AN IMPROVED DC RECOVERY METHOD FROM AC COEFFICIENTS OF DCT-TRANSFORMED IMAGES

Shujun Li, Junaid Jameel Ahmad, Dietmar Saupe*

Dept. of Computer & Information Science
University of Konstanz, Germany

C.-C. Jay Kuo

Ming Hsieh Dept. of EE
University of Southern California, USA

ABSTRACT

Motivated by the work of Uehara *et al.* [1], an improved method to recover DC coefficients from AC coefficients of DCT-transformed images is investigated in this work, which finds applications in cryptanalysis of selective multimedia encryption. The proposed under/over-flow rate minimization (FRM) method employs an optimization process to get a statistically more accurate estimation of unknown DC coefficients, thus achieving a better recovery performance. It was shown by experimental results based on 200 test images that the proposed DC recovery method significantly improves the quality of most recovered images in terms of the PSNR values and several state-of-the-art objective image quality assessment (IQA) metrics such as SSIM and MS-SSIM.

1. INTRODUCTION

The discrete cosine transform (DCT) is an orthogonal transform with sub-optimal performance in terms of de-correlation efficiency [2]. Since DCT is easier to implement than the optimal Karhunen-Loève transform (KLT), it has been widely used in signal and image processing applications, especially for lossy image and video compression. Many well known image and video coding standards, including JPEG and MPEG-1/2/4/H.26x, are based on 2D block DCT [3].

Among all DCT coefficients, the first one, called the DC (direct current) coefficient, plays the most important role. It represents the average intensity of a block and carries most of the energy and the perceptual information. Actually, DC coefficients of all blocks form a thumbnail version of the original image at a lower spatial resolution. When the block-DCT is applied to image/video compression, DC coefficients often consume more bits than other DCT coefficients called AC (alternative current) coefficients. To give an example, to encode the 8-bit gray-scale Lenna image of size 512×512 using the 8×8 block-DCT and the default quantization table of JPEG [4, Table K.1], we estimate that about 16% of all coding bits will be used to encode quantized DC coefficients, where the estimation is obtained by calculating entropies of quantized DCT coefficients at different locations. If we apply differential coding to DC coefficients, the ratio will decrease to 14%, which is still significant.

Due to the importance of DC coefficients, it was argued that encrypting DC coefficients *only* may effectively conceal significant visual information to achieve the goal of light-weight selective encryption [5]. For example, Li *et al.* [6] proposed to encrypt DC coefficients to conceal a rough view of MPEG-encoded video sequences to achieve perceptual encryption. Other hybrid image/video encryption schemes using the combination of DC encryption and secret AC permutations were proposed in [7]. However, since secret

permutations are not secure against plaintext attacks [8], the hybrid encryption scheme can be downgraded to the DC encryption only.

Since DCT is an orthogonal transform, DC encryption is believed to be secure because DC coefficients are independent of AC coefficients. It was however shown by Uehara, Safavi-Naini and Ogunbona [1] that most DC coefficients of a DCT-transformed natural image can be approximately recovered from AC coefficients. Their DC recovery method, called the USO method, exploits two properties of most digital images. First, there is strong correlation between neighboring pixels. Second, the DC coefficient of each block is constrained to an interval defined by the DC-free edition of this block, *i.e.,* pixel values calculated only from AC coefficients.

In this work, we propose an improved DC recovery method, which outperforms the USO method in recovery quality. Our idea is to minimize the under/over-flow rate of pixel values during the recovery process. We provide experimental results to demonstrate the superior performance of the improved DC recovery method in terms of PSNR and nine objective image quality assessment (IQA) metrics in the MeTriX MuX Visual Quality Assessment Package [9], which includes SSIM [10], MS-SSIM [11], IFC [12], VIF [13], etc.

The rest of the paper is organized as follows. The USO method is reviewed in Sec. 2. The proposed DC recovery method is described in detail in Sec. 3. Finally, the last section gives concluding remarks and future work.

2. THE USO METHOD

In this section, we present the implementation detail of the USO method by following [1], where some missing details in the original paper are filled in with our best efforts. Since in [1] it was shown that the USO method is robust against quantization, in this paper we do not consider quantization. To facilitate our discussion, we denote the DC coefficient and the average intensity of an image block B by $DC(B)$ and $\overline{B}$, respectively. We use the orthogonal form of 2D DCT so that $DC(B) = N \cdot \overline{B}$, where N is the block size. Given a block B, define $B^{(d)}$ as the block derived from B by setting $DC(B)$ to d, *i.e.,* $B^{(d)} = (B - \overline{B}) + d/N = B + (d - DC(B))/N$. The valid range of pixel values is denoted by $[t_{\min}, t_{\max}]$.

The USO method is based on the following two properties of most digital images.

Property 1 *The difference between two neighboring pixels is a Laplacian variate with zero mean and a small variance.*

Property 2 *The range of pixel values calculated from $B^{(0)}$ (i.e., only from AC coefficients) constrains the value of $DC(B)$.*

By exploiting Property 1, the unknown DC coefficient of a block can be estimated from its neighboring blocks with known DC coefficients by minimizing the differences of adjacent pixels along the block boundary. Three patterns of adjacent pixels were considered in the USO method: one horizontal or vertical pattern, and two diag-

onal patterns as shown in [1, Fig. 2]. The smoothest pattern (*i.e.*, the one with the smallest gradient) is chosen for DC estimation.

Property 2 can be used to determine a valid range of $\mathrm{DC}(B)$. Since $t_{\min} \le B = B^{(0)} + \mathrm{DC}(B)/N \le t_{\max}$, we obtain

$$N(t_{\min} - \min(B^{(0)})) \le \mathrm{DC}(B) \le N(t_{\max} - \max(B^{(0)})), \quad (1)$$

which gives an estimation of $\mathrm{DC}(B)$ with an accuracy defined by $N(t_{\max} - t_{\min} + \min(B^{(0)}) - \max(B^{(0)}))$.

Uehara *et al.* took the above two properties into account and proposed the following method to recover DC coefficients of a DCT-transformed image with known AC coefficients. Without loss of generality, we assume the input image is a DC-free image, *i.e.*, an image composed of blocks whose DC coefficients are all zeros.[1]

Step 1: Choose a corner block of the DC-free image as the initial reference block B_0 and estimate DC coefficients of other blocks (relative to $\mathrm{DC}(B_0) = 0$) by scanning the whole image from B_0 to the diagonally opposite corner block. Property 1 is used to estimate the DC of each block from its neighboring block. When there are two neighboring blocks, the two estimates are averaged. Ideally, the output of this step is an image identical with the original image except that the average intensity is darker by $\mathrm{DC}(B_0)/N$.

Step 2: The goal of this step is to adjust the whole image so that the recovered image has the same average intensity as the original one. First, calculate the valid DC range of each block B_i from Eq. (1). Then, for each DC range $[d_{i,\min}, d_{i,\max}]$, get a range of the intensity adjustment as follows: $[d_{i,\min}^*, d_{i,\max}^*] = [(d_{i,\min} - \mathrm{DC}^*(B_i))/N, (d_{i,\max} - \mathrm{DC}^*(B_i))/N]$, where $\mathrm{DC}^*(B_i)$ denotes the relative DC coefficient of B_i estimated in Step 1. Next, calculate $d_{\min}^* = \max_i(d_{i,\min}^*)$ and $d_{\max}^* = \min_i(d_{i,\max}^*)$. Finally, the whole image is brightened by $(d_{\min}^* + d_{\max}^*)/2$.

Step 3: Repeat Steps 1 and 2 by choosing the four corner blocks as initial reference blocks, and average the images obtained from four different scans to get the final result.

Step 4: Since the estimated DC values in previous steps may not be accurate, the output pixel values in Step 3 may not be in the valid range. Then, a post-processing operation, which includes the re-scaling of the whole image or the adjustment of under/over-flow pixel values only, can be performed.

No specific postprocessing operation in Step 4 was mentioned in [1]. We found from our experiments that the following post-processing scheme works well.

- If the dynamic range of pixel values is larger than $t_{\max} - t_{\min}$, scale the whole image to $[t_{\min}, t_{\max}]$;

- If the dynamic range of pixel values is not larger than $t_{\max} - t_{\min}$, adjust the average intensity of the image towards $[t_{\min}, t_{\max}]$ until all pixel values are within $[t_{\min}, t_{\max}]$.

The above scheme is therefore adopted in our implementation of the USO method. For the four test images with reported PSNR values in [1], we obtained a higher average PSNR value (21.7 dB vs. 20.9 dB) under the above setting to generate data given in [1, Table III-2].

3. PROPOSED DC RECOVERY METHOD

One major drawback of the USO method is the low accuracy of the DC estimation process for some images. Small estimation errors in Step 1 may propagate to result in a large number of under/over-flows of pixel values. These under/over-flows can also affect Step 2 due to the inaccurately estimated DC ranges of some blocks.

[1]In principle, the DC coefficients can be any fixed value. In [1], the mid-point of the valid range is used. Here, we use zero to simplify our description.

To illustrate this problem, we examine the 8-bit gray scale image in Fig. 1(a). The four images obtained from four different scans (after Step 2) are shown in Fig. 2(a)–(d). Their pixel value ranges are $[-83.0, 338.0]$, $[-90.3, 345.3]$, $[-92.0, 347.0]$, $[-136.3, 391.3]$, respectively. We see that the error propagation effect is very serious. By averaging the four images, the pixel value range is $[-88.6, 303.0]$, which is still far from the valid one. After scaling the range to $[0, 255]$, we get a better image as shown in Fig. 1(b). However, the quality remains poor with a PSNR value of 14.3, an SSIM score of 0.732 and an MS-SSIM score of 0.711.

(a)   (b)

Fig. 1. (a) A test image of size 384×256 and (b) the recovered image using the USO method.
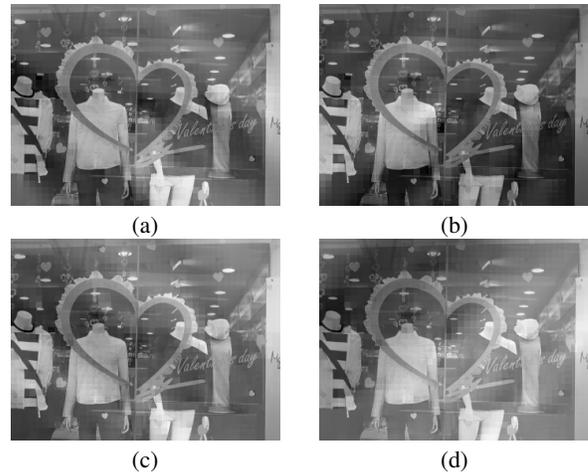

(a)   (b)

(c)   (d)

Fig. 2. Four intermediate images recovered from the DC-free edition of the test image Fig. 1(a) using the USO method.

3.1. Under/Over-Flow Rate Minimization (FRM) Method

To limit the error propagation effect, we apply Property 2 during the relative DC estimation process in Step 1 as follows. For each block, after the relative DC estimation is obtained, we immediately check if the estimated DC coefficient is outside of the valid range. If this occurs, we re-adjust the estimated DC toward the valid range until all pixel values fall in interval $[t_{\min}, t_{\max}]$. After that, the scanning process proceeds to the next block. Since the DC-bounding process is done during the scanning process for all blocks, the output of Step 1 will contain *no* underflow or overflow pixel values. Then, Steps 2 and 4 become unnecessary and can be removed.

The modified DC estimation method can effectively limit error propagation. However, it encounters a new problem. That is, it may not be able to recover the DC coefficient of each block accurately if the estimated value of $\mathrm{DC}(B_0)$ is far from the ground truth. This is because the estimates obtained in Step 1 are actually DC coefficients

of all blocks *relative* to $DC(B_0)$. Thus, it is important to find an accurate estimate of $DC(B_0)$ to ensure that DC coefficients of most blocks are estimated with an acceptable accuracy.

While there is no prior knowledge on the ground truth of $DC(B_0)$, we can exploit a statistical approach to estimate its value for natural images. It is based on an interesting phenomenon of the run-time DC-bounding process in Step 1: when the estimate of $DC(B_0)$ is closer to the ground truth, under/over-flows in pixel values tend to occur less frequently. Although some counter examples exist, this observation holds for most tested natural images.[2]

We use an example to illustrate the above observation. For the test image shown in Fig. 1(a), the relationship between the under/over-flow rate and the estimated $DC(B_0)$ is shown in Fig. 3 for four different scans. The under/over-flow rate is defined as M_{FR}/M_B, where M_{FR} is the number of blocks with under/over-flow pixel values and M_B is the total number of blocks in the image. In Fig. 3, the dotted line gives the ground truth of $DC(B_0)$ while the dashed line gives the $DC(B_0)$ value that has the minimal under/over-flow rate. When there are multiple points that have the minimal under/over-flow rate, the dashed line shows the mid-point of the range defined by the leftmost and the rightmost points. It is clear that the minimal under/over-flow rate always occurs around the ground truth of $DC(B_0)$.
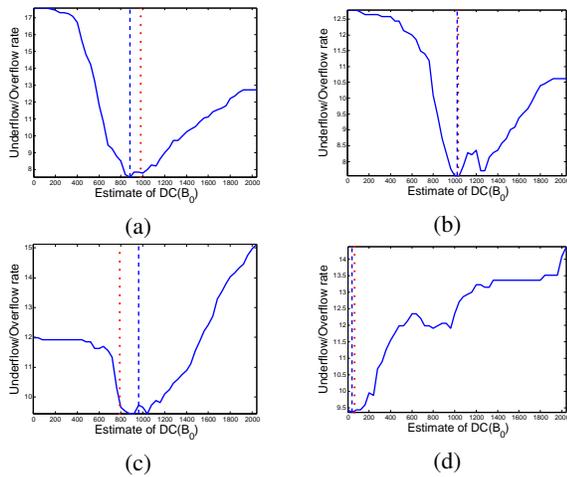

Fig. 3. The relationship between the under/over-flow rate and the estimate of $DC(B_0)$ for the test image Fig. 1(a) where Subfigures (a)-(d) give results for Scans 1-4, respectively.

If the best estimate of $DC(B_0)$ is used for each scan, we obtain the results in Fig. 4. By comparing Figs. 4 with 2, we see that the proposed FRM method performs significantly better. By averaging the four images obtained by four scans, we show the final recovered image in Fig. 5. This recovered image has a PSNR value of 23.2, an SSIM score of 0.900 and an MS-SSIM score of 0.924.

3.2. Computational Complexity

The above FRM method is essentially an optimization process of the observable under/over-flow rate during the DC estimation process. To get the best result, an exhaustive search of all possible values of $DC(B_0)$ could be tried. When the step size is Δ, $N(t_{max} - t_{min} + \min(B_0^*) - \max(B_0^*))/\Delta$ values of $DC(B_0)$ are checked for each

---

[2]It is difficult to theoretically analyze the accuracy and sensitivity of the estimation process of $DC(B_0)$. We conducted experiments on 200 test images to confirm its correctness and efficiency (see Sec. 3.3 for more detail).
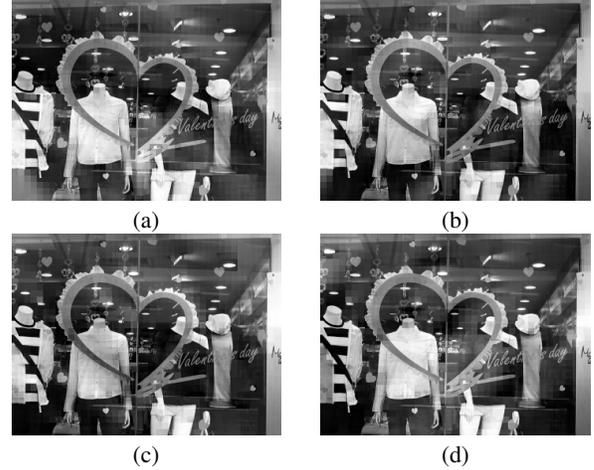

Fig. 4. The images obtained by minimizing the under/over-flow rate for four scans of the test image in Fig. 1(a), where Subfigures (a)-(d) give results for Scans 1-4, respectively.

Fig. 5. The final recovered image using the FRM method from the DC-free edition of the test image in Fig. 1(a).

block in each scan, and the computational complexity of the FRM method is $O\left(M_B N(t_{max} - t_{min} + \min(B_0^*) - \max(B_0^*))/\Delta\right) \leq O\left(M_B N(t_{max} - t_{min})/\Delta\right)$. In contrast, the USO method has a complexity of $O(M_B)$, so it is faster than the FRM method.

To reduce the computational complexity, we notice that the relationship between the under/over-flow rate and $DC(B_0)$ is close to a uni-modal function. Thus, a binary-search strategy can be used to locate the optimal $DC(B_0)$. At the beginning, the two ending points and the midpoint of the whole valid DC range are searched. Then, the DC range is reduced to be the sub-interval corresponding to the two smaller underflow/overflow rates. This process is repeated until a pre-defined precision is reached. The computational complexity of such a binary search process is less than $O\left(M_B \log_2(N(t_{max} - t_{min})/\Delta)\right)$.

Note that the computational complexity of the FRM method is $O(M_B)$ if $\max(B_0^*) - \min(B_0^*) \approx t_{max} - t_{min}$, which may happen for some images. In this case, it is as efficient as the USO method.

3.3. Experimental Results

To verify the effectiveness of the proposed FRM method, we conducted experiments on an image database of 200 test images. Our MATLAB code and the test images are available at `http://www.hooklee.com/default.asp?t=AC2DC`. To measure the quality of recovered images, we compare the performance of the proposed FRM method with that of the USO method using the PSNR values and other nine IQA metrics in the MeTriX MuX Visual Quality Assessment Package [9]. The results are shown in Fig. 6, where the x-axis is the image index, and the y-axis is the
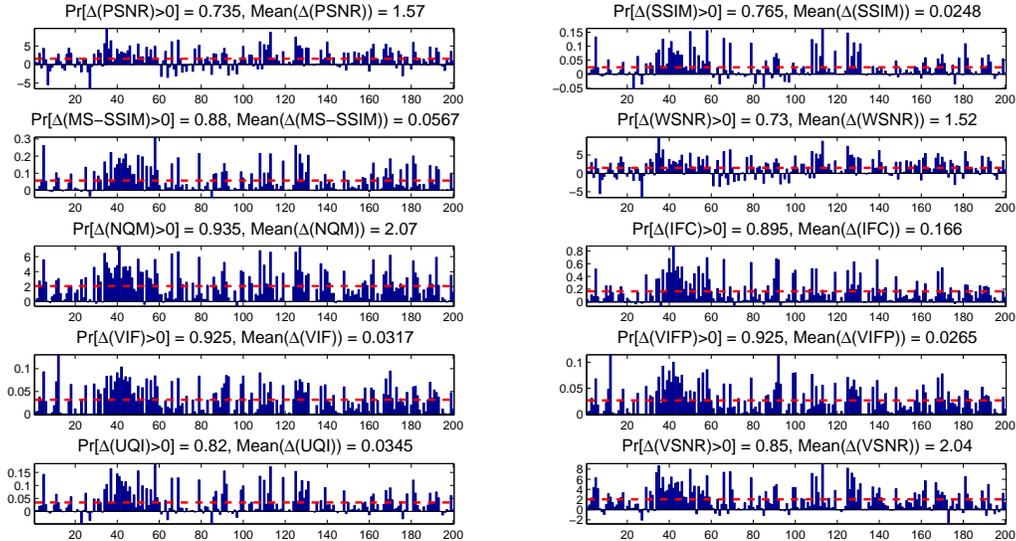
Fig. 6. The performance improvement of the proposed FRM method over the USO method using different IQA metrics. Note that the range of some metrics (SSIM, MS-SSIM, VIF, VIFP, UIQ) is [0,1], so the mean values corresponding to these metrics are relatively small.

difference of IQA scores and the dashed line shows the mean of the IQA differences of all images. A positive y-value means the proposed FRM method outperforms the USO method. We can see that the performance improvement is consistent and significant for most images. Although PSNR and WSNR values become significantly worse for a few images (*e.g.*, Images 7 and 27), their visual quality decreases only slightly or remains nearly the same if measured with other (more accurate) metrics like MS-SSIM and VIF.

4. CONCLUSION AND FUTURE WORK

A method to recover DC coefficients from AC coefficients of DCT-transformed images, called the USO method, was proposed in [1]. In this work, we proposed an improved DC recovery method and called it the FRM method. Experiments showed the performance improvement of the FRM method over the USO method for 200 test images in terms of the PSNR value and nine different IQA metrics.

There are several ways to further improve the proposed FRM method. For instance, if we have some statistics of the image to be recovered, we may be able to get a better estimate of $DC(B_0)$ from the minimized underflow/overflow rate. We may define the underflow/overflow rate in a different way, *e.g.,* modulating the simple definition with the degree of the underflow/overflow. In addition, more advanced DC prediction and averaging schemes may be used by exploiting the prior knowledge on the underlying image. Another possible improvement is to change the locations of the initial reference blocks and/or the scanning pattern. Theoretically speaking, any block can be the initial reference block and any scanning pattern can be used. It may also be beneficial to have multiple initial reference blocks and scanning patterns. One of the implications of the DC recovery method is that a fewer amount of information about DC coefficients should be encoded in image compression. We will investigate this possibility as well in our future work.

5. REFERENCES

[1] T. Uehara, R. Safavi-Naini, and P. Ogunbona, "Recovering DC coefficients in block-based DCT," *IEEE Trans. Image Processing*, vol. 15, no. 11, pp. 3592–3596, 2006.

[2] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Computers*, vol. 23, no. 1, pp. 90–93, 1974.

[3] Y. Q. Shi and H. Sun, *Image and Video Compression for Multimedia Engineering*, CRC Press, 2nd edition, 2008.

[4] ISO/IEC, "Information technology – Digital compression and coding of continuous-tone still images: Requirements and guidelines," ISO/IEC 10918-1 (JPEG), 1994.

[5] L. Weng and B. Preneel, "On encryption and authentication of the DC DCT coefficient," in *Proc. SIGMAP'2007*, 2007, pp. 375–379.

[6] S. Li, G. Chen, A. Cheung, B. Bhargava, and K.-T. Lo, "On the design of perceptual MPEG-video encryption algorithms," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 17, no. 2, pp. 214–223, 2007.

[7] L. Tang, "Methods for encrypting and decrypting MPEG video data efficiently," in *Proc. ACM MM'96*, 1996, pp. 219–229.

[8] S. Li, C. Li, G. Chen, N. G. Bourbakis, and K.-T. Lo, "A general quantitative cryptanalysis of permutation-only multimedia ciphers against plaintext attacks," *Signal Processing: Image Communication*, vol. 23, no. 3, pp. 212–223, 2008.

[9] M. Gaubatz, "MeTriX MuX visual quality assessment package," http://foulard.ece.cornell.edu/gaubatz/metrix_mux.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[11] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Conf. Record Asilomar CSSC'2003*. 2003, vol. 2, pp. 1398–1402, IEEE.

[12] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Processing*, vol. 14, no. 12, pp. 2117–2128, 2005.

[13] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.