Front. Comput. Sci.
DOI

Supplementary File

Perceptual quantization parameter selection for crime scene investigation tool images

Yanchao GONG[1], Zhao LI[2], Zhuang WANG[3], Kaifang YANG (✉)[4], Ying LIU[1], and Keng Pang LIM[5]
1 School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
2 Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China
3 School of Electronic and Information Engineering, Chongqing Three Gorges University, Chongqing 404000, China
4 School of Computer Science, Shaanxi Normal University, Xi'an 710119, China
5 XsecPro Pte. Ltd., Singapore 787820

Abstract Crime scene investigation (CSI) image high efficiency coding is a crucial guarantee for solving criminal cases and court trials. The coding efficiency of CSI images is closely related to the selected quantization parameters for coding. However, the quantization parameter selection currently used in the advanced high efficiency video coding (HEVC) standard are not optimized for CSI images due to the saliency information not being considered. For the CSI images, the foreground regions are significantly important than the other background regions for solving criminal cases and court trials. In this paper, a perceptual quantization parameter selection technique considering the saliency information for CSI tool images, referred to as CSIT-QP, is proposed using a two-step Lagrangian rate-distortion optimization. Compared with the state-of-the-art quantization parameter selection techniques for HEVC, the CSIT-QP can achieve higher reconstructed image perceptual quality using lower coding rate.

Keywords video coding, quantization parameter, crime scene investigation, saliency information, perceptual quality.

Received February 02, 2022; accepted November 03, 2022

E-mail: kfyang@snnu.edu.cn

1 Introduction

Crime scene investigation (CSI) is critical for solving criminal cases and court trials [1, 2]. Fig.1 illustrates the main workflow of a CSI application, where CSI images undergo a sequence of processes, i.e., acquisition, coding, transmission, decoding, display, and analysis. Currently, all CSI data related to cases must be kept as electronic files in specific systems according to relevant regulations. Electronic files, including reconstructed CSI images, provides great convenience for the public security and court to store, manage, display, and analyze CSI data more effectively. Therefore, image coding which pursues higher reconstructed quality of CSI images using lower coding rate plays a very important role in above process [3]. To the best of authors' knowledge, for CSI images, the acquisition [4, 5], transmission [6, 7], display [8, 9], and analysis [1, 2, 10-13], have been widely studied, while the coding has not been addressed in literature.

Currently, high-definition (HD) and ultra HD (UHD) images and equipment have been widely used in CSI. Targeting HD and UHD image/video engineering

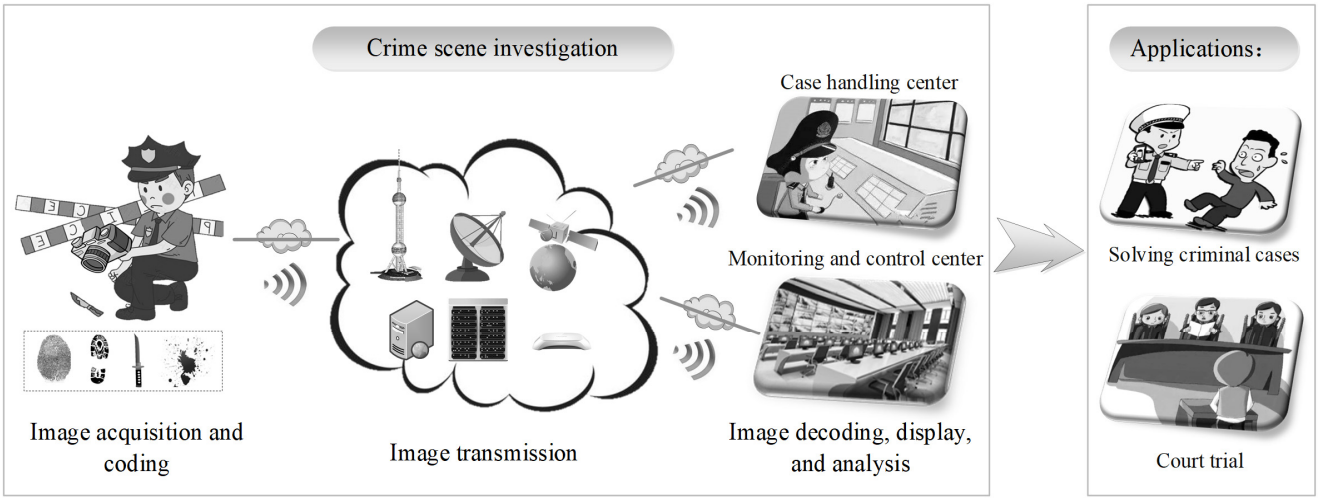


Fig. 1 Illustration of a crime scene investigation application

applications, high efficiency video coding (HEVC) standard shows superior compression performance and has been widely used [14, 15]. The HEVC test model, i.e., HM, supports three profiles [16], called the Main, Main 10, and Main Still Picture profiles, of which the Main Still Picture profile is designed for image coding. Compared with the previous image coding standards, such as JPEG [17] and JPEG2000 [18], HEVC can significantly improves the image coding efficiency [19, 20]. Due to the high coding efficiency, HEVC still image coding technique attracts the attention of researchers [21] and has been widely used in the latest products of many companies, such as Apple (iOS11, macOS High Sierra (10.13), and more recent versions), Microsoft (Windows10 RS4 and more recent versions), Google (Android P and more recent versions), etc.

The coding efficiency of HEVC is closely related to the selected quantization parameter (QP) for coding different basic units. In [22], a QP selection technique is designed, which is generally considered as the default configuration or a recommended solution for HM [16]. But experimental results show that the coding rate-distortion (R-D) performance of the QP selection technique in [22] is not optimal [23]. To improve the coding efficiency, there have been various QP selection techniques [23-35] designed for HEVC.

In [23], a QP selection technique which considers the R-D dependency between adjacent temporal layers was proposed. In [36], a linear inter-frame distortion dependency model which represents the distortion dependency relationship between the current coding picture and its nearest reference picture was initially established with the slope being a constant value. Based on this model, related QP selection techniques were designed in [24-26] for different coding structures.

In [27], an adaptive QP selection technique is proposed for the random-access structure by taking the inter-frame reference dependency into account as well. And similar to [36], another novel linear inter-frame distortion dependency model which only considers the nearest reference picture is proposed. However, unlike the model in [36], the slope of the established model in [27] is not constant but adapt itself to the energy of the prediction residuals. The linear inter-frame distortion dependency model established in [27] is further used to design effective QP selection techniques for the low-delay structure [28, 29].

The QPC techniques in [24-29] only take the nearest reference frame into consideration. Theoretically, the more reference frames are considered, the more accurately reference dependency can be modeled, and the higher R-D performance of a QPC technique can be achieved. Inspired by this, an QPC technique which considered all reference frames was proposed in [30].

In [31] and [32], H. Yin et. al reflected the inter-frame reference dependency by the linear relationship between the QP change and the resulting change of distortion. Then, based on the linear relationship, an adaptive QP selection technique for global rate distortion optimization (RDO) was putted forward. In [33], H. Yin et. al proposed a temporal distortion propagation model by quantitatively evaluating the temporal distortion

dependency. Then, the optimal QP is searched using the simplified dynamic programming to achieve global optimization.

During the standardization of the HEVC, two proposals related to the QP selection technique were also put forward [34, 35]. In [34], the optimal QP is selected according to a logarithmic relationship between QP and Lagrangian multiplier. As known that human eyes are sensitive to degradation in flat regions, but insensitive to complex regions [35]. Based on this perception characteristic, an QP selection technique that describes how to scale the QP according to the spatial activity of the coding blocks was proposed in [35]. Currently, the QP selection techniques in [34] and [35] have been recommended by HEVC and have also been integrated into the latest version of HM.

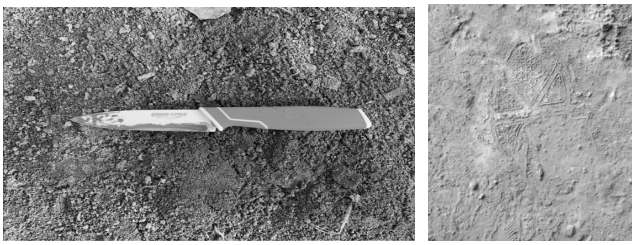
The reported QP selection techniques in [23-35] were all designed for ordinary images, and did not consider the typical characteristics of CSI images. According to the content contained, CSI images are usually divided into the following categories, i.e., “hand (such as fingerprint images, palmprint images, etc.)”, “foot (such as shoeprint images, tire imprint images, etc.)”, “tool (such as dagger images, hammer images, etc.)”, and “others (such as blood images, semen spot images, etc.)”. As widely acknowledged, result from the human eye selective attention mechanism and some special image application requirements (such as CSI application), different regions within an image have significantly different importance [37-39]. In CSI images, the foreground regions, such as the fingerprint regions, shoeprint regions, crime tool regions, etc., contain significantly important and sometimes even unique information for solving crime cases and court trials [39]. So, the foreground regions are usually considered as the salient regions, while the background regions are considered as the non-salient regions. And the saliency information needs to be considered in the design of matched QP selection techniques for CSI images [37- 39].

In [39], a salient region acquisition method for the CSI tool images (named as CSIT-SRA) was proposed. Following this work, in this paper, an effective QP selection technique for CSI tool images, referred to as CSIT-QP, is designed using the Lagrangian R-D optimization theory [40-42]. The main contribution of this paper is to develop a novel two-step Lagrangian R-D optimization method for selecting the matched QPs for different regions within a CSI tool image according to their saliency information and R-D performance.

It should be noted that, theoretically, the proposed two-step Lagrangian R-D optimization method is also suitable for the QP selection of other CSI images, such as the fingerprint images, the shoeprint images, etc. However, as shown in Fig. 2, the fingerprints, shoeprints, and other traces at the crime scene are often vague. And the boundaries and regions of these traces are also seriously discontinuous. The above characteristics pose a great challenge for the accurate acquisition of salient regions in the fingerprint images, shoeprint images, and other traces images. Therefore, for the fingerprint images, shoeprint images, and other traces images, the accurate salient regions acquisition method and the matched QP selection technique using the two-step Lagrangian R-D optimization are valuable researches in the future.

The rest of this paper is organized as follows. Section 2 briefly introduces the salient region acquisition for the CSI tool images. The CSIT-QP is presented in Section 3. Experimental results are provided in Section 4. Conclusions are drawn in Section 5.

2 Salient region acquisition for CSI tool images



(a) crime tool (b) shoeprint

Fig. 2 An example of CSI images

An example of CSI tool images is shown in Fig. 2 (a). For the CSI tool image, the tool region is the salient region. Combined with Fig. 1, the typical content characteristics of the CSI tool images are summarized as follows:

- POSITION: the crime tool in the image as the main target to be photographed usually located in the

near middle of the image;

- SIZE: the crime tool is usually the largest object in the CSI tool image;
- NMUBER: a crime tool is usually photographed into one image. Thus, each CSI tool image usually only contain one crime tool region;
- ANGLE: criminal investigators usually photograph the tool target from an approximate vertical view, so the image of crime tool is usually front;
- BOUNDARY: the crime tools, such as dagger, axe, wrench, etc. are all rigid objects with clear, continuous, and stable boundaries.

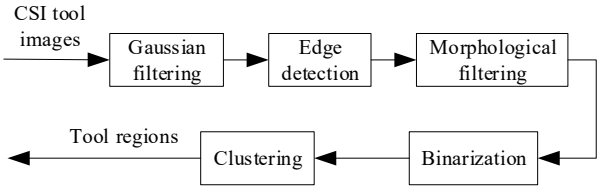


Fig. 3 Flowchart of crime tool regions acquisition

Considering the above content characteristics of the CSI tool images, the crime tool regions acquisition method is designed as shown in Fig. 3 [39]. The input CSI tool images are processed by Gaussian filtering, edge detection, morphological filtering, binarization, and clustering in turn. These techniques are briefly described below.

In order to reduce the influence of image sensor noise on image quality, the CSI tool image is firstly processed by Gaussian filtering using a 3×3 Gaussian kernel function as shown in (1) where (i,j) represents the coordinates of pixels, σ stands for standard deviation parameter and being set to 0.8 in this paper. The filtered image is obtained by convolving the CSI tool image and the Gaussian kernel function.

$$G(i,j)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{i^2+j^2}{2\sigma^2}\right). \tag{1}$$

The edge detection is then used to get the contour of the objects in the filtered image. The edge detection using structured forests proposed by [43] are adopted which shows superior accuracy and fast processing. Although the performance of [43] is excellent, but this method still has the problem that some of the detected edges are not closed for the crime tool regions. In order to further obtain continuous edges, the image is further processed by morphological filtering using a 5×5 structure factor after edge detection. After the morphological filtering, binarization image is obtained by using adaptive binary threshold. In the binarization image, the luminance values of pixels at the edge are set to 255, and the luminance values of other pixels are set to 0. Then the binarization image is further processed by the eight neighboring clustering operation using the pixel as the basic unit [44], and the crime tool region is finally located.

Based on the above process, the pixels belong to crime tool regions can be obtained. In the proposed CSIT-QP, the basic unit for QP selection is largest coding unit (LCU, a 64×64 squared pixel block). For the convenience of following descriptions, the LCUs with any pixels belong to the crime tool regions are defined as the A-LCUs, and the other LCUs are defined as the B-LCUs. The following rules are used to determine whether an LCU belongs to the crime tool regions. First, all A-LCUs belong to the crime tool regions. Second, considering the strong coding reference dependence among neighboring blocks caused by intra prediction technique in HEVC, if a B-LCU locates at the left, top, or left-top of an A-LCU, then the B-LCU also belongs to the crime tool regions.

3 Quantization parameter selection for CSI tool images

In this section, the spirit of the two-step Lagrangian R-D optimization is firstly analyzed. Then, the detailed process of the CSIT-QP is further introduced.

3.1 Spirit of the two-step Lagrangian R-D optimization

The problem of QP selection can be formulated as an R-D optimization problem to minimize the coding distortion at a given rate budget [40]. So, the target of QP selection for a picture is expressed as,

$$\min\sum D_{PQ,P}\quad \text{s.t.}\sum R_{PQ,P}<R_T, \tag{2}$$

where $D_{PQ,P}$ and $R_{PQ,P}$ represent the distortion and rate of the coding the entire picture using a given QP (denoted as $QP_{PQ,P}$ which is usually obtained from the default setting in the configuration file by the coder),

respectively. R_T denotes the rate budget.

As mentioned above, all regions within a CSI tool picture are coded using the same QP are obviously not optimal. The crime tool regions and the background regions need to be separately optimized and more rate should spend on the crime tool regions to present more information compared with the background regions. Based on (2), the target of the CSIT-QP is expressed by (3) and (4),

$$\min \sum D_{CQ,I} \quad \text{s.t.} \sum R_{CQ,I} < R_{T,1}, \tag{3}$$

$$\min \sum D_{CQ,nI} \quad \text{s.t.} \sum R_{CQ,nI} < \left(R_T - R_{T,1}\right), \tag{4}$$

where $D_{CQ,I}$ and $R_{CQ,I}$ represent the distortion and rate for coding crime tool regions using the CSIT-QP, respectively. $D_{CQ,nI}$ and $R_{CQ,\mathrm{nI}}$ represent the distortion and rate for coding background regions using the CSIT-QP, respectively. $R_{T,1}$ denotes the rate budget for crime tool regions.

It is easy to see from (3) and (4) that, the CSIT-QP uses a two-step optimization in which the more important tool regions are firstly optimized. Since the overall rate spent on a picture should better stay on approximate after using different QP selections, the $R_{CQ,I}$ and $R_{CQ,nI}$ satisfy the following relationships,

$$R_{CQ,I} = R_{PQ,I} + \Delta R, \tag{5}$$

$$R_{CQ,nI} = R_{PQ,nI} - \Delta R, \tag{6}$$

where $R_{PQ,I}$ and $R_{PQ,nI}$ represent the rate for coding crime tool and background regions using the given same QP, i.e., $QP_{PQ,P}$, respectively. ΔR denotes a "moved" rate, which is a positive number. The spirit of the CSIT-QP is to "move" some rate from the background regions to the crime tool regions, so as to improve the perceptual quality of crime tool regions without losing or slightly losing the perceptual quality of background regions.

Regarding to the selection of QP, (7) and (8) are obtained according to (5) and (6).

$$QP_{CQ,I} = QP_{PQ,P} - x, \tag{7}$$

$$QP_{CQ,nI} = QP_{PQ,P} + y, \tag{8}$$

where $QP_{CQ,I}$ and $QP_{CQ,nI}$ represent the selected QPs for crime tool and background regions using the CSIT-QP, respectively. x and y stand for the QP offsets, which are positive integers.

3.2 Optimal QP offset for crime tool regions

Applying the Lagrangian optimization [40], the constrained optimization problem in (3) can be converted into the following unconstrained form,

$$J_{CQ,I} = D_{CQ,I} + \lambda_{CQ,I} R_{CQ,I}, \tag{9}$$

where $J_{CQ,I}$ and $\lambda_{CQ,I}$ stand for the cost and the Lagrangian multiplier of the crime tool regions, respectively.

Since the QP and the Lagrangian multiplier have a logarithm relationship [34], i.e.,

$$QP = a \ln(\lambda) + b, \tag{10}$$

where a and b are model parameters. [34] found that the optimal values of a and b are 4.2005 and 13.7122, respectively, for HEVC codec. Solving $\lambda_{CQ,I}$ from (7) and (10) leads to (11) as follows,

$$\lambda_{CQ,I} = \exp\left(\frac{QP_{PQ,P} - x - b}{a}\right). \tag{11}$$

In the following contents, experiments were carried out to explore the relationships between rate, distortion, and Lagrangian multiplier. In this paper, all standard test images are from the CSI image database (named as CIIP-CSID) which are jointly constructing by Xi'an University of Posts and Telecommunications and Shaanxi Provincial Public Security Department for academic research [2]. To our best knowledge, CIIP-CSID is the largest CSI image database at present. Currently, CIIP-CSID includes 19363 CSI images, divided into 52 categories, including crime tool, fingerprint, shoeprint, tyre marks, blood stains, etc.

CSI tool images include crime tool and background regions. And they were further segmented to obtain the crime tool images (corresponds to the crime tool regions in the CSI tool images) and background images

(corresponds to the background regions in the CSI tool images). The crime tool images and background images were both coded using HM16.0 [16], where the main encoding parameters are listed in Table 1, and the other parameters are set as the same as those in the default configuration file of encoder_intra_main [16]. The QPs are set to 27, 32, 37, 42, respectively. After encoding each image, each QP corresponds to a set value of distortion, rate, and Lagrangian multiplier.

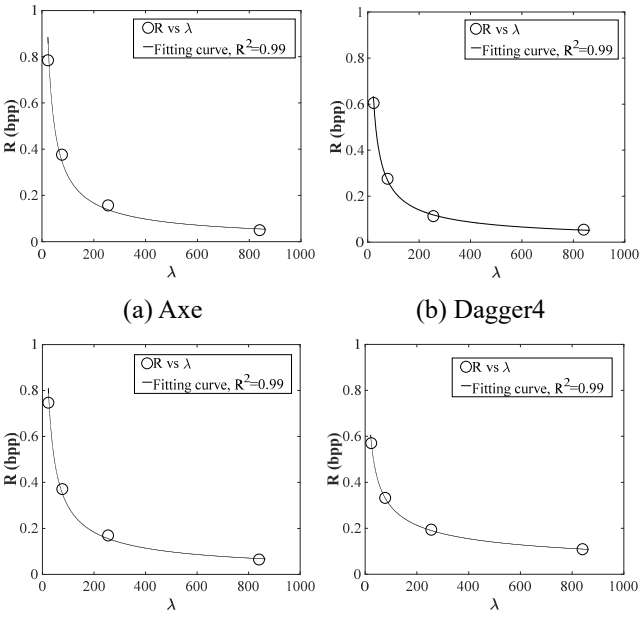
Table 1 Main encoding parameters

parameters	settings
encoding structure	intra-only configuration
profile	main-still-picture
fastsearch	1
searchrange	64
SAO	ON

The rate is measured by the bit per pixel (bpp) [45] as shown in (12).

$$bpp = \frac{N_{bit}}{wh}, \tag{12}$$

where w and h denote the width and height of the image, respectively. N_{bit} means the number of coding bits of the image.



(a) Axe (b) Dagger4

(c) Hammer (d) CutterKnife

Fig. 4 Relationships of rate and Lagrangian multiplier for crime tool images

Fig. 4 intuitively shows the relationships of rate and Lagrangian multiplier with four representative crime tool images. The fitting functions are shown by the dashed lines with the fitting accuracy being evaluated by the square of correlation coefficient (R^2) [45] as shown in the legends. It is observed that the rate and the Lagrangian multiplier are observed to obey a power relationship, i.e.,

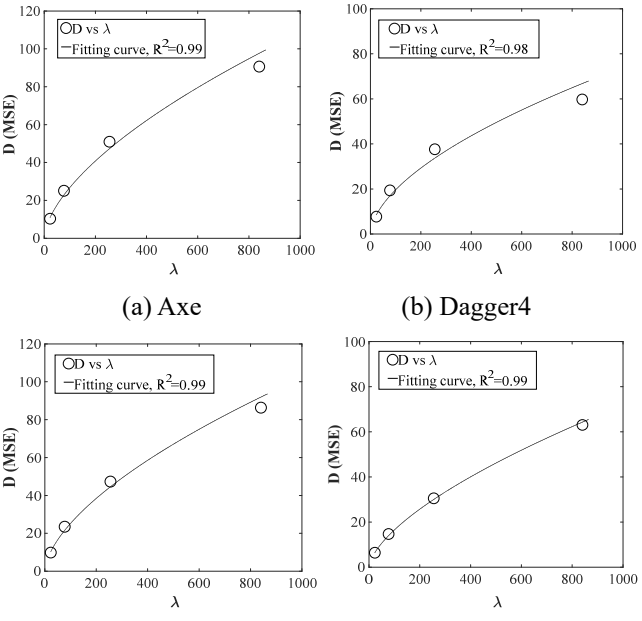
$$R = \gamma\lambda^{\delta}, \tag{13}$$

where γ and δ are the model parameters which are related to the image content characteristics.

The distortion is measured by the mean squared error (MSE) of the luminance component. The relationships of distortion and Lagrangian multiplier are shown in Fig. 5. The relationships shown in Fig. 5 can be fitted by the models in (14),

$$D = \alpha\lambda^{\beta}, \tag{14}$$

where α and β are the model parameters which are related to the image content characteristics.



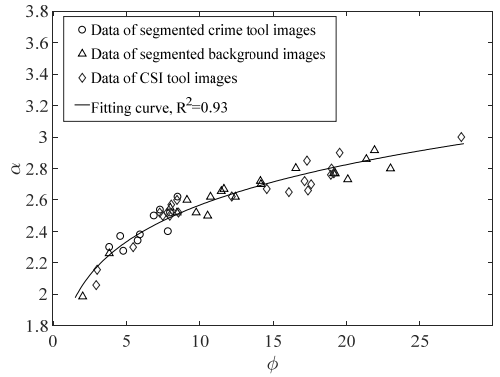
(a) Axe (b) Dagger4

(c) Hammer (d) CutterKnife

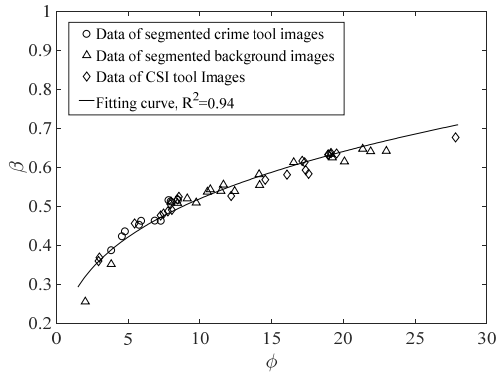
Fig. 5 Relationships of distortion and Lagrangian multiplier for crime tool images.

Substituting (11), (13), and (14) into (9) yields the cost function $J_{CQ,I}$ as shown in (15),

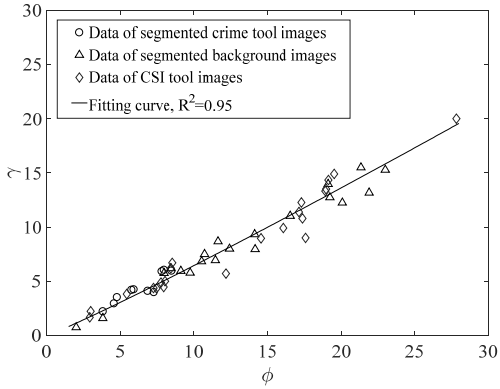
$$J_{CQ,I} = \alpha_{CQ,I}\exp\left(\left(\frac{QP_{PQ,P}-x-b}{a}\right)\beta_{CQ,I}\right) + \gamma_{CQ,I}\exp\left(\left(\frac{QP_{PQ,P}-x-b}{a}\right)\left(\delta_{CQ,I}+1\right)\right). \tag{15}$$



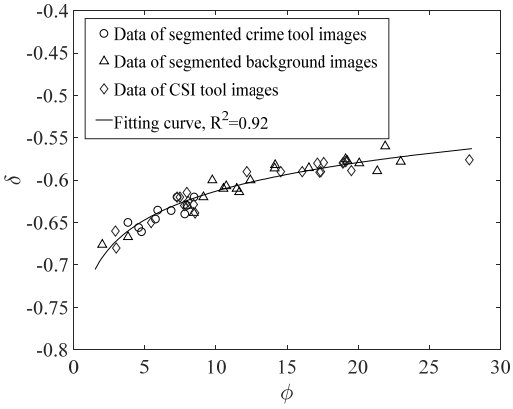
(a) α vs ϕ



(b) β vs ϕ



(c) γ vs ϕ



(d) δ vs ϕ

Fig. 6 Relationships of α, β, γ, δ, and standard deviation (ϕ)

As mentioned above, the model parameters of α, β, γ, and δ are all related to the content characteristics of images. Experiment was further carried out to explore the relationship between these model parameters and the content characteristics of images.

The standard CSI tool images, segmented crime tool images, and segmented background images were coded using HM16.0 [16], where the main encoding parameters are listed in Table 1, and the other parameters are set as the same as those in the default configuration file of encoder_intra_main [16]. The values of D, R, and λ are collected from the coding results with the QPs are set to 27, 32, 37, 42, respectively. α and β are obtained by fitting (14) using the least-squares method with the collected D and λ. Similarly, γ and δ are obtained by fitting (13) using the least-squares method with the collected R and λ.

As shown in Fig. 6, the values of α, β, γ, δ, and the standard deviation satisfy the relationships in (16)-(19). In this paper, the standard deviation of pixels' luminance components is used to measure the texture complexity of images. The basic unit for calculating standard deviation is 4×4 pixel block.

$$\alpha = a_1 \phi^{b_1}, \tag{16}$$

$$\beta = a_2 \phi^{b_2}, \tag{17}$$

$$\gamma = a_3 \phi^{b_3}, \tag{18}$$

$$\delta = a_4 \ln(\phi) + b_4, \tag{19}$$

where ϕ represents the standard deviation. The values of $a_1 \sim a_4$ and $b_1 \sim b_4$ are shown in Table 2, respectively, according to the fitting results in Fig. 6.

Table 2 Values of $a_1 \sim a_4$ and $b_1 \sim b_4$

parameters	values	parameters	values
a_1	1.8715	b_1	0.1374
a_2	0.2589	b_2	0.3023
a_3	0.5384	b_3	1.0782
a_4	0.0487	b_4	-0.7251

Substituting (16)-(19) into (15), the cost function is revised as (20),

$$\hat{x} \approx \left\lfloor \theta_1 + \frac{aa_1a_2\phi_I^{\,b_1+b_2}\exp\left(\left(\frac{QP_{PQ,P}-\theta_1-b}{a}\right)a_2\phi_I^{\,b_2}\right)+aa_3\phi_I^{\,b_3}\left(\begin{matrix}a_4\ln(\phi_I)+\\ b_4+1\end{matrix}\right)\exp\left(\left(\frac{QP_{PQ,P}-\theta_1-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_I)+\\ b_4+1\end{matrix}\right)\right)}{a_1a_2^{\,2}\phi_I^{\,b_1+2b_2}\exp\left(\left(\frac{QP_{PQ,P}-\theta_1-b}{a}\right)a_2\phi_I^{\,b_2}\right)+a_3\phi_I^{\,b_3}\left(\begin{matrix}a_4\ln(\phi_I)+\\ b_4+1\end{matrix}\right)^2\exp\left(\left(\frac{QP_{PQ,P}-\theta_1-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_I)+\\ b_4+1\end{matrix}\right)\right)}\right\rfloor. \quad (21)$$

$$J_{CQ,I}=a_1\phi_I^{\,b_1}\exp\left(\left(\frac{QP_{PQ,P}-x-b}{a}\right)a_2\phi_I^{\,b_2}\right)+ a_3\phi_I^{\,b_3}\exp\left(\left(\frac{QP_{PQ,P}-x-b}{a}\right)\left(a_4\ln(\phi_I)+b_4+1\right)\right), \quad (20)$$

where ϕ_I represents the standard deviation of crime tool regions in the CSI tool images.

Thus, set the partial derivative of $J_{CQ,I}$ with respect to x, i.e., $\partial J_{CQ,I}/\partial x$, to zero to obtain the optimum value of x (i.e., $\hat{x}$) which minimizes the cost function $J_{CQ,I}$. Let $f(x)=\partial J_{CQ,I}/\partial x$, and the approximate solution of $\partial J_{CQ,I}/\partial x=0$ is $x\approx\theta_1-f(\theta_1)/f(\theta_1)'$ (i.e., (21)) using the Newton-Raphson method [46] where $\lfloor\ \rfloor$ indicates rounding down and θ_1 is set according to (22).

$$\theta_1=\theta_{1,b}+\omega_1\exp\left(\left(\ln\left(QP_{PQ,P}\right)-\omega_2\right)^2\omega_3^{-1}\right), \quad (22)$$

where the value of $\theta_{1,b}$ is determined according to Table 3. ω_1, ω_2, and ω_3 are set to -3.1391, 3.8067, and -0.0042, respectively

Table 3 Values of $\theta_{1,b}$

ϕ_I	$\theta_{1,b}$
$[0,5.5)$	-6
$[5.5,9.7)$	-5
$[9.7,17)$	-4
$[17,25)$	-3
$[25,+\infty)$	-2

3.3 Optimal QP offset for background regions

For the background regions, applying the Lagrangian optimization [40], the constrained optimization problem in (4) can be converted into the following unconstrained form as shown in (23),

$$J_{CQ,nI}=D_{CQ,nI}+\lambda_{CQ,nI}R_{CQ,nI}, \quad (23)$$

where $J_{CQ,nI}$ and $\lambda_{CQ,nI}$ denote the cost and the Lagrangian multiplier of the background regions, respectively.

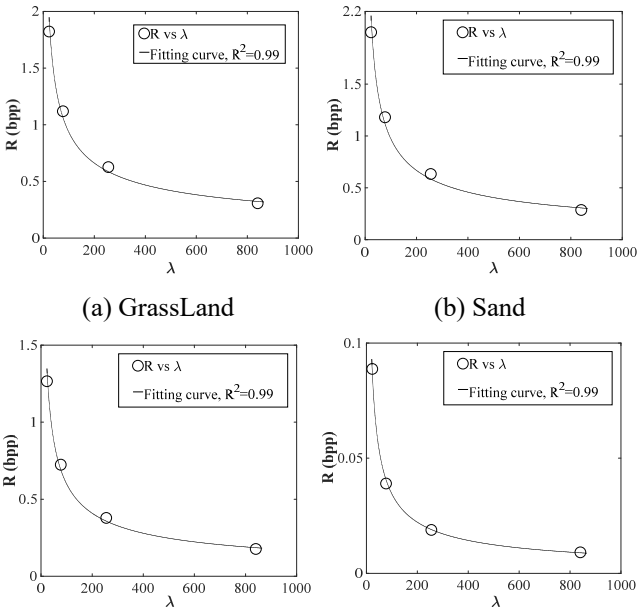
In this paper, the rate is measured by the bit per pixel (bpp) [45]. So the $R_{CQ,nI}$ is expressed as,

$$R_{CQ,nI}=\frac{R_{PQ,P}N_P-R_{CQ,I}N_I}{N_{nI}}, \quad (24)$$

where N_P, N_I, and N_{nI} denote the total number of pixels in the CSI tool image, the crime tool regions, and the background regions, respectively.

Substituting (24) into (23), (23) is further updated to (25),

$$J_{CQ,nI}=D_{CQ,nI}+\lambda_{CQ,nI}\left(\frac{R_{PQ,P}N_P-R_{CQ,I}N_I}{N_{nI}}\right). \quad (25)$$



(a) GrassLand (b) Sand

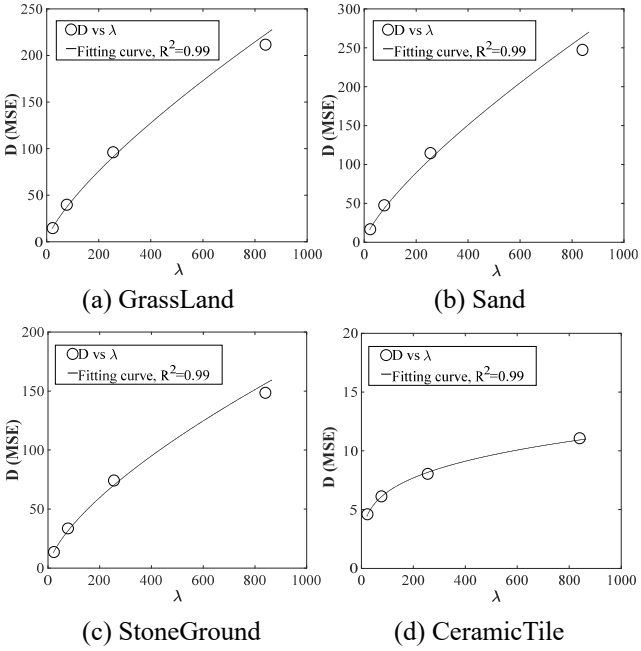
(c) StoneGround (d) CeramicTile

Fig. 7 Relationships of rate and Lagrangian multiplier for background images

Based on the previous experiments which has been described in detail in Section 3.2, the data of rate, distortion, and Lagrangian multiplier have been obtained for the background images. Figs 7 and 8

$$J_{CQ,nI}=a_1\phi_{nI}^{\ b_1}\exp\left(\left(\begin{matrix}QP_{PQ,P}+\\ \dfrac{y-b}{a}\end{matrix}\right)a_2\phi_{nI}^{\ b_2}\right)+\exp\left(\begin{matrix}QP_{PQ,P}+\\ \dfrac{y-b}{a}\end{matrix}\right)\left(\begin{matrix}a_3\phi_P^{\ b_3}\exp\left(\left(\dfrac{QP_{PQ,P}-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_P)\\ +b_4\end{matrix}\right)\right)N_P-\\ a_3\phi_I^{\ b_3}\exp\left(\left(\dfrac{QP_{PQ,P}-\hat{x}-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_I)\\ +b_4\end{matrix}\right)\right)N_I\end{matrix}\right)/N_{nI}. \quad (26)$$

$$\hat{y}\approx\left\lfloor\theta_2-\frac{aa_1a_2\phi_{nI}^{\ b_1+b_2}\exp\left(\left(\dfrac{QP_{PQ,P}+\theta_2-b}{a}\right)\left(a_2\phi_{nI}^{\ b_2}-1\right)\right)N_{nI}+a\left(\begin{matrix}a_3\phi_P^{\ b_3}\exp\left(\left(\dfrac{QP_{PQ,P}-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_P)\\ +b_4\end{matrix}\right)\right)N_P-\\ a_3\phi_I^{\ b_3}\exp\left(\left(\dfrac{QP_{PQ,P}-\hat{x}-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_I)\\ +b_4\end{matrix}\right)\right)N_I\end{matrix}\right)}{a_1a_2^{\ 2}\phi_{nI}^{\ b_1+2b_2}\exp\left(\left(\dfrac{QP_{PQ,P}+\theta_2-b}{a}\right)\left(a_2\phi_{nI}^{\ b_2}-1\right)\right)N_{nI}+\left(\begin{matrix}a_3\phi_P^{\ b_3}\exp\left(\left(\dfrac{QP_{PQ,P}-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_P)\\ +b_4\end{matrix}\right)\right)N_P-\\ a_3\phi_I^{\ b_3}\exp\left(\left(\dfrac{QP_{PQ,P}-\hat{x}-b}{a}\right)\left(\begin{matrix}a_4\ln(\phi_I)\\ +b_4\end{matrix}\right)\right)N_I\end{matrix}\right)}\right\rfloor. \quad (27)$$



(a) GrassLand (b) Sand

(c) StoneGround (d) CeramicTile

Fig. 8 Relationships of distortion and Lagrangian multiplier for background images

intuitively show the relationship of rate and Lagrangian multiplier and the relationship of distortion and Lagrangian multiplier, respectively, of four representative background images. It is easily observed that the rate vs the Lagrangian multiplier and the distortion vs the Lagrangian multiplier are all obey the power relationships, i.e., the model of (13) and (14) also hold for the background images.

Similar to the operation in the previous sub-section, substituting (8), (10), (13), (14), and (16)-(19) into (25) yields the cost function $J_{CQ,nI}$ as shown in (26), where ϕ_{nI} and ϕ_P stand for the standard deviation of the background regions and the CSI tool image, respectively. The $\hat{y}$ is finally obtained using the Newton-Raphson method as shown in (27), where the values of θ_2 is determined according to (28) where ω_4 and ω_5 are set to 9.8923 and 0.000024, respectively.

$$\theta_2=\omega_4+\omega_5 QP_{PQ,P}^{\ 3}. \quad (28)$$

Finally, the implementation of CSIT-QP is described as follows:

1) For a given CSI tool image, the LCUs which belong to the crime tool regions are determined using the operation in Section 2.

2) $QP_{PQ,P}$ are obtained from the default setting in the configuration file by the coder. The LCUs which belong to the crime tool regions are coded using $QP_{PQ,P}-\hat{x}$, and the other LCUs are coded using $QP_{PQ,P}+\hat{y}$.

4 Experimental results

In this section, the CSIT-SRA described in Section 2 are first evaluated in Subsection 4.1. As widely acknowledged, R-D performance is a crucial metric for evaluating the image coding efficiency. Higher R-D performance means higher reconstructed image quality with lower rate. Reconstructed image quality can be evaluated by the objective or the subjective measures [47, 48]. Therefore, the R-D performance of CSIT-QP

proposed in Section 3 is evaluated using both of the objective quality measure and the subjective quality measure in Subsections 4.2 and 4.3, respectively.

4.1 Evaluation of salient region acquisition

Subjective experiments were carried out following the guidelines specified by ITU-R BT 500-13 standard [49], including thirty experts in CSI but unfamiliar with the proposed algorithm were involved as the observers. The observers included 19 men and 11 women, aged between 22-50. The observation distance is approximate 4 times the vertical resolution of the image [44].

As shown in Table 4, ten standard CSI tool images from CIIP-CSID with various kinds of crime tools and backgrounds were selected as the test images. The spatial resolution, color space, and bit depth of the test images are 1080P, YCbCr4:2:0, and 8, respectively. The test images are displayed on the screen one by one in a random order. Each observer was asked to specify the boundaries of crime tool regions as highlighted by curves. And, based on the boundary information, pixels located in the crime tool region are easily determined for each test image. Considering the subjective differences of observers, if equal to or greater than 75% of observers [50] divide a pixel into the crime tool regions, then the pixel is judged to be in the crime tool regions, otherwise it is located in background regions.

For each test images, based on the above subjective experiments, the pixels belong to crime tool regions can be obtained, and these pixels form a set $\mathbf{S}$. Using the CSIT-SRA, the pixels belong to crime tool regions can also be obtained, and these pixels form a set $\mathbf{E}$. Finally, according to the $\mathbf{S}$ and $\mathbf{E}$, the false alarm rate (FAR) and miss rate (MR) [44] are computed to evaluate the performance of CSIT-SRA. FAR represents the percentage of pixels belonging to $\mathbf{E}$ but not to $\mathbf{S}$. MR represents the percentage of pixels belonging to $\mathbf{S}$ but not to $\mathbf{E}$.

The FAR and MR of test images are show in Table 4. For all test images, the average FAR and MR are only 8.1% and 6.5%, respectively, which proves the effectiveness of the CSIT-SRA. Moreover, CSIT-SRA performs better for images with simple backgrounds, such as Dagger3inSoil, because these images have more significant boundaries between crime tool regions and background regions.

Table 4 The FAR and MR of test images

CSI tool image	FAR (%)	MR (%)
AxeinStoneGround	13.8	12.2
CutterKnifeinPlasterBoard	3.5	2.7
Dagger1inCarpet	10.5	9.8
Dagger2inBrickFloor	9.2	3.1
Dagger3inSoil	3.2	3.5
Dagger4inGrassLand	10.6	4.7
HammerinGrassLand	11.1	12.6
KitchenKnifeinSand	5.8	8.7
ScissorsinPlasterBoard	2.8	2.1
WrenchinGranssLand	10.9	5.7
Ave.	8.1	6.5

4.2 R-D performance in terms of PSNR

In order to evaluate the proposed CSIT-QP, the following QP selection techniques are included in the comparison experiments. The first compared method is the QP selection technique in [22] which is generally considered as the default configuration in HM. Currently, the QP selection techniques in [34, 35] have also been integrated into HM. The QP selection techniques reported in [23-27, 29, 30] are also state-of-the-art methods. Therefore, the CSIT-QP presented in this paper were compared with the QP selection techniques in [22-27, 29, 30, 34, 35]. It should be noted that, for the still images (intra frames), the QP selection techniques in [22, 23, 25-27, 29] set QP in the same way. Test images were coded using HM16.0 and different QP selection techniques, where the main encoding parameters are the same as shown in Table 1.

In image coding, one of the most commonly used objective quality measures is peak signal to noise ratio (PSNR). Table 5 shows the PSNR of different QP selection techniques with four coding rate levels. It is easy to see that, compared with other QP selection techniques, the CSIT-QP can improve the PSNR of the crime tool region and reduce the PSNR of the background region. PSNR only considers the influence of pixel value (the difference between the original image pixel value and the reconstructed image pixel value) on

Table 5 PSNR (dB) of different QP selection techniques (T: crime tool regions, B: background regions)

Rate level	CSI tool image	[22, 23, 25-27, 29]		[34]		[24]		[35]		[30]		CSIT-QP	
		T	B	T	B	T	B	T	B	T	B	T	B
High	AxeinStoneGround	37.16	36.57	37.57	37.02	37.96	37.44	37.40	36.49	39.68	39.33	40.49	33.38
	CutterKnifeinPlasterBoard	37.25	36.13	37.60	36.56	38.08	37.04	37.11	36.10	39.92	39.04	40.75	33.95
	Dagger1inCarpet	38.16	36.49	38.52	36.93	38.96	37.35	38.50	36.42	40.64	39.22	42.26	34.22
	Dagger2inBrickFloor	37.36	35.92	37.83	36.48	38.21	36.87	37.68	35.95	40.08	38.92	40.94	33.31
	Dagger3inSoil	37.80	35.89	38.33	36.52	38.74	36.91	38.04	35.98	40.67	39.03	41.55	31.20
	Dagger4inGrassLand	37.42	36.48	37.82	36.96	38.27	37.38	37.45	36.55	39.99	39.30	41.68	33.26
	HammerinGrassLand	37.96	37.18	38.23	37.51	38.67	37.95	38.19	37.14	40.21	39.63	40.96	35.38
	KitchenKnifeinSand	39.05	36.16	39.47	36.78	39.81	37.16	39.37	36.45	41.40	39.26	42.83	32.19
	ScissorsinPlasterBoard	37.13	36.15	37.45	36.58	37.93	37.04	37.13	36.10	39.65	39.02	40.45	34.01
	WrenchinGranssLand	36.88	36.82	37.29	37.24	37.69	37.67	36.70	36.84	39.57	39.49	41.36	33.86
Medium -high	AxeinStoneGround	33.29	32.34	33.67	32.75	34.08	33.17	33.51	32.28	35.54	34.81	36.34	29.79
	CutterKnifeinPlasterBoard	33.40	31.93	33.65	32.25	34.13	32.72	33.26	31.91	35.61	34.33	36.39	30.18
	Dagger1inCarpet	34.36	32.44	34.71	32.85	35.13	33.24	34.63	32.40	36.57	34.79	38.15	29.99
	Dagger2inBrickFloor	33.25	31.37	33.64	31.86	34.04	32.27	33.52	31.40	35.64	34.04	36.47	28.83
	Dagger3inSoil	33.29	30.92	33.80	31.48	34.20	31.93	33.53	30.99	35.95	33.87	36.91	27.71
	Dagger4inGrassLand	33.48	32.08	33.83	32.50	34.25	32.97	33.51	32.12	35.81	34.67	37.45	29.44
	HammerinGrassLand	34.42	33.45	34.69	33.72	35.14	34.21	34.62	33.42	36.49	35.63	37.19	31.60
	KitchenKnifeinSand	34.89	31.10	35.35	31.70	35.78	32.15	35.23	31.37	37.42	34.12	39.06	27.80
	ScissorsinPlasterBoard	33.53	32.02	33.79	32.34	34.21	32.78	33.52	31.98	35.58	34.38	36.33	30.29
	WrenchinGranssLand	32.83	32.73	33.15	33.08	33.59	33.55	32.68	32.73	35.15	35.13	36.03	30.31
Medium -low	AxeinStoneGround	29.87	28.62	30.14	28.92	30.53	29.34	30.01	28.55	31.90	30.79	32.57	26.79
	CutterKnifeinPlasterBoard	30.30	28.76	30.50	28.95	30.85	29.32	30.23	28.73	32.01	30.54	32.65	27.96
	Dagger1inCarpet	30.81	28.77	31.03	29.07	31.49	29.50	30.93	28.72	32.92	30.95	33.57	26.93
	Dagger2inBrickFloor	29.79	27.54	29.98	27.83	30.40	28.24	29.90	27.53	31.78	29.74	32.48	26.06
	Dagger3inSoil	29.25	26.59	29.62	26.99	30.02	27.40	29.44	26.62	31.62	29.10	32.41	24.21
	Dagger4inGrassLand	29.83	28.11	30.11	28.41	30.55	28.89	29.86	28.12	31.99	30.44	32.72	26.12
	HammerinGrassLand	31.20	30.07	31.36	30.26	31.85	30.73	31.32	30.02	33.09	32.06	33.76	28.75
	KitchenKnifeinSand	30.73	26.51	31.05	26.90	31.54	27.39	30.96	26.68	33.20	29.21	34.87	24.12
	ScissorsinPlasterBoard	30.65	28.87	30.79	29.05	31.19	29.42	30.63	28.83	32.27	30.64	32.83	28.09
	WrenchinGranssLand	29.29	29.02	29.51	29.28	29.99	29.75	29.22	29.01	31.31	31.21	32.03	27.26
Low	AxeinStoneGround	26.68	25.19	27.08	25.63	27.28	25.83	26.79	25.16	28.52	27.20	27.85	23.53
	CutterKnifeinPlasterBoard	28.12	26.59	28.36	26.79	28.51	26.94	28.06	26.58	29.34	27.75	29.28	25.97
	Dagger1inCarpet	27.53	25.35	27.84	25.71	28.15	25.98	27.60	25.30	29.45	27.34	29.34	23.84
	Dagger2inBrickFloor	27.41	24.61	27.56	24.87	27.77	25.06	27.46	24.61	28.65	26.19	28.58	23.79
	Dagger3inSoil	25.79	22.89	26.22	23.46	26.39	23.54	25.92	22.89	27.76	24.99	27.69	20.92
	Dagger4inGrassLand	26.43	24.48	26.76	24.86	27.09	25.17	26.41	24.48	28.46	26.60	28.44	22.61
	HammerinGrassLand	28.18	26.87	28.46	27.16	28.77	27.48	28.30	26.84	29.98	28.73	29.96	25.56
	KitchenKnifeinSand	26.86	22.55	27.23	23.03	27.52	23.25	26.97	22.67	29.09	24.80	29.82	20.81
	ScissorsinPlasterBoard	28.46	26.69	28.63	26.89	28.85	27.03	28.42	26.69	29.65	27.87	29.62	26.06

WrenchinGranssLand	26.09	25.55	26.45	25.93	26.73	26.21	26.04	25.57	28.01	27.58	27.96	23.91
Ave.	32.15	30.45	32.48	30.83	32.86	31.21	32.25	30.46	34.31	32.79	35.05	28.35
Ave. of two regions ($PSNR_{TandB}$)	31.30		31.65		32.03		31.35		33.55		31.70	

image quality, but does not consider the influence of pixel position (i.e., the crime tool regions or the background regions) on image quality. In other words, PSNR considers that pixels in different regions are equally important to image quality. Following the idea, the average PSNR of the crime tool region and the background region (noted as $PSNR_{TandB}$) are given in the last row of Table 5 to evaluate the overall image quality. Although the QP selection techniques in [24, 30] can obtain the highest $PSNR_{TandB}$, the corresponding rate are also the highest which increase by 11.5% and 37.1%, respectively, compared with the rate of the QP selection techniques in [22, 23, 25-27, 29]. The CSIT-QP can achieve the third highest $PSNR_{TandB}$, and the rate is also the lowest which reduces by 4.8% compared with the rate of QP selection techniques in [22, 23, 25-27, 29]. Please refer to the Table 7 for the detailed rate data.

From the above analysis, the CSIT-QP achieves high R-D performance. However, the traditional R-D performance analysis in terms of PSNR is not enough for evaluating the performance of the proposed method due to the PSNR has a poor consistency with the image quality perceived by human eyes [47, 48]. In addition, regional importance is not considered in the calculation of the PSNR. As mentioned above, human eye selective attention mechanism and some special image application requirements makes different regions have significantly different effects on the overall image quality. Especially for the CSI application, the effect of different regions on image quality should be considered.

Human eyes as the common receiving end of the reconstructed images makes the quality of human eyes' perception closer to the real image quality. In addition, at present, the comparison, recognition, and other analysis operations for CSI images in the process of crime cases are still mainly completed by criminal investigators, and the automatic information analysis techniques only play an auxiliary role. So, considering the above fact, a subjective image quality evaluation experiment was further carried out for a more accurate evaluation of the perceptual R-D performance using different QP selection techniques.

4.3 R-D performance in terms of subjective experiment evaluation

Subjective experiments were carried out based on the adjectival categorical judgement method recommended by ITU-R BT 500-13 standard [49]. Thirty experts that have been described in Section 4.1 were involved as the observers. In the subjective experiments, the reconstructed images corresponding to the QP selection techniques in [22, 23, 25-27, 29] are set as the benchmark images for observing. And the reconstructed images corresponding to the QP selection techniques in [24, 30, 34, 35], or CSIT-QP are defined as the target images. In each observation, two images of the same scene (i.e., a benchmark image and a target image) were juxtaposed on the screen randomly at left or right. Observers were asked which image (left or right) is better and how much better it is by using quantitative scores listed in Table 6 [50, 51]. Repeating the above observation process for all the image pairs. The average of all observers' scores is taken as the subjective score of each image.

Table 6 Scores for subjective quality evaluation

subjective score	description
0	same quality
1	slightly better
2	better
3	much better

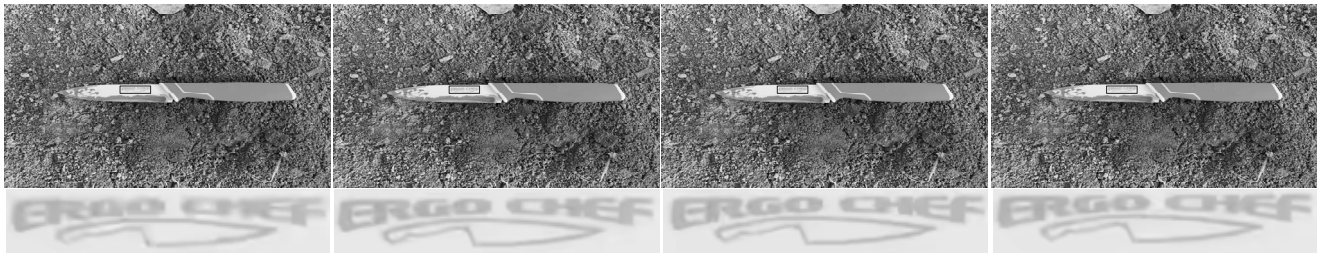
Table 7 shows the perceptual R-D performance using different QP selection techniques. Compared with the QP selection technique in [22, 23, 25-27, 29], the perceptual quality improvement of the QP selection techniques in [24, 34, 35] are not obvious due to they don't consider the influence of regions' difference on image quality. As mentioned above, for the CSI tool images, the perceptual quality improvement of crime tool regions is the key to improve the overall perceptual

Table 7 Perceptual R-D performance of different QP selection techniques (S: subjective score, R: rate (bits))

Rate level	CSI tool image	[22, 23, 25-27, 29]		[34]		[24]		[35]		[30]		CSIT-QP	
		S	R	S	R	S	R	S	R	S	R	S	R
High	AxeinStoneGround	0	3363152	0	3519112	0	3371912	0	3371912	0	4435304	0	3126928
	CutterKnifeinPlasterBoard	0	2528064	0	2673128	0	2840352	0	2552272	0.6	3544872	0.6	2409520
	Dagger1inCarpet	0	3031912	0.1	3179568	0.1	3317040	0	3046608	0.1	3967504	0.2	2980728
	Dagger2inBrickFloor	0	3442912	0	3642584	0	3787080	0	3486248	0.3	4536328	0.3	3210800
	Dagger3inSoil	0	5096560	0	5325440	0.1	5486488	0.1	5149408	0.7	6313424	0.7	4382720
	Dagger4inGrassLand	0	3889936	0	4057456	0.1	4211664	0	3930776	0.7	4945368	0.8	3682312
	HammerinGrassLand	0	2227592	0	2318832	0.2	2449600	0	2237320	0.4	2993840	0.5	2241576
	KitchenKnifeinSand	0	4537360	0	4735456	0.1	4867864	0.1	4653304	0.7	5593856	0.6	4293224
	ScissorsinPlasterBoard	0	2473632	0	2620992	0.1	2787608	0	2499416	0.5	3491208	0.5	2360784
	WrenchinGranssLand	0	3249808	0	3390288	0.2	3550448	0	3270472	0.6	4281808	0.4	3064752
	Ave.	0	3384093	0	3546286	0.1	3698407	0	3419774	0.5	4410351	0.5	3175334
Medium-high	AxeinStoneGround	0	2040776	0.1	2155704	0.3	2278104	0.2	2049448	0.5	2761744	0.9	1940992
	CutterKnifeinPlasterBoard	0	1210928	0	1300136	0.3	1438008	-0.1	1227096	0.6	1935168	0.6	1204144
	Dagger1inCarpet	0	1847520	0.4	1956856	0.5	2058952	0.4	1860984	0.7	2502048	0.8	1709960
	Dagger2inBrickFloor	0	1900336	0.1	2056304	0.3	2189504	0	1935392	1.2	2781624	1.2	1669496
	Dagger3inSoil	0	3285856	0.3	3486688	0.3	3647984	0	3331200	0.7	4337984	0.7	2921144
	Dagger4inGrassLand	0	2497880	0	2619888	0	2756192	0	2522336	0.4	3275000	0.5	2381792
	HammerinGrassLand	0	1320824	0.1	1383400	0.3	1479664	0	1331824	0.6	1807432	0.6	1321272
	KitchenKnifeinSand	0	2966848	0.2	3141592	0.2	3284504	0.2	3066544	0.5	3878728	0.6	2796104
	ScissorsinPlasterBoard	0	1185760	0	1276832	0.5	1405672	0	1199480	1.2	1893064	1.2	1175696
	WrenchinGranssLand	0	2006800	0.1	2103536	0.3	2229744	0	2023160	1.0	2683408	1.0	1872512
	Ave.	0	2026353	0.1	2148094	0.3	2276833	0.1	2054746	0.7	2785620	0.8	1899311
Medium-low	AxeinStoneGround	0	1169200	0.6	1233960	1.1	1313464	0.5	1171880	1.5	1643272	2.5	1186040
	CutterKnifeinPlasterBoard	0	467840	0.2	504000	0.2	577112	0.1	472968	2.2	850168	2.1	588944
	Dagger1inCarpet	0	1037376	0.3	1097736	0.5	1179048	0	1039328	1.6	1487584	1.8	994088
	Dagger2inBrickFloor	0	820392	0.2	889456	1.1	996728	-0.5	831864	1.8	1408536	2.2	834216
	Dagger3inSoil	0	1862688	0.2	1985760	0.3	2106408	0	1883512	1.7	2657448	1.9	1685160
	Dagger4inGrassLand	0	1499424	0	1570896	0.3	1669232	0	1511472	0.6	2045936	0.7	1439024
	HammerinGrassLand	0	764840	0	791776	0.5	852472	0	767520	2.5	1059176	2.5	824904
	KitchenKnifeinSand	0	1664392	0.1	1770152	0.9	1898688	0.3	1723000	2.1	2405112	2.0	1650400
	ScissorsinPlasterBoard	0	464600	0	499528	0.6	570424	0	469256	2.0	835624	2.1	578528
	WrenchinGranssLand	0	1188112	0.2	1242904	0.5	1324768	0.1	1198048	1.7	1629152	1.8	1182176
	Ave.	0	1093886	0.2	1158617	0.6	1248834	0.1	1106885	1.8	1602201	2.0	1096348
Low	AxeinStoneGround	0	596200	0.2	666720	0.5	688984	0	598448	0.9	906384	1.5	577824
	CutterKnifeinPlasterBoard	0	160872	0.1	183320	0.4	198376	0	163408	0.9	304256	1.2	193488
	Dagger1inCarpet	0	495136	0.2	551008	0.5	581072	0	495328	1.7	786248	1.5	479096
	Dagger2inBrickFloor	0	233232	0.2	271752	0.9	308448	0	236800	1.4	520512	1.3	223560
	Dagger3inSoil	0	887808	-0.2	1024624	-0.4	1040304	-0.7	896128	2.0	1410816	2.2	818000
	Dagger4inGrassLand	0	812256	0.5	890432	1.2	929408	0.1	819928	2.1	1189552	2.1	765536

	HammerinGrassLand	0	417288	0.5	451296	0.8	472800	0.2	420216	1.6	601952	1.5	451656
	KitchenKnifeinSand	0	719560	0	825112	0.1	869816	-0.2	748144	1.7	1232536	1.8	753808
	ScissorsinPlasterBoard	0	160864	0	183016	0.2	198584	0	165288	1.4	304080	1.6	193504
	WrenchinGranssLand	0	653352	0	714616	0	741192	0	661072	1.9	944552	1.8	663120
	Ave.	0	513656	0.2	576189	0.4	602898	-0.1	520476	1.6	820089	1.7	511959
Ave.		0	1754497	0.1	1857296 (5.9%)	0.4	1956743 (11.5%)	0	1775470 (1.2%)	1.1	2404565 (37.1%)	1.2	1670738 (-4.8%)

(a) original image, 24883200 (b) CSIT-QP, 2205220 (c) [22-27, 29, 30], 2364728 (d) [34], 2236616



(e) [35], 2401424 (f) [22-27, 29, 30], 4337984 (g) [34], 4561000 (h) [35], 4043728

Fig. 9 Reconstructed images of different QP selection techniques for Dagger3inSoil image

quality of image due to the significantly important or even the only information contained by the crime tool region for solving criminal cases and court trials. Compared with the QP selection technique in [22, 23, 25-27, 29], the QP selection technique in [30] and CSIT-QP significantly improves the perceptual quality of CSI tool images. However, the rate corresponding to [30] is the highest, while CSIT-QP has the lowest rate. In summary, the CSIT-QP achieves higher perceptual R-D performance compared with the other methods.

Compared with the high to medium-high rate levels, the perceptual quality improvement of using the CSIT-QP is significant at medium-low to low rate levels. This phenomenon can be explained from the following two aspects. First, at medium-low to low rate levels, the perceptual quality of the benchmark images is poor, thus CSIT-QP has larger space for the perceptual quality improvement. The "space" can be understood intuitively as the perceptual quality difference between the original images and the benchmark images. Second, due to the Weber's law in the human visual perception [47], when the perceptual quality of benchmark images is worse, it needs a smaller increase in perceptual quality difference between the target images and benchmark images, so that the human eyes can perceive the perceptual quality increment in the target images. The small increment in perceptual quality is easy to implement at the medium-low to low rate levels due to the first aspect mentioned above. It should be noted that most of perception motivated image/video coding algorithms all sharing this phenomenon, i.e., the algorithm is better at the medium-low to low rate levels than that at the high to medium-high rate levels for the perceptual quality improvement [47, 52].

Taking the Dagger3inSoil as and example to further illustrate the practical effect of the proposed algorithm. Fig. 9 (a) shows the original image with the original

rate is 24883200 bits. The original image was coded using the QP selection techniques in [22-27, 29, 30, 34, 35] and CSIT-QP, respectively. And the corresponding reconstructed images with different coding rates are obtained as shown in Figs. 9 (b)-(h). Compared with Figs. 9 (a) and (b), it is easy to see that the CSIT-QP can obtain a high perceptual quality reconstructed image with a significant rate compression (more than 11x). And, all key information of the crime tool, such as the texts and icons, can be easily perceived by the human eye. Compared with the reconstructed image in Fig. 9 (b), the reconstructed images with approximate rates shown in Figs. 9 (c)-(e) have significantly lower perceptual quality. Such as the texts are difficult to be easily recognized in Figs. 9 (c)-(e) by the human eye. As shown in Figs. 9 (f)-(h), in order to obtain the reconstructed images with the approximate perceptual quality as shown in Fig. 9 (b) using the QP selection techniques in [22-27, 29, 30, 34, 35], the rates must be significantly increased.

In conclusion, compared with the selection techniques in [22-27, 29, 30, 34, 35], the proposed CSIT-QP has higher perceptual R-D performance, and is helpful to improve the processing efficiency of CSI tool images in police practice and court trials.

5 Conclusion

This paper presents a saliency motivated perception QP selection technique, i.e., CSIT-QP, for CSI tool images using a two-step Lagrangian R-D optimization. Experimental results have demonstrated that the proposed CSIT-QP is effective in terms of R-D performance. For the other categories of CSI images, such as fingerprint images, shoeprint images, etc, the accurate salient regions acquisition method and the matched QP selection technique using the two-step Lagrangian R-D optimization are valuable researches in the future.

Acknowledgements This work was supported by the National Natural Science Foundation of China (62277036, 61801381, 61801282).

References

1. Abraham J, Ng R, Morelato M, Tahtouh M, Roux C. Automatically classifying crime scene images using machine learning methodologies. Forensic Science International: Digital Investigation, 2021, 39:1-17
2. Liu Y, Hu D, Fan J, Wang F, Li D. Multi-feature fusion based retrieval results optimization for crime scene investigation image retrieval. Acta Electronica Sinica, 2019, 47(2): 296-301
3. Tian T, Wang H. Large-scale video compression: recent advances and challenges. Frontiers of Computer Science, 2018, 12(5): 825-839
4. Madhavan A, Philip S, Ann M, Alsuwaidi M. Camphoric based Nano carbon for the visual enhancement of latent fingerprints. In: Proceedings of Advances in Science and Engineering Technology International Conferences. 2020, 1-4
5. Vagac M, Povinský M, Melichercík M. Obtaining tire tread model from its real world photo. In: Proceedings of IEEE International Scientific Conference on Informatics. 2019, 167-170
6. Li R. Fingerprint-related chaotic image encryption scheme based on blockchain framework. Multimedia Tools and Applications, 2021, 80: 30583-30603
7. Sarkar A, Singh B. A cancelable fingerprint biometric based session key establishment protocol. Multimedia Tools and Applications, 2019, 78: 21645–21671
8. Park K, Oh S, Heo S, Shin S, Bien F. 17-aF_{rms} resolution noise-immune fingerprint scanning analog front-end for under-glass mutual-capacitive fingerprint sensors. IEEE Transactions on Circuits and Systems-I: Regular Papers, 2022, 69(3): 1135-1147
9. Ma H, Liu Z, Heo S, Lee J, Na K, Jin H, Jung S, Park K, Kim J, Bien F. On-display transparent half-diamond pattern capacitive fingerprint sensor compatible with AMOLED display. IEEE Sensors Journal, 2016, 22(16): 8124–8131
10. Ali S, Khan M, Aslam A. Fingerprint matching, spoof and liveness detection: classification and literature review. Frontiers of Computer Science, 2021, 15(1): 157-166
11. Kushwaha R, Nain N. PUG-FB: person-verification using geometric and Haralick features of footprint biometric. Multimedia Tools and Applications, 2020, 79: 2671–2701
12. Kazmi W, Nabney I, Vogiatzis G, Rose P, Codd A. An efficient industrial system for vehicle tyre (tire) detection and text recognition using deep learning. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(2): 1264–1275
13. Liu Y, Ge Y, Wang F, Liu Q, Lei Y. A rotation invariant HOG descriptor for tire pattern image classification. In:

Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton. 2019, 2412-2416

14. Gao W, Jiang Q, Wang R, Ma S, Li G, Kwong S. Consistent quality oriented rate control in HEVC via balancing intra and inter frame coding. IEEE Transactions on Industrial Informatics, 2022, 18(3): 1594-1604
15. Gogoi S, Peesapati R. Design and implementation of an efficient multi-pattern motion estimation search algorithm for HEVC/H.265. IEEE Transactions on Consumer Electronics, 2021, 67(4): 319-328
16. Suehring K, Sharman K. HM sofrware. https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/
17. Zhou G, Tian X, Zhou A. Image copy-move forgery passive detection based on improved PCNN and self-selected sub-images. Frontiers of Computer Science, 2022, 16(4): 1-16
18. Chen T, Liu H, Ma Z, Shen Q, Cao X, Wang Y. End-to-end learnt image compression via non-local attention optimization and improved context modeling. IEEE Transactions on Image Processing, 2021, 30:3179-3191
19. Nguyen T, Marpe D. Objective performance evaluation of the HEVC main still picture profile. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(5): 790-797
20. Lainema J, Hannuksela M, Vadakital V, Aksu E. HEVC still image coding and high efficiency image file format. In: Proceedings of IEEE International Conference on Image Processing. 2016, 71-75
21. Helmrich C, Bosse S, Keydel P, Schwarz H, Marpe D, Wiegand T. A spectrally adaptive noise filling tool for perceptual transform coding of still images. In: Proceedings of IEEE International Conference on Consumer Electronics. 2018, 1-6.
22. McCann K, Rosewarne C, Bross B, Naccari M, Sharman K, Sullivan G. High efficiency video coding (HEVC) test model 16 (HM16) improved encoder description. Jonit Collaborative Team on Video Coding, JCTVC-S1002, 2014.
23. Gong Y, Wan S, Yang K, Yang Y, Li B. Rate-distortion-optimization-based quantization parameter cascading technique for random-access configuration in H.265/HEVC. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(6): 1304–1312
24. Zhao T, Wang Z, Chen C. Adaptive quantization parameter cascading in HEVC hierarchical coding. IEEE Transactions on Image Processing, 2016, 25(7): 2997-3009
25. Xu Y, Li Q, Li X, Liu J, Zhao T. Efficient QP cascading in H.265/HEVC low-delay prediction. In: Proceedings of IEEE International Conference on Multimedia & Expo Workshops. 2017, 151-156
26. Xu Y, Yi S, Lin L, Chen W, Zhao T. GOP structure-independent quantization parameter cascading in video coding. IEEE Access, 2019, 7: 76274-76282
27. He J, Yang E, Yang F, Yang K. Adaptive quantization parameter selection for H.265/HEVC by employing inter-frame dependency. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(12): 3424–3436
28. Amer H, Yang E. Low-delay HEVC adaptive quantization parameter selection through temporal propagation length estimation. In: Proceedings of IEEE International Conference on Image Processing. 2018, 211-215
29. Amer H, Yang E. Adaptive quantization parameter selection for low-delay HEVC via temporal propagation length estimation. Signal Processing: Image Communication, 2020, 84: 115826:1-115826:12
30. Gong Y, Yang K, Liu Y, Lim K P, Ling N, Wu H R. Quantization parameter cascading for surveillance video coding considering all inter reference frames. IEEE Transactions on Image Processing, 2021, 30: 5692-5707
31. Huang X, Li D, Yin H. HEVC quantization parameter selection algorithm based on inter-frame dependency. Multimedia Tools and Applications, 2020, 79: 13951-13966
32. Li D, Yin H, Huang X, Li H. Adaptive quantization parameter selection leveraging the inter-frame distortion propagation for HEVC video coding. In: Proceedings of Data Compression Conference. 2019, 588
33. Yin H, Huang X, Li D, Lu Y, Zhou Y. Distortion propagation modeling and its applications on frame level quantization control for predictive video coding. Signal Processing: Image Communication, 2019, 78: 398–408
34. Li B, Zhang D, Li H, Xu J. QP determination by lambda value. Jonit Collaborative Team on Video Coding. JCTVC-I0426, 2012
35. Sato K. Proposal on large block structure and quantization. Jonit Collaborative Team on Video Coding. JCTVC-C167, 2010
36. Zhao T, Wang Z, Kwong S. Flexible mode selection and complexity allocation in high efficiency video coding. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(6): 1135–1144
37. Khan R, Meyer A, Konik H, Bouakaz S. Saliency-based

framework for facial expression recognition. Frontiers of Computer Science, 2019, 13(5): 183-198

38. Wang X, Zhu L, Tang S, Fu H, Li P, Wu F, Yang Y, Zhuang Y. Boosting RGB-D saliency detection by leveraging unlabeled RGB images. IEEE Transactions on Image Processing, 2022, 31: 1107-1119
39. Liu Y, Li Z, Gong Y, Lim K, Wang F. Characteristics description and recognition of knife images in crime scene investigation. Journal of Xi'an University of Posts and Telecommunications, 2020, 25(1): 49-55
40. Sullivan G, Wiegand T. Rate-distortion optimization for video compression. IEEE Signal Processing Magazine, 1998, 15(6): 74-90
41. Im S, Ghandi M. Improved rate-distortion optimized video coding using non-integer bit estimation and multiple Lambda search. Frontiers of Computer Science, 2016, 10(1): 157-166
42. Yuan H, Wang Q, Liu Q, Huo J, Li P. Hybrid distortion-based rate-distortion optimization and rate control for H.265/HEVC. IEEE Transactions on Consumer Electronics, 2021, 67(2): 97-106
43. Dollár P, Zitnick C. Fast edge detection using structured forests. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(8): 1558-1570
44. Gong Y, Wan S, Yang K, Wu H R, Li B. A visual-masking-based estimation algorithm for temporal pumping artifact region prediction. Circuits, Systems, and Signal Processing, 2017, 36: 1264-1287
45. Gong Y, Wan S, Yang K, Wu H R, Liu Y. Temporal-layer-motivated lambda domain picture level rate control for random-access configuration in H.265/HEVC. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29(1): 156-170
46. Ben-Israel A. A Newton-Raphson method for the solution of systems of equations. Journal of Mathematical Analysis and Applications, 1966, 15(2): 243-252
47. Wu H R, Reibman A, Lin W, Pereira F, Hemami S. Perceptual visual signal compression and transmission. Proceedings of the IEEE, 2013, 101(9): 2025-2043
48. Lin W, Kuo C -C J. Perceptual visual quality metrics: a survey. Journal of Visual Communication and Image Representation, 2011, 22: 297-312
49. ITU-RBT. 500-13. Methodology for the subjective assessment of the quality of television pictures. ITU-R Recommendations, 2012
50. Liu A, Lin W, Paul M, Deng C, Zhang F. Just noticeable difference for images with decomposition model for separating edge and textured regions. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(11): 1648-1652
51. Yang K, Wan S, Gong Y, Wu H R, Feng Y. Perceptual based SAO rate-distortion optimization method with a simplified JND model for H.265/HEVC. Signal Processing: Image Communication, 2015, 31: 10-24
52. Gong Y, Wan S, Yang K, Wu H R, Yang F, Li B. Reduction of temporal distortion in video coding based on detection of just-noticeable temporal pumping artifact. Signal Processing: Image Communication, 2016, 42: 1-18