Three Different Features Based Metric to Assess Image Quality Blindly

Saifeldeen Abdalmajeed

Electrical and Electronics Engineering, Nile Valley University, Atbara, Sudan

Article history: Received 08 August 2019, Received in revised form 08 Feb. 2020, Accepted 09 May 2020

ABSTRACT When creating image quality assessment metric (IQA) no confirmation all distortion types are available. Non-specific distortion blind/no-reference (NR) IQA algorithms mostly need prior knowledge about anticipated distortions. This paper introduces a generic and distortion unaware (DU) approach for IQA with No Reference (NR). The approach uses three different measuring features which are initiated from the gist of natural scenes (NS) using Log-derivatives. These features are; asymmetric general Gaussian distribution (AGGD) model, two sharpness functions, and Weibull distribution. All features were analyzed and compared together to examine their performance. When calibrating the proposed features performance on LIVE database, experiments show they have good contribution to the state of the art IQA and they outperform the popular full-reference peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) methods. Also they show sharpness features are the best when assess both prediction monotonicity and predict accuracy evaluation among the three features categories. Besides they show AGGD based features have the best correlation with differential mean opinion score.

Key words: Weibull distribution, sharpness functions, Image Quality Metric.

1. INTRODUCTION

The image processing techniques; acquisition, compression, transmission, restoration and enhancement are in focus of current research. Therefore, there is a demand for assessing the quality of the consequences of these methods. Known that humans are the ultimate judge of image quality, however, their judgment is time consuming, subjective and at times, impractical. Hence, automatic assessment is needed, which is referred to as objective assessment.

Objective assessment can be categorized into three types. These are full-reference (FR), reduced-reference (RR) and no-reference (NR) image quality assessment. The first models assess image quality by fully accessing the original image. While models assess image quality by extracting some features from the reference image are called RR. In spite of providing a useful and effective way to measure the quality of distorted images, full or even partial of the reference image may not be available. In addition, the purification of reference images can be also uncertain. So the only available choice is NR IQA methods [1-3]. As an example, when assessing the quality of a denoising algorithm on a real-world database the perfect noise-free image is not available.

Distortion-specific NR IQA [4-9] are based on prior knowledge about the type of distortion and they represent most existing NR IQA methods. Application of such algorithms is limited by these
specifications. General distortion algorithms NR IQA algorithms are non-distortion-specific. These are built by obtain a collection of distorted images with co-registering human scores are opinion aware (OA) [10-12]. in contrast, opinion unaware (OU) algorithms do not need training on databases of human judgments of distorted images [13]. Distorted images may not be available during IQA model construction or training, so among OU models that do not require knowledge about anticipated distortions are distortion unaware (DU) [14]. A model for no-reference image quality measurement of Log-derivative of the features: AGGD, two sharpness functions, and Weibull distribution features of natural scenes is developed in this study. It gathered the effective features from gist of image based on sharper regions. The sharper an image the better its quality as claimed by Punit and Damon [15, 16]. Moreover, more heavily weight judgments of image quality given from the sharp image regions [14].

In the field of IQA researches focus on improving prediction accuracy and they ignore algorithmic and micro-architectural efficiency. This study considers both of these problems. As it transfers into application stage from the research environment, the IQA algorithms face issues surrounding efficiency. These are include for example; execution speed and memory bandwidth requirements which began to emerge as equally important criteria. large memory and long runtimes -on the order of seconds for even modest-sized images (e.g.<1MPIxel)-are two factors affecting the algorithms that suffer in terms of lack of efficiency [17]. Such algorithms have to apply local frequency-based decomposition of the input images and would seem to require more computation. The introduced approach avoids both of these two complexities. The presented model work in the spatial domain and no transforms (e.g. DCT, wavelet, etc.) are required [18].

A lot of researches studded Weibull distribution, AGGD, and the two sharpness functions and their relationship with natural images [14, 17, 19]. The authors in [19] found that a considerable information of visual gist information contained by Weibull contrast statistics. The spatial structure of uniform textures of many different origins completely can be characterized by Weibull distribution parameters [20]. Fabian and Erhardt [21] used Weibull distribution for defect detection in textures.

2. RESEARCH MOTIVATION AND AIM
Multimedia content delivered over communication networks go through many processing stages before being provided to a human consumer. Each of these stages may introduce distortions that could reduce the quality of the final display. The economics and/or physical limitations of the devices are the main factors that mostly determine the distortion contribution of each of these stages. Technically, it is important to gauge the distortion that has been added during different stages and then measure the visual contents quality. The image quality assessment algorithms are build to estimate image distortion content. Choosing the appropriate features plays a significant role in constructing these algorithms and measuring image quality. After building a robust model, this study aims to collect three features categories, analyze them, and see the competent one through examine their performance.

3. MATERIAL AND METHOD
The devised natural low level features are composed of locally normalized luminance and contrast values. These features have been modeled as point wise statistics for single pixels. Also the model obtains the pair wise based log-derivative statistics for the relation of adjacent pixels. Before gathering the features, they fit to Multivariate Gaussian Model (MVG). The gathering process includes only the features that corresponding to sharper patches and those of rich of edges. The distance between MVG fit of the features extracted from the distorted image and
MVG model of the natural features extracted from natural (pristine) images is then calculated. This distance is assigned as the distorted image quality score.

A. Normalized luminance and contrast coefficients and their log-derivatives

The model divides the image \( I(i,j) \) into \( 96 \times 96 \) size. The contrast (3) of the distorted and the natural images for each of the patches is then calculated. Besides, The normalized luminance, denoted by, \( \hat{I}(i,j) \), of both images are also computed through local mean subtraction and contrast divisive normalization (MSCN) (1) [22] defined as:

\[
\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + 1} \tag{1}
\]

Where \( i \in \{1, 2, ..., M\} \) and \( j \in \{1, 2, ..., N\} \) are spatial domain indices, \( M \) and \( N \) are the dimensions of the image, and

\[
\mu(i, j) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} I(i+k, j+l) \tag{2}
\]

\[
\sigma(i, j) = \sqrt{\sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} \left[ I(i+k, j+l) - \mu(i, j) \right]^2} \tag{3}
\]

are the estimated local mean and local contrast respectively and \( w = \{w_{k,l}| k = -K, ..., K, l = -L, ..., L\} \) is a 2D circularly-symmetric Gaussian weighting function sampled out to three standard deviations \( (K = L = 3) \) and rescaled to unit volume. After getting coefficients using (1) and (3), features are calculated through them for each patch. The features extraction is done using log-derivative statistics [23]. To acquire the log-derivatives, (4) is used. This is to create new image sub-band \( J \). The small constant \( \varepsilon \) is taken to be 0.1 to prevent \( I(i,j) \) from being zero.

\[
J(i, j) = \log \left( \hat{I}(i, j) + \varepsilon \right) \tag{4}
\]

The five types of log-derivatives are then computed. These include horizontal, vertical, main-diagonal, secondary-diagonal, and combined-diagonal as given in (5-9).

\[
J_x(i, j) = J(i, j+1) - J(i, j) \tag{5}
\]

\[
J_y(i, j) = J(i+1, j) - J(i, j) \tag{6}
\]

\[
J_{xy}(i, j) = J(i+1, j+1) - J(i, j) \tag{7}
\]

\[
J_{yx}(i, j) = J(i+1, j-1) - J(i, j) \tag{8}
\]

\[
J_{xy0}(i, j) = J(i, j) + J(i+1, j+1) - J(i, j+1) - J(i+1, j) \tag{9}
\]

To examine the change in the presence of some distortion in the spatial domain, the coefficients of (1) and their log-derivatives statistics (5-9) are applied [10, 24].

B. The extracted features

In this research the built model used three separate feature categories. These are initiated from the gist of NS using the parameters of: a general Gaussian distribution model, two sharpness functions, and Weibull distribution. The MSCN coefficients in (1) and the five log-derivatives in (5-9) are modeled with these parameters. Only the features corresponding to sharper regions are extracted. The three groups of the gathered features construct three IQA algorithms. These algorithms and their features are examined and compared together.

C. Weibull statistics Based features

The MSCN coefficients in (1) and the five log-derivatives in (5-9) are modeled with Weibull (10), this gives 12 features. The Weibull parameters are doped with maximum likelihood estimation (MLE). Both Weibull parameters and the MLE model parameters used as features. By employing MLE, extra 12 features are obtained at yielding 24 overall. These features are computed at two scales to represent multi-scale behavior, by low pass filtering and down sampling by a factor of two, this process leads to a set of 48 features. All features are extracted in the spatial domain and were fitted with an MVG density (11), to give their rich representation. Here also the gathering process includes only the features corresponding to sharper patches.
is the shape 4 of the MSCN and the five lr main patches. Then the -1 -1 1/2 1/2 re -1 -1 log derivatives in (5-9) are modeled following two sharpness function: grey level “amplitude” and grey level “variance” (14) [17]. The outputs of these two functions represent 12 model features (MSCN and the five log-derivatives used with each sharpness function). These features are computed at two scales to portray multi-scale behavior, by low pass filtering and down sampling by a factor of two, this process leads to a set of 24 features. All features are extracted in the spatial domain and were fitted with an MVG density (11), to give their rich representation.

\[ D(v_1, v_2, \Sigma_1, \Sigma_2) = \left( \frac{1}{(u_1-u_2)^\gamma} \right)^{-1} \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{1/2} \left( u_1 - u_2 \right) \]

4. NATURAL SCENE STATISTIC MODEL

The natural scene statistic (NSS) model computed from 125 natural images, which were selected from Flickr data and from the Berkeley image segmentation database [26]. The features corresponding to sharper are selected. Each patch is divided to sub-patches of 6×6 size and only sub-patches those are sharper (effective sub-patches) are contributed into their main patches. Then the effective sub-patches of each patch were computed. Patches that had an effective sub-patch greater than 75% of the peak patch effective sub-patches over the image are selected. The features corresponding to the selected patches were gathered. The features of NSS extracted using the three categories discussed above in section (3.2). These features were then fitted to MVG model (11).

To compute the quality according to the procedure mentioned above, (15) is used.
The mean vectors and covariance matrices of the NSS MVG and the tested image MVG models are \( \mathbf{u}_1, \mathbf{u}_2 \) and \( \Sigma_1, \Sigma_2 \), respectively.

1. TESTING AND CALIBRATION

LIVE (Laboratory for Image and Video Engineering) IQA database [27] is used to calibrate the proposed algorithms and do performance analysis and comparison. LIVE database contains 29 reference images and 779 distorted ones. These are sorted into five different types of distortions as JPEG and JPEG2000 (JPEG2K) compression or introduced as Gaussian blur (Gblur). The image also can be corrupted through a Rayleigh transmission channel and is termed as fast fading (FF) distortion. One of the commonly known distortion type is the additive white Gaussian noise (WN).

To assess the prediction monotonicity, Spearman’s rank ordered correlation coefficient (SROCC) is used while Pearson’s linear correlation coefficient (PLCC) is employed to evaluate the prediction accuracy of the proposed algorithm. Before PLCC calculated, the objective scores are passed through a logistic non-linear function [28] (where its parameters are found numerically using the MATLAB function ‘fminsearch’ in the optimization toolbox) to maximize the correlations between subjective and objective scores.

2. THE RESULTS AND DISCUSSION

In this section the performance analysis and comparison between the results found using AGGD, Weibull, and sharpness based features mentioned in sections (3.2) will discussed. These include PLCC, SROCC, and subjective opinion scores.

The plots of figures (1) and (2) obtained when the features gathered through the features AGGD, Weibull, and sharpness. The figures show a comparison of SROCC and PLCC for these features respectively. Figures (1) and (2) indicates sharpness features are the best when assess the prediction monotonicity and predict accuracy evaluation which are 0.8183 and 0.8692 for the average of all distortions (indicated by ‘all’ in the figure) respectively.

The figures (3) and (4) compare the SROCC and PLCC for three features categories against popular full-reference peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) respectively. These figures show the studded features have good contribution to the state of the art IQA and they outperform the FR-PSNR and FR-SSIM algorithms.

![Fig. 1. Comparison of SROCC when extracting features using AGGD, Weibull, and sharpness functions.](image1)

![Fig. 2. Comparison of PLCC when extracting features using AGGD, Weibull, and sharpness functions.](image2)

Figure (5) shows scatter plots of differential mean opinion score (DMOS) of: peak signal-to-noise ratio (PSNR) and SSIM index versus the built model with the three different features. The figure indicates that although the metric is a blind/NR, it correlates better than FR PSNR and FR SSIM models with DMOS.
3. CONCLUSIONS

The appropriate features and the way to collect play a significant role in the issue of IQA. In this study, a model for blind IQA built and a performance comparison between asymmetric general Gaussian distribution, two sharpness functions, and Weibull distribution feature introduced. The NR model used in this paper has low computational complexity, and extracted features in the spatial domain where no transforms (e.g. DCT, wavelet, etc.) are required. The results show all the introduced features provides an excellent performances when compared with state of the art algorithms. Besides, sharpness features are the best among the studded features when assess both prediction monotonicity and predict accuracy evaluation.

REFERENCES


Fig. 5. Comparing DMOS with PSNR, SSIM and the model with the three different features. DMOS versus: (a) PSNR, (b) SSIM, (c) The model with Weibull features, (d) The model with sharpness features, and (e) The model with AGGD features.