On Multi-Scale Adaptive Thresholding

Xue Dong Yang and Vipin Gupta
Department of Computer Science
University of Regina
Regina, Sask., Canada S4S 0A2
e-mail: yang @cs.uregina.ca

Abstract

In adaptive thresholding, a threshold is selected according to local characteristics in an image and varies from one part of the image to another. There exist a number of adaptive local thresholding techniques, with many successful applications. However, they also have several problems. First, there is no systematic consideration for the size of the window. Second, the technique, that subdivides an image into a set of subimages and interpolates the thresholds from those found for subimages, tends to create artifacts at subimage boundaries, and can miss boundaries sometimes. Third, they are not applicable to general multi-modal images. This paper presents a new adaptive local thresholding technique which addresses these problems. A variable sized window is scanned over an entire image. An efficient incremental method is developed for discriminating the modalities of a local region by using the mean and variance of graylevel distribution. A threshold hierarchy indexing method is introduced to efficiently compute approximated optimal local thresholds. This preliminary study also investigated the relationship between the window size of adaptive local thresholding and the thresholded image results. It is found that with different window sizes, an image may be segmented at different levels of abstraction.

1 Introduction

Image segmentation is a process of partitioning an image into a set of non-overlapping regions. The purpose of image segmentation is to decompose an image into parts that are meaningful with respect to a particular application. Although it is difficult to tell a computer program what constitutes a "meaningful" segmentation, there are, however, some general rules for segmentation [4] :

1. Regions of an image segmentation should be uniform and homogeneous with respect to some characteristics, such as graylevel or texture.
2. Region interiors should be simple.
3. Boundaries of each segment should be simple and spatially accurate in separating adjacent regions which have significantly different values with respect to the characteristics being considered.

There is no standard approach to different segmentation problems. A variety of techniques have been developed for different applications, which can be broadly classified into three categories: discontinuity-based techniques, region growing techniques, and thresholding techniques. The work reported in this paper belongs to thresholding techniques. Thresholding techniques are computationally simple and never fails to define disjoint regions with closed connected boundaries [11]. Thresholding techniques can be further classified into the following two categories: global thresholding, and adaptive local thresholding.

Image characteristics can change over a broad range of intensity distribution across an image. A threshold value may work well in one region of an image, but may perform poorly in another. This leads to the consideration of adaptive thresholding by which a threshold is selected according to local characteristics in an image and varies from one part of the image to another.

Fernando and Monro [2] suggested a local thresholding technique for X-ray angiograms. The histogram of such images are unimodal with narrow peaks, and so most global thresholding techniques produce unsatisfactory results. According to this method, the image is partitioned into 16 nonoverlapping subimages and the entropic thresholding technique of Pun [8] is applied to determine the threshold value for each of these subimages. This method may yield a thresholded image with graylevel discontinuities at the boundaries of two different subimages. A low-pass filter is then used to reduce these discontinuities.

Chow and Kaneko [1] proposed a variable thresholding method. Their idea is to test for bimodality. If the histogram for a window is bimodal, then a threshold is computed. If the histogram for a window is unimodal, then no threshold can be directly computed. In this case, the threshold will be defined as the interpolated value from the thresholds found in neighboring windows.

Nakagawa and Rosenfeld [6] extended the above method [1] to trimodal cases and found an improvement over bimodal cases.

Though the above methods have been successfully applied to many applications, such as detecting the heart region on chest X-ray [1], they have several problems. First, there is no systematic consideration for the size of a window. We believe that the size of the window should not be fixed, rather it should be related to the scale of features in an image. It is desirable that the size of window is appropriately selected so that it generates a unimodal histogram when the window is inside a homogeneous region, and a bimodal histogram when the window crosses the boundary of two regions. In the case that a window is located near a boundary of two regions, if the window is too small, it may not cover enough regions on both sides of the boundary to produce a well-defined bimodal histogram. On the other hand, if the window is much larger than the scale of features, it may cover too large a area such that a multimodal histogram is produced. Therefore, the size of the window can not be arbitrarily chosen, instead, it should be determined with a consideration about the scales of features in an image. Furthermore, the scale of features may vary across an image. In this case, the size of the window should not be fixed, but, varies across an image according to the scale variation of features.

Second, because of the subdivision of the image into independent windows, there are chances that the threshold values for each window might change abruptly, creating unnecessary artifacts at the boundaries of neighboring subimages. Sometimes smoothing techniques [2] may be applied to remove these artifacts, but the problem is not completely solved. Ideally, a window should scan across an image, and a threshold is determined at each location. But it is computationally expensive. A compromise is to divide the image into subregions and then have thresholds interpolated between regions as presented in [1, 6]. However, it can perform very poorly in situations where the object is between two adjacent windows.

Third, the existing adaptive local thresholding techniques are not directly applicable to general multi-modal images.

Motivated by these considerations, a new adaptive local thresholding technique, based on a variable sized window, is studied. The new technique addresses the first and second problems associated with existing methods described above. Furthermore, the third problem is solved by a novel threshold hierarchy indexing method. The preliminary experimental results reported in this paper indicates that the new method is very promising. The paper is organized as follows. Section 2 introduces the new adaptive local thresholding method based on a variable sized window. Section 3 presents a new threshold hierarchy indexing method for the segmentation of general multi-modal images. Two sets of experimental results are presented and discussed in Section 4. In the final conclusion section, limitations of the presented technique and future research are discussed.

2 Variable Thresholding Based on Local Bimodality.

The first problem concerned is the selection of appropriate window sizes. There is no rule from previous literature as to what the window size should be. From the discussion in the previous section, the size of a window should be large enough to cover sufficient number of pixels for generating a reliable graylevel distribution. A 4×4 or smaller window covers only 16 or less pixels. A histogram from such a small population is very sensitive to noise, and generally not sufficient to defined a reasonably reliable graylevel distribution. Hence, the minimum size of the window is set at 5×5 in this study. With this window size, the associated computational cost is relatively low, and it generates a reliable bimodal histogram when the window is located near or on a boundary of two regions. On the other hand, the maximum size of the window should not exceed the scale of the features under consideration. In addition, if a large window is used, associated computational costs, such as finding the modes, will increase. Therefore, the size of window should be bounded by a minimum and a maximum size so that the uni- or bi-modality could be reliably detected. The effects of different minimum and maximum sizes are experimented in Section 4, and their properties are illustrated by several examples.

The main computation in variable thresholding is to determine the modality of histogram. Conventional techniques for modality determination are usually based on an analysis of histogram distribution, such as identifying valleys and peaks. Since we need to determine modality at multiple scales across an entire image, conventional techniques can be very time consuming. This leads to the consideration of an estimation method introduced below. The principle is based on the mean and variance of the histogram associated with a window. The mean is calculated using the definition:

$$\overline{y}_0 = \frac{1}{N}\sum_{i=x_1}^{x_2}\sum_{i=y_1}^{y_2} x(i,j), \qquad (1)$$

where N is the size of the window; (x_1, y_1) and (x_2, y_2) are the coordinates of the lower left corner and upper right corner of the window, respectively. The variance can be calculated using the definition:

$$V_0 = \frac{1}{N}\sum_{i=0}^{N-1}(y_i - \overline{y}_0)^2, \qquad (2)$$

where N is the number of graylevels.

The calculation starts with a predetermined minimum window size. A threshold value V_T is chosen for the variance. If $V_0 \geq V_T$, the histogram is considered to be bimodal. Otherwise it is unimodal.

It should be noted that this method gives only an approximated estimation of modality, and should not be taken a precise decision. Though it is quite reliable in discriminating uni and bimodes in most situations, some ambiguity may rise in a few cases. For example, some unimodal distribution could have a relatively large V_0 value. To eliminate such an ambiguity, a verification step should be followed. It should also be noted that this method determines only the modality, but does not find a threshold when a bimodal is detected. Threshold selection is discussed in Section 3. As we will see there the verification of modality will be a by-product of the threshold selection process.

With a minimum size window, a small V_0 vales should not lead to the conclusion of unimodal, because there is a possibility that the window is located near the boundary of two regions, but is not wide enough to reach the other side of the boundary. So the window should be expanded and re-tested for modality until it reaches a predefined maximum size. Thus, if the histogram is unimodal, then the window size is increased by one pixel on all sides and tested again for modality.

After the window size is increased, mean and variance have to be calculated again. This will be time consuming if they are re-calculated entirely. In order to improve the efficiency of this calculation, an incremental method was devised to prevent calculation of mean and variance all over again.

In the case of the mean, only those pixels are considered which are new to the old window, i.e. the one-pixel increase on all sides. The new relation becomes :

$$\overline{y}_{new} = \frac{1}{N^*}\left(\overline{y}_0 \times N + \sum_{i=x_{11}}^{x_{21}}\sum_{j=y_2+1}^{y_{21}} x_{ij} + \sum_{i=x_{11}}^{x_{21}}\sum_{j=y_{11}}^{y_1-1} x_{ij} + \sum_{i=x_{11}}^{x_1-1}\sum_{y_1}^{y_2} x_{ij} + \sum_{i=x_2+1}^{x_{21}}\sum_{y_1}^{y_2} x_{ij}\right) \qquad (3)$$

where N^* is the new window size, (x_{11}, y_{11}) is the new left lower corner and (x_{21}, y_{21}) is the new upper right corner of the window.
Also, the new variance can be given as :

$$V_{new} = \frac{1}{N^*}\sum_{i=0}^{N-1}(y_i - \overline{y}_{new})^2. \qquad (4)$$

For variance, an incremental method has been designed. The change in variance can be calculated by finding the difference of the new and old values $(V_{new} - V_0)$. Also the change in mean value can be found using the relation $\overline{y}_{new} = (\overline{y}_0 + \Delta y)$. The above two equations, when solved for $(V_{new} - V_0)$, result in:

$$N^* \times V_{new} - N \times V_0 = \sum_{i=0}^{N-1}\Delta y(2\overline{y}_0 + \Delta y - 2y_i) \quad (5)$$

or

$$V_{new} = \frac{1}{N^*}[N \times V_0 + \sum_{i=0}^{N-1}\Delta y(2\overline{y}_0 + \Delta y - 2y_i)] \quad (6)$$

Our second consideration is the problem associated with the artificial subdivision of an image in previous methods. Ideally, the threshold for a pixel (x,y) should be determined based a subimage centered at (x,y). This will require a window scanning process across the entire image. It is a clearly an expensive computation, because a threshold is to be computed at each window position. To improve the efficiency of this scanning process, an incremental formula similar to (3) is derived. When the window is translated to the right by one pixel, the new mean

Figure 1. Original image of model.

Figure 2. Binary image of modality.

and variance can be also calculated incrementally, rather than be re-calculated entirely for the window. The translation of the window shifts one column out on its left, and includes a new column on its right. Based on this observation, an incremental formula for mean was developed:

$$\overline{y}_{nw} = \frac{1}{N^*}\left(\overline{y}_0 \times N + \sum_{i=x_2+1}^{x_{2new}} \sum_{j=y_2+1}^{y_{2new}} x_{ij} - \sum_{i=x_1}^{x_{1new}-1} \sum_{j=y1}^{y_{1new}-1} x_{ij}\right) \tag{7}$$

where (x_{1new}, y_{1new}) and (x_{2new}, y_{2new}) are the new coordinates of the window. That is, the column being shifted out of the window is subtracted from old y_0, and the column being included is added into the formula. The variance is calculated similarly by the equation (6). This calculation allows us to scan the image pixel by pixel in a relatively efficient manner. Figure 1 is an original image of a model. Figure 2 depicts the credibility of using mean and variance as parameters to determine the bimodality at each pixel in the image. In this black and white image, all pixels in the unimodal region are labeled white and all pixels in bimodal regions are labeled black.

Once a histogram is detected to be bimodal, the precise threshold value needs to be determined. In Chow and Kaneko's method [1], a Gaussian curve fitting process is applied to find the optimal value. The drawback of this method is that it is computationally expensive. There are three variables P, μ, and σ that need to be computed for each dominant mode of the histogram, where P is the probability associated with the mean μ and σ is the standard deviation of the Gaussian distribution. This process would be very time consuming if it were to be performed on all windows. In the next section, an efficient method is presented for finding approximated optimal thresholds. This method also solves the problem of adaptive local thresholding for multimodal image.

3 Indexing into a Threshold Hierarchy

When the variance is above a certain level, the histogram is considered to be bimodal. This bimodal histogram is then convoluted with a Gaussian density function using the relation

$$\begin{aligned} F(x,\sigma) &= f(x) * g(x,\sigma) \\ &= \int_{-\infty}^{\infty} f(u)\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,du. \end{aligned} \tag{8}$$

where f(x) represents the bimodal histogram, and $g(x,\sigma)$ is the Gaussian density function that is used for convolution. The local valley bottoms of the smoothed histogram are chosen as thresholds initially. As it is mentioned in the previous section, a unimodal distribution may have a relatively large V_0 value. In this case, no valley point will be well-defined. Hence, it should be re-classified as unimodal.

Optimal thresholds are desired. Clearly, applying the Gaussian fitting method [1] to each bimodal subimages is computationally expensive. A novel indexing method is studied, which uses a global threshold hierarchy structure [13] to determine an approximated local optimal threshold. At the same time, this method also solves the problem of thresholding general multi-modal images.

A threshold hierarchy structure was introduced by Wang, Yang and Symes (WYS) [13] for reliable selection of thresholds at different scales. We briefly review this technique below.

The histogram of an image is generally noisy, making the selection of local minimum points of a histogram as thresholds unreliable. However, this process can be improved by smoothing the histogram. Smoothing is done by Gaussian convolution, $H(x,\sigma) = h(x) * g(x,\sigma)$, where h(x) is the histogram of the image, and σ is the standard deviation of Gaussian distribution. The convolution can be given by

$$H(x,\sigma) = \int_{-\infty}^{\infty} h(u) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} du. \quad (9)$$

Figure 3.(a) illustrates a sequence of Gaussian smoothings of the original histogram.

$H(x,\sigma)$ defines a surface on the (x,σ) plane, where each profile of constant σ is a Gaussian-smoothed version of $h(x)$, with the amount of smoothing increasing with increasing σ. During the smoothing process the location of local minimum points shifts and a plot of such minimum points at different σ levels results in contours in the scale-space (Figure 3.(b)). The concept of the interval tree was introduced by Witkin [15] to reduce the scale-space contour image into a simple interval tree that concisely, but completely, describes the qualitative structure. Figure 3.(c) shows the interval tree derived from Figure 3.(b). This interval tree gives the exact numbers and values for the thresholds at different scale levels. A coarse segmentation may be obtained by using threshold values for a large scale level. The segmented regions can be further segmented into finer regions by considering threshold values at smaller scale levels.

Optimal thresholds may be computed by fitting a Gaussian distribution to each interval in the hierarchy at a certain scale. Those intervals represent principal modes in the histogram. In order to reduce the number of parameters in the curve fitting process, the mean and magnitude of a Gaussian can be approximated directly by the location and peak of each mode in the smoothed histogram. Thus the curve fitting is subject only to the value of standard deviation of each Gaussian curve. It was found that the intersection of the two Gaussian curves results in a new and better threshold compared to the previous methods [16].

Now, we return to the window scanning process

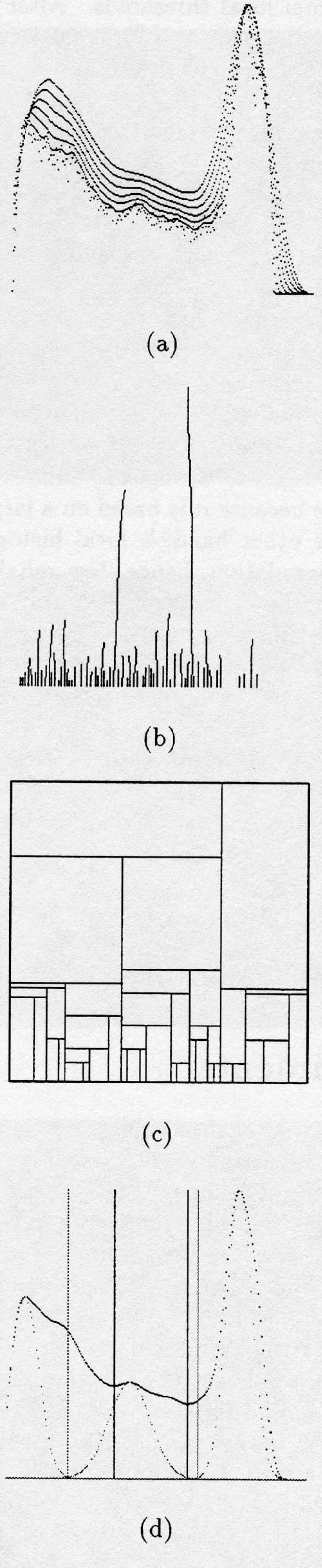


Figure 3: (a) Sequence of Gaussian Smoothing; (b) Contours of local minimum points; (c) Threshold hierarchy; (d) Thresholds : Bold vertical lines - WYS's method, Gray vertical lines - approximated optimal method

for finding optimal local thresholds. After an initial threshold is obtained for a bimodal subimages, the local histogram is divided into two modes and the mean value for each of these modes can be computed. These mean values are used to index into the threshold hierarchy at a chosen scale level. The intersection point between the pair of Gaussian curves corresponding to the two intervals found in the threshold hierarchy provides an approximated local optimal threshold.

The indexing methods has a number of advantages. First, the threshold hierarchy and Gaussian fitting curves need to be computed only once, and can be used throughout the window scanning process. Second, a Gaussian curve for each mode is derived from the global histogram, thus it is less sensitive to noise because it is based on a large population. On the other hand, a local histogram is based a small population, hence, less reliable and more sensitive to noise. Third, by indexing into the hierarchy, the distribution functions for any pair of modes are available immediately, and their intersection point can also be precomputed, thus readily for the thresholding operation.

Once a threshold is obtained for a window, the graylevel associated with the center pixel of that window is modified according to the threshold. The center pixel alone is modified for each window, and after the completion of this procedure, the window is moved one pixel in the horizontal direction. For unimodal cases, a similar procedure is performed except that it is not smoothed and thresholded.

4 Experiments

The new adaptive local thresholding method is applied to the model image (Figure 1). Figure 4.(a) shows thresholded model image by using the thresholds found by threshold hierarchy method [13]. The new technique is tested with two different minimum window sizes. Figure 4.(b) shows the segmented image using minimum window size of 5×5 pixels. This result contains a fewer and larger segmented regions with more smoothed boundaries as compared with the result from global thresholding (Figure 4.(a)). In particular, the narrow gray strips between the cheek and background presented in Figure 4.(a) are successfully eliminated by the new local thresholding method. Also, many other portions, such as hair and background, are segmented into fewer, simpler regions. By changing the minimum window size to 9×9, we expect more small features be eliminated. It is approved by Figure 4.(c). The maximum window size is set at 13×13 for both Figures 4.(b) and 4.(c)

This technique is tested on another landscape image (Figure 5). Figure 5.(a) shows the original image. The results obtained by using the threshold hierarchy method is shown in Figure 5.(b). Figures 5.(c) and (d) show the segmentation results using the new local thresholding method with minimum window sizes of 5×5 and 9×9 respectively. These results are consistent with that of the previous example. That is, the new method generates fewer and simpler segmented regions. Particularly, more small features can be eliminated by increasing the minimum window size.

5 Conclusions

This study presents a new adaptive local thresholding technique. An effective incremental method for discriminating the bimodal regions of the image is developed using mean and variance. A variable sized window is scanned over the entire image, and only the center pixel is changed depending on the graylevel of that pixel as indexed into the threshold hierarchy that is computed from the global histogram. This preliminary study investigated the relationship between the window size of adaptive local thresholding and the thresholded image results. It is found that with different window sizes, an image may be segmented at different levels of abstraction.

Future work may be done to resolve some of the problems associated with the new technique. For example, a fixed maximum window size is used in the current study. However, the scales of features varies across an image in general. An adaptive maximum window size according to the varying local feature scale is desired. An automated method for estimating the scales of features in a local region is both an interesting and a challenge problem. Currently, a fixed V_T is used. It may also be selected adaptively. Any progress made in this direction will make this technique more robust and applicable to a wider range of applications.

Also, the technique uses global histogram for indexing. Sometimes, Gaussian fitting for global histogram may not be similar to local distribution. A technique which can scale Gaussian in the global histogram according to locally derived statistics may generate more accurate thresholding boundaries.

Acknowledgements

The authors appreciate reviewers' valuable and constructive comments. This work is supported in part by NSERC Research Grant OGP0105708.

(a)

(b)

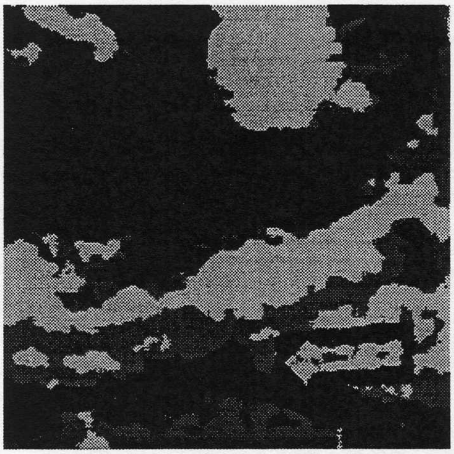
(c)

Figure 4: (a) Segmented image using the threshold hierarchy method; Segmented images with minimum window sizes of (b) 5 × 5 pixels, and (c) 9 × 9 pixels.

(a)

(b)

(c)



(d)

Figure 5: (a) Original image of Landscape; (b) Segmented image using Threshold Hierarchy method; Segmented image using the new method with minimum window sizes of (c) 5 × 5 pixels, and (d) 9 × 9 pixels.

References

[1] Chow, C.K. and T. Kaneko. "Automatic boundary detection of left ventricle form cineangiograms." *Computer Biomedical Research 5*, 1972, 338-410.

[2] Fernando, S.M.X. and D.M. Monro. "Variable thresholding applied to angiography." *Proceedings, 6th International Conference on Pattern Recognition*, 1982

[3] Gonzalez, R.C. and P. Wintz. *Digital Image Processing.* 2nd edition. Addison-Wesley, Reading, MA, 1987.

[4] Haralick, R.M. and L.G. Shapiro. *Computer and Robot Vision, Vol. I*, Addison-Wesley, Reading, MA, 1992.

[5] Kapur, J.N., P.K. Sahoo, and A.K.C. Wong. "A new method for gray level picture thresholding using the entropy of the histogram." *Computer Vision Graphics Image Processing 29*, 1985, 273-285.

[6] Nakagawa, Y. and A. Rosenfeld. "Some Experiments on Variable Thresholding." *Pattern Recognition. Vol. 11*, 1979, 191-204.

[7] Otsu, Nobuyuki. "A Threshold Selection Method from Gray-Level Histograms." *IEEE Transactions on Systems, Man, and Cybernetics. Vol. SMC-9, No. 1*, 1979, 62-66.

[8] Pun, T. "A new method for gray level picture thresholding using the entropy of the histogram." *Signal Processing 2.* 1980, 223-237.

[9] Rao, S.S. *Optimization Theory and Applications.* 2nd edition. Wiley Eastern Limited, New Delhi, 1984.

[10] Reddi, S.S, S.F. Rudin, and H.R. Keshavan. " An Optimal Threshold Selection Scheme for Image Segmentation." *IEEE Transactions on System, Man, and Cybernetics. Vol. SMC-14, No.4*, July/August 1984, 661-665.

[11] Rosenfeld, A. and A.C. Kak. *Digital Picture Processing Vol. 1-2*, 2nd Edition, Academic Press, New York, 1982.

[12] Sahoo, P.K., S. Soltani, A.K.C. Wong, and Y.C. Chen. "A Survey of Thresholding Techniques." *Computer Vision, Graphics, and Image Processing 41*, 1988, 233-260.

[13] Wang, N.Q., X.D. Yang, and L.R. Symes. "Scale-Space Filtering and Threshold Hierarchy for Image Segmentation." *Proceedings of Canadian Conference on Electrical and Computer Engineering*, Sept. 25-27, 1991, 20.4.1-4.

[14] William, H. *Numerical Recipes : The Art of Scientific Computing*, Cambridge, New York : Cambridge University Press, 1986.

[15] Witkin, A.P. "Scale space filtering." *Proc. IJCAI*, Karlsruhe, W. Germany, 1983, 1019-1023.

[16] Yang X.D. and Vipin Gupta. "An Improved Threshold Selection Method for Image Segmentation." *Proceedings of Canadian Conference on Electrical and Computer Engineering*, 1993, 531-534.