

Automatic Extraction of Filled-In Items from Bank-Check Images

Katsuhiko Ueda, Hirotohi Maegawa, and Kenichi Matsuo

Department of Information Engineering, Nara National College of Technology
Yata-cho 22, Yamatokoriyama, Nara 639-1080, Japan
{ueda,matsuo}@info.nara-k.ac.jp

Abstract. In this paper we propose a new method to extract separately filled-in items from Japanese bank-checks based on prior knowledge about their color characteristics and layout. We have analyzed the bank-check characteristics and proposed a model that can be used to extract the filled-in items applicable to any Japanese bank-checks. The areas where the filled-in item is supposed to appear are first extracted through a template. Then the filled-in characters and seal imprints are extracted on the basis of their color characteristics in HSV color space. The results of testing experiments show that this extraction method is capable of extracting the filled-in items from most Japanese bank-checks.

1 Introduction

Automatic bank-check reading and verifying are an active topic in the field of document analysis. Many researches on bank-check processing have been performed up to the present [1], [2]. In order to develop a bank-check processing system, several modules such as those for extraction of filled-in items, for recognition of handwritten and printed characters and for verification of signature or seal imprint must be integrated. Especially, the extraction of filled-in items is an important preprocessing step to facilitate subsequent high accuracy recognition and verification. Therefore there are many reports on the extraction of filled-in items from bank-checks [1], [3], [4].

However these methods cannot be applied directly to Japanese bank-checks due to their specific properties. In the case of Japanese bank-checks, a signature (generally a stamp of payer's name) is written or stamped on a bank-check, which has a various background pattern. Then a seal is stamped superimposed over the signature. Fig. 1 shows an example of Japanese bank-checks. The seal imprint is used as a more important clue than the signature for validity confirmation of bank-checks. Therefore it is important to separately extract the signature and seal imprint from bank-checks.

In this paper we propose a new method to extract accurately the filled-in items from Japanese bank-checks based on the above background. This method extracts the filled-in items based on their color characteristics in HSV color space. We will also present that the proposed method can extract the filled-in items with good quality for subsequent recognition and verification processes.

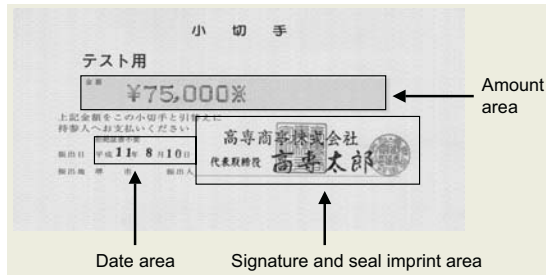


Fig. 1. An example of Japanese bank-checks.

2 Proposed Method

The flow diagram of the proposed method is shown in Fig. 2. Since the size and layout of all items of Japanese bank-checks have been standardized as shown in Fig. 1, we first extract the areas to be processed through a template. Then the extracted areas are converted into HSV color space. Successively filled-in characters (amount, date and signature) and seal imprint are extracted separately based on their color characteristics in HSV color space. Finally isolated noises and preprinted characters are removed from the extracted item images.

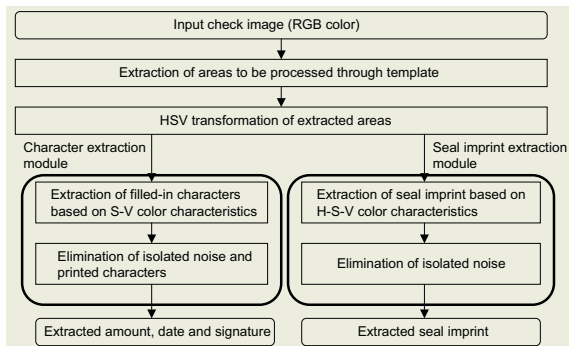


Fig. 2. Flow diagram of the proposed extraction method.

Japanese bank-checks have a common feature in respect to color; the background patterns are various light colors, the signatures are black or blue in color, and the seal imprints are vermilion or red in color. Therefore it is reasonable to assume that pixels of the color check image construct four clusters corresponding to the background B, the characters C, the seal imprint I, and superimposed area of signature and seal imprint IC. In order to obtain color characteristics of Japanese bank-checks, we calculated a histogram of pixels of real bank-checks in S-V color space. Fig. 3(a) shows an example of the histogram of a check image.

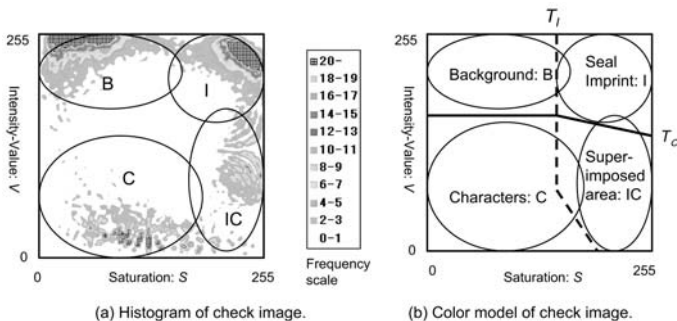


Fig. 3. Histogram and color model of check image in S-V space.

As a result of having analyzed about many checks, we confirmed that most of check images had similar histogram and constructed four clusters as shown in Fig. 3(a). Therefore we define the color model of check images in S-V space as shown in Fig. 3(b).

Character Extraction. The intensity value of characters involving the superimposed area of signature and seal imprint is lower comparing with those of other areas as shown in Fig. 3. Therefore we extract pixels of which intensity value V is threshold value T_C (Solid line in Fig. 3(b)) or less as character pixels. Let V_{max} be the peak position of the histogram of area B on the intensity value axis. Threshold T_C is defined by following equation as an experimental optimum.

$$T_C = \begin{cases} V_{max} - 50 & : S \leq 100 \\ V_{max} - 50 - \frac{1}{3}(S - 100) & : S > 100 \end{cases} \quad (1)$$

After this processing, preprinted characters are removed from the extracted image based on their position, size and line width.

Seal Imprint Extraction. Pixels of seal imprint and blue characters are located in an area of higher saturation value. Therefore we first remove blue pixels using H (Hue) value in order to extract only seal imprint. Then we extract pixels of which saturation value S is threshold value T_I (Broken line in Fig. 3(b)) or higher as seal imprint pixels. The superimposition area of signature and seal imprint can also be extracted as seal imprint, because S of the superimposition area is higher comparing with those of background and characters. Let S_{max} be the peak position of the histogram of area B on the saturation value axis. Threshold T_I is defined by following equation as an experimental optimum.

$$T_I = \begin{cases} T_{I1} : T_{I1} \geq T_{I2} \\ T_{I2} : T_{I1} < T_{I2} \end{cases} \quad (2)$$

where $T_{I1} = S_{max} + 85$, and $T_{I2} = 200 - V$.

3 Experimental Results

We performed the extraction experiment with 19 Japanese bank-checks filled up by various colors of ink which are supposed to appear in practical situations. Although the number of checks used for experiment is few, these represent the properties of most checks used in practical situations. We inspected visually the extraction results. As a result, it was confirmed that all samples in character extraction and 18 samples in seal imprint extraction were extracted precisely. Fig. 4 shows the sample result of filled-in items extracted from the check image shown in Fig. 1. All items were extracted completely. Especially, it is important that the signature and the seal imprint superimposed over each other were able to be extracted precisely by this method as shown in Fig. 4(c) and (d).



Fig. 4. Example of Filled-in items extracted from the check shown in Fig.1.

4 Conclusion

In this paper, we proposed automatic extraction method of filled-in items from Japanese bank-check images. The results of testing experiments showed that this method is applicable for most Japanese bank-checks. We will investigate more precise evaluation of this method with the large number of samples in the future.

References

1. Impedovo S., Wang P.S., Bunke H. (Eds.): Special issues on automatic bankcheck processing Parts I and II, *Int. J. Pattern Recognition and Artificial Intelligence*, Vol.11, No.4/ 5, (1997). (Including 17 papers).
2. Greco N., et al.: Bank-check processing system: Modifications due to the new European currency, *Proc. of 7th International Conference on Document Analysis and Recognition*, (2003) 343-348.
3. Franke K., Köppen M.: Towards an universal approach to background removal in image of bankchecks, *Advances in Handwriting Recognition*, World Scientific (1999) 91-100.
4. Ye X. et al.: Extraction of bankcheck items by mathematical morphology, *Intl. J. on Document Analysis and Recognition*, Vol.2, No.2/3, (1999) 53-66.