MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

8—13 Extraction of Character String Areas from Color Scenery Image using Extended Psychological Potential Field

Masanori Anegawa[1]
Department of Electrical and Computer Engineering
Oyama National College of Technology

Akira Nakamura[2]
Department of Electronics and Computer Network
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 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 extended psychological potential field (imitating the human visual processing system).

The conventional psychological potential field method has the problem that it is hard to be applied to multilevel images such as color scenery images.

This paper proposes a new extended psychological potential field method that has improved the conventional method so that it can be applied to multilevel images such as color scenery images.

The new extended psychological potential field method proved that processing results closer to the human visual processing system could be obtained by adopting the shielding effects in accordance with the pixel value and distance.

1 Flow of Proposed Method

This method first separates a color scenery image into component L^*, a^* and b^* of the $L^*a^*b^*$ colors system, which are believed to be close to the color characteristics that we can sense, and executes subsequent processing.

Next, a potential field is calculated using the extended physiological potential, and a character peripheral area is extracted.

Then a character string area is extracted from this area. Noise is removed and then the area is output as a character string area image for each character string area. The processing stages are carried out for each component L^*, a^* and b^* separately.

2 Separation using L*a*b* Color System

Conventionally strength of the psychological potential field is proportionate to image clearness H (brightness difference between figures and the background). The conventional method omits H because the targets are binary images in which the brightness difference is constant.

However, color scenery images discussed in this paper need to determine H because the brightness differences between figures and the background are not constant. This paper indicates figure clearness H using the color difference (differences between colors) between figures and the background.

The International Commission on Illumination (CIE) recommends the uniform color space and color difference formula by CIE LAB ($L^*a^*b^*$) as methods for quantitatively indicating the color difference between two colors. This paper uses the $L^*a^*b^*$ color system as color space and defines figure clearness H as shown in formula 1 (Fig.1).

$$\Delta E^* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{\frac{1}{2}}$$

$$H = \Delta E^* \qquad (1)$$

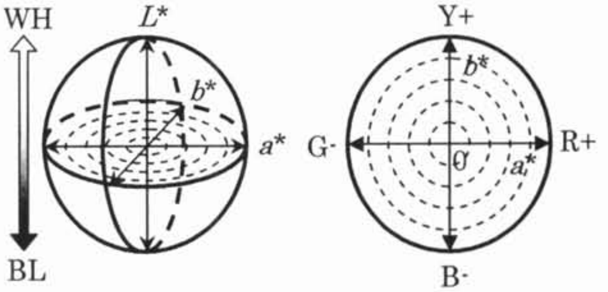


Figure 1: $L^*a^*b^*$ Color System

3 Extended psychological potential field

3.1 Shielding effect of monochrome (binary) image

This paper calculates the strength of psychological potential field of binary images as shown below.

Figure 2 shows characters and figures consisting of black pixels. Scanning is performed from arbitrary peripheral point P in the direction of angle θ_i until the scan first touches the frame of a character or figure. The distance between the pixels from P to the first touch is r_i. If no frame is touched, $r_i = \infty$ so that no potential generates.

This operation is repeated around point P while changing θ_i. When the number of scans for one turn around point P is N, $\theta_i = (2\pi / N) \times i[rad]$ $(i = 0,1,2,\ldots,N-1)$, and potential field strength M_p at point P can be determined by formula 2 as a mean value of reciprocal numbers of r_i.

1 Address: 771 Ooaza Nakakuki, Oyama-shi, Tochigi 323-0806 Japan. E-mail: masanori@anegawa.com

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

$$M_p = \frac{1}{N}\sum_{i=0}^{N-1}\frac{1}{r_i} \qquad (2)$$

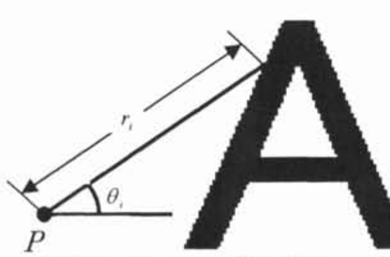


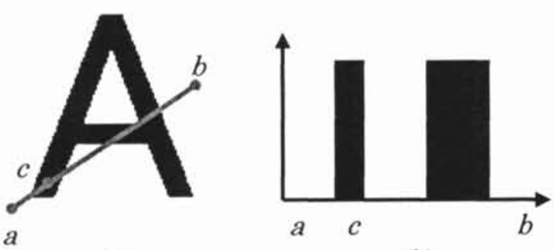
Figure 2: Calculate potential of binary image

3.2. Differences between monochrome (binary) images and color scenery images

Figure 3(a) shows a monochrome (binary) image. Figure 3(b) shows the section view of the segment between *a* and *b* in Figure 3(a). As shown in Figure 3(b), the monochrome (binary) image assumes character pixels as a high wall and stops scanning between *a* and *b* at the point where the scan first touches the wall. Stopping scanning yields non scanned area, named the shielding area, and this effect is called the shielding effect.

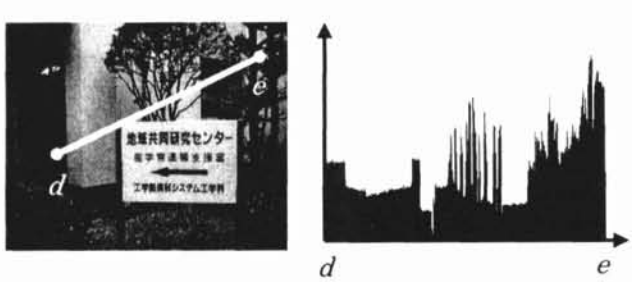
Figure 4(a) shows a color scenery image, and Figure 4(b) shows the section view of the segment between *d* and *e*. As shown, the pixels of the color scenery image have values in varieties of level, which make it difficult to determine what level of values as a wall.

Furthermore, in the scenery image shown in Figure 4(b), the inside of the signboard where characters are written is a separate area from the outside of the signboard even in one image. The potential values inside the signboard must not depend on the information outside the signboard. For preventing this kind of dependence, a wall of shielding effect is needed to be placed on the boarder line between the inside and outside of the signboard. Thus the shielding effect for color scenery images is more important than that for monochrome (binary) images.



(a) (b)

Figure 3: Shielding effect of monochrome (binary) image



(a) (b)

Figure 4: Shielding effect of color image

3.3 Shielding effect for color scenery images

The method proposed in this paper implements shielding effect for color scenery images as follows. We introduce "clearness" which is defined as color difference between the scanning point and its background. When the accumulated total value of clearness for all pixels on a scanning line (one scan direction) reaches a certain level, the applicable pixel is recognized as a wall and scanning in that direction is interrupted.

The shielding effect for color scenery images is explained using Figure 5(a). For this explanation only, color values are represented using the gray scale (0 to 255) and clearness H of figures is represented using the absolute values of the differences between figure background values. This explanation also uses 255 as the threshold of the accumulated total value of clearness at which scanning is to be interrupted. Figure 5(b) shows the section view of scanning in the direction of θ_α from point P in Figure 5(a). The intersections between the scanning line and lines l_1, l_2, and l_3 are respectively $P_{\alpha 60}$, $P_{\alpha 100}$, and $P_{\alpha 150}$.

P_{ij} indicates the j th pixel when scanning is performed in the direction of θ_i from a starting point P, which is selected arbitrarily from pixels not yet scanned. The distance between points P and P_{ij} is r_{ij} and clearness is H_{ij}. In Figure 5(a), the background value is 100 and the values of lines l_1, l_2, and l_3 are respectively 0, 255, and 128. The clearness values of $P_{\alpha 60}$, $P_{\alpha 100}$, and $P_{\alpha 150}$ are thus represented by formula 3.

The value of the background area other than lines l_1, l_2, and l_3 is same as the point P and therefore clearness is 0. This leads to $H_{\alpha 60} + H_{\alpha 100} = 255$. Because threshold 255 is reached, scanning in the direction of θ_α is interrupted at $P_{\alpha 100}$. Consequently, $P_{\alpha 150}$ that is in the shielded area, is not included in calculation.

The formula of extended psychological potential using this shielding effect is given by Formula 4. Where n_i is number of pixels scanned in the direction of θ_i when the accumulated total value of clearness (compared with the point P) reaches the threshold, and N' is the total number of pixels on all scanning lines.

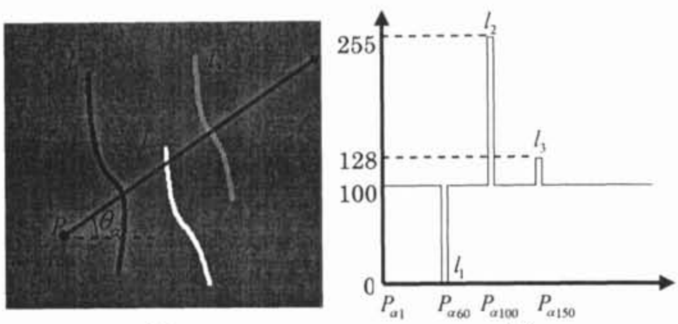
In addition, function $f(x)$ (sigmoid function) that clarifies high or low potential levels is applied to the formula 4 to form a final formula 5 for extended psychological potential.

$$\begin{aligned} H_{\alpha 60} &= |P - P_{\alpha 60}| = |100 - 0| = 100 \\ H_{\alpha 100} &= |P - P_{\alpha 100}| = |100 - 255| = 155 \\ H_{\alpha 150} &= |P - P_{\alpha 150}| = |100 - 128| = 28 \end{aligned} \qquad (3)$$

$$M_p = \mathrm{f}\left(\frac{1}{\mathrm{N}'}\sum_{i=0}^{N-1}\frac{1}{n_i}\sum_{j=1}^{n}\frac{H_{ij}}{r_{ij}}\right) \qquad (4)$$

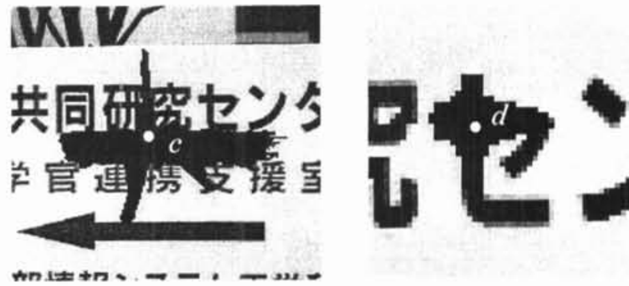
$$f(M_p) = \frac{1}{1+\exp(-aM_p)} \qquad (5)$$

$$a = constant$$



(a) (b)

Figure 5: Explain of shielding effect



(c) (d)

Figure 6: Example of shielding effect

3.4 Example of shielding effect for color scenery images

In Figure 6, black areas are the scanned and calculated areas that the starting points a,b,c,d are chosen as examples out of all pixels in the original image. The shielding effect is effective to cut down the amount of calculation, because the shielding areas are need not calculation. Moreover, Figure 6 tells that the margin of the scanned (black) areas matches well with the boundary of the character areas. Consequently, we concluded that this method, utilizing potential field and the shielding effect, matches well with the human visual characteristics and can be applicable to color scenery image. This method, named the extended psychological potential field methd, is an extension of previously proposed method using potential field.[5]

Figure 7(a) shows the intermediate image of applying extended psychological potentials to the original image. Figure 7(b) shows the final image of that processing.

The characteristics of potentials have been emphasized by *f(x)* (sigmoid function) in the final image of Figure 7(b), compared with the intermediate image of Figure 7(a).

Original image

(a) (b)

(a) M_p image (b) $f(M_p)$ image

Figure 7: Example of extended psychological potential field

4 Extraction of character string areas

4.1 Extraction of character peripheral areas

Figure 7 thus tells that the extended psychological potential field well matches psychological concepts. The potential field shows high values for:

frame lines of walls and signboards,

parts that show high complexity such as wood and soil, and

character strings and figures on signboards.

The figure also shows that the character string areas with high potential in the signboard are surrounded by the character string peripheral area with low potential.

In this background, the areas with low potential (the range $0 < P < 50$ is specified in this experiment) are character string peripheral areas (Fig.8).

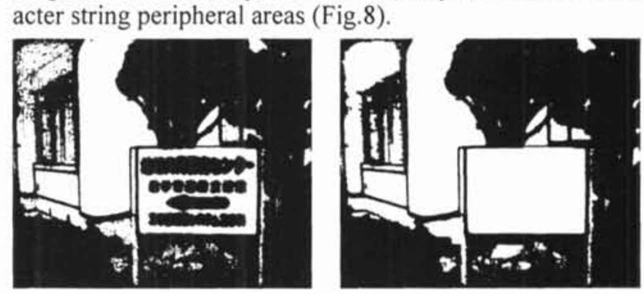


Figure 8: Extraction of character peripheral areas

4.2 Extraction of character string areas

An area with high potential (the range $100 < P < 255$ is specified in this experiment) in the character string peripheral area is extracted as a character string area.

This processing enables more accurate extraction of only character string areas (See Fig.9). However, as seen in Figure 9, the image of the extracted character string areas includes very small noise components.

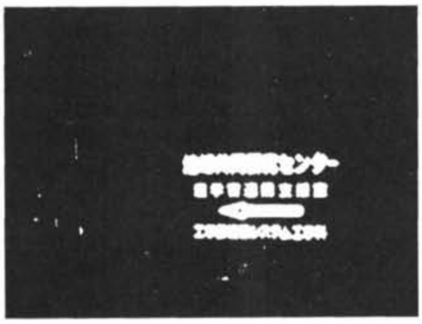


Figure 9: Extraction of Character string areas

4.3 Noises removal

The extracted character string area extracted may contain pixels with a size that is too small for a character. These pixels are removed as noises.

Figure 10 shows the image of a final character string areas after noises removal. This shows that the character string areas has been extracted successfully.

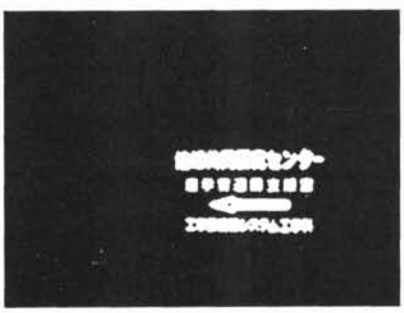


Figure 10: Example of noises removal

4.4 Output of character string extract image

The original image is placed over the character string area extracted in Figure 10 and divided into each area, and the character string extract image is output (Fig.11).

Although the character string areas have been favorably output in Figure 11, some of the output areas include no character strings (indicated by arrows in Figure 11). It will be possible to eliminate these areas by recognition processing, etc.

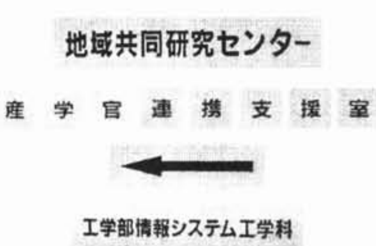


Figure 11: Output of character string extract image

5 Results of experiment

Figure 12 show the results of applying this method to actual images. This experiment used $0 < P < 50$ for the potential of the character string peripheral area and $100 < P < 255$ for the potential of the character string area.

(a) Original image (b) Potential image

(c) Extract image (d) Output image

Figure 12: Example of experiment (1)

6 Conclusions

Figures 6 to 12 show high potential values at character string areas and frame lines of substances. The leaves are relatively dark in the original image but they show relatively high potential values. This result shows that we did well to use the color difference (difference between colors) as a means to represent clearness of figures.

Good results were obtained also for character string extraction.

Examination of character string extraction failures showed that many of the failures occurred where multiple character string areas were fused. It will probably be required in the future to set potential threshold automatically and properly for character string peripheral areas and character areas.

References

[1] Zensho YOKOSE et al.: "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 et al.: "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 et al.: "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 et al.: "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

[5] Masanori ANEGAWA et al.: "Extraction of Character String Areas from Color Scenery Image using Psychological Potential Field", Proceeding of MVA2000 IAPR Workshop on Machine Vision Applications, pp.197-200, 2000