ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/319569738

A new combined PSNR for objective video quality assessment

Conference Paper · July 2017

DOI: 10.1109/ICME.2017.8019494

CITATIONS
7

READS
481

6 authors, including:

Xiwu Shang
Shanghai University of Engineering Science
23 PUBLICATIONS **116** CITATIONS

SEE PROFILE

All content following this page was uploaded by Xiwu Shang on 27 April 2018.

The user has requested enhancement of the downloaded file.

A NEW COMBINED PSNR FOR OBJECTIVE VIDEO QUALITY ASSESSMENT

Xiwu Shang[1], Guozhong Wang[1], Haiwu Zhao[1], Jie Liang[2], Senior Member, Chengjia Wu[1], Chang Lin[1]

[1]School of Communication and Information Engineering, Shanghai University, China
[2]School of Engineering Science, Simon Fraser University, Canada
dxsxw@126.com, wanggz@shu.edu.cn, zhaohaiwu@i.shu.edu.cn, jiel@sfu.ca,
{1565149320, 450213487 }@qq.com

ABSTRACT

In video coding, quality evaluation is important for improving the coding efficiency. Usually Peak Signal-to-Noise Ratio (PSNR) is utilized to measure the performance of different coding techniques. During the video coding process in YCbCr color space, there are three PSNRs, one for each color component. Sometimes they may contradict to each other, which poses a problem for evaluating the coding performance. Several video quality assessment (VQA) metrics have been proposed to measure the video quality with a combined PSNR. However, these combined PSNRs are obtained heuristically without theoretical justification. In this paper, we propose a color-sensitivity-based combined PSNR (CSPSNR) based on extensive subjective tests on the sensitivity of human visual system (HVS) to different color components. Subjective experiment results demonstrate that the proposed combined PSNR correlates well with the mean opinion score (MOS) than existing methods.

Index Terms— video quality assessment, PSNR, color sensitivity, HVS

1. INTRODUCTION

As the explosive growth of video contents in recent years, they need to be compressed and delivered to users. Therefore an accurate assessment of video quality plays a critical role in improving the quality of experience (QoE). Since videos are watched by human, the most reliable quality evaluation is to obtain the judgements from many human observers. However, such a subjective evaluation is time consuming. It is highly desired that an algorithm can automatically provide a video quality measurement which is consistent with human perception.

The peak signal-to-noise ratio (PSNR) is the most commonly used metrics for measuring the quality of reconstructed images in various coding standards, such as H.263 [1], H.264 [2] and high efficiency video coding (HEVC) [3]. It represents the ratio between the maximum power of a signal and the power of distorted signal, although this definition does not always agree with human perception [4, 5].

Various video quality assessment (VQA) metrics have been developed to overcome the drawback of PSNR. The most representative one is the structure similarity (SSIM) [6], which evaluates the perceptual video quality from three aspects: the luminance, contrast and structure, other than the pixel difference between original and reconstructed signals. Recently, more effective metrics have been proposed, such as motion-based video integrity evaluation (MOIVE) [7], feature similarity [8], visual saliency-based index [9] and convolution neural network (CNN)-based approaches [10, 11, 12]. Although these studies demonstrate that PSNR is not always consistent with subjective quality, Huynh et al. in [13] validate that PSNR is a good estimator of the video quality under the condition that the content and codec are fixed across the coding environment during the process of image/video coding. Therefore, the latest HEVC coding standard still uses the average difference in PSNR between two rate-distortion (RD) curves to evaluate the coding performance of different coding tools [14].

Usually videos are represented by YCbCr color space, which results in three PSNRs. Sometimes, they may contradict to each other. For example, coding gain is obtained in Y component but is lost in CbCr components [15]. Therefore, it is desired to combine the three PSNRs into one PSNR. Generally, there are two types of methods to compute the combined PSNR [16, 17, 18]. The weighting factors in the two methods are derived heuristically without solid theoretical basis.

In this paper, we propose a color-sensitivity-based combined PSNR (CSPSNR) to represent the quality of a YCbCr sequence. The basic idea of our approach is to achieve a perceptual weighted PSNR according to the sensitivities to different components. Specifically, by exploiting the relationship between RGB and YCbCr color space, we find the highest contrasts of YCbCr components, and convert them to RGB space for display, which will be used to perform the subjective test of the sensitivity of YCbCr. The weighting factors

This work was supported by National Science Foundation of China under the Grant No. 61271212, National High-tech R&D program of China (863 Program, 2015AA015903) and the China Scholarship Council.

based on the sensitivity are used to derive the proposed combined PSNR. Finally, the relationship between the CSPSNR and subjective scores is investigated to verify the effectiveness of the proposed CSPSNR.

The remainder of the paper is organized as follows. In Sec. 2, we describe the related works. In Sec. 3, we present the CSPSNR based on the perceptual sensitivities of YCbCr components. Sec. 4 demonstrates the experimental results. Sec. 5 concludes the paper and discusses the future work.

2. RELATED WORKS

One of the combined PSNRs is a weighted average of the three PSNRs [16, 17], which is calculated as

$$PSNR_{611} = (6 \cdot PSNR_Y + PSNR_{Cb} + PSNR_{Cr})/8, \quad (1)$$

where $PSNR_Y$, $PSNR_{Cb}$ and $PSNR_{Cr}$ are the PSNRs of three color components. Each of them is defined as

$$PSNR = 10 \cdot \log_{10}((2^{N_b} - 1)^2/MSE), \quad (2)$$

where N_b is the number of bits per sample in the original video signal, and MSE is the MSE of all samples in each channel. The ratio of weighting factors for the three components is 6:1:1, which is chosen empirically, and there is no theoretical justification of the ratio.

Another method adopted in the HEVC reference software HM derives the combined PSNR from the combined MSE MSE_{YCbCr} [18]. We denote this method as $PSNR_{HM}$, which is calculated as

$$PSNR_{HM} = 10 \cdot \log_{10}((2^{N_b} - 1)^2/MSE_{YCbCr}), \quad (3)$$

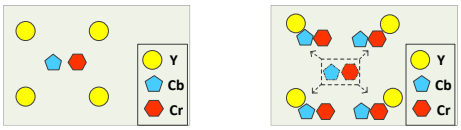
$$MSE_{YCbCr} = w_Y \cdot MSE_Y + w_{Cb} \cdot MSE_{Cb} + w_{Cr} \cdot MSE_{Cr}, \quad (4)$$

$$w_Y = \frac{N_Y}{N_Y + N_{Cb} + N_{Cr}}, \quad (5)$$

$$w_{Cb} = \frac{N_{Cb}}{N_Y + N_{Cb} + N_{Cr}}, \quad (6)$$

$$w_{Cr} = \frac{N_{Cr}}{N_Y + N_{Cb} + N_{Cr}}, \quad (7)$$

where N_Y, N_{Cb} and N_{Cr} are the number of samples of Y, Cb and Cr, respectively. It is based on the assumption that the average MSE is related to the sampling format. For example, for 4:2:0 videos, $w_Y = 2/3$ and $w_{Cb} = w_{Cr} = 1/6$. For 4:4:4 videos, $w_Y = w_{Cb} = w_{Cr} = 1/3$, which means that the importance of the three components is equal. However, this does not consider the fact that human eye is more sensitive to luma component than chroma components. Therefore, when $PSNR_{HM}$ is utilized to evaluate the coding performance of 4:4:4 videos, it may not correlate well with the subjective assessments. In fact, even if we ignore the color sensitivity problem, the weighting factors derived from the sampling format are questionable. Fig. 1 shows an example of 4:2:0 chroma format. It can be seen that four Y components share one Cb and one Cr component. To recover the values of each pixel, each Cb and Cr sample is duplicated four times. Thus, the number of samples for display is the same for the three components, which suggests that the weighting factors should be irrelevant to the sampling format.



(a) (b)

Fig. 1. Example of YCbCr recovery for each pixel from 4:2:0 chroma sampling, (a) 4:2:0 chroma sampling. (b) Recovery of YCbCr value for each pixel.

3. THE PROPOSED COLOR-SENSITIVITY-BASED COMBINED PSNR

To measure the influence of the distortion of one component of the YCbCr color space to the overall distortion, we need to fix the values of the other two components. This fixed value is chosen as 128 in our experiment. One reason is that 128 is the median of 8-bit pixel values. Another reason is that human visual system (HVS) has the highest visual acuity to distinguish the changes of color signal under mid-gray background [19]. This is why in typical subjective assessments of the quality, the mid-gray video is always inserted between the two test sequences.

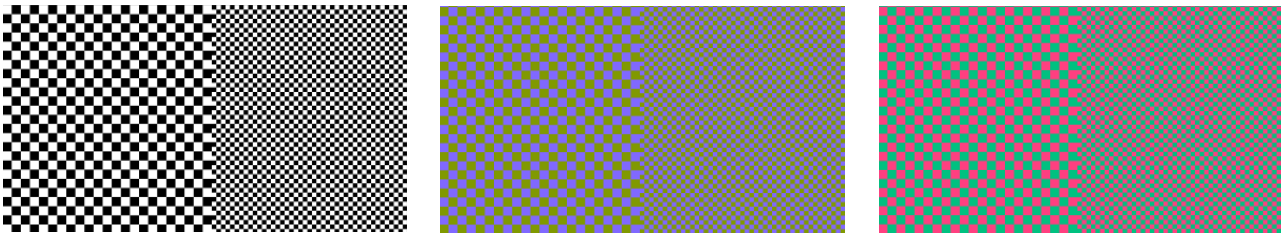
After fixing the values of the other components, the range of the current component is a significant factor in measuring visual resolution. Generally, visual sensitivity measurement uses high contrast patterns [20]. For example, the visual sensitivity test conducted in medical clinics has black letters on white background. To find the range of each component, we need to use the mapping from RGB space to YCbCr space, which is given by the following, according to ITU Rec. 601 [21].

$$\begin{cases} Y = 0.257 \cdot R + 0.504 \cdot G + 0.098 \cdot B + 16 \\ Cb = -0.148 \cdot R - 0.291 \cdot G + 0.439 \cdot B + 128, \\ Cr = 0.439 \cdot R - 0.368 \cdot G - 0.071 \cdot B + 128 \end{cases} \quad (8)$$

and then RGB can be recovered by:

$$\begin{cases} R = 1.164 \cdot (Y - 16) + 1.596 \cdot (Cr - 128) \\ G = 1.164 \cdot (Y - 16) - 0.813 \cdot (Cr - 128) \\ \qquad -0.392 \cdot (Cb - 128). \\ B = 1.164 \cdot (Y - 16) + 2.017 \cdot (Cb - 128) \end{cases} \quad (9)$$

Combining the condition $\{0 \le R \le 255, 0 \le G \le 255, 0 \le B \le 255, Cb = 128, Cr = 128\}$ with Eq. (9), it can be verified that the range of Y is 16 to 235. The two extreme cases are $\{Y = 16, Cb = 128, Cr = 128\}$ and



(a) (b) (c)

Fig. 2. Checkerboard for testing the sensitivity of YCbCr. (a) Y, (b) Cb, (c) Cr.

$\{Y = 235, Cb = 128, Cr = 128\}$. Their corresponding values in RGB space are $\{R = 0, G = 0, B = 0\}$ and $\{R = 255, G = 255, B = 255\}$, which have the highest contrast. For the Cb case, we utilized the condition $\{0 \leq R \leq 255, 0 \leq G \leq 255, 0 \leq B \leq 255, Y = 128, Cr = 128\}$ with Eq. (9) to calculate the range of Cb. For the Cr case, we utilized the condition $\{0 \leq R \leq 255, 0 \leq G \leq 255, 0 \leq B \leq 255, Y = 128, Cb = 128\}$ with Eq. (9) to calculate the range of Cr. These results are listed in Table 1.

Table 1. Transformation between YCbCr and RGB

Y	Cb	Cr	R	G	B
16	128	128	0	0	0
235	128	128	255	255	255
128	64	128	130	155	1
128	190	128	130	106	255
128	128	47	1	196	130
128	128	206	255	67	130

There are three methods to measure the visual sensitivity [20]: double dot target, acuity grating and checkerboard. In this paper, we choose the checkerboard method to measure the sensitivity, as illustrated in Fig. 2, where blocks in the left half of the checkerboard have 10x10 pixels, while blocks in the right half checkerboard have 5x5 pixels. The two sets of RGB values that form the checkerboard are the two extreme cases as listed in Table 1, in order to achieve the highest contrast in each channel. The right half of the checkerboard is used as the reference area.

To conduct the subjective tests, each figure in Fig. 2 is displayed on a Sony PVM-2541 monitor, which has the peak brightness of 120 cd/m^2 and color temperature of D65. Twenty human subjects are involved in the experiment. Before the experiment, they have been screened for normal or corrected normal visual acuity on Snellen chart and for normal color vision using color blindness book. To reduce the measurement errors, we fix the location of observers and just move the monitor. The distance between the observer and the monitor is recorded when the observer can just distinguish the two neighboring blocks in the left side of the checkerboard. From the experiment, we obtain the average distance ratio of 1:0.45:0.51 for Y, Cb and Cr. The distance under the test of Y is longer than that of Cb and Cr, which indicates that resolution threshold of Y is higher. Let T_i ($i = Y$, Cb or Cr) denotes the resolution threshold which is the ratio between the critical distance and the side length of the 10×10-pixel block. Then the ratio of resolution threshold for YCbCr is also 1:0.45:0.51. When the distance between the observers and the test figure is fixed, the side length of the block just distinguished by observers is inversely proportional to the resolution threshold. For simplicity, Y, Cb and Cr components are regarded as a perceptual sub-unit of a pixel with the side length ratio of $(1/T_Y) : (1/T_{Cb}) : (1/T_{Cr})$, respectively.

The area of a sub-unit A_i is then calculated as:

$$A_i = L_i \cdot L_i, \tag{10}$$

where L_i is side length of a sub-unit.

And the number of sub-units N_i in a unit area is:

$$N_i = \frac{1}{A_i}. \tag{11}$$

The percentage of each sub-unit P_i in a unit is given by

$$P_i = \frac{N_i}{N_Y + N_{Cb} + N_{Cr}}. \tag{12}$$

Combined with T_i, we obtain that $P_Y = 0.685$, $P_{Cb} = 0.137$ and $P_{Cr} = 0.178$. The proposed weighted sum of MSE MSE_{P_YCbCr} is thus given by

$$MSE_{P\_YCbCr} = P_Y \cdot MSE_Y + P_{Cb} \cdot MSE_{Cb} + P_{Cr} \cdot MSE_{Cr}. \tag{13}$$

Substituting Eq. 13 to Eq. 2, we can get the proposed $CSPSNR$ as:

$$CSPSNR = 10 \cdot \log_{10}((2^{N_b} - 1)^2 / MSE_{P\_YCbCr}). \tag{14}$$

Eq. 2 can be written as:

$$\frac{1}{10^{PSNR/10}} = \frac{MSE}{(2^{N_b} - 1)^2}. \tag{15}$$

Substituting Eq. 15 into Eq. 14, it can be rewritten to

$$CSPSNR = -10 \cdot \log_{10}\left(\frac{P_Y}{10^{\frac{PSNR_Y}{10}}} + \frac{P_{Cb}}{10^{\frac{PSNR_{Cb}}{10}}} + \frac{P_{Cr}}{10^{\frac{PSNR_{Cr}}{10}}}\right). \tag{16}$$

It can be seen from Eq. 16 that the weighting factors of the proposed CSPSNR are derived from the perceptual sampling rate of YCbCr component, and do not vary with the sampling format.

Table 2. Example of PSNR combinations of different components under test sequences with different sampling formats

	$PSNR_Y$	$PSNR_{Cb}$	$PSNR_{Cr}$
Runners420	42.29	42.26	42.07
	38.30	38.01	38.14
	34.12	33.91	34.24
	29.92	30.05	29.61
Runners422	42.29	42.29	42.22
	38.30	37.96	38.15
	34.12	34.06	33.95
	29.92	29.81	29.74
Runners444	42.29	42.16	42.13
	38.30	38.22	38.05
	34.12	33.81	33.99
	29.92	29.86	29.96

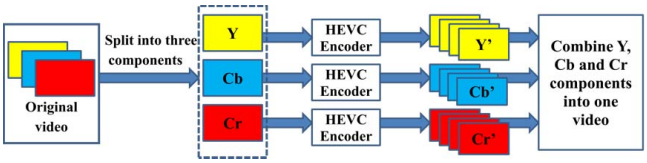


Fig. 3. Structure of the scheme for producing videos with combination of different levels of PSNR of YCbCr.

4. EXPERIMENTAL RESULTS

To verify that the proposed combined PSNR is more consistent with the subjective score, test sequences with different levels of combined PSNRs are required to evaluate the quality of videos. From Eq. 16, we can see that the combined PSNR is related to the PSNR of three components. Thus we need test sequences with different levels of PSNR of YCbCr. However, in the process of video coding, the values of PSNR for Cb and Cr components are always higher than Y component. The reason is that the texture complexity in chroma components is lower, which can achieve higher prediction accuracy and lower prediction error compared to luma component. To solve this problem, we design a new coding scheme which is shown in Fig. 3. Firstly, the original sequence with YCbCr components is split into three sequences, which are Y, Cb and Cr sequences. Then, the three sequences are encoded by HEVC encoder independently. After obtaining the different levels of PSNR for each component, we combine them and produce 4x4x4 distorted sequences. Although we have theoretically analyzed that the combined PSNR is unrelated to sample rate in Sec. 2, test sequences with different sampling rates are required to verify the analysis.

In our experiment, we test three source sequences corresponding to three types of sampling formats [22]. Thus there are totally 192 (4x4x4x3) distorted videos in our test database. Table 2 shows different combinations of the PSNRs of different channels. It can be seen that the range of PSNR values is

Fig. 4. DSIS presentation structure.

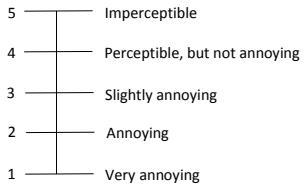


Fig. 5. MOS scales for assessing the video quality.

from 30 to 42 dB. The reason that the lowest value of PSNR is about 30 dB is that the PSNR value derived from coding chrominance component is almost always greater than 30 dB even for sequences with rich colors and large quantization parameters (QPs).

After obtaining 64 distorted videos for each source sequence, subjective tests are conducted, where the participants are asked to score the videos. The double-stimulus impairment scale (DSIS) [23] protocol is conducted to evaluate the subjective quality. The source video and distorted video are displayed in sequence according to DSIS structure, as shown in Fig. 4. Two-second-long mid-gray video is inserted between the two videos. Observers are asked to vote during the voting time after playing the two videos. The MOS scores used for voting range from very annoying to imperceptible, which are shown in Fig. 5. Twenty nonexpert subjects with normal or correct-normal visual acuity are involved in the experiment.

Four commonly used performance metrics, Spearman rank-order correlation coefficient (SROCC), Kendall rank-order correlation coefficient (KROCC), Pearson linear correlation coefficient (PLCC) and the Root Mean Square Error (RMSE) suggested by VQEG [24], are exploited to evaluate the performance of different combined PSNR metrics. SROCC and KROCC are utilized to represent the prediction monotonicity of different methods. PLCC and RMSE are used to evaluate the prediction accuracy and prediction consistency. Values close to 1 for KROCC, SROCC and PLCC, while close to 0 for RMSE, indicate better correlation with the subjective assessment.

Table 3 tabulates SROCC, KROCC, PLCC, and RMSE results of different VQA metrics. Data points highlighted in boldface indicate the best results. In Table 3, we can see that the $PSNR_{HM}$ correlates well with the subjective evaluations for 420 format, but poorly for 444 format. From Fig. 6, we can see that the curve of $PSNR_{HM}$ becomes lower from 420 to 444 format for KROCC, SROCC and PLCC, and gets higher from 420 to 444 format for RMSE. The SROCC can be up to 0.9857 for Runners420, but only 0.8104 for Runners444. The reason is that the weighting factors are 2/3, 1/6 and 1/6 for 4:2:0 format, which are near the optimal weight-

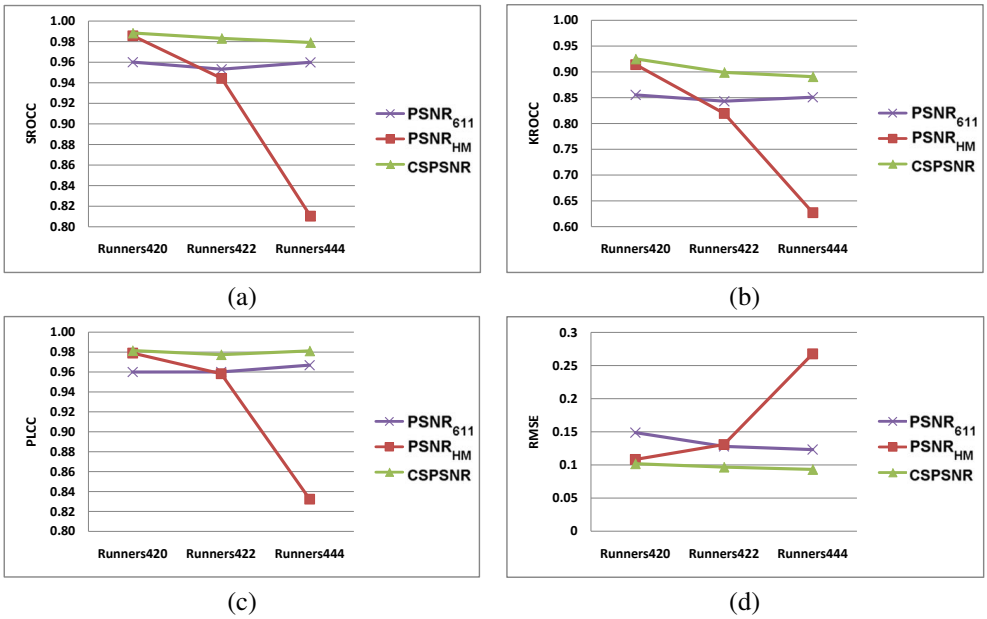


Fig. 6. Diagram comparison of different methods: (a) SROCC, (b) KROCC, (c) PLCC, (d) RMSE.

ing factors derived in this paper. However, for 4:4:4 format, the weighting factors of the three components are all equal to 1/3, which is clearly far from optimal. For the $PSNR_{611}$, its diagram is higher than $PSNR_{HM}$ under 444 format, while is clearly lower under 420 format in terms of correlation coefficients. Compared with the proposed CSPNSR, its correlations with subjective evaluations are lower in most cases. Because its weighting parameters are determined heuristically, whereas the parameters in this paper are derived based on rigorous subjective tests. In Table 3, most of the data points of the CSPSNR are highlighted with boldface indicating the optimal results. Meanwhile, Fig. 6 demonstrates the superior performance of CSPSNR with almost the highest curves of correlation coefficients and the lowest curve of RMSE.

5. CONCLUSION

In this paper, we propose a color-sensitivity-based combined PSNR (CSPSNR) to evaluate the video quality. To the best of our knowledge, this is the first combined PSNR based on subjective experiments and theoretical analysis. To measure the sensitivities of different components, the highest contrast of each component used to perform the subjective test is achieved by exploiting the conversion relationship between YCbCr and RGB color space. Then we deduce theoretically the perceptual sampling rates of YCbCr components in one pixel based on the sensitivities, which is incorporated to characterize the combined distortion of three components. Finally, to verify the effectiveness of CSPSNR, we design a new coding scheme to produce distorted sequences with combinations of various levels of PSNR of YCbCr. Experimental results demonstrate that the CSPSNR has a high correlation with subjective scores, and is superior to the state-of-the-art combined PSNR approaches.

Our future work will modify the rate-distortion (RD) model to improve the perceived quality. In the traditional RD mode, the sensitivities of the distortion of YCbCr components are regarded equally instead of assigning different weighting factors based on the characteristics of HVS. By considering the perceptual sensitivity of different components and lagrange multiplier, a new RD model that conforms well with HVS might be established.

Table 3. Performance of different combined PSNR (The best performed metrics appear in bold)

Sequence		$PSNR_{611}$	$PSNR_{HM}$	$CSPSNR$
Runners420	SROCC	0. 9601	0. 9857	**0. 9884**
	KROCC	0. 8554	0. 9141	**0. 9252**
	PLCC	0. 9600	0. 9790	**0. 9815**
	RMSE	0. 1488	0. 1083	**0. 1019**
Runners422	SROCC	0. 9532	0. 9443	**0. 9832**
	KROCC	0. 8432	0. 8195	**0. 8991**
	PLCC	0. 9601	0. 9583	**0. 9774**
	RMSE	0. 1280	0. 1308	**0. 0968**
Runners444	SROCC	0. 9599	0. 8104	**0. 9792**
	KROCC	0. 8510	0. 6272	**0. 8908**
	PLCC	0. 9669	0. 8324	**0. 9812**
	RMSE	0. 1232	0. 2677	**0. 0933**

6. REFERENCES

[1] T. Wiegand, M. Lightstone, D. Mukherjee, T. G. Campbell, and S. K. Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging h.263 standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 2, pp. 182–190, Apr 1996.

[2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the h.264/avc video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, July 2003.

[3] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[4] F. Zhang, W. Lin, Z. Chen, and K. N. Ngan, "Additive log-logistic model for networked video quality assessment," *IEEE Trans. Image Processing*, vol. 22, no. 4, pp. 1536–1547, April 2013.

[5] L. Xu, W. Lin, L. Ma, Y. Zhang, Y. Fang, K. N. Ngan, S. Li, and Y. Yan, "Free-energy principle inspired video quality metric and its use in video coding," *IEEE Trans. Multimedia*, vol. 18, no. 4, pp. 590–602, April 2016.

[6] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[7] K. Seshadrinathan and A. C. Bovik, "Motion tuned spatio-temporal quality assessment of natural videos," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 335–350, Feb. 2010.

[8] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[9] L. Zhang, Y. Shen, and H. Li, "Vsi: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Oct. 2014.

[10] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, June 2014, pp. 1733–1740.

[11] X. Tian, Z. Dong, K. Yang, and T. Mei, "Query-dependent aesthetic model with deep learning for photo quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 2035–2048, Nov. 2015.

[12] H. Wang, L. Zuo, and J. Fu, "Distortion recognition for image quality assessment with convolutional neural network," in *2016 IEEE International Conference on Multimedia and Expo (ICME)*, July 2016, pp. 1–6.

[13] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of psnr in image/video quality assessment," *Electronics letters*, vol. 44, no. 13, pp. 800–801, 2008.

[14] G. Bjontegaard, "Calculation of average psnr differences between rd curves," Austin, Texas, 2001, ITU-T SC16/Q6, VCEG-M33.

[15] E. Manni J. Xu, A. Tourapis and K. Sato, "Chroma qp extension," in *ITU-T/ISO/IEC joint collaborative team on video coding (JCT-VC)*, 2012, JCTVC-H0400.

[16] J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards-2014;including high efficiency video coding (hevc)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, 2012.

[17] G. J. Sullivan and J. R. Ohm, "Meeting report of the fourth meeting of the joint collaborative team on video coding," in *ITU-T/ISO/IEC joint collaborative team on video coding (JCT-VC)*, 2011, JCTVC-D500.

[18] G. J. Sullivan and J. R. Ohm, "Objective quality metric and alternative methods for measuring coding efficiency," in *ITU-T/ISO/IEC joint collaborative team on video coding (JCT-VC)*, 2012, JCTVC-H0012.

[19] S. H. Bae and M. Kim, "A dct-based total jnd profile for spatio-temporal and foveated masking effects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. PP, no. 99, pp. 1–1, 2016.

[20] H. Kolb, E. Fernandez, and R. Nelson, "Visual acuity–webvision: the organization of the retina and visual system," 1995.

[21] "Recommendation of itu-r bt.601, http://www.itu.int/rec/r-rec-bt.601," .

[22] L. Song, X. Tang, W. Zhang, X. Yang, and P. Xia, "The sjtu 4k video sequence dataset," in *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*, 2013, pp. 34–35.

[23] "Methodology for subjective assessment of the quality of television pictures," in *ITU-R BT. 500-11*, 2012.

[24] "Final report from the video quality experts group on the validation of objective models of video quality assessment," in *VQEG*, 2012.

View publication stats