An Objective Distortion Measure for Binary Document Images Based on Human Visual Perception

Haiping Lu, Jian Wang, Alex C. Kot
School of EEE
Nanyang Technological University
Singapore 639798
EHPLu@NTU.EDU.SG

Yun Q. Shi
Department of ECE
New Jersey Institute of Technology
Newark, NJ07102, USA
Shi@ADM.NJIT.EDU

Abstract

As we are moving to a digital world, digital document image processing is receiving more and more attention. Digital document images are essentially binary images. In applications related to binary document images, such as data hiding and watermarking in binary images, distortion may be present and it is necessary to measure the distortion for performance comparison. However, traditional objective distortion measures cannot describe the distortion in binary images well to have a good match with human visual perception. In this paper, we present a novel objective distortion measure for binary document images that well correlates to the subjective distortion perception. This measure is based on the reciprocal of distance that is straightforward to calculate. Our results show that the proposed distortion measure matches well with subjective evaluation found on human visual perception.

1. Introduction

Digital documents receive more and more attention recently. Digitized documents not only reduce physical storage required but also enable easier backup, searching and retrieval. One related application is data hiding in binary images [13, 6, 8]. One of the important requirements in data hiding is imperceptibility after hiding. In other words, data hiding requires low distortion. Other digital document image processing may also introduce some distortions in the output document images.

There are two ways to measure visual distortions, as discussed in [10]. One is subjective measurement and the other is objective measurement. Subjective measurement is important since human is the ultimate viewer. However, it is very costly and different observers may have different measure of distortions. On the other hand, objective measurement is repeatable and easier to implement. However, such a measure may not be reliable sometimes in the sense that it doesn't always agree with the subjective measurement.

Several authors have discussed the gap between subjective and traditional objective distortion measures and they proposed solutions for objective distortion or quality measures for video or multi-level images [12, 9, 2, 1, 4, 5, 11]. There are also measures proposed to evaluate quality of halftone images [7].

Webster et al. [12] introduced an objective measurement of video quality based on human visual perception. The original video, taken from a library of test scenes is passed to an impairment generator to get a degraded video. Both the original video and the degraded video are then passed to an objective testing, which gives objective test results, and a subjective testing, which gets the assessment from a viewing panel. Statistical analysis is done on the objective test results and the viewing panel results to determine the quality assessment algorithm.

In this paper, we propose an objective distortion measure for binary document images that is based on human visual perception. The distance between pixels is found to play an important role in human perception of distortions in binary document images. Hence, the reciprocal of distance is used to measure distortions in these images. An approach similar to that in [12] is taken to test the distortion measure proposed and the results show a good correlation between the proposed objective measure and human visual perception.

2. Traditional objective distortion measures

There are several traditional objective distortion measures that are widely used. They are mean square error (MSE), signal-to-noise ratio (SNR), and peak signal-to-noise ratio (PSNR) [10, 2]. For an image processing system with $f(x, y)$ as the input image and $g(x, y)$ as the processed output image, the distortion $d(x, y)$ is obtained from the difference between the input and output images:

$$d(x, y) = g(x, y) - f(x, y) \tag{1}$$

Hence,

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} d(x, y)^2 \tag{2}$$

where M and N are the dimensions of the image.

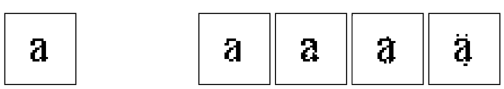
The corresponding SNR and PSNR are defined as [2]:

$$SNR(dB) = 10 \log_{10} \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} d(x, y)^2} \tag{3}$$

$$PSNR(dB) = 10 \log_{10} \frac{P^2 MN}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} d(x, y)^2} \tag{4}$$

where P is the maximum peak-to-peak signal swing. E.g. P is 255 for 8-bit images.

It can be seen that SNR and PSNR are all MSE based and these three measures are essentially equivalent. For binary document images, these traditional distortion measures are not well matched with subjective assessment. For instance, a simple document image is shown in Fig. 1a, and all four images in Fig. 1b have 4-pixel difference from the original image in Fig. 1a, which means that they have the same MSE, SNR, and $PSNR$. However, the distortions perceived by human eyes are quite different for the four distorted images.



(a) Original document image (b) Distorted images

Figure 1. Distortion in document image

3. Distance-reciprocal distortion measure

A number of single-letter images are used to study distortions in binary document images. Each single-letter image is converted from a letter typed in MS Word with font size of 10 or 12, including both uppercase and lowercase, using Adobe Acrobat 5.0 with a resolution of 150 dots per inch (dpi). One of them is shown in Fig. 1a.

It is observed that for a binary document image, the distance between two pixels plays a major role in their mutual interference perceived by human eyes. The nearer the two pixels are, the more sensitive it is to change one pixel when focusing on the other pixel. On the other hand, from a magnified viewing, each pixel is essentially a black or white square. Therefore, a diagonal neighbor pixel is considered to be further away from a pixel in focus than a horizontal or vertical one. Hence, diagonal neighbors have less effect on a center pixel in focus than horizontal or vertical neighbors [3].

Based on these observations, an objective distortion measure is proposed here for binary document images. This method measures distortion using a weighted matrix with each of its weights determined by the reciprocal of a distance measured from the center pixel, and we name it as distance-reciprocal distortion measure (DRDM) method.

The weight matrix $\mathbf{W}_m$ is of size $m \times m, m = 2n + 1, n = 1, 2, 3, 4, 5, ...$ The center of this matrix is at $i_C = j_C = (m + 1)/2$. $\mathbf{W}_m(i, j)$, $1 \leq i, j \leq m$, is defined as following:

$$\mathbf{W}_m(i, j) = \begin{cases} 0 & \text{for } i = i_C \text{ and } j = j_C \\ \frac{1}{\sqrt{(i-i_C)^2 + (j-j_C)^2}} & \text{otherwise.} \end{cases} \tag{5}$$

This matrix is normalized to form the normalized weight matrix $\mathbf{W}_{Nm}$.

$$\mathbf{W}_{Nm}(i, j) = \frac{\mathbf{W}_m(i, j)}{\sum_{i=1}^{m} \sum_{j=1}^{m} \mathbf{W}_m(i, j)} \tag{6}$$

The weight matrices before and after normalization are shown below for $m = 5$:

Table 1. Weight matrix before normalization

$\frac{1}{2\sqrt{2}}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{2\sqrt{2}}$
$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{2}}$	1	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{5}}$
$\frac{1}{2}$	1	0	1	$\frac{1}{2}$
$\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{2}}$	1	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{5}}$
$\frac{1}{2\sqrt{2}}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{2\sqrt{2}}$

Table 2. Weight matrix after normalization

0.0256	0.0324	0.0362	0.0324	0.0256
0.0324	0.0512	0.0724	0.0512	0.0324
0.0362	0.0724	0	0.0724	0.0362
0.0324	0.0512	0.0724	0.0512	0.0324
0.0256	0.0324	0.0362	0.0324	0.0256

Suppose that there are S flipped pixels in $g(x, y)$, each pixel will have a distortion $DRD_k, k = 1, 2, 3, ..., S$. For the kth flipped (from black to white or from white to black) pixel at $(x, y)_k$ in the output image $g(x, y)$, the resulted distortion is calculated from an $m \times m$ block $\mathbf{B}_k$ in $f(x, y)$, which is centered at $(x, y)_k$ and m is the size of the weight matrix used. The distortion measure DRD_k for this flipped pixel $g[(x,y)_k]$ by the proposed DRDM method is given by

$$DRD_k = \sum_{i,j} [\mathbf{D}_k(i,j) \times \mathbf{W}_{Nm}(i,j)] \tag{7}$$

where the elements of the difference matrix $\mathbf{D}_k$ are given by

$$\mathbf{D}_k(i,j) = \left|\mathbf{B}_k(i,j) - g[(x,y)_k]\right| \tag{8}$$

For possibly flipped pixels near the corner, where an $m \times m$ neighborhood may not exist, it is possible to expand the rest of $m \times m$ neighborhood with the same value as $g[(x,y)_k]$.

The distortion in $g(x,y)$ is calculated as:

$$DRD = \frac{\sum_{k=1}^{S} DRD_k}{NUBN} \tag{9}$$

where $NUBN$ is defined as the number of non-uniform (not all black or white pixels) 8×8 blocks in $f(x,y)$. The total pixel number $M \times N$ is not used in the denominator because uniform areas (e.g. all white pixel blocks) are quite common in binary document images and they may have significant effects on the value of the distortion measure.

The proposed DRDM method provides an efficient way to measure distortion in binary document images. On the other hand, it is superior over MSE, SNR or PSNR in the sense that it takes human visual perception into account and hence correlates to subjective assessment, which is the ultimate judge on distortions.

4. Experimental results

To test how well the distortion measure proposed is matched with human visual perception, experiments have been done using an approach similar to that in [12].

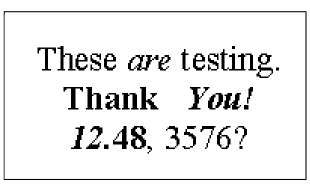


Figure 2. Original binary document image

A test image shown in Fig. 2 is designed to be the original binary document image, converted from MS Word in the same way as in section 3, and the characters in the image are with different fixed-size fonts.

The design of a number of independent test images is important. The design criteria is that under the constraint that the number of flipped pixels is the same in each test image, test images generated should have a wide variety in terms of how noticeable the flipping is. After careful testing, we choose to flip 40 pixels in the original image of size 198×109 with 1763 black pixels as described below:

1. The positions of all 1763 black pixels are recorded in a 2×1763 matrix.

2. 40 black pixels out of 1763 are randomly chosen using a random number generator with uniform distribution.

3. For each black pixel chosen, one pixel is flipped in its neighboring area, including itself. The pixel to be flipped is randomly selected from the black pixel itself, its corresponding eight 3×3 neighbors, or sixteen 5×5 neighbors, with probability of $p1$, $p3$ and $p5$, respectively. For the two latter cases, one neighbor is randomly chosen from the eight or sixteen neighbors.

4. Images generated with the number of flipped pixels less than 40 are ignored. This is possible when one pixel is flipped two or more times.

5. A large number of test images are generated with various $p1$, $p3$, and $p5$.

Since all the test images generated from the same original image have the same number of flipped pixels, they have the same MSE, SNR and $PSNR$. The $PSNR$ value is $27.32dB$. One set of the test images generated is shown in Fig. 3.

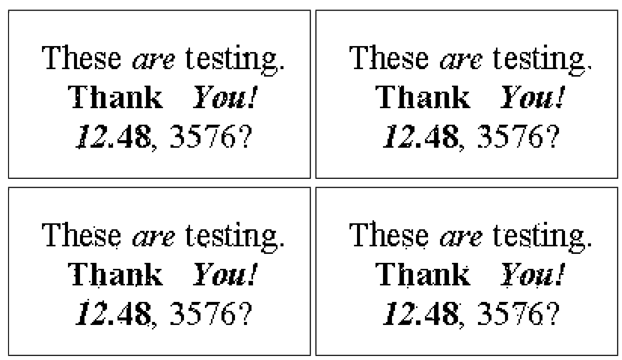


Figure 3. One set of test images

Next, all the test images generated are divided into four groups (1,2,3,4) based on the proposed DRDM method, with group 1 having lower DRD and group 4 having higher DRD. The subjective assessment is done by 60 observers. Each observer is given the original image and four sets of test images, which are printed on an 80 GSM A4-size paper using HP LaserJet 4100 printer. Each set of test images consists of four test images randomly chosen from the four groups. The observer is asked to rank the four images in each set according to the distortions that he or she perceives. There are four rankings (1,2,3,4) with score 1 for the least distortion and 4 for the most distortion perceived.

The ranking scores collected from the 60 observers are analyzed and compared with the rankings according to

DRD using the DRDM method with $m = 5$, as shown in Table 3. In the table, a smaller value means less distortion. The distribution of the subjective ranking scores for each group is shown in Fig. 4. There are 240 scores in total for each group.

Table 3. Experimental results

Test images	Subjective Assessment		$DRD(m = 5)$ $(DRDM)$
	Mean	Standard deviation	
Group 1	1.5333	0.7136	0.1565
Group 2	1.8375	0.7450	0.1869
Group 3	3.0333	0.7253	0.2098
Group 4	3.5958	0.7482	0.2413

From the table, it can be seen that the distortion calculated from the DRDM method correlates well with subjective assessment by human eyes. Although $PSNR$, MSE, and SNR are the same for all the test images, the DRD obtained is different for different distorted images and the DRD measure indicates the amount of distortion perceived by human eyes.

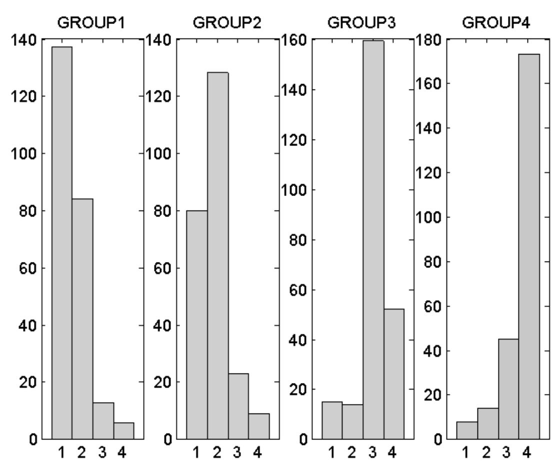


Figure 4. Subjective ranking scores

5. Conclusion

In this paper, we propose an objective distortion measure for binary document images based on human visual perception. This measure is from the observation that for binary document images, the distance between pixels plays a major role in their visual interference, and it is called the distance-reciprocal distortion measure. Experimental results have shown its correlation with subjective assessment. This measure is useful in a wide range of applications involving binary image distortions, such as data hiding in binary images, lossy binary image compression, facsimile transmissions and other digital binary document image processing.

It is worthwhile to point out that the distortion measure proposed in this paper is suitable for binary images excluding halftone images. For halftone images, graininess is desired hence a different measure is needed.

References

[1] N. Chaddha and T. H. Y. Meng. Psycho-visual based distortion measures for monochrome image and video compression. In *1993 Conference Record of the Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 841–845, 1993.

[2] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik. Image quality assessment based on a degradation model. *IEEE Transactions on Image Processing*, 9(4):636–650, Apr. 2000.

[3] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley Publishing Company, Inc., 1992.

[4] S. A. Karunasekera and N. G. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensitivity. *IEEE Transactions on Image Processing*, 4(6):713–724, Jun. 1995.

[5] K. M. Kim, C. S. Lee, E. J. Lee, and Y. H. Ha. Color image quantization using weighted distortion measure of hvs color activity. In *Proc. of IEEE Int. Conf. on Image Processing*, volume 3, pages 1035–1039, 1996.

[6] Y. Liu, J. Mant, E. Wong, and S. H. Low. Marking and detection of text documents using transform-domain techniques. In *Proc. of IS&T/SPIE's 11th Annual Symposium on Electronic Imaging: Science and Technology*, pages 317–328, Jan. 1999.

[7] S. Matsumoto and B. Liu. Analytical fidelity measures in the characterization of halftone processes. *Journal of the Optical Society of America*, 70(10), 1980.

[8] Q. G. Mei, E. K. Wong, and N. D. Memon. Data hiding in binary text documents. In *Proc. of SPIE Security and Watermarking of Multimedia Contents III*, volume 4314, pages 369–375, Aug. 2001.

[9] W. Osberger, A. J. Maeder, and D. McLean. An objective quality assessment technique for digital image sequences. In *Proc. of Int. Conf. On Image Processing*, pages 897–900, 1996.

[10] Y. Q. Shi and H. Sun. *Image and Video Compression for Multimedia Engineering: Fundamental, Algorithm, and Standards*. CRC Press LLC, Boca Raton, FL, 1999.

[11] P. C. Teo and D. J. Heeger. Perceptual image distortion. In *Proc. of IEEE Int. Conf. on Image Processing*, volume 2, pages 982–986, 1994.

[12] A. A. Webster, C. T. Jones, M. H. Pinson, S. D. Voran, and S. Wolf. An objective video quality assessment system based on human perception. In *Proc. of SPIE Human Vision, Visual Processing and Digital Display IV*, volume 1913, pages 15–26, Feb. 1993.

[13] M. Wu, E. Tang, and B. Liu. Data hiding in digital binary image. In *Proc. of IEEE Int. Conf. on Multimedia and Expo*, pages 393–396, New York City, July 31 to August 2, 2000.