Correspondence

A No-Reference Image Blur Metric Based on the Cumulative Probability of Blur Detection (CPBD)

Niranjan D. Narvekar and Lina J. Karam, *Senior Member, IEEE*

***Abstract*—This paper presents a no-reference image blur metric that is based on the study of human blur perception for varying contrast values. The metric utilizes a probabilistic model to estimate the probability of detecting blur at each edge in the image, and then the information is pooled by computing the cumulative probability of blur detection (CPBD). The performance of the metric is demonstated by comparing it with existing no-reference sharpness/blurriness metrics for various publicly available image databases.**

***Index Terms*—Blur detection, blur metric, no-reference, objective, perceptual, sharpness metric, visual quality.**

I. Introduction

Multimedia content has become a very popular means of entertainment and communication. This has resulted in great advancements in compression and transmission techniques. However, impairments are often introduced along the several stages of processing and communication [1]. The visibility of these impairments have a drastic effect on the Quality of Experience (QoE) of the consumers. Hence, quality assessment techniques are necessary in order to measure the perceived quality of the multimedia content [1].

Owing to the cost, complexity, and infeasibility of real-time applications associated with subjective quality assessment techniques, it has become important to develop reliable objective quality metrics. Hence, recently, there has been an increasing interest from the research community and also industry towards developing objective quality assessment techniques for multimedia applications and products [1].

Based on the amount of reference information needed, the objective quality assessment metrics can be classified into full-reference, reduced-reference and no-reference metrics [1]. To obtain a quality score, full-reference metrics utilize the original reference information and reduced-reference metrics utilize features extracted from the original reference information. In contrast, no-reference metrics do not require any reference information and are the most useful in applications where the reference is not available. However, they are equally challenging to design.

This paper deals with no-reference image quality assessment targeted towards blur distortions. Blurring occurs in an image due to the loss of high frequency information which could be caused due to various factors, such as acquisition, processing and compression. The sharpness/blurriness metric can also be combined with other metrics to assess the overall quality of images. Several objective no-reference sharpness/blurriness metrics have been proposed in the literature and are analyzed in [2]. In [2], it was shown that the existing blur metrics cannot predict well the relative blurriness in images with different contents. The metric proposed in [2] could predict the relative blurriness in images better than existing blur metrics, but it does not correlate well with images having nonuniform saliency content. The metric proposed in [3] tries to improve the performance of [2] by incorporating a visual attention model, such that the areas in the images which are most likely noticed by humans are given more weight than the others. However, the metric proposed in [3] does not achieve significant improvement over the method in [2]. More recently, Hassen *et al.* [4] proposed a multiscale sharpness metric based on the local phase coherence (LPC) of complex wavelet coefficients. The computed LPC values are sorted and a weighted averaging method is used to obtain a single sharpness index. The weights are selected such that a higher weight is assigned to the higher LPC values which correspond to sharper image regions. This is done to give emphasis to the sharpest regions in the image in order to consider the situations, for example, when the foreground is sharp and the background is blurred.

In this paper, we build on the work of [2] and propose an improved no-reference blur metric which utilizes the concept of just noticeable blur (JNB) together with a cumulative probability of blur detection (CPBD). A probabilistic framework is developed based on the sensitivity of human blur perception at different contrasts. Utilizing this framework, the probability of detecting blur at each edge in an image is estimated. The blur perception information at each edge is then pooled over the entire image to obtain a final quality score by evaluating the cumulative probability of blur detection (CPBD).

The paper is organized as follows. Section II describes the proposed no-reference blur metric. Performance results are presented in Section III. A conclusion is given in Section IV.

II. Proposed No-Reference Objective Blur Metric

This section describes the proposed no-reference objective blurriness metric. The metric proposed in this work is based on the cumulative probability of blur detection (CPBD) and, as shown in Section III, it exhibits consistently a good performance across blur types (Gaussian blur and JPEG2000 blur) and across databases as compared to existing sharpness/blur metrics.

As discussed in [2], for a given contrast C, the probability of detecting blur takes the form of a psychometric function which is modeled as an exponential given by

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

where $w(e_i)$ is the measured width of the edge e_i, $w_{JNB}(e_i)$ is the "just noticeable blur" (JNB) width, which depends on the local contrast C in the neighborhood of edge e_i, and β is a parameter whose value is obtained by means of least squares fitting. The JNB width w_{JNB} at various contrasts can be modeled as [2]:

$$w_{\mathrm{JNB}} = \begin{cases} 5, & \text{if } C \leq 50 \\ 3, & \text{if } C \geq 51. \end{cases} \quad (2)$$

Manuscript received January 19, 2010; revised September 15, 2010; accepted February 06, 2011. Date of publication March 28, 2011; date of current version August 19, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alex ChiChung Kot.

N. D. Narvekar is with the School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706 USA, and also with Dialogic Research Inc., Eatontown, NJ 07724 USA (e-mail: nnarveka@asu.edu; Niranjan.Narvekar@dialogic.com).

L. J. Karam is with the School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706 USA (e-mail: karam@asu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2011.2131660

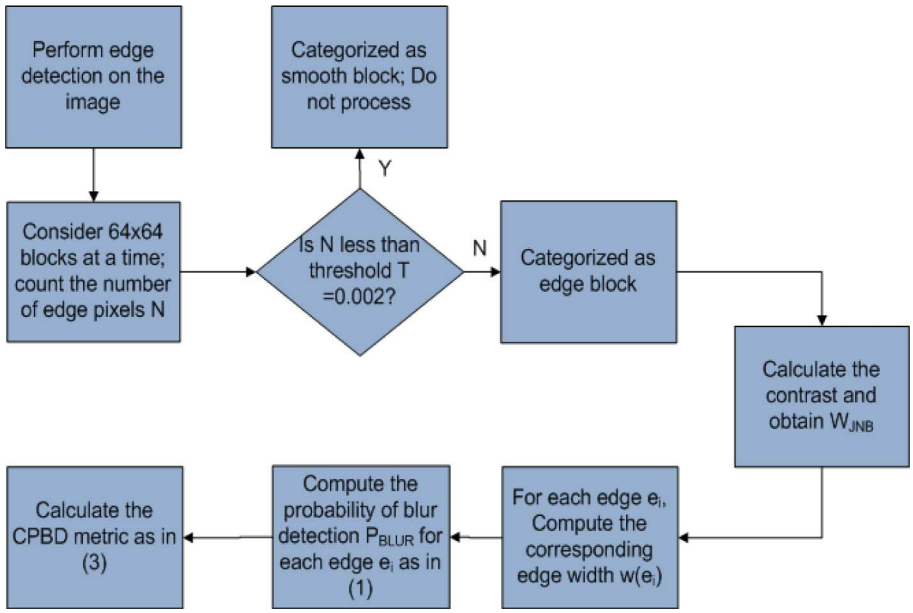

Fig. 1. Block diagram summarizing the computation of the proposed CPBD blur metric.

In (1), at the the JNB, $w(e_i) = w_{\mathrm{JNB}}(e_i)$, which corresponds to the probability of blur detection $P_{\mathrm{BLUR}} = P_{\mathrm{JNB}} = 63\%$.

Natural images consist of either flat regions with uniform intensity values or regions containing edges and texture information. When humans view an image, the information present in the image is pooled in a certain manner such that they can come up with an overall perception of its quality. Equation (1) provides the probability of blur detection for a single edge. But, since natural images consist of a large number of edges, it is important to devise a method for predicting how the information obtained from single edges can be pooled together to get a single quality score. In [2], pooling over edge pixels in individual blocks with significant edge content (referred to as edge blocks) and then over all the considered edge blocks, is accomplished using a Minkowski metric based on a probability summation model [5].

The metric of [2] is based on a model assuming that the blur impairment increases when P_{BLUR} increases. However, the metric of [2] ignores the fact that the blur is not likely to be perceived when it is below the JNB. This work presents an improved metric that exploits this fact by considering the value of P_{BLUR} relative to P_{JNB}. In the proposed metric, the pooling is based on the CPBD, which is obtained from the normalized histogram of the probability of blur detection of the processed edges in the entire image. The CPBD corresponds to the percentage of edges at which the probability of blur detection is below the just noticeable blur detection probability ($P_{JNB} = 63\%$) and, hence, to the percentage of edges at which blur is not likely to be detected.

A block diagram summarizing the computation of the proposed CPBD blur metric is shown in Fig. 1. Edge detection is first performed on the image. In the current implementation, only horizontal edges are detected as in [6]. The metric was also tested by including both horizontal and vertical edges. From the obtained results [7], it was established that including both vertical and horizontal edges in the calculations did not provide any significant improvement in the results for Gaussian-blurred and JPEG2000-compressed images. Hence, only horizontal edges are considered. The image is then divided into 64x64 blocks, approximating the size of foveal regions for a viewing distance of 60 cm and a display resolution of 31.5 pixels/cm, which correspond to a display visual resolution of 32.9 pixels per degree [8]. Depending on the edge information in each block, the blocks are then classified as edge blocks or nonedge blocks. The criterion to be classified as edge blocks is that the number of edges detected in the block should at least be 0.2% of the total number of pixels in the block [9]. The blocks characterized as nonedge blocks are not processed further. For each edge pixel e_i in an edge block, the corresponding edge width $w(e_i)$ is determined as in [6], and the JNB edge width $w_{\mathrm{JNB}}(e_i)$ is obtained depending on the local contrast C of the block using (2). The probability of detecting blur at the edge pixel e_i, $P_{\mathrm{BLUR}}(e_i) = P_{\mathrm{BLUR}}$, is then computed by using (1). A normalized histogram of blur detection probabilities (histogram of P_{BLUR}/Total number of processed edge pixels) is obtained, which gives the probability density function of P_{BLUR}. Finally, from the probability density function of P_{BLUR}, the cumulative probability of blur detection, which is the proposed metric, is calculated as

$$\begin{aligned}\mathrm{CPBD} &= P(P_{\mathrm{BLUR}} \leq P_{\mathrm{JNB}}) \\ &= \sum_{P_{\mathrm{BLUR}}=0}^{P_{\mathrm{BLUR}}=P_{\mathrm{JNB}}} P(P_{\mathrm{BLUR}})\end{aligned} \tag{3}$$

where $P(P_{\mathrm{BLUR}})$ denotes the value of the probability distribution function at a given P_{BLUR}.

The above metric is based on the fact that, at the JNB, $w(e_i) = w_{\mathrm{JNB}}(e_i)$, which corresponds to the probability of blur detection $P_{\mathrm{BLUR}} = P_{\mathrm{JNB}} = 63\%$. Thus, for a given edge e_i, when $P_{\mathrm{BLUR}} \leq P_{\mathrm{JNB}}$, the blur is considered to be not detected at the edge. As an image is increasingly blurred, the spread of the edges increases, which results in a higher value of $w(e_i)$ and, hence, in a higher probability of blur detection at the considered edge. As mentioned earlier, the proposed CPBD blur metric, given by (3), corresponds to the percentage of edges at which the probability of blur detection is below P_{JNB} and, hence, to the percentage of edges at which blur cannot be detected (in a probabilistic sense). Hence, a higher metric value represents a sharper image.

III. Performance Results

Here, performance results for the proposed CPBD blur metric are presented. First, the test sets taken from various publicly available databases, that are used for the metric evaluation, are described. Then, the performance of the proposed metric is tested using the described test sets which consist of various Gaussian blurred and

JPEG2000-compressed images. Results are presented to show how well the proposed metric correlates with the subjective scores as compared to the existing no-reference sharpness/blur metrics.[1]

A. Test Sets

To test the performance of the metric, Gaussian-blurred and JPEG2000-compressed images from the LIVE [10], TID2008 [11], IVC [12], and Toyama [13] databases are used.

The LIVE database [10] consists of 29 24-b/pixel RGB color images (typically 768×512). The images are distorted using different distortion types: JPEG2000, JPEG, Gaussian blur in RGB components, white noise in the RGB components, and bit errors in the JPEG2000 bitstream when transmitted over a simulated fast-fading Rayleigh channel. The subjective experiments for the LIVE database were conducted such that each distortion type was evaluated by different subjects in different experiments. These experiments used the same equipment and viewing conditions. Each image was rated by about 20–29 subjects. The subjects were asked to rate the images on a continuous linear scale which was divided into five different regions namely, "Bad," "Poor," "Fair," "Good," and "Excellent." The raw scores for each subject were converted to difference scores and then z-scores. The scores were then scaled and shifted to a range of 1 to 100. Then the difference mean opinion score (DMOS) for each image was calculated. All of the Gaussian-blurred images (174 images) and all of the JPEG2000-compressed images (227 images) from the LIVE database are used in our experiments.

The TID2008 database [11] consists of 25 reference images (typically 512×384) and 1700 distorted images. The images are distorted using 17 types of distortions, including but not limited to JPEG2000, Gaussian blur, JPEG, impulse noise and mean shift. A different methodology was used to conduct subjective tests for the TID 2008 database. A reference image at the bottom of the screen and a pair of distorted images at the top of the screen were simultaneously presented to the test subjects. Each subject was then asked to select a distorted image that differs less from the reference image. After the first selection, two different (new) distorted images appear in the upper part of screen. Each observer was asked to carry out the subjective test for only one reference image in one experiment. Each observer was given preliminary instructions and was trained on a set of distorted images before carrying out the actual experiments. The evaluation scale used was from 0 to 9 and the final MOS for each image was obtained by averaging all quality evaluations for a given image. The MOS was obtained from the results of 838 experiments carried out by a total of 838 observers from three different countries (251 experiments have been carried out in Finland, 150 in Italy, and 437 in Ukraine). Totally, the 838 observers have performed 256428 comparisons of visual quality of distorted images or 512 856 evaluations of relative visual quality in image pairs. All of the Gaussian-blurred natural images (96 images) and all of the JPEG2000-compressed natural images (96 images) from the TID2008 database are used in our experiments.

The IVC database [12] consists of ten reference images (typically 512×512) and 235 distorted images. The images are distorted using different distortion types: JPEG, JPEG2000, locally adaptive resolution (LAR) coding, and blurring. Fifteen subjects participated in the subjective tests. The viewing distance was set to be six times the picture's height. The tests were conducted using a double stimulus impairment scale method (DSIS), in which both original and distorted pictures were shown sequentially. The subjects were asked to rate the distortion they noticed in the distorted pictures with respect to the original on a five point scale having adjectives namely, 5—imperceptible, 4—perceptible but not annoying, 3—slightly annoying, 2—annoying and 1—very annoying. The mean opinion scores were then calculated after outlier removal. In our experiments, all of the Gaussian-blurred images (24 images) and all pf the JPEG2000-compressed images (60 images) from the IVC database are used.

The Toyama database [13] consists of 14 reference images (typically 768×512) and 168 distorted images. The images are distorted using JPEG and JPEG2000 compression. Sixteen subjects participated in the subjective tests. The viewing distance was set to be four times the picture's height. The tests were conducted using the single stimulus absolute category rating (SSACR) method. The subjects were asked to rate the images on a discrete five-point quality scale namely, bad (1), poor (2), fair (3), good (4), and excellent (5). The test presentation order was randomized. The scores were then converted into difference scores and, finally, the difference mean opinion scores were computed for each image after outlier rejection. In our experiments, all the JPEG2000-compressed images (98 images) from the Toyoma database are used.

B. Performance Results for the CPBD Metric

Here, results are presented to illustrate the performance of the proposed CPBD metric. Fig. 2 illustrates the behavior of the proposed CPBD blur metric for the 512×768 Sailing1 image which was obtained from the UT Austin LIVE database [10]. Fig. 2(a)–(c) show the blurred versions of the Sailing1 image using a circularly symmetric 2-D Gaussian kernel having a standard deviation of 0, 0.9, and 1.7, respectively. Fig. 2(d)–(f) show the probability distribution functions (PDFs) $P(P_{\mathrm{BLUR}})$ corresponding to Fig. 2(a)–(c), respectively. The corresponding cumulative distribution functions are shown in Fig. 2(g)–(i), respectively. From Fig. 2(g)–(i), it can be seen that, as the amount of blur increases, the proposed CPBD metric, which is equal to $P(P_{\mathrm{BLUR}} \leq P_{\mathrm{JNB}})$, decreases as expected. In the proposed implementation, the computed P_{BLUR} values are first quantized using a scalar quantizer with a step size of 0.01 and the quantized values are used for computing the PDFs $P(P_{\mathrm{BLUR}})$ and the proposed CPBD metric as in (3).

As the blurriness in the image increases, the proposed CPBD metric is expected to monotonically decrease. Fig. 3 shows the behavior of the proposed metric for blurred versions of the 512×768 Bikes image. These images were obtained from the LIVE database [10]. It can be noted that the metric behaves as expected.

To measure how well the proposed metric correlates with the provided subjective scores for the various databases, the authors followed the suggestions of the VQEG report [14] where several evaluation metrics are proposed. To account for the quality rating compression at the extremes of the test range, a four parameter logistic function as suggested in [14], is used.

The used logistic function is given by

$$\mathrm{MOS}_{p_i} = \frac{\beta_1 - \beta_2}{1 + e^{(M_i - \beta_3/|\beta_4|)}} + \beta_2 \tag{4}$$

where β_1, β_2, β_3, and β_4 are the model parameters, MOS_{p_i} is the predicted MOS, and M_i is the proposed metric for image i. The values of β_1, β_2, β_3, and β_4 are first obtained using a best fit to the corresponding subjective MOS scores, and are then used to find the predicted MOS (MOS_{p_i}) values using (4). The predicted MOS values are then used in calculating the performance measures including PCC (Pearson correlation coefficient, indicates the prediction accuracy), SROCC (Spearman rank-order correlation coefficient, indicates the prediction monotonicity), RMSE (root mean squared prediction error), MAE (mean absolute prediction error) and OR (outlier ratio, indicates consistency). Note that, for a good metric, the values of the Pearson

[1]The source code for the proposed CPBD metric is available at http://ivulab.asu.edu/Quality.

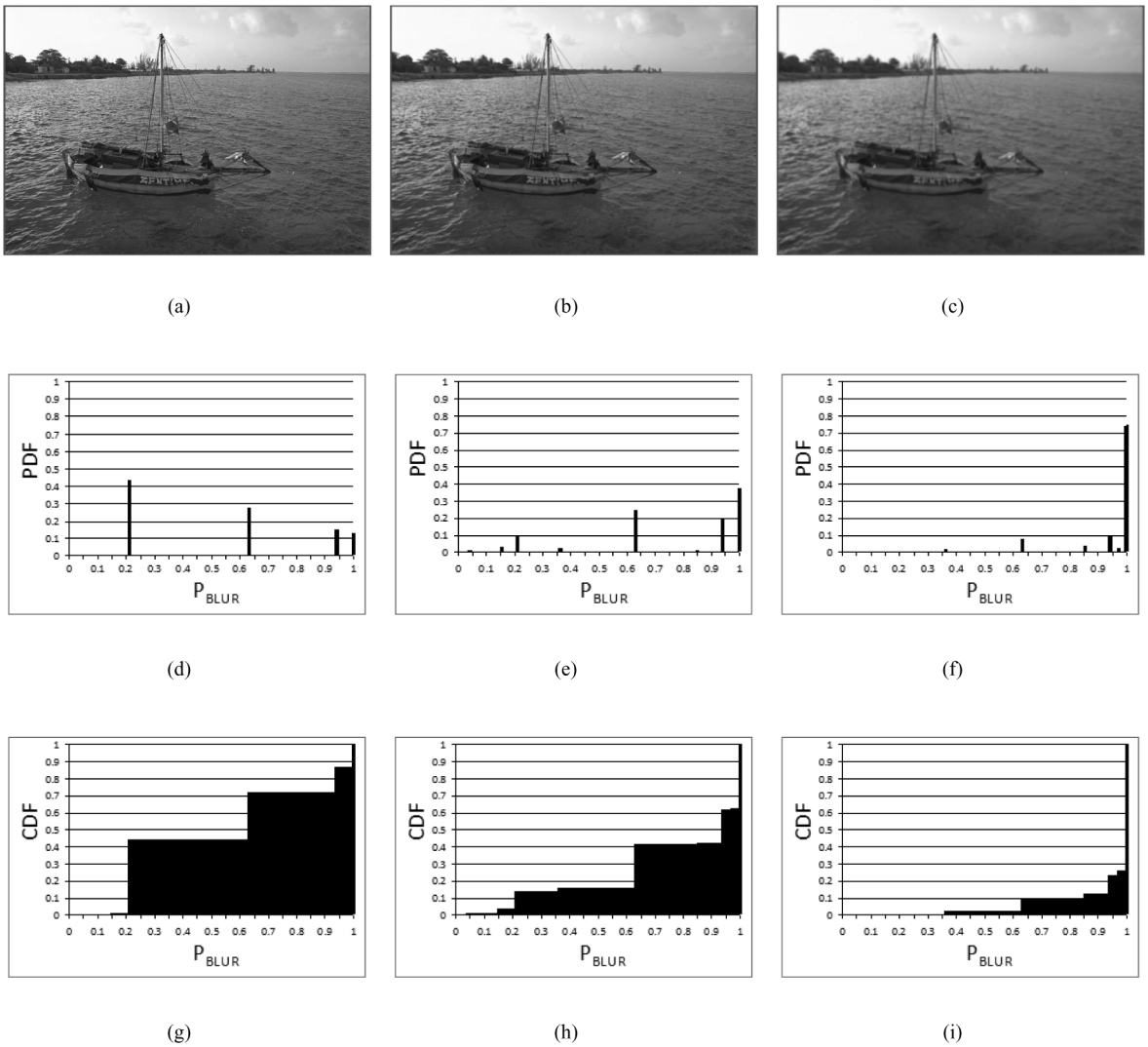

Fig. 2. Effect of blurring on the probability distribution function and the cumulative probability distribution function (CDF) of the localized blur detection probabilities P_{BLUR}. (a) Original 512 × 768 Sailing1 image. (b) Image blurred using a 2-D Gaussian kernel having $\sigma = 0.9$. (c) Image blurred using a 2-D Gaussian kernel having $\sigma = 1.7$. (d) $P(P_{\mathrm{BLUR}})$ of the image in (a). (e) $P(P_{\mathrm{BLUR}})$ of the image in (b). (f) $P(P_{\mathrm{BLUR}})$ of the image in (c). (g) CDF of P_{BLUR} for the image in (a). (h) CDF of P_{BLUR} for the image in (b). (i) CDF of P_{BLUR} for the image in (c).

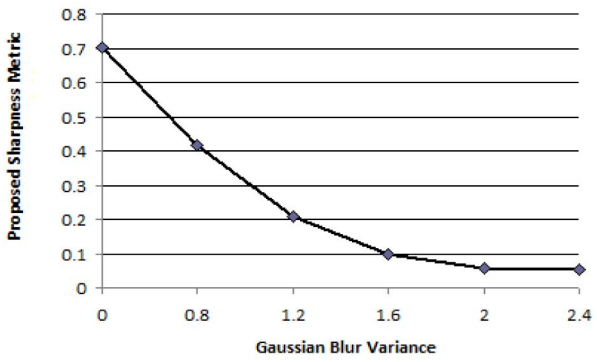

Fig. 3. Monotonic decreasing behavior of the proposed CPBD metric for increasing blur (512 × 768 Bikes image).

and Spearman correlation coefficients should be high and the values of RMSE, MAE, and OR should be low.

Tables I–IV summarize the results of the proposed metric, along with the metrics proposed in [2] and [6], and the more recent LPC-SI metric proposed in [4], for the Gaussian-blurred and JPEG-2000 compressed images obtained from the LIVE, TID2008, IVC, and Toyama databases, respectively. In addition to the recent LPC-SI metric of [4], the metrics in [2] and [6] were chosen for comparison because these blur metrics perform better as compared to other existing blur metrics as discussed in [2]. From Tables I–IV, it can be observed that, while the proposed CPBD metric has a performance that is close to the recent competitive LPC-SI metric of [4] for Gaussian blur, the proposed CPBD metric significantly outperforms the LPC-SI metric [4] for JPEG2000 blur for all databases. In fact, the obtained results show that the LPC-SI metric [4] exhibits a low correlation for JPEG-2000-compressed images for all the databases, and it even results in a negative correlation coefficient for the JPEG2000-compressed images of the Toyama database (Table IV). In addition, the obtained results indicate that the proposed CPBD metric outperforms the other existing state-of-the-art metrics for both Gaussian-blurred and JPEG2000-compressed images across all databases.

It is worth indicating that, as shown in Tables I–IV, the tested metrics exhibit differences in performance across different databases. In [15], the authors analyze the impact of different subjective databases,

TABLE I
EVALUATION OF THE PROPOSED CPBD METRIC PERFORMANCE W.R.T. DMOS SCORES FOR THE LIVE DATABASE

Distortion	Metrics	Pearson	Spearman	RMSE	MAE	OR
Gaussian Blur	CPBD metric	0.9107	0.9437	8.9857	6.8869	0.1609
	LPC-SI metric [4]	0.92	0.9497	8.525	6.9335	0.1724
	JNBM metric [2]	0.839	0.8368	11.8365	9.3485	0.2471
	Marziliano metric [6]	0.8597	0.8659	11.1106	8.2743	0.2184
JPEG-2000	CPBD metric	0.8835	0.8862	11.4288	9.0546	0.326
	LPC-SI metric [4]	0.4233	0.388	22.1024	18.4596	0.6167
	JNBM metric [2]	0.719	0.7255	16.9546	13.7478	0.5022
	Marziliano metric [6]	0.7815	0.7744	15.2196	11.7002	0.4670

TABLE II
EVALUATION OF THE PROPOSED CPBD METRIC PERFORMANCE FOR THE TID2008 DATABASE

Distortion	Metrics	Pearson	Spearman	RMSE	MAE
Gaussian Blur	CPBD metric	0.8316	0.8406	0.6438	0.5019
	LPC-SI metric [4]	0.8113	0.803	0.6778	0.5202
	JNBM metric [2]	0.7171	0.7045	0.8081	0.6254
	Marziliano metric [6]	0.709	0.7165	0.8176	0.6466
JPEG-2000	CPBD metric	0.9223	0.925	0.7406	0.5831
	LPC-SI metric [4]	0.7952	0.7295	1.1621	0.9356
	JNBM metric [2]	0.8798	0.8789	0.9111	0.7213
	Marziliano metric [6]	0.8667	0.8694	0.9561	0.7127

TABLE III
EVALUATION OF THE PROPOSED CPBD METRIC PERFORMANCE FOR THE IVC DATABASE

Distortion	Metrics	Pearson	Spearman	RMSE	MAE
Gaussian Blur	CPBD metric	0.8865	0.8404	0.6629	0.5126
	LPC-SI metric [4]	0.9718	0.9022	0.3378	0.2647
	JNBM metric [2]	0.7992	0.7722	0.8609	0.6374
	Marziliano metric [6]	0.822	0.7896	0.8157	0.6185
JPEG-2000	CPBD metric	0.7784	0.7882	0.8633	0.6196
	LPC-SI metric [4]	0.3974	0.3293	1.2621	1.0749
	JNBM metric [2]	0.6154	0.6206	1.0841	0.8678
	Marziliano metric [6]	0.6451	0.6402	1.0509	0.835

TABLE IV
EVALUATION OF THE PROPOSED CPBD METRIC PERFORMANCE FOR THE TOYAMA DATABASE

Distortion	Metrics	Pearson	Spearman	RMSE	MAE	OR
JPEG-2000	CPBD metric	0.7829	0.7876	0.8202	0.6347	0.1429
	LPC-SI metric [4]	-0.0234	-0.0077	1.3184	1.2059	0.4388
	JNBM metric [2]	0.6095	0.5248	1.0452	0.8374	0.2449
	Marziliano metric [6]	0.6428	0.6254	1.01	0.8064	0.2245

including the LIVE, IVC, and Toyoma databases, on the performance of image quality metrics. From the experiments, it is noted in [15] that the performance of the image quality metrics is quite different across databases, and that this difference in performance can be attributed to the differences in contents, distortions, and quality range across databases. For example, the low-quality images in the LIVE database are significantly more degraded as compared to the low-quality images in the other databases under consideration. Despite this, the results obtained herein (Tables I–IV) show that the proposed metric achieves consistently a good performance across blur types (Gaussian blur and JPEG2000 blur) and across databases as compared to existing sharpness/blur metrics.

IV. CONCLUSION

In this work, a blur metric is proposed based on the CPBD. It involves edge detection followed by estimating the probability of detecting blur at the detected edges. Then a probability density function for the obtained probabilities is calculated from which the final cumulative probability of blur detection is obtained. It is shown that the proposed metric exhibits consistently a good performance across blur types (Gaussian blur and JPEG2000 blur) and across databases as compared with existing sharpness/blur metrics.

Possible directions of research include extending the metric for assessing blur in videos and 3-D visual content by considering additional factors such as temporal and depth effects on blur perception.

REFERENCES

[1] L. J. Karam, T. Ebrahimi, S. S. Hemami, T. N. Pappas, R. J. Safranek, Z. Wang, and A. B. Watson, "Introduction to the issue on visual media quality assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 189–192, Apr. 2009.

[2] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.

[3] N. G. Sadaka, L. J. Karam, R. Ferzli, and G. P. Abousleman, "A noreference perceptual image sharpness metric based on saliency-weighted foveal pooling," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 369–372.

[4] R. Hassen, Z. Wang, and M. Salama, "No-reference image sharpness assessment based on local phase coherence measurement," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 2434–2437.

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

[6] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Applications to JPEG2000," *Signal Process.: Image Commun.*, vol. 19, no. 2, pp. 163–172, Feb. 2004.

[7] N. D. Narvekar, "Objective no-reference visual blur assessment," M.S. thesis, Dept. Electr. Eng., Arizona State Univ., Tempe, 2009.

[8] Z. Liu, L. J. Karam, and A. B. Watson, "JPEG2000 encoding with perceptual distortion control," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1763–1778, Jul. 2006.

[9] D. Hood and M. Finkelstein, *Handbook of Perception and Human Performance*. New York: Wiley, 1986.

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

[11] N. Ponomarenko, M. Carli, V. Lukin, K. Egiazarian, J. Astola, and F. Battisti, "Color image database for evaluation of image quality metrics," in *Proc. Int. Workshop Multimedia Signal Process.*, Oct. 2008, pp. 403–408.

[12] P. Le Callet and F. Autrusseau, "Subjective quality assessment IRCCyN/IVC database," 2005 [Online]. Available: http://www.irccyn.ec-nantes.fr/ivcdb/

[13] Z. M. P. Sazzad, Y. Kawayoke, and Y. Horita, "Image quality evaluation database," [Online]. Available: http://mict.eng.u-toyama.ac.jp/database_toyama/

[14] "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment," VQEG, 2000.

[15] S. Tourancheau, F. Autrusseau, Z. M. Parvez, and Y. Horita, "Impact of subjective datasets on the performance of image quality metrics," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 365–368.

Nonlocal Means With Dimensionality Reduction and SURE-Based Parameter Selection

Dimitri Van De Ville, *Member, IEEE*, and Michel Kocher

***Abstract*—Nonlocal means (NLM) is an effective denoising method that applies adaptive averaging based on similarity between neighborhoods in the image. An attractive way to both improve and speed-up NLM is by first performing a linear projection of the neighborhood. One particular example is to use principal components analysis (PCA) to perform dimensionality reduction. Here, we derive Stein's unbiased risk estimate (SURE) for NLM with linear projection of the neighborhoods. The SURE can then be used to optimize the parameters by a search algorithm or we can consider a linear expansion of multiple NLMs, each with a fixed parameter set, for which the optimal weights can be found by solving a linear system of equations. The experimental results demonstrate the accuracy of the SURE and its successful application to tune the parameters for NLM.**

***Index Terms*—Linear transforms, nonlocal means (NLM), principal component analysis (PCA), Stein's unbiased risk estimate.**

I. INTRODUCTION

Learning from neighborhoods has become an important and powerful data-driven approach for various applications in image processing. Most notably, the nonlocal means (NLM) [1] algorithm applies adaptive averaging based on similar neighborhoods in a search region. Various methods have been proposed to accelerate the initial approach using preselection of the contributing neighborhoods based on average value and gradient [2], average and variance [3] or higher-order statistical moments [4], cluster tree arrangement [5], and [6], [7]. The computation of the distance measure between different neighborhoods itself can be optimized using the fast Fourier transform [8], a moving average filter [9], [10], early termination of the search [11], or by reducing redundant comparisons [12].

Variations of the NLM algorithm have also been proposed to improve the denoising performance; e.g., adaptive neighborhoods [13], iterative application [5], combination with kernel regression [14] and spectral analysis [15], and other similarity measures based on principal component analysis (PCA) [6], [16] or rotation invariance [17]. The smoothing parameter that determines the contributions of the patches has been locally optimized using Mallow's C_p statistic [18]. The most evolved version of the nonlocal principle is probably BM3D [19], which further processes the selected neighborhoods and gives high quality results.

The combination of NLM with dimensionality reduction methods such as PCA [6], [16] and SVD [7] has gained increased interest since the advantages are twofold. First, the computational complexity is highly reduced. Second, measuring the distance between neighborhoods in a lower-dimensional subspace improves robustness to noise;

Manuscript received August 11, 2010; revised December 06, 2010, February 10, 2011; accepted February 21, 2011. Date of publication March 07, 2011; date of current version August 19, 2011. This work has been funded in part by the Swiss National Science Foundation (PP00P2-123438, D. Van De Ville) and in part by the Centre for Biomedical Imaging (CIBM). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Rafael Molina.

D. Van De Ville is with the Department of Radiology and Medical Informatics, University of Geneva, 1211 Geneva 14, Switzerland, and also with the Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015, Switzerland.

M. Kocher is with the Biomedical Imaging Group, EPFL, Lausanne 1015, Switzerland.

Digital Object Identifier 10.1109/TIP.2011.2121083