A NO-REFERENCE OBJECTIVE IMAGE SHARPNESS METRIC BASED ON JUST-NOTICEABLE BLUR AND PROBABILITY SUMMATION

Rony Ferzli and Lina J. Karam

Department of Electrical Engineering
Arizona State University
Tempe, AZ 85287-5706
rferzli@ieee.org, karam@asu.edu

ABSTRACT

This work presents a perceptual-based no-reference objective image sharpness/blurriness metric by integrating the concept of Just Noticeable Blur (JNB) into a probability summation model. Unlike existing objective no-reference image sharpness/blurriness metrics, the proposed metric is able to predict the relative amount of blurriness in images with different content. Results are provided to illustrate the performance of the proposed perceptual-based sharpness metric. These results show that the proposed sharpness metric correlates well with the perceived sharpness.

***Index Terms*—** Image quality, Image Assessment, Perception, HVS, No-reference, Objective, Sharpness Metric.

1. INTRODUCTION

Measuring the quality of images is desirable in various applications such as video compression and image enhancement. Researchers have had to resort to subjective viewing experiments in order to obtain reliable ratings for the quality of digital images or video. While these tests are the best way to measure 'true' perceived quality, they are complex, time-consuming and consequently expensive. Hence, they are often impractical or not feasible at all, for example when real-time online quality monitoring of several video channels is desired. It is desirable to design objective metrics that can predict accurately the quality of the image/video and can be embedded in real-time systems at low cost. Objective full-reference metrics require the original image for calculation while the no-reference ones are independent of the original image. Different types of impairments may exist in an image; for example, image encoders introduce mostly blockiness, blurriness and ringing artifacts while noise is present due to sensor and when transmitting over communication channels. This paper focuses on no-reference image sharpness/blurriness metrics due to their importance in image and video compression, enhancement algorithms, as well as in biomedical applications. Sharpness metrics rely on the fact that edges are steeper for sharp images while they tend to be smooth for blurred images.

Several no-reference objective sharpness metrics were proposed in the literature and are analyzed in [1]. Most of previously proposed objective no-reference metrics neglect the important influence of image content and viewing conditions on the actual visibility of artifacts. Therefore, their predictions often do not agree well with actual perceived quality [1]. So, there is a need for reliable vision models to be incorporated into image and video processing algorithms. The improvement in quality that can be achieved using an approach that exploits the Human Visual System (HVS) characteristics can be significant in a variety of image and video processing applications. Nevertheless, developed perceptually-motivated sharpness metrics [2] concentrate on predicting the quality of images having same content and fail when exposed to images having different content [3]. In this paper, the evaluation of quality for images with various contents, through the evaluation of sharpness, is investigated and an HVS-based metric is proposed based on the concepts of Just Noticeable Blur (JNB) and probability summation over space.

This paper is organized as follows. Section 2 gives a brief overview of available objective no-reference sharpness metrics and the behavior of these metrics when applied to images with different content. Section 3 presents a JNB-based perceptual sharpness metric based on the probability of summation over space. Performance results are presented in Section 4. A conclusion is given in Section 5.

2. BEHAVIOR OF EXISTING SHARPNESS METRICS

A detailed overview of different objective no-reference sharpness metrics was given by the authors in [1]. Table 1 summarizes these methods along with a brief description. These metrics are applied to a testing set consisting of four 512×512 different images having different content and blurred using a 7×7 Gaussian filter with a standard deviation equals to 0.8, 1.6, 2.0 and 2.4, respectively, as shown in Fig. 1. The idea is to test the metrics on a set of images having different characteristics in order to check if these metrics can reliably predict the relative sharpness in images with different content. For example, the 'Peppers' image has large smooth regions while the 'Houses' image contains a lot of edges. Texture is found abundantly in the 'Man' image. The 'Fishingboat' image contains both smooth and high variation areas. The existing metrics (Table 1) were applied to the testing set of Fig. 1 in the following order: FishingBoat ($\sigma_{blur} = 0.8$), Man ($\sigma_{blur} = 1.6$), Peppers ($\sigma_{blur} = 2.0$), and Houses ($\sigma_{blur} = 2.4$). The desired response should be a monotonic decreasing curve for the corresponding sharpness values since increasing blurriness should result in a lower value of the considered sharpness metric (in the case of a blurriness metric, the inverse of the metric value was used to obtain a measure of sharpness). Simulation results [3] indicate that none of the listed metrics can predict correctly the perceived sharpness. In [3], the authors proposed a perceptually-motivated approach for predicting the relative sharpness of images with different content. However, in [3], derived perceptual weights were applied in a heuristic manner. In this work, a perceptual-based no-reference image sharpness assessment method is proposed and justified based on probability summation over space [4, 5], which gives a solid basis for the adopted approach.

(a) 'Fishingboat' ($\sigma_{blur} = 0.8$) (b) 'Man' image ($\sigma_{blur} = 1.6$) (c) 'Peppers' image ($\sigma_{blur} = 2.0$) (d)'Houses' image ($\sigma_{blur} = 2.4$)

Fig. 1. Test images with different content.

Table 1. Existing objective no-reference sharpness metrics.

Noise Immune Metric (NIS) [1]: relies on the Lipschitz regularity property separating the signal singularities from the noise singularities, by applying the dyadic wavelet transform and then measuring the sharpness using the perceptual blur metric [2].
Variance metric [6]: calculates the variance of the whole image.
Autocorrelation-based metric [7]: derived from the auto-correlation function which uses the difference between auto-correlation values at two different distances along the horizontal and vertical directions, respectively.
Derivative-based metrics [8]: include the first-order (gradient) and second-order (Laplacian) derivatives metrics; these metrics act as a high-pass filter in the frequency domain.
Perceptual blur metric [2]: the overall metric is calculated as the average of the edge widths or the local blur values over all edges.
Frequency threshold metric [9]: computes the summation of all frequency component magnitudes above a manually selected threshold.
Kurtosis metric [10, 11]: uses the statistical kurtosis to measure of the peakedness or flatness of a distribution in frequency domain.
Histogram threshold metric [9]: defined as the weighted sum of the histogram bin values above a certain threshold 'T'.
Histogram entropy based metric [12]: the probabilities are calculated by normalizing the obtained histogram, and the entropy is computed.
Histogram frequency based metric [13]: based on the occurrence histogram of non-zero DCT coefficients throughout all 8×8 blocks of the image.

3. PROPOSED SHARPNESS METRIC BASED ON PROBABILITY SUMMATION

This section presents the proposed no-reference objective sharpness metric integrating the concept of Just Noticeable Blur (JNB) into a probability summation model.

3.1. Subjective Blur Detection Experiments

In order to study the response of the HVS to blurriness and sharpness in images, subjective experiments were performed to obtain results in relation to blur perception and just-noticeable blurs (JNBs). The concept of JNB was previously introduced by the authors in [3]. The blurriness is introduced using a 7×7 Gaussian lowpass filter mask. The conducted experiments make use of a foreground square with uniform intensity I_F over a uniform background with intensity I_B. For a given contrast ratio, the foreground square is blurred using a guassian mask with a standard deviation equals to $\{0.2, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.65, 0.8, 1.0\}$. The standard deviation values are selected based on previous experiments conducted by the authors in [3] where the subjects can increase the Gaussian mask variance until detecting blurriness labeled as JNB. In this work, for a given contrast, which should be greater than the Just-Noticeable-Distortion (JND) threshold, the 10 levels of blurriness are displayed randomly to the human observer one after the other. In contrast to the blur detection experiments previously conducted by the authors in [3], in the experiments reported here, the subject cannot control the amount of blurriness. Once exposed with a blurriness level, the subject should select one of two answers: 'Detected' or 'Not Detected'. Each time an answer is received from the subject, another level of blurriness is displayed. The same experiment is repeated for different contrast ratios. Overall, the subject is exposed to 20 different contrast ratios starting at 10 and reaching 200 using a step size of 10. For each contrast ratio, the corresponding collected data is used to compute the σ_{JNB} threshold at the considered contrast. This is done by computing the normalized histogram of the subject responses and selecting the blur detection point corresponding to a probability of detection of 63% [5]. For each blur detection threshold σ_{JNB}, the corresponding edge width is measured and denoted as w_{JNB}.

3.2. Perceptual Blur Detection Model Based on Probability Summation

While the derived JNB thresholds (Section 3.1) provide a localized measure of the blur threshold for a single edge at a given local contrast, a perceptual sharpness metric that also accounts for spatial summation of individual blur distortions is needed. In this work, the probability summation model is adopted [5]. The proposed probability summation model considers a set of independent detectors, one at each edge location e_i. The probability $P(e_i)$ of detecting a blur distortion is the probability that a detector at edge pixel 'i' will signal the occurrence of a blur distortion. $P(e_i)$ is determined by the psychometric function, which is modeled as an exponential having the following form [5]:

$$P(e_i) = 1 - \exp\left(-\left|\frac{w(e_i)}{w_{JNB}(e_i)}\right|^{\beta}\right) \quad (1)$$

where $w(e_i)$ is the measured width of the edge e_i and $w_{JNB}(e_i)$ is the JNB width (Section 3.1) which depends on the local contrast in the neighborhood of edge e_i. The value of β is chosen to increase the correspondence of (1) with the experimentally determined psychometric function for a given type of distortion (blur in our case). From (1), note that a probability of detection of 63% is obtained when the measured width is equal to w_{JNB} as desired.

A less localized probability of error detection can be computed by adopting the probability summation hypothesis which pools the localized detection probabilities over a region of interest $P(e_i)$ over a region of interest R [5]. The probability summation hypothesis

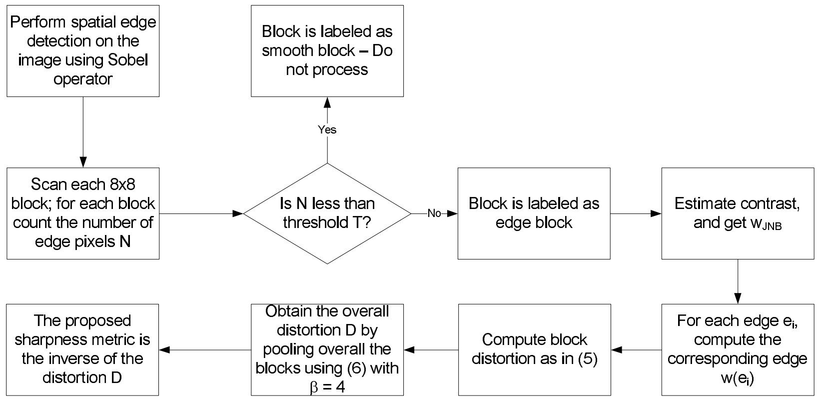


Fig. 2. Flowchart illustrating the computation of the proposed perceptual-based sharpness metric.

is based on the following two assumptions: 1) a blur distortion is detected if and only if at least one detector in R senses the presence of a blur distortion, and 2) the probabilities of detection $P(e_i)$ are independent. The probability of detecting blur in a region R is then given by:

$$P_{blur}(R) = 1 - \prod_{e_i \in R} (1 - P(e_i)) \tag{2}$$

Substituting (1) into (2) yields:

$$P_{blur}(R) = 1 - \exp\left(-D^{\beta}_{(R)}\right) \tag{3}$$

where

$$D_{(R)} = \left(\sum\nolimits_{e_i \in R} \left| \frac{w(e_i)}{w_{JNB}(e_i)} \right|^{\beta} \right)^{\frac{1}{\beta}} \tag{4}$$

In (4), $D_{(R)}$ takes the form of a Minkowski metric with exponent β. From (3), it is evident that a lower $D_{(R)}$ results in a lower probability of blur detection $P_{blur}(R)$. So, $D_{(R)}$ can be used to indicate the amount of perceived blurriness in the considered region R.

In the human visual system, highest visual acuity is limited to the size of the foveal region, which covers approximately 2% of visual angle. In this work, the foveal region is approximated by 8×8 image blocks. Smooth blocks are excluded as they do not contribute to blur perception. For this purpose, a Sobel edge detector is run first on each block, and each block is categorized as a smooth block or an edge block based on the number of edge pixels. In order to account for the maximum perceived blur distortion within a non-smooth block R_b, a $\beta = \infty$ is used in (4) obtaining a maximum probability of blur detection in the block region R_b as follows:

$$D_{R_b} = \max_{e_i \in R_b} \left(\left| \frac{w(e_i)}{w_{JNB}} \right| \right) \tag{5}$$

where w_{JNB} is the JNB width corresponding to the contrast of the considered block region R_b and is obtained as described in Section 3.1.

The perceived blur distortion measure D for the whole image corresponds to the probability of detecting a blur distortion over all possible block regions R_b and is obtained by using a Minkowski metric as follows:

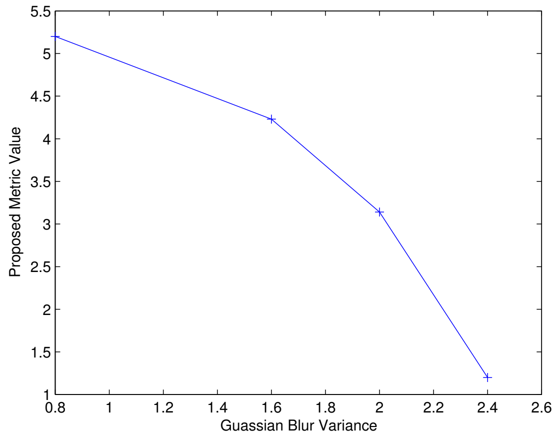
$$D = \left(\sum_{R_b} |D_{R_b}|^{\beta} \right)^{\frac{1}{\beta}} \tag{6}$$

In (6), $\beta = 4$ was found to correlate well with subjective tests. The resulting blur distortion measure D of (6), normalized by the image size, is adopted as the proposed no-reference objective blurriness metric. The proposed no-reference objective sharpness metric is thus taken to be $1/D$. A block diagram summarizing the computation of the proposed sharpness metric is given in Fig. 2.

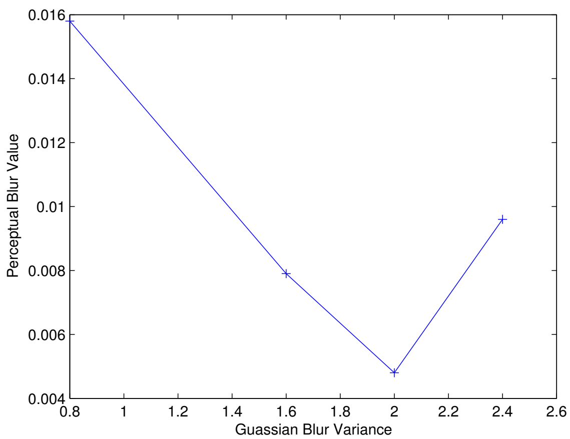
4. SIMULATION RESULTS

In this section, results are presented to illustrate the performance of the proposed JNB-based sharpness metric. Fig. 3(a) illustrates the performance of the proposed perceptual-based sharpness metric when applied to the testing set of Fig. 1. For comparison, Fig. 3(b) shows the performance results when the popular non-perceptually weighted edge-based sharpness metric of Marziliano et al. [2] is applied to the same testing set. Note that the sharpness metric of [2] is also referred to as the perceptual blur metric in the literature but it does not incorporate any perceptual weighting and does not incorporate visual subjective test data. As the blurriness increases, the sharpness metric should decrease monotonically. From Fig. 3, it can be seen that the proposed perceptual JNB-based sharpness metric is decreasing as expected, while the perceptual blur metric of [2] is failing. Also, as indicated in Section 2, all the other existing sharpness metric (Table 1) fail to predict correctly the increase in blurriness in images with different content.

In order to further validate the proposed JNB-based objective no-reference sharpness metric, subjective tests were conducted for assessing the blur in different images. The subjective blur assessment were conducted as follows: 18 images are extracted from the LIVE database [14]. Each image is presented 4 times giving a total of 72 images. Note that the blurred images in the database are generated using a circular-symmetric 2-D Gaussian kernel of standard deviation ranging from 0 to 15. The images are randomly displayed; for each displayed image, the subject is asked to rate the quality of the image in terms of perceived blurriness using a scale from 1 to 5 correpsonding to “Very annoying”, “Annoying”', “Slightly Annoying”, “Perceptible but not annoying”, and “Imperceptible”, respectively. Nine subjects took the test and the Mean Opinion Score (MOS) was computed and compared to the proposed sharpness metric. To measure how well the proposed metric values correlate with the subjective MOS values, the Pearson coefficient [15] was computed. According to [15], a Pearson coefficient value higher than 0.75 indicates well correlated data, while a value higher than 0.9



(a) Performance of proposed metric.



(b)Performance of perceptual blur metric of [2].

Fig. 3. Performance of the proposed perceptual-based sharpness metric and the perceptual blur metric of [2] when applied to the image testing set of Fig. 1.

shows a very strong relation. In our case, the computed Pearson coefficient value is 0.908 showing that the proposed metric is strongly correlated with the subjective MOS values.

5. CONCLUSION

Simulation results showed that none of the existing no-reference objective sharpness metrics will give satisfying results when applied to images with different scenes. A perceptual sharpness metric is derived based on measured Just-Noticeable Blurs (JNBs) and probability summation over space, which takes into account the response of the HVS to sharpness at different contrast levels. Combining the derived model with local image features, it is shown that the proposed metric is able to successfully predict the relative sharpness/blurriness of images, including those with different scenes. Future directions include investigating the effect of color on sharpness perception and incorporating the JNB concept into the noise-immune sharpness metric [1] .

6. REFERENCES

[1] R. Ferzli and L. J. Karam, "No-reference objective wavelet based noise immune image sharpness metric," in *IEEE International Conference on Image Processing*, Sep. 2005, pp. 405–408.

[2] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Applications to JPEG2000," *Signal Proc.: Image Comm.*, vol. 19, pp. 163–172, 2004.

[3] R. Ferzli and L. J. Karam, "Human visual system based on no-reference objective image sharpness metric," in *IEEE International Conference on Image Processing*, Oct. 2006, pp. 2949–2952.

[4] R. R. Quick, "A vector-magnitude model of contrast detection," *Kybernetik*, vol. 16, pp. 65–67, 1974.

[5] J. G. Robson and N. Graham, "Probability summation and regional variation in contrast sensitivity across the visual field," *Vision Research*, vol. 21, no. 3, pp. 409–418, 1981.

[6] S.J. Erasmus and K.C.A. Smith, "An automatic focusing and astigmatism correction system for the sem and ctem," *Journal of Microscopy*, , no. 127, pp. 185–199, 1982.

[7] C. F. Batten, "Autofocusing and astigmatism correction in the scanning electron microscope," *MPhil Thesis, University of Cambridge*, 2000.

[8] Z. Wang, H. R. Sheikh, and A. C. Bovik, ," in *Handbook of image and video processing, Academic Press*, 2000.

[9] L. Firestone, K. Cook, N. Talsania, and K. Preston, "Comparison of autofocus methods for automated microscopy," *Cytometry*, vol. 12, pp. 195–206, 1991.

[10] N. Zhang, A.E. Vladar, M.T. Postek, and B. Larrabee, "A kurtosis-based statistitcal measure for two-dimensional processes and its application to image sharpness," *Proceedings of Section of Physical and Engineering Sciences of American Statistical Society*, pp. 4730–4736, 2003.

[11] J. Caviedes and F. Oberti, "A new sharpness metric based on local kurtosis, edge and energy information," *Signal processing: Image Communication*, vol. 19, no. 2, pp. 147–161, 2004.

[12] N. Chern, P. A. Neow, and H. Marcelo, "Practical issues in pixel-based autofucusing for machine vision," in *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, 2001.

[13] X. Marichal, W.Y. Ma, and H. J. Zhang, "Blur determination in the compressed domain using dct information," *International Conference on Image Processing*, 1999.

[14] H. R. Sheikh, A. C. Bovik, L. Cormack, and Z. Wang, "LIVE image quality assessment database," http://live.ece.utexas.edu/research/quality, 2003.

[15] D. Freedman, R. Pisani, and R. Purves, *Statistics*, 3rd edition, W. W. Norton and Company, 1997.