Thresholding for Change Detection

Paul L. Rosin
Department of Information Systems and Computing,
Brunel University,
Uxbridge,
Middlesex UB8 3PH
UK
email: Paul.Rosin@brunel.ac.uk

Abstract

Image differencing is used for many applications involving change detection. Although it is usually followed by a thresholding operation to isolate regions of change there are few methods available in the literature specific to (and appropriate for) change detection. We describe four different methods for selecting thresholds that work on very different principles. Either the noise or the signal is modelled, and the model covers either the spatial or intensity distribution characteristics. The methods are: 1/ a Normal model is used for the noise intensity distribution, 2/ signal intensities are tested by making local intensity distribution comparisons in the two image frames (i.e. the difference map is not used), 3/ the spatial properties of the noise are modelled by a Poisson distribution, and 4/ the spatial properties of the signal are modelled as a stable number of regions (or stable Euler number).

1 Introduction

Due to its simplicity image differencing is a popular method for change detection. It only requires calculating the absolute values of the difference between the corresponding pixels in two images, and large values in the difference map then indicate locations of change. Common applications of image differencing include object tracking [25], intruder surveillance systems [3, 5], vehicle surveillance systems [7, 8, 11], and interframe data compression [2]. There are also many examples of its use for analysing satellite images [20] to measure land erosion, deforestation, urban growth, crop development, etc., and for analysing medical images to measure cell distribution [10], etc.

The difference map is usually binarised by thresholding it at some pre-determined value to obtain a change/no-change classification. However, the threshold value is critical, since too low a value will swamp the difference map with spurious changes, while too high a value will suppress significant changes. The proper value of the threshold is dependent on the scene, possibly fluctuating camera levels, as well as viewing conditions (e.g. illumination) which may change over time. This indicates that in general the threshold value should be calculated dynamically based on the image content, and that experimentally selecting a value (e.g. Jain [8], Koller *et al.* [11]) is not appropriate for a robust autonomous vision system.

As an extension to global threshold determination there are various other procedures that can improve change detection. Local thresholding can be useful, particularly when the scene illumination varies locally over time. Noisy difference maps can be much improved by removing small isolated change pixels, merging close regions of change, incorporating connectivity, and performing hysteresis thresholding [1, 11, 18, 25]. Rather than differencing adjacent frames in temporal image sequences background images can be dynamically generated [12, 18, 25], and these are differenced with each image instead. However, these issues will not be further explored in this paper.

Several general approaches are possible for determining thresholds for change detection. First the signal, noise, or both can be modelled. Second, either their intensity and/or spatial properties can be modelled. In this paper we consider all of the four combinations of pairs of object and property (signal/noise and intensity/spatial distribution) and describe four techniques for threshold determination that fall into each of these categories (see table 1). Note that most standard intensity image thresholding techniques belong to the signal/intensity class.

	noise	*signal*
intensity	Normal model	intensity distributions
spatial	Poisson model	stable number of regions

Table 1: Techniques for threshold selection described in this paper

2 Previous work on threshold selection

Although there are many techniques for determining image threshold values for binarising intensity images [19] most of them make assumptions that do not hold for standard difference images, thereby invalidating their application. For instance, many methods expect the intensity histogram to be bimodal, and some cannot cope if the modes are too dissimilar in size, or if they are not approximately Normal. We know of rather few thresholding techniques specifically designed to be effective for difference maps; these are described below.

Yang and Levine [25] employed a two stage process to select local and global thresholds for differenced edge maps using the full image sequence. First, for each edgel, the median absolute deviation (MAD_e) of all the edgels at that location over the sequence is calculated. The edgel's local difference map threshold is then set to $2.5 \times 1.4826 \times \text{MAD}_e$. Next, a histogram is formed from those edgels which are above their local threshold values. The median absolute deviation MAD_h calculated over the histogram is used to generate a global threshold $2.5 \times 1.4826 \times \text{MAD}_h$. Edgels now are marked as change only if they exceed both their local and the global thresholds.

Jain and Nagel [9], and later Hsu *et al.* [7], bypassed the difference map and detected change using hypothesis testing. The images were modelled as patches whose intensities were described by bivariate polynomials. For each pixel a likelihood test checked whether the intensities within a local window in each of the two images could have been drawn from a single intensity distribution. Since the *t*-test is used the threshold can then be related to a confidence level.

3 Modelling the noise intensity

3.1 Noise Models

In many instances it is reasonable to assume that the image noise can be modelled by a zero mean Normal distribution $N(0, \sigma^2)$. In this case, analysing the difference in intensity images is straightforward. Differencing followed by taking the absolute value will produce the Normal distribution $2N(0, 2\sigma^2)$ for positive values only.

Sometimes it is preferred to difference edge maps rather than intensity images as they can be more robust for change detection under varying illumination [1, 21, 25]. Unfortunately this makes determining the distribution of the noise more troublesome. First we note that the noise in the edge maps can be modelled by a Rayleigh distribution [24]

$$R(x) = \frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}};\ x \geq 0$$

if the edge response is of the form $\sqrt{\left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2}$. Denoting the noise in two edge maps as independent random variables $\mathbf{a}$ and $\mathbf{b}$ we wish to calculate the density of the noise in the difference image $\mathbf{d} = |\mathbf{a} - \mathbf{b}|$. Initially we consider the symmetric function $\mathbf{c} = \mathbf{a} - \mathbf{b}$. Its density $f_{\mathbf{c}}(c)$ equals the convolution of the densities of $\mathbf{a}$ and $\mathbf{b}$ [14]

$$f_{\mathbf{c}}(c) = \int_{-\infty}^{\infty} f_{\mathbf{a}}(c - x) f_{\mathbf{b}}(-x) dx.$$

For the Rayleigh functions $f_{\mathbf{a}}(x) = f_{\mathbf{b}}(x) = R(x)$ we obtain, using Mathematica,

$$f_{\mathbf{c}}(x) = \frac{\sqrt{\pi}}{8\sigma^2 e^{\frac{x^2}{4\sigma^2}}} \times \left[2\sigma - \frac{x^2}{\sigma} + 2\sqrt{\pi} x L_{\frac{1}{2}}^{\frac{-1}{2}}\left(\frac{-x^2}{4\sigma^2}\right) - \sqrt{\pi} x L_{\frac{1}{2}}^{\frac{1}{2}}\left(\frac{-x^2}{4\sigma^2}\right)\right]$$

where $L_n^a(x)$ is the generalised Laguerre polynomial. Since $\mathbf{d} = |\mathbf{c}|$, then $f_{\mathbf{d}}(x) = 2f_{\mathbf{c}}(x);\ x \geq 0$ (this assumes that $f_{\mathbf{c}}$ is even, although it is actually undefined for negative values in the form given above).

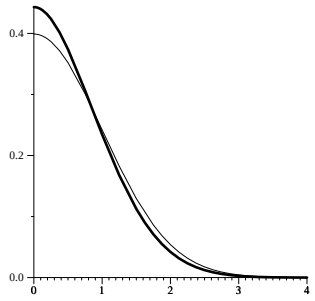


Figure 1: Probability density function of $f_{\mathbf{c}}$ and $N(0, 1)$

Looking at the plot of $f_{\mathbf{c}}$ drawn bold alongside a Normal distribution (figure 1) we see that they are similar, especially near the tails. We are mainly interested in the distribution function towards the tails since we wish to threshold out most of the noise. Therefore we can reasonably approximate $f_{\mathbf{d}}$ by a Normal distribution. When thresholding at $x\sigma$ the probability of incorrectly classifying a pixel as motion is

$$P_F = \text{erfc}\left(\frac{x}{\sqrt{2}\sigma}\right).$$

This enables us to choose a suitable threshold τ for a given acceptable proportion of false motion pixels.

3.2 Noise Estimation

In practise the variance of the noise is often unknown so we need to estimate it from the image. Since the difference image will contain not just noise but also appreciable amounts of signal due to the motion a robust estimation technique is required. Similar to our previous work in estimating noise levels in edge maps [16] we use the Least

Median of Squares (LMedS) method applied to the difference image histogram. Its advantages are that it is efficient (at least for 1D data), and has a high breakdown point. This latter property enables it to return the correct result even when large amounts of outliers (i.e. true motion) are present. It is straightforward [17] to derive the following relation between the LMedS and the expected standard deviation of the noise:

$$\sigma = \frac{\text{LMedS}}{0.33724}.$$

4 Modelling the signal intensity

While it is reasonable to assume that the noise characteristics are known sufficiently such that they can be modelled analytically and estimated from the difference image, we have no information concerning the intensities of the regions of change. Likewise, we know little about the magnitudes of the difference map intensities produced by change, except that we would generally expect them to be significantly larger than zero. Since this provides little help for threshold selection we do not try to analyse difference map intensities, but consider the original (pre-differenced) images instead. Change occurs when corresponding pixels in the two images have significantly different intensities, but it is difficult to quantify what is meant by significant. One solution is to consider in addition to each individual pixel all the intensities in a surrounding window. Comparing two windows is a more tractable task since there are many techniques available for comparing two distributions, and this is the approach taken by Jain and Nagel [9] and Hsu *et al.* [7]. In contrast to them we shall use a non-parametric method so that no assumptions about the intensity distributions need to be made. One of the most popular is the Kolmogorov-Smirnov test which was used for change detection in satellite imagery by Eghbali [4]. It has the nice property that it is invariant to image scaling or offsets. As an alternative we have also experimented with the Cramér-von Mises test which is often more powerful than the Kolmogorov-Smirnov test [22]. Rather than testing the maximum value of the difference between the two cumulative distributions it uses instead the summed squared differences. Thresholding is performed by accepting as motion only those pixels whose distributions are dissimilar, i.e. their test statistic is above the critical value for a selected significance value (e.g. 5%).

Following Eghbali we normalise the data to reduce sensitivity to large scale variations between the images. Each window is rescaled and offset so that it has zero mean and unit standard deviation.

5 Modelling the spatial distribution of the noise

If we assume that the noise is white then its spatial distribution over the image will be random. For the analysis of spatial data there are many measures of randomness [23], often based on the assumption that the observations follow a Poisson distribution. Since a Poisson distribution has its mean equal to its variance then the ratio of the sample variance to the sample mean is a natural test for that distribution, and is called the *relative variance* $V_r = \frac{s^2}{\bar{x}}$. It is calculated by first counting the number of observations (in our case the number of above threshold pixels in the difference map) in n windows, $\{x_i\}_{i=1}^n$, from which the mean, $\bar{x}$, and variance, s^2, of the x_i can be found. Although the test is sensitive to the window size and point density it works adequately as long as $\bar{x}$ is sufficiently large.

For our purposes we do not wish to detect the spatially random noise, but rather to avoid it in our thresholded image. We therefore select the threshold which maximises the relative variance, thereby maximising "clumpiness" (regions of change) and minimising the Poisson distribution (noise).

6 Modelling the spatial distribution of the signal

The location, size, and number of the regions of change are generally unknown. However, we might expect that these properties will remain fairly stable over a wide range of threshold values, whereas down at the noise level small changes in the threshold value can substantially alter the number of regions. Such an observation suggests that if a range of threshold values is found that leads to a stable number of regions, then these regions are unlikely to come from noise, and so a value from this range will provide a suitable threshold. This approach was suggested by O'Gorman [13] for intensity image thresholding (and was recently proposed again by Pikaz and Averbuch [15]), and was applied to difference images by Rosin and Ellis [18]. Rather than counting the number of regions the image's Euler number can be used, and was found to give almost identical results [18]. The advantage of calculating connectivity over region counting is that the Euler number is locally countable [6], and can be determined efficiently in a single raster scan of the image by just a few lines of code.

A stable threshold range will correspond to a plateau in the graph of the Euler numbers against thresholds. Initially we assumed that the plateaus were perfectly flat, and detected them by looking for the longest such range in the graph [18]. However, given the noisy, fragmented nature of images this was not found to be reliable as sometimes the Euler number varied slowly within the stable range. An alternative procedure that we have found more effective is

to model the shape of the graph as a decaying exponential. At low threshold values there will be many regions and holes caused primarily by the noise, and the Euler number will change rapidly with threshold. At high threshold values there will be few regions, and the Euler number will be stable. (We only consider thresholds up until there are no regions remaining, and the Euler number becomes zero). Therefore a suitable partition point between the signal and noise is the "corner" of the curve. which we find as the point on the curve with maximum deviation from the straight line drawn between the end points of the curve.

If we count regions then the curve at low thresholds which is generated primarily by noise will appear roughly Gaussian [15] and the number of regions eventually drops to one at a threshold of zero. When using Euler numbers the shape is different since the number can become negative if there are more holes than connected components. To avoid difficulties for the corner detection we start the straight line from the first positive peak in the curve (see for example figure 2n).

7 Experimental results

We demonstrate the various algorithms on several pairs of frames taken from sequences of 512×512 images. For the techniques that compare intensity distributions the window size was 21×21. The test for the Poisson distribution used 64×64 windows. P_F was set to 0.01 for the Normal noise intensity model, while 5% confidence values were used for the test statistics. Apart from the input images the remaining image are shown inverted to improve visualisation.

Figures 2a&b show two frames of an outdoor scene containing a pedestrian walking across a carpark. The results of thresholding the difference map (figure 2c) using the Normal model for noise intensity are shown in figure 2d. A reasonable fit using the Normal approximation is made to the difference map histogram, and the walker has been thresholded reasonably well although there is some residual noise. Applying the same technique to a difference image constructed from two frames after running the Sobel edge detector produces a noisier result (figure 2e). In this case the difference map histogram is not so well fit near the tail.

Using the methods for comparing intensity distributions, the Kolmogorov-Smirnov and Cramér-von Mises tests produce the confidence values in figures 2f&g. Both performed poorly in this case since either little or no response is retained after thresholding. We also demonstrate a related approach for comparing two sets of values: correlation. Pearson's rank correlation coefficient is shown thresholded at 0.05 (figure 2h) although the coefficient cannot strictly be interpreted as a probability value. Likewise, a non-parametric method, Spearman's rank correlation, is shown after thresholding in figure 2i. Although the correlation results are better than the distribution comparison results, they are still poor.

Both of the spatial methods (the Poisson noise model and the stable Euler number) produce good results (figures 2j&k). The corner finding procedure is demonstrated on the plot of the Euler number in figure 2l.

Finally, the images in figures 3a&b are outdoor images from a test surveillance site, taken under a variety of conditions [5], and are rather more demanding. They are poorly illuminated and contain only two sets of small moving objects (rabbits). Due to the low contrast the threshold obtained by the Normal distribution method retains many spurious noise points. The distribution comparison and correlation methods performed poorly (figure 3e-j), either under- or over-thresholding except for Spearman's correlation coefficient which performed quite well. As before, the two spatial methods produced good results (figures 3k&l).

8 Conclusions

Future work is required to perform more extensive testing, and to incorporate some quantitative assessment. However, from the results above we can form some initial conclusions. First, the Normal model for approximating the noise intensity in combination with the noise estimation procedure works adequately, although the results tended to look speckly. One reason for this is that the method performs the thresholding on a local pixel basis, in contrast to most of the other methods which operate over windows or regions. The speckle could be reduced either by modifying the threshold, but more probably post-processing such as erosion would be simpler and more reliable.

The two techniques we considered for comparing intensity distributions, the Kolmogorov-Smirnov and Cramér-von Mises tests, did not do well. A further disadvantage is that the results are fairly sensitive to the window size which needs to be specified as an additional parameter. The correlation methods for comparing windows, Pearson's rank correlation coefficient and Spearman's rank correlation, also performed poorly. The Kolmogorov-Smirnov and Pearson methods appeared to be particularly sensitive to noise, especially in regions with small dynamic ranges of intensities.

Most promising were the spatial methods. Both the Poisson noise model and the stable Euler number reliably gave good results, which for all the images either matched or bettered the Normal noise intensity model results. However, further tests are necessary to determine the robustness of the peak finding and corner finding procedures used by the two methods.

References

[1] M. Bichsel. Segmenting simply connected moving objects in a static scene. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16:1138–1142, 1994.

[2] S.C. Brofferio. An object-background image model for predictive video coding. *IEEE Trans. Communications*, 37:1391–1394, 1989.

[3] I. Dinstein. A new technique for visual motion alarm. *Pattern Recognition Letters*, 8:347–351, 1989.

[4] H.J. Eghbali. K-S test for detecting changes from Landsat imagery data. *IEEE Trans. on Systems, Man, and Cybernetics*, 9:17–23, 1979.

[5] T.J. Ellis, P. Rosin, and P. Golton. Model-based vision for automatic alarm interpretation. *IEEE Aerospace and Electronic Systems Magazine*, 6(3):14–20, 1991.

[6] S.B. Gray. Local properties of binary images in two dimensions. *IEEE Trans. Computers*, 20:551–561, 1971.

[7] Y.Z. Hsu, H.H. Nagel, and G. Rekers. New likelihood test methods for change detection in image sequences. *Computer Vision, Graphics, and Image Processing*, 26:73–106, 1984.

[8] R. Jain. Extraction of motion information from peripheral processes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 3:489–504, 1981.

[9] R. Jain and H.-H. Nagel. On the analysis of accumulative difference pictures from image sequences of real world scenes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1:206–214, 1979.

[10] T.F. Knoll, L.L. Brinkley, and E.J. Delp. Difference picture algorithms for the analysis of extracellular components of histological images. *J. Histochem. Cytochem.*, 33:261–267, 1985.

[11] D. Koller, J. Weber, and J. Malik. Robust multiple car tracking with occlusion reasoning. In *Europ. Conf. Computer Vision*, pages 189–196, 1994.

[12] A. Makarov. Comparison of background extraction based intrusion detection algorithm. In *Int. Conf. Image Processing*, pages 521–524, 1996.

[13] L. O'Gorman. Binarization and multi-thresholding of document imnages using connectivity. In *Symp. on Document Analysis and Info. Retrieval*, pages 237–252, 1994.

[14] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 1991.

[15] A. Pikaz and A. Averbuch. Digital image thresholding based on topological stable state. *Pattern Recognition*, 29:829–843, 1996.

[16] P.L. Rosin. Edges: saliency measures and automatic thresholding. *Machine Vision and Applications*, 9:139–159, 1997.

[17] P.L. Rosin. Thresholding for change detection. Technical Report ISTR-97-01, Brunel University, 1997.

[18] P.L. Rosin and T. Ellis. Image difference threshold strategies and shadow detection. In *British Machine Vision Conf.*, pages 347–356, 1995.

[19] P.K. Sahoo, S. Soltani, A.K.C. Wong, and Y.C. Chen. A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41:233–260, 1988.

[20] A. Singh. Digital change detection techniques using remotely-sensed data. *Int. J. Remote Sensing*, 10:989–1003, 1989.

[21] K. Skifstad and R. Jain. Illumination independent change detection for real world image sequences. *Computer Vision, Graphics, and Image Processing*, 46:387–399, 1989.

[22] P. Sprent. *Applied Nonparametric Statistical Methods*. Chapman and Hall, 1993.

[23] G.J.G. Upton and B. Fingleton. *Spatial Data Analysis by Example, Volume 1, Point Pattern and Quantitative data*. Wiley, 1985.

[24] H. Voorhees and T. Poggio. Detecting textons and texture boundaries in natural images. In *Int. Conf. Computer Vision*, pages 250–258, 1987.

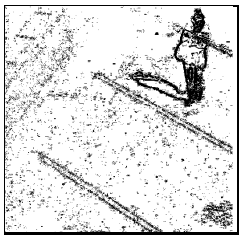
[25] Y.H. Yang and M.D. Levine. The background primal sketch: An approach for tracking moving objects. *Machine Vision and Applications*, 5:17–34, 1992.

(a) frame 1

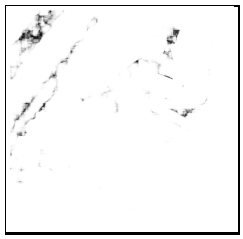
(b) frame 8

(c) difference map

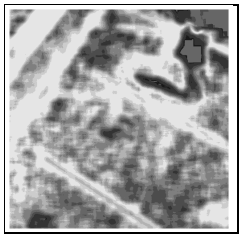
(d) Normal noise intensity



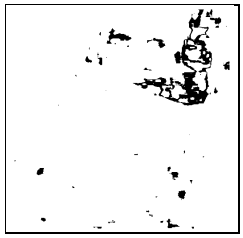
(e) Normal noise intensity edge



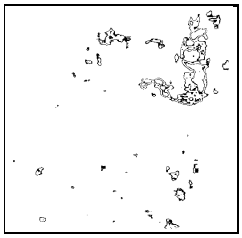
(f) K-S



(g) Cramer



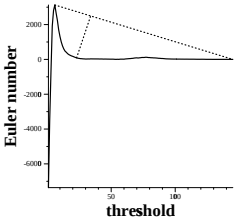
(h) Pearson 5%



(i) Spearman 5%

(j) Poisson model

(k) Euler

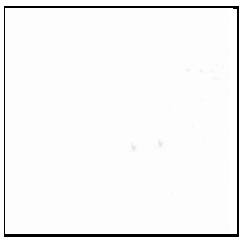


(l) Euler corner

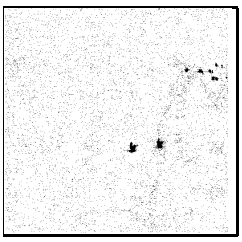
Figure 2: Pedestrian crossing car park

(a) frame 1

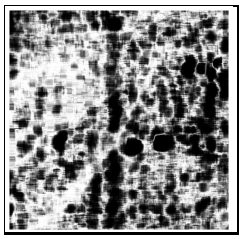
(b) frame 8



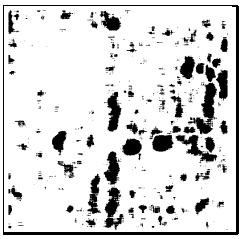
(c) difference map



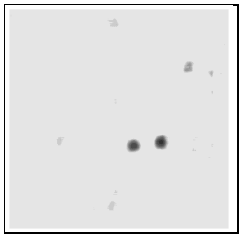
(d) Normal noise intensity



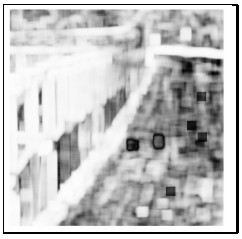
(e) K-S



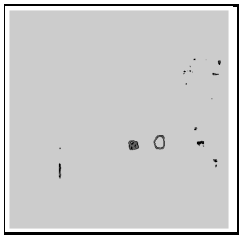
(f) K-S 5%



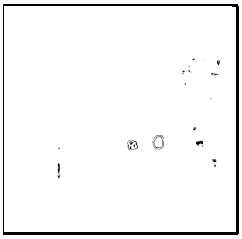
(g) Cramer



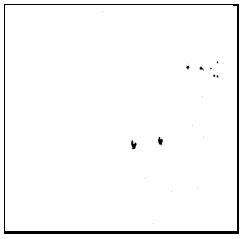
(h) Pearson



(i) Spearman



(j) Spearman 5%



(k) Poisson model

(l) Euler corner

Figure 3: srdb018 image sequence containing moving rabbits