

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Calculating Method of Rate Distortion Function for Binary Source

Zhang YunPeng

College of electric power, North China University of Water Resources and Electric Power,
 450011, Zhengzhou, China

Abstract: Rate-distortion function is the basic information theory which provides the theoretical foundations for lossy data compression; it addresses the problem of the minimal number of bits per symbol given a random variable source and a certain distortion measure. Its calculation is a complex problem. Traditional calculation method is to use the expression parameters. But it is very difficult to calculate by using the expression parameters and especially, the inequality constraints are the biggest obstacle. In general, the analytical expression is difficult to obtain. This study presents a novel method to calculate the rate-distortion function of the binary source under the condition of Hamming distortion based on the nature of rate-distortion function and extremes of conditional entropy. This method avoids the complex calculations, greatly simplifies the derivation and obtains the analytical expressions for the rate-distortion function of binary symbols.

Key words: Rate-distortion function, binary source, conditional entropy

INTRODUCTION

In a communication system, it is always desirable to transmit the information rate as small as possible, under certain distortion measure (Muramatsu and Kanaya, 1994; Yang and Shen, 1993), that is, to seek the minimum rate of information transfer in certain given conditions. At this point the minimum is related to the limit of distortion. From the receiving terminal point of view, the minimum average amount of information for the reproduced source information has to be found under the condition of meeting the fidelity criterion. Because the amount of average mutual information can be used to represent the amount of information receiving terminals obtain, this problem is transformed into seeking the minimum of the average mutual information in the upper bound of the allowed distortion constraint conditions. The average mutual information is concave function, so the minimum must exist. The minimum value is the rate-distortion function (Shannon, 1948).

DISTORTION FUNCTION AND THE DISTORTION MATRIX

Distortion measure and distortion matrix: Let binary source symbols X , with probability space Eq. 1:

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ p(x_1) & p(x_2) \end{bmatrix} \quad (1)$$

X transmitted to the receiving terminal of the symbol Y with probability space Eq. 2:

$$\begin{bmatrix} Y \\ P(Y) \end{bmatrix} = \begin{bmatrix} y_1 & y_2 \\ p(y_1) & p(y_2) \end{bmatrix}$$

Channel transfer matrix is described by Eq. 3:

$$[P(Y|X)] = \begin{bmatrix} p(y_1|x_1) & p(y_2|x_1) \\ p(y_1|x_2) & p(y_2|x_2) \end{bmatrix} \quad (3)$$

Distortion measure is for each pair (x_i, y_j) , to specify a non-negative function $d(x_i, y_j) \geq 0$, $i = 1, 2$, $j = 1, 2$ namely $d(x_i, y_j)$ as a single symbol distortion measure. Distortion matrix can be expressed as Eq. 4.

$$[D] = \begin{bmatrix} d(x_1, y_1) & d(x_1, y_2) \\ d(x_2, y_1) & d(x_2, y_2) \end{bmatrix} \quad (4)$$

When the distortion measure is the Hamming, that:

$$i = j, d(x_i, y_j) = 0$$

$$i \neq j, d(x_i, y_j) = a, a > 0$$

The average measure of distortion and fidelity criterion: The average distortion measure is the mathematical expectation of distortion measure (Gibson et al., 2010), as Eq. 5:

$$\bar{D} = E(x_i, y_j) = \sum_{i=1}^2 \sum_{j=1}^2 P(x_i) P(y_j|x_i) d(x_i, y_j) \quad (5)$$

The average distortion measure \bar{D} is distortion description of the entire communication system. In other words, \bar{D} is the distortion per symbol. Here, the analysis is based on this average distortion measure between sequences.

When $p(x_i)$, $p(y_j|x_i)$ and $d(x_i, y_j)$ are given, the average distortion measure is no longer a random variable, but the amount determined.

D is the upper bound of distortion allowed. $\bar{D} \leq D$ is called fidelity criterion. In Eq. 5 only part of the channel to meet the fidelity criterion. The fidelity criterion as the channel transition probability constraint, to seek a channel information rate has a minimum practical significance (Ziv, 1980).

Rate distortion function R (D): Given source and specifically define the distortion function, communication systems must meet D . The channels of transmission information can be divided into two parts, which are constituted of the two sets:

$$P_D = p\{y_j | x_i, \bar{D} \leq D\} \quad (6)$$

$$P_{\bar{D}} = p\{y_j | x_i, \bar{D} > D\} \quad (7)$$

Among them, the set of P_D meet the criteria channels. Thus, the provision of the distortion measure, the channel is subject to a restriction that: P_D , so that we can calculate $R = I(X, Y)$ reaches a minimum.

A rate distortion function (Berger, 1972) is defined as Eq. 8:

$$R(D) = \min_{p(y_j|x_i) \in P_D} \{I(X; Y)\} \quad (8)$$

$$= \{I(X; Y); \bar{D} \leq D\}$$

Given a random variable and a distortion measure, what is the minimum expected distortion achievable at a particular rate?

The maximum value of the conditional entropy: Conditional entropy, also known as channel equivocation or losses entropy refers to the received symbol Y , the uncertainty of the source of X .

Let the random variable X, Y values in the same set of symbols $\{\alpha_1, \alpha_2\}$, random variable:

$$Z = \begin{cases} 0 & X = Y \\ 1 & X \neq Y \end{cases}$$

Let:

$$f(z_k) = \sum_{i=1}^2 \sum_{j=1}^2 p(y_j) p(z_k | x_i, y_j) \quad k=1, 2$$

and:

$$p(X \neq Y) = p_e, p(X = Y) = 1 - p_e$$

Where:

$$H(p_e) = p_e \ln \frac{1}{p_e} + (1 - p_e) \ln \frac{1}{1 - p_e}$$

$$H(Z) = - \sum_{k=1}^2 p(z_k) \ln p(z_k)$$

$$H(X|Y) = - \sum_{i=1}^2 \sum_{j=1}^2 p(x_i, y_j) \ln p(x_i | y_j)$$

$p(\bullet)$ represents the probability.

Due to:

$$p(x_i, y_j) = \sum_{k=1}^2 p(x_i, y_j, z_k)$$

$$\begin{aligned} H(X|Y) &= \sum_{i=1}^2 \sum_{j=1}^2 p(x_i, y_j) \ln \frac{1}{p(x_i | y_j)} \\ &= \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 p(x_i, y_j, z_k) \ln \frac{1}{p(x_i | y_j)} \\ &= \sum_{k=1}^2 p(z_k) \sum_{i=1}^2 \sum_{j=1}^2 \frac{p(x_i, y_j, z_k)}{p(z_k)} \ln \frac{1}{p(x_i | y_j)} \end{aligned}$$

logarithmic function $\ln(\bullet)$ is convex function (Urruty *et al.*, 2004) and:

$$\frac{p(x_i, y_j, z_k)}{p(z_k)} = p(x_i, y_j | z_k)$$

Then:

$$\sum_{i=1}^2 \sum_{j=1}^2 \frac{p(x_i, y_j, z_k)}{p(z_k)} \ln \frac{1}{p(x_i | y_j)} \leq$$

$$\ln \sum_{i=1}^2 \sum_{j=1}^2 \frac{1}{p(z_k)} \frac{p(x_i, y_j, z_k)}{p(x_i | y_j)}$$

$$\begin{aligned} H(X|Y) &\leq \sum_{k=1}^2 p(z_k) \ln \frac{1}{p(z_k)} \sum_{i=1}^2 \sum_{j=1}^2 \frac{p(x_i, y_j, z_k)}{p(x_i | y_j)} \\ &= \sum_{k=1}^2 p(z_k) \ln \frac{1}{p(z_k)} + \end{aligned}$$

Thereby:

$$\begin{aligned} & \sum_{k=1}^2 p(z_k) \ln \frac{1}{p(z_k)} \sum_{i=1}^2 \sum_{j=1}^2 \frac{p(x_i y_j z_k)}{p(x_i y_j)} p(y_j) \\ & = H(Z) + \sum_{k=1}^2 p(z) \ln f(z) = H(Z) + E[f(z)] \end{aligned}$$

Here, $E[\bullet]$ represents the mathematical expectation:

$$\begin{aligned} f(0) &= \sum_{i=1}^2 \sum_{j=1}^2 p(y_j) p(z=0|x_i, y_j) \\ &= \sum_{j=1}^2 p(y_j) \sum_{i=1}^2 p(z=0|x_i, y_j) \\ &= \sum_{j=1}^2 p(y_j) \sum_{X=Y} p(z=0|X=Y) + \\ & \quad \sum_{j=1}^2 p(y_j) \sum_{X \neq Y} p(z=0|X \neq Y) \end{aligned}$$

Clearly, when $X = Y$:

$$\begin{aligned} p(z=0|X=Y) &= 1, \\ X \neq Y, p(z=0|X \neq Y) &= 0 \end{aligned}$$

Therefore:

$$f(0) = 1$$

Similarly:

$$p(z=1|X=Y) = 0, p(z=1|X \neq Y) = 1$$

Consequently:

$$\begin{aligned} f(1) &= 1 \\ H(Z) &= p_e \ln \frac{1}{p_e} + (1-p_e) \ln \frac{1}{1-p_e} \\ E[f(z)] &= (1-p_e) \ln f(0) + p_e \ln f(1) \\ &= (1-p_e) \ln 1 + p_e \ln 1 = 0 \\ H(X|Y) &\leq p_e \ln \frac{1}{p_e} + (1-p_e) \ln \frac{1}{1-p_e} = H(p_e) \end{aligned} \quad (9)$$

$H(p_e)$ is the upper limit of the conditional entropy. In other words, the maximum value of the conditional entropy is $H(p_e)$.

THE RATE DISTORTION FUNCTION OF BINARY SYMBOL

Let the information source and sink are binary symbols, namely $X, Y \in \{0, 1\}$:

$$\begin{aligned} p(x=0) &= 1-p, \\ p(x=1) &= p \end{aligned}$$

Hamming distortion:

$$d(x_i, y_j) = \begin{cases} 0, & x_i = y_j \\ a, & x_i \neq y_j \end{cases} \quad (10)$$

The traditional solution: The classical method of calculating the rate-distortion function is for discrete information source to choose Lagrange multiplier method (Haghighat *et al.*, 2011; Peng *et al.*, 2005; Hoang *et al.*, 1998).

The average mutual information as defined in Eq. 11:

$$I(X; Y) = \sum_{i=1}^2 \sum_{j=1}^2 P(x_i) P(y_j | x_i) \ln \frac{P(y_j | x_i)}{P(y_j)} \quad (11)$$

Under the constraint in the Eq. 5, 11, 12, 13:

$$\bar{D} < D \quad (12)$$

$$\sum_{j=1}^2 P(y_j | x_i) = 1 \quad (13)$$

$$P(y_j | x_i) \geq 0 \quad (14)$$

Find the minimum average mutual information. Lagrange multipliers s and $\mu_i (i=1, 2, \dots, n)$ to construct the auxiliary functions, the auxiliary function would not consider (11) (13) constraint.

$$\begin{aligned} \Phi &= I(X; Y) - s \left[\sum_{i=1}^2 \sum_{j=1}^2 p(x_i) p(y_j | x_i) d(x_i, y_j) - \right. \\ & \quad \left. \bar{D} \right] - \mu_i [p(y_j | x_i) - 1] \end{aligned} \quad (15)$$

Makes the first order partial derivatives equal to zero, obviously, the solution process more complex and requires a lot of calculation and derivation, because did not put Eq. 11 and 13 constraints into account, the obtained expression does not necessarily meet the requirements, but also need to verify whether the conclusions meet the actual conditions.

Analytical expression of the rate-distortion function: Generally, calculating rate-distortion function $R(D)$ involves three steps:

- Determine the domain of the lower D_{\min} and $R(D_{\min})$
- Determine the upper limit of the domain D_{\max} and $R(D_{\max})$

- Determine the analytical of the domain
 $D_{\min} < D < D_{\max}$

For a given source and binary Hamming distortion function can be determined:

- $$D_{\min} = \sum_{i=1}^2 p(x_i) \min_j \{d(x_i, y_j)\}$$

$$= p(x_1) \min\{d(x_1, y_1), d(x_1, y_2)\} +$$

$$p(x_2) \min\{d(x_2, y_1), d(x_2, y_2)\} = 0$$

So:

$$R(0) = H(X) = p \ln \frac{1}{p} + (1-p) \ln \frac{1}{1-p}$$

- $$D_{\max} = \min_j \left\{ \sum_{i=1}^2 p(x_i) d(x_i, y_j) \right\}$$

$$= \min \left\{ \sum_{i=1}^2 p(x_i) d(x_i, y_1), \sum_{i=1}^2 p(x_i) d(x_i, y_2) \right\}$$

$$= \min\{a(1-p), ap\} \left(p \leq \frac{1}{2}\right) = ap$$

So:

$$R(ap) = 0$$

- When

$$D_{\min} < D < D_{\max}$$

$$R(D) = \min\{I(X; Y); \bar{D} \leq D\}$$

According to the nature of the rate-distortion function, $R(D)$ is a monotonically decreasing function (Sow and Eleftheriadis, 2003; Cover and Thomas, 1991).

So:

$$R(D) = \min\{I(X; Y); \bar{D} = D\}$$

$$\bar{D} = \sum_{i=1}^2 \sum_{j=1}^2 p(x_i) p(y_j | x_i) d(x_i, y_j)$$

$$= \sum_{x_i \neq y_j} p(x_i) p(y_j | x_i) d(x_i, y_j) +$$

$$\sum_{x_i = y_j} p(x_i) p(y_j | x_i) d(x_i, y_j)$$

$$= \sum_{x_i \neq y_j} p(x_i) p(y_j | x_i) d(x_i, y_j)$$

$$\text{(When } x_i = y_j, d(x_i, y_j) = 0)$$

$$= a \sum_{i=1}^2 \sum_{j=1}^2 \{p(x_i) p(y_j | x_i) : x_i \neq y_j\}$$

$$= ap(x_i \neq y_j) = ap_e = D$$

So:

$$p_e = \frac{D}{a}$$

Because of:

$$H(X|Y) \leq H(p_e) = H\left(\frac{D}{a}\right)$$

$$I(X; Y) = H(X) - H(X|Y) \geq H(p) - H\left(\frac{D}{a}\right)$$

$$\min\{I(X; Y)\} = H(p) - H\left(\frac{D}{a}\right)$$

Therefore, there must be a channel, such that the average distortion $\bar{D} = D$. When $a = 1$, these statistical characteristics of channel are:

$$p(x=0|y=0) = 1-D,$$

$$p(x=1|y=0) = D,$$

$$p(x=0|y=1) = D,$$

$$p(x=1|y=1) = 1-D$$

Let:

$$p(y=0) = \alpha$$

$$p(y=1) = \beta$$

So:

$$p(x=0) = p(y=0)p(x=0|y=0) +$$

$$p(y=1)p(x=0|y=1)$$

$$p(x=1) = p(y=0)p(x=1|y=0) +$$

$$p(y=1)p(x=1|y=1)$$

That is:

$$\begin{cases} \alpha(1-D) + \beta D = p \\ \alpha D + (1-D)\beta = 1-p \end{cases}$$

$$\alpha = \frac{p-D}{1-2D}$$

$$\beta = \frac{1-p-D}{1-2D}$$

If:

$$p \leq \frac{1}{2}$$

Then:

$$0 \leq D \leq D_{\max} = p$$

So:

$$\alpha + \beta = \frac{p-D}{1-2D} + \frac{1-p-D}{1-2D} = 1$$

This shows that the channel exists.

In addition, the average distortion:

$$\begin{aligned} \bar{D} &= E[d(x_i, y_j)] \\ &= \sum_{i=1}^2 \sum_{j=1}^2 p(x_i, y_j) d(x_i, y_j) \\ &= \sum_{i=1}^2 \sum_{j=1}^2 p(y_j) p(x_i | y_j) d(x_i, y_j) \\ &= \frac{1-p-D}{1-2D} \cdot D + \frac{p-D}{1-2D} \cdot D = D \end{aligned}$$

$$I(X, Y) = H(X) - H(X|Y)$$

$$\begin{aligned} H(X|Y) &= \sum_{i=1}^2 \sum_{j=1}^2 p(y_j) p(x_i | y_j) \log \frac{1}{p(x_i | y_j)} \\ &= p(y_1) \sum_{i=1}^2 p(x_i | y_1) \log \frac{1}{p(x_i | y_1)} + \\ &\quad p(y_2) \sum_{i=1}^2 p(x_i | y_2) \log \frac{1}{p(x_i | y_2)} \\ &= p(y_1) H(D) + p(y_2) H(D) = H(D) \end{aligned}$$

$$I(X, Y) = H(X) - H(X|Y) = H(p) - H(D)$$

$$R(D) = H(p) - H(D) \quad 0 \leq D \leq p \quad (16)$$

$$\begin{aligned} R(D) &= p \ln \frac{1}{p} + (1-p) \ln \frac{1}{1-p} - D \ln \frac{1}{D} - \\ &\quad (1-D) \ln \frac{1}{1-D} \quad 0 \leq D \leq p \end{aligned} \quad (17)$$

Equation 17 is the analytical expression of rate-distortion function.

CONCLUSION

The Calculating of Rate Distortion Function using partial derivative is very complex. It requires a lot of computation and derivation and the resulting expression does not necessarily meet the actual application requirements. If the result does not match the actual situation, it needs to further adjust parameters and re-evaluate the expression until all of the conditions have been met. This study applies the extreme of conditional entropy, the monotonically decreasing of rate-distortion function and fidelity criteria to solving analytic expression of the rate-distortion function of the binary in the conditions of Hamming distortion measure. This calculation process is simple, greatly saving time of the calculation and derivations.

ACKNOWLEDGMENTS

We thank Professor Xiang and Dr. Sun for helpful comments on drafts of this manuscript. We thank the Project supported by the Major Research plan of the National Natural Science Foundation of China (Grant No. 51190093) for funding this study.

REFERENCES

- Berger, T., 1972. Rate Distortion Theory: A Mathematical Basis for Data Compression. Prentice-Hall, Englewood Cliffs, USA., ISBN: 9780137531035.
- Cover, T.M. and J.A. Thomas, 1991. Elements of Information Theory. Wiley, New York.
- Gibson, J.D., H. Jing and P. Ramadas, 2010. New rate distortion bounds for speech coding based on composite source models. Inform. Theory Appl. Workshop, 31: 1-5.
- Haghighat, M.B.A., A. Aghagolzadeh and H. Seyedarabi, 2011. A non-reference image fusion metric based on mutual information of image features. Comput. Electric. Eng., 37: 744-756.
- Hoang, D.T., P.M. Long and J. Vitter, 1998. Rate-distortion optimizations for motion estimation in low-bitrate video coding. IEEE Trans. Circu. Syst. Video Technol., 8: 488-500.
- Muramatsu, J. and F. Kanaya, 1994. Distortion-complexity and rate-distortion function. Proceedings of the IEEE International Symposium on Information Theory, Jun 27-July 1, 1994, Trondheim, pp: 1224-1229.
- Peng, H., F. Long and C. Ding, 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance and min-redundancy. IEEE Tran. Pattern Anal. Mach. Intell., 27: 1226-1238.

- Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27: 379-423.
- Sow, D.M. and A. Eleftheriadis, 2003. Complexity distortion theory. *IEEE Trans. Inform. Theory*, 49: 604-608.
- Urruty, H., J. Baptiste and C. Lemarechal, 2004. *Fundamentals of Convex Analysis*. Springer-Verlag, Berlin.
- Yang, E.H. and S.Y. Shen, 1993. Distortion program-size complexity with respect to a fidelity criterion and rate-distortion function. *IEEE Trans. Inform. Theory*, 39: 288-292.
- Ziv, J., 1980. Distortion-rate theory for individual sequences. *IEEE Trans. Inform. Theory*, 26: 137-143.