https://doi.org/10.2352/ISSN.2470-1173.2020.9.IQSP-322

Relation Between Image Quality and Scan Resolution: Part I*

Zhenhua Hu[1], Litao Hu[1], Peter Bauer[2], Todd Harris[2], Jan Allebach[1]
[1]School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, 47907
[2]HP Inc., Boise, Idaho, 83714

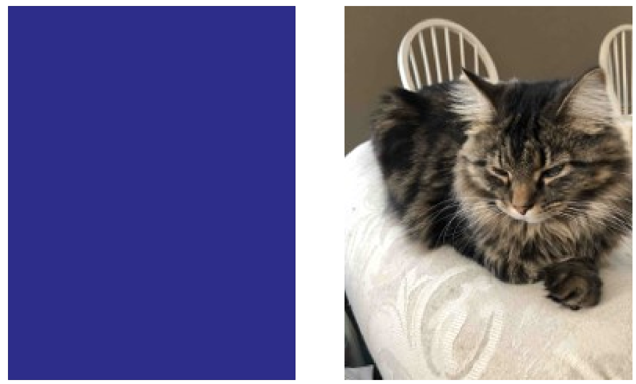
Abstract

Currently, a particular scan resolution has to be defined before a scanner starts working. Two problems arise from this process. Firstly, no matter how different two pages contents are, they will be scanned into the same resolution. For example, after scanning, a blank page and a fine-detailed drawing will have the same resolution. Secondly, for one scanned page, every part of its output would have the same resolution, whatever their contents are. These problems will cause unnecessary waste of memory used to store scanned images. So a method to decide the minimum acceptable scan resolution is needed. But current image quality estimators are not suitable for estimating image quality at different resolutions. This paper proposes four features to assess image qualities at different resolutions, namely 75, 100, 150, 200 and 300 dpi. The features are tile-SSIM mean, tile-SSIM standard deviation, horizontal transition density, and vertical transition density. Tests on images containing different contents show that these features are promising in evaluate image qualities across different scan resolutions.

Introduction

Modern scan routines require a predefined scan resolution, whether it is customer-selected or a default in the scanner's settings. When the scanning process begins, the resolution cannot be changed. This results in all scanned pages, no matter how much their contents may vary, having output images of the same size. For example, if one page has an image area of plain color and another image has many details, as shown in Fig. 1, the two images will have the same size after the scanning process. However, we can clearly see in Fig. 1 that picture (b) needs a larger scan resolution than picture (a). If we scan both pages with a low resolution like 75 dpi, picture (a) will have all the information but picture (b) will lose some details. However, if we scan them both at a high resolution like 300 dpi, picture (b) will have all the details, but we also need more space to store picture (a). As a result, the best strategy for us would be to scan images with different resolutions according to their contents, so that we can still maintain good image quality while lowering the total storage used.

In order to tackle this problem, different scan resolutions for different contents must be chosen. Undoubtedly, higher scan resolutions would keep more details and yield higher-quality images. The quality of scanned images degrades as the scan resolution decreases, which may cause blurriness. And this is quite obvious in scanned text documents, since blurriness can cause a document to be unreadable.



(a) flat image (b) cat image

Figure 1: Example of two pages with different contents. We can see that picture (a) contains a plain blue area, while picture (b) has more detail. So we can see that they need different scan resolutions

Image quality assessment (IQA) is a widely-researched area topic [1] [2]. Typically, there are three categories of quality estimators (QEs): full reference (FR) QEs, reduced reference (RR) QEs, and no reference (NR) QEs. FR QEs exploit the ”perfect” reference image and examine the truthfulness that a new image has with the reference image. The simplest measures are the mean squared error (MSE) or peak signal-to-noise ratio (PSNR). But its nature of heavily relying on pixel correspondences does not reflect human's subject view. Other QEs, like structural similarity index (SSIM) [3] [4], multi-scale structural similarity [5], visual information fidelity (VIF) [6], most apparent distortion (MAD) [7], can all reflect human's perceptual assessment. But they all require a reference image of the same resolution and size. As a result, we cannot use these FR QEs directly.

Reduced reference (RR) approaches are used if the reference image is available, but is too costly to get. Most RR QEs are based on natural scene statistics (NSS) models [8]. For example, the ”divisive normalization” QE [9] uses the Gaussian scale mixture (GSM) model on the wavelet coefficients of the reference image and the image to be tested. Also, the ”quality-aware image” QE [10] exploits information on a general Gaussian distribution (GGD) model. But since we actually have the reference image, and some of the images under study may not be natural scenes, RR QEs are not suitable for our study.

When there is not any information from a reference image, IQA must be processed on the test image only. This is referred to

*Research supported by HP Inc., Boise ID 83714

as the no reference (NR) IQA. NR QEs exploit properties of test images directly, like the blocking from JPEGs [11] [12], ringing from JPEG 2000 [13] and blurring [14]. Some QEs exploit the NSS properties like BRISQUE [15]. There are also some training based NR IQAs that learn the human judgements on a large database of human opinion scores [6] Although our problem has reference images, they cannot be used directly like FR QEs. We could use some NR QEs to help.

In this paper, we propose some QEs that could help perform IQA of images at different resolutions. Although the reference images cannot be used directly as in FR IQA, our QEs can use their information to help assess qualities changes with different resolutions. We also explore NR QEs since they can reflect image qualities without reference images. Further, the image resolutions in consideration are 300 dpi, 200 dpi, 150 dpi, 100 dpi, and 75 dpi. We chose these resolutions because after some research and investigation, we found that 300 dpi can preserve image qualities that are good enough for most scanned materials. Besides, most scanners from the high-end to low-end all support these resolutions. Tests on various kinds of images show that these QEs are promising in predicting image qualities at different resolutions.

Methodology

Tile-SSIM

SSIM (structural similarity index) [3] is the most successful FR QE and is widely used in estimating quality of images. It is composed of three elements over small patches of two images x, y: a luminance similarity term $l(x, y)$, a contrast similarity term $c(x, y)$, and a structure similarity term $s(x, y)$. Their computations are [3]:

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \tag{1}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \tag{2}$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \tag{3}$$

Here μ_x, μ_y are averages of x and y, respectively. σ_x, σ_y are standard deviations of x and y, respectively. σ_{xy} is the covariance of x and y. And c_1, c_2 and c_3 are variables to stabilize the division with a weak denominator. The formula to calculate the SSIM value is

$$SSIM(x,y) = l(x,y) \cdot c(x,y) \cdot s(x,y) \tag{4}$$

SSIM is very powerful in predicting the structural similarity of a test image to a reference image. After some research and investigation, we find that the majority of images have good perceptual quality at 300 dpi. So we can use 300 dpi images as the reference to assess qualities of lower resolution (200 - 75 dpi) images. But images of different resolutions have different sizes, which cannot be assessed using SSIM directly. As a result, we need to do some pre-processing steps before SSIM can be applied.

We noticed that for images of different resolutions, the ratio of their sizes are the same as the ratio of their resolutions. As a result, we can generate images of the same size by utilizing these ratios. Then we can apply the SSIM to assess quality changes across resolutions.

We propose a method called tile-SSIM. According to the size ratios, we separate different resolution images into the same number of non-overlapping tiles. In each tile, we calculate the standard deviation of it. And then we combine all the standard deviations to form a standard deviation map, as is shown in Figure 2. We choose standard deviation to represent each tile because low resolution images (200 - 75 dpi) can be approximated by area-based downsampling from high resolution (300 dpi) images. The number of grayscale levels decreases in this process, which could result in a decrease of the activity in the corresponding tiles. At the end of the preprocessing, since all images, whatever their resolutions are, have the same number of tiles, their standard deviation maps must be the same size.

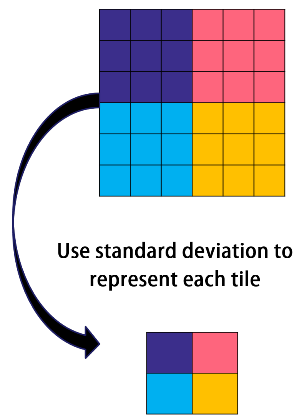


Figure 2: Steps to generate standard deviation map. The above block shows the original image and the tile size of 3×3. After calculating the standard deviation of each tile, we form a new standard deviation map, which has the same size across all resolutions.

The image resolutions and their corresponding tile sizes can be found in Table 1. The tile widths and heights are proportional to their resolutions so that the images have the same number of tiles regardless of their resolutions.

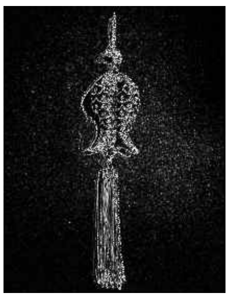
Table 1: Scan resolutions and their corresponding tile size.

Resolution (dpi)	Tile width/height (pixel number)
300	12
200	8
150	6
100	4
75	3

Figure 3 shows the original 300 dpi and 200 dpi images, and their standard deviation maps, respectively. We can see that although the original images are of different sizes, their standard

(a) Original 300 dpi image

(b) 200 dpi image



(c) 300 dpi standard deviation map

(d) 200 dpi standard deviation map

Figure 3: Original 300 dpi and 200 dpi images and their standard deviation maps.

deviation maps have the same size, and keep most of the structural information.

Figure 4 shows the workflow of tile-SSIM. First, color images of all resolutions are turned into grayscale images before being passed into a Gaussian filter to eliminate random noise. Then, these images are separated into non-overlapping tiles according to their resolutions (Table 1), so that the number of tiles will be the same across all resolutions. In each tile, we calculate its standard deviation value. In this way, we have standard deviation maps of the same size. Then, we can calculate the SSIM values between between low-resolution standard deviation map and the high resolution standard deviation map.

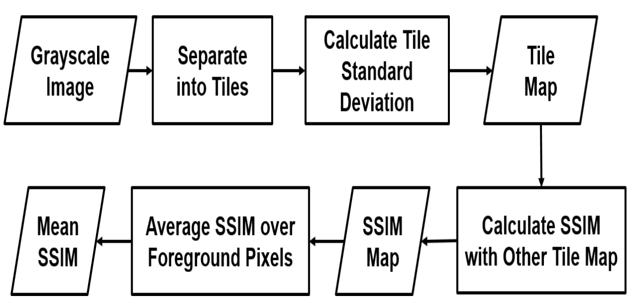
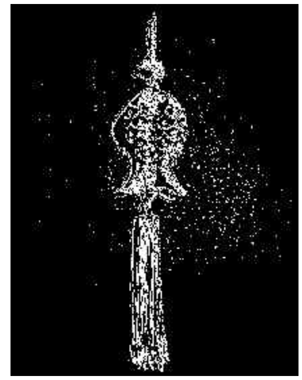


Figure 4: Tile-SSIM workflow.

(a) Full SSIM map of 300 and 200 dpi images



(b) Binary image of standard deviation map

Figure 5: Full SSIM map and the binary image of the standard deviation map. From (a) and (b) we can tell that the binary map provides a good mask to locate foreground pixels, so that background pixels will not affect the quality assessment.

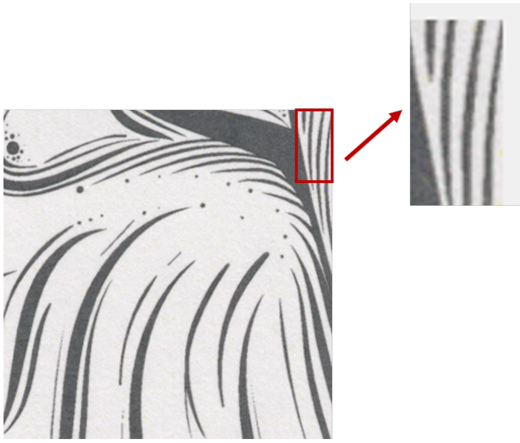
On the standard deviation maps, the SSIM values are calculated on 7×7 image patches with stride equal to *1* pixel. We then could use their average to assess the structural similarities of low resolution images with their references. But when people are looking at an image, we can see that they are mainly focused on the foreground areas. The background areas, like the large white area in a document image, or plain flat areas in a natural scene image, usually attract much less attention. As a result, it would not be reasonable to average over all these background areas.

In order to tackle this problem, we can locate the foreground pixels by binarizing the 300 dpi standard deviation map through some threshold algorithm like Otsu's method [16]. By averaging over foreground pixels only, we can eliminate the effect of background areas and better reflect the viewer's perception of image quality changes at different resolutions, as is shown in Figure 5. With the decreasing scan resolution, we should expect a continuous drop in SSIM mean values.

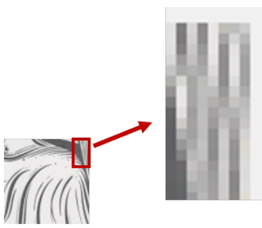
In addition to SSIM average, we can also use its standard deviation as a feature. As the scan resolution goes down, image patches with high variations would see a lot of decrease in variation values, thus resulting in a decrease in SSIM values. In the mean time, the smooth areas would not change much. They can maintain relatively high values across all resolutions. As a result, as the scan resolution goes down, the SSIM standard deviation of image foreground patches would tend to increase. This could be used as an indicator of the relation between scan resolution and image quality.

Transition Density

While tile-SSIM mean and standard deviation may be good QEs to assess the similarity between 300 dpi images and their low resolution counterparts, these QEs cannot assess the quality of the reference 300 dpi images. If the 300 dpi images are of relatively low quality, no matter how good their SSIM value is relative to the low resolution images, we cannot scale them down. So we also need some NR QEs to independently justify the quality of images at each resolution.



(a) Original 300 dpi image and part of its enlarged area



(b) 75 dpi image with the same area enlarged

Figure 6: Change of line drawing with decreasing resolution. The 300 dpi image in (a) has very good details with distinct lines, while the 75 dpi image of (b) has very blurry edges. From the enlarged area, we can even barely see any edges in the 75 dpi image.

As the scan resolution decreases, we can clearly see that the images become more and more blurred. Some sharp edges may even disappear. This is much more obvious with document images or line drawings, as their distinct edges maybe very close to each other. As the resolution decreases, two edges are likely to merge into one edge or become visibly indistinguishable. All of this would cause some of the characters to become unidentifiable, and drawings to become blurry, as shown in Figure 6. With this motivation in mind, we define new features to reflect quality changes at different scan resolutions.

Two features, horizontal and vertical transition densities are defined based on the ideas mentioned above. These transition densities can be used to characterize connected components properties [17]. Based on the discussion above, they may also can be used to test image quality changes at different resolutions. Figure 7 shows the workflow for computing the horizontal and vertical transitions.

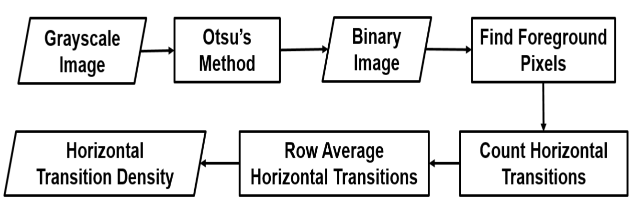


Figure 7: Workflow for calculating horizontal transitions. Calculation of vertical transitions proceeds in a similar manner.

We start with color images which are first converted into grayscale images. Then binary images are obtained through Otsu's binarization and foreground pixels are identified. To count horizontal transitions we draw horizontal lines along with each foreground pixels, as shown in Figure 8. The horizontal transition density is defined as the ratio of total number of horizontal transitions over the number of horizontal lines containing foreground pixels. Similarly, as shown in Figure 9, we draw vertical lines along with each foreground pixels to obtain vertical transitions. And vertical transition density is the ratio of vertical transitions over vertical lines.

Figure 8: Horizontal transition.

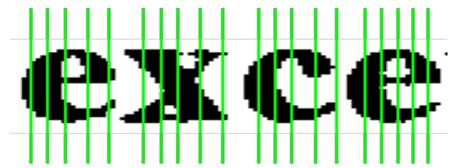


Figure 9: Vertical transition.

Experiment Results

We selected 6 images to test our QEs. The images ”cricket” and ”elephant” are natural images of a cricket and an elephant, respectively. The image ”form” is the picture of a form, which contains only text and lines. The images ”map_purdue” and ”map_nanjing” are two map images. And the image ”mix” is one with both text and pictures.

Tile-SSIM

There are 2 features in tile-SSIM: tile-SSIM mean values and tile-SSIM standard deviations. The tile-SSIM mean values contain *4* elements: between 300 dpi and 200 dpi s_200; between 300 dpi and 150 dpi s_150; between 300 dpi and 100 dpi s_100; and between 300 dpi and 75 dpi s_75. So the feature vector of tile-SSIM mean values s_value is:

$$s\_value = (s\_200, s\_150, s\_100, s\_75) \tag{5}$$

Similarly, the tile-SSIM std values contain *4* elements: between 300 dpi and 200 dpi std_200; between 300 dpi and 150 dpi std_150; between 300 dpi and 100 dpi std_100; and between 300 dpi and 75 dpi std_75. So the feature vector of tile-SSIM std values s_std is:

$$s\_std = (std\_200, std\_150, std\_100, std\_75) \tag{6}$$

Figure 10 shows the result of applying the tile-SSIM algorithm to the 6 different images. In order to provide a better comparison, we also put the tile-SSIM value between 300 dpi and itself in the figure, and that value is always *1*. From Figure 10 we can see that although these images differ a lot in content, their SSIM mean values behave the same way. They slightly decrease at 200 and 150 dpi, then drop quickly at 100 and 75 dpi. These plots reflect the quality of these images as their resolutions decrease. So the tile-SSIM mean can be used as a feature to estimate image quality.

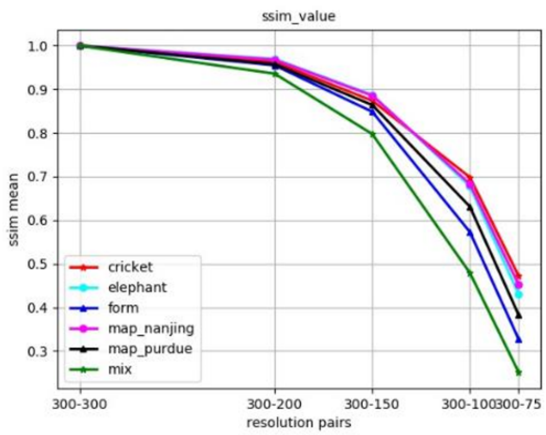


Figure 10: Tile-SSIM mean value.

Figure 11 shows how the standard deviation changes with different resolution pairs. Note that the standard deviations between the 300 dpi images and themselves are always *0*. Except for some minimal decreases (form and mix decrease a little bit from the 300-100 dpi pair to the 300-75 dpi pair), the standard deviations for all images keeps increasing with increasing difference between the pair of resolutions. We could also use this feature to predict image quality.

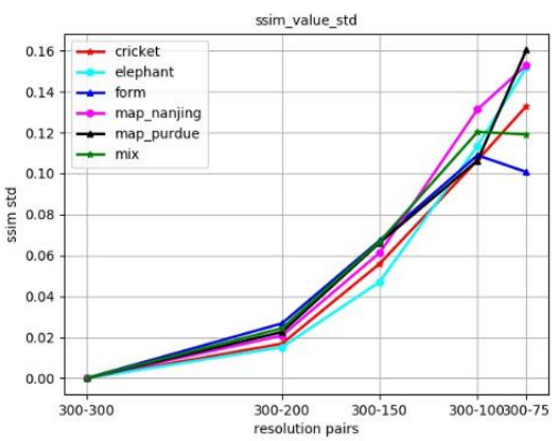


Figure 11: Tile-SSIM standard deviation.

Transition Density

Figure 12 and Figure 13 show the results for horizontal and vertical transition densities, respectively. Note that since these are NR QEs, and each image has 5 resolutions, each feature has 5 elements. The horizontal density *hd* can be denoted as

$$hd = (hd\_300, hd\_200, hd\_150, hd\_100, hd\_75) \tag{7}$$

where *hd_300*, *hd_200*, *hd_150*, *hd_100*, and *hd_75* denote the horizontal density of the 300 dpi, 200 dpi, 150 dpi, 100 dpi, and 75 dpi images, respectively.

Similarly, the vertical density *vd* can be denoted as

$$vd = (vd\_300, vd\_200, vd\_150, vd\_100, vd\_75) \tag{8}$$

where *vd_300*, *vd_200*, *vd_150*, *vd_100*, and *vd_75* denote the vertical density of the 300 dpi, 200 dpi, 150 dpi, 100 dpi, and 75 dpi images, respectively.

We can see in these two figures that the numbers all decrease with decreasing resolution. As expected, the text document (form) has the largest horizontal transition density. The sudden decrease at 100 dpi indicates a significant drop in image quality. The cricket picture has the least intensity variation among all images. Thus, it has the least number of transitions. The big drop from 300 dpi to 200 dpi might indicate an image quality drop.

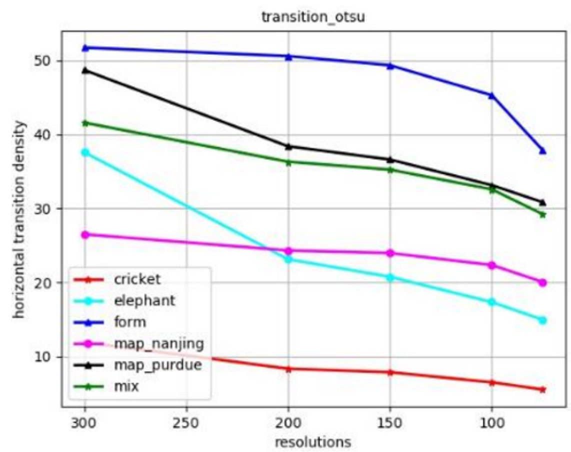


Figure 12: Horizontal transition result.

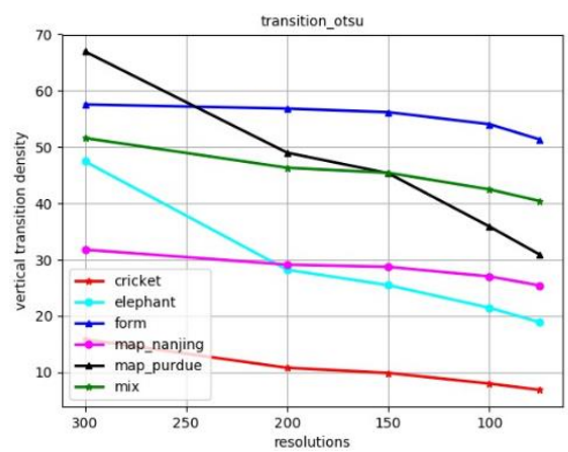


Figure 13: Vertical transition result.

Conclusion

To assess image quality at different scan resolutions, this paper proposes four QEs: tile-SSIM mean, tile-SSIM standard deviation, horizontal transition density, and vertical transition density. Our tests on different content images show that they change consistently with decreasing resolution. With labeled ground truth data, and a machine learning method like support vector machine, we expect to be able to predict the image quality at different resolutions and thus decide the optimal resolution for different scan contents.

References

[1] Alan C. Bovik, Automatic prediction of perceptual image and video quality, in Proc. IEEE, vol 101, pp: 2008-2024, 2013.

[2] Kalpana Seshadrinathan, and Alan C. Bovik, Automatic prediction of perceptual quality of multimedia signalsa survey, Int. J. Multimedia Tools Appl., vol 51, pp. 163-186, 2011.

[3] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process., vol 13, pp. 600-612, 2004.

[4] Richard Szeliski, Computer vision: algorithms and applications, Springer Science & Business Media, 2010.

[5] Demin Wang, Filippo Speranza, Andre Vincent, Taali Martin, and Phil Blanchfield, Toward optimal rate control: a study of the impact of spatial resolution, frame rate, and quantization on subjective video quality and bit rate, Visual Commun. Image Process., vol. 5150, pp. 198-209, 2003.

[6] Hamid R. Sheikh, and Alan C. Bovik, Image information and visual quality, IEEE Trans. Image Process., vol. 15, pp. 430-444, 2006.

[7] Eric C. Larson, and Damon M. Chandler, Most apparent distortion: full-reference image quality assessment and the role of strategy, J. Electronic Imaging, vol. 19, pp. 011006, 2010.

[8] Rajiv Soundararajan, and Alan C. Bovik, RRED indices: Reduced reference entropic differencing for image quality assessment, IEEE Trans. Image Process., vol. 21, pp. 517526, 2012.

[9] Qiang Li, and Zhou Wang, Reduced-reference image quality assessment using divisive normalization-based image representation, IEEE J. Sel. Top. Signal Process., vol. 3, pp. 202211, 2009.

[10] Zhou Wang, Guixing Wu, Hamid R. Sheikh, Eero P. Simoncelli, En-Hui Yang, and Alan C. Bovik, Quality-aware images, IEEE Trans. Image Process., vol. 15, pp. 1680-1689, 2006.

[11] Feng Pan, Xiao Lin, Susanto Rahardja, Weisi Lin, E. Ong, Susu Yao, Zhongkang Lu, and Xiaokang Yang, A locally-adaptive algorithm for measuring blocking artifacts in images and videos, in Proc. IEEE Int. Symp. Circuits Syst. , vol. 3, pp. III-925, 2004.

[12] Alan C. Bovik, and Shizhong Liu, DCT-domain blind measurement of blocking artifacts in DCT-coded images, in Proc. IEEE Int. Conf. Acoust. Speech Signal Process., vol. 3, pp. 1725-1728, 2001.

[13] Xiaojun Feng, and Jan P. Allebach, Measurement of ringing artifacts in JPEG images, Digital Publishing, vol. 6076, pp. 60760A, 2006.

[14] Pina Marziliano, Frederic Dufaux, Stefan Winkler, and Touradj Ebrahimi, A no-reference perceptual blur metric, in Proc. IEEE Int. Conf. Image Process., vol. 3, pp. 57-60, 2002.

[15] Anish Mittal, Anush K. Moorthy, and Alan C. Bovik, Blind/referenceless image spatial quality evaluator, in Proc. Annu. Asilomar Conf. Signals Syst. Comput., pp. 723-727, 2011.

[16] Nobuyuki Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst., Man, Cybern., vol. 9, pg. 62-66, 1979.

[17] Zhenhua Hu, Peter Bauer, Todd Harris, and Jan P. Allebach, Automatic Detection of Scanned Page Orientation, Color Imaging XXXIV: Displaying, Process., Hardcopy, Appl., (Part of IS&T Electron. Imaging 2019), pp.87-1, Jan. 2019.

Author Biography

Zhenhua Hu received his BS in biomedical engineering from the Shandong University (2011) and his MS in biomedical engineering from Zhejiang University (2014). Currently, he is a Ph.D. candidate at Purdue University. His work has focused on image processing, computer vision, and image quality.