Artifacts and Antiforensic Noise Removal in JPEG Compression

Bismitha N[1] Anup Chandrahasan[2] Prof. Ramayan Pratap Singh[3]
[1]M.Tech. Student [2]Assistant Professor
[1,2]Department of Electronics and Communication Engineering
[1,2]Younus College of Engineering and Technology Kollam, Kerala

Abstract— The JPEG stands for "Joint Photographic Expert Group" and, as its name suggests, was specifically developed for storing photographic images. The compression is usually lossy process. The original image can't be restored completely from the compressed object. There are a lot of works held to regain original image without any loss. The aim of this work is to improve the quality of the image by adding anti-forensic noise removal technique to the artifacts removed image. The dither algorithm is used to remove anti-forensic noise. In this paper, each block from 8x8 blocks divided into m patches. Every column in dictionary is taken as an atom. So the obtained matrix is given to next section and form DC transform matrix. The de-quantized result is scanned to minimize total variation and the obtained image will be in artifacts removed manner. Eventually, to improve the quality of the image dither algorithm is used. The different images are used to demonstrate the improved quality of the image and also measured the PSNR and SSIM.

Key words: Compression, Lossy Process, Anti-Forensic Noise Removal, Dither, Patches, Dictionary, Dequantized, Total Variation, Artifacts, PSNR, SSIM

I. INTRODUCTION

The JPEG stands for "Joint Photographic Expert Group" [1], [2] and, as its name suggests, was specifically developed for storing photographic images. It has also become a standard format for storing images in digital cameras and displaying photographic images on internet web page. The JPEG standard specifies both the codec, which defines how an image is compressed into a stream of bytes and decompressed back into an image, and the file format used to contain that stream. The compression is usually losing (lossy) [1] process, which means that most of the compression is obtained by loss of data. The original image can't be restored completely from the compressed object. The purpose of JPEG to compression is to either transmit through a channel or to store and retrieve to decompress the image. The original image is compressed to form compact information. It is used for storage or transmission. Then it is decompressed to form approximation of the original image. It is due to loss of data.[1] There are a lot of works held to regain original image without any loss. The aim of this project is to improve the quality of the image by adding anti-forensic noise removal technique to the artifacts removed image.

Basically, JPEG compression consists of three stages. They are splitting, DCT, quantization. The JPEG image is split [2] into 8x8 blocks. Each block is divided into 15 patches[4], [5],[6]. Every column in dictionary is taken as an atom. Then the obtained matrix is given to next section form DCT. The quantization is a lossy process. Then the matrix is encoded to form compressed result. The encoded result is decoded to make the decompressed result. So the reverse action is needed in each section. The de-quantized result is scanned in four manners. They are Zig - Zag scanning, Inverse scanning, horizontal scanning and vertical scanning. These are the things, which come under total variation (TV) method.

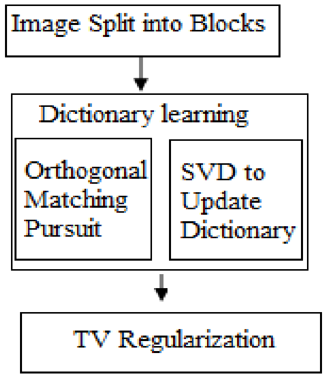


Fig. 1: Block diagram showing the Dic-TV process.

Artifacts can appear when perform block-based coding for quantization. If compression ratio is high, more visible artifacts appear in decompressed images. Different types of artifacts were occur in JPEG decompressed images. In JPEG compression, loss of information is happened under quantization section. There are two classes of methods to reduce the artifacts. They are image enhancement method and image restoration method.

Fig. 2: The input jpeg image of house

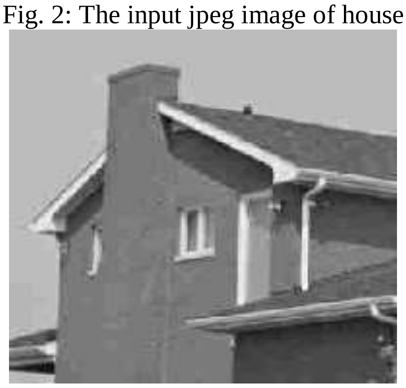


Fig. 3: The compressed jpeg image of house.

So, dictionary TV method is used to improve the quality of image. Even also the quality is not regained purely. To reach the quality of the image is trying through this paper. The main thing considering here is the artifacts produced during JPEG image compression. Along with the learned dic TV method, Anti-forensics noise removal method is adopted for improving the quality. The learned Dic-TV keep on the feature of original one.

II. PREVIOUS WORKS

The colour image is converted to gray scale image through YUV [2]. If gray scale image is used, then directly give to processing compression. The image is split into blocks. Each block is considered as a dictionary to form matrix for DCT. The quantization stage is to divide the above cosine transform coefficients by a quantization table point wisely, and the quantized values are rounded to their nearest integers. The final stage is to use lossless compression coding to generate a compressed data file. The decompression for JPEG images involves lossless decoding, de-quantization and computing the inverse DCT toeach block.

Fig. 4: Image obtained through Dic-TV method

Every coloumn in Dictionary is an atom. The vector α is generated randomly with few non zeros in random locations and with random values. Then find the one atom that best matches the signal.The Dictionary Learning [4] is a two-step iterative method. The first step is to use the orthonormal matching pursuit (OMP) algorithm to update the encoding coefficients and the second step is to use SVD to update the dictionary. The combination of the first and second terms requires the restored image is a sparse linear combination of elements in the dictionary. The restored image can keep the features of the original input image.

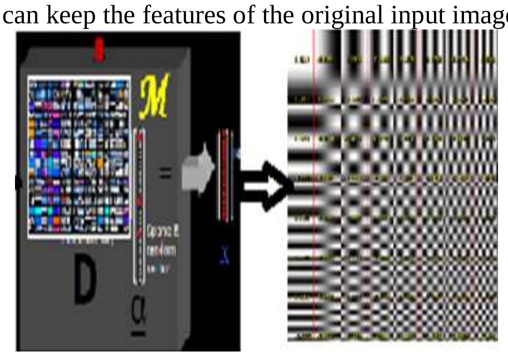


Fig. 5: Sparse linear combination of elements in the dictionary.

The combined equation of the above two steps are given below,

$$(\gamma_{i,j}^{*}, D^{*}) = \arg_{\gamma_{i,j}} \min_{D} \sum_{(i,j)\varepsilon P} \left(\frac{1}{2}\|R_{i,j}u^{0} - D_{\gamma_{i,j}}\|_2^2 + \mu_{i,j}\|\gamma_{i,j}\|_0\right) \quad (1)$$

The first term $\|R_{i,j}u^0 - D_{\gamma_{i\,j}}\|_2^2$ is related to the representation of their stored image in the dictionary. The second term $\|\gamma_{i,j}\|_0$is used to require the encoding coefficients vector to be sparse.$\gamma_{i,j}^*$ is the encoding coefficients and D^* is the dictionary.

The D is a dictionary of size m^2-by-c attached to their stored image with c atoms in the dictionary; $R_{i,j}$ is the sampling matrix of size m^2-by-N [5] to construct a patch for the part of $\boldsymbol{u}$; $\gamma_{i,j}$ is a vector of size c-by-1 containing the encoding coefficients for the patch $\boldsymbol{u}$ of represented in the dictionary; $P = \{1, 2, \ldots, n - m + 1\}^2$ denotes the index set for different patches of $\boldsymbol{u}$; $\|\cdot\|_2$ denotes the Euclidean norm of a vector; $\|\cdot\|_0$ denotes the number of non-zero elements; The parameter λ is a positive parameter of data fitting term, $\mu_{i,j}$ and is the positive patch-specific weight.

The existing total variation based model assumed that the minimizer was piecewise constant. Different artifacts [7] (i.e., staircase artifacts) are introduced, especially for the images with more textures. Based on the ideas of sparse representation and energy minimization methods, decompress the JPEG images with less artifacts and better textures.

To reducing artifact model for restoring JPEG decompressed images in the discrete setting, they do only one dictionary learning step and then solve the TV model [7] with fixed dictionary.

$$u^{*} = \min_{u\varepsilon U} \lambda TV(u) + \sum_{(i,j)\varepsilon P} \frac{1}{2}\|R_{i,j}u - D^{*}\gamma_{i,j}^{*}\|_2^2 \quad (2)$$

There are number of methods and equations were adopted for improving the quality of the image. Even also the decompressed images never regain the quality of original one [1]. So the difficulty behind image processing is different from each method.

III. THE PROPOSED MODEL

The fixed dictionary [6] is used for making forward discrete cosine transformation matrix. Thus improved quality of input matrix was obtained at the DCT block. To reduce compression artifact, filtering was used at encoded side. The filterings used are zig-zag [1], inverse zig-zag, horizontal and vertical manners filtering. As enhancement here propose to improve the quality of the final image from TV method, by using the same quantization matrix used in jpeg compression step.The DCT – AC [1] component,

$$P(X_i = x) = \begin{cases} -\frac{1}{\lambda_i} * \mathrm{sgn}(x) * \log\left(C(1 - 2|x|)\right) + e^{-\lambda_i\left(\frac{q}{2}\right)}, & x = 0 \\ -\frac{1}{\lambda_i} * \log\left(1 - \left(1 - e^{-\lambda_i q}\right)\right) * x - \left(\frac{q}{2}\right), & x > 0 \end{cases} \quad (3)$$

where

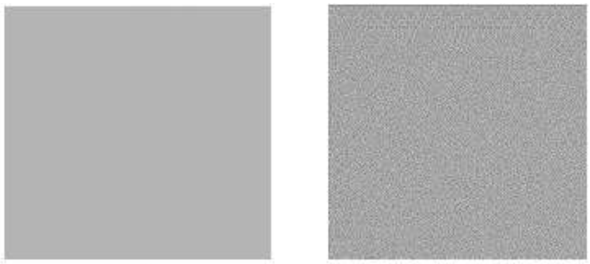
$$C = 1 - e^{-\lambda_i\left(\frac{q}{2}\right)}$$

The DCT – DC component,

$$P(X_i = x) = q * x - \left(\frac{q}{2}\right) \quad (4)$$

By doing so, the high frequency information is also regained after quantization and we obtain improved psnr and ssim values. The quality of the image was improved by adding anti-forensic noise removal technique to the artifacts removed image.

The reduced artifact output contains noise. The output consists of the transmitted image and the noise. That noise unclear the gray and white pixels present in the output [8]. To removing the noise dithering algorithm were implemented. The DC is preserved and the positive AC coefficients are multiplied by negative one to remove the noise.



(a) (b)

Fig. 6: (a) Plain image; (b) appears shade gray because of dithering.

The dither image is used to remove anti-forensic noise. Every column in dictionary is taken as an atom. So the obtained matrix is given to next section and form DC transform matrix.

Fig. 7: Image obtained through Anti-Forensic Noise removal method.

The quantization matrix consists of DC and AC coefficients. The AC coefficients are of lower frequency coefficients and higher frequency coefficients. The human eye responds to lower frequency coefficients. The high frequency coefficient as zero is not determined. So artifacts were occurred.

The high frequency coefficients were regained through minimizing total variation. The learned Dic-TV consists of dictionary learning from each block and minimizing total variation of the restored image. The OMP consist of two iterative steps. The orthogonal matching pursuit algorithm consists of dictionary learnind and minimizing total variation. The dictionary learning consists of two algorithms orthogonal matching pursuit algorithm and SVD to update dictionary.

The de-quantized result is scanned in four manners to form total variation output and the obtained image will be perfect in artifacts removed manner. Eventually, to improve the quality of the image dithering algorithm is required. The different images are used to demonstrate the improved quality of the image and also measured the PSNR and SSIM.

IV. THE COMPARISON RESULTS

This paper proposes two parameters for evaluating the quality of compressed image. They are PSNR value and SSIM VALUE. The PSNR value can ensure the output is received with the same or good quality at the receiver side. The PSNR value,

$$PSNR(u,u_r)=10\log10\frac{255^2}{\sum_{1\le i\le N}\frac{([u]i-[u]i)^2}{N}} \qquad (5)$$

The average SSIM index is used to evaluate the overall image quality. The larger the value is, the better the restoration result. The local SSIM index is defined by,

$$SSIM(u, u_r)=\frac{1}{N}\sum_{i=1}^{N} SSIM_{local}(u(i), u_r(i)) \qquad (6)$$

Where

$$SSIMlocal(u(i), (ur(i)=\frac{[2\mu(u(i)\mu(u_r(i))+C_1][2\sigma(u(i)u_r(i)+C_2}{[\mu^2(u(i)+\mu^2(u_r(i)+C_1][\sigma^2(u(i))+\sigma^2(u_r(i)+C_2}$$

(7)

The dither image helps the plian image to seen like gray. There are different types of dithering. The Fig.6.(b) shows the dithered image. In Fig.7.shows Antiforensic noise removed image.

q	Name	Before compression	After compression	Reducing artifacts (PSNR)	Anti-Forensic Noise removal (PSNR)
10	House	08.107	03.732	34.6080	40.2145
	Pepper	11.339	06.122	34.2842	37.2271
	Glomeruli	06.611	03.382	37.2366	40.0355
	Cameraman	10.718	05.297	35.1362	37.1465
	Barbara	44.859	21.764	34.4266	36.6824
	Village	42.004	17.096	34.3744	37.7546
	Lena	32.640	15.175	35.0719	37.8483
50/3	House	08.107	04.438	35.3409	40.0843
	Pepper	11.339	07.237	34.7011	37.2094
	Glomeruli	06.611	03.973	37.1109	39.4269
	Cameraman	10.718	06.360	35.0573	36.7183
	Barbara	44.859	28.693	34.5403	36.0103
	Village	42.004	22.828	34.3110	37.3048
	Lena	32.640	18.416	35.8458	38.4418
25	House	08.107	05.569	35.5333	40.0817
	Pepper	11.339	08.811	34.8543	36.9806
	Glomeruli	06.611	04.955	34.9826	40.5349
	Cameraman	10.718	08.061	35.0629	36.2515
	Barbara	44.859	37.305	34.6867	35.5594
	Village	42.004	31.087	34.3056	37.0340
	Lena	32.640	23.357	36.1096	38.2852

				Average =35.122	Average =37.940

Table I: The PSNR Comparison Results

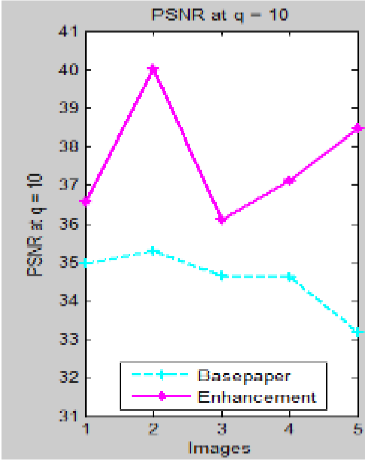


Fig. 8: Analysis graph for PSNR at q=10

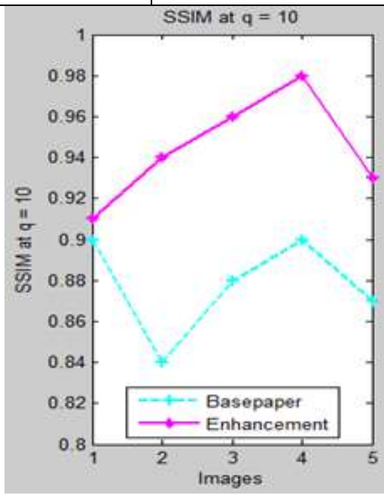


Fig. 9: Analysis graph for SSIM at q=10

q	Name	Before compression	After compression	Reducing artifacts (SSIM)	Anti-Forensic Noise removal (SSIM)
10	House	08.107	03.732	0.8455	0.9326
	Pepper	11.339	06.122	0.9022	0.9821
	Glomeruli	06.611	03.382	0.9250	0.9712
	Cameraman	10.718	05.297	0.8956	0.9281
	Barbara	44.859	21.764	0.8975	0.9640
	Village	42.004	17.096	0.8719	0.9572
	Lena	32.640	15.175	0.8718	0.9526
50/3	House	08.107	04.438	0.8435	0.9423
	Pepper	11.339	07.237	0.9130	0.9822
	Glomeruli	06.611	03.973	0.9152	0.9648
	Cameraman	10.718	06.360	0.9050	0.9156
	Barbara	44.859	28.693	0.9092	0.9589
	Village	42.004	22.828	0.8691	0.9523
	Lena	32.640	18.416	0.8890	0.9570
25	House	08.107	05.569	0.8435	0.9334
	Pepper	11.339	08.811	0.9143	0.9803
	Glomeruli	06.611	04.955	0.8830	0.9497
	Cameraman	10.718	08.061	0.9060	0.9002
	Barbara	44.859	37.305	0.9176	0.9537
	Village	42.004	31.087	0.8670	0.9490
	Lena	32.640	23.357	0.8948	0.9522
				Average =0.8895	Average =0.9514

Table 2: The SSIM Comparison Results

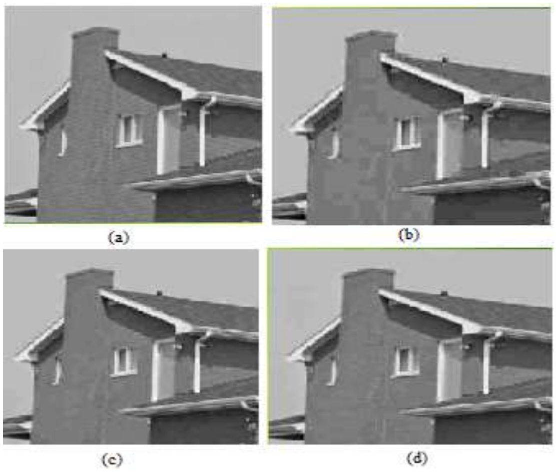


Fig. 10: (a) Original image; Image compressed (b) at q=10 (c) q=50/3 (d) at q=25.

The analysis graph using the images in the order of cameraman, glomeruli, house, parrot, pepper.The psnr and ssim value is better than the previous dic-tv method.

Fig. 11: The first column: Original image; the second column: compressed at q=10 and q=50/3; the third column: obtained through Dic-TV method; the fourth column: Dithering method.

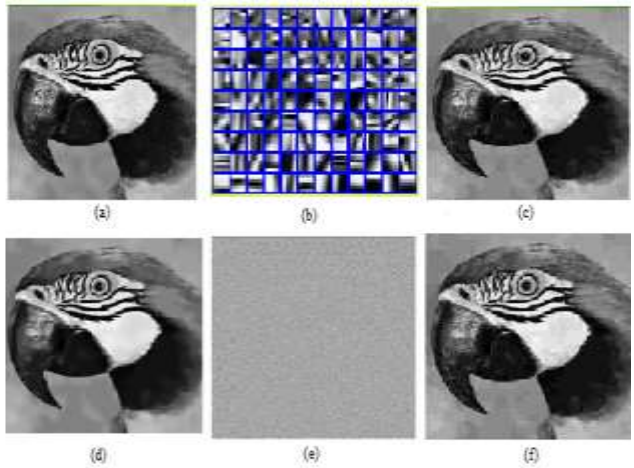


Fig. 12: (a) Original image; Image obtained through (b) Dic method; (c) compressed at q=10; (d) Dic-TV method; (e) dithered plain image appears gray; (f) Anti-Forensic Noise removal method.

The Dic-TV and Anti-forensic noise removal methods are all coded in Matlab and numerical tests are done by Matlab2012a on laptop. By comparing the quality obtained through Dic-TV and Anti-forensic noise removal methods. The PSNR value is 2.818 and SSIM value is 0.0619 better than previous one. The graphical analysis for PSNR and SSIM is also provided in this section for more clarification. The average PSNR and SSIM values were in the table. The images will also provide the quality details.

V. Conclusion

In this paper, the approach is to reduce the artifacts and remove the anti-forensic noise. The dithering method is used along with Dic-TV. The JPEG images are used to demonstrate the improved quality. The high frequency information is also regained after quantization. The PSNR and SSIM values are used to evaluating the quality of compressed image. As a future research, explore how to design an efficient dictionary algorithm to reduce the computational time.

Acknowledgement

The authors would like to thank anonymous reviewers for their constructive comments and valuable suggestions that helped in the improvement of this paper and helped improved the manuscript.

References

[1] Huibin Chang, Michael K. Ng, and Tieyong Zeng, “Reducing artifacts in JPEG decompression via a learned dictionary”, IEEE Transactions on Signal processing, Vol. 62, No. 3, February 1, 2014.
[2] Gregory K. Wallace, “The JPEG still picture compression standard”, IEEE Transactions on Consumer Electronics, vol. 34, no. 4, pp. 30–44, December 1991.
[3] S.Karthik1,Hemanth V K2, K.P. Soman3, V.Balaji4, Sachin Kumar S5, M. Sabarimalai Manikandan6, “Directional total variation filtering based image denoising method”, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012.
[4] Yu-Mei Huang, Lionel Moisan, Michael K. Ng, and Tieyong Zeng, “Multiplicative noise removal via a learned dictionary”, IEEE Transactions on image processing, Vol. 21, No. 11, November 2012.
[5] Antonin Chambolle, and Thomas Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging”, Institute for Computer Graphics and Vision, Graz University of Technology, 8010 Graz, Austria, June 9, 2010.
[6] Matan Protter and Michael Elad, Senior Member, IEEE, “Image sequence denoising via sparse and redundant representations”,IEEE Transactions on image processing, Vol. 18, No. 1, January 2009.
[7] Mei-Yin Shen and C.-C. Jay Kuo, “Review of postprocessing techniques for compression artifact removal”, Journal of visual communication and image representation Vol. 9, No. 1, March, pp. 2–14, 1998.
[8] Matthew C. Stamm, Steven K. Tjoa, W. Sabrina Lin, and K. J. Ray Liu, “Anti-Forensics of JPEG compression”, IEEE Transactions on ICASSP 2010.

