Thresholding based on linear diffusion for feature segmentation

Marc-Olivier Baradez[1,2,3], Colin P. McGuckin[2,3], Nicolas Forraz[2,3], Ruth Pettengell[3], Andreas Hoppe[1,3,]*
[1]Digital Imaging Research Centre, Kingston University, UK
[2]School of Life Sciences, Kingston University, UK
[3]King George Laboratory, St. George's Hospital Medical School, University of London, UK
`*a.hoppe@kingston.ac.uk`

Abstract

Segmentation is an important step in image vision. Computationally efficient methods, such as histogram based algorithms, are preferable for real-time applications. In this paper we developed automatic algorithms based on linear diffusion to segment both unimodal and plurimodal histograms. These algorithms are computationally efficient and do not use default value parameters or other heuristics. We demonstrate their potential use to a variety of applications, such as the automatic detection of fluorescence signal in confocal microscopy images and the segmentation of clusters in 2D feature space.

1 Introduction

Image segmentation describes a process in which regions or objects sharing similar features are identified. Thresholding images from their histogram (first order statistics) is a computationally efficient way to segment images and is being used extensively in many image processing applications, in particular for segmenting large data sets or video streams in real-time vision systems. Simple feature segmentation can be performed using a single threshold on image intensity histograms. Other features such as texture, gradient, optic flow can also be used. These images (feature maps) typically express either unimodal or plurimodal histogram distributions.

Unimodal distributions are particularly difficult to segment as no common shape features in the form of modes, minima or maxima are present. Segmenting such images is important, for example in identifying fluorescent signals in microscopy or in the segmentation of gradient images to locate edges in image vision systems. This study was initiated by the need for a robust and computationally efficient segmentation algorithm for images expressing strongly unimodal histogram distributions. In the first part of this paper we present a new thresholding approach based on the analysis of

BMVC 2003 doi:10.5244/C.17.30

dynamic features in scale-space by facilitating linear diffusion. This algorithm is implemented in an extremely computationally efficient way, which should make it suitable for real-time vision applications. The performance of the algorithm is demonstrated on segmenting fluorescent signals in laser scanning confocal microscopy (LSCM), and edge detection in gradient images. This algorithm was also applied iteratively to investigate the segmentation of plurimodal images. In a second approach, linear diffusion was performed directly at histogram level to segment plurimodal histogram distributions of 2D feature spaces.

2 Background

Many histogram-based approaches have been described to segment images, particularly in the general case of plurimodal histograms. Typically, plurimodal histograms display several modes, each of these being usually associated with a class, or object. Non-parametric methods imply thresholding the histogram according to the maximisation, or minimisation, of criterion functions [1] or geometrical considerations. In this case, valleys between modes are considered to indicate the best thresholds that minimise the error between the different classes. In all cases, these methods can reveal to be quite sensitive to small peaks in the histograms. Therefore, these methods are generally applied to smoothed histograms.

However, these approaches do not work on unimodal histograms and other approaches have been developed to segment unimodal histograms. Some of these assume Poisson-like distributions and rely on finding specific inflexion points [2] or define emperically the threshold at intensity $m + k\sigma$ (m = mean intensity, k = constant, σ = standard deviation). These approaches are often heuristic and cannot be applied universally to images as the shape of the histogram may vary unpredictably [3]. In summary, these observations imply that an algorithm based only on shape-information would be unsuitable for segmenting unimodal histograms.

3 Linear diffusion and histogram thresholding

Our approach is based on the principle of linear diffusion. Given a predominant background signal, weaker foreground signals will over time diffuse into the background until equilibrium is reached. The mathematical formulation of such an approach was first described in [4], where any image is shown to be possibly embedded in a one parameter family of derived images, with resolution as the parameter. The dynamic behaviour in scale-space of the images *I(x)* at decreasing scales *t* was demonstrated to be governed by the diffusion equation, which has the general form

$$\frac{\partial I}{\partial t} = c\,\Delta I(x,t)\,. \qquad (1)$$

If c is a constant the diffusion is referred to as linear diffusion. Eventually, the diffusion process reaches a state of equilibrium at $t \to \infty$ where it converges towards the mean signal value.

We present two different techniques based on linear diffusion to segment (3.1) unimodal and (3.2) plurimodal histograms.

3.1 Unimodal histogram thresholding

Unimodal histograms are commonly found in images where a predominant background signal is overlaid with a weaker foreground signal. The occurrence of signal pixels in such images is far less than the occurrence of background pixels and signal pixels typically tend to form clusters. Furthermore, many of these pixels in signal areas can share a wide range of intensities with background pixels as shown in Figure 2. From these observations, we make the hypothesis that signal values will express a different dynamic behaviour in scale-space compared to background values.

Linear diffusion is often implemented by consecutive convolutions of an image with a Gaussian kernel, either in the spatial domain or in the Fourier domain. In that perspective, starting with an original image $I(x)$, the transformation of the images in the scale-space can be described as,

$$I(x,\sigma) = I(x) \otimes G(x,\sigma) \tag{2}$$

where G is the Gaussian convolution kernel of standard deviation σ.

However, this process is computationally expensive and would diminish the computational advantage from using thresholding for image segmentation in real-time systems. Iterating the smoothing of the image with a Gaussian kernel can be approximated by successively reducing its size using intensity interpolation. In this process, the intensities of the image pixels quickly converge to the mean intensity of the whole image. This re-scaling was implemented by vectorising the image I first as $C_{t=0}$

$$C_{t=0} = \{C(1), C(2), \ldots, C(n)\} \tag{3}$$

and then re-ordering C’ such that

$$C'_{t=1} = \begin{bmatrix} C(1) & C(3) & \ldots & C(n-1) \\ C(2) & C(4) & \ldots & C(n) \end{bmatrix}. \tag{4}$$

The values of the vector $C_{t=1}$ are calculated by using mean interpolation down the columns of the matrix C. A normalised histogram H_t is calculated at each step t from vector C_t and ordered to form a resultant matrix H,

$$H = \begin{bmatrix} H_0 \\ \ldots \\ H_{k-1} \end{bmatrix} \tag{5}$$

containing as many rows as re-scaling steps. Equilibrium is reached when the image has been reduced to a single pixel after k steps

$$k = \frac{2\log(M)}{\log(2)} \tag{6}$$

in an M x M image. In other words, the computational efficiency of this algorithm is proportional to the logarithm of the square root of the number of pixels. The dynamic behaviour of each histogram intensity in matrix H over the scales is analysed. The convergence rate R_t of intensity values is calculated as

$$R_t = \left(\frac{1+H_t}{1+H_{t-1}} - 1 \right)^2 \tag{7}$$

and expresses dynamic changes for each intensity value in scale-space along the columns of R.

We expect high convergence rates of signal values and background values at the beginning of the re-scaling process, quickly shifting towards an intensity value at equilibrium after k steps. We chose to describe the overall dynamic behaviour of intensity values over scales by calculating the median convergence rate R_{med} along the columns of matrix R. Background signals express the strongest convergence and hence the lower bounds a of a possible threshold value is defined as

$$a = arg\ max(R_{med}). \tag{8}$$

The upper bounds b of a possible threshold is determined at a position where the majority of *signal* pixels at a particular intensity have been diffused, hence

$$b = arg\ [R_{med}\ (i) = 0] \quad \text{with} \quad i \in [\text{a}; 255]. \tag{9}$$

Due to the unimodal nature of the histogram, the interval [a; b] is generally small and thus the threshold T is defined at the centre of the interval [a; b]. The presented algorithm does not rely on user parameters.

3.2 Plurimodal histogram thresholding

We also investigated the use of linear diffusion to analyse the structure of plurimodal histograms in scale-space, as described by [5] in order to automatically segment N classes. We assume that the different modes observed in the histogram are related to different classes. Often several features are combined in feature space and the differentiation of each feature into several classes is a useful pre-segmentation step. Cluster analysis of different image domains (e.g. local texture in images, entropy, kurtosis) supports this hypothesis, as different clusters represent different classes or features (class-specific peaks or clusters).

Since valleys between modes are the locations to be identified in the histograms, we implemented a scale-space representation by successively convoluting the original 1D histogram with a 3 x 1 averaging filter. Ideally, such process should be iterated an infinite number of times, until total convergence of the histogram to a flat curve (zero-curvature at all points). Practically, for a histogram of 256 intensity values, we found that 500 iterations were sufficient. The size of this scale-space histogram matrix *HS*, was therefore a 256 x 500 matrix, the first row containing the values of the original histogram (the finest resolution). At each (scale) step *i*, the position *j* of all valleys were detected by computing the first derivative of the histogram for locating minima and these were subsequently marked in a second matrix *VS*, i.e. at the position (*i*, *j*) of each detected minima, we assigned *VS(i,j)*=1. As a consequence of the linear diffusion, the smallest minima in the original histogram quickly disappear at coarser scales.

The matrix *VS* contains a complete description of the position of the valleys in the original histogram at all observed scales, therefore embedding the usual approaches which use heuristics smoothing prior to minima detection. A similar approach was described in [5] where the inflexion points of the profile of a scale-space image where identified to characterise the image. The distribution of the ones in *VS* is a natural tree-like distribution. Since this process was applied to a discrete space, the smallest minima over the scales quickly collapse and disappear, therefore not branching to the remaining minima. Such points are referred to as saddle-points in the continuous domain [4], and we use the same name here for convenience. Figure 1 displays *VS* obtained from an arbitrary signal. At the collapsing positions of minima, indicated by the vertical curves, we also plotted a horizontal line joining both closest remaining minima. Therefore, the areas obtained indicate the location of the different classes at all observation scales, the smallest areas being successively merged into a bigger area at the saddle-points, in a tessellation pattern of the scale-space.

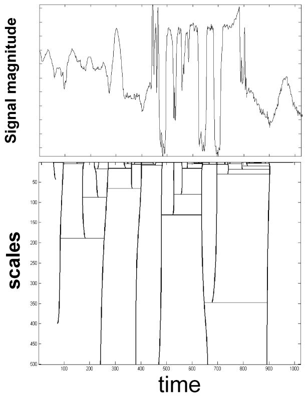

Figure 1: Signal and its corresponding minima tree in scale-space

This simple yet powerful tree-like class description can be used to automatically segment a histogram in N classes, under the assumption that peaks characterise classes (as for cluster analysis in feature space, where measurements discriminate classes). For this purpose, the scale at which only N-1 minima are remaining is identified. The thresholds can be located on the original histogram through a coarse-to-fine tracking procedure, also described in [5].

The positions of the minima over the scales were quite stable, especially at the higher scales. This observation is used for the identification of the N thresholds, which is performed by summing *VS* down the column

$$f(j)=\sum_{i=1}^{500} VS(i,j)\,. \tag{10}$$

The function ***f*** displays several peaks whose heights are proportional to the occurrence of minima down the scales. Therefore, the highest peaks indicate the strongest minima at all scales. Due to the slight shift of the minima over the scales, the peaks in ***f*** could have different shapes, and identifying those peaks by locating their maxima did not provide a sufficient result. Instead, we integrated the area under each peak and identified their mean position. The positions for the N-1 main peaks consequently provide the thresholds to segment the desired N classes in the histogram.

4 Results

In this paper thresholding is used to distinguish particular features in images. In the case of unimodal histograms, there is usually one predominant signal (background) and a second, weaker signal (feature). Depending on the investigated images, these may represent fluorescent spots in microarrays, edges in gradient images, fringe patterns in shearography imaging or interferometry. In this section, we show a particular application of our algorithm for segmenting 198 unimodal Laser Scanning Confocal Microscopy images [LSCM]. These results were compared with those obtained by a bayesian approach we previously described, which relies upon the identification of a representative background region. This algorithm was further applied to segment phase contrast microscope images and gradient images.

The algorithm designed to automatically threshold plurimodal histograms was applied to a variety of histograms where peaks were class-specific. To demonstrate a more elaborate segmentation approach, we also extended the application of this algorithm to automatically segment clusters in a 2D feature space

4.1 Example of unimodal histogram segmentation

In laser scanner confocal images, the histogram of the whole image and the histogram of the background area strongly overlap over a wide range of intensities, which greatly complicates any approach by directly working from the original image histogram. Smoothing the original image is therefore a necessary initial step, justified as it allows merging signal pixels together. This produces more homogeneous signal areas and makes the background more uniform while reducing noise. This is consistent with other published studies confronted to similar problems [6]. Figure 2 shows a typical LSCM image expressing weak fluorescent signals and its corresponding histogram.

Figure 3 shows a comparison between the mean intensities of signal areas from a set of 198 LSCM images, obtained with the linear diffusion approach detailed in this paper versus those obtained with a Bayesian approach. There is good agreement between both methods (r = 0.98). However the diffusion algorithm is computationally highly efficient and does not require the identification of a representative background area.

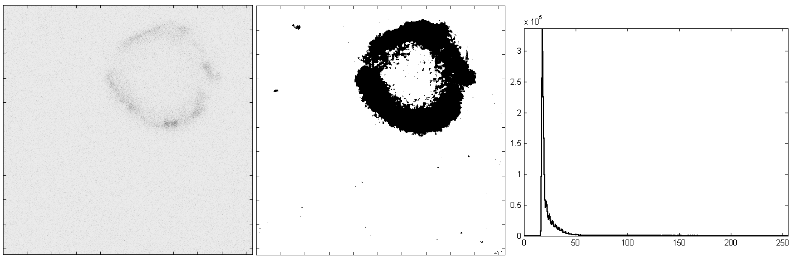

Figure 2: Original negative LSCM image, thresholded signal area, unimodal histogram

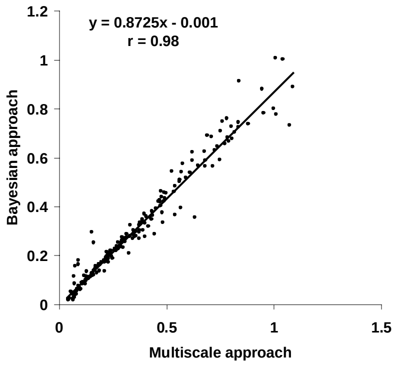

Figure 3: Comparison of the mean intensities in segmented signal areas from a set of 198 LSCM images. The X-axis represents the values obtained with the diffusion approach detailed in this paper and the Y-axis those obtained with a Bayesian approach.

Figure 4 shows the application of this algorithm to segment cells in phase-contrast images of neuron cells. Global thresholding of such images would not segment the cell areas correctly. However, identifying and subsequently segmenting the high frequency components using our algorithm enables to give a first estimate of the cell locations.

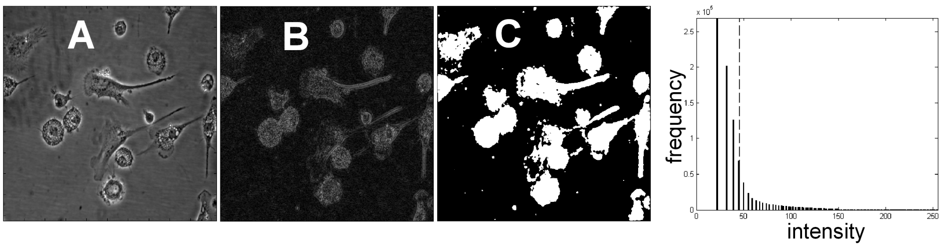

Figure 4: A – Bright field microscope image of neuron cells derived from blood stem cells in culture medium, B – high frequency areas, C – Binary image of B identifying cell areas, automatically obtained as described in 3.1, D – unimodal histogram of B.

In figure 5, this algorithm was applied to automatically binarise a gradient image, therefore enabling further processing such as edge or corner detection.

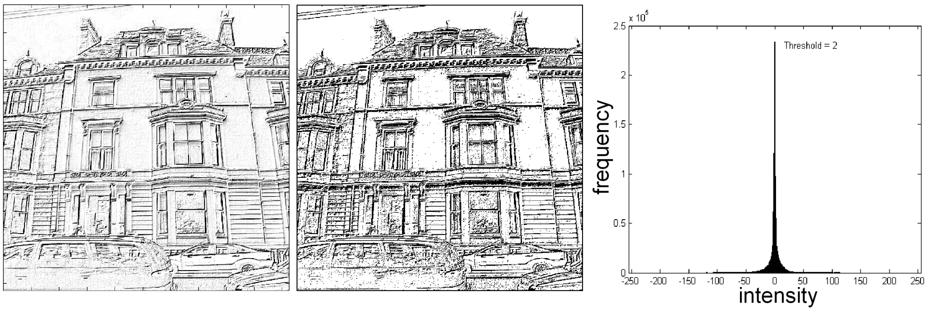

Figure 5: Automatic binarisation of gradient image using linear diffusion segmentation.

Finally, we investigated the possibility of applying this algorithm to segment plurimodal images. The algorithm was applied iteratively, removing the segmented signal areas after each iteration, as shown in figure 6.

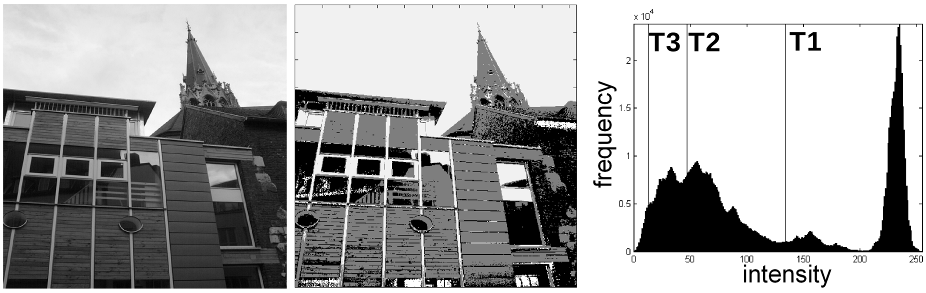

Figure 6: Application of the diffusion algorithm initially designed to segment unimodal images (3.1) to a plurimodal image. This image is segmented into 3 classes, respectively white, grey and black colors. T1, T2 and T3 indicate the thresholds, as calculated by the algorithm.

4.2 Examples of plurimodal histogram segmentation

Figure 7 shows a typical grey-intensity histogram displayed expressing several modes. This histogram was segmented into N = 6 classes. The bigger arrows indicate the position of the N-1 thresholds automatically calculated by the algorithm. It can be verified that these thresholds correlate well with the main valleys of this histogram curve. Smaller arrows (pointing up) indicate the position of other, less significant, thresholds. Similar results were obtained on histograms from a variety of images.

To demonstrate a more elaborate segmentation approach, we also extended the application of this algorithm to automatically segment clusters in a 2D feature space. For example, the image shown in Figure 8 was scanned with a moving window in which both the standard deviation and the mean value of these intensities were used as features.

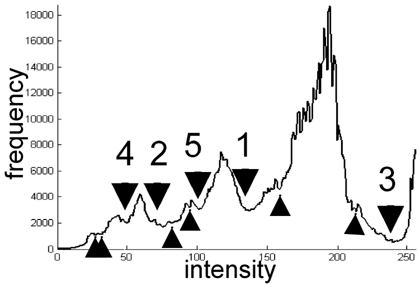

Figure 7: Typical plurimodal histogram segmented in 6 classes. The five thresholds are indicated by arrows, and the numbers represent the order in which they were detected.

We generated a scatter plot of the standard deviations versus the mean intensities and both data sets were normalised in the range [0; 256]. The histograms H_x and H_y along each axis were computed with a bin size of 1. These histograms were subsequently processed with the plurimodal histogram thresholding algorithm, and the first six main thresholds obtained for both the horizontal and the vertical directions were displayed as dashed lines overlapping the scatter plot in Figure 8. The width of these lines is proportional to the order in which the thresholds were detected (therefore, the threshold detected at the coarsest scale is the widest one, the second threshold detected at a finer scale is slightly thinner, and so on). It can be observed that this allowed to segment regions of the scatter plot, well separating most predominant clusters. Since the algorithm is not very sensitive to the height of the peaks, it can be observed that the small clusters (located in a vertical parabolic way) are mostly contained in a separate segmented square.

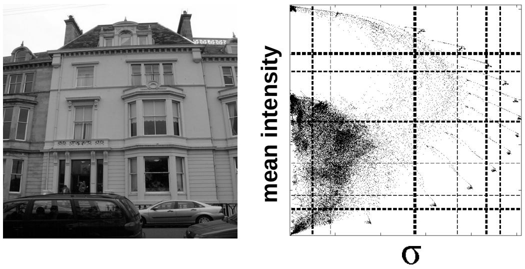

Figure 8: Original image and its 2D feature space

The dense region of the scatter plot is magnified in Figure 9 and is shown together with the corresponding regions in the original image, shown as binary images surrounding the scatter plot. They appear to correlate very well with coherent regions, e.g. smooth-bright, smooth-dark, coarse-bright.

5 Conclusions

In this paper we used linear diffusion to investigate dynamic features in scale-space for thresholding images expressing unimodal and plurimodal histogram distributions. We have implemented two computationally efficient scale-space approaches in which diffusion of all signal values is completed after a short number of iterations. In both methods, threshold values are obtained automatically from analysing dynamic features in scale-space.

In the case of unimodal histograms, the obtained results were in good agreement with human observers, as demonstrated by segmenting confocal microscope images. Plurimodal histograms were segmented using 1D scale-space analysis in the histogram domain. This was extended for a 2D feature-space analysis for cluster segmentation resulting in proper identification of coherent regions in images.

We think that these algorithms could be adapted to find a broader range of applications in image vision systems, especially for real-time segmentation.

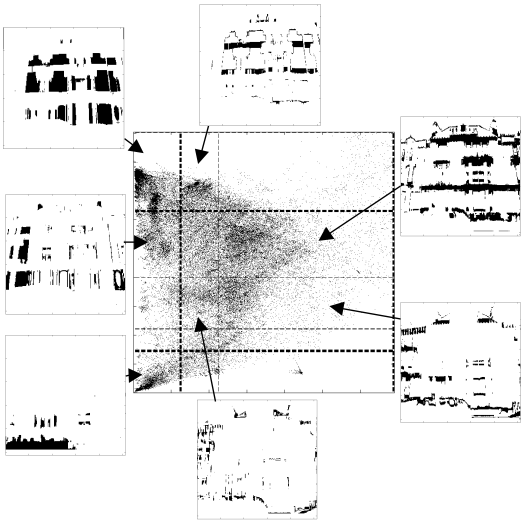

Figure 9: Segmented clusters and corresponding regions in the original image

6 References

[1] N Otsu. A threshold selection method from gray-level histograms. *In IEEE Transactions on Systems, Man, and Cybernetics* 9 (1979) pages 919-926

[2] C Antoine, M D Lloyd, J Antoine. A robust thresholding algorithm for halftone dots. *Journal of Pulp and Paper Science* 8 (2001) pages 268-272

[3] D-M Tsai. A fast thresholding selection procedure for multimodal and unimodal histograms. *Pattern Recognition Letters* 16 (1995) pages 653-666

[4] J Koenderink. The structure of images. *Biol. Cybern.* 50 (1984) 363-370

[5] A P Witkin. Scale filtering . *In Proc. 8th Int. Joint Conf. on Artificial Intelligence*, (1983), pages 1019-1022

[6] H O Nyongesa, A W Otieno, P L Rosin. Neural fuzzy analysis of delaminated composites from sherography imaging. *Composite Structures* 54 (2001) pages 313-318