Threshold Selection Using Second Derivatives of the Gray Scale Image†

Theo Pavlidis
Image Analysis Laboratory
SUNY
Stony Brook, NY 11794-4400

Abstract

It is known that when a bilevel image is blurred, the intensity of the original pixels is related with the sign of the curvature of the pixels of the blurred image. A technique for threshold selection is presented where a partial histogram is constructed solely from the pixels where curvature achieves extrema values.

Introduction

In 1986 Pavlidis and Wolberg [1] proposed an algorithm for binarizing images based on the observation that when a bilevel image is convolved with a bell-shaped kernel, then most areas that were dark (low values) have positive curvature and areas that were light (high values) have negative curvature. Since the blurring caused by a scanner usually corresponds to convolution with a Gaussian point spread function, detection of the sign of curvature can be used for deblurring (in effect binarizing) an image. However there are certain practical obstacles to the straightforward application of the method caused by the presence of noise and areas of constant color that are small compared to the support of the point spread function. In order to overcome these difficulties an iterative method was proposed in [1]. It produced satisfactory results but it was slow because of the need for iterations.

More recently, Wang and Pavlidis [2, 3] proposed a related method for the direct feature extraction from gray scale data. The key observation was that the initial step of the method of [1] determined many of the features and most of the iterations simply filled an area that was later removed during a thinning process. Hence the focus was on a more careful estimation of the curvature of the gray scale surface. Since the method includes feature extraction it requires its own subsequent classifier and cannot be interfaced easily with existing character recognition systems.

This paper presents a method that determines a threshold on the basis of the values of pixels which have significant curvature. While the results may not be as good as those obtained by [1] or [2, 3], it is much faster and it has the advantage of being easily interfaced with existing recognition systems. The proposed method is compared with the histogram analysis method of Otsu as described in [4]. Otsu’s method selects as threshold the value that minimizes the between group variance. Since the proposed method also computes a histogram (but only over certain pixels), it is also possible to combine the two methods under certain conditions.

The main weakness of methods relying on histogram analysis for selecting a threshold is that they usually assume that there only two levels of intensity in the image and try to split the histogram into two parts. Unfortunately, this assumptions is often not true in practice. Since the proposed method uses the curvature information it is expected to be better than histogram analysis methods in situations where there are more than one levels (for example, textured background) or where certain bold typefaces are used. Because of the point spread function of the scanner the character images are blurred and some of the dark gray area correspond to narrow gaps that a histogram analysis might consider as black.

Figure 1: Relationship between intensity values and curvature of the smoothed signal.

Description of the Algorithm

Figure 1 shows an ideal profile of a bilevel image at the top and its blurred version at the bottom. Points $A_i$ are places where the curvature achieves a maximum.
value while points $B_i$ are places where the curvature achieves a minimum value. The key idea is that by identifying pixels where the curvature achieves an extremum, we can use the intensity values there to compute a threshold. The method of [2, 3] requires the computation of the principal directions of curvature at each point of the surface. In order to speed up the computation, we calculate approximations to the second derivatives with respect to $x$ and $y$ only. For most points, if there is a direction where the principal curvature is large, then it is likely that either of these two derivatives will also be large. In order to reduce the effects of noise the derivatives are approximated by finite difference taken at $b$ pixels apart. (The selection of this and other parameters will be discussed in the section on experimental results.) Therefore the first step is to compute two images from a given one $f(x,y)$ according to the equations below

$$f_x(x,y) = f(x+b,y) + f(x-b,y) - 2f(x,y) \quad (1a)$$

$$f_y(x,y) = f(x,y+b) + f(x,y-b) - 2f(x,y) \quad (1b)$$

In order to save memory the results may be stored in single bytes by: (a) replacing all negative values by zero, since such values are of no further interest; (b) dividing $f_x$ and $f_y$ by 2.

The second step involves the search for extrema by checking whether the inequalities

$$f_x(x,y) > f_x(x+b,y) + c \quad (2a)$$

and

$$f_x(x,y) > f_x(x-b,y) + c \quad (2b)$$

are both true, if so, then the value of $f(x,y)$ is added to a histogram. Similar checks are made for $f_y$ and for both functions at the orthogonal directions. A value of $c$ greater than zero is needed to avoid the effects of noise.

The resulting histogram contains only the values of pixels where either $f_x$ or $f_y$ is positive and significantly greater than in the neighboring pixels. To reduce the effects of quantization noise, we divide each pixel value by a constant $h$ so that a histogram with fewer levels is formed. (Quite often scanned images have histograms with gaps due to improper function of the A/D converter. If the results affect a high order bit, they are clearly visible and the device is fixed but this may not be the case when low order bits are affected.) Ideally, that histogram will contain only values occurring at dark pixels and should have a very narrow support. The threshold could be set around its upper nonzero value.

However complications arise because of images where the original was not bilevel or where originally white pixels need be interpreted as black. The former are images with textured background and they may have histograms with more than one peak. The latter are dot matrix images and they may have histograms with very broad support. Because the areas between the dots are usually saddle points (see [2, 3]), they contribute to the curvature maxima and their values are also included in the histogram. Therefore the third (and last) step of the algorithm requires the examination of the histogram according to the following rules.

---

**Estimation of the threshold from the histogram.**

1. Compute the average value of the histogram and identify the zones where the histogram exceeds the average value. Set the threshold depending on the number of such zones as follows.

2. If there is only one zone, set the threshold at the first value after the zone where the histogram drops below $1/g$ of its average value, where $g$ is a preselected constant.

3. If there is more than one zone, then Otsu's method is applied to the histogram.

---

Note that it would be wrong to apply Otsu's method in the case of one zone because in that case the populations is assumed to be homogeneous. On the other hand if the histogram has a flat, nearly rectangular shape, Otsu's method might be applicable. There was only one case of such a histogram in the tests and splitting the histogram gave better results than by selecting its upper bound.

**Experimental results**

The method was tested on samples of the U.S. Postal Service data base and some non-text images that contain a single object against a supposedly flat background. (A total of 76 images.) It was also compared with results obtained by the histogram analysis method of Otsu as described in [4]. The latter method selects as threshold the value that minimizes the between class variance in the histogram and it does not require any parameter selection. The proposed method requires the setting of various parameters which help produce better results, if properly selected. The following are guidelines and the values used in the experiments.

- $b$, the step between pixels for the approximation of the second derivatives. It was set equal to 3.
- $c$, the constant that defines maxima as significant. It was selected adaptively according to the expression

$$c = 10 + r/4$$

where $r$ is the range of observed values occurring in at least 10 pixels. $c$ was clipped at 60.
the histogram bin size expressed in bits was set to 3. This imposes a granularity of 8 units in the selection of the threshold.

The fraction of average value beyond which data are ignored was set to 3. The results do not appear to depend much on this value, since in most cases the histogram rise sharply over a narrow support area.

Some examples are shown in Figures 2 to 6. In each one the leftmost figure is a halftone of the gray scale original (300dpi at 8 bits), the middle is the thresholded image obtained by the method of this paper and the rightmost is the thresholded image obtained by Otsu's method.

For most images both methods give similar threshold values, as expected for the case where there are distinct differences in the values of pixels that are supposed to be dark or light. However significant differences were observed in the difficult cases such as those shown here. These include bold faced (Figure 2), images with textured background (Figures 3), etc.

Figure 4 to 6 show three cases where the high curvature histogram yields respectively about the same, higher, and lower threshold than the global histogram. These images were not used during the development of the method and the same parameters as for the postal addresses were used. Therefore they represent an objective test of the method. The new method (middle) has done much better in one (Figure 6), slightly worse in another (Figure 5), and about the same in a third (Figure 4).

Conclusions

The proposed method has higher computational requirements than a purely histogram analysis method because of the need to compute and store $f_{xx}()$ and $f_{yy}()$. and the subsequent masking operation. For many images the results are similar with that a histogram analysis method [4] to the point that the selected thresholds are very close if not identical. For other images both methods are inferior to the direct extraction from gray scale. This is because the method of [2, 3] uses saddle point information. The method described here cannot find saddle points because it computes the second derivatives along the co-ordinate axes rather than the principal directions. When both principal curvatures are of the same sign this is not a serious impediment. But when they are of opposite signs they may be very small (because of cancellation) in the directions of the co-ordinate axes and therefore such points are ignored. Thus there is only a narrow domain of pictures where this method outperforms methods based on histogram analysis.

If low contrast images with textured backgrounds (but not sparse dot matrices) are a large fraction of the input population, then it may be worthwhile to use the method described in this paper.

References


Figure 2: Original (left). New Method (center). Otsu's Method (right).

Figure 3: Original (left). New Method (center). Otsu's Method (right).
Figure 4: Original (left). New Method (center). Otsu’s Method (right).

Figure 5: Original (left). New Method (center). Otsu’s Method (right).

Figure 6: Original (left). New Method (center). Otsu’s Method (right).