Template for DCC Manuscripts

Author One*, Author Two†, and Author Three*

*Institution One
Street Address One
City, State, ZIP, Country
email@address

†Institution Two
Street Address Two
City, State, ZIP, Country
email@address

Abstract

The abstract describes concisely and clearly the main contributions of the paper. It should not contain math equations or citations, to ensure the abstract is self-contained and readable if converted to ASCII text. The abstract should contain about 100 to 150 words.

1 Introduction

Along with the explosive increase of multimedia content, effective and efficient compression algorithms have been always in demand for decades. Catering for such demand, various video coding standards have been proposed and developed to compress the raw video data [1–3], in the way of minimizing bit-rates or optimizing quality with certain constraints according to different applications. Generally speaking, rate control [4] aims at controlling bit-rates to meet different requirements, e.g., minimizing distortion for storage application or reducing bit-rate fluctuation for communication usage. Moreover, in some cases when the perceived quality is highly crucial, quality control is adopted to compress a video at a certain and constant quality, thus obtaining a more desirable quality of experience (QoE).

Similar to constant bit-rate control which aims at smoothing bit-rates for each frame to avoid buffer overflow or underflow, constant quality control (Figure 1) in video coding provides smooth quality in compressed frames to avoid overall perceived quality degradation caused by some intense quality fluctuation. Toward this end, a direct way is to adopt a two-pass procedure, which pre-analyzes all the frames in advance and calculates the global quantization parameters (QPs) for each frame [5] [6] [7]. In addition, in some cases the calculated QPs are not integers, and [8] thus proposed a solution by adjusting different QPs at block level to reach the desired QP value at frame level. Obviously, the pre-analysis strategy is not applicable to real-time streaming applications, such as video conferencing and live video streaming.

Later, He *et.al.* [9] proposed a low-pass filter with geometric factor, and then applied it in quality control to achieve smoothed frame quality, meanwhile satisfying buffer status with low delay. Besides, a PSNR adjustment was proposed in [10] to maintain the group of picture (GOP) level quality constant by empirically moving up/down QPs, according to previous frames. Instead of the GOP-based control, [11] proposed a sequence-based method, which tracks scene changes and then achieves smoother video quality adaptive to scene changes. Recently, a trellis-based method has been proposed in [12] to consider both bit budge and PSNR variance. However, although the above works can reduce quality

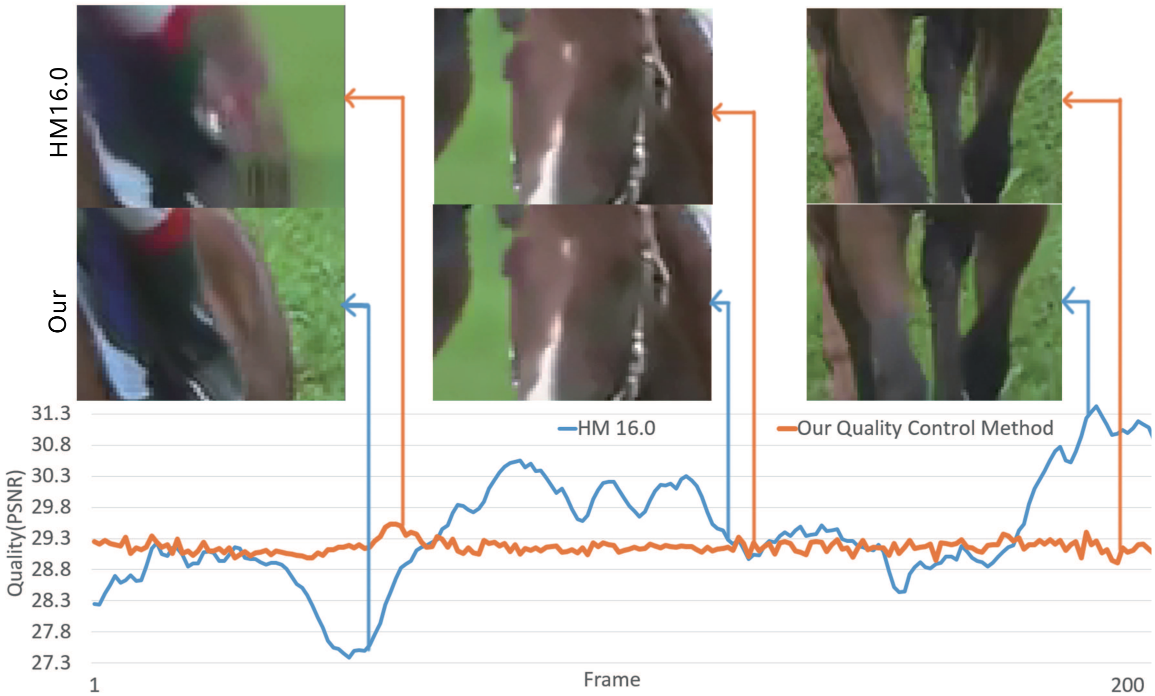


Figure 1: An example of quality control in video coding. In this figure, the state-of-the-art video coding standard high efficiency video coding (HEVC), which is implemented by the latest reference software HM16.0, is used as an anchor. This figure shows that the fluctuation of the HEVC encoded video may result in poor perceived quality for some frames (e.g, frames 55, 120 and 190). By contrast, our quality control method is able to yield smooth quality, better than the poor quality of frames 55, 120 and 190 in HEVC. Besides, our method is able to make the distortion of the encoded video approach to the target, i.e., 29.2 dB.

fluctuation with constant quality, they are unable to control quality to a certain level that users demand.

Most recently, Hu *et.al.* [13] have proposed to control the quality of encoded videos to a certain target for H.264. In their method, QP of each frame is adjusted by modelling consecutive frames, in the assumption of Laplacian distribution on transform coefficients. Also based on the same assumption, [14] proposed a mixture model on Laplacian function to tailor distortion-quantization (D-Q) and rate-quantization (R-Q) relationships in HEVC, thus controlling quality in a constant manner. However, those works are restricted to either different standards or certain assumptions, and therefore they are not effective when being applied to the latest r-λ model [4] of HEVC.

In this paper, we propose a novel quality control method for video coding, which does not rely on some specific encoder or assumption. Therefore, our method is suitable for different encoders. To our best knowledge, our method is the first work on encoder-free quality control, which enables certain and constant distortion across all frames of encoded videos. Specifically, our method is inspired by the basic idea of the proportion-integral-derivative (PID) controller, thus called the PID-based quality control (PQC) method. We first propose a formulation of quality control in video coding for our PQC method, with objectives of minimization on both control error and fluctuation of distortion. Then, the relationship between QP and distortion is modeled for the formulation of out PQC method, such that the PID controller can be used to solve our formulation. Next, a PID-based solution is provided to our quality control formulation with modelled relationship between QP and distortion. As such, the distortion of video coding can be controlled to a target distortion with smooth quality fluctuation. Finally, we implement our PQC method in the latest HEVC encoder (HM 16.0), and our experiment results show that our method achieves the state-of-the-art quality control performance in terms of both control error and quality fluctuation.

2 Overview of PID Controller

The PID controller [15] is widely used to minimize the error between a desired step point and the measured process variable. Using the terms of proportional operator (P), integral operator (I) and derivative operator (D), a PID controller maintains a well trade-off among response speed, static error correction and overreacting repression. Generally speaking, the PID controller performs robustly with little overhead of computational complexity. Thus, our quality control approach for video coding is also based on the PID controller.

To be more specific, assume that there exists errors between the current and target positions until time $(t-1)$, which are denoted by $\{e_{t-1}, e_{t-2},...,e_0\}$. The PID controller focuses on minimizing error e_t by adjusting a control variable o_t. In the PID controller, o_t can be calculated by

$$o_t = K_p e_{t-1} + K_i \int_0^{t-1} e_\tau d\tau - K_d \frac{de_{t-1}}{dt}, \tag{1}$$

where e_{t-1}, $\int_0^{t-1} e_\tau d\tau$ and $\frac{de_{t-1}}{dt}$ are the proportion (P), integral (I) and derivative (D) values; K_p, K_i and K_d (all ≥ 0) denote their corresponding weights. As can be seen from (1), P is decided by the most recent error, whilst I accounts for long lasting previous errors and D predicts error in the future. For more detials about the setting of K_p, K_i and K_d, refer to [16]. To sum up, the PID controller outputs an optimal predicted value of the control variable to minimize error e_t, using all the errors incurred until $(t-1)$. Next, we propose to control quality of video coding by incorporating the PID controller.

3 PID-based quality control method

In this section, we present our encoder-free PQC method for video coding, which is achieved by predicting the optimal QP before encoding each frame. Specifically, we first establish in Section 3.1 the formulatioin of quality control in video coding. In Section 3.2 we further model the relationship between QP_t and e_t for the proposed quality control formulation. In Section 3.3, we solve the proposed quality control formualtion using the PID controller.

3.1 Formulation of quality control

In video coding, there are two main objectives for quality control:

Objective I: Minimizing the error between the actual and target quality, averaged over all frames.

Objective II: Minimizing the fluctuation of quality along with frames.

The above two objectives can be achieved by predicting the optimal QP before encoding each frame. In other words, before encoding the t-th frame, we need to estimate the best QP value for this frame, which is denoted by QP_t. Assuming that T is the target distortion and D_t is the distortion of the t-th frame, the quality control can be formulated by

$$\mathrm{QP}_t = \underset{\mathrm{QP}}{\mathrm{argmin}} \{ \underbrace{\lambda \cdot \underbrace{(D_t(\mathrm{QP}) - T)}_{\textbf{Objective I}} + (1-\lambda) \cdot \underbrace{\frac{dD_t(\mathrm{QP})}{dt}}_{\textbf{Objective II}}}_{e_t} \}, \tag{2}$$

where $(D_t(\mathrm{QP}) - T)$ models the error between the actual quality and target quality (**Objective I**), whereas $\frac{dD_t(\mathrm{QP})}{d_t}$ models the fluctuation of quality (**Objective II**). In addition, λ represents the trade off between the two objectives. In addition, e_t denotes the overall error to be minimized.

Since PSNR is a widely used distortion evaluation metric for video coding, it is applied to model distortion D_t in (2) for this paper. However, our PQC method can be simply extended to adopting other distortion evaluation metrics, like Structural Similarity Index (SSIM) and Video Signal-to-Noise Ratio (VSNR). To solve formulation (2), we first need to model the relationship between QP_t and e_t, to be discussed in the following.

3.2 Relationship between QP_t *and* e_t

In a coding system, the relationship between $\mathrm{QP_t}$ and e_t can be formulated by the following function Ψ,

$$\Psi(\mathbf{I}_t, \mathbf{I}_{t-1}, ..., \mathbf{I}_0, \mathrm{QP}_t, \mathrm{QP}_{t-1}, ..., \mathrm{QP}_0) = e_t, \tag{3}$$

where $\mathbf{I}_t$ is the frame content at frame t. This formulation shows that content and QPs of all frames until currently encoded frame contribute to quality control error e_t. In fact, $\mathbf{I}_t, \mathbf{I}_{t-1}, ..., \mathbf{I}_0$ is a set of images from the video sequence. Thus, we denote them by a single tensor $\mathbb{I}$. Since our intention here is to analyze the relationship between QP_t and e_t. Given a sequence, we have a fixed $\mathbb{I}$, and thus (3) can be rewritten by

$$\Psi_{\mathbb{I}}(\mathrm{QP}_t, \mathrm{QP}_{t-1}, ..., \mathrm{QP}_0) = e_t. \tag{4}$$

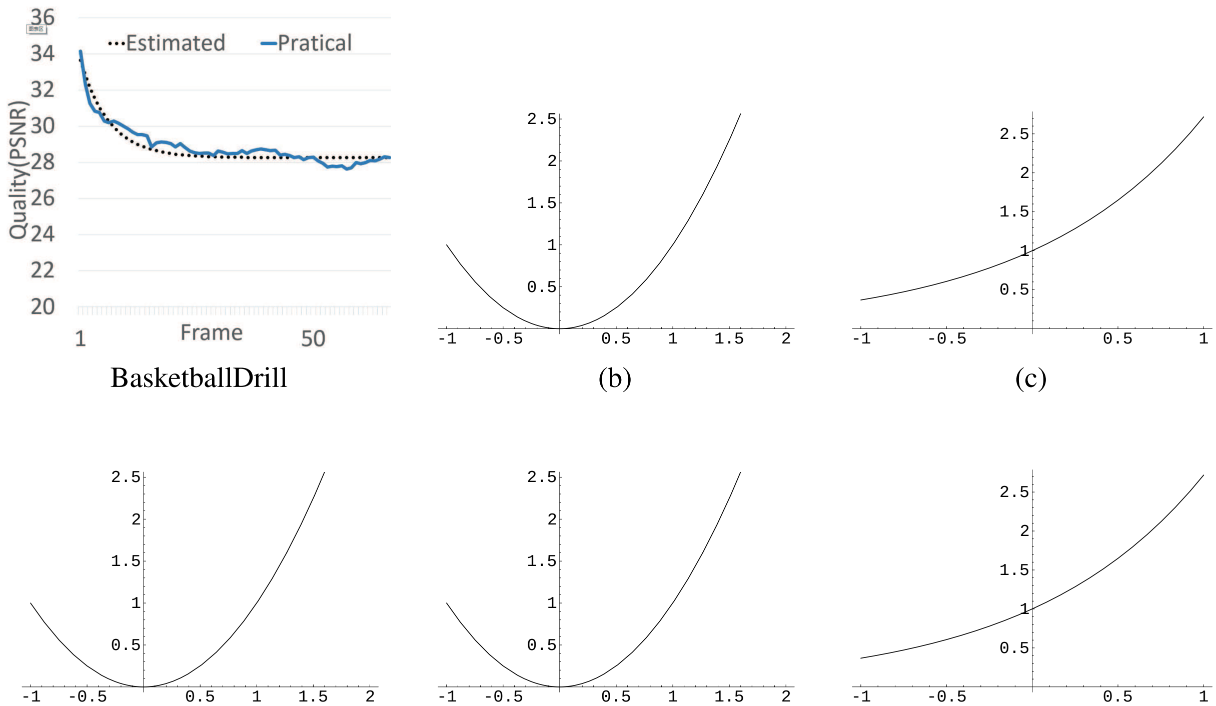
Obviously, $\Psi_{\mathbb{I}}$ describes the relationship between QP_t and e_t for given $\mathbb{I}$. To obtain $\Psi_{\mathbb{I}}$, we propose a simple and pratical way from the viewpoint of signal processing: Treating $\Psi_{\mathbb{I}}$ as an unknown linear time invariant(LTI) system (Input:QP_t; Output:e_t), we process time domain analysis on it. The analysis can be achieved via the following steps:

$\mathbf{Step\ 1}$: $\mathbf{Input}$ a impulse signal to the encoding system with $\mathbf{QP} = \{\mathrm{QP}_0 = \mathrm{QP}_{min}, \mathrm{QP}_1 = \mathrm{QP}_{max}, \mathrm{QP}_2 = \mathrm{QP}_{max}, ..., \mathrm{QP}_t = \mathrm{QP}_{max}\}$.

$\mathbf{Step\ 2}$: $\mathbf{Obtain}$ system response: $\Psi_{\mathbb{I}}(\mathbf{QP})$.

$\mathbf{Step\ 3}$: $\mathbf{Model}$ $\Psi_{\mathbb{I}}(\mathbf{QP})$ in the form of

$$\Psi_{\mathbb{I}}(\mathbf{QP}) = C_0 \cdot \delta(t) + C_1 \cdot e^{s_1 t} + C_2 \cdot e^{s_2 t} + ... + C_n \cdot e^{s_n t} \tag{5}$$



BasketballDrill (b) (c)

(b) (b) (c)

Figure 2: Model $\Psi_{\mathbb{I}}(\mathbf{QP})$ in the form of $\Psi_{\mathbb{I}}(\mathbf{QP}) = C_0 \cdot \delta(t) + C_1 \cdot e^{s_1 t} + C_2 \cdot e^{s_2 t} + ... + C_n \cdot e^{s_n t}$.

where $\{C_0, C_1, ...C_n\}$ and $\{s_0, s_1, ..., s_n\}$ are the coefficients and index terms to be estimated. Next, we investigate the coefficients of (5) in HEVC encoding footnote $\{XXXX\}$ by the data analysis of Figure 2 . Figure 2 shows that in inter frame coding, the $\Psi_{\mathbb{I}}(\mathbf{QP})$ can be modeled as:

$$\textbf{Inter Frame}: \ \Psi_{\mathbb{I}}(\mathbf{QP}) = C_1^{\mathbb{I}} \cdot e^{s_1 t} + C_2^{\mathbb{I}} \cdot e^{s_2^{\mathbb{I}} t}. \tag{6}$$

where $C_0^{\mathbb{I}}$ and $C_1^{\mathbb{I}}$ are coefficients estimated from Figure 2. Similarly, we have the following equation for the intra frame coding:

$$\textbf{Intra Frame}: \ \Psi_{\mathbb{I}}(\mathbf{QP}) = C_0^{\mathbb{I}} \cdot \delta(t), \tag{7}$$

Since $\Psi_{\mathbb{I}}$ is an LTI system, following equation holds:

$$\Psi_{\mathbb{I}}(\mathbf{QP}) * \mathrm{QP}_t = e_t \tag{8}$$

Step 4 : **Derive** the relationship between QP_t and e_t via implementing Laplace transform on (8), and repalce $\Psi_{\mathbb{I}}(\mathbf{QP})$ with (6) and (7) respectively:

$$\textbf{Inter Frame}: \ A_1^{\mathbb{I}} \cdot e_t + A_2^{\mathbb{I}} \cdot \frac{de_t}{dt} = \mathrm{QP}_t. \tag{9}$$

where $A_1^{\mathbb{I}}$ and $A_2^{\mathbb{I}}$ are the coefficients derived from $C_1^{\mathbb{I}}$ and $C_2^{\mathbb{I}}$ in (6), respectively. Similarly, we have the following equation for the intra frame coding:

$$\textbf{Intra Frame}: \; A_0^{\mathbb{I}} \cdot e_t = \mathrm{QP}_t, \tag{10}$$

where $A_0^{\mathbb{I}}$ is the coefficient derived from $C_0^{\mathbb{I}}$ in (7).

In following section, we focus on our solution to the proposed quality control formulation(2) on the basis of e_t and QP_t relationship of (9)(10) and the PID controller of (1).

3.3 PID-based solution to the quality control formulation

In this section, we apply the PID controller to solve (2). As metioned in Section 2, the PID controller minimizes the error e_t alongside time t, via adjusting the control variable o_t. However, the PID controller can perform well, only when applied to a second-order system [17]. In other words, e_t and o_t in (1) need to satisfy the following differential equation:

$$M_2 \cdot \frac{d^2 e_t}{dt^2} + M_1 \cdot \frac{de_t}{dt} + M_0 \cdot e_t = o_t. \tag{11}$$

where C_2, C_1 and C_0 are constants. Next, we use the following way to make the modelled relationship between e_t and QP_t (i.e, (9) and (10)) satisfy the above requirement of the PID controller (i.e, (11)). As such the PID controller can be applied to solve our quality control formulation of Section 3.1. Specifically, by applying the differential operation on (9), the following holds:

$$A_1^{\mathbb{I}} \cdot \frac{d^2 e_t}{dt^2} + A_0^{\mathbb{I}} \cdot \frac{de_t}{dt} + 0 \cdot e_t = \frac{d\mathrm{QP}_t}{dt}. \tag{12}$$

Thus, we can see that (12) satisfies the requirement (11) of the PID controller with $M_2 = A_1^{\mathbb{I}}$, $M_1 = A_0^{\mathbb{I}}$ and $M_0 = 0$. As a result, we have

$$\frac{d\mathrm{QP}_t}{dt} = o_t. \tag{13}$$

Then, (13) can be rewritten as follows,

$$\textbf{Inter Frame}: \; \mathrm{QP}_t = \int_0^t o_\tau d\tau. \tag{14}$$

Similarly, on the basis of (10) and (11), we have the following equation for the intra frames of video coding:

$$\textbf{Intra Frame}: \; \mathrm{QP}_t = \int_0^t \int_0^\tau o_\rho d\rho d\tau. \tag{15}$$

Finally, by replacing o_t with (1), the QP values of each frame can be estimated as follows,

$$\textbf{Inter Frame}: \; \mathrm{QP}_t = \int_0^t o_\tau d\tau = \int_0^{t-1} (K_p e_{\rho-1} + K_i \int_0^{\rho-1} e_\gamma d\gamma - K_d \frac{de_{\rho-1}}{d\rho}) d\rho, \tag{16}$$

$$\textbf{Intra Frame}: \; \mathrm{QP}_t = \int_0^t \int_0^\tau o_\rho d\rho d\tau = \int_0^{t-1} \int_0^\tau (K_p e_{\rho-1} + K_i \int_0^{\rho-1} e_\gamma d\gamma - K_d \frac{de_{\rho-1}}{d\rho}) d\rho d\tau, \tag{17}$$

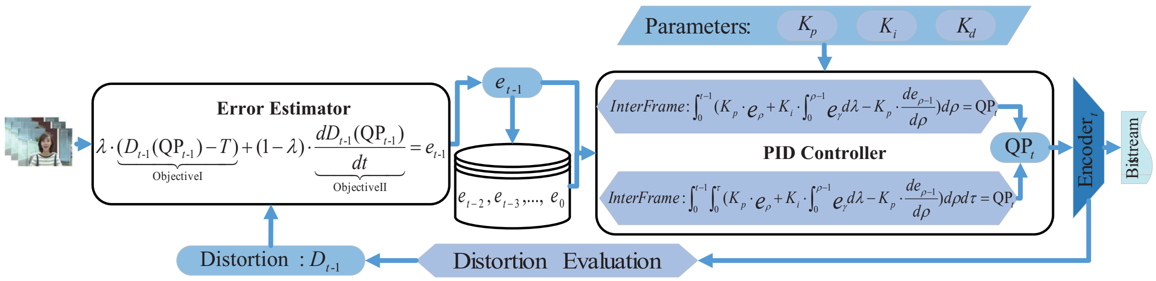


Figure 3: Overview of our PQC method.

Table 1: Average PSNR in dB for the “Coastguard” video sequence

	2D	3D	MC-BCS-SPL	
S_{NK}	BCS-SPL	BCS-SPL	$S_{\mathrm{K}} = S_{\mathrm{NK}}$	$S_{\mathrm{K}} = 0.7$
0.1	22.69	22.76	23.06	25.29
0.2	24.70	24.76	25.78	27.94
0.3	26.37	26.45	28.29	30.15
0.4	27.99	27.95	30.88	32.30
0.5	29.60	29.57	33.58	34.42

for controlling quality of video coding. Obviously, we can see that QP_t is only related with the control error $\{e_0, e_1, ..., e_{t-1}\}$ and parameters K_p, K_i and K_d, such that our method can be seen as an encoder-free quality control method. Figure 3 summarizes the overall framework of our PQC method.

Headings

The LaTeXclass for DCC includes formatting for sections ...

A Subsection Heading

and also for subsections. The use of sub-subsections is discouraged.

Figures and Tables

The proceedings are published in black and white; all figures and charts should be clear when printed in grayscale. Figures and tables should be concise and easy to read. Avoid making complex graphics and then reducing them so much that they become hard to read.

Position illustrations at the top of the page rather than in the middle or at the bottom. Caption and number every illustration. Fig. 4 shows an example illustration. Table 1 shows an example table.

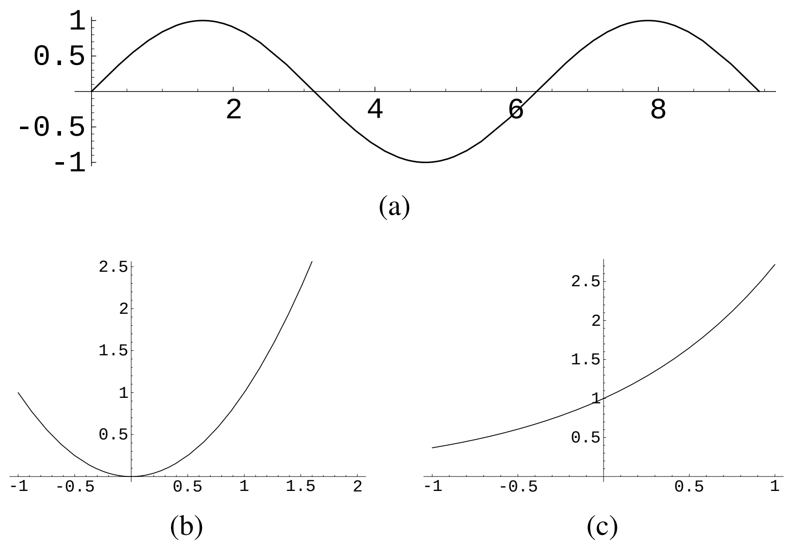


Figure 4: An example figure.

Reference to Prior Literature

List and number all bibliographical references at the end of the paper. The references can be numbered in alphabetic order or in order of appearance in the document. When referring to them in the text, type the corresponding reference number in square brackets as shown at the end of this sentence [18]. The reference list below shows an example of citing a journal article [18], a conference paper [19], a book chapter [20], and a book [21]. Add your citations to the `refs.bib` file.

References

[1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the h. 264/avc video coding standard," *IEEE Transactions on circuits and systems for video technology*, vol. 13, no. 7, pp. 560–576, 2003.

[2] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[3] D. Mukherjee, J. Bankoski, A. Grange, J. Han, J. Koleszar, P. Wilkins, Y. Xu, and R. Bultje, "The latest open-source video codec vp9-an overview and preliminary results," in *Picture Coding Symposium (PCS), 2013*. IEEE, 2013, pp. 390–393.

[4] S. Li, M. Xu, Z. Wang, and X. Sun, "Optimal bit allocation for ctu level rate control in hevc," 2016.

[5] Y. Yu, J. Zhou, Y. Wang, and C. W. Chen, "A novel two-pass vbr coding algorithm for fixed-size storage application," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 345–356, 2001.

[6] L.-J. Lin and A. Ortega, “Bit-rate control using piecewise approximated rate-distortion characteristics,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 4, pp. 446–459, 1998.

[7] J. Lee and B. W. Dickinson, “Temporally adaptive motion interpolation exploiting temporal masking in visual perception,” *IEEE Transactions on image processing*, vol. 3, no. 5, pp. 513–526, 1994.

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

[9] Z. He, W. Zeng, and C. W. Chen, “Low-pass filtering of rate-distortion functions for quality smoothing in real-time video communication,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 8, pp. 973–981, 2005.

[10] F. De Vito and J. C. De Martin, “Psnr control for gop-level constant quality in h. 264 video coding,” in *Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology, 2005.* IEEE, 2005, pp. 612–617.

[11] B. Xie and W. Zeng, “A sequence-based rate control framework for consistent quality real-time video,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 1, pp. 56–71, 2006.

[12] K.-L. Huang and H.-M. Hang, “Consistent picture quality control strategy for dependent video coding,” *IEEE Transactions on Image Processing*, vol. 18, no. 5, pp. 1004–1014, 2009.

[13] S. Hu, H. Wang, and S. Kwong, “Adaptive quantization-parameter clip scheme for smooth quality in h. 264/avc,” *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1911–1919, 2012.

[14] C.-W. Seo, J.-H. Moon, and J.-K. Han, “Rate control for consistent objective quality in high efficiency video coding,” *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2442–2454, 2013.

[15] K. J. strm and T. Hgglund, “Advanced pid control,” 2006.

[16] P. Cominos and N. Munro, “Pid controllers: recent tuning methods and design to specification,” *IEE Proceedings-Control Theory and Applications*, vol. 149, no. 1, pp. 46–53, 2002.

[17] C. Zhao, D. Xue, and Y. Chen, “A fractional order pid tuning algorithm for a class of fractional order plants,” in *IEEE International Conference Mechatronics and Automation, 2005*, vol. 1. IEEE, 2005, pp. 216–221.

[18] J. E. Fowler, “Compressive-projection principal component analysis,” *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2230–2242, October 2009.

[19] ——, “Compressive-projection principal component analysis for the compression of hyperspectral signatures,” in *Proceedings of the Data Compression Conference*, J. A. Storer and M. W. Marcellin, Eds., Snowbird, UT, March 2008, pp. 83–92.

[20] J. E. Fowler and Q. Du, “Reconstructions from compressive random projections of hyperspectral imagery,” in *Optical Remote Sensing: Advances in Signal Processing and Exploitation Techniques*, S. Prasad, L. M. Bruce, and J. Chanussot, Eds. Springer, 2011, ch. 3, pp. 31–48.

[21] B. N. Parlett, *The Symmetric Eigenvalue Problem.* Philadelphia, PA: Society for Industrial and Applied Mathematics, 1998.