# Local Blur Estimation Based on Toggle Mapping

Théodore Chabardès and Beatriz Marcotegui[✉]

CMM - Centre for Mathematical Morphology,
MINES ParisTech, PSL Research University,
35 rue Saint Honoré - Fontainebleau, Paris, France
{theodore.chabardes,beatriz.marcotegui}@mines-paristech.fr
http://cmm.mines-paristech.fr

**Abstract.** A local blur estimation method is proposed, based on the difference between the gradient and the residue of the toggle mapping. This method is able to compare the quality of images with different content and does not require a contour detection step. Qualitative results are shown in the context of the LINX project. Then, quantitative results are given on DIQA database, outperforming the combination of classical blur detection methods reported in the literature.

**Keywords:** No-reference blur estimation · Local blur estimation · Toggle mapping · DIQA database

## 1  Introduction

With the proliferation of handheld devices equipped with high resolution cameras and increasing computational power, many mobile applications become possible nowadays. Text is present everywhere in our everyday's life. Visually impaired people have no access to it. In the framework of LINX project we aim at developing a smartphone application making access to textual information possible within everyone's reach, including visually impaired people. This project is funded by the French Interministerial funds (FUI) for competitive clusters.

LINX conditions of use are not under control. The user being visually impaired cannot check the acquired image quality: bad lighting conditions, blur or noise can degrade the acquired image. We focus here on blur, which is a common problem linked to handheld devices and low cost objectives of smartphone cameras.

Blur estimation is an extensively studied topic in the literature. Image quality assessment techniques can be classified in 3 groups:

- Full reference techniques, which compare the image to be evaluated with the undegraded version of it. This comparison can be based on the mean square error, the PSNR, the cross-correlation between both images, the structural similarity index [25] or any other fidelity measure.
- Reduced reference: the comparison is based on some features and not on the whole image.

– No-reference: the image quality assessment is based only on the degraded
image itself. This is the most difficult case, and corresponds to the LINX
context, where only the acquired image is available.

No-reference sharpness metrics have been defined in the spatial and in the
frequency domain. Some spatial metrics are based on variance computation [9],
total variation, gradient [8], Brenner's gradient [3], multiscale gradient [5], lapla-
cian [16]. Others are based on edge width estimation [15,4] or on histogram
measures such as histogram entropy [11,21,17].

Examples of spectral focus measures are the relative high frequency power
compared to the low frequency one, the sum of frequency components above a
certain threshold [8] or the correlation of wavelet coefficients over scales [6,24].
Other works rely on the phase coherence  [10,14,1].

Most of these methods have been designed in an autofocus context. They
can rank the quality of an image from an image set as long as their content is
similar. Therefore, they are not appropriate for our application. Some attempts
have been made in order to drop this restriction. Local approaches have been
proposed, computing the metric in a given neighborhood [19], that has to be
defined, or after a contour detection step [4,18].

Several papers address the specific problem of document image quality and
try to correlate the quality measure with an OCR accuracy [2,18,26]. An inter-
esting database, DIQA, has been introduced in this context [13]. 25 documents
have been acquired with a mobile-phone camera. 6 to 8 different images of each
document were taken. The camera was focused at varying focal lenghts to gen-
erate a series of images with focal blur. The quality of each image is estimated
as the OCR accuracy on it.

In this paper we introduce a local blur estimator that does not require a
prior contour detection, nor a neighborhood definition. The rest of the paper is
organised as follows: section 2 presents the toggle-mapping based blur estimator,
section 3 evaluates its performance on DIQA dataset [13]. Finally section 5
concludes the paper and discusses some perspectives of this work.

## 2   Toggle Mapping Based Blur Estimation

Toggle mapping (TM) operator was introduced in [12]. From an image $f$, two
transformations are computed: the dilation, $\delta_B(f)$ and the erosion $\epsilon_B(f)$. Equa-
tion 1 describes the toggle mapping operator:

$$TM_B(f) = \begin{cases} \delta_B(f) & if \ \delta_B(f) - f \ < \ f - \epsilon_B(f) \\ \epsilon_B(f) & otherwise \end{cases} \tag{1}$$

where $\delta_B(f)$ and $\epsilon_B(f)$ correspond respectively to the dilation and the erosion
with structuring element $B$.

Thus, each pixel is replaced by one of these two transformations, selecting
among them the one that is closest to the original pixel value. This process can
be iterated until convergence is reached. This principle was generalized in [22],
with the use of other transformations involved in the toggle process.

TM was designed as a contrast enhancement operator. It was also used as scene text segmentation tool in [7]: pixels replaced by the dilation are set to 1 and those replaced by the erosion are set to 0. In this paper a local blur estimation is introduced based on TM.

We define the residual image as the absolute difference between the original image and the TM:

$$TM\_Residue_B(f) = \begin{cases} \delta_B(f) - f & if \ \delta_B(f) - f \ < \ f - \epsilon_B(f) \\ f - \epsilon_B(f) & otherwise \end{cases} \tag{2}$$

A blur boundary leads to a high residue, and this residue increases with the size of the structuring element involved in the TM. This is true up to a TM size equivalent to the blur edge width. A blur estimation could then be built on the evolution of the TM residues, with a series of structuring elements of increasing size. However a sharp boundary and a homogeneous region would have the same series of low value TM residues. In order to distinguish between these two situations we compare these residues with the gradient value. The morphological gradient is defined as the difference between the maximum and the minimum in the neighborhood of a pixel ($\delta_B(f) - \epsilon_B(f)$), while the TM replaces each pixel with the closest extremum in the neighborhood, that is either the maximum or the minimum. Thus, by definition, the residue of the TM is lower than the gradient. However if the TM uses a larger structuring element ($B_M$ instead of $B_N$, with $M > N$) its residue can reach the gradient value. In which case, the pixel is classified as blur. Therefore we define the Q image as:

$$Q_{B_N, B_M}(f) = max \left(0, gradient_{B_N}(f) - TM\_Residue_{B_M}(f)\right) \tag{3}$$

Q image is filtered: first it is thresholded (threshold = 3 for all our experiments), then small regions (less than 5 pixels) are removed and finally Q is averaged in each connected component. The result is $Q_{filter}$ image.

Figure 1 shows an example with intermediary images. Figure 1(a) is a crop of an image acquired by a LINX end user with a mobile phone. The image is sharp but we can observe some noise that can be reinforced by the toggle mapping operator. A bilateral filter [23] is applied to get rid of this noise, see figure 1(b). Figure 1(c) shows the toggle mapping of figure 1(b) and figure 1(d) its corresponding residue. This residue is subtracted from the gradient image shown in figure 1(e), leading to pixel-wise quality estimation in figure 1(f). Figure 1(g) illustrates the $Q_{filter}$ image. Figure 2 shows the flow chart of the algorithm. The letters of the diagram (from (a) to (g)) correspond to intermediary images of figure 1. The average of selected pixels for this image is equal to 34, and it can be considered as the quality score $q_{score}$ of the crop, assuming it is homogeneous.

## 3   Results

In this section we will show qualitative results from LINX examples and quantitative results when applying our algorithm to DIQA database.
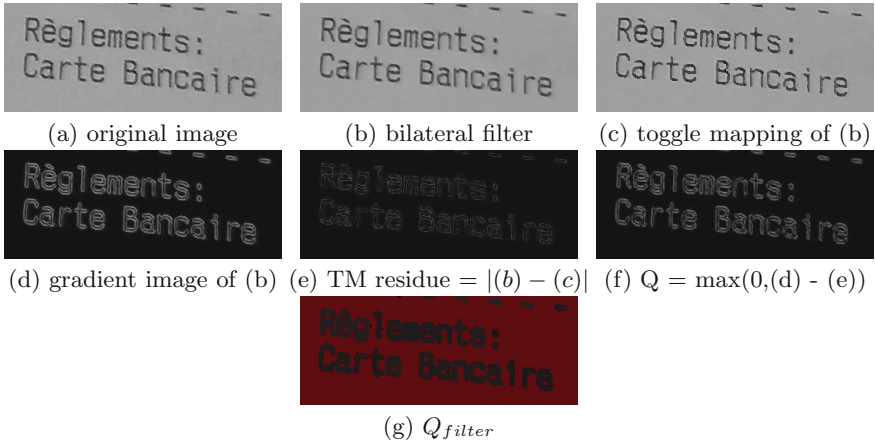
(a) original image      (b) bilateral filter      (c) toggle mapping of (b)

(d) gradient image of (b)    (e) TM residue $= |(b) - (c)|$    (f) Q = max(0,(d) - (e))

(g) $Q_{filter}$

**Fig. 1.** Intermediary steps of local quality estimation: (b) Bilateral filter $\sigma_{spatial} = 2, \sigma_{gray} = 20$; (c) Toggle mapping of size 2; (d) Gradient of size 1; (e) $|(b) - (c)|$; (f) Q = max(0,(d)-(e)); (g) $Q_{filter}$ with threshold equal to 3, regions larger than 5 pixels and average $Q$ on each selected connected component
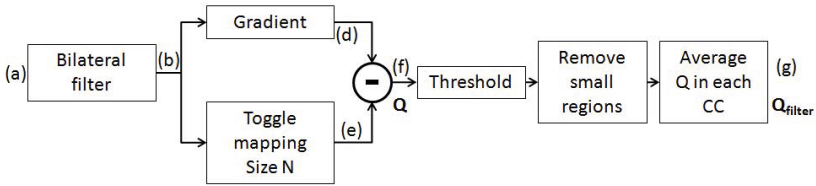


**Fig. 2.** Flow chart of the algorithm

The parameters used in all our experiments (if not specified otherwise) are:

– Bilateral filter $\sigma_{spatial} = 2, \sigma_{gray} = 20$
– Structuring element: cross (4 neighbors)
– Gradient of size 1 (3x3 pixels)
– Toggle mapping residue of size 2 (5x5 pixels)
– Threshold: 3
– Small regions: 5 pixels

Figure 3 shows some examples with different levels of blur from LINX database. In the first row we can see a good quality image example, with a quality score of 30. In the second row a good quality image with a score of 20 and in the third row a blurred image with a score of 4. We can observe a good correspondence of our score and the observed quality.
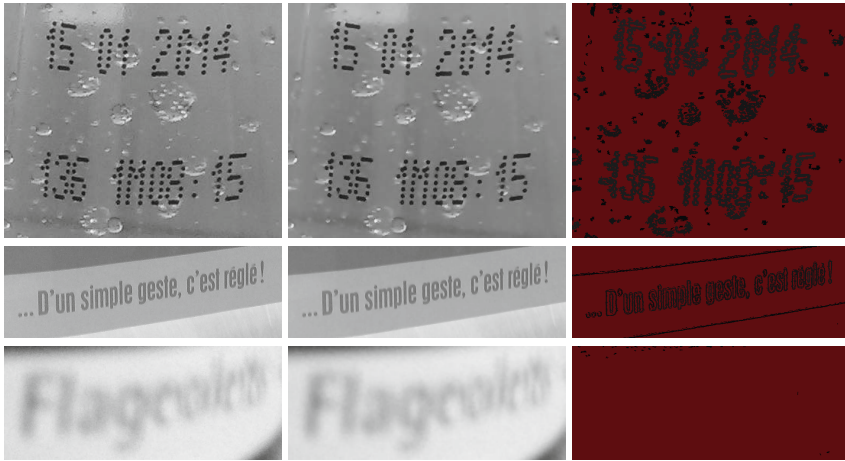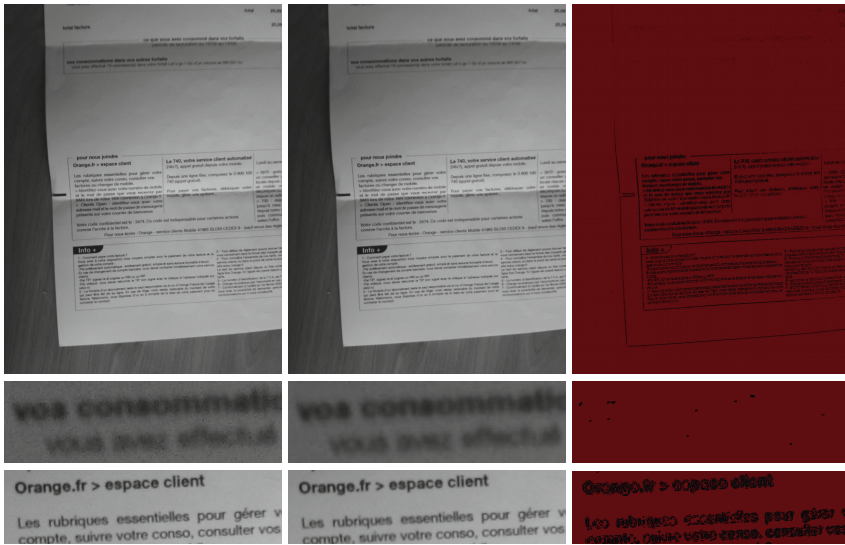
**Fig. 3.** Example of LINX images with different levels of blur. First row: sharp image ($q_{score}$=30). Second row: Good quality image ($q_{score}$=20). Third row: Blurred image ($q_{score}$=4). Left column: original image, Middle column: bilateral filter, Right column: $Q_{filter}$.



(a) original image          (b) bilateral filter          (c) $Q_{filter}$

**Fig. 4.** Example with heterogenous blur. First row, the whole image; second row a crop from the top of the image, in a blurred area ($q_{score}$=5); third row a crop from the middle of the image, in a good quality area ($q_{score}$=10).

One of the advantages of our method is that the measure is local. Figure 4 shows an example with heterogeneous blur. The upper part of the image is more out-of-focus than the lower part. We can observe a higher result in the lower part. Two crops from different areas have been extracted, and we confirm a good correlation of our measure with the image quality: the crop from the upper part has a quality score equal to 5 while the crop from the lower part has a quality score equal to 10. Note that the crop has been extracted for visualization purposes. The measure is local and each region has its own score.

In order to have quantitative results we apply our algorithm to the recently published DIQA image database [13]. It is an interesting dataset, composed of 175 document images acquired with a mobile phone. It is noteworthy that images are directly acquired by a camera, so the blur is not simulated. 25 documents have been acquired, 6 to 8 times each, in different conditions: from perfect focus to completely out-of-focus images. The quality of each image is given by the OCR accuracy on the processed image. Three different OCRs have been used: Tesseract, FineReader and Omni. The overall accuracy of these OCRs on the whole database are:

|                  | Tesseract | FineReader | Omni |
|------------------|-----------|------------|------|
| OCR accuracies:  | 0.53      | 0.76       | 0.72 |

FineReader leads to the best overall accuracy (0.76) and will be our reference in the rest of the paper, unless stated otherwise. Figure 5 shows 3 crops from this dataset, with an OCR accuracy ranging from 0.97 to 0.24. Our quality score for these images ranges from 5 to 55. Spearman correlation is commonly used to assess the coherence of two variables without taking into account their precise values. It compares their ranking indexes instead of the variables themselves. In our example, the ranking based on the OCR accuracy or on our quality score is the same ( 1.- (a); 2.-(b); 3.-(c)). Thus, for this simple example with only three images, our quality score is perfectly correlated with the OCR accuracy: the Spearman correlation between them is then equal to 1.

Some works [13,20] have tried to estimate the OCR accuracy based on some features extracted from the image. The score used for comparing different methods is the median of the Spearman correlation applied to each document set. The scores reported are over 0.9. We think that the proposed score overestimates the quality of the OCR accuracy prediction. Indeed, if the Spearman correlation is computed for each document set, the content of images to be compared is similar. The context is then close to an autofocus situation which is much easier than comparing the quality of images with different content. This problem is pointed out in [20]: a failure is reported for an image with a low focus measure whereas the OCR accuracy is over 0.9, provoked by the large white space in the page. Another reported issue concerns an image with a high focus measure with a low OCR accuracy due to a text out-of-focus but a huge headline leading to a focus measure optimistically high. Those are exactly the situations that we address in this paper. Moreover, as stated in [20], by reporting the median value, outlier classes in which the methods might not perform well are disregarded. They propose to compute the Spearman correlation directly for the 175 images
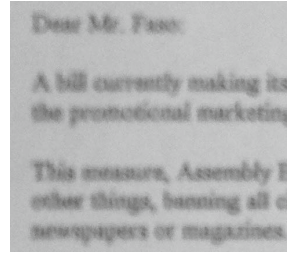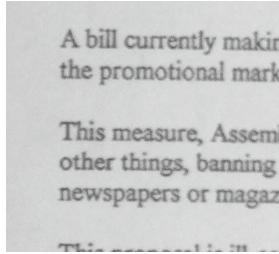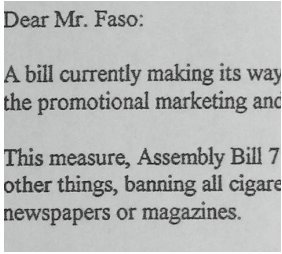
(a) 2012-04-16_17-34-12_68    (b) 2012-04-16_17-32-27_295    (c) 2012-04-16_17-33-23_805
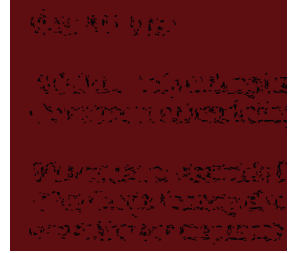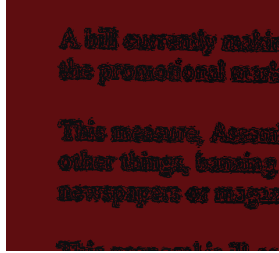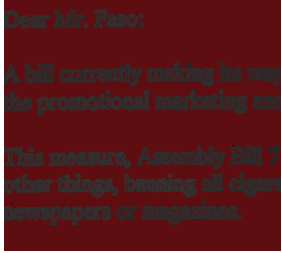
(d)    OCR    accuracy= 0.9738; $q_{score}$=55

(e)    OCR    accuracy = 0.8861; $q_{score}$=14

(f)    OCR    accuracy = 0.2412; $q_{score}$=5

**Fig. 5.** DIQA examples. First row: original images. Second row: $Q_{filter}$ Left (FineReader accuracy = 0.9738; $q_{score}$=55), Middle (FineReader accuracy =0.8861; $q_{score}$ = 14). Right (FineReader accuracy = 0.2412; $q_{score}$ = 5).

**Table 1.** Spearman correlation of our quality score against OCR accuracy, with different pre-processing filter

| Filter | FineReader | Tesseract | Omni |
|---|---|---|---|
| original | 0.745 | 0.711 | 0.622 |
| AF 1 | 0.784 | 0.748 | 0.645 |
| bilateral 1 | 0.816 | 0.813 | 0.692 |
| bilateral 2 | 0.849 | 0.892 | 0.725 |
| bilateral 3 | 0.843 | 0.902 | 0.727 |
| median 1 | 0.787 | 0.756 | 0.659 |
| median 2 | 0.822 | 0.808 | 0.690 |

of the database. We adopt this proposal that we find more reliable. They report in [20] a Spearman correlation of 0.6467 for the whole database as a single set for FineReader OCR.

Table 1 summarizes the results of Spearman correlation between our quality score against the OCR accuracy for 175 images in DIQA dataset. A higher Spearman correlation value indicates the method's ability to rank images according to the expected OCR accuracy. Our score is 0.745 compared to 0.6467 reported in the [20] combining several classical descriptors from [19]. As we observed

**Table 2.** Spearman correlation on document 1 from DIQA dataset. Spearman correlation(our_score,tesseract) = 0.53. Spearman correlation(our_score,FineReader)=0.89.

|  | imA | imB | imC | imD | imE | imF | imG |
|---|---|---|---|---|---|---|---|
| Our score: | 16 | 17 | 35 | 38 | 19 | 6 | 37 |
| Our rank: | 6 | 5 | 3 | 1 | 4 | 7 | 2 |
| Tesseract accuracy: | 0.77 | 0.78 | 0.87 | 0.86 | 0.70 | 0.12 | 0.71 |
| Tesseract rank: | 4 | 3 | 1 | 2 | 6 | 7 | 5 |
| Finereader: | 0.95 | 0.97 | 0.99 | 0.98 | 0.97 | 0.13 | 0.98 |
| Finereader rank: | 6 | 5 | 1 | 2 | 4 | 7 | 3 |



**Fig. 6.** Example with large characters in a blurred context, but still readable

a qualitative improvement in LINX database using a prefiltering step, we tried different filters (alternate filter AF, bilateral or median filter) on DIQA database and the best result was 0.849, obtained by a bilateral filter of size 2 ($\sigma_{gray}$ was fixed to 20, and we have not verified yet the sensitivity of this parameter in our quality measure).

## 4   Discussion

Spearman correlation estimates the correlation between two ranked variables. For example, if we take the 7 images of the first document of DIQA dataset we get a quality score ranging from 6 to 38. The Spearman correlation between our score and Tesseract accuracy is 0.53. It can be considered as a relative low correlation value. However analyzing the result closer we can observe that imF is classified as the worst image in both cases. imC and imD have quite similar values (0.87 and 0.86 for Tesseract accuracy and 35 and 38 for our score respectively). The ranking is different but the difference in quality between the two images is small. Both measures are able to estimate the different groups of image quality but Spearman correlation does not catch these similarities between images. Somehow it would

be interesting to quantize the scores and do not consider an error for ranking differences in the same quantile.

Another issue is related to the scale. Images of figure 6 contain large blurred characters. As they are blurred, no good quality pixel is selected and the image is then rejected. However despite the blur the characters are still readable.

## 5     Conclusions and Perspectives

In this paper we introduce a local quality estimation approach. It is used as a pre-processing step in a mobile phone application, aiming at avoiding further processing of bad quality images. The method compares the residue of a toggle mapping of size $N$ with the gradient of size $M$, with $M < N$.

The interest of the approach is shown first through several LINX images. The quality scores obtained are correlated with the perceived quality of the images. Then, in order to evaluate quantitatively the efficiency of our method we rank the quality of the DIQA images. A Spearman correlation of 0.745 is reached, between our score and FineReader accuracy. This performance is to be compared to 0.6467 reported in the literature, combining several classical blur descriptors. Our performance raises to 0.849 if a bilateral filter is applied before estimating the quality. The OCR accuracy is given for the whole image. It would be interesting to verify if the errors appear when the quality is locally lower, for some images with heterogeneous focus.

In the future we will set the quality threshold values for accepting or rejecting a region in the context of the LINX project. Intermediate quality values could be considered: contrast enhancement would be applied in those cases, before further processing. We will also address the problem of large blurred characters, in a multi-scale approach.

## References

1. Blanchet, G., Moisan, L.: An explicit sharpness index related to global phase coherence. In: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1065–1068. IEEE (2012)
2. Blando, L.R., Kanai, J., Nartker, T.A.: Prediction of OCR accuracy using simple image features. In: Proceedings of the Third International Conference on Document Analysis and Recognition, vol. 1, pp. 319–322. IEEE (1995)
3. Brenner, J.F., Dew, B.S., Horton, B., King, T., Neurath, P.W., Selles, W.D.: An automated microscope for cytologic research a preliminary evaluation. Journal of Histochemistry & Cytochemistry 24(1), 100–111 (1976)
4. Cao, G., Zhao, Y., Ni, R.: Edge-based blur metric for tamper detection. Journal of Information Hiding and Multimedia Signal Processing 1(1), 20–27 (2010)

5. Chen, M.-J., Bovik, A.C.: No-reference image blur assessment using multiscale gradient. EURASIP Journal on Image and Video Processing 2011(1), 1–11 (2011)

6. Ciancio, A., da Costa, A.L.N.T., da Silva, E.A.B., Said, A., Samadani, R., Obrador, P.: No-reference blur assessment of digital pictures based on multifeature classifiers. IEEE Transactions on Image Processing 20(1), 64–75 (2011)

7. Fabrizio, J., Marcotegui, B., Cord, M.: Text segmentation in natural scenes using toggle-mapping. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 2373–2376. IEEE (2009)

8. Firestone, L., Cook, K., Culp, K., Talsania, N., Preston, K.: Comparison of autofocus methods for automated microscopy. Cytometry 12(3), 195–206 (1991)

9. Groen, F.C.A., Young, I.T., Ligthart, G.: A comparison of different focus functions for use in autofocus algorithms. Cytometry 6(2), 81–91 (1985)

10. Hassen, R., Wang, Z., Salama, M.: No-reference image sharpness assessment based on local phase coherence measurement. In: 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 2434–2437. IEEE (2010)

11. Jarvis, R.A.: Focus optimization criteria for computer image-processing. Microscope 24(2), 163–180 (1976)

12. Kramer, H.P., Bruckner, J.B.: Iterations of a non-linear transformation for enhancement of digital images. Pattern Recognition 7(1), 53–58 (1975)

13. Kumar, J., Ye, P., Doermann, D.: A dataset for quality assessment of camera captured document images. In: Iwamura, M., Shafait, F. (eds.) CBDAR 2013. LNCS, vol. 3857, pp. 113–125. Springer, Heidelberg (2014)

14. Leclaire, A., Moisan, L., et al.: No-reference image quality assessment and blind deblurring with sharpness metrics exploiting fourier phase information (2014)

15. Marziliano, P., Dufaux, F., Winkler, S., Ebrahimi, T.: Perceptual blur and ringing metrics: application to jpeg2000. Signal Processing: Image Communication 19(2), 163–172 (2004)

16. Nayar, S.K., Nakagawa, Y.: Shape from focus. IEEE Transactions on Pattern analysis and machine intelligence 16(8), 824–831 (1994)

17. Chern, N.N.K., Neow, P.A., Ang, V.M.H.: Practical issues in pixel-based autofocusing for machine vision. In: Proceedings of 2001 IEEE International Conference on Robotics and Automation, ICRA, vol. 3, pp. 2791–2796. IEEE (2001)

18. Peng, X., Cao, H., Subramanian, K., Prasad, R., Natarajan, P.: Automated image quality assessment for camera-captured OCR. In: 2011 18th IEEE International Conference on Image Processing (ICIP), pp. 2621–2624. IEEE (2011)

19. Pertuz, S., Puig, D., Garcia, M.A.: Analysis of focus measure operators for shape-from-focus. Pattern Recognition 46(5), 1415–1432 (2013)

20. Rusiñol, M., Chazalon, J., Ogier, J.-M.: Combining focus measure operators to predict OCR accuracy in mobile-captured document images. In: 2014 11th IAPR International Workshop on Document Analysis Systems (DAS), pp. 181–185. IEEE (2014)

21. Schlag, J.F., Sanderson, A.C., Neuman, C.P., Wimberly, F.C.: Implementation of automatic focusing algorithms for a computer vision system with camera control. Carnegie-Mellon Univ., the Robotics Inst. (1983)

22. Serra, J.: Toggle mappings. From pixels to features,61–72 (1988)

23. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Sixth International Conference on Computer Vision, pp. 839–846. IEEE (1998)

24. Wang, Z., Bovik, A.C.: Reduced-and no-reference image quality assessment. IEEE Signal Processing Magazine 28(6), 29–40 (2011)

25. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004)
26. Ye, P., Doermann, D.: Learning features for predicting OCR accuracy. In: 2012 21st International Conference on Pattern Recognition (ICPR), pp. 3204–3207. IEEE (2012)