2

Multilevel Saliency Enabled Compression for Camera-Content Images

In view of the computational efficiency of DCT based compression methods [Wallace, 1992; Douak *et al*., 2011; Dhara and Chanda, 2007; Messaoudi and Srairi, 2016] as discussed in Chapter 1, our focus has been on designing a method in a similar domain. The major challenge was to prevent over-quantization of high-frequency regions of an image. Moreover, it was discussed in Chapter 1 that the human eyes steer an image according to its saliency, i.e., the major focus of the eyes are on the salient regions [Borji and Itti, 2013]. It was also observed that for camera content images (CCIs), our brain stores a multi-level saliency map in order to store image information. This provided us with the motivation to work on a multi-level saliency-based compression technique. The main issue in designing such an algorithm was to reduce the overhead information in order to send the shape of the salient regions and their corresponding rank. As discussed in Chapter 1, there is a trade-off between sending the exact region of interest (ROI) information where the ROI needs to be reconstructed on the decoder side, and sending the rectangular approximation where although the ROI is not needed to be reconstructed on the decoder side, compression ratio (CR) is compromised. We aimed to develop an ROI encoding which can provide the ROI information more accurately than that of rectangular approximation and yet needs low computation and space overhead.

In this Chapter we develop a multi-level saliency-based image compression algorithm that provides a statistically optimal trade-off between the overhead, perceptual quality, and CR [Rahul and Tiwari, 2018]. The devised method chooses variance as a basis to classify and rank the image into an optimal number of classes. The aim is to enable JPEG standard to judiciously retain high-frequency regions in the image that provides perceptual homogeneity in compression, particularly in the case of high compression requirement.

The rest of the chapter is organized in the following way. Section 2.1 explains the developed multi-level saliency detection and image coding algorithm. Section 2.2 presents the experimental results in order to show the performance of the developed method. The performance of our method is also compared with the state-of-the-art methods in terms of the perceptual quality of the reconstructed image and the rate-distortion. Concluding remarks are given in Section 2.3.

2.1 PROPOSED COMPRESSION METHOD

The proposed image encoding process is shown in Figure 2.1. For ease of implementation, various steps are briefly explained as follows:

The encoder processes a given image through two paths. The first path generates multi-level saliency map for the input image. The optimal number of classes is adaptively calculated by using efficiency of segmentation [Otsu, 1979]. The image is then segmented into the same number of salient classes by maximizing between-class variances [Huang and Wang, 2009]. Every class is then given a rank based on its importance by using their weighted-variances. The class having high weighted-variance is given higher rank i.e. more important, and vice-versa.

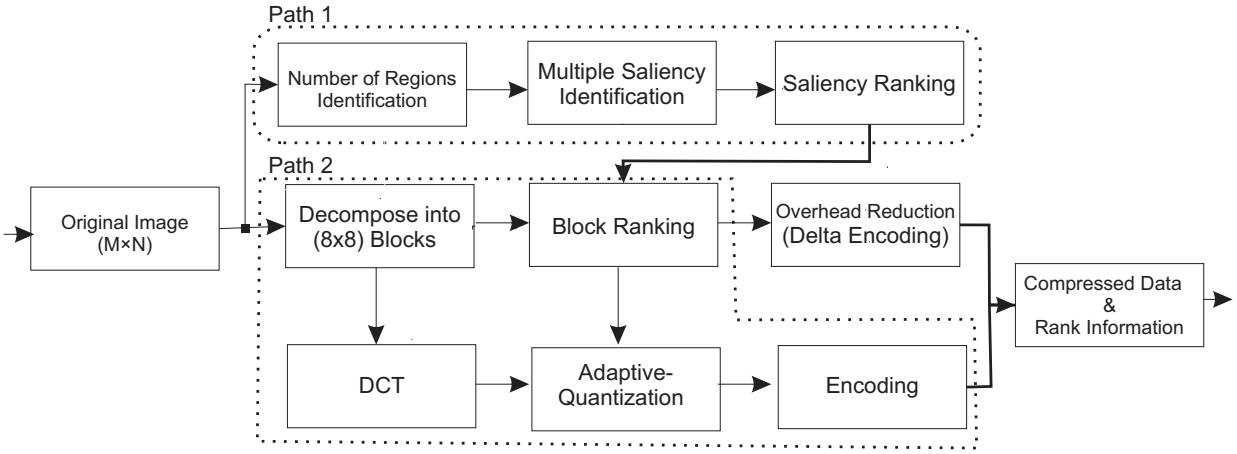

Figure 2.1 : Encoding Process of Proposed Method

The second path is used for the adaptive quantization. The image is decomposed into blocks of size 8×8 and each block is ranked based on the input saliency map obtained from the first path by applying probability bound. 2D-DCT coefficients of each block are quantized adaptively by the quantization parameters modified by the rank of the block. The higher rank block (i.e. more salient) will be lightly quantized and vice-versa. The quantized coefficients are then entropy coded and the overhead for the rank of the blocks is reduced by using delta encoding method [Schindler, 1970].

Unlike the method in [Christopoulos *et al.*, 2000], where decoder requires reproducing the ROI mask, making the decoder complex, the decoder of our method is simple as ROI information is sent to the decoder by the encoder. The reconstruction of the image is the inverse of the encoding steps. The detailed description of the key steps in the encoding process are given as follows:

2.1.1 Number of Regions Identification & Multiple Saliency Identification

Salient regions are identified by segmenting the image into K number of classes, to be discussed later, by maximizing between-class variances. For segmenting the image, the Otsu's segmentation method [Otsu, 1979; Huang and Wang, 2009; Lagarias *et al.*, 1998] is extended for K classes. Let these K classes be arbitrarily bounded by $K+1$ intensity levels $(t_0, t_1, t_2, ..., t_K)$ as $t_0 < t_1 < t_2 < ... < t_{K-1} < t_K$. For an image with L intensity levels, t_i $(0 \leq i \leq K)$ is intensity value of pixels with $t_0 = 0$, and $t_K = L-1$. Let i^{th} class $(1 < i < K-1)$ consists of all the pixels with their intensities in the range $[t_{i-1}, t_i - 1]$. Whereas, the K^{th} class consists of pixels with intensity values in the range of $[t_{K-1}, t_K]$. With these initial assumptions, probability of i^{th} class occurrence (ω_i), and the class mean (μ_i) are obtained as follows:

$$\omega_i = \sum_{j=t_{i-1}}^{t_i} p_j, \quad \mu_i = \frac{1}{\omega_i} \sum_{j=t_{i-1}}^{t_i} j p_j, \quad \mu_T = \sum_{i=1}^{K} \omega_i \mu_i \tag{2.1}$$

Here p_j is probability of the pixels with intensity value j, and (μ_T) is the mean of the image. Thereafter between-class variance (σ_K^2) can be obtained using (2.2), given below.

$$\sigma_K^2 = \sum_{i=1}^{K} \omega_i (\mu_i - \mu_T)^2 \tag{2.2}$$

σ_K^2 is function of ω_i and μ_i and these parameters, in turn, are functions of the chosen class boundaries $t_1, t_2, ...t_{K-1}$. It is desired to obtain an optimal set of the class boundaries that results into a maximum value of σ_K^2. This can be obtained by iteratively solving (2.2) for possible boundary values given in (2.1). The maximum value of σ_K^2 is called maximum between-class variance. Algorithm 2 shows the step by step procedure for threshold values optimization.

Algorithm 1 Number of Regions Identification

Input: Image to be Segmented, and required goodness-of-segmentation (GoS) (η_r)
Step1: Initialize a loop with $K = 1$
Step2: Use **Algorithm 2** to get optimized $[t_0, ..., t_K]$
Step3: Evaluate $\omega_i, \mu_i, \mu_T, \sigma_i^2$
Step4: Evaluate Between Class Variance as:
$\sigma_K^2 = \sum_{i=1}^{K} \omega_i (\mu_i - \mu_T)^2$
Step5: Calculate total variance (σ_T^2):
$\sigma_T^2 = \sigma_K^2 + \sum_{i=1}^{K} \omega_i \sigma_i^2$
Step6: Evaluate actual GoS (η_K):
$\eta_K = \frac{\sigma_K^2}{\sigma_T^2}$
Step7: Compare the required GoS (η_r) and actual GoS (η_K):
if $\eta_K \geq \eta_r$ **then**
exit with K regions
else
$K = K + 1$, and go back to **Step2**
end if
Output: Number of Regions (K)

To identify the total number of classes (K), it is proposed to first obtain the goodness-of-segmentation (GOS) (η_K), given in (2.3). Between class variance, σ_K^2, given in (2.2), and weighted-variance, S_i, given in (2.5), are used to calculate the total variance σ_T^2 and η_K, for initial value of $K = 2$. Value of K is incremented till the inequality given in (2.4) is satisfied for required value of (η_r), typically in the range of 0.8 to 0.99.

$$\sigma_T^2 = \sigma_K^2 + \sum_{i=1}^{K} S_i, \qquad \eta_K = \frac{\sigma_K^2}{\sigma_T^2} \tag{2.3}$$

As it can be observed from (2.3), η_K will be less than 1.

$$\eta_K \geq \eta_r \qquad (0 \leq \eta_r \leq 1) \tag{2.4}$$

Choosing number of classes (K) based on GOS helps to avoid the over-segmentation and under-segmentation situations. Algorithm 1 shows the proposed number of regions identification process. For a good segmentation, the required GOS η_r should be at least 0.80, as observed in the proposed work. The effect of η_r parameter is discussed in Section 2.2.

Algorithm 2 Threshold Optimization

Input: Number of regions (K)
Step1: Initialize the threshold values $(t_0,...,t_K)$ as:
$t_0 = 0$, $t_1 = (\frac{255}{K})$, $t_2 = (\frac{255}{K} * 2)$, $t_i = (\frac{255}{K} * i)$,... $t_K = 255$
Step2: Initialize two variables $max = 0$, and $j = 1$
Step3: Create a Function $Optimize([t_0, t_1, t_2, ..., t_K])$
Step4: Evaluate $\omega_i, \mu_i, \mu_T, \sigma_i^2$, for $0 < i < K$
Step5: Evaluate Between Class Variance as:
$\sigma_K^2 = \sum_{i=1}^{K} \omega_i(\mu_i - \mu_T)^2$
Step6: Compare the value of between class variance σ_K^2 and max:
if σ_K^2 is higher than the variable max and $j < K-1$ **then**:
$max = \sigma_K^2$
if (σ_K^2 with $t_j = t_j - 1$) is higher than the (σ_K^2 with $t_j = t_j + 1$) **then**
Call the Function $Optimize$ with $t_j = t_j - 1$ until $t_j > t_{j-1}$
else Call the Function $Optimize$ with $t_j = t_j + 1$ until $t_j < t_{j+1}$
end if
else If ($\sigma_K^2 > max$ and $j >= K-1$)
Exit and Keep the current optimized values $[t_0,...,t_K]$
else If ($\sigma_K^2 < max\_variance$ and $j < K-1$)
$j = j+1$, and call the Function $Optimize$ with previous values of $[t_0,...,t_K]$
else If ($\sigma_K^2 < max\_variance$ and $j >= K-1$)
Exit and keep the previous values $[t_0,...,t_K]$
end if
Step7: End Function $Optimize$
Output: $[t_0,...,t_K]$

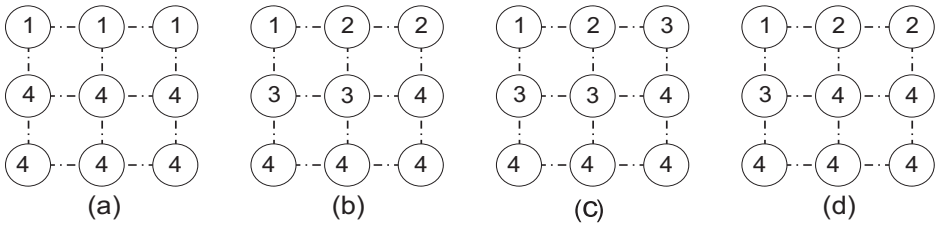

Figure 2.2 : Block ranking process: (a) to (d) represents pixel with different rank situations.

2.1.2 Saliency Ranking

In order to rank each class, consisting of randomly distributed pixels, weighted-variance of the pixels are obtained, using (2.5).

$$S_i = \omega_i \sigma_i^2, \qquad i = 1, 2, ..., K \tag{2.5}$$

where σ_i^2 is the variance of i^{th} class, as given in (2.6)

$$\sigma_i^2 = \frac{1}{\omega_i} \sum_{j=t_{i-1}}^{t_i} (j - \mu_i)^2 p_j \tag{2.6}$$

S_i for $1 \leq i \leq K$ is sorted in descending order. The i^{th} class pixels will get rank q where q is the position of the sorted S_i. Highest weighted-variance refers to the most salient class and gets the highest rank and vice versa, i.e. pixels corresponding to $max(S_i)$ will get rank $(r = 1)$ and $min(S_i)$ will be ranked $(r = K)$. The aim is to give more importance to the class with a considerable area having high variance. As expressed in (2.5), a class with a high variance but very small area (small ω_i) may get less importance than a class with relatively lower variance but a larger area.

2.1.3 Block Ranking

After classifying pixels based on the threshold values t_1, t_2, ..., t_{K-1} and ranking them according to sorted series of S_i, given in (2.5), we propose to use probability mass function (PMF) to rank every 8×8 blocks used in JPEG. Blocks having pixels with more than one rank may be the one that is on the border, or sometimes on the edges. The block is assigned rank r, whenever the empirically proposed probability bound (2.7) is satisfied, starting with $r = 1$. The motivation behind choosing this empirical formula was to provide an advantage to the lower ranked or more salient region when more than one rank pixels are present in the block. This approach saves the edges of the higher ranked regions, which are important for an end user.

$$\sum_{i=1}^{r} p_i \geq \frac{1}{K - r + 1}, \qquad r = 1, 2, ..., K \tag{2.7}$$

Where p_i is the probability of i^{th} ranked pixels in the block. To illustrate the use of probability bound and ranking of blocks, lets assume $K = 4$ and apply (2.7) on four different blocks of size 3×3 shown in Figure 2.2 (a) to (d). Considering Figure 2.2 (a), for example, it is found that $p_1 = 0.33$, $p_2 = p_3 = 0$, and $p_4 = 0.66$. The probability bound (2.7) is then applied, starting with $r = 1$.

The probability bound is satisfied for $r = 1$, resulting block rank to be 1. Here it can be observed that although p_4 is higher than p_1, i.e. pixels with rank 4 are more than pixels with rank 1, but the block was ranked as 1. The motivation behind choosing such empirical formula was to provide more importance to the lower ranked pixels (higher saliency) if their frequency is adequate. Similarly, blocks in Figure 2.2(b) to (d) gets rank 2, 3, and 4 respectively.

2.1.4 Adaptive Quantization of DCT Coefficients

DCT coefficients of every ranked block of size 8×8 is adaptively quantized as per their importance, estimated in terms of their rank values (r). The quantization table (T_{50}) used in JPEG baseline [Wallace, 1992] is proposed to be scaled by a factor F_r for r^{th} ranked block $(1 \leq r \leq K)$, and the same is controlled by two variables V_{ar} and Q_{am} as given in (2.8).

$$F_r = V_{ar} + (r-1)Q_{am} \tag{2.8}$$

$F_r = V_{ar}$ for $r = 1$ i.e., for the most salient blocks and value of F_r increases by $(r-1)Q_{am}$ as the saliency of the block decreases (i.e., r increases). Details of choice of V_{ar} and Q_{am} are given in Section 2.2. It can be observed from (2.8) that the scaling in the quantization table T_{50} is least for the most salient regions $(r = 1)$ in the image. The scaling increases gradually by a factor of Q_{am} for every subsequent region with lower saliency value (higher rank). The motivation behind doing this was to prevent an abrupt change in the reconstructed image as it is very likely that the rank of the adjacent blocks will be in close vicinity.

2.1.5 Overhead Reduction

The number of bits for sending the rank (r) information associated with various blocks will be $\lceil \log_2(K) \rceil$ and its value in terms of bpp will be n_o, as given in (2.9).

$$n_o = \frac{\lceil \log_2(K) \rceil}{8 \times 8} \tag{2.9}$$

For example, if an image is segmented in 8 regions, the overhead will be $\frac{\lceil \log_2(8) \rceil}{64} = 0.0469$. Although, the overhead value per-pixel is very less, the effect of this can still be present in case of a very low bit requirements. To further lower the overhead for sending the saliency map, the correlation between the adjacent blocks are used. Due to the Markovian property in images, it was observed that there was high correlation in the ranks of adjacent blocks. Figure 2.3 shows the histogram of the delta encoded rank matrix [Schindler, 1970] where the error between the current rank and the previous rank values evaluated. This helped in further reduction of overhead by applying delta encoding [Schindler, 1970] on rank matrix. In delta encoding, only the error between the current-block rank and the previous-block rank is encoded.

2.2 RESULTS AND ANALYSIS

In order to evaluate the robustness and efficiency of the proposed method, USC-SIPI Image data-set [Weber, 1997] is used. The data set contains 397 images in 5 different volumes as 155 texture images, 91 rotated texture images, 38 high altitude aerial images, 44 miscellaneous outdoor images and 69 frames taken from 4 video sequences. Images in each volume are of the dimensions between 256×256 pixels to 2048×2048 pixels. These images are 8 bits/pixel, for gray-scale images,

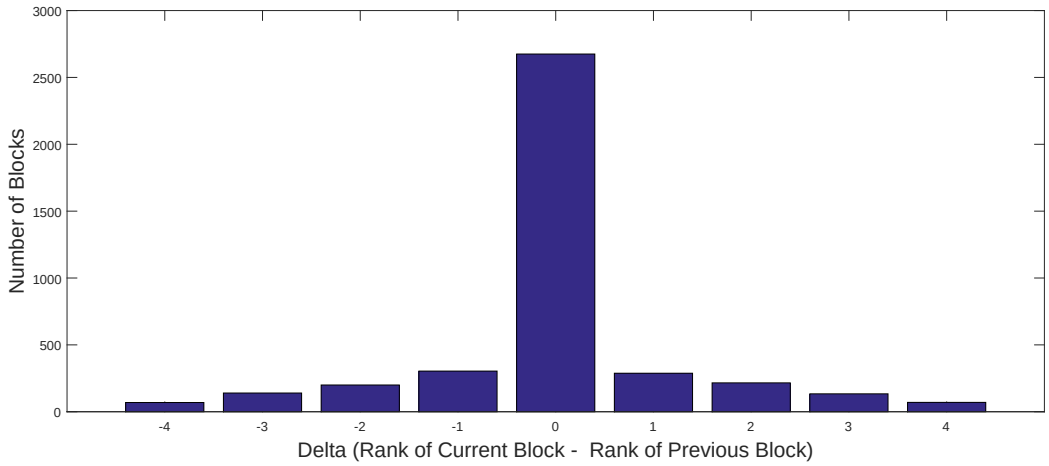

Figure 2.3 : Histogram of delta encoded rank matrix.

Table 2.1 : Effect of varying parameter (η_r) on the image data-set [Weber, 1997]

η_r	Number of Regions	R_1 area in %	R_2 area in %
0.78	2.83	36.18	42.23
0.81	3	35.38	30.14
0.84	3.33	34.37	25.37
0.87	3.83	36.79	33.98
0.90	4.33	38.17	30.71
0.93	5.17	29.66	28.46
0.96	7.33	38.49	23.82
0.99	11.5	28.88	15.52

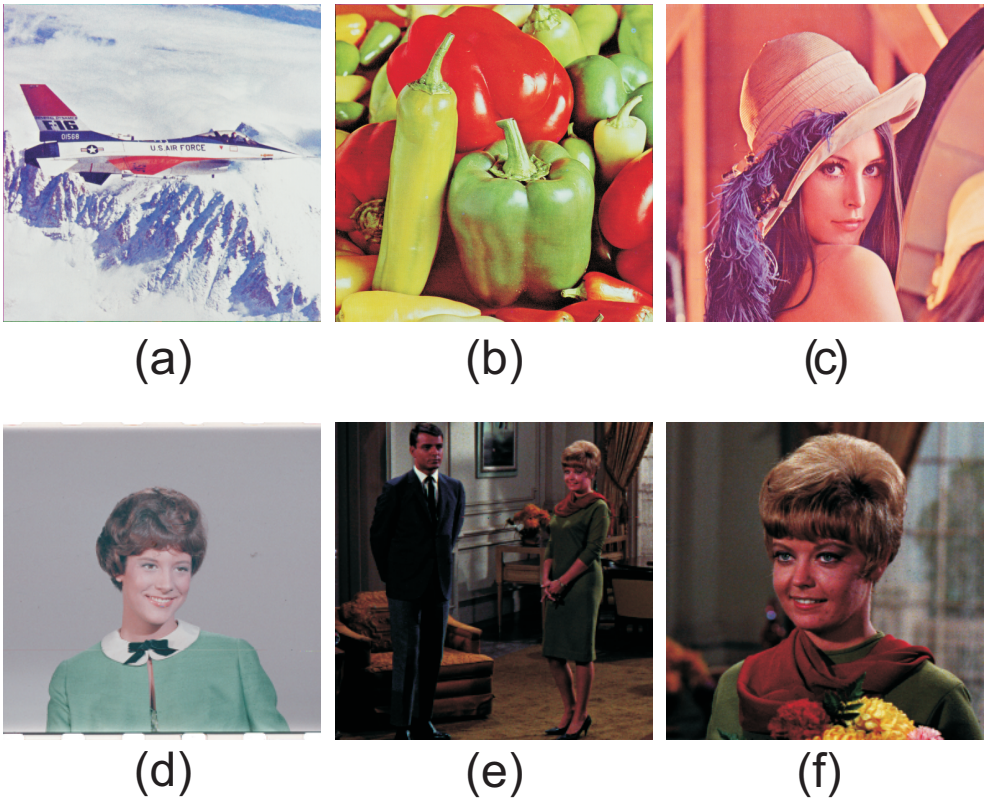

Figure 2.4 : Test images: (a) Airplane, (b) Peppers, (c) Lena, (d) Girl, (e) Couple, (f) Zelda. [a-c] dimension 512×512, [d-f] dimension 256×256.

and 24 bits/pixel, for color images. We used $4:4:4$ chroma sub-sampling, as in [Barua *et al.*, 2015], on the color images to quantize on all luminance and chrominance components. To compare the quality of the reconstructed image, PSNR [Wang and Bovik, 2002] and SSIM [Wang *et al.*, 2004] are used.

The key parameters in the proposed method are η_r, V_{ar} and Q_{am}. The number of classes in which the image is to be segmented depends on η_r. Table 2.1 shows average number of regions and the percentage of area with rank $r = 1$, and 2 as a function of η_r, for the image data-set in [Weber, 1997]. From Table 2.1, it is observed that number of classes obtained are 3 and 12 for $\eta_r = 0.8$ and 0.99, respectively. The value of η_r can be chosen as per the target bit-rate. Choosing lower value for η_r yields less number of salient regions, the overhead per block is less, which is suitable to achieve lower bit-rate. In order to control the overhead, an appropriate value for η_r between 0.8 to 0.99 can be selected for an image.

Region-wise effect of changing the parameters V_{ar} and Q_{am}, at 0.5 bpp, for Lena test image is shown in Table 2.2. To achieve 0.5 bpp, the quantization table T_{25} for JPEG baseline is used. Increasing Q_{am} yields improved quality at the most salient regions (i.e. $r = 1$) which is 38% of the image. By decreasing Q_{am}, the results behave like JPEG baseline. So, by changing the parameters V_{ar} and Q_{am}, we get adequate flexibility in controlling quality and compression ratio than that can be obtained in JPEG baseline method, as shown in Table 2.2.

The rate-distortion curve in Figure 2.5, shows that the PSNR of the proposed method is always higher compared to JPEG for the image's most salient regions $(r = 1)$ which constitutes an average of 31.1% of total area (or the number of pixels) of the test images. However, the overall

Table 2.2 : Effect of varying parameters V_{ar} and Q_{am} on 24 bpp Lena image compressed to 0.5 bpp and at $\eta_r = 0.95$. R_i denotes the regions in the image with rank i

V_{ar}	Q_{am}	PSNR R_1	PSNR R_2	PSNR Overall	SSIM Overall
1.4	4	31.28	26.43	26.12	0.9351
1.6	3.5	30.95	26.48	26.48	0.9450
1.8	3	30.59	26.63	26.40	0.9443
2	2.5	30.33	26.98	27.05	0.9509
2.3	2	29.87	27.18	27.39	0.9520
2.6	1.5	29.47	27.32	27.59	0.9555
2.8	1	29.25	27.54	28.14	0.9623
3	0.5	29.01	27.89	28.75	0.9666
JPEG Baseline		29	28.41	29.42	0.9714
Region Weight (%)		38	20	100	100

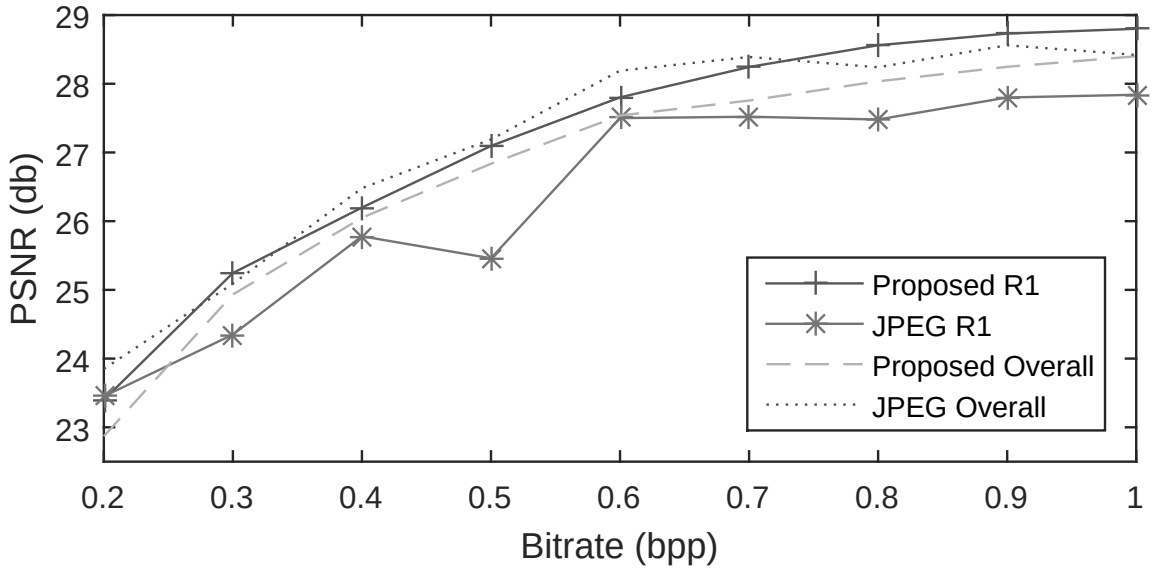

Figure 2.5 : Rate-distortion comparison (at $\eta_r = 0.95$) between proposed method and JPEG baseline on the data-set [Weber, 1997]. R1 denotes most salient regions in the image.

Figure 2.6 : Multiple salient regions after applying proposed saliency detection technique at $\eta_r = 0.92$. Brighter region indicates more saliency and vice versa.

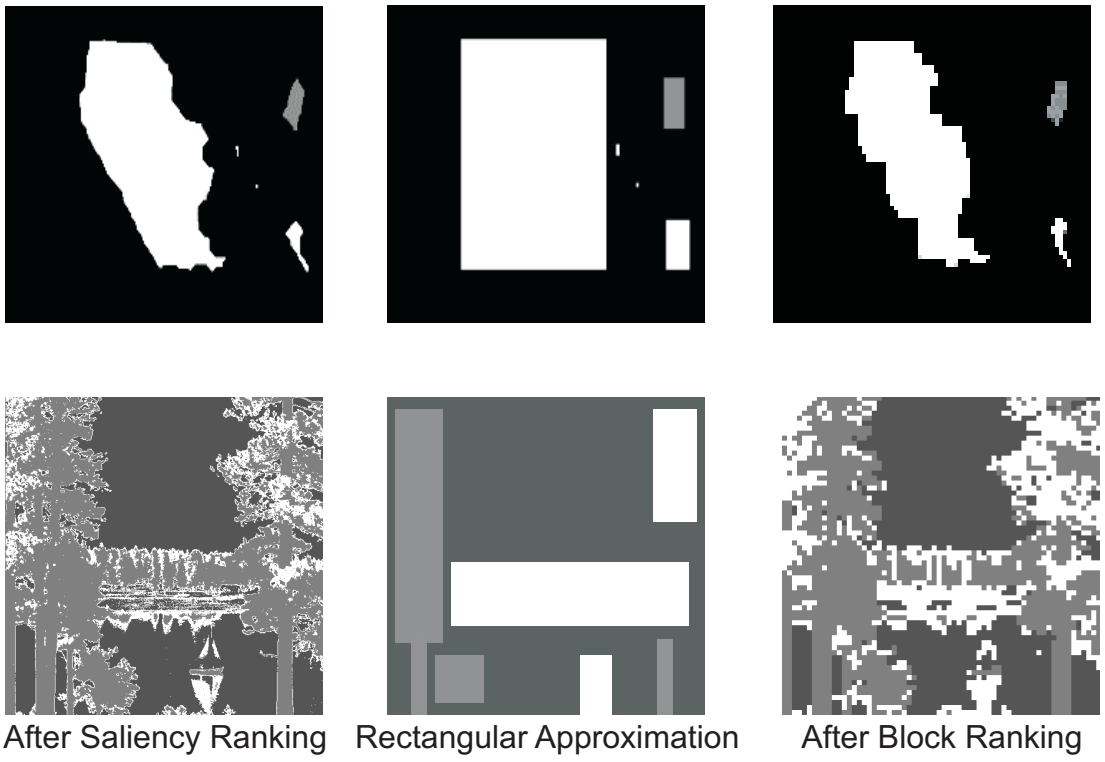

Figure 2.7 : Comparison of ROI reconstruction at the decoder side by using rectangular approximation and the proposed method.

quality has an average degradation of 1.57% compared to JPEG baseline. Since, there is a significant improvement in quality at the salient regions i.e. the regions of perceptual importance, the overall visual quality of the images has improved. It can also be observed from Figure 2.5, that between 0.4 to 0.6 bpp, there is a fluctuation in JPEG baseline curve for the most important regions i.e. a region with a large variance. The reason for this is non-homogeneous compression in JPEG, leading to compression artifacts. On the other hand, the proposed method gives a stable curve for the most important regions as well as for the overall image.

Table 2.3 presents the performance comparison of the proposed method with the recently published DCT and DWT based algorithms in [Dhara and Chanda, 2007; Douak *et al.*, 2011; Boucetta and Melkemi, 2012; Messaoudi and Srairi, 2016] and the JPEG baseline [Wallace, 1992] on the same test images as shown in Figure 2.4. The value of PSNR (in db) shown in Table 2.3 for proposed method, and JPEG baseline is for the most important regions $(r = 1)$ of the image and for rest of the methods for the whole image. When any DWT or DCT methods in [Wallace, 1992; Dhara and Chanda, 2007; Douak *et al.*, 2011; Boucetta and Melkemi, 2012; Messaoudi and Srairi, 2016] are applied on the most important regions $(r = 1)$, the mean-square error (MSE) of these regions is higher than overall MSE. The reason for this is while applying these transform based methods in a region with high variance, the energy compaction is lesser compared to a region with lower variance [Yang *et al.*, 2016], which results in higher MSE after quantization to achieve lower bit-rate. A similar example can be referred from rate-distortion curve in Figure 2.5, where the PSNR after applying JPEG on the most important regions of the image is always lower than the overall image. This information suggests that the PSNR values provided in Table 2.3 for the methods in [Dhara and Chanda, 2007; Douak *et al.*, 2011; Boucetta and Melkemi, 2012; Messaoudi and Srairi, 2016], which is for the whole image will have a lower value of PSNR for the regions with (r=1). It is clear that the proposed method outperforms those reported in [Wallace, 1992; Dhara and Chanda, 2007; Douak *et al.*, 2011; Boucetta and Melkemi, 2012; Messaoudi and Srairi, 2016].

The Figure 2.7 shows the performance of the proposed multi-level saliency ranking and block ranking on Lena. For illustration purpose, the saliency of the regions is shown in gray-scale with brighter region implies more salient and vice-versa. The face region of Lena image gets high importance regardless of having comparatively low variance. This is because the face region has large area compared to other regions. It is observed that the proposed approach of block ranking provides effective means of labeling Region of Interest (ROI).

Figure 2.7 shows the comparison in terms of accuracy of reconstructed ROI at the decoder side, between the proposed method of sending the ROI and the rectangular approximation used in state-of-art saliency enabled methods [Barua *et al.*, 2015; Christopoulos *et al.*, 2000]. The reference images can be seen in Figure 1.1. It is observed that the proposed approach of sending ROI information to the decoder by using block ranks, retains the ROI structure better than the rectangular approximation of ROI. The average overhead found to be 0.00038 bpp while using the rectangular approximation, and 0.0091 bpp while using the proposed method.

Although, the rate-distortion curve in Figure 2.5 shows that at 0.2 bpp, the PSNR of the proposed method converges to JPEG for the image's most salient regions $(r = 1)$, however, the overall perceptual quality of the proposed method is significantly better than JPEG. To illustrate this, Figure 2.9, and Figure 2.10 provide a better visual comparison of the reconstructed images after applying JPEG baseline and the proposed method. In Figure 2.9 (a), Lena image of dimension 512×512 is shown with portion of its face area highlighted. Figure 2.10 (b), and (c) are the reconstructed images after applying JPEG baseline and the proposed method, respectively, to achieve bit-rate of 0.2 bpp. Similarly, in Figure 2.10 (a), Baboon image of dimension 512×512 is shown and a high variance region is highlighted. Figure 2.10 (b), and (c) are the reconstructed images after applying JPEG baseline and the proposed method, respectively, to achieve bit-rate of

Table 2.3 : Performance comparison between the proposed method, JPEG Baseline [Wallace, 1992], CBTF-PF [Dhara and Chanda, 2007], CDABS [Douak *et al.*, 2011], GA-DWT [Boucetta and Melkemi, 2012], and dLUT [Messaoudi and Srairi, 2016] algorithms

Image	R_1 Area in%	JPEG		CBTC – PF		CDABS		GA – DWT		dLUT		Proposed	
		PSNR	bpp	PSNR	bpp	PSNR	bpp	PSNR	bpp	PSNR	bpp	PSNR	bpp
Airplane	19.6	29.71	0.97	30.36	1.04	31.40	0.72	31.16	0.49	31.16	0.48	31.43	0.45
Peppers	45.2	30.16	1.47	30.15	1.5	30.33	0.88	31.20	0.83	31.19	0.88	31.49	0.85
Lena	51.1	32.57	1.03	31.93	1.17	32.77	1	32.76	0.66	32.65	0.74	33.37	0.72
Girl	21.58	34.98	0.62	35.13	0.6	36.96	0.69	35.90	0.41	35.86	0.38	36.26	0.37
Couple	44.4	31.49	0.94	32.44	1	33.07	1.13	32.87	0.89	32.62	0.79	32.51	0.81
Zelda	30.8	31.24	1	31.31	1.12	32.05	1.09	31.98	0.76	32.01	0.82	32.89	0.80
Average	**35.45**	**31.69**	**1.01**	**31.89**	**1.07**	**32.76**	**0.92**	**32.65**	**0.67**	**32.58**	**0.68**	**32.99**	**0.67**

0.31 bpp. The overall perceptual quality of the reconstructed images, obtained from the proposed method is significantly better than that of the reconstructed images from JPEG baseline. Also, the highlighted area of the images obtained from the proposed method shows very less distortion compared to the reconstructed images from JPEG baseline.

Usually, aerial images are compressed at a lower bit rate compared to the other multimedia data set. Figure 2.8 can also be referred for better visual comparison, where an aerial image is compressed at 0.31 bpp by the proposed and the JPEG method. A region with high variance is highlighted in Figure 2.8 to show the efficacy of the proposed method on such regions. To achieve bit-rate of 0.31 bpp, JPEG quantizes the highlighted area heavily, yielding significantly higher artifacts compared to the proposed method.

The overhead of sending the ROI information to the decoder is dependent on η_r. We set the parameter $\eta_r = 0.95$ and got average overhead for the data set without any post-processing as 0.047 bpp. Although this overhead seems to be low, however, at high CR requirements (i.e. low bpp), reduction in this overhead is desired. For this, after applying delta encoding, the average overhead reduced to 0.0299 bpp.

2.3 CONCLUSIONS

For the requirement of high compression ratio without any significant quality loss in salient regions of the reconstructed image, we propose to classify a given image into multiple ranked ROI's and quantize the corresponding DCT coefficients judiciously. The goodness-of-segmentation (GOS) is used as a parameter to adaptively identify the number of classes in the image. The multiple ROI's are obtained by maximizing between-class variances and ranked by using within-class variances. Coefficients of the quantization table used in JPEG are adaptively changed as a function of the rank of the ROI, and its application in JPEG's framework resulted in significant improvement in compression performance.

We achieved (average) 2.88% better quality at the most salient regions, which contained an average of 31.1% area in 397 test images. Due to improvement in the salient regions, the overall perceptual quality of the reconstructed image is better than JPEG. The experimental results obtained on different color images clearly showed that our proposed method outperforms the

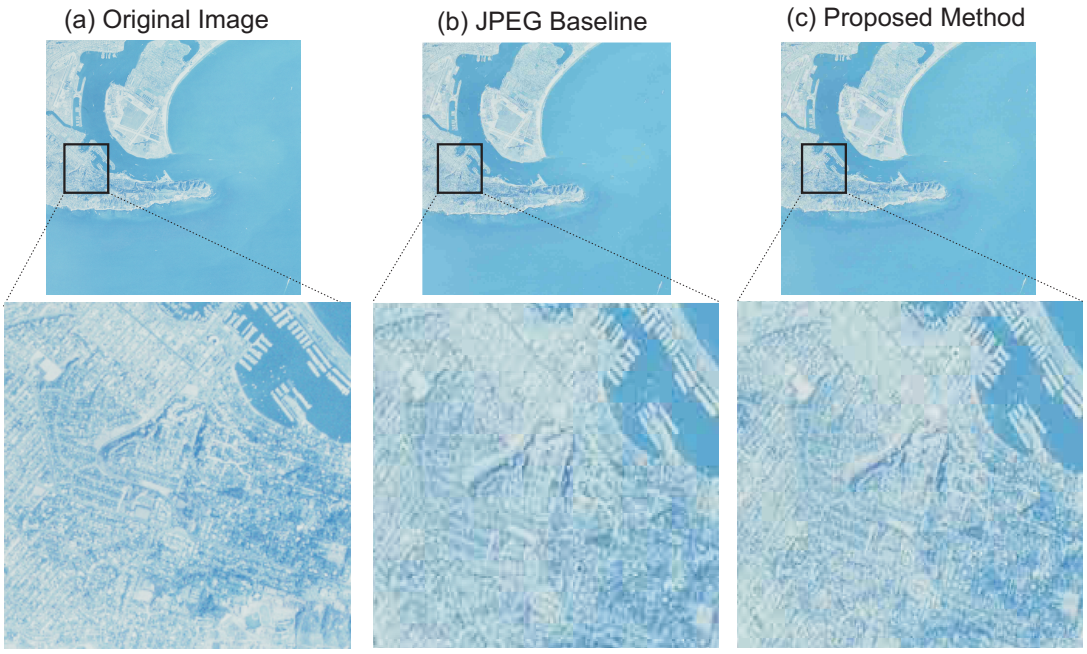

Figure 2.8 : Reconstructed aerial image from JPEG baseline and the proposed compression technique. (a) Original Image (24 bpp). (b) Reconstructed image compressed at 0.31 bpp, using JPEG baseline. (c) Reconstructed images at 0.5 bpp, using the proposed method.

recently published similar methods in terms of the perceptual quality of the reconstructed images. By ranking 8×8 blocks, we were able to reconstruct the ROI at the decoder side more accurately than the recent state-of-the-art works where the ROI is approximated by a rectangular bounding box. The average overhead is also found to get reduced by 36.38% by using delta encoding as a post-processing on rank information matrix.

...

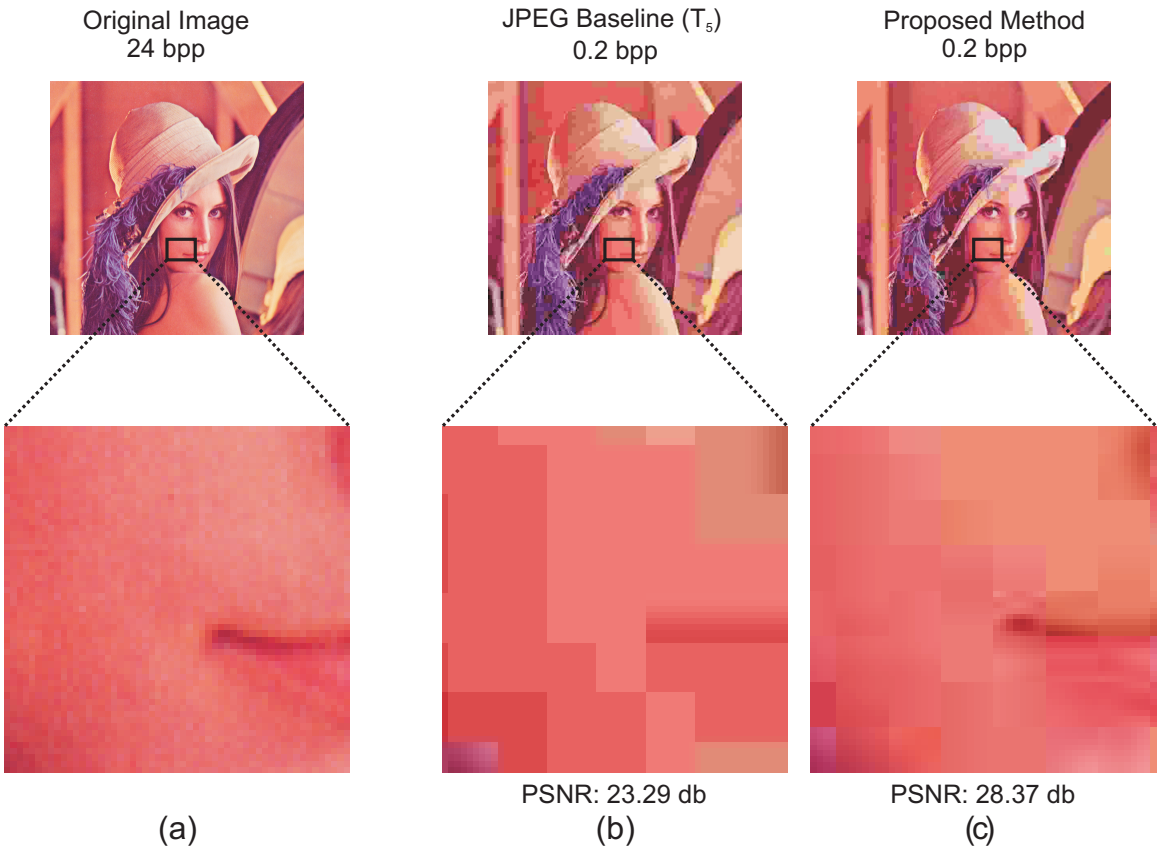

Figure 2.9 : Reconstructed Lena image from JPEG baseline and the proposed compression technique. (a) Original Image (24 bpp). (b) Reconstructed image compressed at 0.2 bpp, using JPEG baseline. (c) Reconstructed images at 0.2 bpp, using the proposed method.

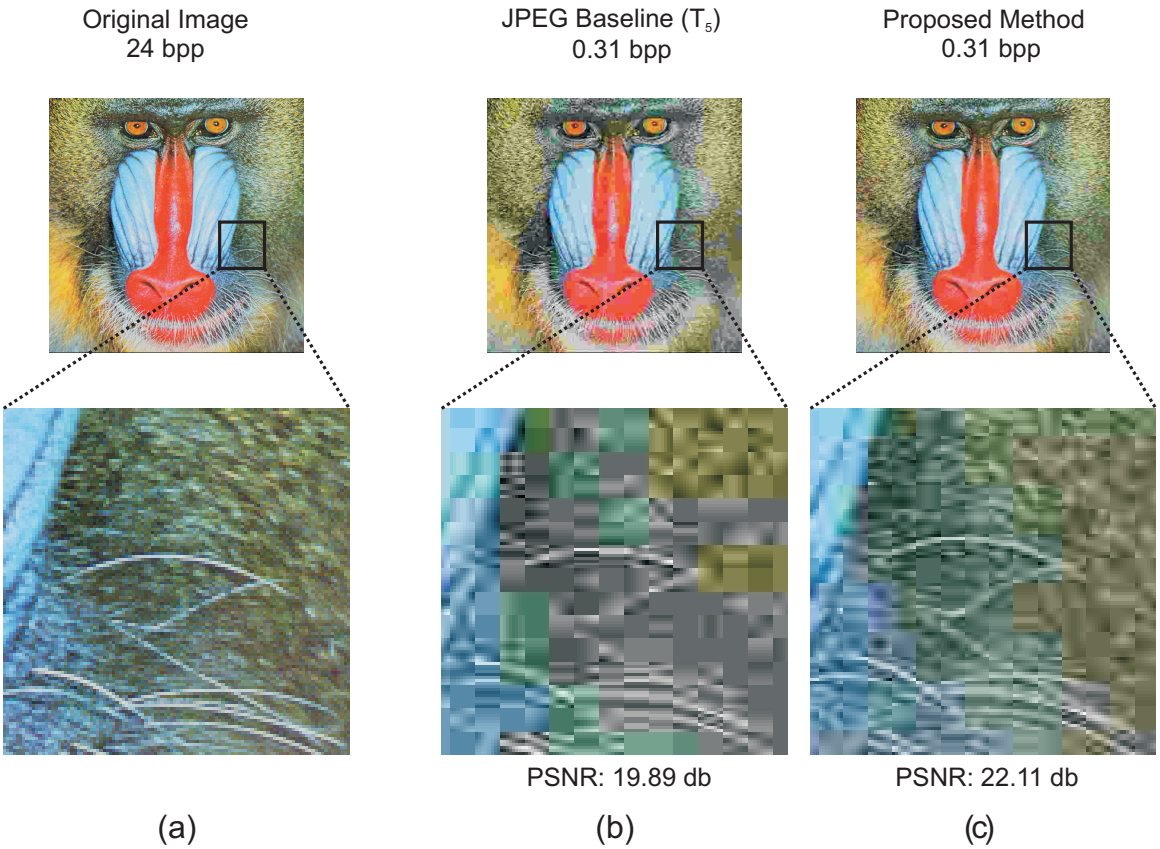

Figure 2.10 : Reconstructed Baboon image from JPEG baseline and the proposed compression technique. (a) Original Image (24 bpp). (b) Reconstructed image compressed 0.31 bpp, using JPEG baseline. (c) Reconstructed image at 0.31 bpp, using the proposed method.