

A NEW PERCEPTUALLY ADAPTIVE METHOD FOR DEBLOCKING AND DERINGING

Quoc Bao DO, Marie LUONG, Azeddine BEGHDAI

L2TI, Université Paris 13, Villetaneuse, France

ABSTRACT

In this paper, a new perceptually adaptive method for reducing the blocking and ringing artifacts encountered in image compression is proposed. The method consists of three steps: (i) blocking-ringing artifacts detection, (ii) perceptual distortion measure and (iii) blocking-ringing artifacts reduction. The performance of the proposed method is evaluated objectively and subjectively in terms of image fidelity and blocking, ringing and blur effects reduction. The obtained results are very promising and confirm once more the efficiency of perceptual approaches in image processing.

1. INTRODUCTION

Although the internet bandwidth is continuously increasing, most communication networks could not support the transmission of a digital film in real time. To cope with this issue, many compression techniques such as JPEG, MPEG, H.264 have been proposed. These techniques bridge an important gap between the user's demands and the limited capacity of transmission networks and storage devices. However, at high compression rate, they could introduce several artifacts such as blocking and ringing effects. Blocking artifact, which is mostly visible in textureless areas, presents itself as artificial vertical and horizontal transitions between adjacent blocks. This is mainly due to the independent quantization of each block. On the other hand, ringing artifact which behaves as spurious oscillations in the vicinity of contours, is the consequence of a high frequency components decimation. The major challenge lies in how to effectively smooth out the artifacts without affecting image details. To this end, several post-processing techniques have been then proposed in the literature such as POCS (projection onto convex sets) [1,2] and Total Variation (TV) [3, 4]. In [5], Nosratinia proposes a new method by reapplying JPEG on shifted versions of the degraded image. The enhanced image is finally obtained by taking the average of the re-compressed images. In [6], the authors propose to filter the artifacts by averaging the similar blocks and to retain the image details by using a projection onto a narrow quantization set. It is worth noticing that the use of some characteristics of the Human Visual System (HVS) in the

design of deblocking-deringing methods could help to better filter out these artifacts [4,6]. In this paper, a perceptually adaptive filtering method is proposed to remove blocking and ringing artifacts encountered in lossy image compression JPEG. It consists of three steps, namely (i) blocking-ringing artifacts detection, (ii) perceptual distortion measure and (iii) blocking-ringing artifacts reduction. In the first step, a new method is proposed to accurately identify the location of the artifacts yielding a Distortion Map (DM). In the second step, a perceptual measure is used to quantify the distortion level which is then used to control the filtering strength in the last step.

The paper is organized as follows: the blocking-ringing artifacts detection is presented in section 2 followed by the perceptual distortion measure in section 3. The blocking-ringing artifacts reduction is then described in section 4. The experimental results are reported in section 5. Section 6 is devoted to the conclusion.

2. BLOCKING-RINGING DETECTION

Note that most existing methods detect blocking artifacts based on the observation that they appear at regular locations, i.e. between 8x8 blocks. However, these methods could consider image contours as blocking artifacts if these contours are accidentally localized at boundaries of 8x8 blocks. A contour detection step should be then performed prior to blocking artifact detection to avoid this issue. On the other hand, while blocking artifact has been extensively studied, very few works have been devoted to ringing artifacts detection. One of the main reasons is that, unlike the blocking artifact, ringing artifact does not appear as oriented and regular patterns. It could appear in any direction near object contours and sharp transitions in the image. Indeed, to detect ringing effect, most existing methods rely on the a priori knowledge, i.e. this artifact appears in the vicinity of the image contours. So that a contour detector should be carried out before starting ringing detection. Based on the above analysis, contour detection is one of the most important steps to better identify the location of blocking and ringing artifacts. However, contour detection is not easy especially in the presence of other components such as texture or ringing artifacts around the contours. In order to cope with this issue, in

the first step, an image decomposition technique [7] is used to decompose the degraded image into two parts, i.e. oscillating image and cartoon image where the latter contains only homogenous regions and well-defined contours (see Fig.1). The Contour Map (CM) can then be easily extracted from the cartoon image by thresholding the gradient norm. The overview of detection scheme is shown in Fig.1. To demonstrate the usefulness of the image decomposition step, the CM which is extracted directly from the degraded image using the same thresholding technique is shown Fig.2c. As can be seen, much textural and ringing information are considered as image contours. The CM is the backbone of the proposed detection scheme. The Ringing Map (RM) is then simply derived from the CM by using a dilation operator. It is considered in small surrounding regions of radius $T_{ringing}$ ($T_{ringing}=6$ pixels in our experiments) around the CM. In order to estimate the Blocking Map (BM), we propose to analyze the variations at boundaries and inside the 8×8 blocks. The detection of vertical blocking effect between (m,n) and $(m,n+1)$ blocks is illustrated in Fig.3 (the coordinates of the upper left pixel in (m,n) block are $(8m,8n)$ – the upper left pixel of the image locates at $(0,0)$). There are mainly four steps in this procedure:

- (1) Estimate the absolute horizontal gradient $D_h(x,y)=|I(x,y)-I(x+1,y)|$ where I is the intensity of the degraded image.
- (2) For each line y , estimate the average $M_L(y)$ and $M_R(y)$ of D_h inside the left and right blocks (see Fig.3)
- (3) For each line y , if the variation at boundary $D_h(8m+7,y)$ is greater than $M_L(y)$ or $M_R(y)$, the counter cpt is incremented by one.
- (4) For each block pair, if $cpt > 6$, we consider that there is a blocking effect between two blocks.

A similar process is carried out to detect horizontal blockiness. Note that the above procedure may consider some image contours as blocking effects. To solve this problem, all pixels belonging to the CM are removed from the BM. Finally, the Distortion Map (DM) is obtained by combining the RM and BM, i.e. $DM = RM \cup BM$. An example of RM, BM and DM is shown in Fig.4.

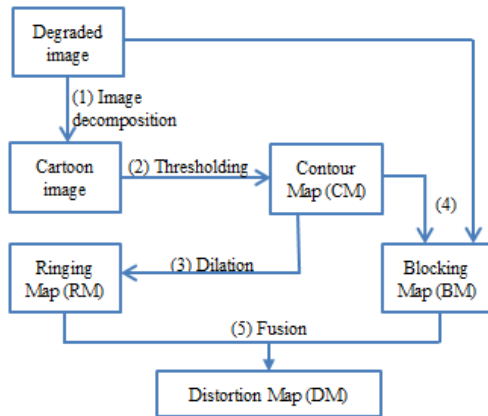


Figure 1. Overview of detection scheme

3. PERCEPTUAL DISTORTION MEASURE

Once the degraded regions have been well identified and localized, the next step is to estimate the distortion level of the artifacts using three factors, i.e. local distortion, texture masking effect and luminance masking effect.

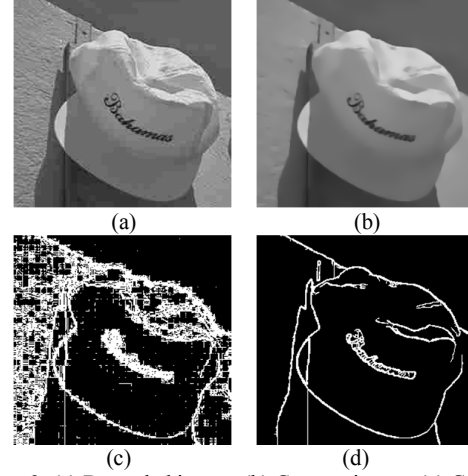


Figure 2: (a) Degraded image, (b) Cartoon image, (c) Contour map extracted directly from the degraded image, (d) Contour map estimated from the cartoon image

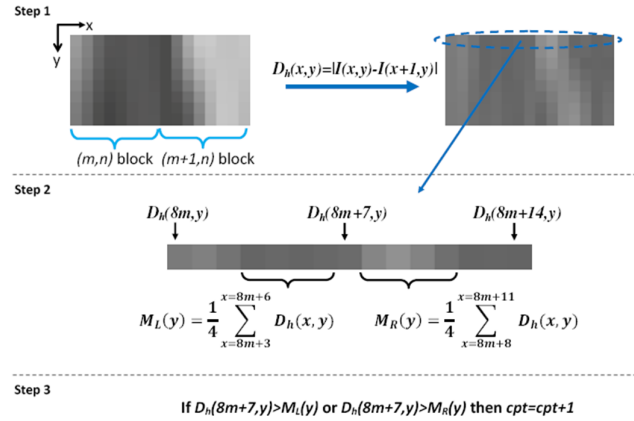


Figure 3: Overview of blocking and ringing artifacts detection

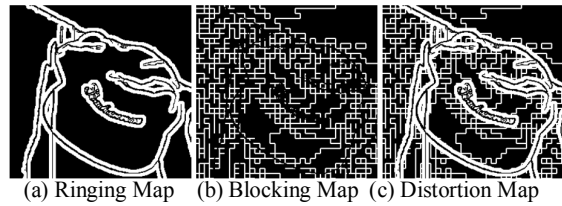


Figure 4: Example of Ringing, Blocking, Distortion Maps

In order to use this estimation to control filtering strength in the next step, we propose to normalize it in the range $[0...1]$ where 0 (1) corresponds to the lowest (highest) level of distortion, respectively. The overview of this step is illustrated in Fig.5. The first factor, Local Distortion (LD), is estimated using the gradient information in the sense that the more variation is, the more distortion is. It is given by:

$$LD(x,y) = \frac{ld(x,y)}{MAXI} \quad (1)$$

where $ld(x,y) = D_h(x,y) + D_v(x,y)$ and $MAXI$ is the maximal value of ld . However, only the LD does not reflect the level of annoyance. As can be seen, in Fig. 8a,d, LD indicates that blocking and ringing artifacts appearing in the wall are very annoying. However, the artifacts in this region are less visible due to masking effects. Based on these observations, some properties of the HVS should be taken into account to estimate the distortion level [8]. In this work, we consider two fundamental characteristics of the HVS called texture masking and luminance masking effects.

For the texture masking effect, it is well known that the visibility of the artifacts depends on the activity of the signal in the local background. The higher activity of local background is, the less visible the artifacts are. In our work, the procedure of modeling texture masking consists of three steps:

1. Identify the local background region for a pixel affected by the artifacts.
2. Estimate the activity of this local background region.
3. Estimate the texture masking effect.

For a pixel situated on a vertical blocking, the local background is composed of two parts: the right half of the left block and the left half of the right block as shown in Fig. 6. For horizontal blocking artifact, the local background is defined in a similar way. In the case of ringing artifact, the determination of the local background for a pixel affected by this artifact is more complex due to the fact that we have to distinguish between two ringing regions on the two sides of the contour. This discrimination is necessary because the visibility of ringing artifact on one side may significantly differ from the other. In many cases, a contour is just a segment then two sides of ringing region determined by dilation operator are merged. To cope with this issue, in the first step, we extract pixels on the outward border of ringing region (see Fig. 7). For each pixel on this border, a window of radius $T_{ringing}$ is considered. Note that since contour region has at least one pixel, this window never reaches the ringing region in the other side of the contour. The region not-affected by ringing artifact is considered as the local background region for all ringing pixel in this window. Note that, each pixel inside the ringing region is covered by several windows resulting in several local background regions. The final local background region noted $\Omega(x,y)$ of each ringing pixel (x,y) is the fusion of all these regions. Once local background has been well determined, the activity ACT of this region is calculated using gradient information. In the case of vertical blockiness, considering two pixels laying on boundary of (m,n) and $(m+1,n)$ blocks, the activity of the surrounding regions is the average of D_h within the blocks. It is expressed as follows:

$$ACT(x,y) = \frac{1}{8} \left(\sum_{i=8m+3}^{8m+6} D_h(x,y) + \sum_{i=8m+8}^{8m+11} D_h(x,y) \right) \quad (2)$$

Similar calculation is done for horizontal blocking artifact. In the case of ringing artifact, the activity $ACT(x,y)$ of surrounding region $\Omega(x,y)$ defined for ringing pixel (x,y) is the average of the gradient norm inside $\Omega(x,y)$. It is expressed as follows:

$$ACT(x,y) = \frac{1}{N} \left(\sum_{(i,j) \in \Omega(x,y)} D_h(x,y) + D_v(x,y) \right) \quad (3)$$

where N is the number of pixel in $\Omega(x,y)$. Finally, the Texture Masking (TM) effect is estimated based on the model proposed in [8] as follows:

$$TM(x,y) = \frac{1}{(1 + ACT(x,y))^5} \quad (4)$$

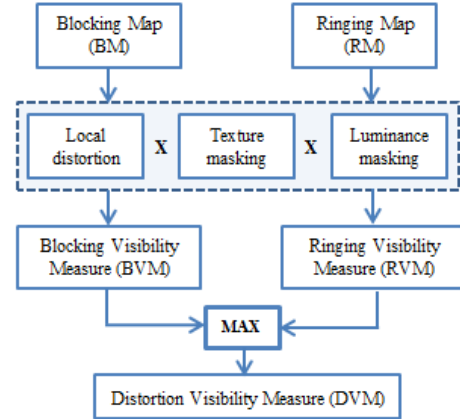


Figure 5: Distortion visibility measure scheme.

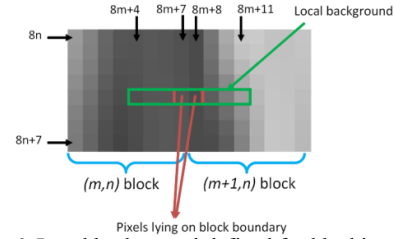


Figure 6: Local background defined for blocking artifact

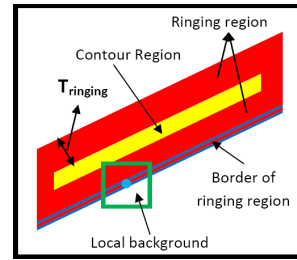


Figure 7: Local background defined for ringing artifact

For Luminance Masking effect (LM), it has been shown that the visibility of the artifacts depends on the local mean luminance. Similar to the above procedure, the modeling of the luminance masking consists of three steps:

1. Identify the local background region for a pixel affected by the artifacts.
2. Estimate the average luminance of this local background region.
3. Estimate the luminance masking effect.

The first step is exactly as described above. In the second step, the average luminance (ALM) of vertical blocking artifact is expressed as follows:

$$ALM(x,y) = \frac{1}{9} \sum_{i=8m+4}^{8m+11} I(i,y) \quad (5)$$

Similar computation is performed for horizontal blockiness. The *ALM* of ringing pixel (x,y) is determined as follows:

$$ALM(x,y) = \frac{1}{N} \sum_{(i,j) \in \Omega(x,y)} I(x,y) \quad (6)$$

The *LM* effect used in this work is inspired from [8]. It is expressed as a nonlinear function which models two important aspects, i.e., (i) an artifact in a dark background is less visible than the one in bright background and (ii) an artifact is most visible in background whose luminance average is in middle range. This effect could be modeled through the function given below:

$$LM(x,y) = \begin{cases} \frac{\sqrt{ALM(x,y)}}{9}, & 0 \leq ALM(x,y) \leq 81 \\ 0.0017(81 - ALM(x,y)) + 1, & \text{otherwise} \end{cases} \quad (7)$$

The Blocking Visibility Measure (*BVM*) and Ringing Visibility Measure (*RVM*) are finally obtained by multiplying the three factors, i.e. the local distortion *LD*, the texture masking effect *TM* and the luminance masking effect *LM*. Since one pixel could be affected by both blocking and ringing artifacts, the Distortion Visibility Measure (*DVM*) of this pixel is the maximum between *BVM* and *RVM*. An example of this estimation is shown in Fig.9. As can be seen, all distortions including blocking and ringing artifacts in high activity region such as the wall on the left of the image are much less visible than these distortions in flat regions.

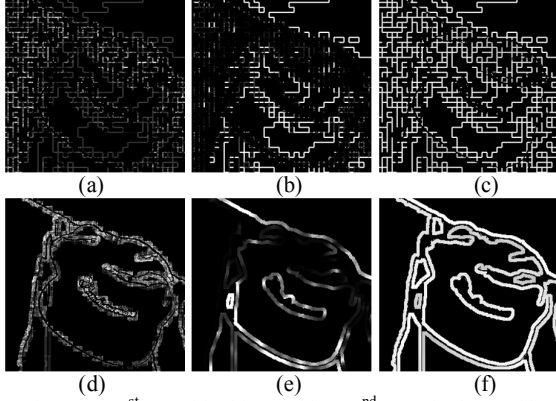


Figure 8: (1st row: blocking artifact, 2nd row ringing artifact), from left to right: Local distortion (LD), texture masking effect (TM), luminance masking effect (LM)

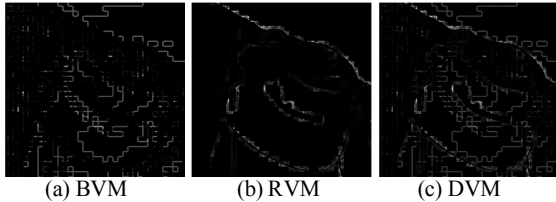


Figure 9: (a) blocking visibility measure, (b) ringing visibility measure, (c) distortion visibility measure

4. BLOCKING-RINGING REDUCTION

Based on the estimated Distortion Visibility Measure *DVM*, a new perceptually adaptive method is proposed to

remove blocking and ringing artifacts. This method consists of two steps: (i) an adaptive filtering step performed in the spatial domain and (ii) a projection onto adaptive quantization set in the DCT domain.

In the first step, we adopt Non-Local Means (NLM) filter [9] which is expressed as follows:

$$\hat{I}(x,y) = \frac{\sum_{i=x-r}^{x+r} \sum_{j=y-r}^{y+r} \exp\left(\frac{-\|G(i,j) - G(x,y)\|^2}{h^2}\right) I(i,j)}{\sum_{i=x-r}^{x+r} \sum_{j=y-r}^{y+r} \exp\left(\frac{-\|G(i,j) - G(x,y)\|^2}{h^2}\right)} \quad (8)$$

where $G(i,j)$ and $G(x,y)$ are two small patches of size 7×7 around the pixel (i,j) and (x,y) , respectively, $\|\cdot\|^2$ is the l^2 norm, the search window r is set to 4 in our work and h is a decay parameter which controls the smoothness level. Note that the initial purpose of NLM is denoising in which the optimal h is set equal to the standard deviation of noise σ . To adapt denoising methods to deblocking purpose, some authors propose to estimate the parameter σ in the image compression domain. This idea has been introduced in [10] by Foi et al. It has been pointed out that the estimation of σ is related to the quantification matrix Q by a nonlinear function. This estimation is expressed as follows:

$$\sigma = \sqrt{0.69 \left(\frac{1}{9} \sum_{u=0}^2 \sum_{v=0}^2 Q(u,v) \right)^{1/3}} \quad (9)$$

In our work, we adopt this estimation and set σ as the optimal decay parameter of NLM method. Moreover, in order to remove blocking and ringing artifacts and preserve more true details, we propose to adjust h adaptively according to each image regions. The tuning of h is guided by the Distortion Visibility Measure *DVM*. This parameter is expressed as follows:

$$h(x,y) = (1 + DVM(x,y))\sigma \quad (10)$$

It means that the more distortion a pixel is affected, the stronger it is filtered.

In the second step, the filtered image \hat{I} is projected onto a narrow quantization constraint set determined in the DCT domain. Let us denote $\hat{F}^{(m,n)}(u,v)$ the DCT coefficient of block (m,n) of the filtered image \hat{I} . The projection is defined as follows:

$$\tilde{F}^{(m,n)}(u,v) = \begin{cases} F_{\max}^{(m,n)}(u,v), & \hat{F}^{(m,n)}(u,v) > F_{\max}^{(m,n)}(u,v) \\ F_{\min}^{(m,n)}(u,v), & \hat{F}^{(m,n)}(u,v) < F_{\min}^{(m,n)}(u,v) \\ \hat{F}^{(m,n)}(u,v), & \text{otherwise} \end{cases} \quad (11)$$

where $F_{\max}^{(m,n)}(u,v)$ and $F_{\min}^{(m,n)}(u,v)$ are determined from the DCT coefficient $F^{(m,n)}(u,v)$ of the degraded I as follows:

$$F_{\max}^{(m,n)}(u,v) = F^{(m,n)}(u,v) + \lambda^{(m,n)} Q(u,v) \quad (12.a)$$

$$F_{\min}^{(m,n)}(u,v) = F^{(m,n)}(u,v) - \lambda^{(m,n)} Q(u,v) \quad (12.b)$$

where λ is a parameter to adjust the extent of the constraint set. The smaller λ is, the more details are retained but at the expense of less blocking/ringing reduction. The larger λ is, the more blocking and ringing artifacts are eliminated but the more details are altered. To adapt to the image content, this parameter should be tuned adaptively to the distortion level in each block. It is given by:

$$\lambda^{(m,n)} = \max(0.5, 0.25 + \lambda_0^{(m,n)}) \quad (13)$$

where

$$\lambda_0^{(m,n)} = \frac{\sum_{x=8m}^{8m+7} \sum_{y=8n}^{8n+7} DVM(x,y)}{\sum_{x=8m}^{8m+7} \sum_{y=8n}^{8n+7} DM(x,y)} \quad (14)$$

It means that if a block is free from distortion, $\lambda^{(m,n)}$ is set equal to 0.25. If a block is heavily affected by the artifacts, it should be strongly filtered using large $\lambda^{(m,n)}$. The restored image \tilde{I} is finally obtained by using IDCT on the projected values \tilde{F} .

5. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed method, several tests are carried out on image database of Kodak [11]. Each image is compressed at 10 levels using Matlab's *imwrite* command with quality parameter $q \in [5, 50]$. Small q corresponds to high compression rate resulting in low image quality. A comparative evaluation using objective and subjective measures has been performed to demonstrate the advantages of our method over the methods proposed in [1-6]. To objectively evaluate the results, we use a fidelity measure SSIM [12] as well as three distortion measures, i.e. Blocking [13], Ringing [14] and Blur [15] measure. In contrast to SSIM metric, the three distortion measures increase with the level of distortion, i.e. high value corresponds to a low level of image quality. A good method must yield high SSIM and is capable to reduce blocking, ringing artifacts (indicated by small value of Blocking and Ringing measures) while avoiding oversmoothing of the images (indicated by small value of Blur measure). To better compare the performance of the methods, we show the improvement, i.e. the difference of objective measures between the restored image and the degraded image. A positive difference of SSIM indicates that the quality of the degraded image has been improved. In the case of the distortion measures, a negative difference means that the method has succeeded in reducing the artifact. Due to limit of space, we only report here the result of typical image "Caps" containing various structures.

The improvements in terms of blocking and ringing reduction are shown in Fig.10c and Fig.10d, respectively. As can be seen, all methods have reduced the blocking and ringing artifacts as expressed by negative difference values. Two methods of Nosratinia and Zakhor yield the best results for blocking and ringing reduction. Kim's method is the worst one in this competition. The proposed method is in the middle ranking. In terms of blur measure (see Fig.10b), the restored images of all the methods are more blur than the degraded ones as expressed through positive difference values. This observation seems to be obvious since all methods are based on filtering principle which usually results in a blurring effect. Note that, two methods of Nosratinia and Zakhor result in the best values in terms of blocking and ringing reduction but at the expense of oversmoothing effect. Kim's method yields less blur but it does not effectively eliminate blocking and ringing effects. The

proposed method is also in the middle ranking in terms of blur measure.

The improvements of SSIM are illustrated in Fig.10a. As can be seen, the proposed method results in the best improvements for all compression levels. To summarize, the proposed method achieves a good compromise between blocking-ringing reduction and oversmoothing effect. On the other hand, it outperforms all other methods in terms of SSIM measure.

For subjective comparison, the results of seven methods are shown in Fig.11. As can be seen, Zakhor's method results in oversmoothed images. Many blocking and ringing artifacts still remain in the case of Kim's method. These artifacts have been noticeably removed in the case of Alter's method but image details have been smoothed out as well (see texture region on the wall). The method in [4] can keep more details but there is still kind of staircase artifact along the edges (for example around the caps). Zhai's method substantially eliminates blocking effects but fails in removing ringing artifacts. As can be seen, ringing artifacts are still present around the letters, the caps. Similar remarks could be drawn for Nosratinia's method. The proposed method yields the best results since all blocking and ringing artifacts have been effectively eliminated while image contrast and details are well preserved. Please use your monitor to view the images and visit <http://www-l2ti.univ-paris13.fr/~do/isspa2012/index.html> to get more results.

6. CONCLUSIONS

In this paper, a new perceptually adaptive filtering method is proposed for blocking and ringing effects reduction. An accurate blocking and ringing artifacts detection has been also proposed. The proposed method has been evaluated on the Kodak base image. The obtained results clearly demonstrate the advantage over some of the state-of-the-art methods in terms of objective and subjective measures. As a future work, we intend to make some tuning parameters data-dependants by introducing an optimization strategy constrained by an appropriate image quality improvement index.

REFERENCES

- [1] A. Zakhor, "Iterative procedure for reduction of blocking effects in transform image coding," IEEE Transaction on Circuits System and Video Technology, pp. 91-94, 1992.
- [2] Y. Kim, C.-S. Park, and S.-J. Ko, "Fast POCS Based Post-Processing Technique for HDTV," IEEE Transaction on Consumer Electronics, pp.1438 - 1447, 2003.
- [3] F. Alter, S.Y. Durand, and J. Froment, "Deblocking DCT-based compressed images with weighted total variation," IEEE ICASSP, vol.3, pp. 17-21, 2004.
- [4] Q.B. Do, A. Beghdadi, M. Luong, "A new adaptive image post-treatment for deblocking and deringing based on Total Variation method," IEEE ISSPA pp.464-467, 2010.
- [5] A. Nosratinia, "Denoising of jpeg images by re-application of jpeg," Journal of VLSI Signal Processing, vol. 27, pp. 69-79, 2001.
- [6] G. Zhai, W. Zhang, X. Yang, and W. Lin, "Efficient Image Deblocking Based on Postfiltering in Shifted Windows," IEEE

Transaction on Circuits System and Video Technology, vol. 18, pp. 122-126, 2008.

[7] L. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms", Physica D, pp 259-268, 1992.

[8] H. Liu and I. Heynderickx, "A perceptually relevant no-reference blockiness metric based on local image characteristics," EURASIP Journal on Advances in Signal Processing 2009.

[9] A. Buades, B. Coll, and J.M. Morel, "A non-local algorithm for image denoising," IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 60-65, 2005.

[10] A. Foi, V. Katkovnik, and K. Egiazarian, Pointwise shape-adaptive det for high-quality denoising and deblocking of grayscale and color images, IEEE Transactions on Image Processing, vol. 16, no. 5, pp. 1395-1411, 2007.

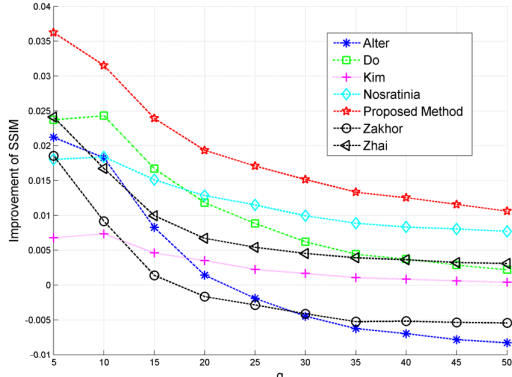
[11] <http://www.r0k.us/graphics/kodak/>

[12] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE ICIP, pp. 600-612, 2004.

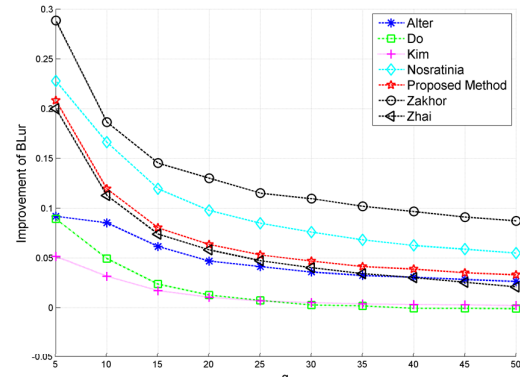
[13] Z. Wang, A.C. Bovik, and B.L. Evans, "Blind measurement of blocking artefacts in images," IEEE ICIP, pp. 981-984, 2000.

[14] M.C.Q. Farias, "No-Reference and Reduced Reference Video Quality Metrics: New Contributions," thesis in University of California, Santa Barbara, 2004.

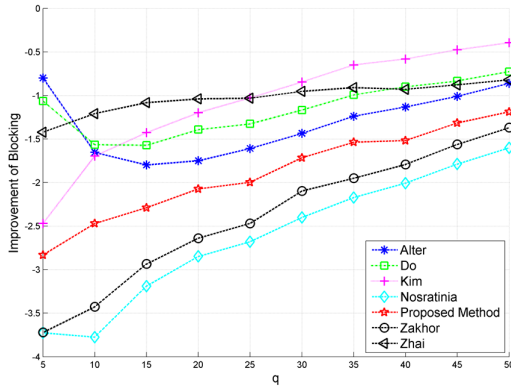
[15] F. Crété-Roffet, T. Dolmiere, P. Ladret, M. Nicolas, "The blur effect: perception and estimation with a new no-reference perceptual blur metric," Proc. SPIE, vol. 6492, 2007.



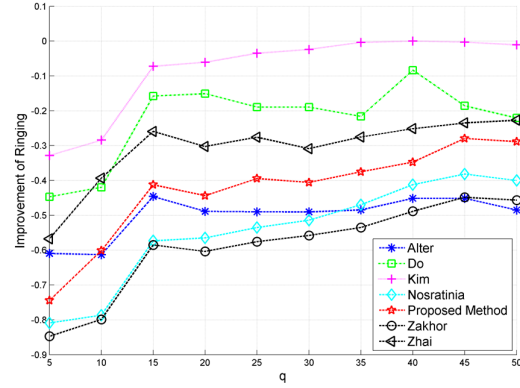
(a) SSIM



(b) Blur



(c) Blocking



(d) Ringing

Figure 10: (left to right, top to bottom) SSIM, Blur, Blocking and Ringing measures

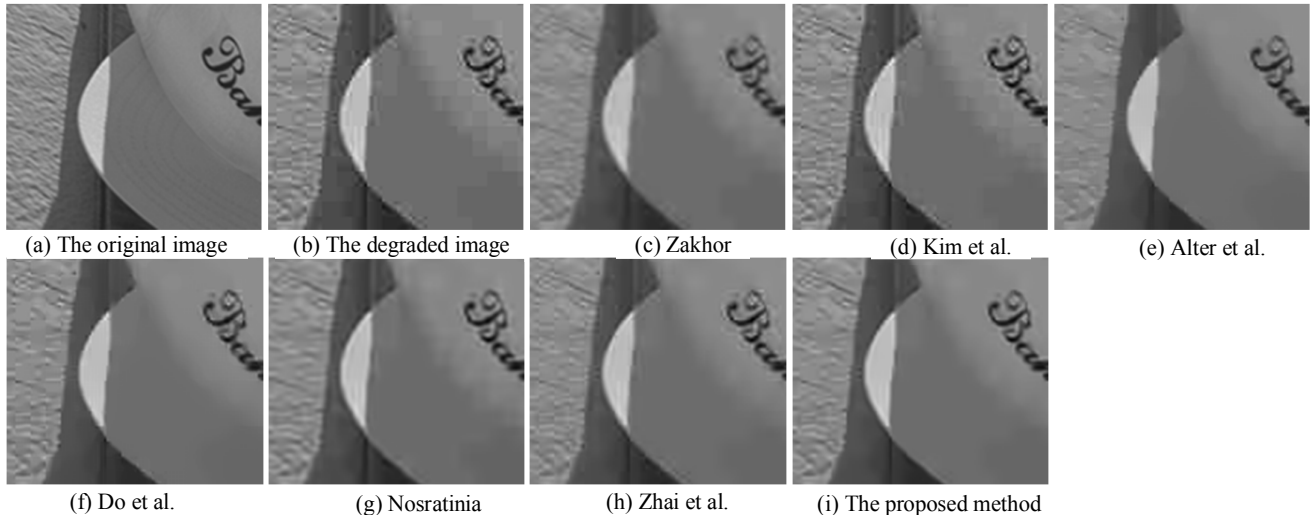


Figure 11: Results of Caps image