MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

5—3

Extraction of Character String Areas from Color Scenery Image using Psychological Potential Field

Masanori ANEGAWA, Sueharu MIYAHARA *
Department of Computer and Information Sciences,
Faculty of Engineering, Nagasaki University

Akira NAKAMURA †
Department of Electronic Engineering,
Faculty of Engineering, Sojo University

Abstract

As digital cameras have been used widely, a technology to extract features of character string areas from scenery (nature) image is required. However, it is very difficult to extract character string areas from a complex background of scenery image as compared with a relatively simple background of documentary image.

Human can extract of character string areas from scenery image having complex background. Therefore, a fairly high degree of accuracy in the extraction of character string areas may be attained by imitating the human visual processing system.

In this paper we proposed the method to extract character string areas from color scenery images using the psychological potential field [1],[2] (imitating the human visual processing system). We have also executed the experiment and obtained the favorable processing results that are close to the human sense and unlikely to be affected by amounts of color information and photographing conditions. And, this method have high degree of accuracy in the extraction of character string areas when it is applied to the color scenery image.

1 Introduction

This method first separates a color scenery image into the Munsell color system, which are believed to be close to the color characteristics that we can sense, and executes subsequent processing (Fig.1(a)). Next, remove noises. And, to imitate the process of the horizontal retinal cells, features are enhanced, using local complexity level (Fig.1(b)). Then, features are extracted, using the psychological potential field that is relatively agrees with human (sensory) subjective image processing results (Fig.1(c)). After these processing are completed, character string areas are extracted with the those feature extraction results judged and merged generally (Fig.1(d)).

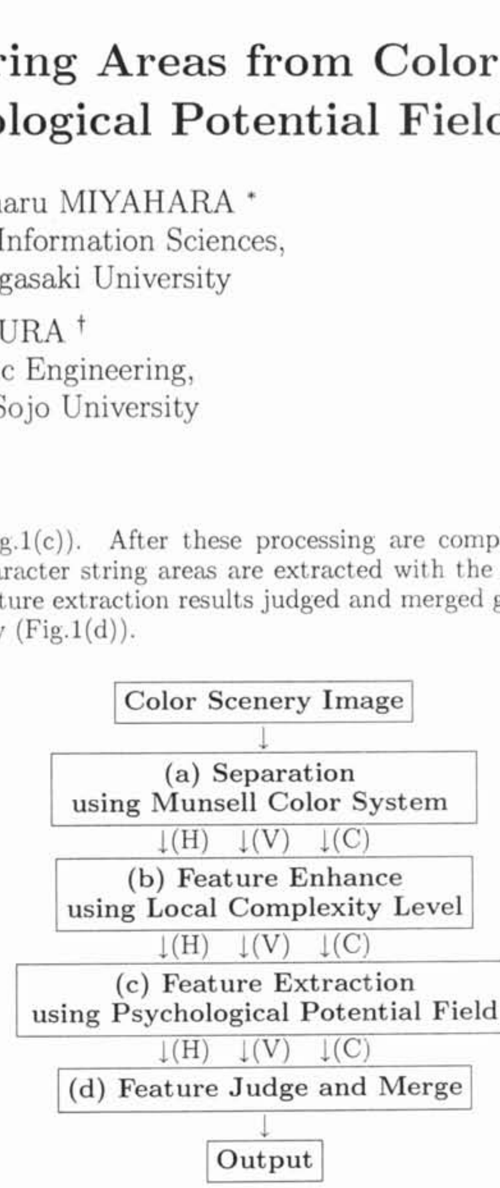


Figure 1: Flow of proposed method

2 Separation using Munsell Color System

The human color characteristics are believed to be close to the Munsell color system represented by hue (H), value (V), and chromatic (C), no to the RGB color system used generally in computers.

This method separates an input color scenery image into the Munsell color system and executes the subsequent processing for each information hue (H), value (V) and chromatic (C). The value information is mainly used for images having high contrasts, and hue and chromatic information is mainly used for colorful images. If input image is gray scale the value information is only used. This makes it possi-

*Address: 1-14 Bunkyou-machi, Nagasaki-shi, Nagasaki 852-8521 Japan. E-mail: masanori@anegawa.com

†Address: 4-22-1 Ikeda, Kumamoto-shi, Kumamoto 860-0082 Japan. E-mail: akira@ee.sojo-u.ac.jp

ble to reduce effects by amounts of color information and photographing conditions.

Fig.2(a)-(c) show the separation images hue (H), value (V) and chromatic (C) in Munsell color system.

(a)Hue image (b)Value image (c)Chromatic image

Figure 2: Separation using Munsell color system

3 Feature Enhance using Local Complexity Level

The edge elements in an image are very important characteristics. It is also generally known that the horizontal retinal cells of human are extracting and enhancing the edges.

At first, this method removed noises for each information item hue (H), value (V), and chromatic (C) separated in Chapter.2. Next, this method imitates the horizonal retainal cells process by enhancing the edge elements using local complexity level[3]. This method defines and measures the local complexity level in steps from (i) to (iii) below.

(i) The local pixel density ($D_{(x,y)}$) is measured for each local area of an input image. Then, normalization is performed for each local pixel density ($D_{(x,y)}$) to seize a local density image ($\tilde{D}_{(x,y)}$).

(ii) The local average frequency ($F_{(x,y)}$) is measured, as an approximate high frequency element, for each local area of the input image. Then, normalization is performed for each local average frequency ($F_{(x,y)}$) to seize a local average frequency image ($\tilde{F}_{(x,y)}$).

(iii) From the local pixel density image ($\tilde{D}_{(x,y)}$) and the local average frequency image ($\tilde{F}_{(x,y)}$), the local complexity level ($C_{(x,y)}$) is defined as expression (1). Then, normalization is performed for each local complexity level ($C_{(x,y)}$) to seize a local complexity level image ($\tilde{C}_{(x,y)}$).

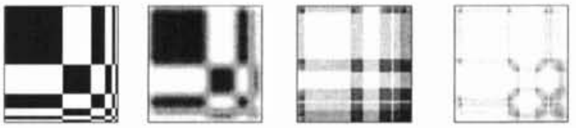
Fig.4 shows the relationship among the normalized local pixel density, local average frequency, and local complexity level.

Fig.5(a) shows an example of the local complexity level image created from the input image in the above procedure. This figure shows that the local complexity level is high in an area that contains many edge elements.

Fig.5(a) also shows that the character string areas have certain ranges (medium) of local complexity level. Therefore, the local complexity level of the range, where character string areas may exist, is emphasized (See Fig.5(b)).

The enhanced edge elements specially help to enhance character areas for the hue(H), value(V), and chromatic(C) information.

$$C_{(x,y)} = \tilde{D}_{(x,y)} \cdot \tilde{F}_{(x,y)} \quad (1)$$



(a)input (b)density (c)frequency (d)complexity

Figure 3: Example of local complexity level

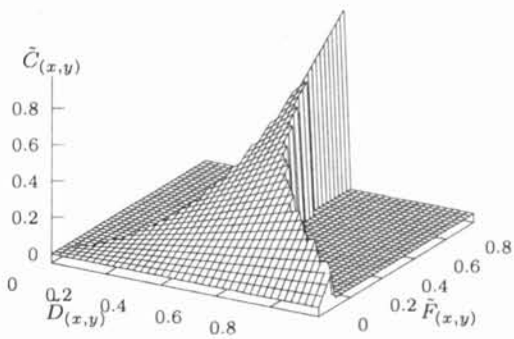
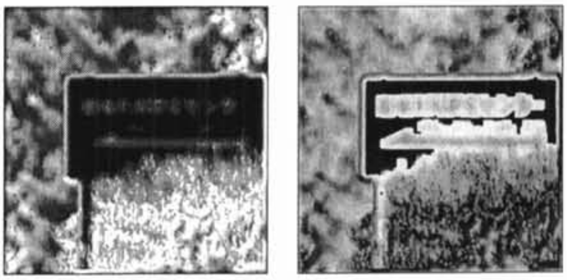


Figure 4: Relation of $\tilde{D}_{(x,y)}$, $\tilde{F}_{(x,y)}$ and $\tilde{C}_{(x,y)}$



(a)Complexity image (b)Enhanced image

Figure 5: Example of feature enhance

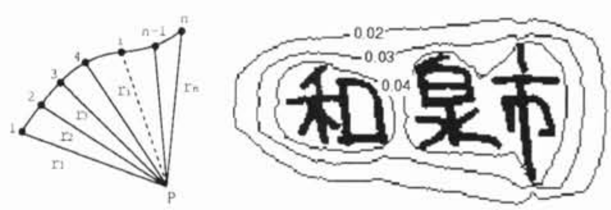
4 Feature Extraction using Psychological Potential Field

It has been proposed that DOG and LoG (Laplacian Of Gaussian) filter processing are similar to human initial image processing. But, from the standpoint of human subjective image processing, these methods are different from the present image processing method in human sense. On the other hand in cognitive science field, it is proposed that the method called psychological potential field (Field of induction on the retina) is similar to human subjective image processing[1],[2].

This method used psychological potential field to feature extraction. Psychological potential field is defined by expression(2), where r_i is distance to measurement points, n is number of measurement points (See Fig.6(a)). As shown in the psychological potential field in Fig.6(b), the potential field match to human subjective processing.

Fig.7 shows the example of the image using psychological potential field that corresponds to Fig.5(b).

$$P = \frac{1}{n}\sum_{i}^{n}\frac{1}{r_i} \qquad (0 \le P < 1) \qquad (2)$$



(a) (b)

Figure 6: Psychological potential field

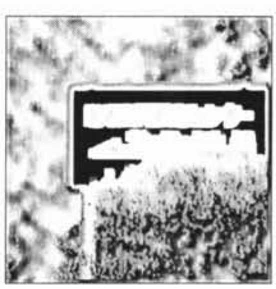


Figure 7: Example of feature extraction

5 Future Judgment and Merge

The results of preliminary experiment indicate that character string areas existing in color scenery images generally have the following features:

- The potential of each character string area is high.
- The potential around each character string area is medium (A background area exists).
- Each character string area is existing continuously.

This method, in future judgment and merge stage, extracts the areas that meet the above conditions as candidates for character string areas. Then, the hue(H), value(V), and chromatic(C) information of these areas are compared, and if the following conditions are met, the corresponding areas are extracted as character areas (See Fig.8):

- One of the information items indicates a high value.
- Two or more of the information items have the same results.

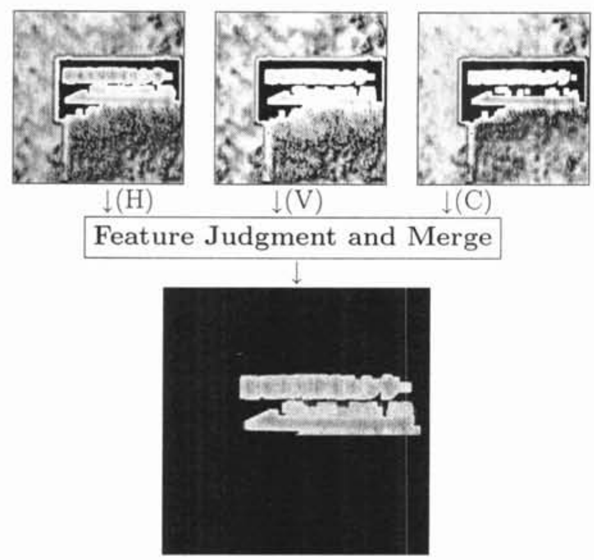


Figure 8: Feature judgment and merge

6 Results of Experiments

The results of experiments on color scenery images are shown in Fig.9. This method is similar to human subjective image processing(See Fig.9(a)).

Even if the contrast of the character string areas is too low to seize the value(V) information, those areas can be extracted by effectively using the color information (See Fig.9(b)).

(a)High contrast image (b)Low contrast image

Figure 9: Example of extraction character string areas

7 Conclusions

In this paper we proposed the method to extract character string areas from color scenery images using the psychological potential field (imitating the human visual processing system). We have also executed the experiment and obtained the favorable processing results that are close to the human sense and unlikely to be affected by amounts of color information and photographing conditions.

An investigation of unsuccessful examples showed that almost of all the failed examples were too small character or character string areas(See Fig.10). Presently, we are studying the character segmentation[4] from color scenery image using this proposed method.

Moreover, we are now comparing the effects of this technique when it is applied to color systems other than the Munsell color system ($L^*a^*b^*$ etc.) with the effects when the technique is applied to the Munsell color system.

Figure 10: Example of unsuccessful image

References

[1] Zensho YOKOSE: "A Study on Character Patterns Based Upon the Theory of Psychological Potential Field", Japanese Psychological Research, Vol.12, No.1, pp.18-25, 1970

[2] Michihiro NAGAISHI: "A Consideration on Segmentation of Handwritten Characters using Field of Induction on the Retina", Technical Report of IEICE. Japan, PRU92-120, pp.17-24, Jan.1993

[3] Masanori ANEGAWA: "A Division Algorithm of a Chinese Character into Its Component Elements using Local Complexity of Image", IEICE Transactions on Information and System. Japan, Vol.J79-D-II, No.1, pp.146-150, Jan 1996

[4] Masanori ANEGAWA: "Handwritten Chinese Character Segmentation using Local Potential Threshold and Minimum Potential Search", Proceeding of MVA' 98 IAPR Workshop on Machine Vision Applications, pp.57-60, 1998