Generative Learning for Character Recognition of Uneven Lighting

Tomohiko Tsuji†, Masakazu Iwamura†, and Koichi Kise†
†Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuencho, Naka, Sakai, Osaka 599-8531, Japan
E-mail: tsuji@m.cs.osakafu-u.ac.jp, {masa,kise}@cs.osakafu-u.ac.jp

1. Introduction

Recently, development of a text input system based on a text recognizer with a digital camera or a camera phone is highly demanded. However, character images captured by a camera are often degraded by low resolution, blurring, uneven lighting, camera shake and so on. A promising method to prevent decrease in recognition performance is generative learning. It enhances the recognition performance by creating various artificially degraded images that can be obtained in real scenes from degradation-free images. The created images are used for learning together with the degradation-free images, and compensate for lack of actual degraded images. The degradation processes should be well analyzed to emulate them as real as possible. Ishida et al. proposed degradation models of low resolution and blurring[1]. In this paper, we propose a degradation model of uneven lighting and experimentally confirm its effectiveness.

2. Proposed Method

The degradation of uneven lighting is modeled with two parameters: an intensity l, $0 \le l \le 256$, and an angle θ, $0 \le \theta < 2\pi$. Examples of the proposed uneven lighting filter are shown in Figs. 1 and 2. The filter $f(p,q)$ is defined as follows.

$$f(p,q) = 1 - \frac{l}{256} \frac{\left(p - \frac{P}{2}\right)\sin\theta + \left(q - \frac{Q}{2}\right)\cos\theta + \frac{Q}{2}}{Q}, \tag{1}$$

where P and Q are the width and height of the image. $f(p,q)$ is defined in the range of $0 \le f(p,q) \le 1$ and takes 1 at a well-lighted place and 0 at a dark place. Finally, the degraded character image is created by multiplying pixel values of the degradation-free character image and the filter $f(p,q)$ for each pixel (p,q).

3. Experiments

We carried out an experiment of comparing four methods. Methods A and C are conventional methods and methods B and D are the proposed method. Fig. 3 shows a part of training samples of the methods.

Method A used one degradation-free character image per character for training. This method does not use degraded images.

Method B used 65 degraded images created by the proposed method from the degradation-free character image shown in Fig.3(a). The used parameters were combinations of nine intensities, $l = 0, 32, \ldots, 256$, and eight angles, $\theta = 0, \pi/4, \ldots, 7\pi/4$. Since $l = 0$ means even lighting and the angle parameter θ does not make sense at $l = 0$, only one image was created for $l = 0$.

Method C used 120 degraded images created by the degraded models of low resolution and Gaussian blurring proposed by Ishida et al.[1]. The used parameters were combinations of six standard deviations σ

(a) $l = 256$ (b) $l = 192$ (c) $l = 128$

Fig 1. Change of the proposed illumination filter for the intensity l. θ is fixed to 0.

(a) $\theta = 0$ (b) $\theta = \pi/4$ (c) $\theta = \pi/2$

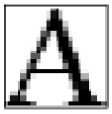
Fig 2. Change of the proposed illumination filter for the angle θ. l is fixed to 256.

of Gaussian blurring, $\sigma = 0, 0.2, \ldots, 1$, and 20 scales, $s = 13, 14, \ldots, 32$.

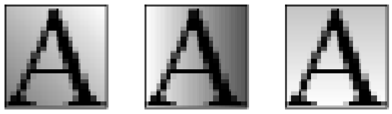
Method D used 510 degraded images created by the proposed method and the degraded models used in the method C. The proposed method was used first and then Ishida's degradation models were used. The used parameters were combinations of three intensities, $l = 0, 128, 256$, eight angles, $\theta = 0, \pi/4, \ldots, 7\pi/4$, three standard deviations σ of Gaussian blurring, $\sigma = 0, 0.5, 1$, and 10 scales, $s = 13, 15, \ldots, 31$.

For the degradation-free training images, 62 character images in Times New Roman whose size was 256pt were used; 62 characters consist of 10 figures, 26 lowercase alphabets and 26 capital alphabets.

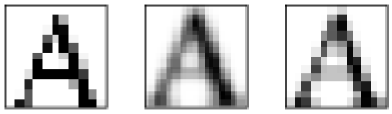
Test images were extracted from a movie captured by a digital camera. The movie was 320×240 pixels and its frame rate was 30 fps. Each character was printed on an A4 paper and put on a dark place. Then, it was unevenly illuminated by a fluorescent bulb. Changing angles and intensities of the light, three movies were captured per character at the distance of 120cm. Then, 10 still test images were extracted from a movie. Finally, they were segmented by hand and normalized to fit to a 32×32 square. Examples of test images are shown in Fig. 4.



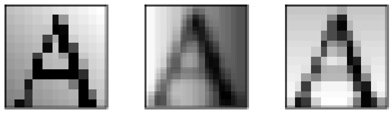
(a) Method A



(b) Method B



(c) Method C



(d) Method D

Fig 3. A part of training images.

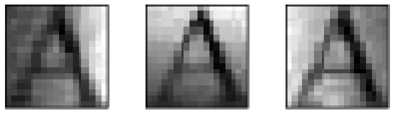


Fig 4. A part of test images extracted from movies. The movies were captured by a digital camera under various illuminations.

The recognition result was determined by voting; each of 10 still images extracted from a movie was recognized by subspace method using 1024 dimensional feature vectors consists of pixel values of the image. Fig. 5 shows the recognition result. By comparing the results of the methods A and B, we can confirm that the proposed method using the generative learning achieved a higher recognition rate. In addition, comparison of the methods C and D shows the combination of the proposed model and other models also succeeded because the method D achieved higher recognition rate. These results show that the proposed method is effective for recognition of degraded character images.

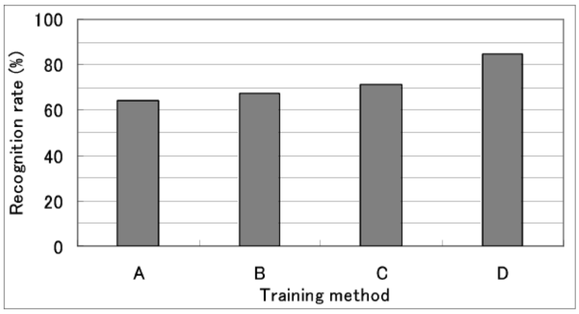


Fig 5. Recognition result.

4. Conclusion

In order to improve recognition performance, we proposed a degradation model of uneven lighting for generative learning. We confirmed the proposed method increases recognition rates by experiments. Future work includes use of automatic segmentation methods.

Acknowledge This research was supported in part by the Grant-in-Aid for Young Scientists (B) (19700177) from Japan Society for Promotion of Science.

Key words degraded character recognition, generative learning, uneven lighting, degradation model.

References

[1] H. Ishida, S. Yanadume, T. Takahashi, I. Ide, Y. Mekada, and H. Murase, "Recognition of low-resolution characters by a generative learning method," Proc. 1st Intl. Workshop on Camera-Based Document Analysis and Recognition, pp.45-51, Seoul, Korea, August 2005.