A QUALITY PREDICTION MODEL FOR JPEG2000-BASED COLOR IMAGES

Ling Li[1,2], *Zhensong Wang*[1]

[1]Institute of Computing Technology, Beijing, China
[2]Graduate University of Chinese Academy of Sciences, Beijing, China
{liling, zswang@ict.ac.cn}

ABSTRACT

A quality prediction model is proposed for color images coded with JPEG2000. This model estimates the quality (PSNR) of a color image at a given compression ratio without coding. The image activity measure as the image feature and compression ratio are taken as the input of the model. Experimental results show that the prediction error is less than 1dB for more than 75% images, and less than 2dB for over 95% images. The computation of the prediction process is much lower than that of the compression and quality calculation.

***Index Terms*—** Quality prediction, JPEG2000, image activity measure

1. INTRODUCTION

Predicting image quality at the pre-encoding stage is very necessary for rate allocation[1, 2], filter selection[3], encoding parameter decision[4, 5], etc. So far as we know, image quality is obtained after encoding and decoding for most coding algorithms, such as JPEG, AVC, etc. While with EBCOT in JPEG2000, the image distortion can be estimated during encoding. However, the image needs to be transformed and coded in tier-1 stage.

Methods and models of quality prediction were proposed in recent years. [4, 5] target MPEG-based encoded video sequences, classify video clips into several categories, and establish quality vs. bit-rate (BR) curve for each category. The quality vs. BR curve of a specific video is determined after the category is decided. However, accurate quality may not be obtained with coarse classification. Variance is used to predict quality for regions at a given bit-rate for DCT-based algorithms [2]. [3] established the relationship of the spatial and frequency indexes of filter banks with the quality (PSNR) experimentally for a fixed bit-rate (BR). Saha[6, 7] found that for grey images only the gradient-based image activity measure (IAM) has a direct effect on PSNR values, among the tested features including variance, energy, entropy, IAM, etc. The logarithmic equation is fitted as $PSNR\big|_{CR} = g(IAM) = \alpha \ln(IAM) + \beta$ for a fixed compression ratio (CR). Other researchers, as [1, 8], have also disclosed the strong correlation between the gradient-based image activity measure and the coding performance.

Nowadays, color images and videos are widely used, and no quality prediction model has been found for JPEG2000. In this paper, we study the effect of image features and compression ratios on the coding quality for JPEG2000 system. A prediction model is proposed based on image activities. The quality of the image can be predicted without coding for a given compression ratio.

The paper is organized as follows. The quality prediction model is explained in Section 2. Estimation of the model parameters is discussed in Section 3. Section 4 demonstrates the prediction performance. Section 5 draws the conclusion.

2. QUALITY PREDICTION MODEL

2.1. Quality metrics

Various quality metrics have been proposed, such as PSNR, MOS, MSSIM. Among all the metrics, PSNR is widely used for its simplicity. And it can evaluate the compression quality for various images coding with the same codec. Thus, PSNR measure is used in this paper, and is defined as

$$PSNR = 10\log_{10}\frac{\left(2^{B}-1\right)^{2}}{MSE}. \tag{1}$$

Where, MSE is the mean square error between the original and reconstructed images. We have defined the MSE for color images as

$$MSE = \frac{1}{M \times N \times C}\sum_{c=1}^{C}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[\mathbf{x}_c(i,j) - \hat{\mathbf{x}}_c(i,j)]^2 .$$

Where, $\mathbf{x}_c(i,j)$ and $\hat{\mathbf{x}}_c(i,j)$ are the (i,j) pixel values of the component c in the original and reconstructed images, respectively. M and N are the height and width of the image.

2.2. Model

Color images consisting of three components are studied in this paper. In JPEG2000, color images are often transformed into YCbCr space, then transformed with wavelets, and coded with EBCOT. The Y component presents the luminance information, while the Cb and Cr components generally present chrominance information.

To predict the quality ($PSNR_{pred}$) of a color image $\mathbf{x}$, a set of features f_k $(k=1,...,m)$ is used. We select the IAM of the components in YCbCr space as the features, since IAM correlates well with PSNR for grey images [6, 7]. Besides, the image distortion has close connections with these components. Since the distortion is also affected by the compression ratios, the predicted quality is modeled as a function of the feature values and the CR

$$PSNR_{pred} = \Phi(f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x}), CR) . \quad (2)$$

$f_1(\mathbf{x}) = IAM_Y$, $f_2(\mathbf{x}) = IAM_{Cb}$ and $f_3(\mathbf{x}) = IAM_{Cr}$ are the IAM values of Y, Cb and Cr components, respectively.

Next, based on the logarithmic relationship between IAM and PSNR at a fixed compression ratio, we would give the detailed expression of (2).

For grey images, the relationship between IAM and PSNR at a fixed compression ratio is fitted as a logarithmic formula [6, 7]

$$PSNR\big|_{CR} = g(IAM) = \alpha \ln(IAM) + \beta . \quad (3)$$

Where, α and β are constants dependent on the required compression ratio and the wavelet filters. PSNR is the resulted quality. IAM is the image activity measure for the grey image, calculated as [6, 7]

$$\begin{aligned} IAM = & \frac{1}{(M-1)\times N}\sum_{i=0}^{M-2}\sum_{j=0}^{N-1}|\mathbf{x}(i,j)-\mathbf{x}(i+1,j)| \\ & +\frac{1}{M\times(N-1)}\sum_{i=0}^{M-1}\sum_{j=0}^{N-2}|\mathbf{x}(i,j)-\mathbf{x}(i,j+1)| \end{aligned} . \quad (4)$$

$\mathbf{x}(i,j)$ is the (i,j) pixel value of an grey image. Let IAM_Y , IAM_{Cb} and IAM_{Cr} be calculated with (4) replacing $\mathbf{x}$ with the corresponding components.

For color images coding with JPEG2000, the three components contribute different lengths of bit-stream to the final codestream, and Cb and Cr components generally have similar compression ratios. Thus, we set the final prediction quality equation as:

$$\begin{aligned} PSNR_{pred} &= \Phi(IAM_Y, IAM_{Cb}, IAM_{Cr}, CR) \\ &= g(IAM_{weight}, CR) = \alpha \ln(IAM_{weight}) + \beta \end{aligned} . \quad (5)$$

Where, α and β are functions of CR.

Let us define

$$IAM_{weight} = \theta IAM_Y + (IAM_{Cb} + IAM_{Cr})(1-\theta)/2 . \quad (6)$$

θ is a function of CR, and satisfies $\theta \in [0,1]$.

Inserting (6) in (5), yield that

$$\begin{aligned} PSNR_{pred} &= \Phi(IAM_Y, IAM_{Cb}, IAM_{Cr}, CR) \\ &= \alpha \ln\left(\theta IAM_Y + (IAM_{Cb} + IAM_{Cr})(1-\theta)/2\right) + \beta \end{aligned}$$

$$\theta \in [0,1] . \quad (7)$$

Since this model integrates image features and CR together, the image quality at various compression ratios can be predicted at a pre-encoding stage.

As seen in Section 3 below, θ is a monotone increasing function of CR. And $\alpha(CR)$ and $\beta(CR)$ are two piecewise functions.

3. PARAMETER ESTIMATION

In this section, we would introduce the estimation of the parameters in the quality prediction model (7).

To establish the model, we select 105 widely used color images, including people, animals, sea, plants, motors, etc. These images are compressed with the well-known JPEG2000 coder kakadu_v6.0[1], at the compression ratios ranging from 10:1 to 100:1. 5 levels of 9/7 and 5/3 DWT are applied. 64x64 codeblock size is used. The quality is calculated for the reconstructed images.

For each fixed compression ratio, we solve the constrained linear optimal problem (7) to get parameter θ with Matlab aiming at maximizing the R^2 values. R^2 is the square of the correlation between the response values and the predicted response values. R-square with a value closer to 1 indicates a better fit. The resulted R^2 values of IAM_{weight}-PSNR fits for different CR are higher than 0.92.

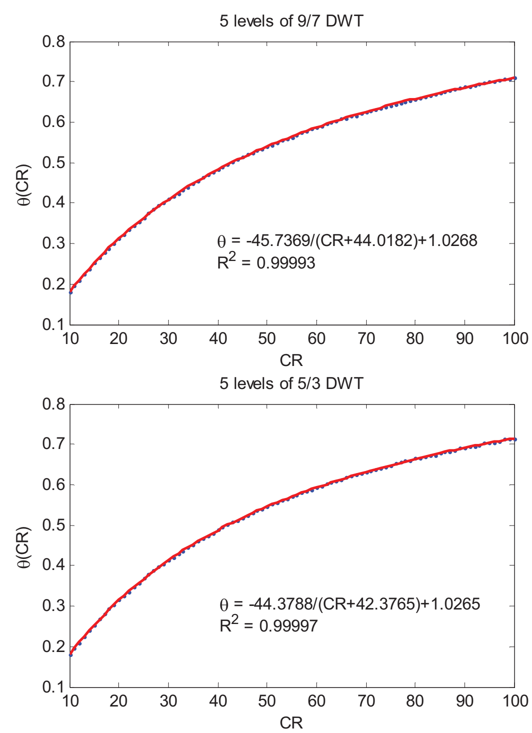


Fig. 1 $\theta(CR)$ **for 5 levels of 9/7and 5/3 DWT.**

[1] The following parameters have been used in Kakadu: "-full -no_weights -no_info". Default application parameters are used elsewhere.

We plot the resulted θ against CR as shown in Fig. 1. And $\theta(CR)$ can be perfectly modeled with the inverse function as

$\theta(CR) = -45.7369/(CR + 44.0182) + 1.0268$ for 5-level 9/7 DWT.

$\theta(CR) = -44.3788/(CR + 42.3765) + 1.0265$ for 5-level 5/3 DWT.

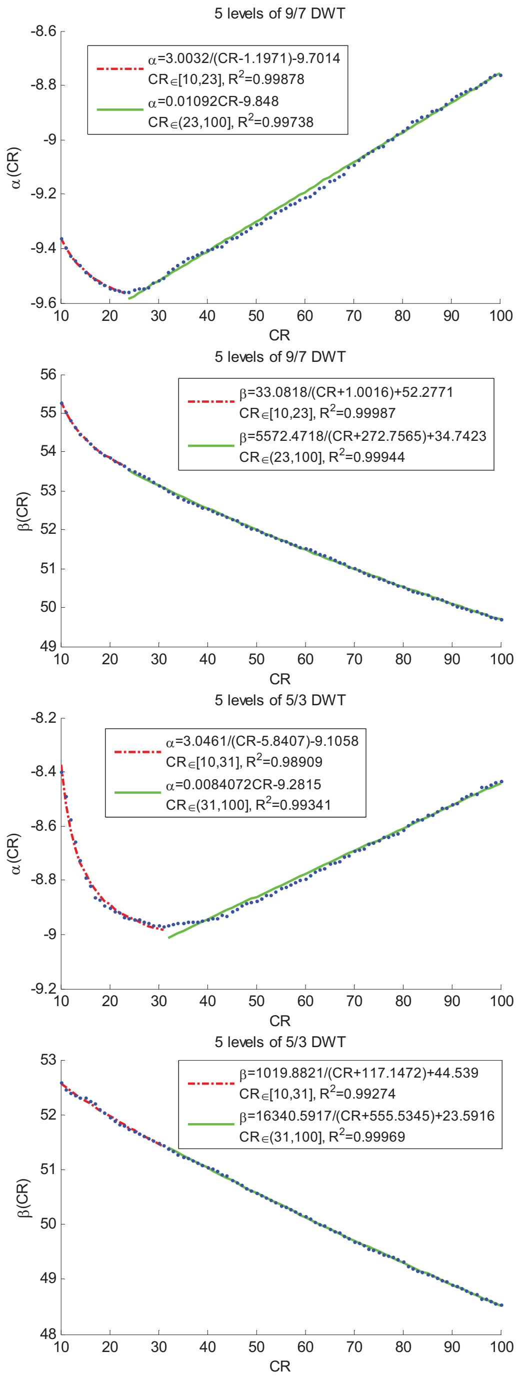


Fig. 2 $\alpha(CR)$ **and** $\beta(CR)$ **for 5 levels of 9/7 and 5/3 DWT.**

While solving the optimal θ, we also obtain α and β in (7) for each CR, as plotted in Fig. 2. $\alpha(CR)$ is a piecewise function and can be written as an inverse function and a linear function. $\beta(CR)$ can be modeled with two inverse functions.

4. QUALITY PREDICTION PERFORMANCE

We select sample images from two image sets to verify the prediction performance of the model. The two sets contain different images. The first image set includes 105 images used to estimate the model parameters. The other set consists of 50 images with similar scenes as the first set, as well as images with quite different contents. Limited to pages, only the experimental results with 5 levels of 9/7 and 5/3 DWT are shown in this paper.

We first verify the prediction accuracy with image set 1, and then verify the prediction performance for selected images from set 1 and set 2.

4.1. Prediction accuracy

With the quality model, PSNR values are predicted for 105 images from set 1. The differences between the predicted and actual PSNR values are calculated. We define $\eta = \frac{N_{diff\_n}}{N} \times 100\%$ to quantify the ratio of the PSNR difference. N is the number of sample images, and N_{diff_n} is the number of images with the PSNR difference lying in the $diff\_n$ range.

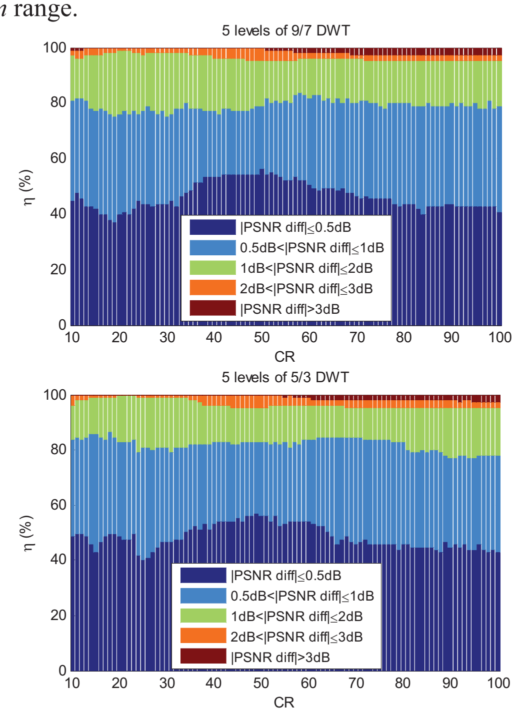
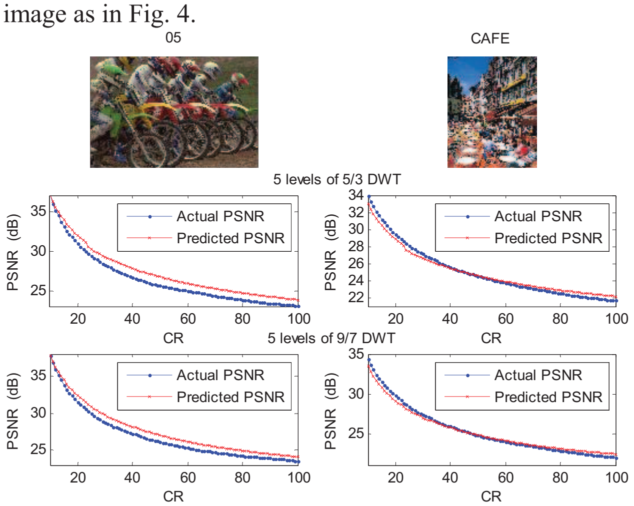


Fig. 3 Prediction accuracy for 5 levels of 9/7 and 5/3 DWT.

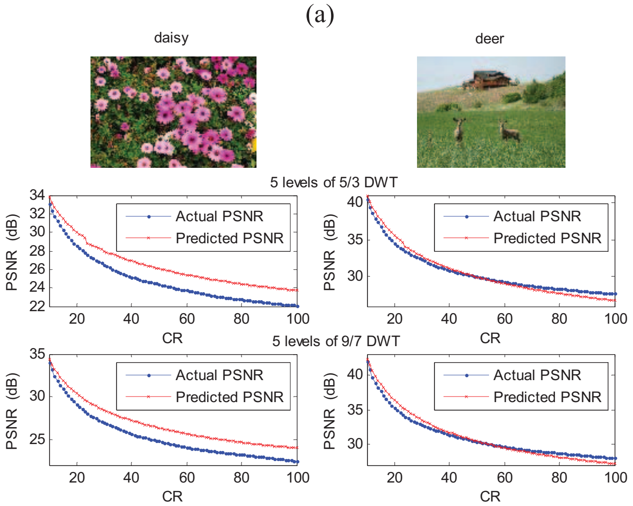
Fig. 3 shows η at different compression ratios for 5 levels of 9/7 and 5/3 DWT. The PSNR difference is less than 1dB for more than 75% images. The PSNR difference is less than 2dB for over 95% images. The prediction error is acceptable for usual applications.

4.2. Prediction performance

We compress the selected sample images from the two sets at the compression ratios ranging from 10:1 to 100:1 with 5 levels of 9/7 and 5/3 DWT, and calculate the quality ($PSNR_{actual}$) for the reconstructed images. Then we estimate the quality ($PSNR_{pred}$) with the proposed model (7). The two kinds of quality values are plotted against the CR for each image as in Fig. 4.



(a)



(b)

Fig. 4 Prediction performance. (a) Selected images from set 1. (b) Selected images from set 2.

As shown in Fig. 4 (a), the predicted PSNR values are close to the actual quality for selected images from set 1. And the largest difference is less than 1.5dB for the '05' image. The prediction performance for selected images from set 2 is shown in Fig. 4 (b). The predicted PSNR values of the images are close to the actual PSNR values with prediction errors less than 2dB.

5. CONCLUSION

We proposed a quality prediction model for color images coded with JPEG2000. This model uses the image activity measure as the image feature, and can predict the quality for an image at a given compression ratio without coding. The prediction error is less than 1dB for over 75% images, and less than 2dB for more than 95% images.

The advantages of the proposed model are that it predicts the quality at a pre-encoding stage, which can be useful for the filter selection and encoding parameter decision. Moreover, the computation cost of the prediction process is much less than that of the compression and quality evaluation.

6. REFERENCES

[1] W. J. Kim, J. W. Yi, and S. D. Kim, "A bit allocation method based on picture activity for still image coding," *IEEE Trans. Image Process.,* vol. 8, pp. 974-977, July 1999.

[2] P. Fleury, J. Reichel, and T. Ebrahimi, "Image quality prediction for bitrate allocation," in *Proc. Int. Conf. Image Processing*, 1996, pp. 339-342.

[3] A. Kourzi, D. Nuzillard, G. Millon, and F. Nicolier, "Quality estimation in wavelet image coding," in *13th European Signal Processing Conf.* Antalya, Turkey, 2005.

[4] H. Koumaras, A. Kourtis, C.-H. Lin, and C.-K. Shieh, "A Theoretical Framework for End-to-End Video Quality Prediction of MPEG-based Sequences," in *3rd Int. Conf. Networking and Services*: IEEE Computer Society, 2007, pp. 62-62.

[5] H. Koumaras and A. Kourtis, "Video Quality Prediction based on the Spatial and Temporal Classification of the Uncompressed Content," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, 2007, pp. 1-5.

[6] S. Saha and R. Vemuri, "An analysis on the effect of image features on lossy coding performance," *IEEE Signal Process. Lett.,* vol. 7, pp. 104-107, 2000.

[7] S. Saha and R. Vemuri, "How do image statistics impact lossy coding performance?," in *Proc. Int. Conf. Information Technology: Coding and Computing*, 2000, pp. 42-47.

[8] R. Kumar, S. Mitra, T. Karp, and B. Nutter, "Adaptive wavelet filter design for optimized image source encoding," in *Proc. Int. Conf. Information Technology: Coding and Computing*, 2003, pp. 478-482.