Faster Detection of Independent Lossy Compressed Block Errors in Images and Videos

Md. Mehedi Hasan and Oksam Chae

Image Processing Lab, Department of Computer Engineering,
Kyung Hee University, 446-701, Youngin-si,
Gyeonggi-do, Republic of Korea.
mehedi@khu.ac.kr, oschae@khu.ac.kr

Abstract

An effecting compression algorithm removes the redundancy of image signal, when we want to represent high quality videos and images with lower bit rate. Removal of statistical correlation and insignificant components of image signal make the corresponding videos and images highly compressed. This paper represents a new algorithm to measure the blocking artifacts of videos by analyzing the distortions of local properties of image signals like dominant edge magnitude and direction. We also want to incorporate light weighted human vision measurement system like edge information to measure video artifacts in real time. Then simple bucket filling approach is applied, where the particular bucket contains the maximum value also indicating the block boundaries that are passed to the report module. After computing distortion measure a proposed detection approach is used to capture the distorted frames. Extensive experiments on various videos show that the new algorithm is very much efficient and faster to measure the independent lossy compressed block errors in real time video error detection applications.

Keywords: *Block error, Kirsch Mask, Discrete Cosine Transform, Edge Enhancement, Human Visual Sensitivity.*

1. Introduction

Block transform coding is the most popular approach for image and video coding. There are many standards exist in image and video coding, like JPEG, H.263, MPEG-1,MPEG-2, MPEG-4 etc. The BLOCK-based discrete cosine transform (B-DCT) [1, 2, 3] is the fundamental component of many image and video compression standards, used in a wide range of applications. The B-DCT scheme takes into account the local spatial correlation property of the images by dividing the image into 8×8 blocks of pixels. Then transforming each block from the spatial domain to the frequency domain is done by using the discrete cosine transform (DCT) and quantizing the DCT coefficients. In discrete cosine transform (DCT) each block pixels are treated as single entities and coded separately. But the correlation among spatially adjacent blocks is not taking into account in coding and as a result block boundaries are visible when decoded image is reconstructed. As for example, a slight change of luminance in border area can cause a step in the decoded image if the neighbor block fall into different quantization intervals. Therefore, the decompressed image and video exhibits various kind of artifacts. One of the most obtrusive artifacts is the "Blocking Artifact"[1, 2, 4, 5, 6]. In Figure 1, we have shown some highly compressed images with blocking artifacts.

Figure 1. Blocking Artifact Images due to B-DCT Coding

There are many potential methods of measuring discrete cosine transform (DCT) based codec degradations involve directly examining the coarseness of the compressed video stream at the time of quantization scaling. This kind of video streams optionally combined with a measure of the complexity of the original image transmitted by means of outside the video – a form of compressed reference. It is computationally expensive, not accurate and in any case can only make measurements on compressed video that has been already decompressed and potentially passed through other systems, including other additional codecs, prior to end-user delivery.

With the rapid development of the application of video surveillance and broadcast systems, the evaluation of video quality becomes especially significant. Error detection is an important technique to measure the quality of images/videos transmitted over unreliable networks particularly in the wireless channel. At the time of acquisition and transmission, video frames are always distorted by various artifacts. In many video processing applications, accurate knowledge of these errors or distortions present in the input video sequence is very important. In a real system, noises are mainly introduced by the camera and the quantization step of decoding process as showed in Figure 2. But the distortions [7] occur when the videos are [8] transmitted through analog or digital medium. Some errors may be introduced when the analog video signal transmits in wired channel, but in wireless communication it cannot be ignored as occurring frequently.

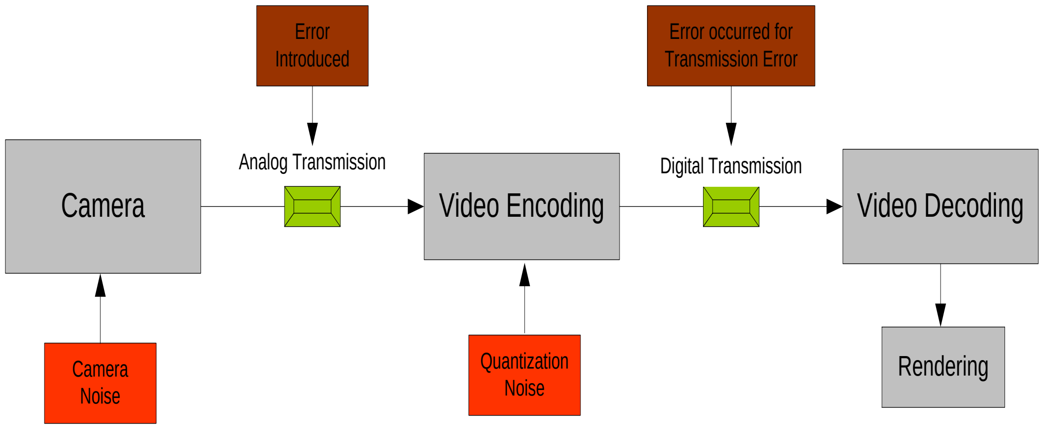


Figure 2. Error and Noise Model for Broadcasting and Surveillance System

Objective image quality matrices are divided into Full-reference, Reduced-reference and No-reference. In Full and Reduced reference approach the access to original image is required. By using the original and reproduced image as inputs the system outputs a numerical value to show the quality of the reproduced image. This approach is not useful in applications like image and video communication, where original image and video is not accessible which is called No-reference approach, which is most experimented approach because of its computational efficiency, wider scope of applications, in-service visual monitoring and post-processing of decoded signal. In image coding techniques, Chou et. al. [8] addressed a key concept of perceptual coding considering human visual system, namely just-noticeable-distortion (JND). JND provides each signal being encoded with threshold level or error visibility. Actually JND is a function of local signal properties, such as background intensity, activity of luminance changes, dominant spatial frequency and changes in edge gradients. Once the JND profile of an image is obtained, the energy of the perceptible distortion like blockiness can be measured. But this kind of HVS [9, 10, 11] measurement system is computationally expensive and cannot be applicable for fast real time cases.

In this paper we actually incorporate light weighted human vision measurement system like edge information [5] to measure blocking artifacts in real time .To detect the distortions in edges we use kirsch mask in eight direction to make the edge detection process more perfect. Then simple bucket filling approach is applied, where the particular bucket contains the maximum value also indicating the block boundaries that are passed to the report module. So, to design an algorithm in this arena, what is desired is a method of monitoring distortions in block-based codecs in a manner that is truly reference-less and it uses minimum amount of hardware acceleration like processing speed to perform such measurements in real time.

2. Proposed Blockiness Metric to Detect Video Artifacts

Most of the compression systems today separate video into blocks of data and separately compress each block using a discrete cosine transform (DCT). Research shows that in video which has been subjected to such kind of compression, image or, picture quality is strongly correlated with the visible blockiness. In our approach we first assume that our video is reference free and so we have to blindly measure the blockiness of videos or images.

2.1. Blind Measurement Approach

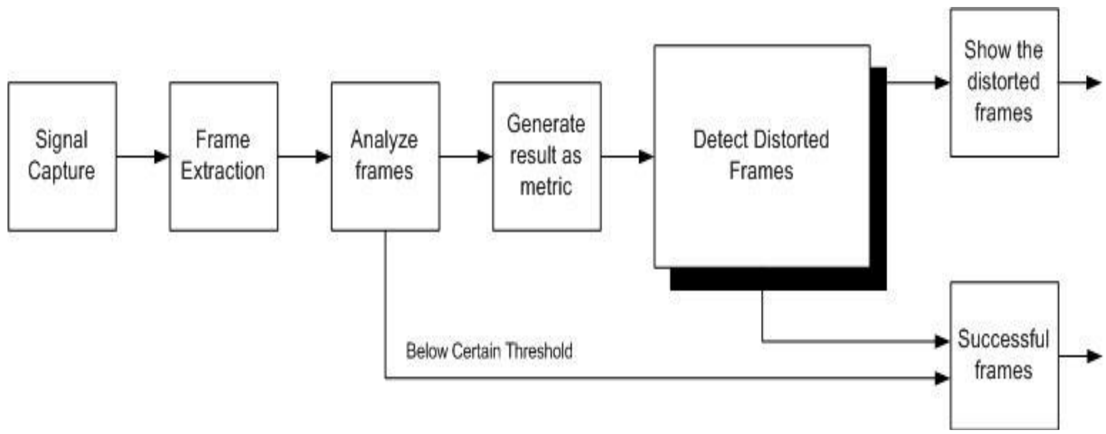


Fig. 3. Overview of the Proposed System

The measurement is shown in Figure 3. Where firstly the video signal that has been compressed is to be measured for picture quality degradation is input to a signal module. Then the fame is captured and converted to component video if necessary. Now the contents of the extracted frames are analyzed to measure the blockiness of the frame and the picture quality is provided to a report results module.

2.2. Edge Enhancement Filtering

The A frame or field of a proposed video signal representing an image or picture is captured and converted, if necessary into luminance and color components. Traditionally, one or more of the components is analyzed by appropriate vertical and horizontal edge enhancement filtering. But if we consider the high gradient information in eight directions and use this information for the distortion measurement, we will gain apparently more accurate results.

So, to consider the different directional information we use Kirsch Masks [12] in eight directions. In Figure 4, the kirsch masks for detecting eight directional edges information are shown. Taking a single mask and rotating it to eight major compass orientations: N, NW, W, SW, S, SE, E, and NE. The edge magnitude is equal to the maximum value found by the convolution of each mask with the image. The edge direction is defined by the mask that produces the maximum magnitude.

$$G_1=\begin{bmatrix}-3&-3&5\\-3&0&5\\-3&-3&5\end{bmatrix}\ G_2=\begin{bmatrix}-3&5&5\\-3&0&5\\-3&-3&-3\end{bmatrix}\ G_3=\begin{bmatrix}5&5&5\\-3&0&-3\\-3&-3&-3\end{bmatrix}\ G_4=\begin{bmatrix}5&5&-3\\5&0&-3\\-3&-3&-3\end{bmatrix}$$

$$G_5=\begin{bmatrix}5&-3&-3\\5&0&-3\\5&-3&-3\end{bmatrix}\ G_6=\begin{bmatrix}-3&-3&-3\\5&0&-3\\5&5&-3\end{bmatrix}\ G_7=\begin{bmatrix}-3&-3&-3\\-3&0&-3\\5&5&5\end{bmatrix}\ G_8=\begin{bmatrix}-3&-3&5\\-3&0&5\\-3&5&5\end{bmatrix}$$

Fig. 4. Eight Directional Kirsch Masks

The $g(x,y)$ across the pixel at (x,y) is determined by calculating the edge changes in eight directions. As shown in Figure 2. Eight operators, $G_k(x,y)$, for k = 1, ……., 8 and i, j = 1,2,3, are employed to perform the calculation.

$$g(x,y)=\max_{k=1,\dots,8}\{|\,grad_k(x,y)\,|\} \tag{1}$$

$$grad_k(x,y)=\sum_{i=1}^{3}\sum_{j=1}^{3}P(x-2+i,y-2+i).G_k(i,j) \tag{2}$$

$$for \quad 0\le x\le H, \quad 0\le y\le W$$

Where $p(x,y)$ denotes the pixel at (x,y). H *and* W are the height and width of the frame consecutively. The resulting edges are correlated with an infinite grid having boundaries corresponding to the block boundaries used in the video compression. Optionally, a second correlation may be made with boundaries slightly different than the block boundaries used in the video compression, with this result being subtracted from the first value. Further the locations of the compression block boundaries may be detected by observing where the maximum correlation value occurs. The resulting correlation results are proposed to generate a picture quality rating for the image which represents the amount of human-perceivable block degradation that has been introduced into the proposed video signal. Figure 5 shows the consequence after using Kirsch masks to detect the edges.

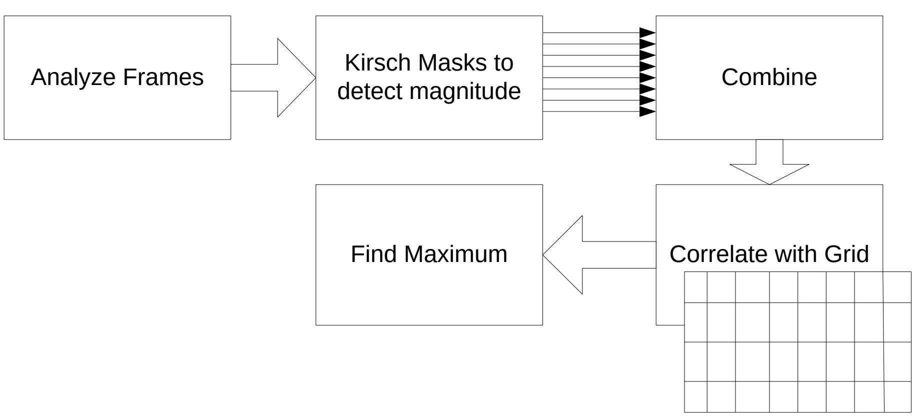


Fig. 5. An Illustration of the Extraction of Blockiness from a Frame

2.3. Generate Blockiness Metric

To generate the blockiness metric for real time systems we have to consider a faster and efficient approach to measure and also detect the location of the blockiness occurred in a frame. Many of the existing algorithms are based on HVS (Human Visual Sensitivity). But these kinds of algorithms are not faster enough to use in real time. So, to use light weighted metric in real time applications we have develop a simpler approach. Figure 6 shows a block diagram view of a block detection apparatus for a picture quality measurement system.

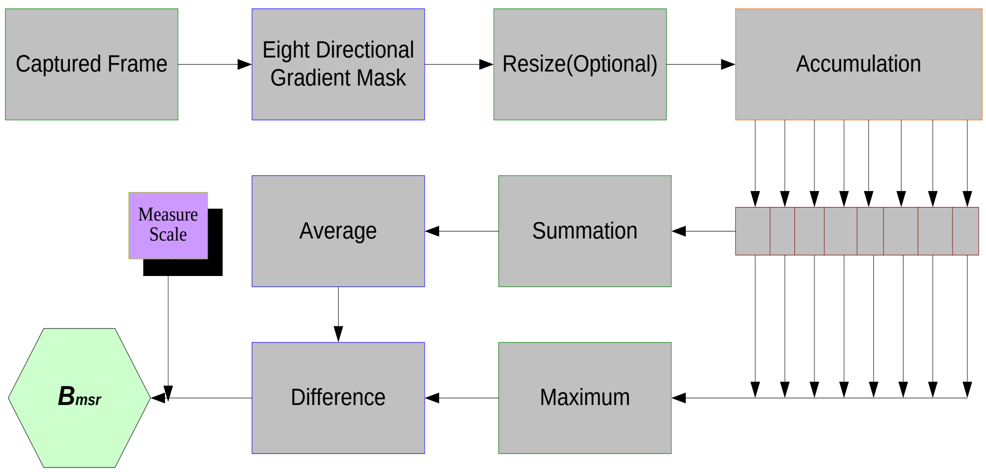


Fig. 6. Schematic Diagram to Measure Blockiness

Here is the Matlab pseudo code is given to show the algorithm to detect blockiness after using the kirsch masks. After taking the maximum magnitude from the masks we will take the absolute value of the maximum edge direction:

PSEUDOCODE:

```
AbsB = abs(B)
CB = AbsB(10:1920 , 10:1080)                //Optional clipping//

% Here is the efficient cross-correlation algorithm:

Buck    = zeroes(1,8);                      //Create buckets and initialize with zero//
ImgR    = length(V(:1));                     //Get picture Height//
ImgC    = length(V(1:));                    //Get Picture Weidth//
ImgC8  = floor(ImgC/8)*8-8;                // Round width down to nearest x8//

for i=1:ImgR
   for j=0:8:ImgC8
          for k=1:8
               Buck(k) = Buck(k) + V(i, j+k);
          end
    end
end
```

Here we, sum the values into eight buckets with each bucket containing the total of all the columns by modulo eight (bucket #1 contains the sum of columns 1, 9, 17……….., bucket #2 contains the sum of columns 2, 10 ,18…….etc.). And finally we achieve the measurement values by the following equations:

$$S = Sum(Buck)/8 \tag{3}$$

$$Block\_msr = (\max(Buck) - S) \times msr\_scl \tag{4}$$

The average values which are acquired from equation (3) are subtracted from the maximum values of the buckets and multiplied by *msr_scl*, a scale factor which is a constant value (4) and which range can be achieved as the Picture Quality Rating (PQR) produced by a device like PQA200. Nevertheless, the bucket which contains the maximum value is also indicates where the compression block boundary is. It is also output to the report module. To improve the algorithm accuracy, measurement can be changed correlating with the block or, macro block spacing and then simultaneous calculation using a non-block correlated kernel size. This significantly reduces false positives in blockiness that a noisy image might otherwise produce. The location of the compressed blocks whether they are human-perceivable or not, may be determined and separately reported.

3. Experimental Result

The JPEG image dataset in the LIVE image quality assessment database release 2 [13, 14, 15] and the MPEG-2 video dataset in the LIVE video quality database [16, 17] are used. The JPEG image dataset includes 29 color reference images (typically 768 × 512 in size) and 204 JPEG distorted images. The LIVE Video Quality Database uses ten uncompressed high-quality videos with a wide variety of content as reference videos. A set of 150 distorted videos were created from these reference videos (15 distorted videos per reference) using four different distortion types-MPEG-2 compression, H.264 compression, simulated transmission of H.264 compressed bit streams through error-prone IP networks and through error-prone wireless networks. Distortion strengths were adjusted manually taking care of ensuring that the different distorted videos were separated by perceptual levels of distortion. Each video in the LIVE Video Quality Database was assessed by 38 human subjects in a single stimulus study with hidden reference removal, where the subjects scored the video quality on a continuous quality scale. Please notice, only the luminance component of each image or video sequence is used for blockiness measurement.

3.1. Experiments on JPEG Images

In the following experiments, we used a number of still images, as well as frames from the test video sequences. These images have different resolutions, ranging from 176×144 to 1920×1080. We also compared our results with those from other objective quality metrics such as PSNR, the quality metrics M_{GBIM} of [5], and the NR quality metrics S of [3]. In order to plot all these metrics in the same figure, we scale PSNR by dividing a factor of 5. According to [5], there is no defined range for the M_{GBIM} and if M_{GBIM} values are greater than one, then blocking effect turns out severe. On the other hand, according to [3], the smaller the S is, the greater the severity of the blocking effect is .Table 1 shows the Pearson Correlation and Spearman rank order Correction between the proposed blockiness measure and the subjective ratings of QCIF video sequences (obtained from subjective video quality experiments similar to that conducted for the evaluation of the JVT sequences [18]). It can be seen that compared to the metrics of [5] and [3], the *Block_msr* of this paper has a better correlation with subjective test results. Table 2 shows the Pearson Correlation and Spearman rank-order Correlation between various quality metrics and the subjective ratings of the JEPG database provided by LIVE [14]. Table 1 and Table 2 show that, our metrics have a comparable correlation with other approaches using subjective data. Additionally, the advantages of our algorithm are that, it is locally adaptive, fast response to blocking artifacts and most of all, it is suitable for real-time implementation. These good technicalities of our algorithm can make it a good choice for practical usage and possibly outweigh the slight drop in correlation values [Table 1, Table 2].

Table 1. Pearson Correlation and Spearman for FUB Database

Algorithm	Pearson Correlation	Spearman Correlation
Block_msr	-.721	.685
M_{GBIM} [5]	-.597	.584
S[3]	.614	.570

Table 2. Pearson Correlation and Spearman for LIVE Database

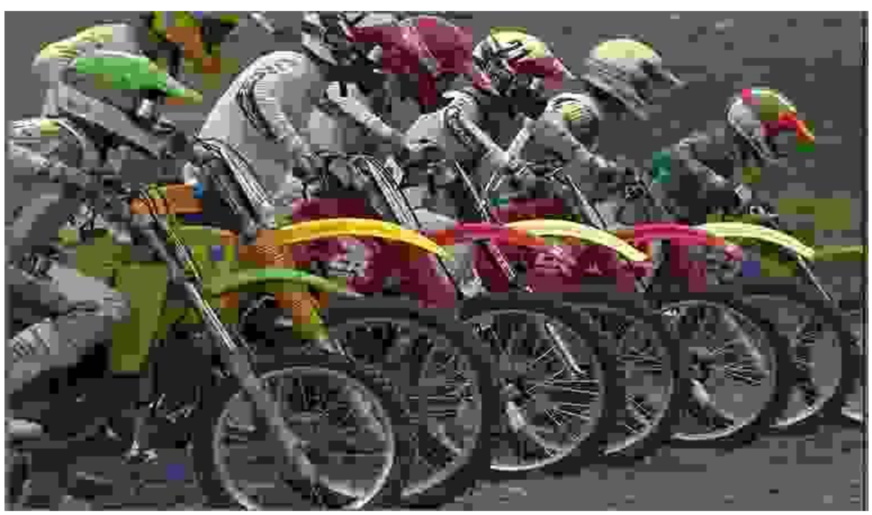
Algorithm	Pearson Correlation	Spearman Correlation
Block_msr	-.843	.838
M_{GBIM} [5]	-.727	.925
S[3]	.944	.937

In Figure 7 we also have shown some results of distortion measurement for different degraded images from LIVE database.

Bike.bmp → *Block_msr* = 1112

Img81.bmp → *Block_msr* = 3048

Img29.bmp → *Block_msr* = 5835



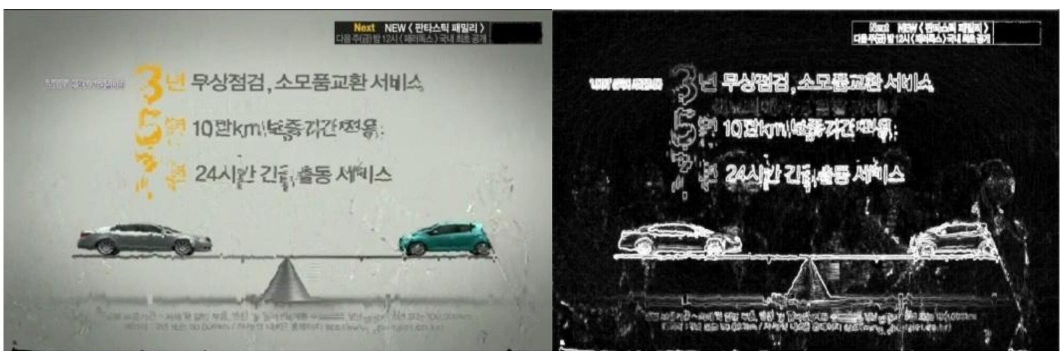
Img188.bmp → *Block_msr* = 8423

Figure 7. Schematic Diagram to Measure Blockiness

3.2. Experiments on MPEG-2 Video Frames

The proposed approach can be applied to a video sequence on a frame-by-frame basis. The blockiness measure for a sequence is defined as the mean value of the blockiness measures over all the video frames in the sequence. Testing results on the MPEG-2 video dataset are given in Table 3. In the first step we are showing experiment results on our own video datasets. For experimental result, we mention different frames of a video, where different frames have blocking artifacts. We can observe the reliability of our algorithm by comparing *Block_msr* values between them.

Frame No: 115 *Block_msr*: 2455



Frame No: 116 *Block_msr*: 7552

Frame No: 430 *Block_msr*: 4485

Frame No: 536 *Block_msr*: 10844

Frame No: 982 *Block_msr*: 3409.34

Frame No: 1045 *Block_msr*: 9804.73

Figure 8. Experimental Result of Picture Quality by using Blockiness Measurement

In Figure 8 Left column is the color frame and right column is the frame by using kirsch masks. By observing the edge enhanced images we can observe that how the edge direction and magnitude is changed by the blockiness artifacts. Experimental results on the same video dataset using Wu and Yuen's [5], Vlachos' [19], Pan et al.'s [20], Perra et al.'s [21], Pan et al.'s [22], and Muijs and Kirenko's [6] are also reported. From Table 3, we can observe that most of these methods give very satisfactory performance while the proposed outperforms the state of the arts.

Table 3. Test Blockiness Result using Different Approaches on the MPEG-2 Video Dataset

Approaches	Pearson Corr.	Spearman Corr.	RMSE
Wu and Yuen's [5]	.6344	.7365	7.1869
Vlachos' [19]	.5378	.7930	7.0183
Pan et al.'s [20]	.6231	.6684	8.4497
Perra et al.'s [21]	.6916	.6531	8.4357
Pan et al.'s [22]	.5008	.6718	8.1979
Muijs & Kirenko's [6]	.7875	.6939	7.9394
Proposed Method	.8627	.7104	7.0236

4. Conclusion and Future Work

The proposed metric involves a grid detection phase, which is used to account for a block size change or grid shift, and intrinsically ensures the subsequent local processing of the blocking artifacts. For each pre-detected blocking artifact the blockiness is individually calculated as there is a signal discontinuity relative to its local content and its visibility because the masking is locally estimated. Combining the results in a simple way yields a metric that shows a promising performance with respect to practical reliability, prediction accuracy, and computational efficiency. Incorporating HVS to blockiness with this method is under implementation to make the algorithm to be more perfect depending on human visual characteristics.

Acknowledgement

This work was supported by National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2011-0003183).

References

[1] S. Liu and A. C. Bovik, Efficient DCT-Domain Blind Measurement and Reduction of Blocking Artifacts. IEEE Trans. on CSVT, vol. 12, no. 12, **(2002)**, pp. 1139-1149.

[2] Z. Wang, A.C. Bovik and B.L. Evan, Blind measurement of blocking artifacts in images. International Conference on Image Processing, **(2000)**,Vancouver, BC, Canada .

[3] Z. Wang, H. R. Sheikh, and A. C. Bovik, No-Reference Perceptual Quality Assessment of JPEG Compressed Images. IEEE International Conference on Image Processing, **(2002)**, pp. 477- 480.

[4] K. T. Tan and M. Ghanbari, Frequency domain measurement of blockiness in MPEG- 2 coded video. International Conference on Image Processing, **(2000)**, Vancouver, BC, Canada.

[5] H. R. Wu, and M. Yuen, A generalized block-edge impairment metric for video coding. IEEE Signal Processing Letters, vol. 4, no. 11, **(1997)**, pp. 317–320.

[6] R. Muijs and I, Kirenko, A no-reference blocking artifact measure for adaptive video processing. European Signal Processing Conference, **(2005)**, Antalya, Turkey.

[7] Y. Wang and Q. Zhu, Error control and concealment for video communication: A review. Proceeding of the IEEE, vol. 86, no. 5, **(1998)** May, pp. 974-997.

[8] B. Kang, A Review on Image and Video processing. International Journal of Multimedia and Ubiquitous Engineering, vol. 2, no. 2, **(2007)** April, pp. 49-63.

[9] C. Chou and L. Li, A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile. IEEE Trans. on CSVT, vol. 5, no. 6, **(1995)**, pp. 467-476.

[10] S. A. Karunasekera and N. G. Kingsbury, A Distortion Measure for Blocking Artifacts in Images Based on Human Visual Sensitivity. IEEE Transactions on Image Processing, vol. 4, no. 6, **(1995)**, pp. 713-724.

[11] H. Liu and I. Heynderickx, A no-reference perceptual blockiness metric. International Conference on Acoustics, Speech, and Signal Processing, **(2008)**, Las Vegas, NV, USA

[12] M. M. Hasan, K. Ahn and O. Chae, Measuring Blockiness of Videos Using Edge Enhancement Filtering. International Conference on Signal Processing, Image Processing and Pattern Recognition, CCIS, vol. 260, **(2011)** December, pp. 10-19.

[13] Z. Yu, H. R. Wu, S. Winkler and T. Chen, Vision-Model-Based Impairment Metric to Evaluate Blocking Artifacts in Digital Video. Proceedings of the IEEE , **(2002)**, pp. 154-169.

[14] H. R. Sheikh, , Z. Wang, L. K. Cormack and A.C. Bovik, LIVE image quality assessment database release 2. **(2005)**. http://live.ece.utexass.edu/research/quality.

[15] H. R. Sheikh, M. F. Sabir, A. C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Transactions on Image Processing, vol. 15, no. 11, **(2006)**, pp. 3440–3451.

[16] K. Seshadrinathan, R. Soundararajan, A.C. Bovik and L. K. Cormack, Study of subjective and objective quality assessment of video. IEEE Transactions on Image Processing, vol. 9, no. 6, **(2010)**, pp. 1427–1441.

[17] K. Seshadrinathan, R. Soundararajan, A.C. Bovik and L. K. Cormack, A subjective study to evaluate video quality assessment algorithms. SPIE Proceedings Human Vision and Electronic Imaging, **(2010)**.

[18] Vittorio Baroncini, ISO/IEC JTC 1/SC29/WG 11, **(2001),** 4240, Sidney.

[19] T. Vlachos, Detection of blocking artifacts in compressed video. IET Electronics Letters, vol. 36, no. 13, **(2000)**, pp. 1106–1108.

[20] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, , S. Yao, Z. Lu and X. Yang, A locally-adaptive algorithm for measuring blocking artifacts in images and videos. International Symposium on Circuits and Systems, **(2004)**, Vancouver, BC, Canada.

[21] C. Perra, F. Massidda and D. D. Giusto, Image blockiness evaluation based on Sobel operator. In: International Conference on Image Processing, **(2005),** Genova, Italy.

[22] F. Pan, X. Lin, S. Rahardja, E.P. Ong and W.S. Lin, Using edge direction information for measuring blocking artifacts of images. Multidimensional Systems and Signal Processing, **(2007)**, Vol. 18, No. 4, pp. 279–308.

[23] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, vol. 13, no. 4, **(2004)**, pp. 600–612.

[24] VQEG: Final report from the video quality experts group on the validation of objective quality metrics for video quality assessment, http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseI.

Authors

Md. Mehedi Hasan

Md. Mehedi Hasan has received his B.Sc. and M.Sc. degree in Computer Science and Engineering from University of Dhaka, Bangladesh in 2009 and 2010, respectively. From 2009 to 2010, he worked as a Lecturer for Dhaka International University, Bangladesh. Now, he is enrolled as a graduate student in the Department of Computer Engineering, Kyung Hee University, Rep. of Korea. His research interest interest include Image processing, noise estimation, Video error detection, intelligent filters, moving object detection, texture classification, face detection, expression recognition and scene recognition, He is a member of IEEE.

Oksam Chae

Oksam Chae received his BS in electronics engineering from Inha University, Rep. of Korea, in 1977. He completed his MS and PhD in electrical and computer engineering from Oklahoma State University, USA, in 1982 and 1986, respectively. From 1986 to 1988, he worked as research engineer for Texas Instruments, USA. Since 1988, he has been working as a professor in the Department of Computer Engineering, Kyung Hee University, Rep. of Korea. His research interests include multimedia data processing environments, intelligent filters, motion estimation, intrusion detection systems, and medical image processing in dentistry. He is a member of IEEE, SPIE, KES, and IEICE.