

Method for Character Domain Extraction from Image Using Wavelet Transform

Taiki Taniguchi

Zensho Holdings Co., Ltd., 2-18-1 Konan Minato-ku Tokyo 108-0075, JAPAN

Yasunari Yoshitomi

Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,

1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan

E-mail: yoshitomi@kpu.ac.jp

http://www2.kpu.ac.jp/ningen/infsys/English_index.html

Abstract

There has been an increase in the number of images on the Internet that contain private information and/or the URL of an illegal Web site. This information could result in a violation of human rights and/or a crime. We present a method for extracting from an image the region(s) that contain characters. Our method uses the discrete wavelet transform and the empirical knowledge that characters have strong vertical and/or horizontal elements. The experimental results show the usefulness of our method.

Keywords: Character domain extraction, Image processing, Internet, Private information, URL of illegal Web site, and Wavelet transform.

1. Introduction

There has been an increase in the number of images on the Internet that contain private information and/or the URL of an illegal Web site. This information could result in a violation of human rights and/or a crime. It is very time-consuming and inefficient to detect them by inspecting each individual Web page, but there is no existing automatic method for checking images for such information. To achieve this, the first step is to extract from the image the region(s) that contain characters. An algorithm for extracting such regions has received considerable attention in the field of computer vision research,¹⁻³ however, it falls far short of practical application.

In this paper, we present a method for extracting from an image the region(s) that contain characters. Our

method uses the discrete wavelet transform (DWT), and it is based on the empirical knowledge that characters have strong vertical and/or horizontal elements.

2. Wavelet Transform of Image Signals

The hierarchical decomposition of the DWT has been widely used in image compression and other image processing techniques. The image is first decomposed into four subbands, 1LL, 1LH, 1HL, and 1HH. The subbands labeled 1LH, 1HL, and 1HH represent the finest-scale wavelet coefficients. Fig. 1 shows the level 1 decomposition, in which the image is decomposed into four subbands that share the same scale.

Taiki Taniguchi, Yasunari Yoshitomi

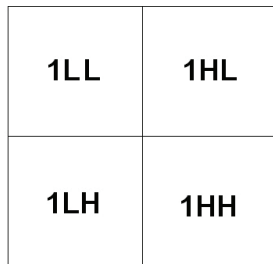


Fig. 1. Mallat division.

In general, the wavelet coefficients for these three domains (HH, HL, and LH) are called the elements of the multiresolution representation, whereas the wavelet coefficients of the other domain (LL) are called the elements of the multiresolution analysis. For further information on the DWT, see Ref. 4.

3. Proposed Method

Fig. 2 shows the flowchart of the proposed method. The algorithm of the proposed method is as follows:

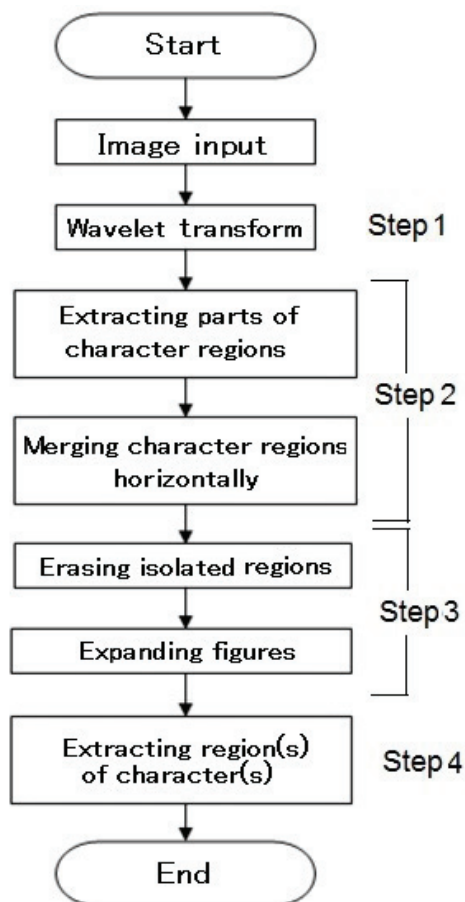


Fig. 2. Flow chart of the proposed method.

Step 1: The RGB elements of the image (Fig. 3) are transformed into YCrCb elements (Fig. 4). This is followed by applying DWT to each element to obtain the 1HL and 1LH elements, which correspond to the vertical and horizontal elements, respectively (Fig. 4).

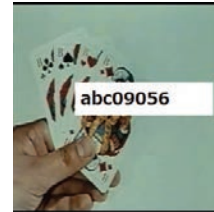


Fig. 3. Example of an input image.



Fig. 4. 1HL (top row center) and 1LH (top row right) elements for the Y element (top row left); 1HL (middle row center) and 1LH (middle row right) elements for the Cr element (middle row left); and 1HL (lower row center) and 1LH (lower row right) elements for the Cb element (lower row left); all obtained from the input image (Fig. 3).

Step 2: A segmentation is performed for both the 1HL and 1LH elements, using each threshold (Fig. 5). Then, the scattered regions that have values of 255 in the binary image for each of YCrCb elements are horizontally merged by changing to the value 255 the value 0 of pixels having the same horizontal coordinate with those with the value 255 (Fig. 6).

Step 3: Isolated regions are erased and regions with the value 255 are expanded (Fig. 7).

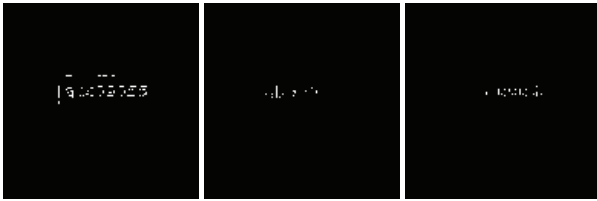


Fig. 5. Binary images obtained for both 1HL and 1LH elements for the Y, Cr, and Cb elements of the related images in Fig.4; left: Y; center: Cr; right: Cb.

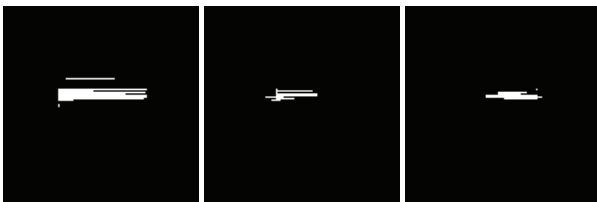


Fig. 6. Binary images after horizontally merging the Y, Cr, and Cb elements of the related images in Fig.5; left: Y; center: Cr; right: Cb.

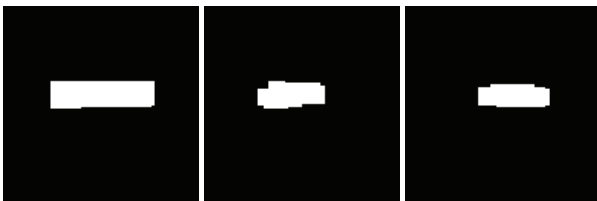


Fig. 7. Binary images after completing Step 3 for the Y, Cr, and Cb elements of the related images in Fig.6; left: Y; center: Cr; right: Cb.

Step 4: A mask is generated for extracting from the initial RGB image those areas that contain characters. This is done by using a logical summation to merge the regions with the value 255 in the binary images obtained for the Y, Cr, and Cb elements. Finally, the output RGB image is generated by changing to white the color of any pixel for which the value of the corresponding pixel on the mask image is 0; this leaves only the region(s) containing character(s) on the RGB image (Fig. 8).



Fig. 8. Input image (left) and output image (right).

4. Experiments

4.1. Conditions

Three standard images (Earth, Couple, and Mandrill; 24-bit BMP form; 256×256 pixels) were prepared for evaluation of the proposed method. The images used in this experiment were selected from the Standard Image Database (SIDBA). We used the Daubechies wavelets for the DWT. Several black characters (“abc09056”) were written on the original images; the backgrounds of the characters included one each of white, nonwhite color, and transparent (Fig. 9). The experiment was performed in the following environment: personal computer: DELL OPTIPLEX780 (CPU: Intel(R) Core(TM) 2 Duo CPU E8400 3.00GHz; main memory: 4.00GB); OS: Microsoft Windows 7 Professional; development language: Microsoft Visual C++ 6.0.

In the process of segmentation in Step 2, which was described in Section 3, the pixels in the binary images were set to 255 when the values of the 1HL or 1LH elements were below the threshold for that element; otherwise, the value was set to 0. The threshold was set to be half the minimum value for each element, but not higher than -40. In the process of segmentation in Step 3, as described in Section 3, regions with less than or equal to 50 pixels were erased; this expansion of the larger regions was performed five times.



Fig. 9. Examples of an original image (far left) and the input images with characters added (others).

4.2. Results and discussion

Figs. 10 and 11 show the input images and the output images obtained using the proposed method. When the background of the characters was not transparent, all regions with characters were successfully extracted by the proposed method (Fig. 10). However, when the background of the characters was transparent, some regions with characters were not extracted, and some regions without characters were extracted (Fig. 11).

Taiki Taniguchi, Yasunari Yoshitomi

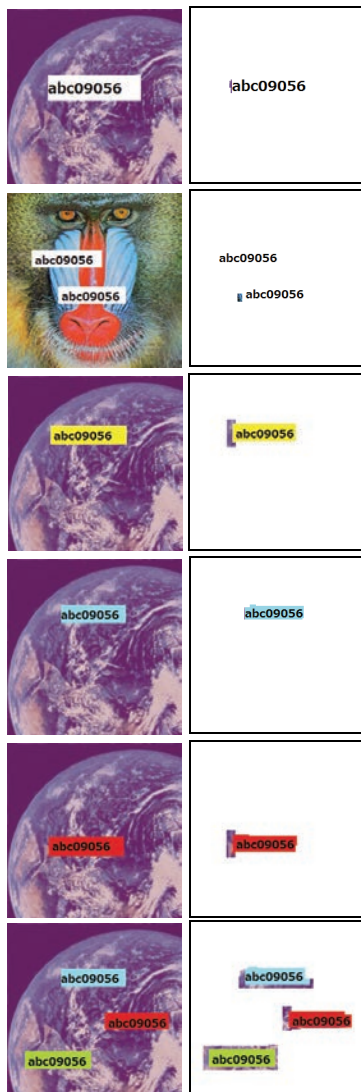


Fig. 10. Input images (left) with characters on color backgrounds, and output images (right).

5. Conclusion

A method for extracting from an image those regions that contain characters was developed; this method uses the discrete wavelet transform and the empirical knowledge that characters have strong vertical and/or horizontal elements. Our experimental results show the usefulness of the proposed method. We intend to improve the proposed method so that it can be successfully applied when the background of the characters is transparent.

Acknowledgements

The authors would like to express their sincere thanks to Associate Professors M. Tabuse and T. Asada of Kyoto Prefectural University for their useful advice and support of this research.

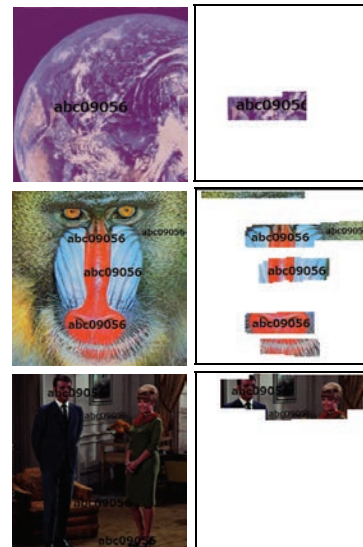


Fig. 11. Input images (left) with characters on transparent backgrounds, and output images (right).

References

1. H. Hirano and M. Okabe, Detection of text area inserted by image editing (in Japanese), in *IPSJ SIG Notes*, vol. 2012-CVIM-183, no. 9, (Japan, Koganei, 2012), pp. 1-8.
2. K. Ashida, H. Nagai, M. Okamoto, H. Miyao, and H. Yamamoto, Extraction of characters from scene images (in Japanese), *Trans IEICE J88-D-II(9)* (2005) 1817-1824.
3. S. Masuko, T. Hoshi, and H. Hirano, A system for searching texts in images by using reliability factor and saliency (in Japanese), *J. IPSJ* **51(9)** (2010) 1520-1529.
4. D. Inoue and Y. Yoshitomi, Watermarking using wavelet transform and genetic algorithm for realizing high tolerance to image compression, *J. IEEJ* **38(2)** (2009) 136-144.