

$R(D)$ Definition Domain Deriving of Limited Distortion Source Coding With Different Sources

Chunhua Zhu, Huan Mi, Mengfei Kan, Yi Song

School of Information Science and Engineering, Henan University of Technology, Zhengzhou, China

Email: zhuchunhua@haut.edu.cn

How to cite this paper: Zhu, C.H., Mi, H., Kan, M.F. and Song, Y. (2017) $R(D)$ Definition Domain Deriving of Limited Distortion Source Coding With Different Sources. *Int. J. Communications, Network and System Sciences*, 10, 263-268.
<https://doi.org/10.4236/ijcns.2017.108B028>

Received: July 3, 2017

Accepted: August 11, 2017

Published: August 14, 2017

Abstract

For limited distortion source coding, it is generally considered that the minimum value of the coding average distortion is 0, and the maximum value is the minimum distortion value of making $R(D) = 0$. This is the definition domain of the information rate distortion function. In this paper, the upper and lower bounds of the information rate distortion function $R(D)$ are derived and computed for the typical sources. The results show that the lower bound of the coding average distortion D is related to the symbol distortion function, which can further improve the theory of limited distortion source coding.

Keywords

The Information Rate Distortion Function, Limited Distortion Coding, Definition Domain

1. Introduction

From the point of view of information processing, no distortion source coding is entropy preserving. And the entropy preserving coding is not always necessary. Such as the human eyes to accept the visual signal, then there is no need to carry out the distortion-free entropy coding. But entropy preserving coding is not always possible. For example, when the continuous signal is subjected to digital processing, it is impossible to fundamentally remove the quantization error. At this point $\lim_{D \rightarrow 0} R(D) \rightarrow \infty$. Although the relative entropy of the continuous source is limited, the amount of information is infinite [1]. But the actual channel always has interference, and its capacity is limited, so the distortionless transmission of this continuous information is impossible. Only when certain distortion is allowed, the value of $R(D)$ is finite, and the transmission is possible.

Reducing the rate of information is beneficial to transmission and processing, so it is necessary to perform entropy compression coding. Therefore, it is necessary to introduce the source coding with limited distortion. $R(D)$ is a very important parameter in the source code is the information rate distortion function [2] [3]. In the case of allowing a certain degree of distortion, the rate-distortion function of $R(D)$ the source can be used as a measure of the performance of various compression coding methods. It is generally assumed that the minimum value D_{\min} of the coding average distortion is 0 and the maximum value D_{\max} is the minimum distortion value of $R(D)=0$, that is the definition field of the $R(D)$ function is $[0, D_{\max}]$. In fact, in many cases, D_{\min} is not necessarily zero [4] [5], its value is related to the single symbol distortion function. Only when the distortion matrix has at least one zero element in each row, the average distortion of the source can reach zero. In this paper, the upper and lower bounds of the definition domain of $R(D)$ are deduced, and the upper and lower bounds of the definition domain of $R(D)$ under different source and different distortion function are given. In this way, we can further improve the limitless source coding theorem.

2. The Infer of Domain Definition of the Information Rate Distortion Function $R(D)$

The information rate distortion function $R(D)$ independent variable D is the average distortion caused by a limited information source coding algorithm. The average distortion allowed is the upper limit of the stipulated average distortion D . The definition domain question of information rate distortion function is that the minimum and maximum values of average distortion D are studied when the information source and the distortion function are known. The value of D must satisfy the fidelity criterion

$$\bar{D} \leq D \tag{1}$$

The average distortion \bar{D} is determined according to the statistical properties of a information source X , Statistical property $P(X)$, coding mapping function $p(y_j/x_i)$ and distortion function $d(x_i, y_j)$, that is [1]

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

2.1. The Defined Domain Lower Bound of Information Rate Distortion Function $R(D)$

The mathematical expectation of nonnegative real numbers $d(x_i, y_j)$ is \bar{D} , therefore, \bar{D} is a nonnegative real number, with a lower limit of 0, so that the permissible lower bound of the average distortion D must be 0, that is $D_{\min} = 0$, this is the case where no any distortion is allowed. As for whether the permissible average distortion can reach its lower limit value of 0, it is related to the distortion function of a single symbol. When solving the minimum mean distortion D_{\min} of the information source, for each x_i , find a corresponding

y_j that makes the minimum $d(x_i, y_j)$, and the corresponding $d(x_i, y_j)$ minimum values for different x_i are also different. This is equivalent to finding a smallest $d(x_i, y_j)$ in each line of the distortion matrix, and the minimum $d(x_i, y_j)$ of each line is different. The mathematical expectation of all these different minima $d(x_i, y_j)$ is the minimum mean distortion of the information source [1], that is

$$D_{\min} = \sum_{i=1}^n p(x_i) \min d(x_i, y_j) \quad (3)$$

It can be seen by (3) formula, When each line of the distortion matrix has at least one zero element, the average distortion of the information source can reach the lower bound value, otherwise D_{\min} must not be equal to zero. When $D_{\min} = 0$, that is, the information source does not allow any distortion to exist, which is equivalent to noiseless channels. At this point, the amount of information transmitted by the channel is equal to the entropy of the information source. That is

$$R(D_{\min}) = R(0) = H(X) \quad (4)$$

If the formula (4) is established, each line in the distortion matrix has at least one zero, and each column can have at most one zero. Otherwise, $R(0)$ can be less than $H(X)$, which means that the information source symbols are concentrated, and some symbols can be compressed and merged without any distortion.

2.2. The Defined Domain Upper Bound of Information Rate Distortion Function D_{\max}

For the maximum average distortion of the information source D_{\max} , the smaller the required compression rate, the greater the tolerance distortion. When $R(D)=0$, the corresponding average distortion is maximal, D_{\max} is the upper bound of the domain of the function $R(D)$ definition. Since the information rate distortion function $R(D)$ is the minimum of average mutual information $I(X;Y)$. So when $R(D)=0$, the minimum value of the average mutual information at that time is equal to 0, this moment $D = D_{\max}$. Because $I(X;Y)$ is a nonnegative function, $R(D)$ is also a nonnegative function. When $D > D_{\max}$, $R(D)$ can only be equal to 0. Therefore, the upper bound D_{\max} of the $R(D)$'s defined domain is the minimum value for all D that satisfy $R(D)=0$.

$R(D)=0$ equivalent input X and output Y statistical independence, which means that any information sent not received at the receiving end, equivalent to the information source does not send any information, or transmission of information source symbol information rate can be compress to 0. Therefore, y_j has nothing to do with x_i . That is

$$p(y_j/x_i) = p(y_j) \quad (5)$$

By Equation (5), D_{\max} can be solved.

$$\begin{aligned}
 D_{\max} &= \min_j \sum_{i=1}^n \sum_{j=1}^m p(x_i)p(y_j/x_i)d(x_i, y_j) \\
 &= \min_j \sum_{i=1}^n \sum_{j=1}^m p(x_i)p(y_j)d(x_i, y_j) \\
 &= \min_j \sum_{j=1}^m p(y_j) \sum_{i=1}^n p(x_i)d(x_i, y_j)
 \end{aligned}
 \tag{6}$$

From the Equation (6) can be observed, In $j = 1, 2, \dots, m$, you can find the j that minimizes the $\sum_{i=1}^n p(x_i)d(x_i, y_j)$ value. When j corresponds to the $p(y_j) = 1$, and the rest of the $p(y_j) = 0$, the right of the formula is minimized. At this point, the upper Formula can be simplified into

$$D_{\max} = \min_j \sum_{i=1}^n p(x_i)d(x_i, y_j)
 \tag{7}$$

Formula (7) is for different, using the input probability distribution $p(x_i)$, the mathematical expectation of the $d(x_i, y_j)$ is taken, the minimum of the mathematical expectation corresponding to the j is taken as D_{\max} , at this point, making the j is the corresponding output probability of the $p(y_j) = 1$, while the other $p(y_j) = 0$.

Thus, the domain of definition $R(D)$ can be defined as $[D_{\min}, D_{\max}]$.

2.3. Domain of the Different Sources Rate Distortion Function $R(D)$

The information source is n element equal probability distribution, the distortion is Hamming distortion

$$d(x_i, y_j) = \begin{cases} 0, & x_i = y_j \\ a, & x_i \neq y_j \end{cases}
 \tag{8}$$

Then the rate distortion function of the information source is [2]

$$R(D) = \begin{cases} \log_2 n - \frac{D}{a} \log_2(n-1) - H\left(\frac{D}{a}\right) & \text{if } 0 \leq D \leq a\left(1 - \frac{1}{n}\right) \\ 0 & \text{if } D > a\left(1 - \frac{1}{n}\right) \end{cases}
 \tag{9}$$

By the Equation (9), the value of $R(D_{\min})$ and $R(D_{\max})$ can be computed, the corresponding simulation diagram of $R(D)$ based on MATLAB is shown in **Figure 1**.

For the information source of unequal probability distribution, the several typical information sources are considered, and the corresponding upper bound D_{\max} and lower bound D_{\min} are obtained by the above derivations as **Table 1**.

3. Conclusion

In this paper, the upper and lower bounds of the definition domain of the information rate distortion function $R(D)$ are deduced and analyzed. The information source type and the distortion function $d(x_i, y_j)$ determine the ex-

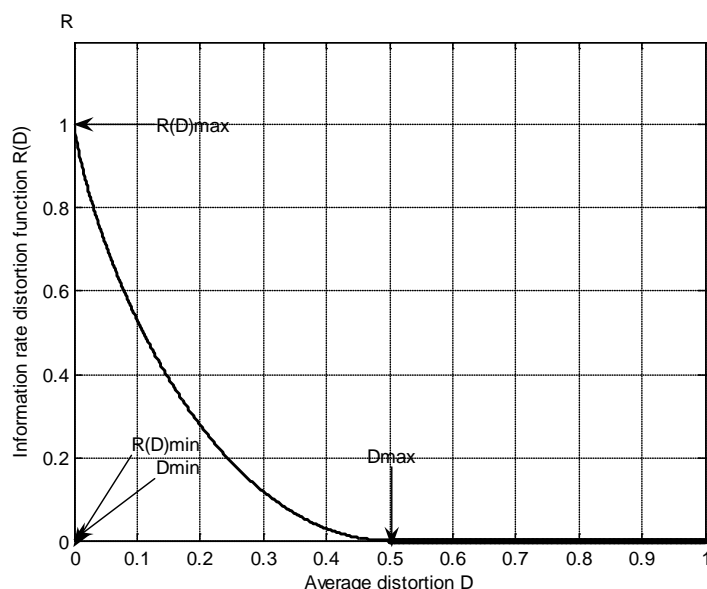


Figure 1. $R(D)$ curve.

Table 1. The upper and the lower bound of typical information source.

| | | | | |
|---|--|--|--|--|
| Source | $\begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$ | $\begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$ | $\begin{bmatrix} 0 & 1 \\ 1/3 & 2/3 \end{bmatrix}$ | $\begin{bmatrix} 0 & 1 \\ 1/3 & 2/3 \end{bmatrix}$ |
| $\begin{bmatrix} X \\ P(X) \end{bmatrix}$ | | | | |
| Distortion matrix d | $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ | $\begin{bmatrix} 1/2 & 1 \\ 2 & 1 \end{bmatrix}$ | $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ | $\begin{bmatrix} 1/2 & 1 \\ 2 & 1 \end{bmatrix}$ |
| D_{\min} | 0 | 3/4 | 0 | 5/6 |
| D_{\max} | 1/2 | 1 | 1/3 | 1 |
| $R(D_{\min})$ | 1 | -- | 0.9183 | -- |
| $R(D_{\max})$ | 0 | 0 | 0 | 0 |

pression of the information rate distortion function $R(D)$ and the upper and lower bounds of the definition domain. Ideally, we all think that the lower bound of the definition domain of the information rate distortion function $R(D)$ is 0, but in practical applications, it's necessary to consider if the lower bound of the definition domain of $R(D)$ can be taken as zero. Accurately speaking, the definition domain of $R(D)$ is $[D_{\min}, D_{\max}]$. The value of D_{\min} is related to whether the transmission rate can reach the entropy of the source. Therefore, it is necessary to calculate the maximum information transmission rate $R(D_{\min})$ after the source coding, so as to solve all kinds of practical problems in information transmission. In this paper, by defining the precise deduction of the lower bound of the $R(D)$ definition domain, it is explained in detail whether or not it can get zero value, and under what conditions zero value can be reached. It further perfects the theory of distortion limited source coding, and it has practical significance for data compression and data transmission of real information source.

Acknowledgements

This work was funded by National Natural Science Foundation of China—Research on no proportion coding cooperative transmission method based on dynamic antenna selection in large scale MIMO systems (61601170) and Henan Provincial Department of science and technology project—Study on wireless channel characteristics of bulk grain reactor (172102210230).

References

- [1] Cao, X.-H. and Zhang, Z.-C. (2009) Information Theory and Coding. 2nd Edition, Tsinghua University Press, Beijing.
- [2] Li, Y.-L. and Zhao, J. (2006) Calculation Method of Discrete Source Rate Distortion Function. *Water Conservancy Science and Technology and Economy*, **12**, 564-566.
- [3] Lu, C.-G. (2012) The Relation between GPS Information and Error Limited Information Rate and Information Rate Distortion and Complexity Distortion. *Journal of Chengdu University of Information Technology*, **6**, 615-622.
- [4] Jiang, S.-F. (2006) Information Rate Distortion Function $R(D)$ Prediction-Correction Calculation Method. *Journal of Shanghai University of Electric Power*, **2**, 187-191.
- [5] You, X.-X. and Wang, J.-H. (2014) Calculation Method of Information Rate Distortion Function Based on Reverse Test Channel. *Journal of Hubei Normal University (Natural Science Edition)*, **4**, 12-16.



Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact ijcns@scirp.org

