An evaluation model for character quality on scanned image

Seul Ki Jang, Choon-Woo Kim*, Yousun Bang**, Heui-Keun Choh**, Sang Ho Kim***

**School of Information and Communication Eng., Inha University, Incheon, Korea*

***Digital Printing Division, Samsung Electronics Co., Suwon, Korea*

Abstract

Scanners, standalone or embedded in MFP, have been widely utilized for document scanning/copying in office environments. There are many factors affecting the quality of scanned image. They include quality of characters, size of gamut, degree of noises, color reproduction, tone characteristics, and gray balance, etc. This paper is focused on quantitative evaluation of character quality on scanned images. Objective of this paper is to derive a quantitative evaluation model that faithfully matches with the human perception. Key attributes determining character quality are extracted first based on subjective human visual experiments. For each of the identified attributes, a quantitative metric is developed. The proposed evaluation model is designed as a linear combination of the quantified metrics. Coefficients for linear combination are estimated by applying linear regression. Various experiments are performed to verify the performance of the proposed evaluation model.

1. Introduction

There are many factors affecting the quality of scanned image. They include quality of characters and lines, size of gamut, degree of noises, color reproduction, tone characteristics, and gray balance etc. Among these factors, this paper is focused on the quality of characters on scanned image. Various methods have been proposed to enhance the character quality for scanners and copiers. However, the assessment method for character quality is scarce. Most of previous works were to evaluate the legibility and/or quality of characters generated by printers and copiers[1-4]. Legibility and quality of characters are closely related. However, it should be noted that legibility is often referred to the smallest size of recognizable characters whereas quality of characters represents level of faithful reproduction of originals. In [5], legibility of characters on the scanned image was determined by counting legible characters. Also, the quantitative measures for the character quality were mainly based on the quality of vertical, horizontal and diagonal lines[1-4]. In [4], a linear combination of blurriness, contrast and stroke width is proposed based on the observers' comments collected after visual experiments. However, experiments were performed based on characters having vertical and horizontal lines such as "I", "L" and "T". Furthermore, defects on characters were not considered.

This paper is focused on quantitative evaluation model of character quality on scanned image. Background density on the scanned image shows wider variation than those on the printer/copier outputs. Also, characters appear blurred due to the low pass nature of scanning procedure. Procedure to develop a quantitative evaluation model is similar to [4]. Five attributes determining character quality are extracted first based on subjective human visual experiments. They are sharpness, contrast, thickness, darkness and non-homogeneity. These attributes represent the level of clear reproduction of characters as well as defects of characters such as discontinuities, voids, graininess, and missing serif, etc. For each of five attributes, a quantitative metric is developed. The proposed evaluation model is designed as a linear combination of the quantified metrics.

In section 2, the proposed method is described. Subjective human visual experiments and calculation of five attributes are explained. Also, design of a quantitative evaluation model is described. In section 3, experimental results of the proposed method are presented. Performance of the proposed method is discussed. Finally, the conclusion is addressed in section 4.

2. Proposed evaluation method for character quality

Figure 1 shows the flow chart of procedure to develop the proposed evaluation model for character quality. Subjective human visual experiments were performed. Offset printed characters with 10pt Times new roman font were scanned by various scanners. Scanners were set at the factory default mode and scanning resolution was 300dpi. Visual experiments were performed in typical office viewing environment. Scanned images were displayed on the sRGB monitor for the paired comparison. Ten observers were asked to grade the quality (1=good, 0.5=same, 0=bad). Figure 2 shows example of the scanned images arranged in descending order of the averaged scores.

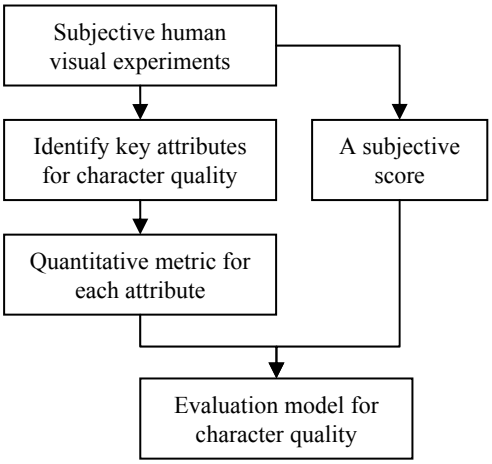


Figure 1. *Flow chart of the proposed evaluation method.*

A. Key attributes for character quality

Based on the comments by the human observers participated in visual experiments, five attributes are identified for the quantification of character quality. They are sharpness, contrast, thickness, darkness and non-homogeneity. These attributes represent the level of clear reproduction of characters as well as defects of characters such as discontinuities, voids, graininess, and

missing serif, etc. Scanned image is segmented into 3 subareas. They are edge, character, and background area. Edge area consists of the boundary pixels of characters and their immediate neighbors in background. Character and background areas are extracted to maximize the inclusion of pixels in their own subareas. Figure 3 shows examples of these subareas. Figure 4 shows how 3 subareas are utilized for calculation of 5 identified attributes.

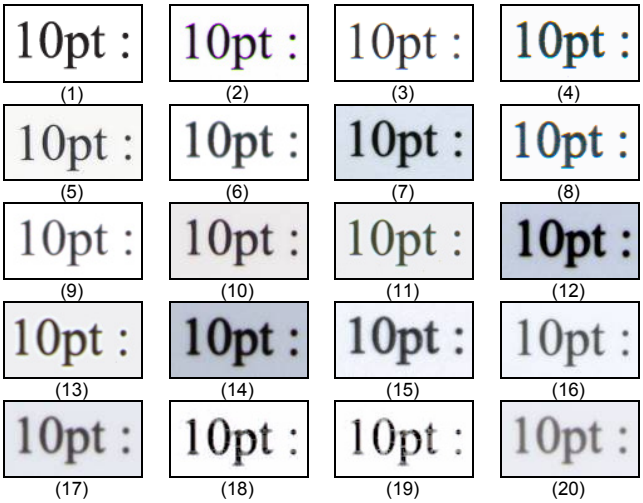


Figure 2. *Scanned image: 10 pt Times new roman*

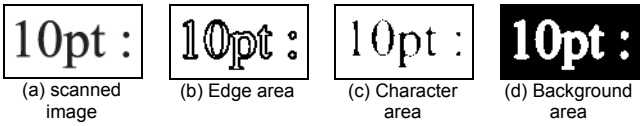


Figure 3. *Example of scanned image and its subareas.*

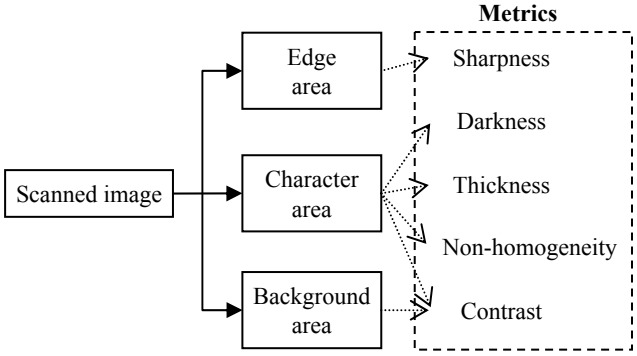


Figure 4. *Flow chart for the calculation of five metrics.*

Five key attributes are defined as follows ;

• Sharpness is defined as the standard deviation of gray levels in edge area. Higher value of standard deviation implies sharper edges of characters.

• Contrast is calculated based on the average gray levels in character and background areas. It is defined as

$$C = \frac{g_b - g_c}{g_b + g_c} \quad (1)$$

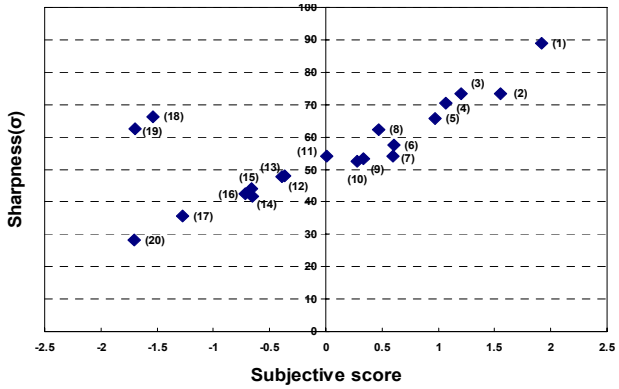
where g_b and g_c represent the average gray levels in background and character areas, respectively.

• Darkness is defined in this study to overcome the limitation of contrast. As shown in Figure 2, background density on the scanned image shows wide variations. Even though two scanned images have similar value of contrast, density levels of character and background area may appear different. For examples, Figure 2 (10) and (11) have similar values of contrast. However, density levels of character and background are quite different. Darkness is defined as the average of gray levels in character area.

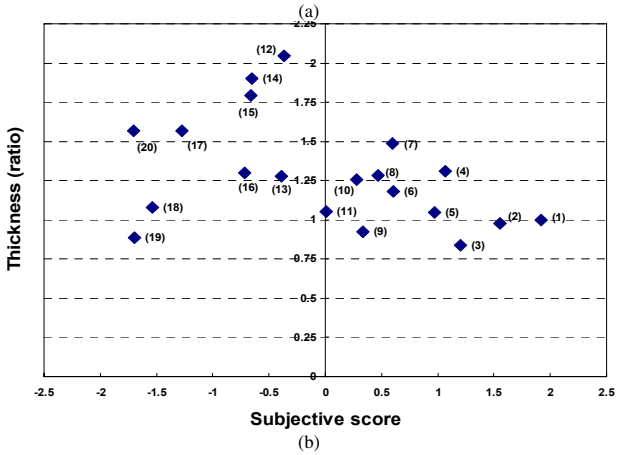
• Thickness is defined in terms of number of pixels instead of width of strokes. It is defined as a ratio of the number of ideal character pixels to the number of pixels in real character area. In this way, defects such as discontinuities, voids, and missing serif can be considered in metrics.

• Non-homogeneity is defined as the standard deviation of gray levels in character area. Smaller values of non-homogeneity represent homogeneous density distribution in character area. Defects such as graininess can be incorporated in this attribute.

Figure 5 shows the relation between the averaged value of subjective scores and the calculated metrics. From Figure 5 (a), it can be noted that the sharpness and subjective scores are highly correlated except Figure 2 (18) and (19) exhibiting the defects. For Figure 2 (18) and (19), the values of thickness and darkness are quite small. However, they have large values of the non-homogeneity. Based on Figure 2 and Figure 5, it can be said that visual characteristics of the scanned images are faithfully reflected in the five calculated metrics.



(a)



(b)

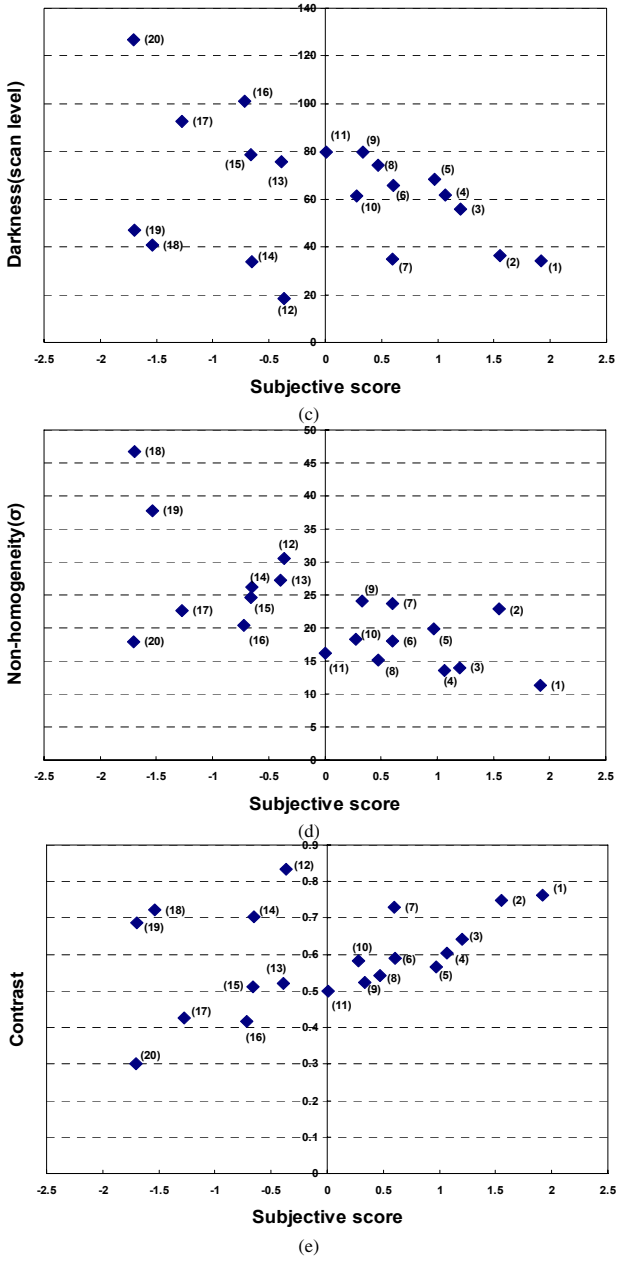


Figure 5. *Correlation between quantified metrics and subjective scores.*

B. Evaluation model

In order to design evaluation models, five key attributes are identified. Quantified metrics for five attributes are proposed. An evaluation model is designed as a linear combination of the quantified metrics. Coefficients for linear combination are estimated by applying linear regression. An evaluation model for character quality is defined as

$$M = 0.523\times D + (-1.293)\times N + 4.485\times S + (-0.848)\times T + 5.111\times C \quad (2)$$

where D, N, S, T, and C represent Darkness, Uniformity, Sharpness, Thickness and Contrast, respectively. Figure 6 illustrates the calculated values by Eq. (2) and the averaged subjective scores. Correlation is 0.93. When Figure 2 (18) and (19) are excluded, the correlation improves to 0.98.

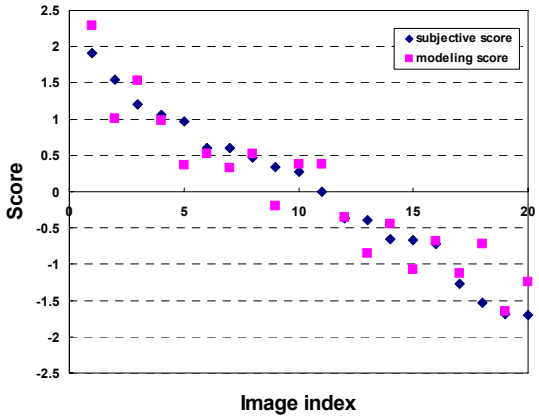


Figure 6. *Calculated value of the proposed model and averaged subjective scores.*

3. Experimental results

In this section, experimental results of the proposed method are described. Stability of the proposed model is examined by applying Eq. (2) to test images having different characters. Also, dependency on the size of characters is evaluated by applying Eq. (2) to the scanned images of 5 pt characters.

A. Experiment with Test Samples

Figure 7 shows the test images that are utilized in this experiment. Five images in Figure 7 are arranged in descending order of the averaged scores of visual tests. These images are not utilized for designing the proposed evaluation model in Eq. (2). Figure 8 illustrates the calculated values of five test images by Eq. (2) and the averaged subjective scores. As shown in Figure 8, the calculated values by the proposed evaluation model faithfully match with the averaged scores of the human visual tests. Moreover, the correlation between the quantified measure and subjective scores for testing images is 0.99. It implies that the proposed evaluation model can be applied to different characters.

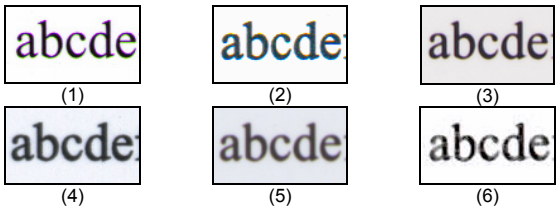


Figure 7. *Test image: 10 pt Times new roman for testing*

B. Experiment with 5 pt characters

Feasibility of applying Eq. (2) to scanned images with different sized characters is examined in this experiment. Figure 9 shows test images having 5 pt characters. The proposed model in Eq. (2) is applied to the scanned images in Figure 9. Correlation between the quantified measure and subjective scores for the 5 pt images is 0.85. The value of correlation is somewhat decreased. However, it shows the feasibility of applying the proposed model to the scanned images having different sized characters.

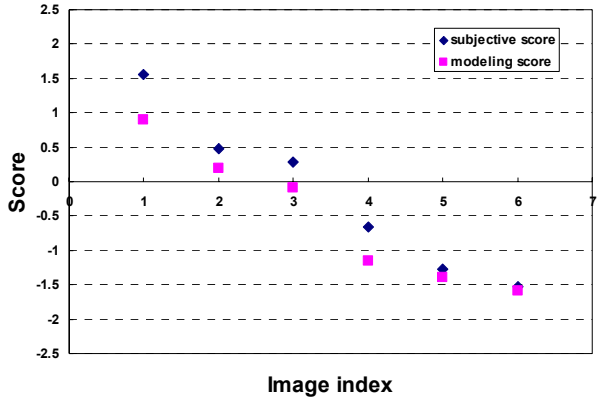


Figure 8. *Calculated value of the proposed model and averaged subjective scores for test images.*

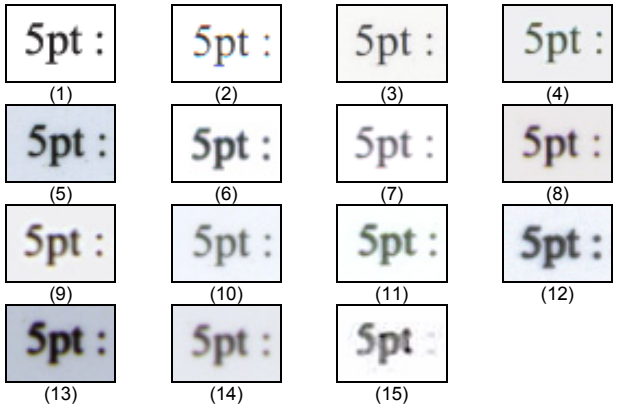


Figure 9. *Test image: 5 pt Times new roman for testing*

4. Conclusion

In this paper, a quantitative evaluation model for character quality on scanned image is proposed. In order to design the evaluation model, subjective human visual tests are performed. Five attributes determining character quality are extracted first based on subjective human visual experiments. They are sharpness, contrast, thickness, darkness and non-homogeneity. These attributes represent the level of clear reproduction of characters as well as defects of characters such as discontinuities, voids, graininess, and missing serif, etc. For each of five attributes, a quantitative metric is developed. The proposed metrics can be applied regardless of the shape of characters. The proposed evaluation model is designed as a linear combination of the quantified metrics. Experimental results indicate that the proposed evaluation model is stable and yields high value of correlation.

References

[1] T. Itoh, S. Hirota, Evaluation of Legibility, PICS, pg. 213. (1999)
[2] T. Itoh, S. Hirota, Evaluation of electro photographic legibility using line reproduction quality and MTF, JIST, vol. 43, pg. 478-483. (1999)
[3] ISE/IEC 13660:2001 Information Technology – Office equipment – Measurements of image quality attributes for hardcopy output – Binary monochrome text and graphic images.
[4] M. K. Tse, A predictive model for text quality: analysis case study, NIP, pg. 419-423. (2007)
[5] ISE/IEC 12653-2:2000, Electronic imaging – Test target for the black and white scanning of office documents – Part2: Method of use.

Acknowledgement

This work was supported by research grant from Digital Printing Division, Samsung Electronics.