Global Journal of Computer Science and Technology
Volume 11 Issue 20 Version 1.0 December 2011
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

The Impact of Spatial Masking in Image Quality Meters

By M.T. Qadri, K.T. Tan, M. Ghanbari
University of Essex, UK

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 DCT-based JPEG or intra- coded 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.

Keywords : Image Quality Assessment, Spatial Masking, Blockiness Measurement, Frequency Domian Analysis.

GJCST Classification : I.4.8, I.4.7

THE IMPACT OF SPATIAL MASKING IN IMAGE QUALITY METERS

Strictly as per the compliance and regulations of:

The Impact of Spatial Masking in Image Quality Meters

M.T. Qadri[α], K.T. Tan[Ω], M. Ghanbari[β]

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 DCT-based JPEG or intra- coded 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.

Keywords : Image Quality Assessment, Spatial Masking, Blockiness Measurement, Frequency Domian Analysis.

I. Introduction

The 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 [1-11] 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 and sharp edges due to blockiness artifact. The edge cancelled image which contains edges only due to blockiness is then processed for frequency domain analysis to estimate the distortion.

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

We have tested various masking functions with the three FR, RR and NR models on LIVE image

Author [α β] : School of Computer Science and Electronic Engineering, Fellow IEEE, University of Essex, UK.
E-mail : mtqadr@essex.ac.uk, ghan@essex.ac.uk
Author [Ω] : School of Electrical and Electronic Engineering Singapore Polytechnic, Singapore. E-mail : kttan@sp.edu.sg

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

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

II. Overview of Blockiness Measurement

Since, the Blockiness artifact is due to the discontinuities in pixel intensity across DCT block boundaries, for a fixed DCT block size i.e. 8x8, this discontinuity appears as a periodic pattern. Transforming the pattern into a frequency domain, it will appear as harmonics at certain frequencies. The periodicity of the pattern makes it easy to discriminate blockiness edges from the natural ones. The strength of these harmonics are proportional to the amount of blockiness in the coded picture. The sum of energy concentrated in these harmonics is therefore used for blockiness estimation [12]. Figure below illustrates the concept of blockiness measurement through harmonic analysis in the frequency domain.

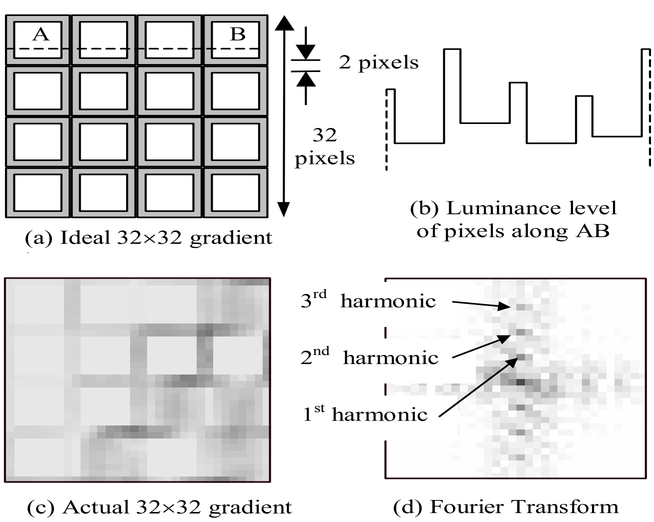

Fig 1 : Basic concept of Harmonic Amplitude Analysis

Figure 1(a) shows the gradient of block of 8x8 pixels within a 32x32 pixel image. Scanning of the gradient image results in a rectangular waveform as shown in figure 1(b). Fourier transform of this waveform generates harmonics at the spatial frequencies of multiple of 8/32. For real images, figure 1(c) shows the gradient of a real blocky image at block dimensions of 8x8 pixels. The two dimensional harmonics of the image with a 2D Fourier transform is shown in figure 1(d). Naturally if the image didn't have such a regular pattern, then the image energy in these harmonics will be nonexistent. Comparing the image energy at these harmonics with those of whole block energy at the ac coefficients can give an indication of blockiness [12].

III. Full Reference Blockiness Meter

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

a) Overview of Blockiness Estimation in FR Mode

The block diagram of the simplest full reference meter without using any spatial masking is given in figure below.

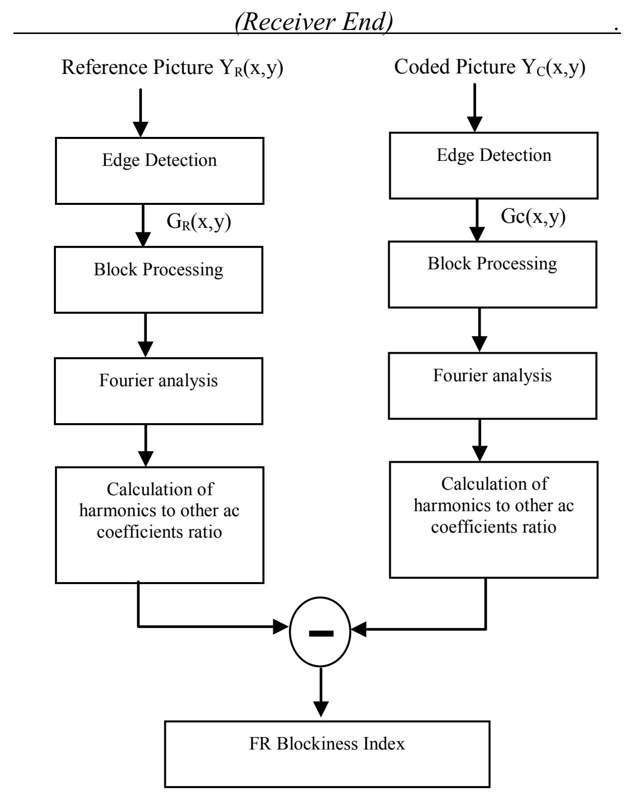

Fig 2 : Block diagram of full reference blockiness meter

Due to the availability of complete reference image at user end, the first step is to apply the edge detection process on both reference and coded images. It is used to determine the sharp luminance edges which might be because of the details present in the image or because of the blockiness artifact, as blockiness results in sharp luminance discontinuities at DCT block boundaries. The edge detection process is performed in horizontal and vertical directions on both images separately. The chances of misreading textual details to be treated as blockiness artifact are reduced by using gradient versions of both images.

Since, the blockiness distortion is not likely to be the same in all parts of the image, therefore both of the resulted edge detected gradient images are then divided into blocks in spatial domain to determine the blockiness locally for each block. The block size must be any multiple of 8x8 (DCT block size). Here 32x32 block size is selected to keep the appropriate distance between harmonics. After block processing, each block is transformed into frequency domain and harmonics are calculated for each block for both gradient images. Each block of 32x32 will have a DC coefficient and two repetitive groups of 15 ac components in horizontal and vertical directions denoted by H1-H15 and V1-V15 respectively. Of these ac components, the particular frequency components are extracted which have same spatial frequency as the blockiness pattern (i.e. 8 pixels/cycle) and its multiples, called harmonics. These harmonics can be determined using the following equation.

$$f_k = k * \left(\frac{w}{8}\right), \quad \text{for k} = 1,2,\ldots, \left(\frac{w}{8} - 1\right) \qquad (1)$$

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

$$Rh = \frac{\sum(H_4+H_8+H_{12})}{\sum_{n=1}^{16} H_n} \qquad (2)$$

and

$$Rv = \frac{\sum(V_4+V_8+V_{12})}{\sum_{n=1}^{16} V_n} \qquad (3)$$

The above ratio is calculated for each 32x32 pixels block of reference and coded images and then compared to determine the amount of blockiness locally for every block of the coded image. The blockier an image is, the higher will be the harmonics to AC components ratio. Finally, the blockiness of each block is accumulated at the end for a single quality metric.

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

b) Effect of Spatial Masking in Full Reference Mode

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

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

The edge cancelled image is obtained by taking the difference of reference and coded gradient images. To determine the blockiness locally, the edge cancelled gradient image is divided into blocks and then each block is transformed into frequency domain. If the coded image is affected by blockiness then the edge cancelled gradient image must have harmonics at spatial frequencies and by comparing the harmonics with other ac coefficients, the blockiness can be estimated accurately. The Pearson's correlation coefficient of 96.12% is obtained by including the edge cancellation process which improves the above results by 0.3%. The block diagram of the full reference meter with edge cancellation process is given below.

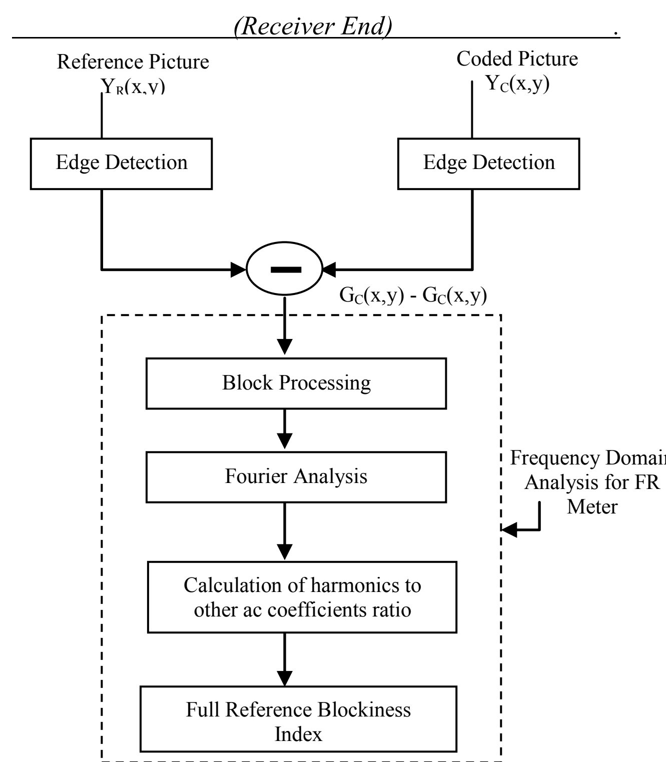

Fig 3 : Block diagram of FR blockiness meter with edge cancellation process

Global Journal of Computer Science and Technology Volume XI Issue XX Version I December 2011

As the human visual system is non-linear and sensitive to different frequency variations this means that users observe different amount of distortion depending upon the adjacent details. The distortion will be masked by the nearby details and be less visible to human observers. Applying more sophisticated masking technique includes the features of human visual. This property is implemented by using the masking technique by Fiorentini and Zoli [13]. For this purpose, the local masking function is calculated for each pixel using the adjacent pixel values to implement the property of human visual system. For each pixel in gradient reference image, the adjacent horizontal and vertical pixels (5 adjacent pixels) are compared and the minimum value is chosen which represents the maximum masking effect for that pixel. The process is applied for each pixel of gradient reference image and the final masking function is determined which contains the spatial activity of the reference image.

The masking function is then convolved with the edge cancelled coded gradient image to mask the distortion. Then the block processing is applied on masked coded image and each block is transformed into frequency domain analysis. For each block, the harmonics to other ac coefficients ratio is calculated to determine the blockiness distortion locally. Finally, the distortion of each block is accumulated in the end for single value. The Pearson's correlation coefficient is improved to 96.25% by using this masking technique. The block diagram including above masking technique is given below.

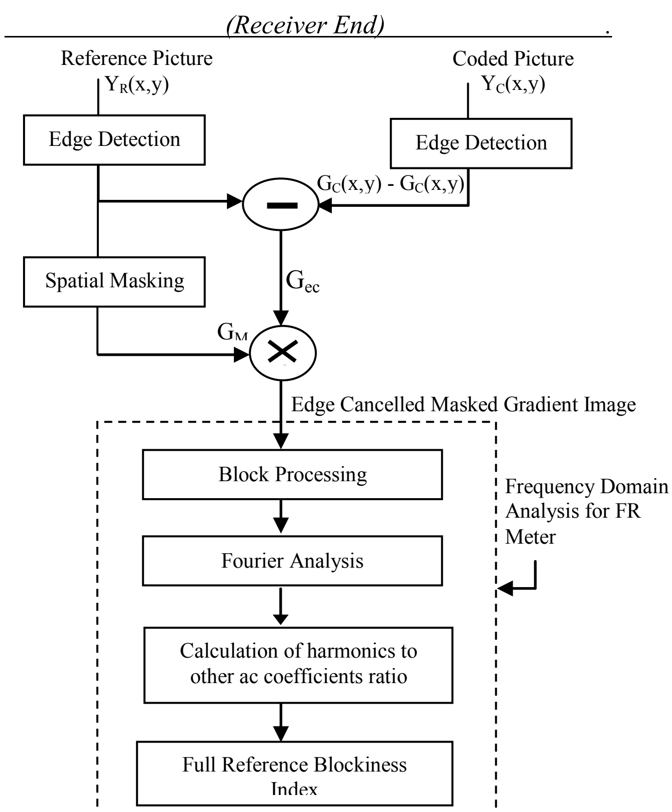

Fig 4 : Block diagram of FR blockiness meter with spatial maskings

To conclude the discussion for full reference distortion meter, the table below summarizes the effect of spatial masking on full reference meter.

Table 1 : Effect of different maskings on FR blockiness meter

FR Blockiness Meter	Processing Time	Correlation Coefficient
Without Any Masking	1 Unit	95.75%
Edge Cancellation	1.33 Units	96.12%
Fiorentini and Zoli Masking	4.18 Units	96.25%

Due to the availability of the complete reference image at user end in full reference mode, the spatial masking has not much importance in full reference mode. However, it helps to improve the accuracy of the meter but it increases the complexity and processing time of the system. Since masking also increases some processing burden to the quality meter, the table also shows how much this is increased due to the calculation of masking function.

IV. Reduced Reference Blockiness Meter

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

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, 8^{th} 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.

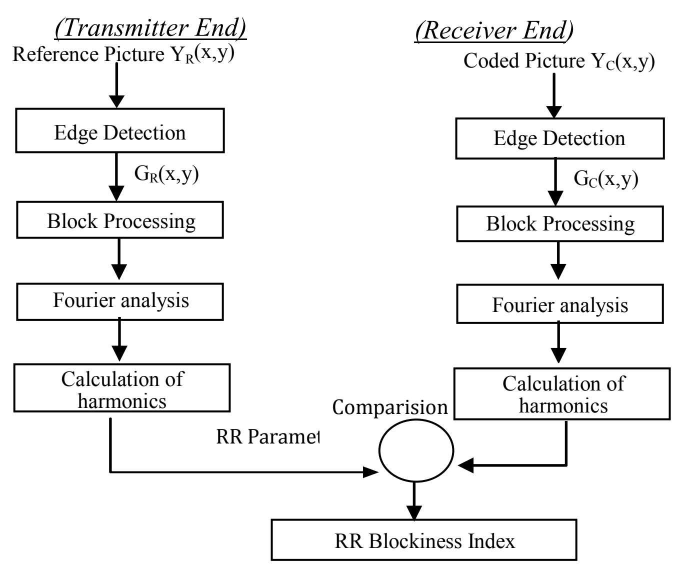

Fig 5 : Block diagram of reduced reference meter without any masking

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.

$$R_{ref} = \sum(H_4 + H_8 + H_{12} + V_4 + V_8 + V_{12})_{\,reference} \quad (4)$$

and

$$R_{coded} = \sum(H_4 + H_8 + H_{12} + V_4 + V_8 + V_{12})_{\,coded} \quad (5)$$

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.

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 weather 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. $R_{coded} > R_{ref}$, 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.

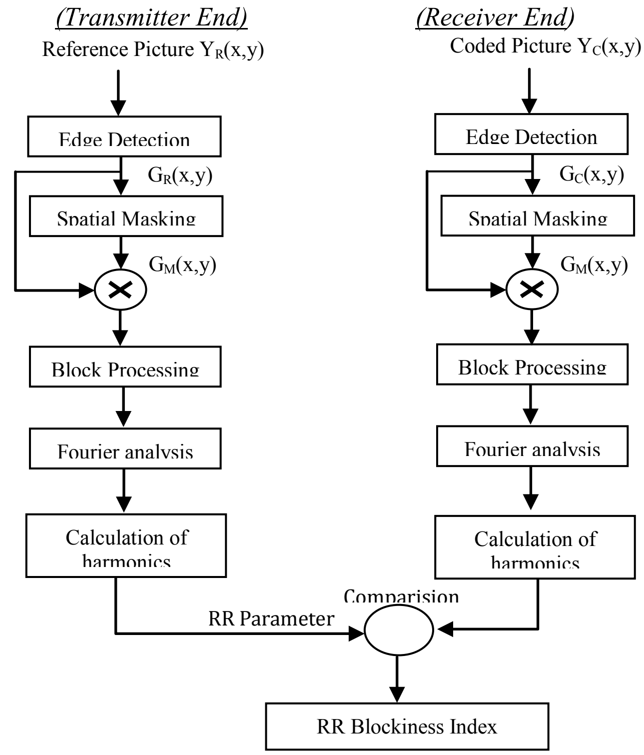

Fig 6 : Block diagram of RR meter with spatial masking

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.

Table 2 : Effect of different maskings on RR blockiness meter

RR Blockiness Meter	Processing Time	Correlation Coefficient
Without any Masking	1 Unit	88.2%
Edge Cancellation	1.2 Units	93.59%
Fiorentini and Zoli Masking	6.7 Units	94.89%

Due to unavailability of full reference image at user end in RR mode, the edge cancellation process helps to discriminate between natural edges and edges due to blockiness. But using complex masking techniques may improve the results but they require more processing time as shown in table 2 above.

V. 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 techniques 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.

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.

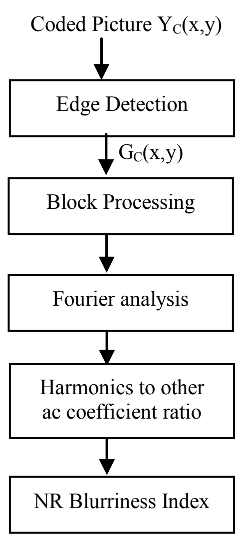

Fig 7 : Block diagram of NR meter without any masking

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 not as good as FR and RR approaches because there are chances of interpreting natural edges as blockiness as the 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.

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 would roughly be same as of reference picture. The rest of the procedures are same as used for above techniques. The block diagram of the NR meter with spatial masking is given below.

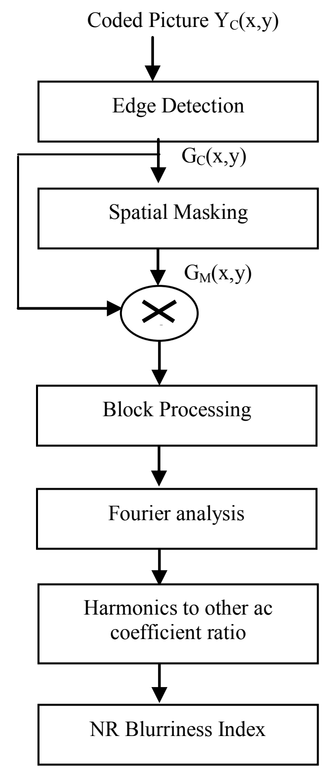

Fig 8 : Block diagram of NR meter with masking

The correlation coefficient of 90.30% is obtained by including the spatial masking from Fiorentini and Zoli. The table below summarizes the effect of spatial masking on no reference meter.

Table 3 : Effect of different maskings on NR blockiness meter

NR Blockiness Meter	Processing Time	Correlation Coefficient
Without any Masking	1 Unit	83.72%
Edge Cancellation	Not Possible due to absence of any reference	
Fiorentini and Zoli Masking	3.6 Units	90.30%

The above results show that the spatial masking is essential in NR mode as natural edges of the image may be treated as blockiness edges because of the absence of any reference information.

VI. Conclusion

The aim of the work is to determine the importance of spatial masking in FR, RR and NR modes. Different masking techniques, from simple to complex, are applied on all three types of meters. The algorithm is tested on LIVE image database of blocky images compressed at different compression rates. The database consists of 233 images with their 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 blockiness detection is possible using the gradient versions of reference and coded images. However, in RR mode, in table 2, the features of reference image can be used to distinguish natural and blockiness edges to improve the quality of distortion meter. Simple cancellation of natural edges at DCT block boundaries helps to develop a reliable distortion meter and there is no need for any sophisticated spatial masking.

Finally, in NR mode, as in table 3, complex masking techniques are required to compensate the absence of any required information, this is because some part of natural edges may be miscalculated as blockiness and the masking will reduce this miscalculation. However, the accuracy is not as good as for FR and RR meters. It shows 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.

References Références Referencias

1. 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.19, no.6, pp.1427-1441, June 2010.
2. 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, Jan. 2010.
3. A.A. Webster, C.T. Jones, M.H. Pinson, S.D. Voran, S. Wolf, "An objective video quality assessment system based on human perception," SPIE Human Vision, Visual Processing, and Digital Display IV, San Jose, CA, February 1993, pp. 15–26.
4. Z. Wang, A. C. Bovik, and B. L. Evans, "Blind measurement of blocking artefacts in images," in IEEE Int. Conf. Image Processing, September 2000, vol. 3, p. 981-984, IEEE.
5. H. R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," IEEE Signal Processing Letters, Vol. 4, Nov. 1997, pp. 317-320.
6. L. Meesters and J. B. Martens, "A single-ended blockiness measure for JPEG-coded images", Signal Process., Vol. 82, pp. 369-387, 2002.
7. Z. M. Parvez Sazzad, Y. Kawayoke, and Y. Horita, "No-reference image quality assessment for jpeg2000 based on spatial features," Signal Processing: Image Communication, vol. 23, no. 4, pp. 257–268, April 2008.
8. Zhu and P. Milanfar, "A no-reference sharpness metric sensitive to blur and noise", QoMEX, 2009. July 29–31, 2009, San Diego, California, U.S.A.
9. M. G. Choi, J. H. Jung, and J. W. Jeon, "No-Reference Image Quality Assessment using Blur and Noise," in International Journal of Computer Science and Engineering 3:2 2009, pp. 76–80.
10. R. Barland and A. Saadane, "A new reference free approach for the quality assessment of mpeg coded videos," in 7th Int. Conf. Advanced Concepts for Intelligent Vision Systems, Sep. 2005, vol. 3708, pp. 364–371.
11. H. R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," IEEE Signal Processing Letters, Vol. 4, Nov. 1997, pp. 317-320.
12. K. T. Tan, M. Ghanbari, "Frequency domain measurement of blockiness in MPEG-2 coded video," Proceedings of ICIP 2000, Vol. 3, pp.977 – 980.
13. Fiorentini, A. and Zoli, M. T., "Detection of a Target Superimposed to a Step Pattern of Illumination. II. Effects of a Just-Perceptible Illumination Step", Atti. Fond. G. Ronchi, Vol. 22, pp. 207-217,1967.
14. Live website for subjective scores MOS., http://live.ece.utexas.edu/research/quality/