Rate and Distortion Models Derivation Process in Detail

Figure 1 shows the temporal coding structure of our lossy compaction scheme. There are total five temporal levels shown in Figure 1, where each increasing temporal level corresponds to a double of frequency at which collections of sensor data at certain time instance are included in lossy coded data set. Thus, the highest temporal level shall contain all of data sampled in line with temporal dimension. Figure 1 also displays the temporal prediction structure shown by arrows.

**Figure 1.** Temporal coding and prediction structure of our lossy compaction scheme.

It is quite intuitive to reckon that the size of compressed data $R$ is reduced by half as the temporal level decreases by one step. However, due to the temporal prediction structure shown in Figure 1, the amount of reduction becomes less than half per one temporal level decrease. We can model this relation as:

$$ R = \alpha(\Delta) \cdot \exp(\beta(\Delta) \cdot T), $$

where $\alpha(\Delta)$ and $\beta(\Delta)$ are model parameters dependent on the quantization step size $\Delta$, and $T \in \{0, 1, 2, 3, 4\}$ denotes the temporal level. Comparison between actual data size and the model in (1) with a fixed $\Delta$ is shown in Figure 2a, where we can confirm the model effectively follows the varying size of actual sensor data with respect to the temporal level.

In (1), two model parameters $\alpha(\Delta)$ and $\beta(\Delta)$ have to be estimated from the real data based on the quantization step size, which are represented by

$$ \alpha(\Delta) = a_\alpha \cdot \exp(b_\alpha \cdot \Delta) + c_\alpha \cdot \exp(d_\alpha \cdot \Delta), $$
$$ \beta(\Delta) = a_\beta \cdot \exp(b_\beta \cdot \Delta) + c_\beta, $$

where $a_\alpha$, $b_\alpha$, $c_\alpha$, and $d_\alpha$ are data-dependent constants supplementary to $\alpha(\Delta)$ in (1), and similarly, $a_\beta$, $b_\beta$, and $c_\beta$ are constants for $\beta(\Delta)$ in (1). It should be noted that $b_\alpha$ and $d_\alpha$ in (2) and $b_\beta$ in (3) are all negative valued parameters that reflect decay of $\alpha(\Delta)$ and $\beta(\Delta)$ with an increasing $\Delta$. Combining (2) and (3) with (1), we can represent the total rate as a function of both the quantization step and the temporal level. The resulting model function is plotted in Figure 2b, where five lines represent each temporal levels and actual data points are also plotted for comparison.
Figure 2. (a) Rate curve as the function of temporal level estimated by (1) with \(QP=20\).
(b) Rate curves as the function of quantization step size for different temporal levels estimated by (1).

\[
D_{\text{quant}} = a_{\text{quant}} \cdot \exp(b_{\text{quant}} \cdot QP) + c_{\text{quant}},
\]

where \(a_{\text{quant}}\), \(b_{\text{quant}}\), and \(c_{\text{quant}}\) are data-dependent constants. It should be noted that (4) is a function of \(QP\), whose relationship with the quantization step size \(\Delta\) can be expressed by \(\Delta = 0.625 \cdot 2^{QP/6}\). Figure 3 shows actual distortion points and its approximation using (4).

Figure 3. Distortion curve as the function of QP estimated by (4).
Although (4) effectively models the distortion caused by quantization, the source of distortion is not limited to the quantization. As the temporal level $T$ varies, the amount of sampled data along temporal dimension varies as well, which causes another distortion. Recalling the temporal coding structure shown in Figure 1, as $T$ decreases by one step, half of data are excluded from data set, which leads to the condition that omitted data should be estimated using previous data samples. As a result, the total distortion increases as $T$ decreases.

In order to incorporate the temporal distortion into total distortion along with the quantization distortion, we assume that temporal distortion is measured by mismatch between actual data samples and omitted data samples that are replaced by previous data samples. Although the combination of these two different types of distortion seems tightly coupled, they can be separated as proven by the following lemma.

**Lemma 1:** The joint distortion $D_{total}$ caused by the quantization from lossy coding and the omission of data samples along temporal dimension is separable and can be expressed by sum of both distortions.

**Proof:** First we assume an arbitrary pdf of distance between actual data samples and reconstructed data samples, in which missing samples are covered by previous existing data samples. This pdf is denoted by $f_{e_L}(e_T)$, where random variable $E_T$ reflects the near continuity of distance between data samples.

On the other hand, the quantization step size $\Delta$ is controlled by QP of a lossy compaction scheme that usually performs quantization operation in discrete cosine transform (DCT) domain. However it is well known that in an ideal encoder-decoder system, spatial-domain distortion and DCT-domain distortion are equal, which enables us to render the probability mass function (pmf) of quantization error from lossy coding as in Figure 4. Assuming $\Delta$ is an odd number without loss of generality, we can express the pmf of the quantization error from lossy coding as follows:

$$P_{e_L}(e_L) = \begin{cases} \frac{1}{\Delta} & e_L = -\left\lfloor \frac{\Delta}{2} \right\rfloor, -\left\lfloor \frac{\Delta}{2} \right\rfloor + 1, ..., \left\lfloor \frac{\Delta}{2} \right\rfloor, \\ 0 & \text{otherwise} \end{cases}$$

(5)
where $E_L$ is a discrete random variable that denotes an amount of quantization error in integer domain.

**Figure 4.** Pmf of quantization error from lossy coding.

When a specific $E_T$ is given by $e_T$, the conditional pmf of the quantization error from lossy coding is given by

$$
P_{e_L|e_T}(e_L | e_T) = \begin{cases} \frac{1}{\Delta} & e_L = e_T - \left\lfloor \frac{\Delta}{2} \right\rfloor, e_T - \left\lfloor \frac{\Delta}{2} \right\rfloor + 1, \ldots, e_T + \left\lfloor \frac{\Delta}{2} \right\rfloor, \\ 0 & \text{otherwise} \end{cases}$$

which indicates that the pmf shown in Figure 4 can be shifted to left or right according to given $E_T$.

We can express $D_{\text{total}}$ using joint distribution:

$$D_{\text{total}} = \int \sum_{e_T} f_{E_T,E_L}(e_T, e_L) \cdot e_L^2 \cdot de_T. \quad (7)$$

Then we have

$$D_{\text{total}} = \int \sum_{e_T - \left\lfloor \frac{\Delta}{2} \right\rfloor}^{\left\lceil \frac{\Delta}{2} \right\rceil} f_{E_T}(e_T) \cdot P_{e_L|e_T}(e_L | e_T) \cdot (e_L + e_T)^2 \cdot de_T = \int f_{E_T}(e_T) \cdot \frac{1}{\Delta} \cdot \sum_{e_L - \left\lfloor \frac{\Delta}{2} \right\rceil}^{\left\lceil \frac{\Delta}{2} \right\rceil} (e_L + e_T)^2 \cdot de_T, \quad (8)$$

which continues in

$$D_{\text{total}} = \int_{-\infty}^{\infty} f_{E_T}(e_T) \cdot \left(e_T^2 + \frac{\Delta^2 - 1}{12}\right) \cdot de_T \approx \int_{-\infty}^{\infty} e_T^2 \cdot f_{E_T}(e_T) \cdot de_T + \frac{\Delta^2 - 1}{\beta}, \quad (9)$$
where $\beta$ is a denominator which is larger than 12 in case of a larger quantization step size compared to the signal variance. This is because when the quantization step size becomes large, quantization errors can no longer be treated as uniformly distributed, where $\beta$ is 12.

In the right-hand side of (9), the first term is the distortion in temporal dimension and the second term is the lossy coding distortion.

Using the above lemma, the total distortion can be simply expressed by summing distortions from two different sources. Now we turn to the problem of estimating the temporal distortion model. Specifically, we have found that the temporal distortion $D_{\text{temp}}$ is a linear function of the temporal level $T$, which is given by

$$D_{\text{temp}} = a_{\text{temp}} \cdot T + b_{\text{temp}},$$

(10)

where $a_{\text{temp}}$ and $b_{\text{temp}}$ are constants. The accuracy of (10) can also be verified by Figure 5.

**Figure 5.** Temporal distortion as the function of temporal level estimated by (10).

Thanks to the separation property proven in the lemma 1, we can combine both distortions in (4) and (10) to yield the joint distortion $D_{\text{total}}$ as follows:

$$D_{\text{total}}(QP, T) = D_{\text{quant}} + D_{\text{temp}} = a_{\text{quant}} \cdot \exp(b_{\text{quant}} \cdot QP) + a_{\text{temp}} \cdot T + a_{\text{total}},$$

(11)

where $c_{\text{quant}}$ in (4) and $b_{\text{temp}}$ in (10) are absorbed into one constant $a_{\text{total}}$. 