, $\int_{b_{i-1}}^{b_i} x g_X(x)dx = E_g X$, and $E_g(X - E_g X)^2 \geq 0 \Rightarrow (E_g X)^2 \leq E_g X^2$. Thus,

$$EQ(X)^2 = \sum_{i=1}^{M} y_i^2 Pr(Q(X) = y_i) \tag{22}$$

$$= \sum_{i=1}^{M} (\frac{\int_{b_{i-1}}^{b_i} x f_X(x)dx}{\int_{b_{i-1}}^{b_i} f_X(x)dx})^2 \int_{b_{i-1}}^{b_i} f_X(x)dx \tag{23}$$

$$= \sum_{i=1}^{M} (\int_{b_{i-1}}^{b_i} x \frac{f_X(x)}{\int_{b_{i-1}}^{b_i} f_X(x)dx} dx)^2 \int_{b_{i-1}}^{b_i} f_X(x)dx \tag{24}$$

$$\leq \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} x^2 \frac{f_X(x)}{\int_{b_{i-1}}^{b_i} f_X(x)dx} dx \int_{b_{i-1}}^{b_i} f_X(x)dx \tag{25}$$

$$= \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} x^2 f_X(x)dx \tag{26}$$

$$= \int_{-\infty}^{+\infty} x^2 f_X(x)dx \tag{27}$$

$$= EX^2 \tag{28}$$

4. $\sigma_q^2 = EX^2 - EQ(X)^2$

Lloyd Algorithm

The Lloyd algorith is another method to find $\{b_i\}_{i=0}^M$ and $\{y_i\}_{i=1}^M$. The distribution $f_X(x)$ is assumed known.
Assume $y_1^{(0)}, y_2^{(0)}, \cdots, y_M^{(0)}$ is an initial sequence of reconstruction values $\{y_i\}_{i=1}^M$. Select a threshold ϵ.
1.By Eqn (14). Find $b_0^{(1)}, b_1^{(1)}, \cdots, b_M^{(1)}$.
2.By Eqn (8). Find $y_1^{(1)}, y_2^{(1)}, \cdots, y_M^{(1)}$. And compute $\sigma_q^{2^{(1)}} = \sum_{i=1}^M \int_{b_{i-1}^{(1)}}^{b_i^{(1)}} (x - y_i^{(1)})^2 f_X(x)dx$.
3.By Eqn (14). Find $b_0^{(2)}, b_1^{(2)}, \cdots, b_M^{(2)}$.
4.By Eqn (8). Find $y_1^{(2)}, y_2^{(2)}, \cdots, y_M^{(2)}$. And compute$\sigma_q^{2^{(2)}} = \sum_{i=1}^M \int_{b_{i-1}^{(2)}}^{b_i^{(2)}} (x - y_i^{(2)})^2 f_X(x)dx$.
5.If $|\sigma_q^{2^{(2)}} - \sigma_q^{2^{(1)}}| = \begin{cases} < \epsilon, & \text{then} \quad stop \\ \geq \epsilon, & \text{then} \quad continue \quad the \quad procedure \end{cases}$

In summary, for each time j, the mean sqaured quantization error $\sigma_q^{2^{(j)}} = \sum_{i=1}^M \int_{b_{i-1}^{(j)}}^{b_i^{(j)}} (x - y_i^{(j)})^2 f_X(x)dx$ is calculated and compare it with previously error value $\sigma_q^{2^{(j-1)}}$. Stop iff $|\sigma_q^{2^{(j)}} - \sigma_q^{2^{(j-1)}}| < \epsilon$; otherwise, continue the same procedure of computing $b_0^{(j+1)}, b_1^{(j+1)}, \cdots, b_M^{(j+1)}$ and $y_1^{(j+1)}, y_2^{(j+1)}, \cdots, y_M^{(j+1)}$ by Eqn (14) and (8) for next time $j+1$.

Vector Quantization

The idea of vector quantization is that encoding sequences of outputs can provide an advantage over the encoding of individual samples. This indicates that a quantization strategy that works with sequences or blocks of outputs would provide some improvement in performance over scalar quantization. Here is an example. Suppose we have two uniform random variables height $X_1 \sim Unif[40, 80]$ and weight $X_2 \sim Unif[40, 240]$ and 3 bits are allowed to represent each random variable. Thus, the weight range is divided into 8 equal intervals and with reconstruction levels $\{52, 77, \cdots, 227\}$; the height range is divided into 8 equal intervals and with reconstruction levels $\{42, 47, \cdots, 77\}$. The two dimensional representation of these two quantizers is shown in Figure 2(a).

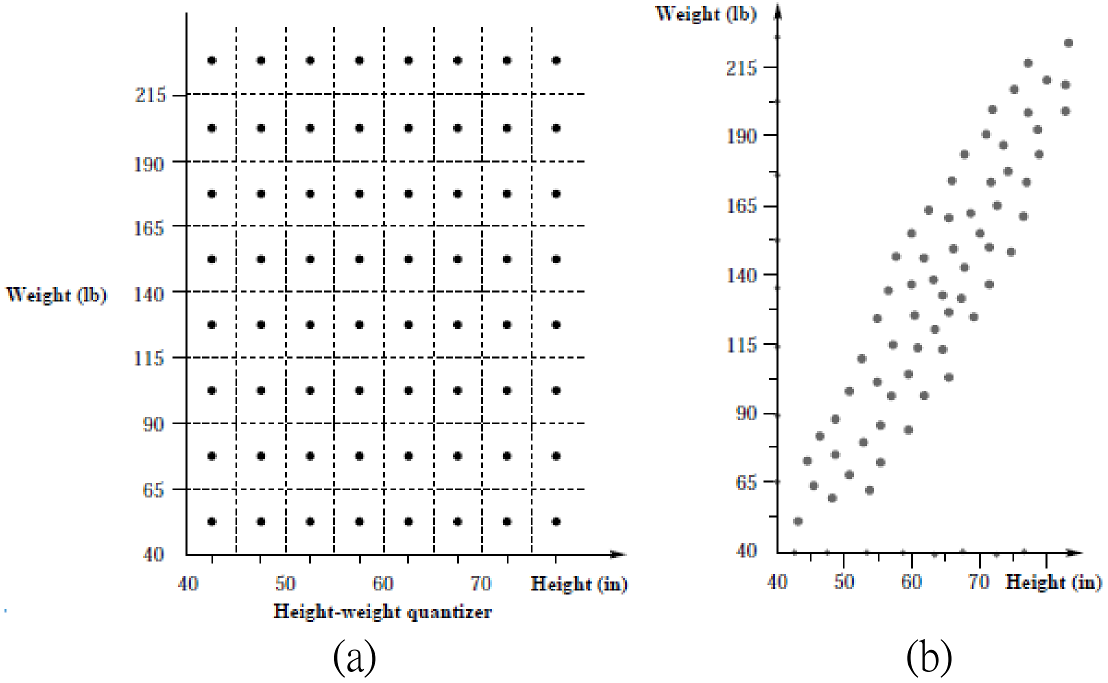

Figure 2: (a)The dimensions of the height/weight scalar quantization. (b)The height-weight vector quantization

However, the height and weight are correlated. For example, a quantizer for a person who is 80 inches tall and weights 40 pounds or who is 42 inches tall and weights 200 pounds is never used. A more sensible approach will use a quantizer like the one shown in Figure 2(b). Using this quantizer, we can no longer quantize the height and weight separately. We will consider them as the coordinates of a point in two dimensions in order to find the closest quantizer output point.