Objective no-reference image blur metric based on local phase coherence

A. Ciancio, A.L.N.T. da Costa, E.A.B. da Silva, A. Said, R. Samadani and P. Obrador

A novel algorithm for no-reference blur assessment in digital pictures is proposed. This metric is based on the evaluation of local phase coherence across the scales of an overcomplete wavelet transform. We compare the performance of the proposed method with four state-of-the art metrics using a large database containing both simulated and real blur.

Introduction: Popularisation of digital cameras has created a demand for objective algorithms for image quality assessment, as a means to select images for a final application (e.g. printing), automatically discarding the ones that do not meet a pre-established quality criterion. In this scenario, blur can be considered as one of the most common causes for quality degradation in digital pictures. While blur can be generated by a number of different optical and motion phenomena, it can be generally modelled by a smoothing of the high frequencies present in the image.

Quality assessment techniques can be divided into three main categories: full-reference, reduced-reference and no-reference measurements [1]. In full-reference techniques, both the original undegraded image and the distorted images are known; in reduced-reference methods, only a few features of the original image are known; and in no-reference techniques, neither the reference image nor any of its features is known.

In [2], the authors proposed a new model in which the blur process causes a disruption of phase coherence among scales in the wavelet transform domain. In this Letter, based on this model and inspired by the work in [3], we propose a novel no-reference objective metric for image blur assessment and evaluate its performance against four well-known quality evaluation metrics in different blurring scenarios.

Overview of typical no-reference quality metrics: In [4], the authors propose a quality evaluation algorithm that exploits the distribution of null discrete cosine transform (DCT) coefficients instead of the values themselves (blurred images tend to have a large number of their high-frequency coefficients set to zero). The quality measure is obtained by using a weighting grid that gives more importance to the coefficients on the central diagonal of the DCT coefficient matrix, since they better characterise global (circular, non-directional) blur. We refer to this method as the frequency domain metric.

In [5], the authors propose a method that takes advantage of the ability of the Haar wavelet transform (HWT) in discriminating edge types. After performing an HWT with three levels of decomposition, an edge map is constructed in each scale. After that, this edge map is partitioned, and local maxima in each window are found. The quality grade is estimated based on the relative occurrence of different types of edges. Since it operates on the edges, we refer to this method as the spatial domain metric.

In [6], a perceptual blur, no-reference metric, based on edge length is proposed. First, a Sobel operator is used to detect edge locations on the luminance component of the image. Then, the edge lengths corresponding to the distance between the start and end positions of the edge are computed. The global blur measure is obtained by averaging the lengths over all edges found.

In [7], the authors propose an algorithm that uses human visual system (HVS) features to improve the performance of an objective metric. In this method, the image is divided into blocks of 8×8 pixels. Blocks are marked for processing based on their edge count. Then, the average edge length for each block is computed based on the method proposed in [6], and a perceptual weight based on the contrast of the block is multiplied by each block's average length. The final blur measure is the weighted average edge length. This metric is referred to as the HVS based metric.

Proposed metric: In [2] the authors show that precisely localised features such as steep edges result in strong local coherent phase structures across scale and space in the complex wavelet transform domain, and that blurring causes loss of such phase coherence. They demonstrate that the only signals that have a continuous spectrum and have such a scale invariance are the step function and its derivatives, such as the delta function. They also develop a measure of phase coherence based on coarse-to-fine phase prediction and show that this measure can serve as an indication of blur in natural images.

In [8], a similar idea is used to develop a wavelet-based multiscale products thresholding scheme for noise suppression of magnetic resonance images. The authors exploit the wavelet interscale dependencies by multiplying the adjacent subbands of the dyadic wavelet transform of the image to enhance edge structures while weakening noise.

In the proposed algorithm, an overcomplete wavelet transform [3] of an input image is computed. Because of the wavelet properties, it is expected that coefficients of subbands with the same orientation are located in similar positions. As shown in [2], the presence of blur will introduce phase incoherence, causing these positions to slightly change from subband to subband. Inspired by the work in [8] we created an algorithm that separates the bands into coherent wavelet coefficients and incoherent coefficients.

Coefficients are classified as coherent or incoherent based on an adaptive threshold. In what follows, we consider $W^{\mathrm{D}s}$, the Sth scale of the wavelet decomposition of the image in the direction D (where D can be H or V, corresponding to horizontal or vertical direction, respectively), with $S = 1$ being the highest frequency. Each scale of the wavelet can therefore be spatially partitioned as $W^{\mathrm{D}s} = C^{\mathrm{D}s} + I^{\mathrm{D}s}$, where $C^{\mathrm{D}s}$ contains the coherent coefficients (with the incoherent equal to zero) and $I^{\mathrm{D}s}$ contains the incoherent coefficients. To find the appropriate threshold value for the considered image, we follow a similar approach to the one found in [8]. We start by multiplying adjacent wavelet subbands, which enhances coherent coefficients and weakens incoherent ones (due to matching). The adaptive threshold value is found by the following algorithm:

For D = H and V, do:
1) Set the threshold $thr = 0$;
2) For all pixels (i, j), if $[W^{\mathrm{D}1}(i,j)W^{\mathrm{D}2}(i,j)] > thr$, then $C^{\mathrm{D}1}(i,j) = W^{\mathrm{D}1}(i,j)$ and $I^{\mathrm{D}1}(i,j) = 0$; else, $C^{\mathrm{D}1}(i,j) = 0$ and $I^{\mathrm{D}1}(i,j) = W^{\mathrm{D}1}(i,j)$;
3) Compute σ^2, the variance of $I^{\mathrm{D}1}(i,j)$, and set $thr_{\mathrm{new}} = \sigma 2$;
4) Compute new $W^{\mathrm{D}1} = C^{\mathrm{D}1} + I^{\mathrm{D}1}$;
5) If $|thr - thr_{\mathrm{new}}| > 1/255^2$, then set $thr = thr_{\mathrm{new}}$ and go to step 2; else, end algorithm.

It typically takes around three iterations for the algorithm to converge. Once it has converged, the blur estimation is calculated as the mean value of standard deviations of $I^{\mathrm{H}1}$ and $I^{\mathrm{V}1}$.

We have also employed a mapping of the metric based on a logistic function in order to enhance correlation with subjective grades. The mapping has been computed based on a training set, by dividing the set into three and applying a k-fold strategy [9].

Simulations and results: To assess the proposed metric's performance, 120 high-resolution (1280×960 to 2272×1704 pixels), high-quality images were distorted with two types of simulated blur: global isotropic blur and linear motion blur. The global blur case corresponds, in general, to an incorrect optical arrangement of the camera, resulting in an out-of-focus image, and can be efficiently modelled by the convolution of the original image and a Gaussian kernel, where the amount of blur is proportional to the standard deviation of the Gaussian function. The linear motion blur models the movement of the camera during the exposure time when the picture is taken. For relatively small time intervals, this movement can be approximated as linear, and the resulting distortion can be simulated by filtering the original images with kernels containing a single line of a certain length.

To generate the simulated blur database we convolved high-quality images with Gaussian filters of different standard deviations (to simulate out-of-focus blur) and with linear kernels (simulating linear motion blur), resulting in a set of 6000 degraded images (3000 corrupted with Gaussian blur and 3000 with linear motion blur). Also, a set of 580 high-resolution pictures taken by human users was used to assess the metrics' performances in a realistic blurring scenario. For these pictures, subjective tests inspired on the ITU-R 500 Recommendation [10] were conducted in order to obtain the ground truth for each image. In the tests, at least ten subjects graded each image based on their perceived quality. The final grade for each image consisted of the average grade given by all subjects (the images, as well as the subjective grades can be found at http://www.lps.ufrj.br/profs/eduardo/ImageDatabase.htm). We used, as a measure of effectiveness, the Pearson correlation between reference grades (for artificial blur, we used the variance in the case of the Gaussian filter and the kernel length in the linear blur case) and the metric's assessments.

The correlation values between the metrics' assessments and the ground truths for the proposed metric and the four metrics discussed previously can be seen in Table 1 (the values shown are the averages and standard deviations for the three folds [9]).

Table 1: Correlation values between metrics' assessments and ground truths

Metric	Gaussian blur	Linear blur	Gaussian linear	Real
Frequency domain [2]	0.631 ± 0.005	0.46 ± 0.02	0.55 ± 0.01	0.46 ± 0.07
Spatial domain [5]	0.72 ± 0.05	0.600 ± 0.003	0.64 ± 0.02	0.25 ± 0.03
Perceptual blur [6]	0.82 ± 0.03	0.29 ± 0.02	0.56 ± 0.03	0.32 ± 0.02
HVS based [7]	0.80 ± 0.04	0.41 ± 0.03	0.61 ± 0.03	0.43 ± 0.01
Proposed metric	0.75 ± 0.04	0.62 ± 0.01	0.67 ± 0.03	0.50 ± 0.05

It can be seen from the obtained results that the proposed metric significantly outperforms these metrics in the linear motion blur case, and provides the best overall results for simulated blur when both types of distortion are taken into account in the simulation. Linear motion blur can be frequently found in real scenarios owing to slow shutter speed during exposure and camera shake. For the specific case of Gaussian blur, the proposed metric shows an inferior performance to both perceptual blur and HVS based metrics.

For the real blurred images, although not as good as the simulation results, the proposed method still shows the best results among the tested metrics. In this case, pictures may contain hard-to-model blurring conditions, such as localised blur (only part of the picture is out of focus) or complex motion blur, affecting the performance of quality metrics.

The best overall performance for the proposed metric may be attributed to the use of blur modelling in terms of phase coherence. Typical methods usually assume a particular model for the blurring process (e.g. Gaussian or linear model), which makes their performance drop when that model does not correctly reflect the nature of the blur. In our case, the local phase coherence approach can be considered a more general strategy, which better models blur for generic scenarios.

Conclusions: In this Letter, we propose a novel no-reference quality assessment metric for digital pictures. This metric is based on the evaluation of local phase coherence across the scales of an overcomplete wavelet transform. We have compared the performance of the proposed metric against four well-referenced metrics from the literature by using a large, high-resolution database containing 6000 images corrupted with simulated blur and 580 images taken by humans, containing a variety of real blurring scenarios. Results have shown that the proposed metric offers the best overall performance for the simulated blur case. Although this is also true for the real blur case, absolute results can still be improved. We believe that a better study of the metrics' deficiencies or a deeper understanding of real blur may help to develop modifications to reflect even better the human quality assessment process.

30 June 2009
doi: 10.1049/el.2009.1800

A. Ciancio, A.L.N.T. da Costa and E.A.B. da Silva (*PEE/COPPE and DEL/Poli, Universidade Federal do Rio de Janeiro, C. P. 68504 – Rio de Janeiro – RJ, 21941-972, Brazil*)

E-mail: ciancio@lps.ufrj.br

A. Said and R. Samadani (*Multimedia Communications and Networking Lab, HP Labs, 1501 Page Mill Road, MS 1811, Palo Alto, CA 94306, USA*)

P. Obrador (*Telefonica Research, Via Augusta 177, Barcelona 08021, Spain*)

E-mail: pere@tid.es

References

1 Wang, Z., Sheikh, H.R., and Bovik, A.C.: 'Objective video quality assessment', *in* Furht, B. and Marqure, O. (Eds.): 'The handbook of video databases: design and applications' (CRC Press, 2003, pp. 1041–1078)

2 Wang, Z., and Simonceli, E.P.: 'Local phase coherence and the perception of blur' *in* 'Adv. neural information processing systems' (MIT Press, 2004, pp. 786— 792)

3 Mallat, S., and Zhong, S.: 'Characterization of signals from multiscale edges', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1992, **14**, (7), pp. 710–732

4 Marichal, X., Ma, W.-Y., and Zhang, H.: 'Blur determination in the compressed domain using DCT information'. Proc. Int. Conf. on Image Processing (ICIP'99), 1999, Vol. 2, pp. 286–390

5 Tong, H., Li, M., Zhang, H., and Zhang, C.: 'Blur detection for digital images using wavelet transform'. IEEE Int. Conf. on Multimedia and EXPO (ICME'04), June 2004, Vol. 1, pp. 17–20

6 Marziliano, P., Dufaux, F., Winkler, S., and Ebrahimi, T.: 'Perceptual blur and ringing metrics: application to JPEG2000', *Signal Process., Image Commun.*, 2004, **19**, (2), pp. 163–172

7 Ferzli, R., and Karam, L.J.: 'A human visual system based no-reference objective image sharpness metric'. Int. Conf. on Image Processing (ICIP), 2006

8 Bao, P., and Zhang, L.: 'Noise reduction for magnetic resonance images via adaptive multiscale products thresholding', *IEEE Trans. Med. Imaging*, 2003, **22**, (9), pp. 1089–1099

9 Kohavi, R.: 'A study of cross-validation and bootstrap for accuracy estimation and model selection'. Proc. Int. Joint Conf. on Artificial Intelligence, Montreal, Canada, August 1996

10 ITU-R 500 Recommendation: 'Methodology for the subjective assessment of quality of television pictures'