IMAGE ARTIFACTS AND NO-REFERENCE FEATURE EXTRACTIONS: A REVIEW

Tubagus Maulana Kusuma

Graduate Program – Department of Electrical Engineering ,
Gunadarma University
Jl. Margonda Raya 100, Depok 16424, Indonesia
E-mail: mkusuma@staff.gunadarma.ac.id

ABSTRACT

Image attributes, such as shape, orientation, size, texture, gradient, luminance, brightness, and contrast can be easily recognized by human visual system. All these attributes contribute to the characteristics of an image or image features, which reveal uniqueness of the image. Once the images get corrupted due to compression or transmission, the unexpected artificial features, which are not existing in the original images are introduced. The characteristics of these unexpected features known as artifact characteristics vary from one artifact to another. By knowing the original image features, un-expected features can be detected. There are a number of feature extraction techniques that can be used to extract image features, in order to observe the presence of artifacts. These techniques can be generally classified into two categories, as spatial domain based and transform domain based. This paper presents a review of the existing spatial domain artifact extraction techniques using no-reference (without comparison to original image) as they are computationally inexpensive and suitable for in-service quality monitoring.

Keywords: Artifact, feature extraction, error, blocking, blur, ringing, masking, lost blocks.

INTRODUCTION

Human vision has the ability to recognize image attributes, such as shape, orientation, size, texture, gradient, luminance, brightness, and contrast. All these attributes contribute to the characteristics of an image or image features, which reveal uniqueness of the image. An undistorted image carries its natural features. However, once these images get corrupted due to compression or transmission, new artificial features, which are not existing in the original images are introduced. The characteristics of these new artificial features known as artifact characteristics vary from one artifact to another. By knowing the original image features, artificial features introduced by compression errors or transmission errors can be detected. Moreover, the level of severeness of these artifacts can also be measured. The impact of the presence of artifacts to the perceivability of an image is based on the fact that the details decrease as the viewing distance increases. As a result, the HVS fails to capture the artifact when the viewing distance is large. Conversely, once the viewing distance is decreased to a certain extent, the whole details of the image at the pixel level can easily be perceived. The size of the image also affects the perceivability of the artifacts. Assuming the viewing distance is constant, the larger the size of an image, the easier the human vision detects the artifact. In a small terminal such as a mobile phone with small screen, artifacts are even more difficult to perceive. There are a number of feature extraction techniques that can be used to extract image features, in order to observe the presence of artifacts. These techniques can be generally classified into two categories, as spatial domain based and transform domain based. In this paper, only spatial domain artifact extraction techniques using no-reference (without comparison to original image) will be considered as they are computationally inexpensive. Furthermore, these techniques are suitable for low delay tolerance applications, such as in-service quality monitoring and link adaptation.

Common image artifacts that exist in the image acquisition, processing, and transmission are shown in Table 1. However in this paper, only the artifacts, blocking, blur, ringing/false edge, masking, and lost block (Winkler, 2000;

Jakulin, 2000; Rane, 2002), that have been observed during the simulations and carried out on wireless communication system simulator in Matlab environment are considered. Moreover, as HVS is more sensitive to luminance component, distortions in the chrominance component are not discussed.

Table 1.
Common image artifacts.

Artifact	Origin
Blocking	coarse quantization of frequency components
Blur	loss of high frequency components
Ringing/False edge	improper truncation or quantization of high frequency components
Masking	a transmission distortion where low frequency components are corrupted
Lost blocks	corrupted bits in the very beginning parts of the image bitstream
Salt and pepper noise	maximum-amplitude pixels replace original image pixels in a spatially random manner
Speckle noise	the interaction of out-of-phase waves with a target in ultrasound images
Gaussian noise	thermal noise during image acquisition
Colour smearing	the reduction of high frequency coefficients of the chroma components

ARTIFACT CHARACTERISTICS AND EXTRACTIONS
Blocking

Blocking appears in all block-based compression techniques and is caused by coarse quantization of frequency components. These artifacts can be observed as surface discontinuity or an edge at block boundaries. The problem with a block-based scheme such as JPEG, is that the image is sub-divided into sub-blocks with the size of 8-by-8 pixels each. The transformation and the quantization processes are then applied to the sub-blocks individually and independently. Correlation among spatially adjacent sub-blocks is not taken into account during the encoding process. As a consequence, smooth transitions between edge boundaries of each sub-block are reduced. During the decoding process, the edge boundary cannot be fully recovered as it appeared in the original image. The block boundaries are now being visible. Therefore, blockiness or blocking artifact can be easily observed on the image as shown in Figure 1. It is even more noticeable when the bit rate or the number of bits to represent the image is reduced.

Figure 1. Sample original image "Lena" and its distorted version with blocking artifact.

Non-overlapped block-based scheme, such as DCT, gives rise to the appearance of blocking artifact. The extraction process of this artifact will be targeted on the activity on block boundaries. Several techniques, which take HVS into account, have been proposed to extract the information regarding the above artifact (Wu and Yuen, 1997; Wang et al., 2002).

In the study by Wu and Yuen (1997), a non-intrusive or no-reference block-edge extraction technique for video called generalized block-edge impairment metric (GBIM) has been proposed. However, this technique could also apply to still images since it focuses only on intra-coded frames. The basic idea of this method is to find the artifacts in the vertical and horizontal edges. The algorithm uses weighted mean-squared difference along block boundaries.

The block diagram of this technique is illustrated in Figure 2. The process starts by calculating the difference between two adjacent pixel values at the block boundaries. The differences are measured in vertical and horizontal direction to obtain horizontal edge artifacts, Δf_c, and vertical edge artifacts, Δf_r, with the assumption that 8-by-8 pixel blocks are used. Here, for a given image f, the column and row vectors f_c and f_r, respectively are given by

$$f_c = [f_{c1} f_{c2} \cdots f_{ci} \cdots f_{cN_c}]$$

and

$$f_r = [f_{r1} f_{r2} \cdots f_{rj} \cdots f_{rN_r}]$$

where N_c and N_r represent the width and the height of the image. Here f_{ci} and f_{rj} are respectively the ith

column and the jth row of the image f. The edge artifacts Δf_c and Δf_r are then weighted, in order to take into account the human visual sensitivity to a particular luminance value. This luminance value is typically between 70 and 90, centered approximately on 81. The weights are obtained from a weighting function, whose parameters could be adjusted to provide better correlation to user perception. The horizontal blockiness, M_h, and the vertical blockiness, M_v, can then be defined as (Wu and Yuen, 1997).

$$M_h = \|W_c \Delta f_c\|$$
$$= \left[\sum_{k=1}^{N_c/8-1} \|W_{ck}[f_{c(8\times k)} - f_{c(8\times k+1)}]\|^2 \right]^{1/2}$$

$$M_v = \|W_r \Delta f_r\|$$
$$= \left[\sum_{k=1}^{N_r/8-1} \|W_{rk}[f_{r(8\times k)} - f_{r(8\times k+1)}]\|^2 \right]^{1/2}$$

Here, $\|\cdot\|$ denotes the $l_2 - norm$ and

$$W_c = diag[w_{c1}, w_{c2}, \cdots, w_{ck}, \cdots, w_{cN_c/8-1}]$$

and

$$W_r = diag[w_{r1}, w_{c2}, \cdots, w_{ck}, \cdots, w_{cN_c/8-1}]$$

are diagonal weighting matrices which take into account the local spatial characteristics of a given image. Each w_{ck} and w_{rk} are the diagonal matrices of the weighting functions.
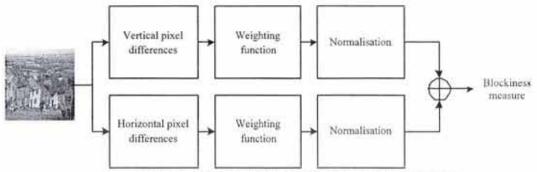
Figure 2. Block diagram of Wu and Yuen extraction technique.

Next, the M_h and M_v values are further normalized by average inter-pixel differences outside the block boundaries resulting in

M_{hGBIM} and M_{vGBIM} as

$$M_{hGBIM} = \frac{M_h}{E_c}$$

$$M_{vGBIM} = \frac{M_v}{E_v}$$

where

$$E_c = \frac{1}{7} \sum_{n=1}^{7} S_{cn}$$

$$E_r = \frac{1}{7} \sum_{n=1}^{7} S_{rn}$$

and

$$S_{cn} = \left[\sum_{k=1}^{N_c/8-1} \| W_{ck} [f_{c(8 \times k - n)} - f_{c(8 \times k - n + 1)}] \|^2 \right]^{1/2}$$

$$S_{rn} = \left[\sum_{k=1}^{N_r/8-1} \| W_{rk} [f_{r(8 \times k - n)} - f_{r(8 \times k - n + 1)}] \|^2 \right]^{1/2}$$
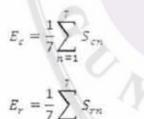
Finally, the overall blockiness measure M_{GBIM} is obtained using the weighted sum of the horizontal and vertical components as

$$M_{GBIM} = \alpha M_{hGBIM} + \beta M_{vGBIM}$$

where α and β denote the prediction parameters.

Similar to Wu and Yuen, Wang et al. have also proposed a no-reference blocking artifact extraction and measurement model (Wang et al., 2002). The model, as depicted in Figure 3 is relatively simpler than the Wu and Yuen model.
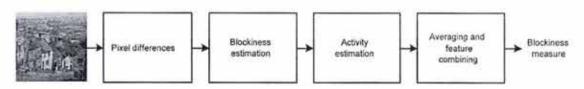

Figure 3. Block diagram of Wang's extraction technique.

It is based on the knowledge that blocking artifact has signatures in the spectral domain, which are called the signature of blockiness. They are the peaks in the power spectrum curve of the image at the feature frequencies of L/8, 2L/8, 3L/8, and 4L/8, where L is the length of the

segment extracted from the signal. The basic idea of this algorithm is to detect the blocky signal and to estimate its power using the assumption that a blocky image is a non-blocky image interfered by an ideal blocky signal. The algorithm starts with calculating the differences between pixels along the row to obtain the

horizontal difference d_h and along the column to

get the vertical difference d_v. Once these differences are calculated, the difference image can be obtained. Next, the blockiness is computed by taking the average differences

across block boundaries. For an $M \times N$ image, the horizontal measurement is defined by Wang et al. (2002).

$$B_h = \frac{1}{M(\lfloor N/8 \rfloor - 1)} \sum_{i=1}^{M} \sum_{j=1}^{\lfloor N/8 \rfloor - 1} |d_h(i, 8j)|$$

where $(\cdot, \cdot)$ denotes the index of pixel location. Similarly, the vertical measurement, B_v, can be obtained. In addition to extracting the blocking artifact, this algorithm can also be used to identify the bluriness by observing the reduction of signal activity. It is believed that some insight about the relative blur can be obtained by combining blockiness with the activity measurements (Wang et al., 2000). The activity can be measured using two parameters, namely the average absolute difference between in-block image samples, A, and the zero-crossing (ZC) rate, Z. The first parameter in the horizontal direction can be obtained as

$$A_h = \frac{1}{7} \left[\frac{8}{M(N-1)} \sum_{i=1}^{M} \sum_{j=1}^{N-1} |d_h(i, j)| - B_h \right]$$

The horizontal zero-crossing rate is defined as

$$Z_h = \frac{1}{M(N-2)} \sum_{i=1}^{M} \sum_{j=1}^{N-2} z_h(i, j)$$

where

$$z_h(i, j) = \begin{cases} 1 & \text{horizontal ZC at } d_h(i, j) \\ 0 & \text{otherwise} \end{cases}$$

Similarly, the vertical measures A_v and Z_v can also be obtained. Then, the mean blockiness, B, the mean absolute difference, A, and the mean ZC rate, Z, are obtained using the vertical and the horizontal measures. The final blockiness measure, S, can then be obtained using

$$S = \alpha - \beta B^{\gamma_1} A^{\gamma_2} Z^{\gamma_3}$$

Where α, β, γ_1, γ_2, and γ_3 denote the prediction parameters that were obtained using curve fitting of the subjective experiment data.

Blur
Artifact characteristic
Blur is observed as edge smoothness or lack of fine detail as shown in Figure 4. It is caused by the loss of high frequency components when compared with the original image. Mathematically, image blur can be modeled as

$$g(x, y) = h(x, y) * f(x, y) + n(x, y)$$

where $g(\cdot, \cdot)$, $f(\cdot, \cdot)$, and $h(\cdot, \cdot)$ represent the blurred image, the original image and the point spread function (PSF) or blur function, respectively. The function $n(\cdot, \cdot)$ denotes additive noise from image acquisition if exist. The symbol * is a convolution operator. The artifact of blur mostly occurs in wavelet-based

compressed images, such as JPEG2000. It is due to the multi-resolution decomposition of wavelet transform. If an image is highly compressed, then only low frequency coefficients are preserved in the compressed image. As a result, the image looses fine detail that are associated with high frequency

components. The shape information is basically preserved while the texture information is severely smoothened. However, the DCT-based compression scheme such as JPEG also exhibits blur, despite it is not a major artifact (Ebrahim, et al., 2000).

Figure 4. Sample original image "Goldhill" and its distorted version with blur artifact.

Artifact extraction

Extracting information about image blur can be done by using various approaches. Marziliano *et al.* (2002, 2004) believed that blur is perceptually apparent along edges or in textured regions. Therefore, their algorithm is focused on how to observe the edge spread in spatial domain. Despite the observation can be performed in both vertical and horizontal orientations, only the vertical edge processing is observed, as it is sufficient to identify the blur. It also depends upon the type of application that this algorithm is going to be used. Adding a horizontal edge processing will, of course, increase the processing power. Initially, the vertical edge of the blurred image is detected using first-order derivative edge detection using Sobel operator (weeks, 1998) shown in Figure 5.

scanned to find the edge. If an edge is found, then the local extreme (local maximum and local minimum) is searched to the left and right direction of the edge pixel. To find the distance or the length of the edge, the difference between local maximum and local minimum is calculated. The result is identified as the local blur measure for the current edge location. Finally, the global blur measure is computed by calculating the average of the local blurs over all edge locations. The block diagram of this technique is shown in Figure 6.

Similar to the technique in Marziliano (2002, 2004), Ong *et al.* (2003) attempt to measure the average edge spread. In this technique, edge spread is identified by looking into the slope spread of an edge in the gradient directions. The block diagram of this technique is illustrated in Figure 7.

$$Vertical = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \qquad Horizontal = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Figure 5. Sobel edge operator.

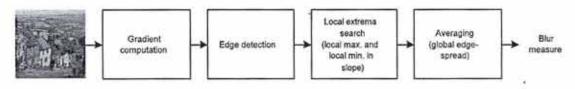Following the edge detection, each row is

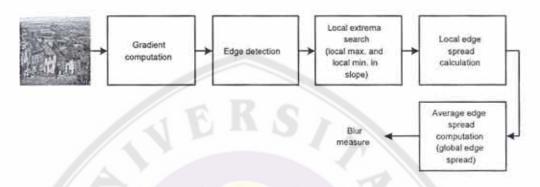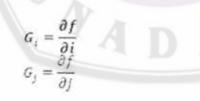Figure 6. Block diagram of Marziliano's extraction technique.

Figure 7. Block diagram of Ong's extraction technique.

The algorithm starts with computing the gradient magnitudes and gradient directions. It provides pixels with gradient information, including the signs '+' or '-' which represent gradient direction. For image $f(i,j)$, the gradient direction computation is given by

$$\theta = tan^{-1}\left(\frac{G_j}{G_i}\right)$$

where

$$G_i = \frac{\partial f}{\partial i}$$
$$G_j = \frac{\partial f}{\partial j}$$

Next, the Canny edge detection algorithm is used to find the edges on the image. In order to obtain the edge spread, each pixel is scanned until an edge pixel is found. Once an edge pixel is found, the scanning process continues to the direction of the gradient as well as to the opposite direction to find the nearest local extreme location, relative to the current edge pixel. The number of pixels to the '+' direction are counted until the maximum intensity level is reached. The same procedure is performed to the '-' direction until the minimum intensity level is reached. The local edge spread is then calculated by summing up the positive slope and negative slope edge spread. The overall edge spread, s, then becomes the total edge spread divided by the number of edge pixels found. This value gives the indication of the image bluriness. In order to predict the subjective experiment data, a prediction model is applied to obtain the image quality as

$$Q = \alpha + \beta s^\gamma$$

where α, β and γ are the curve fitting parameters.

Ringing
Artifact characteristic
Ringing is observed as periodic pseudo edges around original edges. This artifact is also known as the Gibbs Phenomenon or Gibbs effect. It occurs due to improper truncation or quantization of high frequency components. In the spatial domain, a digital signal is represented by a finite number of basis functions. The use of finite series of basis functions approximations to represent the

discontinuous waveforms produces the Gibbs effect, which is an overshoot near the point of discontinuity. It appears around edges because they contain many high frequencies. This artifact is perceived as ripples or oscillations around sharp edges in spatial domain as shown in Figure 8. Improper truncation of high frequency components due to channel errors in image transmissions creates similar effect. The edges from high resolution components can be shifted further away from the original edge locations, which create a ghost-edge effect. This effect is observed as false edge.
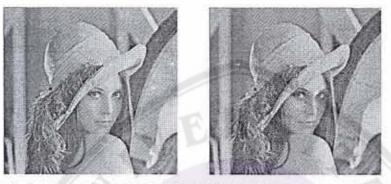

Figure 8. Sample original image ``Lena" and its distorted version with ringing artifact

Artifact extraction

The ringing artifact is caused by the quantization or truncation of the high frequency transform coefficients as discussed before. It is not an easy task to identify ringing on an image using a no-reference method. However, since ringing can be observed as an increased edge activity, image activity detection technique can be applied to the image to extract the effect of ringing. The activity measurement gives some insight into the presence of ringing artifact on the image as shown in Figure 9 and Figure 10.

Saha and Vemuri have proposed several image activity measurement (IAM) techniques, such as IAM from edge information and IAM from gradient. For a $M \times N$ image, the IAM from edge information is defined as (Sahan and Vemuri, 2000).

$$IAM_{edge} = \left[\frac{1}{MN} \sum_{i=1}^{MN} B(i) \right] \times 100$$

Where B(i) represents an edge at pixel location i. Based on the gradient activity, the IAM can be defined as

$$IAM_{grad} = \frac{1}{MN} \left[\sum_{i=1}^{M-1} \sum_{j=1}^{N} |f(i,j) - f(i+1,j)| - \sum_{i=1}^{M} \sum_{j=1}^{N-1} |f(i,j) - f(i,j+1)| \right]$$

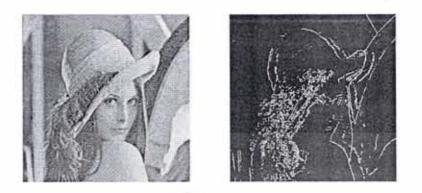Figure 9. Sample original image ``Lena" and its edge representation.

Figure 10. Sample image ``Lena'' with ringing artifact and its edge representation.

Masking and lost block
Artifact characteristic

Masking is the reduction in the visibility of one image component (the target) due to the presence of another (the masker). This occurs due to transmission distortions where low frequency components are corrupted. There are two types of masking that are observed during the simulations. In the first, a darker or a brighter intensity creates a layer on top of the image covering a certain area and this type is referred to as the intensity masking. With the second type called texture masking, the mask is a complex texture. Examples of both masking artifacts are shown in Figure 11.

Figure 11. Sample original image ``Lena'' and its distorted version with intensity masking and texture masking.

A lost block is observed due to corrupted bits in the very beginning parts of the image bit stream. The distortion causes the subsequent parts in the bit stream become useless, since they are related to each other. In common wireless transmission scenario, the image is transmitted block-by-block over wireless channel. Due to severe multi path propagation, entire image blocks can be lost as shown in Figure 12.
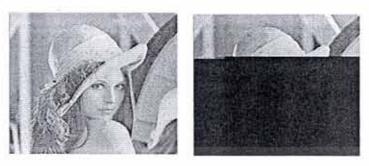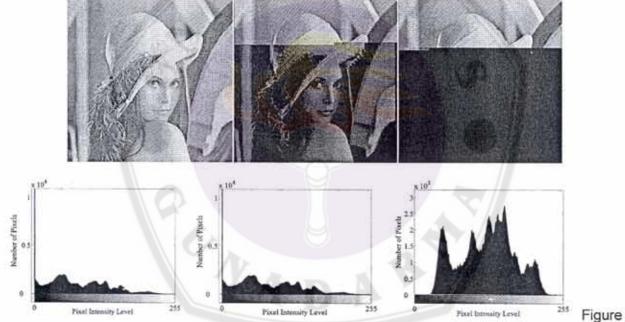
Figure 12. Sample original image "Lena" and its distorted version with lost blocks.

Artifact extraction

Masking and lost blocks are purely caused by transmission distortion. These artifacts affect the pixel distribution of an image. Therefore, the effective way to detect the presence of these two artifacts are the use of image statistics, such as the image histogram. The shape of the histogram curve gives us a clue that is related to pixel distribution and further can be used as masking and lost block detections as illustrated in Figure 13.

Figure 13. Samples for test image "Lena" with histogram information: (a) Original Image, (b) Image with lost block, and (c) Image with masking.

Combined artifact

In wireless imaging, the image may suffer from a combination of blocking, blur, ringing, masking and lost block as shown in Figure 14. In this case, artifact identification and extraction can be more complicated. The use of the previously mentioned artifact extraction techniques in a proper arrangement may be sufficient to identify and to extract the combined artifact. The amount of contributions of each artifact to the overall distortion also need to be considered as some artifacts may not be too distracted to the HVS

Figure 14. Sample original image "Lena" and its distorted version with combined artifact.

REFERENCES

Winkler, S., "Vision models and quality metrics for image processing applications," Ph.D. dissertation, École Polytechnique Fédérale De Laussane (EPFL), Dec. 2000.

Jakulin, A.. Baseline JPEG and JPEG2000 artifacts illustrated. [Online]. Available: http://ai.fri.uni-lj.si/~aleks/jpeg/artifacts.htm

Rane, S. D., J. Remus, and G. Sapiro, "Wavelet-domain reconstruction of lost blocks in wireless image transmission and packet-switched networks," in Proc. IEEE International Conference on Image Processing, Sept. 2002, pp. 309-312.

Wu, H. R. and M. Yuen, "A generalized block edge impairment metric for video coding," IEEE Signal Processing Letters, vol. 4, pp. 317-320, Nov. 1997.

Wang, Z., A. C. Bovik, and B. L. Evans, "Blind measurement of blocking artifacts in images," in Proc. IEEE International Conference on Image Processing, Sept. 2000, pp. 981-984.

Wang, Z., H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in Proc. IEEE International Conference on Image Processing, Sept. 2002, pp. 477-480.

⸻ Perceptual Blur and Ringing Metrics: Application to JPEG2000," Elsevier Journal of Signal Processing: Image Communication, vol. 19, pp. 163-172, Feb. 2004.

A. Weeks, Fundamentals of Electronic Image Processing. Chichester: Wiley-IEEE Press, 1998.

Ong, E., W. Lin, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, and F. Moschetti, "A no-reference quality metric for measuring image blur," in Proc. Seventh International Symposium on Signal Processing and Its Applications, July 2003, pp. 469-472.

Saha, S. and R. Vemuri, "An analysis on the effect of image features on lossy coding performance," IEEE Signal Processing Letters, vol. 7, pp. 104-107, May 2000.