Development of an Extension of the Otsu Algorithm for Multidimensional Image Segmentation of Thin-film Blood Slides

Bernard F. Buxton, Houari Abdallahi
Department of Computer Science, UCL
Gower Street, London, WC1E 6BT
{B.Buxton , H.Abdallahi} @cs.ucl.ac.uk

Delmiro Fernandez-Reyes, William Jarra
National Institute for Medical Research, MRC
Mill Hill, London, NW7 1AA
{dfernan , wjarra} @nimr.mrc.ac.uk

Abstract

Commencing with the Fisher discrimminant, the Otsu algorithm for image segmentation is reviewed and a multi-dimensional extension proposed. Identification of the pixels belonging to red-blood cells in 1300×1030 *24-bit colour images of thin-film microscope slides of laboratory samples of malarial infected blood is used to illustrate the Otsu algorithm and the proposed extension. The Otsu algorithm is applied to the image intensity and independently to each of the colour channels in a set of 142 images and the best combination of the colour channel outputs that maximizes the Fisher discriminant selected. ROC and MRROC curves and error rates are used to evaluate performance. The multi-dimensional extension of the Otsu algorithm is similarly assessed. Preliminary results obtained from sequential application of the multi-dimensional extension of the Otsu algorithm are presented and directions for further research discussed.*

1 Introduction

Segmentation is a classic problem in image processing and computer vision. Thus, it is discussed in most text books, both in image processing [18, 29, 14] and in computer vision [19, 40, 13], though there are often differences in approach. A wide variety of techniques are used in both kinds of applications, ranging from hand crafted algorithms based on user evaluation or interaction, model based methods, and machine learning.

Model-based approaches, including the construction of models of the object of interest, the imaging geometry and photometry [9, 24], models of the object's shape and appearance [3] and models of the effect of changing viewpoint [5, 44] are popular in computer vision research and applications. The application of machine learning methods have also become more popular recently both in image processing and in computer vision. This is a result of increased computer power, improved learning algorithms and improved statistical methods [2]. Of particular interest is the application of capacity-controlled supervised machine learning algorithms, such as those based on the support vector machine technology [28, 1]. Capacity control combats the need for the huge numbers of training examples that would otherwise be required owing to the very-high dimensionality of the pixel space of an image. Multi-resolution, pyramid and wavelet transform filtering methods are often employed which have the additional benefits of introducing the spatial coherence between neighbouring pixels in an image and thereby making the statistics of the filtered images predictable [21, 15]. Reducing the resolution also helps combat the computational complexity of such supervised machine learning algorithms.

Principal component analysis [20] is also a very effective way of reducing the dimensionality of the image representation, in particular when applied to specific types of object as by Cootes and Taylor [4] or, when incorporating hidden landmark point variables as in the flexible appearance methods, also developed by Cootes and Taylor [3]. The characteristics of an object or class of objects may then often be learnt from a few hundred or thousand examples, even when the object has internal degrees of freedom, as in the expression of a face [6].

1.1 Supervised and unsupervised learning

In spite of these advances, an unavoidable disadvantage of such supervised learning approaches is the requirement for manual interaction to obtain training data. Even though an initial system may be constructed from a small training set to be used as a tool for assisting manual labelling of a large training set [4], acquiring training sets can be time consuming and laborious. Furthermore, if the training data is to be accurate, segmentation of the training images should be carried out by an expert. Their time is often a scarce resource and different experts may not always agree [41].

Proceedings of the International Conference on Computing: Theory and Applications (ICCTA'07)
0-7695-2770-1/07 $20.00 © 2007 IEEE

In machine learning terms, in unsupervised methods, labels are assigned according to a criterion that reflects the underlying characteristics of the problem. In image processing, this often amounts to the choice of a predicate according to which segmentation is to be carried out and is most common in the clustering of pixels in a colour space. The choice of colour space, involving a linear or non-linear transformation of the image data, clustering criterion, initialisation and the algorithm employed are all important in the design of such systems. Often unsupervised algorithms are based on statistical criteria which may be parametric, as in the application of Gaussian mixture models in colour image segmentation, or non-parametric when a statistical criterion may be used that reflects overall properties of the pixel data. Amongst the latter is the classical Fisher discriminant or linear discriminant analysis (LDA) according to which one seeks a labelling or classification of image pixels that maximises the ratio of a between class variance to a within class variance [11, 8]. When applied to a one-dimensional feature space, such as pixel intensity, this leads to the Otsu algorithm whose implementation [19] seems rather better known than its origin [27].

We have recently been exploring the application of such methods for segmentation of blood cells and malarial parasite infections in the analysis of images taken from thin-film slides of stained blood samples. Though the work is in its early stages this approach and, in particular, an extension of the Otsu algorithm to the three-dimensional colour space appears to be remarkably successful. In the remainder of this paper we therefore in section 2 briefly introduce the malarial application and the long-term objectives of our work. In section 3, we then describe an application of the well-known one-dimensional Otsu algorithm applied to the three colour channels, independently and in combination.

In section 5, we describe our extension of the Otsu algorithm, which to the best of our knowledge is novel, and a sequential application of it to obtain a segmentation or labelling of blood plasma (and some uninteresting artefacts), of the red blood cells themselves, and of their boundaries (i.e. of the cell membrane). Some preliminary results on the performance of the algorithms are given in sections 4 and 5 using ROC and MRROC (Maximum Realisable ROC [35] analyses). In section 6 we discuss the implications of our work and describe a number of directions for further research. Section 7 contains a brief summary and conclusions.

2 Image Processing for the Analysis of Microscope Slides of Malarial Infected Blood

Two kinds of slides may be prepared for microscopic examination of malarial blood samples, "thin" and "thick". A thick slide or smear is just a drop of blood on the microscope slide, but for laboratory work, scientists prefer to make thin slides in which one sees only a single layer of cells in plasma. The samples are usually Giemsa stained so that malarial DNA is visible as a distinctive dark blue colouration. Red blood cells have no DNA and are thus not stained. Currently, we have two sets of data: 100 images in which human red-blood cells are infected with *P. flaciparum* and a second set of 25 more recent images. We also have 17 images of malarial infected mouse blood. All the images are 1300×1030 24-bit colour.

The goal of our work is to develop, implement and evaluate an automatic system for the detection, location, identification, characterisation, and counting of red-blood cells in images of thin slides produced in the laboratory. By identification and characterisation in the above we mean determination of whether a cell is healthy or infected and, if infected, what the type of malarial infection is and what the stage of the parasite lifecycle is. To build such a system and be confident it will meet our requirements, we must be able to evaluate its performance on suitable test data and, in principle, be able reliably to predict its performance on images that may be taken in the future when the system is used in practice.

In order to carry out such work efficiently, to make the best use of the available time of expert malaria researchers and, more importantly, to enable us to carry out the required research and development effectively, software tools are required, for image capture and recording of experimental protocols and the image capture conditions, for annotation and storage of the images in a database, their retrieval, manipulation, processing etc. Such tools will not be discussed here, as building them is nowadays a routine matter, especially for research and laboratory purposes if high-level tools such as Matlab are used or if image processing modules from many of the available extensive libraries are employed.

3 Otsu's algorithm

We noted when first mentioning Otsu's algorithm in the introduction that the procedure itself seems to be more familiar than its relationship to Fisher's discriminant. We therefore briefly review this. Suppose we have a one-dimensional feature space, such as the intensity of the pixels in an image, or the signal in a particular channel of a colour image, denoted for convenience by x. Suppose further that we wish to classify the pixels as belonging to foreground and background (say), which might be the red-blood cells and plasma respectively in the analysis of thin-film slides. According to Fisher, we should then seek to maximise the ratio of the between class variance σ_B^2 to the within class variance σ_W^2 by choosing a threshold, T, i.e.

$$\max_T \left\{ \frac{\sigma_B^2}{\sigma_W^2} \right\}. \quad (1)$$

Since,

$$\sigma_T^2 = \sigma_B^2 + \sigma_W^2, \quad (2)$$

and σ_T^2, the total variance, is constant for a given sample or distribution

$$\sigma_T^2 = \frac{1}{N}\sum_{i=1}^{N}(x(i)-\mu)^2 = \int dx p(x)(x-\mu)^2 \quad (3)$$

with mean μ, equation 1 is equivalent to:

$$\max_T \left\{ \frac{\sigma_B^2}{\sigma_T^2} \right\} \Leftrightarrow \max_T \left\{ \sigma_B^2 \right\}. \quad (4)$$

In practice, the distribution, $p(x)$ is unknown, so we estimate it from the histogram $h(x)$ by assuming, for N samples, that

$$p(x) = \frac{1}{N}h(x). \quad (5)$$

Since, after thresholding, values of x above or below T will be assigned to the foreground or background classes respectively (or vice-versa), it is useful to introduce the indicator density,

$$z(x|I) = p(x) \text{ for } T_{I-1} < x < T_I \quad (6)$$

where $I = 1, 2$ stands for the two classes. T_1 is the threshold T, and T_0 and T_2 are the upper and lower limits of the values of the feature variable, x, often 0 and 255 in image processing applications. Note that we refer to $z(x|I)$ as an indicator density since, after thresholding, values of x in the foreground and background are distinct, whereas the conditional densities of the foreground and background, $p(x|foreground)$ and $p(x|background)$, say, may allow foreground and background values of x to overlap. Such overlap would be allowed in a *mixture model* in which we would set

$$p(x) = \sum_I p(x|I)P(I) \quad (7)$$

with class conditional densities $p(x|I)$ and mixing probabilities $P(I)$. With a thresholding procedure, such overlap is not allowed and is, of course, a source of classification errors.

It is also useful to introduce an indicator variable, $Z(I)$ for each class produced by the thresholding, defined as

$$Z(I) = \int dx z(x|I). \quad (8)$$

$z(x|I)/Z(I)$ then behaves as a probability density. The means of the distributions above and below threshold may then, for example, be represented as

$$\mu(I) = \frac{1}{Z(I)}\int dx z(x|I)x. \quad (9)$$

The variance of class I conditioned on the thresholding, $\sigma(I)^2$ is similarly

$$\sigma(I)^2 = \frac{1}{Z(I)}\int dx z(x|I)(x-\mu(I))^2, \quad (10)$$

and,

$$\sigma_W^2 = \sum_I \int dx z(x|I)(x-\mu(I))^2. \quad (11)$$

Differentiation of equation 4 in order to maximise the between class variance, σ_B^2, then yields

$$\frac{\partial \sigma_B^2}{\partial T} = p(T)[\mu(1)-\mu(2)][2T-(\mu(1)+\mu(2))]. \quad (12)$$

It can be seen from equation 12 that the between class variance and hence, for problems characterised by a single feature variable x, the Fisher discriminant in equation 1 will be extremal if $p(T) = 0$, or $\mu(1) = \mu(2)$, or

$$T = \frac{1}{2}(\mu(1)+\mu(2)). \quad (13)$$

Obviously, if $\mu(1) = \mu(2)$, σ_B^2 is zero and is minimised. Choosing a threshold T at a "gap" in a histogram where $p(T)$ vanishes is a familiar segmentation procedure, but one that can rarely be used rigorously in practice. Of more interest is the third condition expressed by equation 13. Since the means $\mu(1)$ and $\mu(2)$ depend on the choice of T, this is a self consistent equation for T whose iterative solution generates Otsu's algorithm. It is usually initialised by choosing a starting value for the threshold equal to the mean μ and, in practice, converges quickly in a few iterations $(3-4)$. An example is shown in figure 2 applied to the intensity histogram of the whole image. One ninth of the image is displayed in figure 1 together with the resulting segmentation in which foreground pixels above threshold are set to white and those below threshold in the background to black.

3.1 A multi-class generalisation

In reviewing the Otsu algorithm above, we have developed a notation that enables us to generalise to a multi-class problem in which the histogram is to be segemented into C classes by use of several thresholds T_I for $I = 1 \ldots C-1$. The result is that

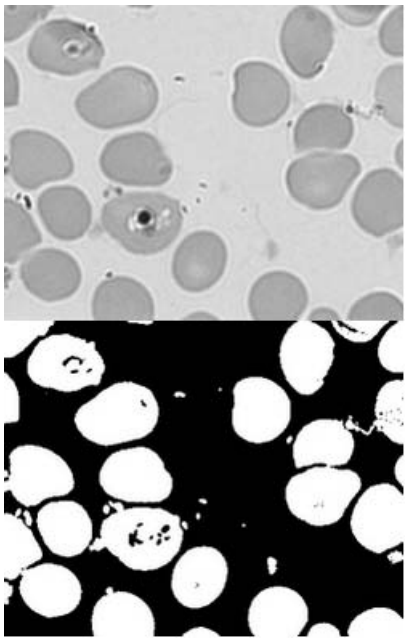

Figure 1. A one-ninth part of an example thin-film slide image shown as grey-level intensity (top), together with the segmentation obtained from the intensity (bottom).

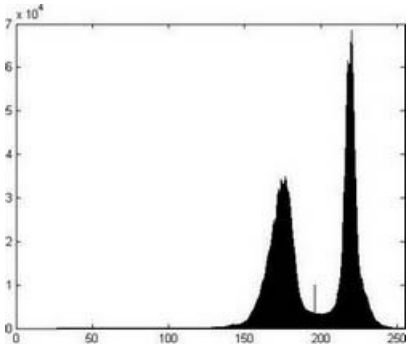

Figure 2. Application of the Otsu algorithm to the intensity histogram of the whole image of figure 1 converges rapidly to a threshold T = 196.

$$\frac{\partial \sigma_B^2}{\partial T_I} = p(T_I)[\mu(I) - \mu(I+1)][2T_I - (\mu(I) + \mu(I+1))] \tag{14}$$

which may be interpreted in a similar way to equation 12. In particular we obtain a multi-class version of Otsu's algorithm for the thresholds T_I from the set of $C-1$ self-consistent conditions

$$T_I = \frac{1}{2}(\mu(I) + \mu(I+1)). \tag{15}$$

Initialisation may be by equally spacing the thresholds over the range $[T_0, T_C]$, where T_C is the upper limit of the values taken by x. Convergence in the multi-class case is much less certain than with a single threshold, though experiments with illustrative distributions indicate it is a useful procedure. In practice, the available histogram data may be smoothed, for example, by using a Gaussian mixture model for curve fitting and a powerful optimisation algorithm such as a particle swarm method used [42]. In fact, it is straightforward to obtain formulae for the Hessian of second derivatives with respect to the thresholds. This opens up the possibility of a variety of implementations, ranging from multi-dimensional Newton-Raphson and Levenberg-Marquardt methods, conjugate gradients, etc., as well as the application of modern stochastic methods such as differential evolution and SOMA (Self Organising Migratory Algorithm) [37, 43]. In addition, it is easy to envisage the application of continuation methods and, given the close similarity to mixture models, development of EM-like algorithms.

This is an interesting generalisation but, it must be emphasized, one that is restricted to problems where there is a one-dimensional feature space x. Higher dimensional problems are much more complicated and perhaps for this reason not discussed much in standard texts [8].

4 Application to thin film slide images

We have applied the Otsu thresholding procedure to the analysis of images of thin-film microscope slides for the development of automated procedures in the study of malaria and, possibly in the longer term, for the diagnosis and treatment of the disease. As indicated in section 2 our first aim has been segmentation of images such as that shown in part in figure 1 to identify the pixels belonging to the red-blood cells and to label, locate and count the red blood cells. This is an image processing problem that has been the subject of several recent works, most notably the series of papers by Di Ruberto *et al* [32, 31, 33, 34] who employed a variety of thresholding and morphological techniques. In their later papers, Di Ruberto *et al* address the problem of parasite detection and characterisation, a problem also studied recently by [30, 38, 36]. We have yet systematically to study such problems, but the extension of the Otsu algorithm to be described in section 5 is promising.

4.1 Thresholding single colour channels

We began with independent thresholding of the red, green and blue channels. The images, having been taken

by clinical researchers who optimised the microscope and illumination settings for human inspection of the slides were not pre-processed or standardised, for example, by histogram equalisation. Typical results obtained with the thresholds produced by the application of the Otsu algorithm are shown in figure 3. It is notable that there is signal in each colour channel, so use of a colour transformation, such as the HSV representation used in [17] may be useful.

ROC curves were calculated from comparison of the results obtained when the threshold was systematically varied from a low to a high value with a hand-segmented ground truth. A typical ROC resulting from processing the green channel is shown in figure 4, together with that obtained from the earlier, intensity-based segmentation. The foreground (red-blood cells) is regarded as the positive class. The hand segmentation had to be carried out with great care as not only the red-blood cells are visible, but also the cell membranes (see figure 1), and the latter excluded from the foreground. This was extremely time consuming, so it was carried out on one complete 1030×1300 image and on one ninth of each of eight other similar images divided into 3×3 arrays of sub-images. The latter were used to provide a means of assessing performance across the first set of 100 images which vary considerably in their appearance (for example the mean colour varied from $(125, 124, 153)$ to $(223, 212, 236)$ and in the density of red-blood cells (from 109 to 250 per image). The image which was annotated in toto was similarly divided into nine sub-images to provide a means of assessing variation within an image. The 17 ROC curves produced from the 17 image patches are plotted superimposed in figure 4.

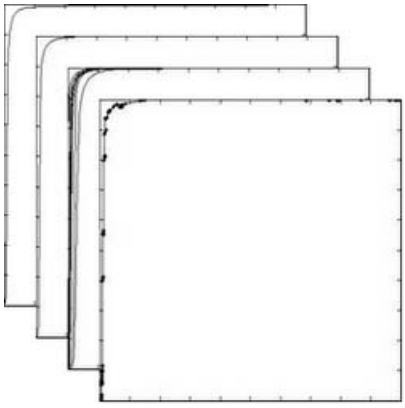

Figure 4. A typical ROC curve obtained from the intensity of of the sub-image shown in figure 1 (back), and from the green channel (second). The 17 ROC curves obtained as described in the text are shown (third). The MRROC (section 4.2) is at the front.

The areas under the ROC curves, A, obtained from the intensity and from the red, green and blue channels of the whole image are given in table 1. The lower area under the ROC indicates the red channel is slightly less discriminating than the green or blue. Close inspection of the individual ROC curves (not shown) reveals that the green channel produces a slightly better true positive rate TP than the blue. The operating points produced by the Otsu algorithm also vary. That in the green channel gives the highest TP, the red channel gives a fairly good balance of the two types of error, whilst the blue channel gives a very low false positive rate FP but its error rate $\epsilon = 1 - TP + FP$ is the highest and the operating point is further from the ideal of $TP = 1$, $FP = 0$ than those in the other colour channels. Segmentation based on the intensity is competitive with those obtained from the green and blue channels and, at the Otsu operating point gives a slightly better balance of the two types of error.

	I	R	G	B
A	0.9848	0.9760	0.9860	0.9813
TP	0.9604	0.9523	0.9777	0.9294
FP	0.0297	0.0345	0.0450	0.0203
ϵ	0.0693	0.0821	0.0673	0.0909

Table 1. Classification performance obtained by thresholding the intensity, red, green and blue channels denoted I, R, G, B respectively.

4.2 Using all three colour channels

The fact that there is signal in each colour channel indicates that all three should be utilised. One way is to combine the outputs as in a multi-classifier system. Since there are only the three classifiers based on the red, green and blue channels to be combined, the sophistication of general-purpose multi-classifier combination methods [23], such as the genetic programming techniques developed by Langdon and Buxton [25] are not needed. Nor have we implemented the EM approach developed by Warfield *et al* [41] or merging methods more familiar in image processing applications [7]. Experiments indicated that the latter did not work well so, noting that there are only 162 relevant inequivalent partitions of the eight possible outcomes of using independent thresholds on each colour channel, we carried out an exhaustive search for the best combinations that gave the highest value of the Fisher discriminant on each image. The results are summarised by the maximum realisable ROC (MRROC) curve [35] shown in figure 4 and the area under the curve and other parameters given in table 2. The value of TP and FP are those obtained from application of the best classifier combination over the whole image

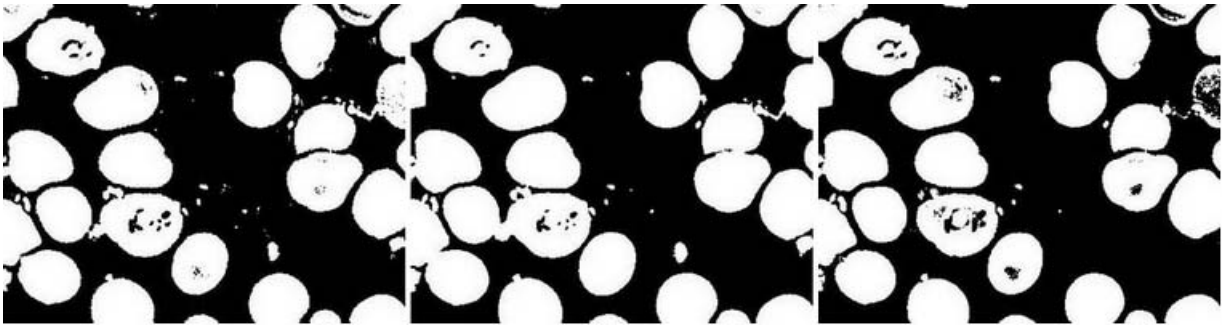

Figure 3. Results obtained by application of the Otsu algorithm to the red (right), green (middle) and blue (left) channels of a thin-film slide image, displayed as in figure 1.

shown in part in figure 1.

Amongst the total of 142 images tested, the most frequent optimal combination was that the red-blood cells were detected in all three colour channels (54 of the 142 images), with detection in at least one channel the second most frequent optimal combination (32 times), and detection in at least two channels the third (22 times).

5 A multi-dimensional extension of Otsu's algorithm

Although the results obtained by thresholding the colour channels singly or in combination are all quite good, in this approach, as in many other multi-classifier systems, the types of classifiers to be combined are defined from the outset. In particular the geometry of the decision surfaces is constrained, in this case for each of them to be normal to the axis of a primary colour. It would seem better to retain as much freedom as possible in choosing the decision surface. This we can do, within the context of a linear model, by utilising the multi-dimensional version of the Fisher discriminant according to which, in feature space $\mathbf{x}$ with between-class covariance S_B and within-class covariance S_W, we seek to

$$\max_A \left\{ \frac{tr(A^T S_B A)}{tr(A^T S_W A)} \right\} \text{ or } \max_A \left\{ \frac{\det |A^T S_B A|}{\det |A^T S_W A|} \right\}. \quad (16)$$

For a two-class problem, when S_B is rank one, A is a vector and solution of either of the above criteria leads to the vector, $\mathbf{a}$ say, which defines the most discriminating direction in the feature space $\mathbf{x}$ [22]. It is given by:

$$\mathbf{a} \propto \mathbf{S_W^{-1}}(\boldsymbol{\mu}(2) - \boldsymbol{\mu}(1)) \quad (17)$$

To locate the optimal position of the decision surface, we also need to know how close to locate it to $\boldsymbol{\mu}(1)$ or $\boldsymbol{\mu}(2)$. This is essentially the one-dimensional problem of finding the optimal thresholding of the histogram of feature vectors projected onto $\mathbf{a}$ which may be solved by use of the Otsu algorithm.

For a two-class problem, we thus have a very simple generalisation of the Otsu algorithm in which we iteratively recompute the most discriminating direction $\mathbf{a}$ and the threshold defining the location of the decision plane. To initialise the algorithm, we carry out PCA to determine the direction of greatest variance in the data and take this as our first approximation to $\mathbf{a}$, project the data onto this direction and apply Otsu's algorithm. The resulting segmentation into two classes is used to construct the covariance matrix S_W for the next iteration and so on.

5.1 Application to colour image segmentation

Application of the above multi-dimensional extension of Otsu's algorithm to segmentation of the red-blood cells in images of thin-film slides is straightforward and leads to very pleasing results such as those illustrated in figure 5. Visualisation of vector $\mathbf{a}$ (computed from the whole image) as iteration proceeds indicates considerable change from the direction of principal variance used to initialise the process. This arises because there are many more pixels belonging to the blood plasma background than to the foreground of the red-blood cells, resulting in the distribution in colour space being heavily weighted towards the top of the pixel distribution and the initial decision surface being similarly displaced towards this point as shown in figure 6. This configuration produces only a quite imperfect segmentation as illustrated in figure 7. As iteration proceeds, the decision surface swings around until it is slicing the distribution of pixel colours more evenly as shown in figure 6. This produces a very good segmentation as depicted in figure 5. Convergence of the direction of $\mathbf{a}$ is quite rapid in $10 - 15$ iterations.

Once the direction of vector $\mathbf{a}$ has settled down, it is easy

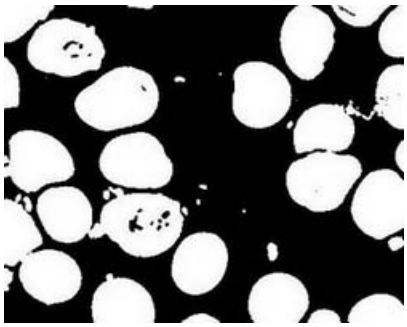

Figure 5. The segmentation produced by applying the 3D Otsu algorithm to the portion of the colour image displayed in monochrome in figure 1.

	3×1D	3D	I
A	0.9858	0.9863	0.9848
TP	0.9574	0.9669	0.9604
FP	0.0324	0.0345	0.0297
ϵ	0.0750	0.0676	0.0693

Table 2. Performance of the system obtained by combination of the separately thresholded colour channels (3×1D) and of the multi-dimensional extension of the Otsu algorithm (3D). Results from the intensity-based segmentation, I, are included for comparison.

to produce an ROC curve for segmentation of the whole image. The area under the ROC curve, operating point, and error rate are given in table 2. The results are very similar to those obtained from combination of the outputs of the separate thresholding on the three colour channels and also similar to those obtained from the image intensity (repeated for ease of comparison) and the best from a single colour channel (green) shown in table 1. The computation time is small (2 secs for processing a 1030×1300 image on a Pentium M 1.8 GHz processor using a Matlab implementation. In contrast, the search for the optimal combination described in section 4 was computationally intensive and took 57 secs on the same machine.

To assess how well the various algorithms cope with variability within and between images, we used the subdivision of one image into nine regions as described in section 4 and similar regions extracted from 8 other images selected at random from the first set of 100. Area under the ROC curves, the operating point selected by the Otsu algorithms, and corresponding error rate were used as performance indicators. There was little variation within the selected image (shown in part in figure 1) that was analysed in detail

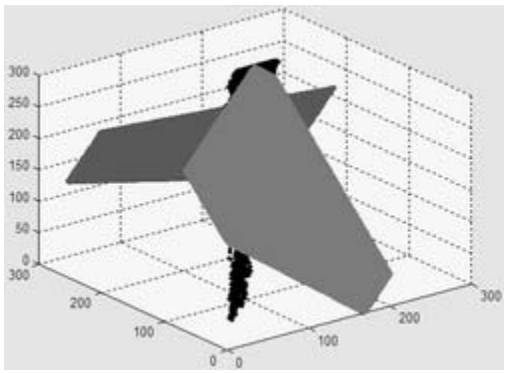

Figure 6. The initial decision surface in 3D obtained from PCA of the pixel distribution in RGB colour space (front) together with the final decision surface obtained by iteration (back) superimposed on the distribution of pixels in the colour space shown by occupancy of the RGB triplets in black.

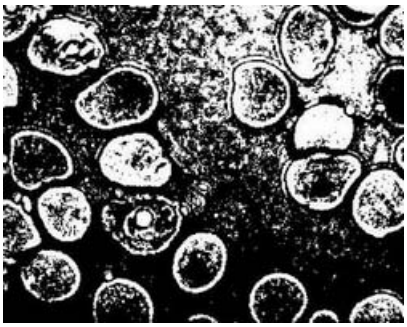

Figure 7. Segmentation obtained for the initial decision surface shown in figure 6.

but, as we might expect from the fact the mean colour of the images varied by almost a factor of two, greater variation between images. Results, averaged over the selected regions from the 9 images are summarized in table 3. On average, performance is a little better with higher values of TP than on the single image analysed in detail. The error rates are a littler lower and the two types of error somewhat better balanced. Evidently the image analysed in detail was somewhat atypical.

5.2 Sequential application to a multi-class problem

We mentioned in section 4.1 that, as can be seen from figure 1, careful inspection of the images reveals that the red-blood cell membrane is visible as a separate region dis-

tinct from the body of the cell and the background plasma. Also, some of the blood cells are infected by malarial parasites and display structures within them where the malarial DNA has been Giemsa stained a dark-blue colour. These structures have a complicated and somewhat variable shape, but to the expert observer indicate the type of infection and stage of the parasite's life-cycle [12]. Di Ruberto *et al* used the characteristic colour and size analysis to detect the malarial infections [31] and later, incorporated a shape classification step [33]. Tek *et al* [38] also worked on the detection of malarial parasites using a Bayes classifier to detect stained pixels and a distance weighted K-nearest neighbour classifier to detect the parasites. Ross *et al* similarly used a neural network classifier system [30].

We have yet systematically to study the detection and characterisation of malarial parasites, but present here a preliminary observation illustrating that sequential application of the multi-dimensional extension of the Otsu algorithm described above can produce interesting results. Having used the multi-dimensional extension of the Otsu algorithm to produce a colour image segmentation, we re-ran the process once on the background class and once on the foreground class and then once again on each of the two segmentation classes obtained from the foreground class to produce six final segmentation classes. For some images, merging the minority class from the segmentation of the background with one of the four leaf classes obtained from the foreground produced a very good segmentation of the red-blood cell membrane (figure 8). The remaining three leaf classes represent healthy red-blood cells and cells infected with malarial parasites, plus a small number of artefacts. However, such a pleasing result was not always obtained and seems to be sensitive to the image colour and contrast. For example, carrying out this procedure on the image shown in part in figure 1 produced the rather mixed and difficult to interpret results shown also in figure 8. It is also somewhat disappointing, but consistent with earlier work in particular that of Di Ruberto *et al* [31, 33], that the parasites cannot easily be detected from the stain colouration alone. Further work is required and it seems essential to standardize the images as discussed in section 6 below.

6 Discussion

This paper describes work in progress. Thus we are unable either to present a fully working, tested and evaluated system for the analysis of thin film microscope slides for the determination of malarial parasitemia, or to give a theoretical analysis of the multi-dimensional extension to Otsu's algorithm. For example, we cannot be sure it converges, though it appears to do so in practice, and to do so quickly.

Of greater interest is an understanding as to why the technique works well and how its performance may be optimised. The former raises the question as to why the Fisher discriminant should be so effective, for example, for segmentation of red-blood cells when applied in the RGB colour space of images of thin film slides. There are two ways this might be addressed and the system's performance improved. One is by transforming the images to some other representation (HSV has been used by others as noted in section 4.1) and by standardising the images. The latter is most easily carried out by means of a Procrustes colour image alignment. This could utilise one of the several possible linear or affine transformations of the colour space. Non-linear transformations such as histogram equalisation and chromatic, grey-world and comprehensive colour normalisation [10] may also be important. Such non-linear transformations are particularly important because they change the way Fisher criterion affects the resulting segmentation. Such transformations are used by Trotter and Buxton [39] in their application of the Otsu algorithm to the analysis of bioinformatics microarray data.

Another way that the reasons for the system's good performance might be understood and improved is by studying the extent to which the classification is linearly separable. Linear systems have the advantage of low capacity and therefore tend to be robust and not over-sensitive to noise or uninteresting details, but the geometry of the decision surface is constrained. A way of generalising the decision surface would be to introduce kernel functions and use KDA, a non-linear extension of linear discriminant analysis or LDA [26]. It must be noted however, that transformation of the colour space and use of kernel functions are not independent techniques. Both are ways of tackling the general problem of non-linear feature selection and optimisation.

It is clear from the above that much remains to be done. A particular strength of the approach is that it is a data-driven, unsupervised technique and that it does not assume any particular structure of the image objects. Whilst this is very attractive as the red-blood cells do not assume perfect geometrical shapes, such as circles or ellipses, and the parasite stains exhibit a complicated and somewhat variable shape, it is also a weakness as classifiers that were able correctly to take into account spatial structure in the images would undoubtedly perform better. It is possible that flexible shape and appearance models [3] could therefore be used to good effect and a comparative evaluation using such techniques, in particular, to model the red-blood cells should be carried out. The extent to which the parasite shape can be modelled remains unknown and we note that it is better to have no model than to use an incorrect or inadequate model.

Finally, we note that the algorithms and systems developed must be thoroughly tested and evaluated. The techniques required are mostly in place, including ROC and MRROC analyses and classifier combination techniques as

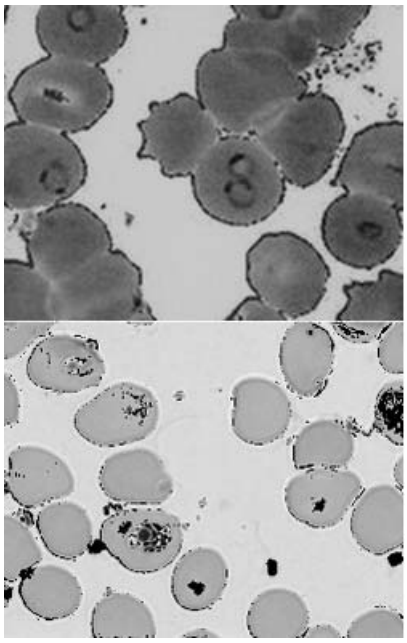

Figure 8. Segmentation of the red-blood cells, cell boundaries and background plasma obtained by sequential application of the 3D Otsu algorithm, shown superimposed on the grey-level image (top). A similar procedure on the image shown in part in figure 1 produced a rather different result (bottom).

described above, though we note that ROC methods must be extended for application to multi-class problems [16]. We have begun to build a database of annotated thin-film slide images of malarial infected blood for this purpose, but it is time consuming and labour intensive. For a world-wide problem as important as malaria, it would be beneficial if a centralised, freely available database were prepared.

	R	G	B	3×1D	3D
A	0.9880	0.9917	0.9929	0.9928	0.9930
TP	0.9700	0.9808	0.9740	0.9768	0.9777
FP	0.0345	0.0364	0.0270	0.0327	0.0274
ϵ	0.0646	0.0556	0.0530	0.0560	0.0497

Table 3. Performance averaged over regions selected from nine images

7 Conclusions

The Otsu algorithm for thresholding a histogram was briefly reviewed as an unsupervised machine learning procedure, with emphasis on its relationship to Fisher's discriminant and its application in image processing and computer vision. The latter was illustrated by using the algorithm for the identification of pixels belonging to the red-blood cells in images of thin-film microscope slides of laboratory samples of malarial infected blood. Segmentation of a set of 142 1300×1030 24-bit RGB images was carried out using the intensity , by independently thresholding each of the colour channels, and by selecting the combination of the results of this thresholding which maximized the Fisher discriminant. ROC and MRROC curves were used to assess performance together with the true-positive, false positive and error rates at the operating points obtained from the algorithms.

A multi-dimensional extension of the Otsu algorithm was proposed and applied to the same set of thin-film colour images and performance analysed in a similar manner. All algorithms performed well with average error rates of less than $\approx 6\%$, with the multi-dimensional algorithms just superior to the best of the results obtained on a single channel (blue). Sequential application of the multi-dimensional extension of the Otsu algorithm produced interesting results which may enable a more detailed segmentation to be carried out with identification of the red-blood cells and of the membranes separating them from the blood plasma. Further work is required to verify this and to explore ways of segmenting the malarial parasites, identifying the type of infection and stage of the parasite life-cycle.

Though there was little variation in performance of the Otsu algorithm across the set of images investigated, a number of ways of optimizing the colour representation, standardizing the images, and of improving performance by using more powerful discriminant criteria were briefly discussed.

References

[1] A. Barla, F. Odone, and A. Verri. Old fashioned state-of-the-art image classification. *ICIAP*, 00:566–571, 2003.

[2] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[3] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE PAMI*, 23(6):681–685, 2001.

[4] T. F. Cootes, G. J. Page, C. B. Jackson, and C. J. Taylor. Statistical grey-level models for object location and identification. In *British Machine Vision Conference*, pages 533–542, 1995.

[5] T. F. Cootes, G. V. Wheeler, K. N. Walker, and C. J. Taylor. View-based active appearance models. *Image and Vision Computing*, 20:657–664, 2002.

[6] N. P. Costen, T. F. Cootes, G. J. Edwards, and C. J. Taylor. Automatic extraction of the face identity-subspace. *Image and Vision Computing*, 20:319–329, 2002.

[7] Y. Du, C.-I. Chang, and P. D. Thouin. Unsupervised approach to color video thresholding. *Optical Engineering*, 43(2):282–289, February 2004.

[8] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification Second Edition*. John Wiley & Sons, Inc., 605 Third Avenue, New York, 10158-0012, 2001.

[9] J. M. Ferryman, A. D. Worrall, G. D. Sullivan, and K. D. Baker. A generic deformable model for vehicle recognition. In *6th British Machine Vision Conference*, pages 127–136, September 1995.

[10] G. D. Finlayson, B. Schiele, and J. Crowley. Comprehensive colour image normalisation. In H. Burkhardt and B. Neumann, editors, *5th European Conference on Computer Vision*, Lecture Notes in Computer Science, 1998.

[11] R. A. Fisher. The statistical utilization of multiple measurements. *Annals of Eugenics*, 8:376–386, 1938.

[12] N. C. for Infectious Diseases. Division of Parasitic Diseases.

[13] D. A. Forsyth and J. Ponce. *Computer Vision A Modern Approach*. Prentice Hall, Upper Saddle River, 2003.

[14] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addoson-Wesley, Reading Massachusetts, 1992.

[15] U. Grenander and A. Srivastava. Probability models for clutter in natural images. *IEEE PAMI*, 23(4):424–429, April 2001.

[16] D. J. Hand and R. J. Till. A simple generalisation of the area under the roc curve for multiple class classification problems. *MACH LEARN*, 45:171–186, 2001.

[17] H. Hengen, S. Spoor, and M. Pandit. Analysis of blood and bone marrow smears using digital image processing techniques. *Proceedings of SPIE*, 4684:624–635, 2002.

[18] A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[19] R. Jain, R. Kasturi, and B. G. Schunck. *Machine Vision*. NcGraw-Hill, New York, 1995.

[20] I. T. Jolliffe. *Principal Component Analysis Second edition*. Springer, New York, 2002.

[21] J.Sullivan, A. Blake, M.Isard, and J.MacCormick. Bayesian object localisation in images. *Int. J. Computer Vision*, 44(2):111–136, 2001.

[22] J. Kittler. Feature selection and extraction. *Handbook of Pattern Recognition and Image Processing*, pages 60–83, 1986.

[23] J. Kittler and F. Roli. Multiple classifier systems. volume 1857 of *Lecture Notes in Computer Science*. Springer, 2000.

[24] D. Koller, K. Daniilidis, and H. H. Nagel. Model-based object tracking in monocular image sequences of road traffic scenes. *International Journal of Computer Vision*, pages 257–281, 1993.

[25] W. B. Langdon and B. F. Buxton. Genetic programming for improved receiver operating characteristics. In J. Kittler and F. Roli, editors, *2nd International Conference on Multiple Classifier System*, volume 2096 of *Lecture Notes in Computer Science*, pages 68–77. Springer, July 2001.

[26] Y. Li, S. Gong, and H. Liddell. Recognising trajectories of facial identities using kernel discriminant analysis. In *British Machine Vision Conference*, pages 613–622, 2001.

[27] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.

[28] M. Pontil and A. Verri. Support vector machines for 3-d object recognition. *IEEE Trans. Patt. Anal. Machine Intell.*, 20:637–646, 1998.

[29] W. K. Pratt. *Digital Image Processing Second edition*. John Wiley & Sons, Inc., New York, 1991.

[30] N. E. Ross, C. J. Pritchard, D. M. Rubin, and A. G. Duse. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Med Biol Eng Comput*, 44:427–436, 2006.

[31] C. D. Ruberto, A. Dempster, S. Khan, and B. Jarra. Automatic thresholding of infected blood images using granulometry and regional extrema. *ICPR*, pages 445–448, 2000.

[32] C. D. Ruberto, A. Dempster, S. Khan, and B. Jarra. Segmentation of blood images using morphological operators. *ICPR*, pages 401–404, 2000.

[33] C. D. Ruberto, A. Dempster, S. Khan, and B. Jarra. Morphological image processing for evaluating malaria disease. *IWVF*, pages 739–748, 2001.

[34] C. D. Ruberto, A. Dempster, S. Khan, and B. Jarra. Analysis of infected blood cell images using morphological operators. *Image and Vision Computing*, 20(2):133–146, February 2002.

[35] M. J. J. Scott, M. Niranjan, and R. W. Prager. Realisable classifiers: Improving operating performance on variable cost problems. In *British Machine Vision Conference*, pages 306–315, September 1998.

[36] S. W. S. Sio, W. Sun, S. Kumar, W. Z. Bin, S. S. Tan, S. H. Ong, H. Kikuchi, Y. Oshima, and K. S. W. Tan. Malariacount: An image analysis-based program for the accurate determination of parasitemia. *Journal of Microbiological Methods*, (July), 2006.

[37] R. Storn and K. Price. Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. Technical report, International Computer Science Institute, Berkeley, 1995.

[38] F. B. Tek, A. G. Dempster, and I. Kale. Malaria parasite detection in peripheral blood images. In *Medical Image Understanding and Analysis*, 2006.

[39] M. W. B. Trotter and B. F. Buxton. Unsupervised thresholding of affymetrix microarray data. In *ICCTA*, 2007.

[40] E. Trucco and A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, Upper Saddle River, 1998.

[41] S. K. Warfield, K. H. Zou, and W. M. W. III. Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. *IEEE Transactions on Medical Imaging*, 23(7):903–921, July 2004.

[42] E. Zahara, S.-K. S. Fan, and D.-M. Tsai. Optimal multithresholding using a hybrid optimisation approach. *Pattern Recognition Letters*, 26:1082–1095, 2005.

[43] I. Zelinka. Soma - self organizing migrating algorithm. *New Optimization Techniques in Engineering*, pages 167–217, 2004.

[44] V. Zografos and B. F. Buxton. Pose-invariant 3d object recognition using linear combination of 2d views and evolutionary optimisation. In *ICCTA*, 2007.