

The Impact of Spatial Masking in Image Quality Meters

M.T. Qadri¹, M. Ghanbari² and K.T. Tan³

¹ Univerity of Essex

Received: 27 October 2011 Accepted: 22 November 2011 Published: 2 December 2011

Abstract

Compression of digital image and video leads to block-based visible distortions like blockiness. The PSNR quality metric doesn't correlate well with the subjective metric as it doesn't take into consideration the impact of human visual system. In this work, we study the impact of human visual system in masking the coding distortions and its effect on the accuracy of the quality meter. We have chosen blockiness which is the most common coding distortion in DCTbased JPEG or intracoded video. We have studied the role of spatial masking by applying different masking techniques on full, reduced and no reference meters. As the visibility of distortion is content dependent, the distortion needs to be masked according to the spatial activity of the image. The results show that the complexity of spatial masking may be reduced by using the reference information efficiently. For full and reduced reference meters the spatial masking hasn't much importance, if the blockiness detection is accurate, while for the no reference meter spatial masking is required to compensate the absence of any required reference information.

Index terms— Image Quality Assessment, Spatial Masking, Blockiness Measurement, Frequency Domian Analysis.

1 INTRODUCTION

he coding of images and video introduces artefacts like blockiness, blurriness, ringing etc. Blockiness is the most common artefact in block based compression techniques. For fixed size DCT blocks, in high compressions, luminance changes do appear as regular patterns at the DCT block boundaries. Many researchers [1][2][3][4][5][6][7][8][9][10][11] have designed objective quality meters which use engineering approaches to determine the image quality. They are divided into full reference (FR), reduced reference (RR) and no reference (NR) meters. Full reference requires the entire reference image to be available at user end while reduced reference requires some features of the reference image. No reference meters don't require any information of reference image to be available at user side.

In literature various kinds of spatial maskings are used by the researchers ranging from very simple tools to the very complex psychological models. Many of them have studied the non-linear behavior of human visual system and its sensitivity with frequency variations. The aim of this paper is to investigate how important is the sophistication of a spatial masking in the accuracy of a quality meter and whether it is or is not equally effective for all types of the reference models. Our experiments show that if the blockiness detection is accurate then by using the reference information the role of spatial masking is reduced. However, in no reference mode the spatial masking is required to compensate the absence of any reference information.

For the full and reduced reference meters, different masking techniques are applied which improve the quality of meter but improvement is not much significant in weight if compared with the complexity and processing time consumed for masking.

The simplest masking includes edge cancellation between reference and coded images but it is only applicable for FR and RR modes only. Edge cancellation process helps to distinguish between natural edges of the image

44 and sharp edges due to blockiness artifact. The edge cancelled image which contains edges only due to blockiness
 45 is then processed for frequency domain analysis to estimate the distortion.

46 For NR meter, more complex masking technique is required to mimic the behavior of human visual system.
 47 Edge cancellation process is not possible in no reference mode therefore a more reliable blockiness detector is
 48 required to discriminate natural edges from artificial edges appeared due to blockiness. Fortunately we have shown
 49 that through Fourier analysis one can easily separate the periodic pattern resulting from the regular blockiness
 50 edges from the irregular natural edges [12]. For the effective spatial masking, one can use Fiorentini and Zoli [13]
 51 masking technique after the edge detection process to mask the distortion before quantifying the blockiness in
 52 frequency domain. They calculate the masking function for each pixel by using the adjacent pixel values in order
 53 to compensate the distortion locally to apply the property of human visual system. The perception is that the
 54 humans observe less distortion in detailed areas because the distortion is masked by the details present in adjacent
 55 pixels. This technique of course has no masking effect on the regular edges resulting from the blockiness. This
 56 masking technique can be very useful in the no reference meter but it requires more processing time as discussed
 57 in section V.

58 database [14]. There are 233 images in this database, where they are DCT coded and compressed at various
 59 quality. There is a mean opinion score (MOS) assigned to each compressed image, where we can test the accuracy
 60 of our meters against these scores.

61 The rest of the paper is organized as follows: Section II gives an overview of blockiness measurement in
 62 frequency domain and the details of blockiness detection for each model is given in its own section. Section III
 63 explains blockiness detection and the role of spatial masking in Full Reference model. These for the Reduced
 64 Reference model are given in section IV and those for No Reference model are given in section V. Finally the
 65 concluding remarks are given in section VI.

66 2 II.

67 3 OVERVIEW OF BLOCKINESS MEASUREMENT

68 Since, the Blockiness artifact is due to the discontinuities in pixel intensity across DCT block boundaries, for a
 69 fixed DCT block size i.e. 8×8 , this discontinuity appears as a periodic pattern. Transforming the pattern into
 70 a frequency domain, it will appear as harmonics at certain frequencies. The periodicity of the pattern makes it
 71 easy to discriminate blockiness edges from the natural ones. The strength of these harmonics are proportional to
 72 the amount of blockiness in the coded picture. The sum of energy concentrated in these harmonics is therefore
 73 used for blockiness estimation [12]. Figure below illustrates the concept of blockiness measurement through
 74 harmonic analysis in the frequency domain. . Fourier transform of this waveform generates harmonics at the
 75 spatial frequencies of multiple of $8/32$. For real images, figure 1(c) shows the gradient of a real blocky image
 76 at block dimensions of 8×8 pixels. The two dimensional harmonics of the image with a 2D Fourier transform is
 77 shown in figure 1(d). Naturally if the image didn't have such a regular pattern, then the image energy in these
 78 harmonics will be nonexistent. Comparing the image energy at these harmonics with those of whole block energy
 79 at the ac coefficients can give an indication of blockiness [12].

80 4 III. FULL REFERENCE BLOCKINESS METER

81 In full reference meter, the complete reference image is used with the coded image to determine the quality.
 82 Initially the distortion meter is designed using the frequency domain approach without using any spatial masking
 83 technique. Section A of this part discusses the blockiness estimation in full reference mode using harmonics
 84 analysis in frequency domain and in section B the impact of different spatial masking is discovered to improve
 85 the quality of distortion meter.

86 5 a) Overview of Blockiness Estimation in FR Mode

87 The block diagram of the simplest full reference meter without using any spatial masking is given in figure below.
 88 Due to the availability of complete reference image at user end, the first step is to apply the edge detection
 89 process on both reference and coded images. It is used to determine the sharp luminance edges which might
 90 be because of the details present in the image or because of the blockiness artifact, as blockiness results in
 91 sharp luminance discontinuities at DCT block boundaries. The edge detection process is performed in horizontal
 92 and vertical directions on both images separately. The chances of misreading textual details to be treated as
 93 blockiness artifact are reduced by using gradient versions of both images.

94 Since, the blockiness distortion is not likely to be the same in all parts of the image, therefore both of the
 95 resulted edge detected gradient images are then divided into blocks in spatial domain to determine the blockiness
 96 locally for each block. The block size must be any multiple of 8×8 (DCT block size). Here 32×32 block size is
 97 selected to keep the appropriate distance between harmonics. After block processing, each block is transformed
 98 into frequency domain and harmonics are calculated for each block for gradient images. Each block of 32×32
 99 will have a DC coefficient and two repetitive groups of 15 ac components in horizontal and vertical directions
 100 denoted by H1-H15 and V1-V15 respectively. Of these ac components, the particular frequency components are

101 extracted which have same spatial frequency as the blockiness pattern (i.e. 8 pixels/cycle) and its multiples,
102 called harmonics. These harmonics can be determined using the following equation.

103 (1)

104 where, is the frequency of harmonics in cycles per window (cpw), and w is the width of the FFT window. For
105 32×32 FFT window, the interested harmonic frequencies are 4 cpw, 8 cpw, and 12 cpw, which are in fact the
106 H4, H8 and H12 for the horizontal direction, and V4, V8 and V12 for the vertical direction. The equations for
107 the horizontal and vertical blockiness parameter are given below:

108 (2) and

109 (3)

110 The above ratio is calculated for each 32×32 pixels block of reference and coded images and then compared
111 to determine the amount of blockiness locally for every block of the coded image. The blockier an image is, the
112 higher will be the harmonics to AC components ratio. Finally, the blockiness of each block is accumulated at the
113 end for a single quality metric.

114 The objective results are compared with the mean opinion scores of images from LIVE image database. The
115 Pearson's correlation coefficient of 95.75% is obtained by using the above full reference frequency domain analysis
116 without using any complex spatial masking technique.

117 6 b) Effect of Spatial Masking in Full Reference Mode

118 The spatial masking helps to compensate the distortion according to the local spatial activity of the image. In
119 this section, different spatial maskings techniques are applied with the above full reference meter to compare its
120 effect and highlight the importance of spatial masking in quality estimation.

121 The simplest spatial masking is implemented by cancelling the textual edges between gradient versions of
122 reference and coded images. For blockiness estimation, the detection of sharp luminance edges at DCT block
123 boundaries are needed whereas simply edge detection process may treat natural edges and textual details as
124 blockiness distortion. To avoid misinterpreting textual edges as blockiness, Edge Cancellation process is included
125 which cancels the natural edges and leaves only edges because of blockiness artifact. It improves the efficiency
126 of distortion meter.

127 The edge cancelled image is obtained by taking the difference of reference and coded gradient images. To
128 determine the blockiness locally, the edge cancelled gradient image is divided into blocks and then each block
129 is transformed into frequency domain. If the coded image is affected by blockiness then the edge cancelled
130 gradient image must have harmonics at spatial frequencies and by comparing the harmonics with other ac
131 coefficients, the blockiness can be estimated accurately. The Pearson's correlation coefficient of 96.12% is obtained
132 by including the edge cancellation process which improves the above results by 0.3%. The block diagram of the
133 full reference meter with edge cancellation process is given below. As the human visual system is non-linear and
134 sensitive to different frequency variations this means that users observe different amount of distortion depending
135 upon the adjacent details. The distortion will be masked by the nearby details and be less visible to human
136 observers. Applying more sophisticated masking technique includes the features of human visual. This property
137 is implemented by using the masking technique by Fiorentini and Zoli [13]. For this purpose, the local masking
138 function is calculated for each pixel using the adjacent pixel values to implement the property of human visual
139 system. For each pixel in gradient reference image, the adjacent horizontal and vertical pixels (5 adjacent pixels)
140 are compared and the minimum value is chosen which represents the maximum masking effect for that pixel.
141 The process is applied for each pixel of gradient reference image and the final masking function is determined
142 which contains the spatial activity of the reference image.

143 The masking function is then convolved with the edge cancelled coded gradient image to mask the distortion.
144 Then the block processing is applied on masked coded image and each block is transformed into frequency domain
145 analysis. For each block, the harmonics to other ac coefficients ratio is calculated to determine the blockiness
146 distortion locally. Finally, the distortion of each block is accumulated in the end for single value. The Pearson's
147 correlation coefficient is improved to 96.25% by using this masking technique. The block diagram including above
148 masking technique is given below. To conclude the discussion for full reference distortion meter, the table below
149 summarizes the effect of spatial masking on full reference meter.

150 Table ?? : Effect of different maskings on FR blockiness meter Due to the availability of the complete reference
151 image at user end in full reference mode, the spatial masking has not much importance in full reference mode.
152 However, it helps to improve the accuracy of the meter but it increases the complexity and processing time of
153 the system. Since masking also increases some processing burden to the quality meter, the table also shows how
154 much this is increased due to the calculation of masking function.

155 7 IV. REDUCED REFERENCE BLOCKINESS METER

156 The availability of some reference features at user end in reduced reference mode helps to reduce the complexity
157 of spatial masking. Part A of this section discusses the blockiness estimation in RR mode without using any
158 spatial masking technique while part B examines the role of spatial masking in reduced reference meter.

8 a) Overview of Blockiness Estimation in RR Mode

Since, blockiness has periodic patterns which are represented by harmonics in frequency domain, therefore the strength of ac coefficients at harmonics frequency i.e. 4th, 8th and 12th ac coefficient of each block in frequency domain are used as reduced reference parameter. The block diagram of the simplest reduced reference blockiness meter without using any spatial masking is given below.

9 Calculation of harmonics

For RR meter, a similar approach to FR is used except that, instead of using complete reference image, some of its features are used which reduces the quality of RR meter. The procedures include edge detection, block processing (of 32x32 pixels in spatial domain) and then frequency domain analysis of each block to determine the strength of harmonics of reference image. For each block, vertical and horizontal harmonics of the reference and coded images are compared to calculate the amount of distortion locally. The equations of vertical and horizontal harmonics of reference and coded images are given below.

(4) and (5) Finally, the average of blockiness distortion is calculated at the end for a single quality metric. The RR meter is tested on LIVE image database and the Pearson's correlation coefficient of 88.2% is obtained with no spatial masking used.

10 b) Effect of Spatial Masking in Reduced Reference Mode

In the above RR distortion meter, natural edges in the reference picture located at DCT block boundaries may be treated as blockiness which may be mistakenly considered as harmonics in frequency domain. To improve the quality of RR meter, the sharp edges at DCT block boundaries must be distinguished with edges due to blockiness artifact. For this purpose, edge cancellation process is applied to cancel natural edges of reference image. However, edge cancellation in RR mode is not as effective as in FR mode because it only cancels edges at block boundaries and leaves the remaining natural edges within the block. For edge cancellation in RR mode, the check is applied on each block to determine that whether the corresponding block is affected by blockiness or contains natural edges. Due to the fact that the blockiness increases the harmonics amplitude therefore if the strength of harmonics of coded image is greater than the harmonics of reference image i.e. then consider that block as blocky and determine the amount of blockiness by the difference in harmonics amplitudes. Using the edge cancellation process the correlation coefficient is improved by 5.39% which becomes as 93.59%.

The quality of distortion meter may be improved by including the spatial masking of Fiorentini and Zoli [13] as performed in FR mode. This technique is discussed in detail in section III. Here the masking is performed separately on reference and coded images in spatial domain before block processing. The harmonics of masked reference image are used as reduced reference parameter and compared with the harmonics of masked coded image to determine the amount of blockiness in coded image. The block diagram of the meter after including Fiorentini and Zoli is given below. By using this masking technique the correlation coefficient is increased to 94.89% but it increases the complexity and processing time of the system. The table below summarizes the effect of spatial masking on reduced reference meter. V.

11 NO REFERENCE BLOCKINESS METER

The spatial masking has more importance in no reference mode to compensate the absence of any reference information. The basic idea of blockiness estimation is the same as used for FR and RR above. Part A explains the method for measuring blockiness distortion in no reference mode and in part B, the role of spatial masking is discussed using different masking techniques.

12 a) Overview of Blockiness Estimation in NR Mode

There is only coded image available to determine the quality therefore all processes which includes edge detection, block processing and frequency domain analysis must be applied on the coded image itself. The block diagram of the NR meter without using any masking is given below. After the edge detection and block processing in spatial domain, the blockiness is estimated in frequency domain by comparing the amplitude of harmonics to other ac coefficients as blockiness results in harmonics in frequency domain due to their periodic patterns at DCT block boundaries. Due to absence of any reference, the accuracy of NR meter is as good as FR and RR approaches because there are chances of interpreting natural edges as blockiness as their cancellation is not possible without any reference information. The Pearson's correlation coefficient of 83.72% is achieved for no reference blockiness meter without using any masking.

13 b) Effect of Spatial Masking in No Reference Mode

The absence of reference information reduces the quality of NR meter as there is no other way to distinguish between natural and blockiness edges except the distortion must be masked before frequency domain analysis. For this purpose, the spatial masking from Fiorentini and Zoli [13], discussed in section III, is applied on edge detected version of the coded image itself to mask the distortion according to the local spatial activity. Here, the same coded image is used to determine the spatial activity because the overall spatial activity of the coded image

215 would roughly be same as of reference picture. The rest of the procedures are same as used for above techniques.
216 The block diagram of the NR meter with spatial masking is given below. The correlation coefficient of 90.30%
217 is obtained by including the spatial masking from Fiorentini and Zoli. The table below summarizes the effect
218 of spatial masking on no reference meter. The above results show that the spatial masking is essential in NR
219 mode as natural edges of the image may be treated as blockiness edges because of the absence of any reference
220 information.

221 14 VI.

222 15 CONCLUSION

223 The aim of the work is to determine the importance of spatial masking in FR, RR and NR modes. Different
224 masking techniques, from simple to complex, are applied on all three types of meters. The algorithm is tested
225 on LIVE image database of blocky images compressed at different compression rates. The blockiness detection is
226 possible using the gradient versions of reference and coded images. However, in RR mode, in table 2, the features
227 of reference image can be used to distinguish natural and blockiness edges to improve the quality of distortion
228 meter. Simple cancellation of natural edges at DCT block boundaries helps to develop a reliable distortion meter
229 and there is no need for any sophisticated spatial masking.

230 Finally, in NR mode, as in table 3, complex masking techniques are required to compensate the absence of
231 any required information, this is because some part of natural edges may be miscalculated as blockiness and the
232 masking will reduce this miscalculation. However, the accuracy is not as good as for FR and RR meters. It shows
233 that the spatial masking is necessary to compensate the reference information but if the reference information is
available, the role of spatial masking is not vital. ^{1 2}



Figure 1: Fig 1 :

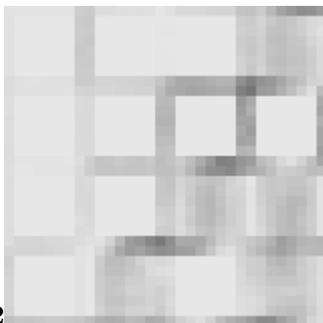


Figure 2: Fig 2 :

234

¹© 2011 Global Journals Inc. (US) Global Journal of Computer Science and Technology Volume XI Issue XX
Version I

²DecemberThe Impact of Spatial Masking in Image Quality Meters database consists of 233 images with their

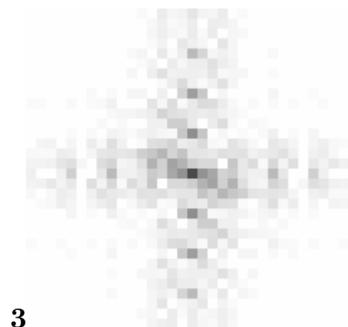


Figure 3: Fig 3 :

mean opinion scores. The results show that the complexity of spatial masking can be reduced by using the reference information. For FR meter, in table 1, due to availability of complete reference image, there is virtually no need of using any masking techniques. The accurate

| The Impact of Spatial Masking in Image Quality Meters | | | |
|---|---|----------------------------|--|
| | | | FR Blockiness Meter Processing Time |
| | | | Without Any Masking 1 Unit |
| | | | Edge Cancellation 1.33 Units |
| | | | Fiorentini and Zoli Masking 4.18 Units |
| 2011 December 72 | | | |
| | (Receiver End) | | |
| Reference Picture | Coded Picture | | |
| $YR(x,y)$ | | $YC(x,y)$ | |
| Edge Detection | Edge Detection | | |
| | $GC(x,y)$ | | |
| | $-GC(x,y)$ | | |
| Spatial Masking | G M G ec | (Transmitter End) | (Receiver End) Coded Picture |
| | | Reference Picture | Picture |
| | | $YR(x,y)$ | $YC(x,y)$ |
| | Edge Cancelled Masked Gradient Image | Edge Detection | Edge Detection |
| | | $GR(x,y)$ | $GC(x,y)$ |
| | Block Processing | Frequency Block Processing | Block Processing |
| | | Do-main | |
| | | Anal-ysis | |
| | | for | |
| | | FR | |
| | | Meter | |
| | Fourier Analysis | Fourier analysis | Fourier analysis |
| | | | sis |
| | Calculation of harmonics to other ac coefficients ratio | Calculation of harmonics | |
| | | | |
| | Full Reference Blockiness Index | | RR Blockiness Index |

Figure 4:

-
- 235 [Wu and Yuen (1997)] ‘A generalized block-edge impairment metric for video coding’. H R Wu , M Yuen . *IEEE*
236 *Signal Processing Letters* Nov. 1997. 4 p. .
- 237 [Wu and Yuen (1997)] ‘A generalized block-edge impairment metric for video coding’. H R Wu , M Yuen . *IEEE*
238 *Signal Processing Letters* Nov. 1997. 4 p. .
- 239 [Barland and Saadane (2005)] ‘A new reference free approach for the quality assessment of mpeg coded videos’.
240 R Barland , A Saadane . *7th Int. Conf. Advanced Concepts for Intelligent Vision Systems*, Sep. 2005. 3708
241 p. .
- 242 [Zhu and Milanfar (2009)] *A no-reference sharpness metric sensitive to blur and noise*, P Zhu , Milanfar . 2009.
243 July 29-31, 2009. San Diego, California, U.S.A.
- 244 [Meesters and Martens ()] ‘A single-ended blockiness measure for JPEG-coded images’. L Meesters , J B Martens
245 . *Signal Process.* 2002. 82 p. .
- 246 [Seshadrinathan et al. (2010)] ‘A Subjective Study to Evaluate Video Quality Assessment Algorithms’. K
247 Seshadrinathan , R Soundararajan , A C Bovik , L K Cormack . *SPIE Proceedings Human Vision and*
248 *Electronic Imaging*, Jan. 2010.
- 249 [Webster et al. (1993)] ‘An objective video quality assessment system based on human perception’. A A Webster
250 , C T Jones , M H Pinson , S D Voran , S Wolf . *SPIE Human Vision, Visual Processing, and Digital Display*
251 *IV*, (San Jose, CA) February 1993. p. .
- 252 [Wang et al. (2000)] ‘Blind measurement of blocking artefacts in images’. Z Wang , A C Bovik , B L Evans .
253 *IEEE Int. Conf. Image Processing*, September 2000. IEEE. 3 p. .
- 254 [Fiorentini and Zoli ()] ‘Detection of a Target Superimposed to a Step Pattern of Illumination. II. Effects of a
255 Just-Perceptible Illumination Step’. A Fiorentini , M T Zoli . *Atti. Fond. G. Ronchi* 1967. 22 p. .
- 256 [Tan and Ghanbari] ‘Frequency domain measurement of blockiness in MPEG-2 coded video’. K T Tan , M
257 Ghanbari . *Proceedings of ICIP 2000*, (ICIP 2000) 3 p. .
- 258 [Live website for subjective scores MOS] *Live website for subjective scores MOS*, [http://live.ece.utexas.
259 edu/research/quality/](http://live.ece.utexas.edu/research/quality/)
- 260 [Sazzad et al. (2008)] ‘No-reference image quality assessment for jpeg2000 based on spatial features’. Z M Sazzad
261 , Y Kawayoke , Y Horita . *Signal Processing: Image Communication*, April 2008. 23 p. .
- 262 [Choi et al. ()] ‘No-Reference Image Quality Assessment using Blur and Noise’. M G Choi , J H Jung , J W Jeon
263 . *International Journal of Computer Science and Engineering* 2 2009. 3 p. .
- 264 [Seshadrinathan et al. (2010)] ‘Study of Subjective and Objective Quality Assessment of Video’. K Seshadri-
265 nathan , R Soundararajan , A C Bovik , L K Cormack . *IEEE Transactions on Image Processing* June 2010.
266 19 (6) p. .