

Chunhua Chen, Wen Chen, and Jeffrey A. Bloom, " A Universal, Reference-Free Blurriness Measure," Image Quality and System Performance VIII, Susan P. Farnand and Frans Gaykema, Editors, Proc. SPIE 7867 (2011).

Copyright 2011 Society of Photo Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic electronic or print reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

[http://dx.doi.org/ 10.1117/12.872477](http://dx.doi.org/10.1117/12.872477)

A Universal, Reference-Free Blurriness Measure

Chunhua Chen, Wen Chen and Jeffrey A. Bloom

Dialogic Media Labs

12 Christopher Way, Suite 104

Eatontown, NJ 07724

{edward.chen, wen.chen, Jeffrey.bloom@dialogic.com}

ABSTRACT

The perceptual quality of digital imagery is of great interest in many applications. Blur artifacts can be among the most annoying in processed images and video sequences. In many applications of perceptual quality assessment, a reference is not available. Therefore no-reference blurriness measures are of interest. In this paper, we present a universal, reference-free blurriness measurement approach. While some other methods are designed for a particular source of blurriness such as block-based compression, the proposed is universal in that it should work for any source of blur. The proposed approach models the gradient image of the given image as Markov chain and utilizes transition probabilities to compute a blurriness measure. This is the first time that transition probabilities are applied to perceptual quality assessment. Specifically, we first compute the transition probabilities for selected pairs of gradient values and then combine these probabilities, using a pooling strategy, to formulate the blurriness measure. Experimental studies compare the proposed method to the state-of-the-art reference-free blurriness measurement algorithms and show that the proposed method outperforms the commonly used measures.

Keywords: Perceptual quality assessment, blurriness, gradient image, Markov chain, transition probability

1. INTRODUCTION

In many applications, such as image acquisition, coding, and distribution, the perceptual quality of digital imagery is of great interest. Perceptual quality assessment of digital images and video sequences has been studied for years.

Human observers may be the best to assess the perceptual quality of digital images and videos. This is called subjective assessment. Some international standard organizations, for example, the International Telecommunication Union (ITU) with the help of the Video Quality Experts Group (VQEG 1), have documented a few standard techniques for assessing the perceptual quality, or the change in perceptual quality of digital imagery.

However, subjective assessment is expensive in terms of time and subjects involved, difficult to repeat, and infeasible for real-time applications. As an alternative, people have been actively studying objective assessment methods to seek computational device based, accurate (i.e., well correlated to the human visual system), and computationally economic approaches.

Depending on the information that is available during the measurement, objective perceptual quality assessment methods can be categorized into three classes: full reference (FR) methods, reduced reference (RR) methods, and no reference (NR, or reference-free) methods.

Full reference methods make use of the original imagery for a comparison. It can compare each pixel of the reference to each corresponding pixel of the reconstructed still image or video sequence. FR measurements deliver the highest accuracy and repeatability but tend to be processing intensive.

Reduced reference methods use a side channel between the sender and the receiver which may not be capable of transmitting the full reference signal. Parameters, extracted at the sender side, help predict the quality at the receiver side. RR measurements may be less accurate than FR methods and represent a working compromise if the side channel bandwidth is limited.

Reference-free methods are appropriate when there is no side channel. These methods only use the degraded signal for the quality estimation and have no information about the original. In many communication applications, no reference is available. Reference-free perceptual quality assessment techniques are therefore required.

A number of factors have influence over the perceptual quality of still images and video sequences. Blurriness is one that has strong influence on the overall perceptual quality. In this paper, we address the task of measuring the blurriness in digital still images and video sequences without reference.

The prior art in no-reference blur measurement can be classified into two categories. The first category is based on a measurement of edges in the image. The spread of the edge in the image is used to measure the blurriness of the image. In this category, some methods make assumptions about the location of the edges and others first apply an edge detection technique. Methods in the second group do not require edge detection and are often machine-learning based. In this category, some global features such as the gradient histogram, or local features such as the power spectrum, are input to a trained classifier. Further processing based on the output of the classifier gives blurriness measure.

Blurred images generally do not have sharp edges, so blur is perceptually apparent along edges. P. Marziliano et al. [2] propose a blurriness measurement method based on the spread of the edges in the spatial domain. First an edge detector is applied. Then, for each edge pixel, two pixels corresponding to the local maxima and the local minima around the edge are detected. The edge width is given by the difference between the positions of the two extremes. Finally the image blur is defined as the averaged edge width.

E. Ong et al. present a blur measure also based on the spread of an edge in [3]. First they determine the direction of the gradient at each pixel and apply a Canny edge detector to the image in order to identify the edge pixels. For each edge pixel, they search along the gradient direction to find the first local extremum on each side of the edge. The distance between these two points is defined as the edge spread for that edge pixel. The quality score is then computed by using the average edge spread in a parameterized exponential model.

In [4], R. Ferzli and L. J. Karam measure the image blur by integrating the concept of just noticeable blur (JNB) into a probability summation model in which models of the human visual system are exploited. The JNB is defined as the minimum amount of blur around an edge that can be perceived for a given contrast. It is determined experimentally via subjective experiments to find the standard deviation of Gaussian filter corresponding to the JNB threshold for a specific contrast. The corresponding edge width (JNB width) is measured and used as a parameter for the blurriness measure. To measure the image blur, the image is divided into 64×64 pixel blocks, then edge detection and edge width calculation are performed on each block using the same methods as in Marziliano et al. For blocks containing many edge pixels (termed *edge blocks*), the local contrasts in the neighborhood of the edges are used to recover the subjectively obtained JNB widths. A comparison of the measured widths with the subjectively obtained JNB widths is used to classify the blur on each edge as perceptible or imperceptible. These values are pooled over all edges in an edge block and then over all edge blocks in the image.

The approach in [4] works for moderate levels of blur. However, in images with a high amount of blur, edges might go undetected. This is a common problem with methods that rely on edge detection. To overcome this problem, the authors developed an iterative edge refinement algorithm in [5] to detect blurry edges. For every 64×64 block, the Sobel edge detector is applied with an initial gradient threshold to obtain an initial estimate of the number of edge pixels. Iteratively, threshold is reduced and the corresponding number of edges is recomputed. The iteration stops when the computed number of edges is essentially the same as the number found at the previous iteration. The iterative refinement improves the accuracy of blur measurement by recovering relevant edge pixels in images with a high amount of blur, but also greatly increases the computational complexity.

Caviedes et al. introduce the kurtosis concept into image sharpness measurement in [6]. First Canny edge detection is performed in the spatial domain to obtain an edge image. The DCT of an 8×8 pixel block centered at each edge pixel is normalized to form the probability density function (PDF), which is used to calculate the kurtosis. Finally, all the local kurtoses are averaged to give the final sharpness measure in the image. This algorithm has high computational cost as it requires a block DCT for each edge pixel in addition to the Canny edge detector.

In [7], D. Liu et al. design a specific blurriness measure for H.264/AVC video. Different from the algorithms in [2], [3], and [4], which use edge detection to locate the edge pixel position, Liu et al.'s algorithm is designed to measure the blur introduced by the de-blocking filter of an H.264/AVC encoder and therefore considers the pixels that lie along the macroblock boundaries as edge pixels. Local blur is measured in regions around edge pixels that have a local variance falling into the experimentally determined range [2, 30]. This local blur calculation is a modified version of that used in Marziliano et al. Finally all the local blurs are averaged to give the final blurriness measure. Because no edge detection is required, the computational cost is low compared to [2], [3], and [4].

M. Chen and A. C. Bovik [8] present a no-reference blur measure using a machine learning classifier and a multi-resolution wavelet decomposition. A coarse blur is first assessed using a two-class support vector machine. The feature for this classification is the gradient histogram, selected due to the observation that natural images have a heavy tailed distribution of gradients while the gradient distribution tail of blurred images tends to be small. For a given image, the classifier outputs a label indicating its class and a confidence score indicating the certainty of the classification result. The coarse blur score, defined as a function of confidence score, is one component of the final blur score. The other component is obtained by analyzing a wavelet decomposition of the image. At each layer of the wavelet decomposition, a detail map is formed by the horizontal and vertical high subbands. A detail score, defined as a function of gradient magnitudes of the detail map, is combined with the coarse blur score to give the final blur score.

R. Liu et al. [9] propose a framework to detect blurry regions and classify the blur type as either motion blur or out-of-focus blur. Three blur features are measured. First, *local power spectrum* is used because the amplitude spectrum slope of a blur region tends to be steeper than that of a non-blur region due to loss of high frequency components. Second, blur regions rarely contain sharp edges, resulting in small gradient magnitude, so the *distribution of gradient magnitude* serves as an important clue for blur detection. Finally, non-blur regions are likely to have more vivid colors than blur regions, so small values of the *maximum value of saturation* is an indication of a blur region. The image is partitioned into blocks and these three features are measured in each block. Based on these features, a trained classifier labels each block as blurred or non-blurred. The image blur is defined as the percentage of blur blocks in the image.

Many of the existing reference-free blur measures, such as [2], [3], [4], and [6], are based on evaluating the edge spread which normally requires edge detection. The reliable detection of edge is a challenge, as indicated by Marziliano et al [2]. As the blurriness increases, several edge pixels go undetectable by standard edge detectors such as Sobel and Canny. As a result, only a portion of the relevant pixels are used to measure the blurriness of the image. More advanced edge detection methods can improve the robustness of the blur detection, but usually with a corresponding increase in the computational cost. Although the method proposed in [7] does not require edge detection, it cannot be used to detect blur from general sources other than block-transform-based compression.

The non-edge-based methods in [8] and [9] use a learning framework to detect blur. Like all classifiers, these depend on the effectiveness of the features and the availability of sufficient training samples. The method in [8] suffers from high computational cost since multi-resolution wavelet decomposition is involved and the method in [9] does not provide information on the extent to which each region is blurred.

In this paper, we present a universal reference-free blurriness measurement approach. The method is reference-free because it does not rely on any pre-distortion information (such as the original imagery, information about the original, or information about the processing). The method is universal because it makes no assumptions about the source of the blurriness. The proposed method is designed to measure blurriness in still images. It can detect blurriness due to any source including JPEG2000 compression and Gaussian filtering. When applied to frames of a video sequence, it can again detect blur due to any source including H.264 compression. The proposed method does not require edge detection. Experimental studies on a few publicly available datasets have demonstrated the effectiveness of the proposed scheme. The proposed method is also computationally economic, which makes its real-time application possible.

The rest of this paper is organized as follows. In Section 2, we take a close look at blurriness and how it changes some commonly used image statistics. Then we propose an image statistical model for blurriness measurement in Section 3. The proposed blurriness measurement scheme is presented in Section 4. Experimental results on still images and that on video sequences are given in Sections 5 and 6, respectively. Discussion and future works are addressed in Section 7. Finally, the paper is concluded in Section 8.

2. BLURRINESS AND IMAGE STATISTICS

2.1 Blurriness

Blur can be introduced in a picture from a number of different sources. Blur can be introduced during image acquisition (capture blur) or after image acquisition as a result of processing. Typical causes of capture blur include motion blur caused by the relative motion between the camera and the scene, out-of-focus blur, and blur owing to lens aberration. Processing blur can be introduced by image processing such as editing, denoising, or compression. In general, blur indicates an attenuation of the high spatial frequencies. For example, in a lossy JPEG2000 compression, the quantization truncates the wavelet coefficients in the high frequency subbands and the loss appears as blur in the reconstructed image.

Figure 1 illustrates two blurred images due to capture blur (Figure 1(a) and (b)) and one due to JPEG2000 compression (Figure 1(c)).



Figure 1. Sample blurred images: (a). motion blur, (b). out-of-focus blur, and (c). JPEG2000 blur

Measurement of processing blur is particular interest as part of an assessment of the impact or side effects of a particular process. For example, the amount of blur can be used to evaluate the perceptual quality degradation due to image compression or to assess the effectiveness of image enhancement algorithms.

A color image (image source: [10]) and blurred versions associated with several image/video processing operations are shown in Figure 2. To highlight the perceptual variations among these images, the partially magnified image (within the white rectangle) of each of them is also shown to the right. The Gaussian blurred image in Figure 2(b) looks evenly blurred, the blurriness in the JPEG2000 compressed image in Figure 2(c) looks less even and less regular, and the H.264 compressed image in Figure 2(d) displays some blockiness in addition to the dominant blurriness distortion. It is a challenge to find a universal measure for all these kinds of blurriness since different blurs due to different sources look quite different.





Figure 2. An image and its blurred versions (left), and partially magnified images (right): (a). original image, (b). Gaussian low-pass filtered, (c). JPEG2000 compressed, and (d). H.264 compressed

2.2 Image Statistics for Blurriness Measure

As discussed in Section 1, some of the blurriness measures in the literature rely on edge-detection and others do not. Given the challenge of detecting edges in a severely blurred image, we seek a method that does not rely on edge detection. Instead, motivated by the observation that blurriness is a perceptual effect of image area rather than one of individual pixels, we seek a method based on image statistics conveying information among adjacent pixels.

The most commonly used image statistic is the image histogram, which is used to estimate the pixel value distribution or probability mass function. The one-dimensional histogram is a global statistic. The histogram of the original image of Figure 2(a) and that of the Gaussian blurred image of Figure 2(b) are shown in Figure 3(a). Clearly, the Gaussian filtering has changed the histograms and perhaps there is some way to measure the blurriness by examining the histogram. However, the relationship between neighboring pixels is not represented in the 1-D histogram.

As described in Subsection 2.1, there are many sources of blur in an image. In most cases, a blurred version of an image will have reduced high spatial frequencies compared to a non-blurred sharp version. In other words, blur often indicates loss of sharpness. Intuitively speaking, that means the difference between adjacent pixels is expected to decrease due to blurring.

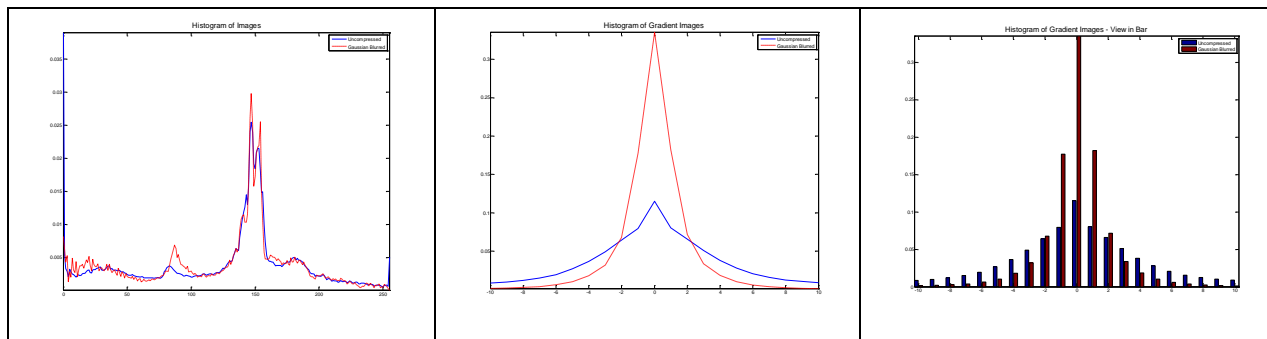


Figure 3. Histograms (a) of two images, (b) of two gradient images, and (c) of the gradient images as a bar graph

In order to capture this effect in an image statistic, we consider the histogram of the *gradient image*, often called *gradient histogram*. A gradient image is the difference between an image and a 1 pixel shifted version of that image. In other words, the values in a gradient image are the differences between adjacent pixels. In this work, we consider the vertical gradient image capturing the difference between each pixel and its vertical neighbor. The histograms of the gradient images of Figure 2(a) and Figure 2(b) are shown in Figure 3(b). To better illustrate the difference between these two gradient histograms, they are shown using bars in Figure 3(c).

The gradient histogram of an image shrinks towards zero as the image is blurred. That is, bars far away from zero decrease while the bars close to zero (including the zero bar) increase greatly. This is due to the attenuation of high spatial frequencies caused by low-pass filtering and it is this change we seek to measure.

3. MODELING BLURRED IMAGES

The gradient histogram has been used in [8] and [9] as the basis of blurriness measurement. However, it lacks information about the correlation between adjacent elements in the gradient image. We believe that blurriness can be better captured by examining how the gradient image changes spatially.

In this section, we propose a novel method for modeling the introduction of blurriness to an image. To do this, the gradient image is first modeled as a Markov chain specified by transition probabilities. We then represent the effect of blurring as a specific, easily measurable change to the set of transition probabilities.

3.1 Markov Chain and Transition Probability Matrix

In mathematics, a Markov chain [11] is a discrete stochastic process with the property that the probability of the system at the next step depends only on the current state of the system and not on previous states. A transition probability is the probability that the system will change from one state to another. A transition probability matrix, a table containing all the transition probabilities, is often used to specify a Markov chain statistically.

3.2 Modeling Gradient Images

A (vertical) gradient image has the same number of columns as the image and one fewer row. The range of values is $-N$ to $+N$ where N is the maximum pixel value. Each column of this gradient image is modeled as a first-order Markov chain with the value of one element treated as the current state and the value of the element directly below treated as the next state.

The transition probability matrix represents the probabilities that, for a current state P , the next state will be Q . Note that there are $(2N+1) \times (2N+1)$ such probabilities in the matrix. In analyzing a gradient image we can count the occurrences of each transition to estimate the probabilities. For many transitions, there will be too little data for a reliable estimate. However the estimate can be useful for some of the more common transitions.

3.3 Effect of Blurriness on Transition Probability Matrix

Figure 4 illustrates the effect of blurriness on the transition probability matrix. For the image of Figure 2(a) and the blurred version of that image (that of Figure 2(b)), the transition probabilities were estimated for values of P and Q between -5 and $+5$. The graph in Figure 4 shows the ratio between these two sets of transition probabilities. A value greater than 1 indicates that the blur caused an increase in the corresponding transition probability and a value less than 1 indicates that the blur caused a decrease in the corresponding transition probability.

It is clear from Figure 4 that transition probabilities between states with the same sign tend to increase due to blur and transition probabilities between states with different signs are decreased due to blur. This is not a surprise. Blurring tends to expand object edges. Due to this expansion, the positive elements in the gradient image tend to have more positive neighbors and the negative elements in the gradient image tend to have more negative neighbors along the gradient direction. As a result, the transition probabilities between two positive states and that between two negative states will tend to increase. As the blurriness increases, the same sign transition probabilities also increase.

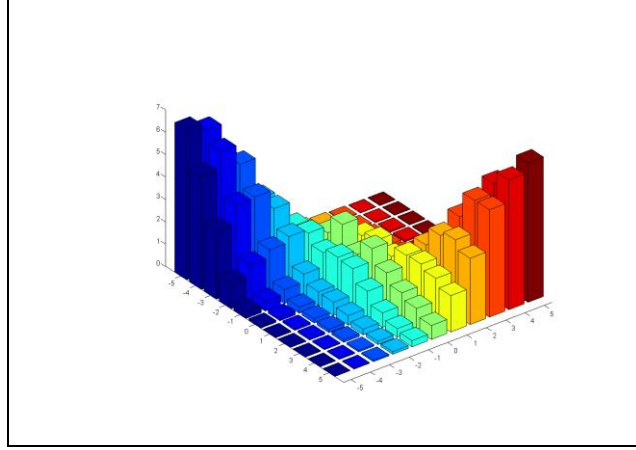


Figure 4. Change between transition probability matrices.

4. MEASURING BLURRINESS

The process of computing the blurriness measure is depicted in Figure 5. The gradient image is calculated, the one-step transition probabilities are estimated for a subset of the transition pairs. Finally, we pool the transition probabilities to formulate a blurriness measure for the given image.

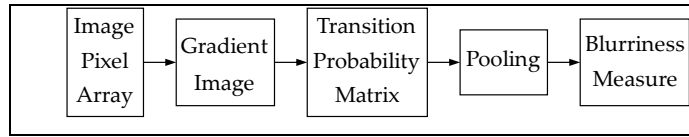


Figure 5. Blurriness measure flow chart.

Given a still image or one frame of a video sequence, which may have one or more than one component, we can process each component in a similar fashion. In this paper, we just describe the process of calculating the blurriness from the luminance component of the given image, which is called the *image pixel array* in Figure 5.

4.1 Gradient Image

Denote the given image pixel array as $I(x, y)$, where $0 \leq x \leq S_x - 1$, $0 \leq y \leq S_y - 1$, and S_x is the width and S_y is the height of the given image. We can calculate the (vertical) gradient image $D(x, y)$ as follows

$$D(x, y) = I(x, y) - I(x, y + 1), \quad (1)$$

where $x \in [0, S_x - 1]$ and $y \in [0, S_y - 2]$

4.2 Transition Probabilities

The transition probability $\Pr_{p \rightarrow q}$ of the gradient image is the probability that the current element equals P while its neighbor below equals Q , where P and Q range from -255 to 255 , in the case that the image pixel bit depth is 8.

This process can be expressed mathematically as

$$\Pr_{P \rightarrow Q} = \Pr(D(x, y+1) = P \mid D(x, y) = Q). \quad (2)$$

Equation (2) can be expanded according to Bayes Law:

$$\Pr_{P \rightarrow Q} = \frac{\Pr(D(x, y) = P, D(x, y+1) = Q)}{\Pr(D(x, y) = P)}. \quad (3)$$

The transition probability $\Pr_{P \rightarrow Q}$ is estimated by

$$\Pr_{P \rightarrow Q} = \frac{\sum_{x=0}^{S_x-1} \sum_{y=0}^{S_y-3} \delta\{D(x, y) = P, D(x, y+1) = Q\}}{\sum_{x=0}^{S_x-1} \sum_{y=0}^{S_y-3} \delta\{D(x, y) = P\}}, \quad (4)$$

where $\delta\{\bullet\}$ is the Kronecker delta function, which is defined as

$$\delta\{A\} = \begin{cases} 1, & \text{while condition A holds} \\ 0, & \text{otherwise} \end{cases}. \quad (5)$$

4.3 Pooling Strategies

In Subsection 3.2 we discussed the fact that the transition probability estimates will only be useful for the more common transitions. As mentioned in the beginning of this section, the transition probabilities are only estimated for a subset of the $(2N+1) \times (2N+1)$ transition pairs. In order to conserve computation, only those transition probabilities used in the pooling will be estimated.

The histograms of gradient images before and after blurring, shown in Figure 3, are both symmetric about zero as are the changes to the transition probability matrices due to blurring as shown in Figure 4.

In order to exploit this symmetry and limit the size of the subset, we select a single pair of values, one value P_0 and one value Q_0 , and calculate the four transition probabilities: $\Pr_{P_0 \rightarrow Q_0}$, $\Pr_{Q_0 \rightarrow P_0}$, $\Pr_{-P_0 \rightarrow -Q_0}$, and $\Pr_{-Q_0 \rightarrow -P_0}$, all of which we expect to increase with blurriness.

These four transition probabilities are pooled to obtain the blurriness measure BLUR as

$$BLUR = \Pr_{P_0 \rightarrow Q_0}^\beta + \Pr_{Q_0 \rightarrow P_0}^\beta + \Pr_{-P_0 \rightarrow -Q_0}^\beta + \Pr_{-Q_0 \rightarrow -P_0}^\beta, \quad (6)$$

where β is a factor experimentally determined to maximize the correlation with subjective quality assessment scores

Note that, as illustrated in Figure 3(b) and Figure 3(c), the histogram counts decrease quickly as the element value increases. Probability estimates will only be useful when the transitions are common which can only occur when the

values themselves are common. Thus, the values of P_0 and Q_0 should be chosen to be relatively small. For example, the histograms of Figure 3 suggest that there would be insufficient data for values larger than about 6.

4.4 Computational Complexity

In Subsection 4.3, we have presented a concrete implementation using only four transition probabilities drawn from the transition probability matrix. There are two approaches to calculating the transition probabilities. One approach is to calculate the nominator and denominator of Equation (4) for each of the four transition probabilities in Equation (6) and then divide the nominator by the corresponding denominator. Another approach is to calculate the nominator of Equation (4) for all the $(2N+1) \times (2N+1)$ transition probabilities, from those nominators calculate the denominator of Equation (4) for each of the four transition probabilities in Equation (6), and then divide the nominator by the denominator accordingly. Although the latter is more costly, given that we only use four transition probabilities in this concrete implementation, it can be easily extended to include more transition probabilities without incurring too much extra cost. Here we analyze the computational complexity of the proposed method based on the second approach, which gives an upper bound of computational cost.

The computational cost mainly consists of three parts. The first part is computation of the gradient: one subtraction per image pixel. The second part is computation of the nominator of Equation (4): one addition per gradient element (or approximately one per image pixel). The third part is computation of the denominator of Equation (4): $2N$ additions per selected (P, Q) pair. After these steps, there is one division, one power operation, and one addition per selected (P, Q) pair in the final pooling. For an $S_x \times S_y$ image with k selected transition pairs, the proposed scheme at most needs $(2 \times S_x \times S_y + (2N + 1) \times k)$ additions/subtractions, k divisions, and k power operations.

Compared to those methods requiring edge detection, which is much more expensive, the proposed blurriness measurement method is computationally economic, which makes it promising for real-time applications.

5. EXPERIMENTS AND RESULTS ON STILL IMAGES

The proposed blurriness measurement approach has been implemented, applied to still images, and compared to a number of other algorithms from the literature.

5.1 Image Dataset

Still image experiments were conducted using the Gaussian blurred image dataset and the JPEG2000 image dataset from the LIVE image database release 2 (2005) [12] [13] [14]. The Gaussian blurred image dataset includes 174 Gaussian low-pass filtered or reference images and the JPEG2000 image dataset includes 227 JPEG2000 distorted or reference images. The datasets include subjective DMOS scores for each distorted image.

5.2 Accuracy Criteria

The accuracy of the objective blur measurements are quantified using three different correlation coefficients and root mean square error (RMSE).

The best known correlation coefficient is the Pearson's product-moment correlation coefficient (PCC), which is obtained by dividing the covariance of the two variables X and Y by the product of their standard deviations, as given by ([15])

$$r_p = \frac{n \sum_i X_i Y_i - (\sum_i X_i)(\sum_i Y_i)}{\sqrt{n \sum_i X_i^2 - (\sum_i X_i)^2} \sqrt{n \sum_i Y_i^2 - (\sum_i Y_i)^2}}, \quad (7)$$

where n is the number of elements in each data set (same for both input datasets). In our experiments, X is a vector of subjective scores and Y is a vector of objective scores

The PCC is computed both before and after performing a non-linear regression. For the former, the PCC is computed between the non-fitted objective scores and the subjective DMOS. For the latter, we use a logistic function and follow

the procedure outlined in [16] (pp 32) to perform the non-linear regression. Thereafter, the PCC (denoted as PCC-f in this paper) is computed between the fitted objective scores and the subjective DMOS.

In statistics, Spearman's rank correlation coefficient [17], or SROCC in short, is a non-parametric measure of correlation. It assesses how well an arbitrary monotonic function could describe the relationship between two variables, without making any other assumptions about the particular nature of the relationship between the variables. In principle, SROCC is simply a special case of the PCC, in which two sets of data X and Y are converted to rankings x and y before calculating the coefficient.

In this calculation, one has to assign the same rank to each of the tied values. It is an average of their positions in the ascending order of the values.

The RMSE is calculated between the fitted objective score (SO) and the subjective score (SSO as

$$R_{MSE} = \sqrt{\frac{1}{n} \sum_i (SS_i - SO_i)^2}. \quad (8)$$

5.3 Accuracy Testing Results on Gaussian Blurred Images

The accuracy testing results of the proposed approach on the Gaussian blurred image dataset are summarized in **Table 1**. For this test the parameters were set as follows: $P_0=4$ and $Q_0=3$. In order to select β , a set of values were considered and 10 trials were performed. For each trial, half of the images in the dataset were randomly chosen and each value in the set was used to calculate a blur score. The value that yielded the best PCC and SROCC was recorded. After all 10 trials, the 10 recorded values of β were averaged, yielding a final $\beta=0.869$.

Table 1. Accuracy testing results on Gaussian blurred images

Algorithm	Accuracy Metric			
	PCC	PCC-f	SROCC	RMSE
Marziliano et al.'s	0.8048	0.9286	0.9383	8.0740
Proposed approach	0.8931	0.9295	0.9370	8.0236

For comparison, experimental results on the same image dataset using Marziliano et al.'s algorithm [2] are also given in **Table 1**. The proposed approach is slightly better as measured by PCC and RMSE and Marziliano et al.'s is slightly better as measured by SROCC

5.4 Results on JPEG2000 Images

The accuracy testing results of the above two algorithms on the JPEG2000 image dataset are given in **Table 2**. For the proposed approach, we selected $P_0=4$, $Q_0=3$, and $\beta = 0.467$.

Table 2. Accuracy testing results on JPEG2000 images

Algorithm	Accuracy Metric			
	PCC	PCC-f	SROCC	RMSE
Marziliano et al.'s	0.6792	0.9311	0.9422	8.9083
Proposed approach	0.9481	0.9516	0.9396	7.4994

From in **Table 2**, one can observe the proposed approach is better as measured by PCC and RMSE and Marziliano et al.'s is slightly better as measured by SROCC

5.5 Results on Combined Image Dataset

In practice, often people do not have a priori information of the processing history the given image has been undergone. We thus conducted experiments to test these algorithms on the combined image dataset, i.e., we put all the images from the Gaussian blurred image dataset and those from the JPEG2000 image dataset into one image dataset. The accuracy testing results of these two algorithms on this combined image dataset are given in **Table 3**. For the proposed approach, we selected $P_0=4$, $Q_0=3$, and $\beta = 0.653$. The proposed approach has outperformed Marziliano et al.'s algorithm on the combined image dataset.

Table 3. Accuracy testing results on combined image dataset

Algorithm	Accuracy Metric			
	PCC	PCC-f	SROCC	RMSE
Marziliano et al.'s	0.6733	0.8688	0.8676	11.5447
Proposed approach	0.9168	0.9386	0.9319	8.0424

5.6 Monotonicity Test

We also conducted experiments to test the monotonicity of the above algorithms on images. By filtering the image in Figure 2(a) using Gaussian low-pass filter with different standard deviation σ , specifically, $\sigma = 0.4, 0.8, 1.2, 1.6,$ and 2.0 , we obtained 5 images with different level of Gaussian blurriness distortion. Similarly, by compressing the image in Figure 2(a) using JPEG2000 with different compression ratio (CR), specifically, $CR = 40, 80, 120, 160,$ and 200 , we obtained 5 images with different level of JPEG2000 blurriness distortion. In Figure 6(a) and Figure 6(b), we draw the plot of the blurriness measure versus the standard deviation (for Gaussian blur) and the plot of the blurriness measure versus the compression ratio (for JPEG2000), respectively. The blurriness measure has been scaled for illustration purpose. It is seen that both of the above two algorithms give blurriness measure in general monotonically increasing with the CR and/or σ .

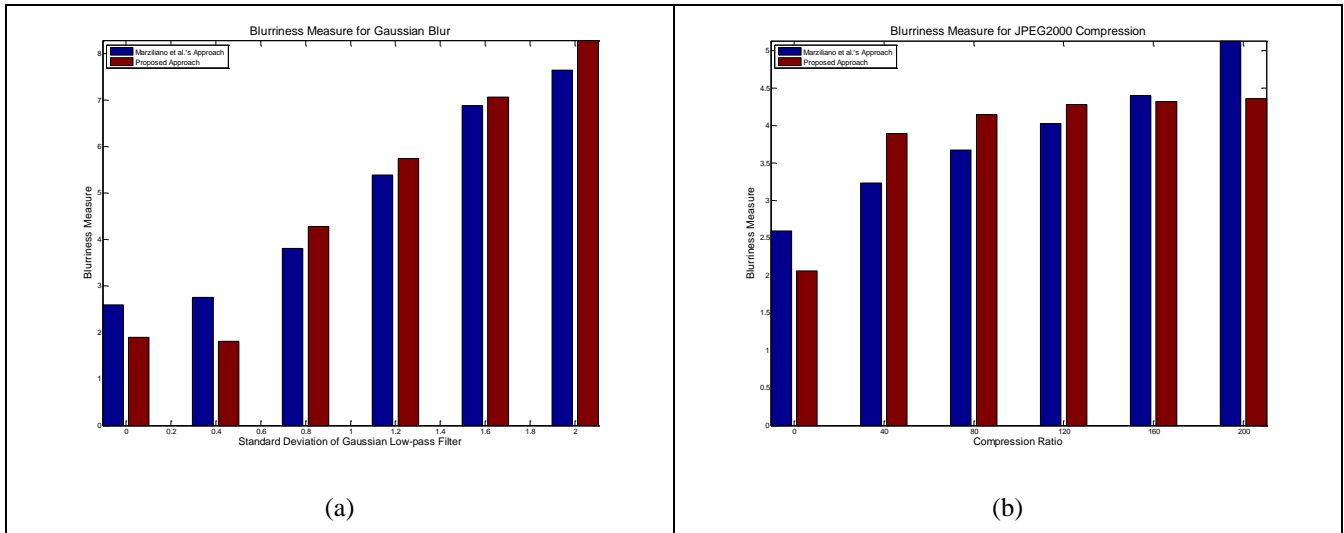


Figure 6. Blurriness measure versus compression ratio.

6. EXPERIMENTS AND RESULTS ON VIDEOS

The proposed approach can be applied to video sequences on a frame-by-frame basis to measure the blurriness in the video sequence. To do this, the blurriness score for a sequence is defined as the mean value of the blurriness measures over all the video frames in the sequence. We use blurriness introduced by H.264 compression.

6.1 Video Dataset

The dataset for this experiment comes from the LIVE video quality database [18] [19]. There are ten video sequences, each with spatial resolution of 768×432. For each sequence there are 5 versions: a reference version and 4 compressed versions encoded at different bitrates. For each of the 40 compressed sequences, the database has subjective DMOS scores

6.2 Accuracy Test

The accuracy testing results of the proposed approach are shown in **Table 4**. In our experiments, we select $P_0=3$ and $Q_0=2$. The parameter β is selected as 2.331. The process of determining β is the same as has been described in Section 5.

Table 4. Accuracy testing results on H.264 video

Algorithm	Accuracy Metric			
	PCC (nonfitted)	PCC (fitted)	SROCC	RMSE
Marziliano et al.'s	0.5032	0.9041	0.6076	9.5860
Liu et al.'s	0.8151	0.9199	0.7521	8.7998
Proposed approach	0.8484	0.9470	0.8396	7.2094

For comparison, experimental results on the same video sequences using Marziliano et al.'s algorithm [2] and Liu et al.'s algorithm [7] are also given in **Table 4**. The proposed approach outperforms both of these algorithms.

6.3 Monotonicity Test

We also conducted experiments to test the monotonicity of the above algorithms. By compressing the image in Figure 2(a) using H.264 encoder with different quantization parameter (QP), specifically, $QP = 15, 21, 27, 33, 39,$ and 45 , we obtained 6 images with different level of blurriness distortion. In Figure 7, we draw the plot of the blurriness measure versus the quantization parameter. The blurriness measure has been scaled for illustration purpose. It is seen that all of the above three algorithms give blurriness measure in general monotonically increasing with the quantization parameter QP.

7. DISCUSSION AND FUTURE WORK

7.1 Utilizing Chrominance Components

All of the experimental studies presented in this paper were conducted on the luminance component of the image. In practice, color distortion may introduce blur to images. Future work examines methods to incorporate measurement of blur in the chrominance components into an overall blur measure.

7.2 Introducing Machine Learning Approaches

The current implementation performs a simple pooling of four transition probabilities to generate a blurriness measure. With the aid of feature selection tools, we may include all the transition probabilities and utilize machine learning approaches, specifically, regression, to select more transition probabilities in order to improve the accuracy of blurriness measure.

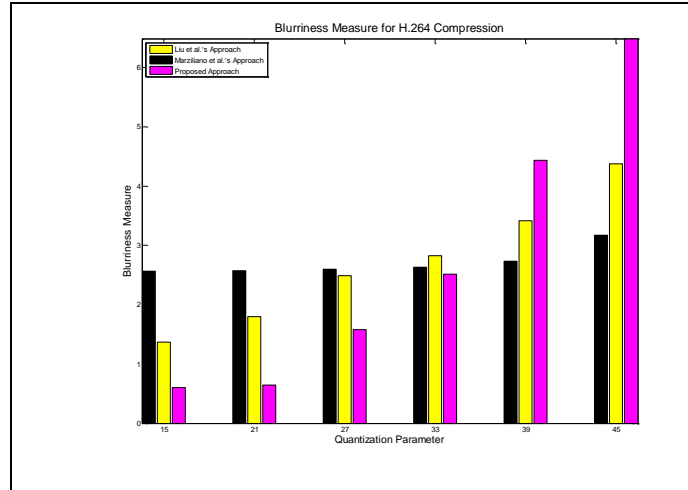


Figure 7. Blurriness measure versus quality parameter.

7.3 Incorporating Points of Interests

In this paper, a few selected transition probabilities are pooled to give a blurriness measure for the image. Transition probabilities are global statistics of the image. However, in perceptual quality assessment, salient points, or points of interest, are often more important than other areas in an image. Thus, it would be interesting to weight blurriness in areas of interest more heavily in a final blurriness measure.

7.4 Extending from Image to Video

We have proposed a blurriness measurement method for images in this paper and extended it to measure blurriness in video sequences on a frame-by-frame basis. We have observed that the accuracy testing results of the proposed approach on video sequences is not as good as that on still images. This is because motion and other temporal properties play a key role in the perception of blur. Future work investigates some of these temporal properties, how they influence subjective video quality assessment, and how they can be properly accounted for in a video blur measure.

8. CONCLUSION

In this paper, a novel blurriness measurement scheme has been presented. This measure utilizes the transition probabilities of the gradient image representing the first time that transition probabilities have been utilized to measure perceptual quality. The proposed method makes no assumptions about the source of the blurriness and can be used for measuring blurriness in image or video without a reference signal. Experimental studies have shown that the proposed method performs as well as or better than some of the state-of-the-art reference-free blurriness measurement algorithms in terms of correlation to the subjective score, on data from two publicly available image/video databases. With relatively low computational complexity, the proposed scheme is promising for real-time image and video blurriness measurement.

REFERENCES

1. <http://www.its.bldrdoc.gov/vqeg/>.
2. P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," *IEEE International Conference on Image Processing*, Rochester, New York, USA, September 2002.
3. E. Ong, W. Liu, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, and F. Moschetti, "A no-reference quality metric for measuring image blur," *IEEE Seventh International Symposium on Signal Processing and its Applications*, Paris, France, July 2003.

4. R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on just-noticeable blur and probability summation," *IEEE International Conference on Image Processing*, San Antonio, Texas, USA, September 2007.
5. S. Varadarajan and L. J. Karam, "An improved perception-based no-reference objective image sharpness metric using iterative edge refinement," *IEEE International Conference on Image Processing*, San Diego, California, USA, October 2008.
6. J. Caviedes and S. Gurbuz, "No-reference sharpness metric based on local edge kurtosis," *IEEE International Conference on Image Processing*, Rochester, New York, USA, September 2002.
7. D. Liu, Z. Chen, H. Ma, F. Xu, and X. Gu, "No reference block based blur detection," *IEEE First International Workshop on Quality of Multimedia Experience*, San Diego, California, USA, July 2009.
8. M. Chen and A. C. Bovik, "No-reference image blur assessment using multiscale gradient," *IEEE First International Workshop on Quality of Multimedia Experience*, San Diego, California, USA, July 2009.
9. R. Liu, Z. Li, and J. Jia, "Image partial blur detection and classification," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Anchorage, Alaska, USA, June 2008.
10. ISO 12640: Standard color image data (SCID).
11. C. M. Grinstead and J. L. Snell, *Introduction to probability*, second edition, American Mathematical Society, July 1997.
12. H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2". Available online: <http://live.ece.utexas.edu/research/quality>.
13. H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms", *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3440-3451, November 2006.
14. Z. Wang, A.C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol.13, no.4, pp. 600- 612, April 2004.
15. J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician*, vol 42, no 1, pp 59-66, February 1988.
16. VQEG, "Final report from the video quality experts group on the validation of objective quality metrics for video quality assessment". http://www.its.bldrdoc.gov/vqeg/projects/frtv_phase1.
17. J. L. Myers and A. D. Well, *Research Design and Statistical Analysis*, second edition, Routledge, November 2002.
18. K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video", in press, *IEEE Transactions on Image Processing*, 2010.
19. 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*, January 2010.