Optik 126 (2015) 626–629

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Image segmentation based on gray stretch and threshold algorithm

Leiming Liu [a], Ning Yang [a], Jinhui Lan [a,*], Juanjuan Li [b]

[a] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, PR China
[b] Beijing Wanji Techonology Co., Ltd, Beijing 100085, PR China

ARTICLE INFO

Article history:
Received 24 January 2014
Accepted 27 January 2015

Keywords:
Image segmentation
Wavelet transform
Gray stretch

ABSTRACT

When the gray levels between the foreground and background do not change significantly in the image, the Otsu's method cannot determine the segmentation threshold accurately. In this paper, an image binarization method based on wavelet domain gray stretch is proposed. In order to emphasize the gray stretch and edge enhancements on the image, it combines the discrete wavelet transform with the principle of the largest between-class variance. Experimental comparisons between the method and other methods show that the proposed approach has a satisfactory image segmentation and strong robustness to noise.

1. Introduction

Image segmentation is a key step from image processing [1] to image analysis [2], since a qualified segmentation is the prerequisite and base of the consequent processing such as object extraction, parameter measurement and object recognition [3]. Image segmentation [4–7] has always been an unsolved and challenging topic in image engineering for many years.

Separating the foreground out from the background of an image is an important preprocessing in image analysis. Its purpose is to acquire some useful information in the image for higher level image processing. Thresholding [8–13] or binarization [14–16] is such a widely used method, and generally, its process is to first determine a gray threshold according to some objective criteria and then assign each pixel to one class (such as the foreground) if its gray level or gray value is greater than the determined threshold and otherwise to the other class (such as the background).

Threshold based image segmentation techniques discriminate regions on the basis of intensity value difference between pixels. Thresholds for image segmentation have been calculated based on interclass variation [17] and histogram [18,19]. The limitation of threshold based segmentation technique is that it performs well for images, which have only two components.

In order to determine the segmentation threshold accurately, this paper introduces an automatic segmentation method based on discrete wavelet transform (DWT). Using 2D wavelet transform, an image could be decomposed into four sub-images via high-pass and low-pass filter. Low frequency components represent the basic figure of an image. They are extracted as the frame matrix. Low frequency components are less sensitive to varying images and are the most informative sub-images gearing with highest discriminating power.

This paper is organized as follows: In Section 2, we review briefly discrete wavelet transform, and then propose a threshold segmentation method, which takes not only the wavelet transform but also the principle of maximum between-class variance separation. The experimental results are shown in Section 3. Finally, Section 4 concludes this paper.

2. The proposed approach

2.1. Basic discrete wavelet transform

Wavelet transform provides a special basis that a signal can express features easily and efficiently. The wavelet transform connotes hierarchical features. Two-dimensional DWT [20] can be used to decompose an image into four sub-images (as shown in Fig. 1). The two-dimensional wavelet transform is performed by consecutively applying one-dimensional wavelet transform to the rows and columns of the two-dimensional data. As can be seen in Fig. 2, filters are first applied in one dimension (e.g. horizontally) and then in the other (e.g. vertically). These sub-images contain coarse approximations of the image as well as horizontal, vertical and diagonal details of the image at various scales. For an original image, it can be decomposed into four different bands (LL, HL, LH and HH) by using the discrete wavelet transform. These sub-bands contain different frequency characteristics by high-pass and low-pass filter. The high-pass filter extracts the high frequency part and the low-pass filter gives the low frequency information representing the most energy of an image.

* Corresponding author. Tel.: +86 10 62334961; fax: +86 10 62334491.
E-mail address: ustblanjinhui@gmail.com (J. Lan).

http://dx.doi.org/10.1016/j.ijleo.2015.01.033
0030-4026/

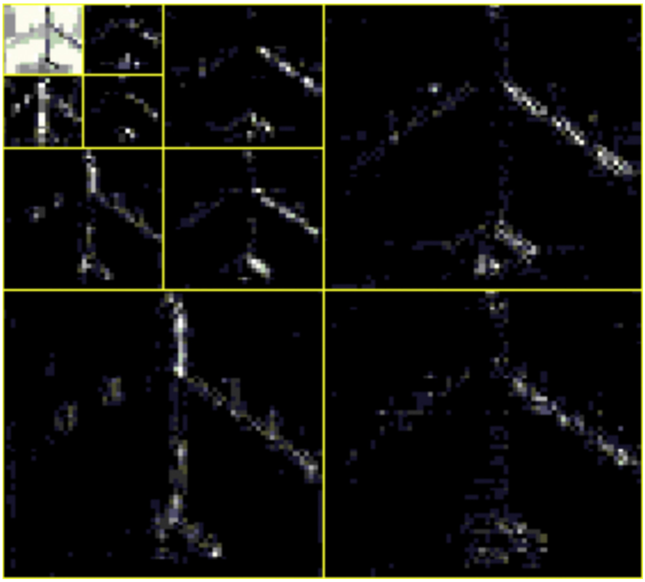

Fig. 1. Three-level DWT decomposition of the image.

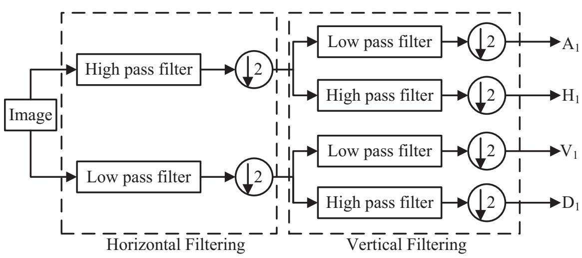

Fig. 2. The structure of two-dimensional DWT.

2-D DWT is derived from one-dimensional case. The scaling function and the wavelets are

$$\begin{cases} \Phi(x,y) = \phi(x)\phi(y) \\ \psi^{H}(x,y) = \psi(x)\gamma(y) \\ \psi^{V}(x,y) = \gamma(x)\psi(y) \\ \psi^{D}(x,y) = \psi(x)\psi(y) \end{cases}$$

where $\phi(\bullet)$ and $\psi(\bullet)$ denote the scaling function and wavelet in one-dimensional case respectively, $\hat{\gamma}(\omega) = \frac{1}{\sqrt{2}}\hat{q}\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right)$

Then for a finite signal named as $f(x,y) \in L^2(R^2)$, the value of (m,n) is calculated by the decomposition of wavelet as follows:

$$\begin{cases} A_{2^j}(m,n) = \sum_{k,p} \bar{h}_k \bar{h}_p A_{2^{j-1}}(m - 2^{j-1}k, n - 2^{j-1}p) \\ D^1_{2^j}(m,n) = \sum_{k,p} \bar{g}_k \bar{q}_p A_{2^{j-1}}(m - 2^{j-1}k, n - 2^{j-1}p) \\ D^2_{2^j}(m,n) = \sum_{k,p} \bar{q}_k \bar{g}_p A_{2^{j-1}}(m - 2^{j-1}k, n - 2^{j-1}p) \end{cases}$$

where g and h represent high-pass and low-pass filter, respectively.

So we have four sub-images: A, the scaling component containing global low-pass information, and three wavelet components, H, V and D, corresponding respectively to the horizontal, vertical and diagonal details. The proposed method decreases noise and the background interference by using one-level DWT.

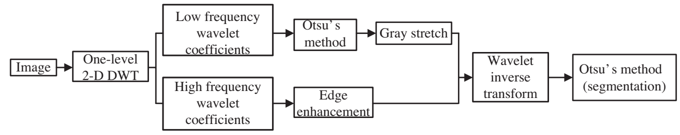

Fig. 3. The procedure of the proposed method.

2.2. Gray stretch based on different frequency components

Otsu's method [21] only takes the gray levels of the pixels into account but ignores their spatial distribution and contextual relationship of the pixels themselves belonging to the foreground or background. It only can maximize the between-class separation in the foreground and background accurately. In order to remedy such shortcomings of Otsu's method, in this paper we will define a new criterion.

The procedure of the proposed method is illustrated in Fig. 3, which consists of the following steps:

Step 1: One-level 2-D DWT is used to decompose the original image of every channel into some sub-image sets at different frequency range, which are groups of approximation, horizontal, vertical and diagonal details.

Step 2: For the low frequency part, this paper uses Otsu's method to get the threshold t as a boundary. Then gray stretch is applied to the low-frequency component, as follows:

$$A^*_{2^j}(m,n) = \begin{cases} A_{2^j}(m,n) - pA_{2^j}(m,n), A_{2^j}(m,n) < t \\ A_{2^j}(m,n) + pA_{2^j}(m,n), A_{2^j}(m,n) \geq t \end{cases} \quad (1)$$

where p represents the scale factor.

Step 3: The high-frequency channel of the transformed image is actually the first derivative of the original image. The step edge in the airspace of the original image is manifested as the roof edge of the high-frequency channel in wavelet domain. It can be regarded as a truncated Gaussian function approximately. The edge can be enhanced by reducing the variance of the Gaussian function. While the roof edge in the airspace is manifested as a pair of odd symmetry roof edges of in the high-frequency channel in wavelet domain. It can be approximately as a single-cycle sine function. And then we can reduce the period of the single-cycle to achieve edge enhancement.

Assume the period of the edge in the airspace is T, $T/3$ is the period of the single-cycle sine function. The function is used to fit the roof edge of a high-frequency channel in wavelet domain as follows:

$$y = \exp\left[-\frac{(x-u)^2}{2\sigma^2}\right] \quad (2)$$

where u and $\sigma/2$ are the mean value and the variance of the Gaussian fitting function respectively, so adjusted edge:

$$D^* = D * y \quad (3)$$

We find the step edge and the roof edge in HL and LH wavelet domain by detecting the local extreme point and the point whose derivative is zero, respectively. And then the Gaussian function and the single-cycle sine function are used to fit the edges above.

Step 4: Discrete wavelet inverse transform is applied to the modified wavelet coefficients. The wavelet reconstruction formula as follows:

$$a^j = \frac{1}{4}\left[a^{j+1^*} * (h^j, h^j) + d^{j+1^*} * (\tilde{g}^j, l^j) + d^{j+1,2^*} * (l^j, \tilde{g}^j)\right] \quad (4)$$