R-Q model analysis

In this document, we will introduce the R-Q deduced from the R-lambda model.

As we all know, in a typical video encoder, the QP will have influence on both the residue bits and distortion, so most early rate control algorithms focused on the researches on the R-Q model [1]. Later He et al. [2] proposed a rho-domain based rate control algorithm and propose a linear relationship between R and rho. The rho here means the percentage of non-zero coefficients after transform and quantization. In [2], it was shown that the rho is a more robust factor compared with the QP to determine the residue bitrate R. However, no matter the R-Q model or the R-rho model, they can only characterize the residue bitrate. The non-residue bits cannot be accurately characterized.

Along with the fast development of video coding standards, the non-residue bits can no longer be ignored. Especially under the low bitrate case of the most recent video coding standard High Efficiency Video Coding (HEVC), the non-residue bits may take about half of the total bits. So, the QP or rho is unsuitable for the newest video coding standard. Recently, Li et al. [3] proposed that the lambda is the key factor to determine the bitrate and can determine both the residue and non-residue bits. Since the lambda is the Lagrange Multiplier which determines the optimization target, it will also determine the non-residue bits during the rate distortion optimization process.

Since the lambda is the slope of the R-D curve, to determine the R-lambda relationship is equivalent to determine the R-D relationship. In the previous researches, several types of R-D model have been proposed to characterize the relationship between R and D. One typical of the R-D model is the Exponential function [4].

 （1）

The other type of the R-D model is the Hyperbolic function [5][6].

 （2）

In both the two models, the C and K are the parameters related the video content. We verified the R-D model under the HEVC context as shown in Fig. 1. From Fig. 1, we can obviously see that the Hyperbolic function is more suitable for the modern video coding standard compared with the Exponential function.

Then as lambda is the slope of the R-D curve, the lambda can then be expressed as

 （3）

According to such a R-lambda relationship, we can then derive the R-QP relationship according to a relationship between lambda and QP. The researches of the lambda and QP relationship can be seen from [7]. They got the lambda and QP relationship in inter frames according to multiple QP optimization by trying different QPs for each block and obtaining the optimal one. The lambda and QP relationships are approximated by a linear relationship as shown in Fig. 2. The relationship can be described as follows.

 （4）

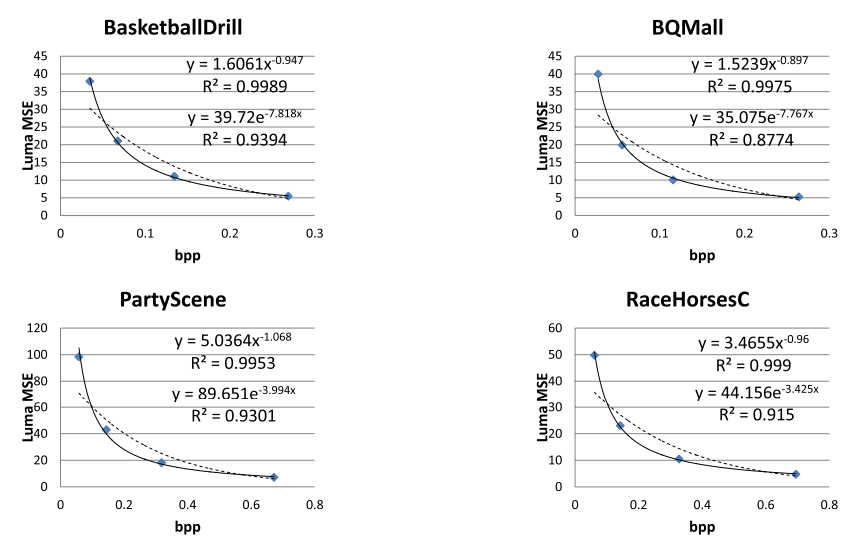


Fig. 1 R-D Curves Fitting According to the Two Different Models.

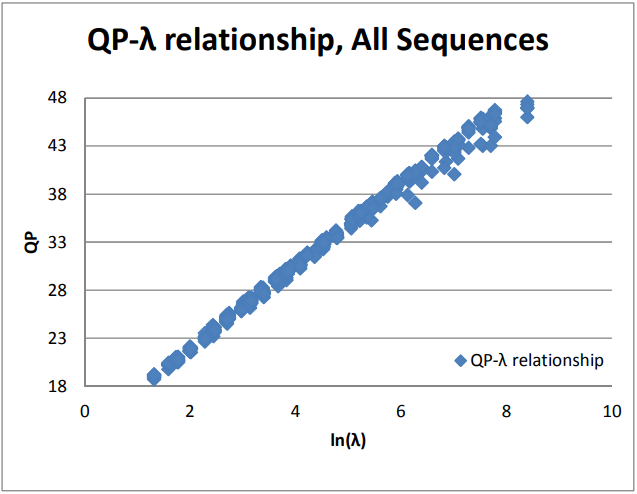


Fig. 2 The QP-lambda relationship

Due to the existence of the skip mode in inter frames, the relationship in (4) may not be that accurate. Most recently, some researches deep dive the coefficients before and after quantization and propose a new linear model to better characterize the QP and lambda relationship [8]. They still use the linear relationship but with different linear coefficients. Overall, they keep the linear relationship between QP and lambda.

 （5）

If we combine (3) and (5), we can organize a linear relationship between QP and R.

 （6）

Where a and b are the model parameters related to the video content.

We then try to adapt the above theory analysis to our proposed CRF-based two-pass transcoding scheme. The process can be divided into two problems. We will first determine the lambda according to the R-D model. Then we will calculate the CRF according to the lambda.

For the first problem, in the first pass, we collect the bitrate R, distortion D, and the average lambda of a specified sequence. If we divide (2) by (3), we can obtain the following equation to derive the model parameters.

 （7）

Using the above equation, we can easily obtain the model parameter K. Then the model parameter C can also be easily derived according to (2). In the second pass, we will calculate the lambda according to (3) and the target bitrate since we already know the model parameters.

To solve the second problem, we try to find a similar linear relationship between lambda and CRF. Different from the lambda and QP relationship, which can be derived from the data block by block, the CRF is a sequence level parameter. So here, we try to find a linear relationship between sequence level lambda and the CRF. We collect a few samples of the test data and derive the model using the linear fitting method. The results can be seen from Fig. 3 and the equation as follows.

 （8）

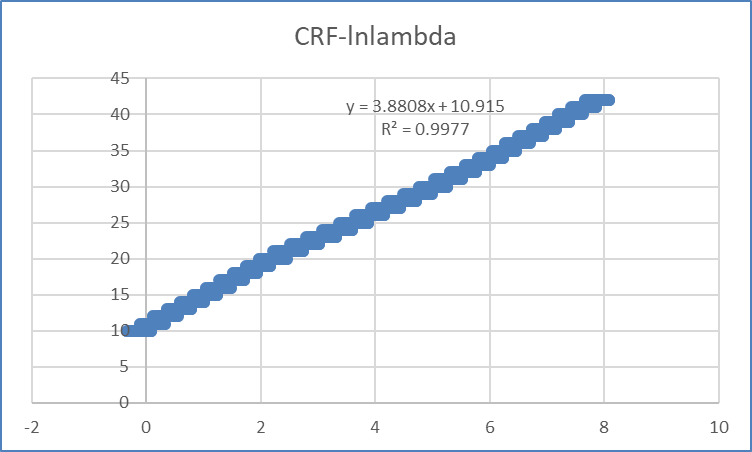


Fig. 3 The CRF-ln(lambda) relationship

We test the performance on some selected sequences. We encode a sequence with CRF equal to 10 in the first pass to derive the bitrate, distortion, and the average lambda to obtain the model parameters. In the second pass, we will use the target bitrates of the actual CRF 11 to 42 to calculate the estimated CRF. The differences between the actual CRF and the calculated CRF from the above equations are shown in Table 1 for some selected sequences. From the results, we can see that the differences are satisfiable for some sequences. However, for the other sequences, the differences can be quite large which may lead to very bad results.

Since the idea of calculating the CRF according to bitrate in theory cannot work will for the actual data, we then come up with an idea to fitting the linear model according to content itself directly. The linear model we tried to use is as follows, which is quite similar to the relationship between QP and R.

 （9）

And at the same time, a large dataset with lots of sequences has been ready. The large dataset including various video contents with four video resolutions including 270p, 360p, 720p, and 1080p. In the following, we will use the large dataset to show the fitting results of the linear model.

Table 1 the differences between the actual CRF and the calculated CRF

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 92ban25M | | | 127xiaoshi35M | | | Aomenfengyun2 | | | BarelonaFC | | |
| A | C | AbsDiff | A | C | AbsDiff | A | C | AbsDiff | A | C | AbsDiff |
| 11 | 11.21 | 0.21 | 11 | 11.89 | 0.89 | 11 | 11.34 | 0.34 | 11 | 11.89 | 0.89 |
| 12 | 12.25 | 0.25 | 12 | 12.63 | 0.63 | 12 | 12.04 | 0.04 | 12 | 13.42 | 1.42 |
| 13 | 13.36 | 0.36 | 13 | 13.34 | 0.34 | 13 | 12.77 | 0.23 | 13 | 14.98 | 1.98 |
| 14 | 14.54 | 0.54 | 14 | 14.03 | 0.03 | 14 | 13.51 | 0.49 | 14 | 16.59 | 2.59 |
| 15 | 15.78 | 0.78 | 15 | 14.77 | 0.23 | 15 | 14.26 | 0.74 | 15 | 18.11 | 3.11 |
| 16 | 17.04 | 1.04 | 16 | 15.51 | 0.49 | 16 | 15.02 | 0.98 | 16 | 19.61 | 3.61 |
| 17 | 18.30 | 1.30 | 17 | 16.32 | 0.68 | 17 | 15.80 | 1.20 | 17 | 21.03 | 4.03 |
| 18 | 19.57 | 1.57 | 18 | 17.18 | 0.82 | 18 | 16.61 | 1.39 | 18 | 22.45 | 4.45 |
| 19 | 20.80 | 1.80 | 19 | 18.12 | 0.88 | 19 | 17.43 | 1.57 | 19 | 23.77 | 4.77 |
| 20 | 22.03 | 2.03 | 20 | 19.38 | 0.62 | 20 | 18.32 | 1.68 | 20 | 25.20 | 5.20 |
| 21 | 23.15 | 2.15 | 21 | 20.85 | 0.15 | 21 | 19.19 | 1.81 | 21 | 26.48 | 5.48 |
| 22 | 24.37 | 2.37 | 22 | 22.85 | 0.85 | 22 | 20.11 | 1.89 | 22 | 27.79 | 5.79 |
| 23 | 25.44 | 2.44 | 23 | 24.80 | 1.80 | 23 | 20.97 | 2.03 | 23 | 29.03 | 6.03 |
| 24 | 26.56 | 2.56 | 24 | 27.14 | 3.14 | 24 | 21.86 | 2.14 | 24 | 30.28 | 6.28 |
| 25 | 27.60 | 2.60 | 25 | 29.52 | 4.52 | 25 | 22.75 | 2.25 | 25 | 31.34 | 6.34 |
| 26 | 28.59 | 2.59 | 26 | 31.70 | 5.70 | 26 | 23.73 | 2.27 | 26 | 32.54 | 6.54 |
| 27 | 29.55 | 2.55 | 27 | 33.59 | 6.59 | 27 | 24.71 | 2.29 | 27 | 33.75 | 6.75 |
| 28 | 30.43 | 2.43 | 28 | 35.25 | 7.25 | 28 | 25.68 | 2.32 | 28 | 34.91 | 6.91 |
| 29 | 31.30 | 2.30 | 29 | 36.75 | 7.75 | 29 | 26.65 | 2.35 | 29 | 36.10 | 7.10 |
| 30 | 32.05 | 2.05 | 30 | 38.13 | 8.13 | 30 | 27.59 | 2.41 | 30 | 37.18 | 7.18 |
| 31 | 32.73 | 1.73 | 31 | 39.48 | 8.48 | 31 | 28.57 | 2.43 | 31 | 38.35 | 7.35 |
| 32 | 33.44 | 1.44 | 32 | 40.60 | 8.60 | 32 | 29.49 | 2.51 | 32 | 39.18 | 7.18 |
| 33 | 34.05 | 1.05 | 33 | 41.78 | 8.78 | 33 | 30.47 | 2.53 | 33 | 40.20 | 7.20 |
| 34 | 34.71 | 0.71 | 34 | 42.82 | 8.82 | 34 | 31.41 | 2.59 | 34 | 41.15 | 7.15 |
| 35 | 35.36 | 0.36 | 35 | 44.00 | 9.00 | 35 | 32.43 | 2.57 | 35 | 42.17 | 7.17 |
| 36 | 36.01 | 0.01 | 36 | 44.96 | 8.96 | 36 | 33.41 | 2.59 | 36 | 42.97 | 6.97 |
| 37 | 36.57 | 0.43 | 37 | 46.11 | 9.11 | 37 | 34.44 | 2.56 | 37 | 43.76 | 6.76 |
| 38 | 37.19 | 0.81 | 38 | 47.12 | 9.12 | 38 | 35.40 | 2.60 | 38 | 44.49 | 6.49 |
| 39 | 37.74 | 1.26 | 39 | 48.26 | 9.26 | 39 | 36.46 | 2.54 | 39 | 45.43 | 6.43 |
| 40 | 38.32 | 1.68 | 40 | 49.32 | 9.32 | 40 | 37.55 | 2.45 | 40 | 46.27 | 6.27 |
| 41 | 38.89 | 2.11 | 41 | 50.34 | 9.34 | 41 | 38.58 | 2.42 | 41 | 47.08 | 6.08 |
| 42 | 39.48 | 2.52 | 42 | 51.44 | 9.44 | 42 | 39.58 | 2.42 | 42 | 48.06 | 6.06 |
| Avg. |  | **1.50** |  |  | **4.99** |  |  | **1.89** |  |  | **5.55** |

We fit the linear model (9) by generating many pairs of CRFs and bitrates from 10 to 42 and obtain a model for each specified sequence. Then we use the model to calculate the estimated bitrate according to the CRF within the range of 10 to 42. The bitrate error will then be calculated using the difference between the actual bitrate and the estimated bitrate.

 （10）

The fitting results are shown in Table 2. From Table 2, we can see that the bitrate errors are still not satisfiable using the content-dependent model parameters. The bitrate error within 10% is only about 69% in average, and the bitrate error within 20% is about 91% in average

Table 2. The linear model fitting results

|  |  |  |
| --- | --- | --- |
| Sequence resolution | Bitrate error within 20% | Bitrate error within 10% |
| 270p | 98% | 87% |
| 480p | 96% | 80% |
| 720p | 89% | 62% |
| 1080p | 82% | 49% |
| Average | **91%** | **69%** |

Also, according to some observations in Fig. 4, the linear relationship in some sequences is far from good enough.

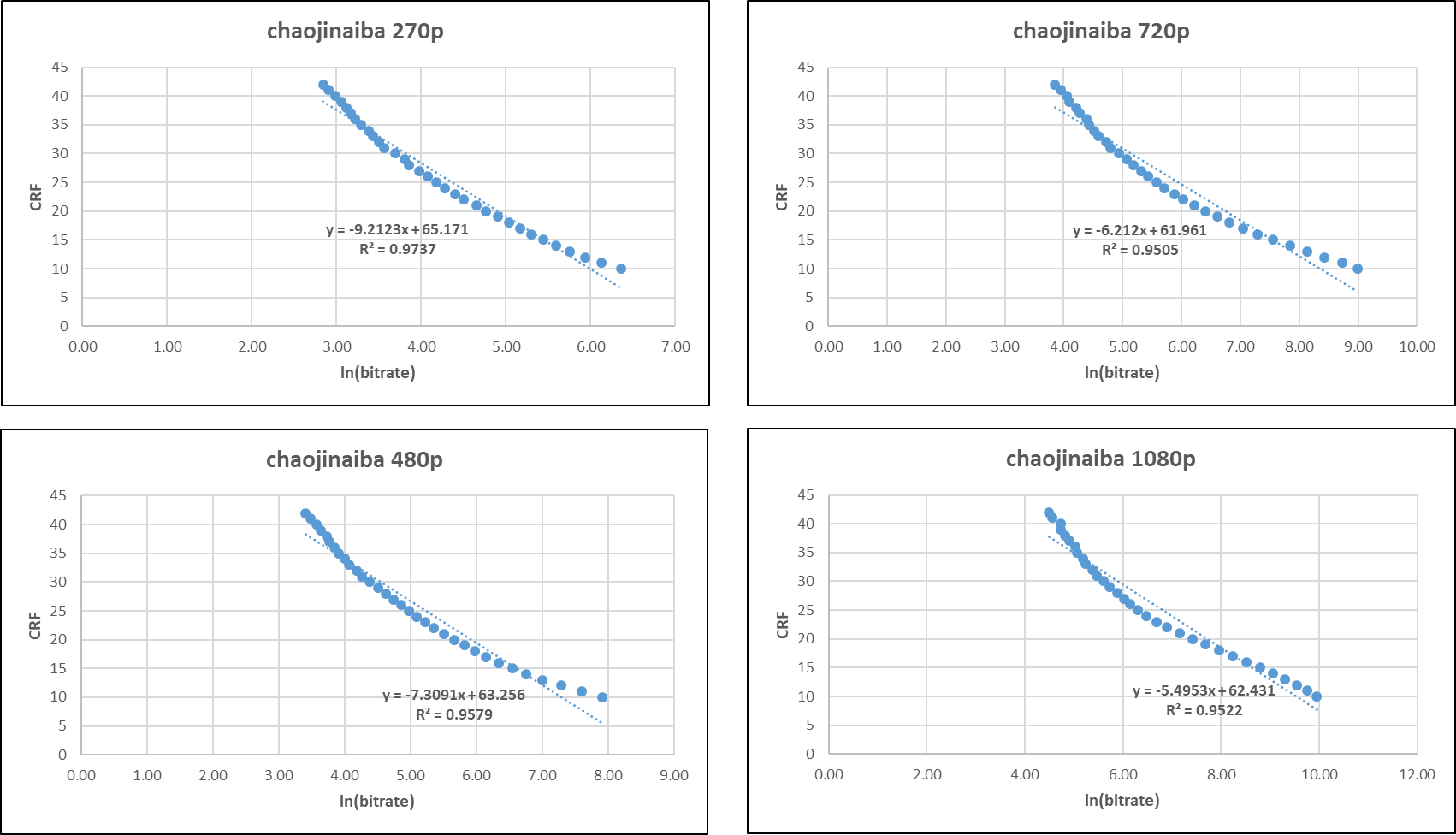


Fig. 4 The linear relationship failure case

Then motivated by our recent researches that the second order model between QP and bitrate can be more adaptable to a larger range of QPs. And some figures in Fig. 5 also show that the second order model can achieve better accuracy. We try to use the following second order model to fit the relationship between CRF and ln(R).

 （11）

We follow the same criteria to show the performance of the proposed second order model compared with the linear model. The comparison results are shown in Table 3. The numbers in the brackets show the improvements of the second order model over the linear model. As can be seen from Table 3, the proposed method can obviously improve the bitrate errors within 10% by 27%, and the bitrate errors within 20% by 8%. Therefore, we finally use the second order model to characterize the relationship between CRF and bitrate.

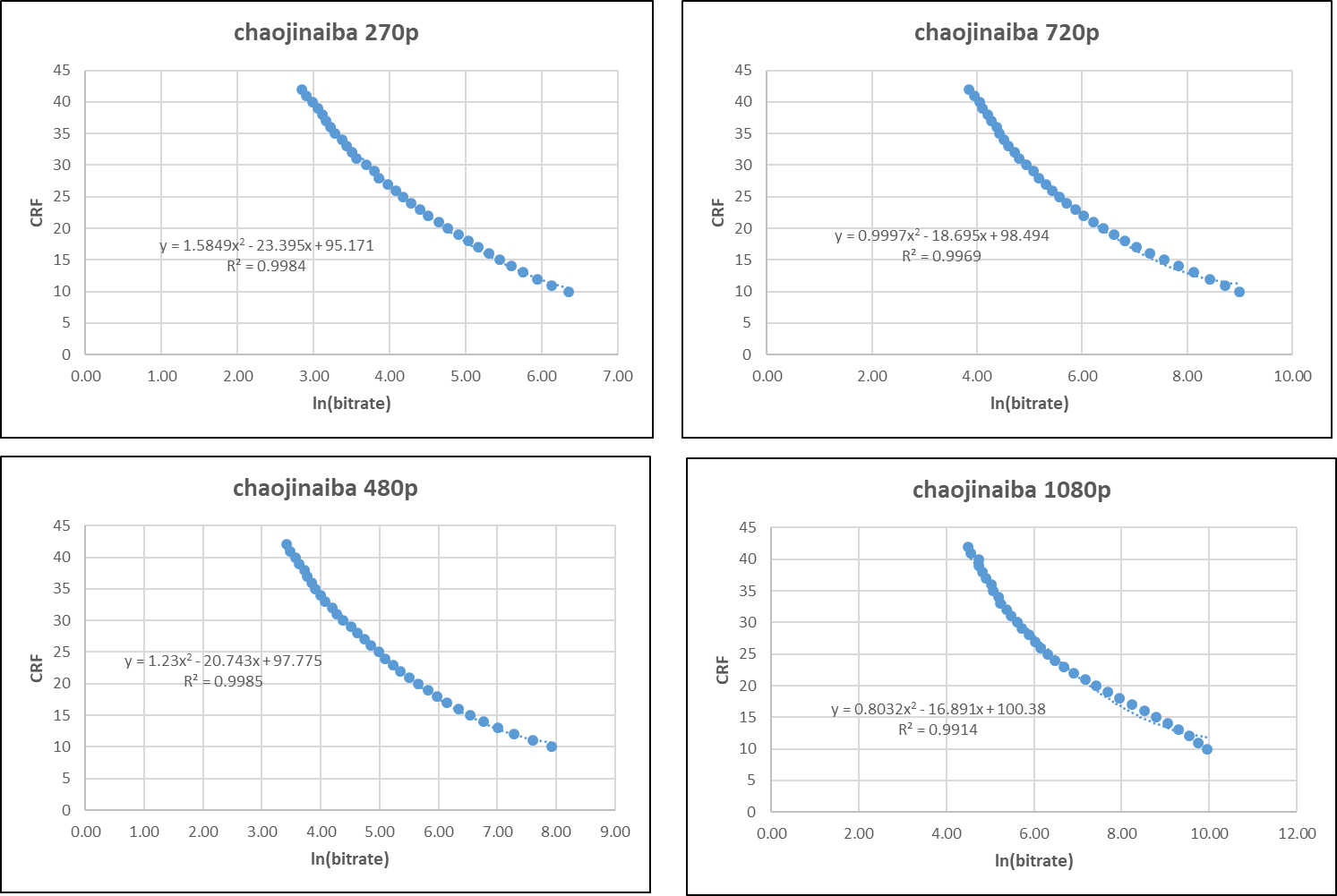


Fig. 5 The second order relationship of typical sequences

Table 2. The second order model fitting results

|  |  |  |
| --- | --- | --- |
| Sequence resolution | Bitrate error within 20% | Bitrate error within 10% |
| 270p | 100% (2%) | 99% |
| 480p | 100% (4%) | 99% |
| 720p | 99% (10%) | 97% |
| 1080p | 97% (15%) | 89% |
| Average | **99% (8%)** | **96% (27%)** |

**References**

[1] S. Ma, W. Gao and Y. Lu, “Rate-distortion analysis for H.264/AVC video coding and its application to rate control,” in IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 12, pp. 1533-1544, Dec. 2005.

[2] Z. He, Y. K. Kim and S. K. Mitra, “Low-delay rate control for DCT video coding via ρ-domain source modeling,” in IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 8, pp. 928-940, Aug. 2001.

[3] B. Li, H. Li, L. Li and J. Zhang, " $\lambda$ Domain Rate Control Algorithm for High Efficiency Video Coding," in IEEE Transactions on Image Processing, vol. 23, no. 9, pp. 3841-3854, Sept. 2014.

[4] G. J. Sullivan and T. Wiegand, “Rate-distortion optimization for video compression,” IEEE Signal Process. Mag., vol. 15, no. 6, pp. 74–90, Nov. 1998.

[5] S. Mallat and F. Falzon, “Analysis of low bit rate image transform coding,” IEEE Trans. Signal Process., vol. 46, no. 4, pp. 1027–1042, Apr. 1998.

[6] M. Dai, D. Loguinov, and H. Radha, “Rate-distortion modeling of scalable video coders,” in Proc. Int. Conf. Image Process. (ICIP), vol. 2. Oct. 2004, pp. 1093–1096.

[7] B. Li, J. Xu, D. Zhang and H. Li, "QP refinement according to Lagrange multiplier for High Efficiency Video Coding," 2013 IEEE International Symposium on Circuits and Systems (ISCAS2013), Beijing, 2013, pp. 477-480.

[8] K. Andersson, P. Wennersten, R. Sjöberg, J. Samuelsson, J. Ström, P. Hermansson, M. Pettersson, “Non-normative HM encoder improvements”, JCTVC-W0062, San Diego, USA, 19–26 February 2016.