

Model-based automated detection of mammalian cell colonies

Rok Bernard, Maša Kandušer and Franjo Pernuš

Faculty of Electrical Engineering, University of Ljubljana, Tržaška cesta 25, 1000 Ljubljana, Slovenia

E-mail: rok.bernard@fe.uni-lj.si, masa@svarun.fe.uni-lj.si and franjo.pernus@fe.uni-lj.si

Received 12 June 2001

Published 17 October 2001

Online at stacks.iop.org/PMB/46/3061

Abstract

Manually counting cell colonies, especially those that originate from fibroblast cell lines, is a time-consuming, eye-straining and tedious task in which consistency of counting is difficult to maintain. In this paper we present a novel model-based image segmentation method, which employs prior knowledge about the shape of a colony with the aim to automatically detect isolated, touching and overlapping cell colonies of various sizes and intensities. First, a set of hypothetical model instances is generated by using a robust statistical approach to estimate the model parameters and a novel confidence measure to quantify the difference between a model instance and the underlying image. Second, the model instances matching the individual colonies in the image are selected from the set by a minimum description length principle. The procedure was applied to images of Chinese hamster lung fibroblast cell line DC3F, which forms poorly defined or ‘fuzzy’ colonies. The correlation with manual counting was determined and the cell survival curves obtained by automated and manual counting were compared. The results obtained show that the proposed automatic procedure was capable to correctly identify 91% of cell colonies typical of mammalian cell lines.

1. Introduction

The reproductive potential of cells undergoing chemical or physical treatment is commonly assessed by the cell colony formation method. Despite new methods being devised for viability measurements in mammalian cell systems, a colony-forming ability test is still the most consistent, relevant and reproducible, especially for cytotoxic studies (Cook and Mitchell 1989). First, a certain number of cells is plated onto petri dishes where each cell that remains viable after a certain treatment divides and eventually gives rise to a colony of daughter cells; second, the cells are fixed and stained; third, cell colonies are counted; and finally, the relative

survival rate is calculated from the number of colonies obtained from treated and untreated cells (Cook and Mitchell 1989). High cell numbers are required to achieve acceptable statistical accuracy. Counting cell colonies manually is a time-consuming, eye-straining and tedious task in which consistent objectivity is hard to achieve. Computer vision methods, which in many applications have proven to be accurate, reliable, robust and fast, may overcome the deficiencies of manual counting. The crucial part of automated detection of cell colonies is image segmentation, by which individual colonies are identified in an image, regardless of their position, size and intensity. The proper separation of touching and overlapping colonies, often forming larger clusters, is an especially demanding task. Although significant inter-operator differences have to be taken into account, a colony detection algorithm is usually judged by how well it emulates a well-trained, well-rested and motivated human operator.

Various methods have been proposed in the past to segment images of cell colonies. Under optimal illumination and material preparation, a careful selection of one or more thresholds, usually assisted by an operator, may extract the colonies from the background (Corkidi *et al* 1998). Because clustered colonies cannot be separated by simple thresholding, a method using thresholding followed by a distance transform was proposed (Mukherjee *et al* 1995). Region-based methods, e.g. the watershed algorithm, may also separate touching and overlapping colonies, but require manual interaction (Malpica *et al* 1997). An operator must enter a unique seed for each cell colony, which is a task similar to manual counting. An automated counter being able to identify less-discrete cell colonies typical of fibroblast cell lines was recently proposed by Barber *et al* (2001). The method uses the edge information of the image and a compact Hough transform. This model-based segmentation method, which employs prior knowledge about the object's shape, may fail when cell colonies are small or of low contrast. Other model-based methods recognize objects in an image as instances from a database of models or as model instances generated on-line. An example of the former approach, similar to the well-known template matching, is the method of Noordmans and Smeulders (1998). They approximate the shape of a spot or colony by the Gaussian model and determine the parameters of the model in two phases. In the detection phase, models with different parameters are matched to the underlying image and the model that produces a minimal match error is retained. The match is performed at each image position and only those models, which match their underlying images better than the models in the neighbourhood, are kept. In the subsequent characterization phase, the parameter vector of each detected colony model is refined to reduce the match error.

In this paper we present a novel model-based method for automated detection of cell colonies, which improves on the approach by Noordmans and Smeulders (1998). To avoid exhaustive testing of numerous parameter instantiations, a robust statistical approach is used to estimate the model parameters. To efficiently quantify the difference between a model instance and the underlying image a novel confidence measure is applied. The identification of touching and overlapping colonies is improved by employing a global optimization criterion, based on the minimum description length (MDL) principle (Rissanen 1983). Experiments were performed, which showed that the proposed method is capable of correctly identifying isolated, touching and overlapping colonies of various sizes and intensities which are typical of fibroblast cell lines.

2. Material and methods

2.1. Petri dish preparation

2.1.1. Cell cultures. Transformed Chinese hamster lung fibroblast cells, DC3F (Biedler and Riehm 1970) were grown in Eagle's minimum essential medium with Earle's salts

(Sigma–Aldrich, UK) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, sodium bicarbonate, 200 000 IE benzil penicylin and 16 mg gentamicin sulfate. The cell cultures were grown in an incubator at 37 °C in a humidified atmosphere that contained 5% CO₂. From the confluent culture, cell suspension was prepared with 0.05% trypsin solution in Hank's salt solution that contained 0.02% EDTA.

2.1.2. Experiments

- (a) *Preliminary experiment.* To test the ability of the automated colony detection procedure to accurately detect isolated, touching or overlapping cell colonies of various sizes and intensities, untreated cells were plated in 17 round petri dishes, dimensions (diameter × height) 60 × 15 mm (Corling, USA). To study the ability of the method to properly treat overlapping colonies, high proportions of overlapping colonies were obtained by plating 500 cells in each petri dish. The colonies were grown for 5 days in an incubator under the conditions described above.
- (b) *Electroporation.* Cell survival was determined after the exposure of the cell suspension to short intense electric pulses of different amplitudes, a treatment known as electroporation (Kotnik *et al* 2000). For this purpose a cell suspension was prepared in SMEM culture medium (Life Technologies, USA), which is a Spiner modification of Eagle's minimum essential medium that does not contain calcium. A 50 μl drop of cell suspension that contained 10⁶ cells was placed between stainless steel electrodes spaced 2 mm apart. The train of eight pulses, duration of 100 μs and repetition frequency 1 Hz was employed. The amplitudes of applied pulses were 0, 80, 120, 160, 200, 240, 280, 320, 360 and 400 V. Electroporated cells were incubated at room temperature for 30 min to allow the cell membrane to reseal.
- (c) *Surfactant C₁₂E₈.* To study the effect of non-cytotoxic concentration of surfactant C₁₂E₈ on electroporated cells, C₁₂E₈ was added immediately after electroporation. Cells which had been only electroporated and cells which had been electroporated and treated with C₁₂E₈ were plated in round petri dishes of 60 × 15 mm dimensions (Corling, USA) at a concentration of 200 cells per petri dish. A smaller number of cells was plated to reduce the time of manual counting. The colonies were grown for 5 days in the incubator under the same conditions described above.

2.1.3. *Fixation and staining.* In all experiments the growth medium was removed and colonies were fixed with methanol (Merck, Germany) for 10 min and stained with an aqueous solution of crystal violet (Sigma–Aldrich, UK) in concentration of 5 mg ml⁻¹ for 10 min. The stain was removed and the petri dishes were rinsed with water.

2.1.4. *Cell survival.* Cell survival of only electroporated cells was defined as the number of cells that survived electroporation with a specific amplitude divided by the number of surviving cells not undergoing electroporation (0 V). Survival of electroporated cells, to which C₁₂E₈ had been added, was determined as the number of cells that survived electroporation and treatment with C₁₂E₈ divided by the number of cells not undergoing electroporation but treated with C₁₂E₈. Three repetitions were performed for electroporated as well as electroporated and C₁₂E₈-treated cells, yielding altogether 60 petri dishes. To determine the correlation between automatic and manual count, the cell colonies were counted manually and automatically.

2.2. Hardware

The cell colony detection system was composed of a custom-made homogeneous diffuse illumination unit to which a petri dish holder was attached (figure 1). A 2/3" monochrome

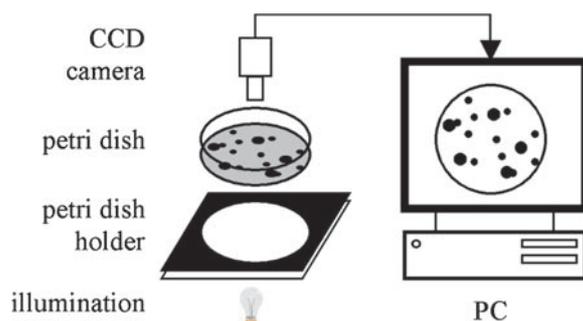


Figure 1. Cell colony detection system.

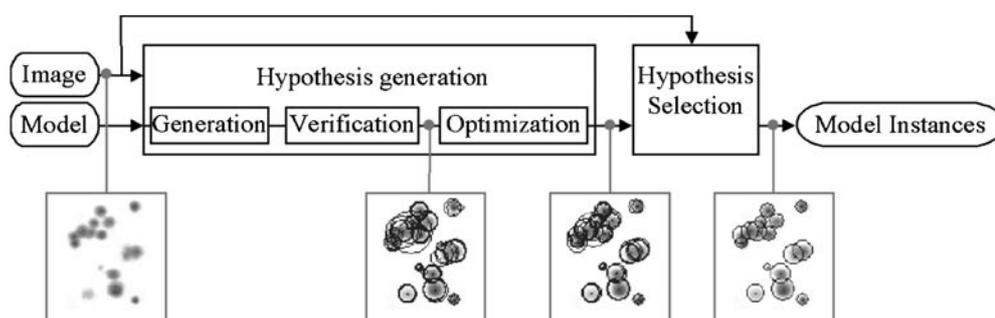


Figure 2. Cell colony detection flow chart.

CCD camera (Sony XC-77) with C-mount lens ($f = 50$ mm) was installed above the petri dish and connected to a Pentium III based PC via a frame grabber. The spatial resolution of the obtained eight-bit images was 512×512 pixels.

2.3. Software

The image processing software, implemented in C++, masks the petri dish edge before counting cell colonies. As the petri dish position is fixed, a simple image masking with a circular region of interest is applied to each petri dish image. Colonies lying completely under the mask are discarded from further image analysis, while the colonies that are only partially covered by the mask are further analysed taking into account only their unmasked part. The flow chart of the cell colony detection method is presented in figure 2. Only small portions (90×100 pixels) of processed images of cell colonies are shown in order to illustrate the main steps of the procedure. The method is based on a hypothesis generation–selection principle (Leonardis *et al* 1995). Prior knowledge about the shape of a colony is incorporated into a parametric model. In the hypothesis generation step, a model instance is generated at every image position, describing a possible colony in the underlying image. Instead of exhaustively trying out all possible model amplitudes and sizes, a robust statistical approach is used to estimate these parameters from the underlying image. Next, a confidence measure is used to verify the obtained model instances. The model instances that pass the verification process form a set of possible cell models or hypotheses, containing at least one model instance for a detectable cell colony. Finally, the position, amplitude and size of the model instances in the set are optimized to better match the underlying image. In the hypothesis selection step, the

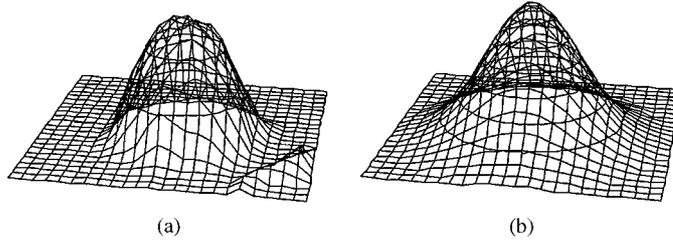


Figure 3. (a) Wireframe representation of a cell colony and (b) corresponding Gaussian model instance $M(\mathbf{x}, \mathbf{p}_i)$.

minimum description length (MDL) principle (Rissanen 1983) is employed to select a unique model instance for each colony. Each step in this process is outlined in more detail in the following subsections.

2.3.1. Cell colony model. A cell colony model incorporates prior knowledge about the shape of a colony, which is the result of both the mechanism that drives the formation of a colony and the transfer function of the imaging system. The shape can be well approximated by the parametric Gaussian model $M(\mathbf{x}, \mathbf{p}_i)$:

$$M(\mathbf{x}, \mathbf{p}_i) = \begin{cases} a_i \exp(-2^{-1} \|\mathbf{x} - \mathbf{x}_i\|^2 / s_i^2) & \|\mathbf{x} - \mathbf{x}_i\| \leq 3s_i \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{x} = (x, y)$ is an arbitrary point in the image, $\mathbf{p}_i = (\mathbf{x}_i, a_i, s_i)$ is the parameter vector consisting of the model's position $\mathbf{x}_i = (x_i, y_i)$, amplitude (intensity) a_i , and size s_i . A specific example of a model, uniquely defined by \mathbf{p}_i , is named a model instance. Figure 3 shows the wireframe image of a cell colony and its corresponding Gaussian model instance.

2.3.2. Model instance generation. The initial values of the amplitude a_i and size s_i of a model instance positioned at point \mathbf{x}_i are first estimated. The initial amplitude is simply set to the intensity value $I(\mathbf{x}_i)$, while the initial size is defined by using the intensities and intensity gradients of pixels

$$s_i = \begin{cases} 2^{-1/2} \|\mathbf{x}' - \mathbf{x}_i\| & \|\mathbf{x}' - \mathbf{x}_i\| \leq s_{\max} \\ s_{\max} & \text{otherwise} \end{cases}$$

where s_{\max} is the size of the largest expected cell colony and \mathbf{x}' are points equidistant from \mathbf{x}_i , $\|\mathbf{x}' - \mathbf{x}_i\| = \text{constant}$, satisfying

$$\frac{I(\mathbf{x}') \cdot w(\mathbf{x}')}{w(\mathbf{x}')} = \frac{I(\mathbf{x}_i)}{e}.$$

The weight $w(\mathbf{x})$ is proportional to the image gradient magnitude $|\nabla I(\mathbf{x})|$. It decreases with angle $\gamma(\mathbf{x}, \mathbf{p}_i)$, defined by the scalar product between the image gradient $\nabla I(\mathbf{x})$ and model gradient $\nabla M(\mathbf{x}, \mathbf{p}_i)$, normalized by the parameter γ_g

$$w(\mathbf{x}) = |\nabla I(\mathbf{x})| \exp(-2^{-1} \gamma^2(\mathbf{x}, \mathbf{p}_i) / \gamma_g^2)$$

$$\gamma(\mathbf{x}, \mathbf{p}_i) = \arccos \frac{\nabla I(\mathbf{x}) \cdot \nabla M(\mathbf{x}, \mathbf{p}_i)}{|\nabla I(\mathbf{x})| |\nabla M(\mathbf{x}, \mathbf{p}_i)|}.$$

In the case of the Gaussian model, $\nabla M(\mathbf{x}, \mathbf{p}_i)$ depends only on \mathbf{x}_i , $\nabla M(\mathbf{x}, \mathbf{p}_i) \propto \mathbf{x} - \mathbf{x}_i$.

From the initial model parameters, a better estimation of a_i and s_i is obtained by a hard redescender M-estimator (Mirza and Boyer 1993), which is robust against Gaussian noise and outliers introduced by touching and overlapping colonies. The objective function of the M-estimator is Tukey's biweight function (Mirza and Boyer 1993, Meer *et al* 1991, Zhang 1998) of the difference between a model instance and the underlying image. A downhill simplex method (Press *et al* 1992) is employed for its minimization. The scale of the M-estimator is calculated as suggested by Rousseeuw and Leroy (1987).

2.3.3. Model instance verification. To eliminate model instances matching non-circular structures, we use a confidence measure $o(x_i)$, which quantitatively expresses the difference between a model instance and the underlying image. The measure is based on the assumption that gradient directions at geometrically corresponding points of the image and model instance are equal only if (a) the centre of the model instance corresponds to the centre of the colony and (b) both colony and model are circular with monotonically increasing or decreasing intensities. The confidence measure $o(x_i)$ is

$$o(x_i) = \frac{1}{N} \sum_{\substack{x \text{ where} \\ M(x, p_i) > 0}} \exp(-2^{-1} \gamma^2(x, p_i) / \gamma_g^2)$$

where N is the number of points under the model, i.e. points satisfying condition $M(x, p_i) > 0$.

When a model instance and the underlying image perfectly match, the value of the confidence measure $o(x_i)$ will be 1. In practice, lower values than 1 are obtained because (a) the colony under the model instance is not perfectly circular, noise is present in the image and there are numerical gradient calculation errors, and (b) the model instance covers more than one colony.

After recovering a model instance at each pixel and obtaining its confidence measure, a verification test is performed to eliminate model instances, which poorly match the underlying image. First, model instances for which $a_i < a_{\min}$ and $s_i < s_{\min}$, where a_{\min} and s_{\min} are pre-selected values, are eliminated. Next, a non-maxima criterion is used to eliminate the model instances whose confidence measure does not represent a local maximum in the two-dimensional function $o(x)$. The model instances that pass the verification process form a set of hypotheses containing at least one model instance for a colony.

2.3.4. Model instance optimization. To achieve a better segmentation accuracy, the parameters of model instances in the set of hypotheses are optimized. The same optimization routine is used as in section 2.3.2, with the exception that now besides amplitude and size the model's position is also optimized using the values estimated in the generation step as an initial guess. Figure 4 is a simple one-dimensional illustration of the hypothesis generation part of the algorithm. Suppose that two functions are given (figures 4(a), (d)) and we would like to obtain the Gaussian model instances at points x_1 and x_2 . The dotted functions in figures 4(b), (c) and 4(e), (f) show the initial model instances obtained in the model generation step, while the dark solid functions in these images show the model instances after the optimization step. This example shows that as long as we are close to an object's centre, the object will be correctly modelled even if it was only roughly initialized.

2.3.5. Hypothesis selection. To select a unique model instance for a colony in the image the minimum description length (MDL) principle was employed (Rissanen 1983, Leonardis *et al* 1995). Model instances are selected sequentially, according to their contribution to the criterion function F evaluating their mutual overlap and confidence measure. Let m_i be a

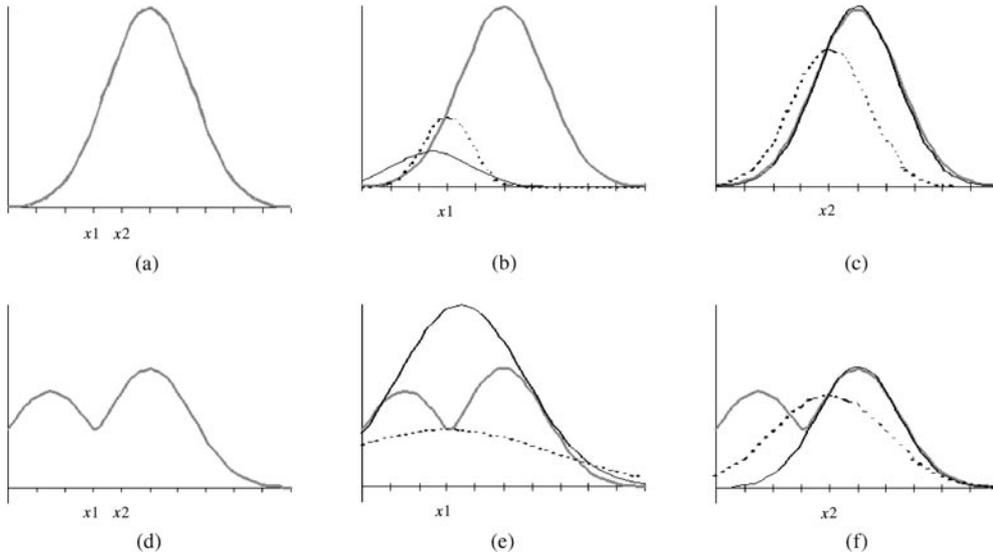


Figure 4. Intensity profile of a Gaussian object (a–c) and two overlapping Gaussian objects (d–f) with initial (dotted functions) and optimized (dark solid functions) model instances at points x_1 and x_2 .

Boolean variable for model instance i . The value of m_i is 1 when a model instance is selected and 0 otherwise. Already selected model instances compose a vector $\hat{\mathbf{m}}$. The contribution $\Delta F(\hat{\mathbf{m}}, m_i)$ of a model instance i to the criterion function F consists of two terms

$$\Delta F(\hat{\mathbf{m}}, m_i) = \left(\frac{o(\mathbf{x}_i) - o_K}{1 - o_K} \right) + \left(1 - \frac{(1 - V_i/V_i^0)}{V_K} \right).$$

The first term denotes the match of the model instance i to the corresponding colony and is based on its confidence measure $o(\mathbf{x}_i)$. The constant o_K is the lowest allowed value for the model instance i to be selected. The second term denotes the volume overlap of the model instance i with previously selected model instances $\hat{\mathbf{m}}$. V_i^0 denotes the whole volume under the model instance i , whereas V_i denotes the volume under the model instance i not occupied by previously selected model instances. The constant V_K is a pre-selected proportion of volume overlap.

3. Results and discussion

Two experiments, carried out on images of DC3F cell colonies, were conducted to evaluate the performance of the proposed automated cell colony detection procedure. First, its ability to accurately detect cell colonies in the images was examined. Second, the correlation with the manually counted colonies was determined and the cell survival curves, obtained by automated and manual counting, were compared. The following image analysis parameters were used in all experiments: $\gamma_g = 30^\circ$, $s_{\max} = 10$, $s_{\min} = 1$, $a_{\min} = 20$, $o_K = 0.25$ and $V_K = 0.7$. The parameter γ_g has a direct influence on the sensitivity and smoothness of the confidence measure $o(\mathbf{x}_i)$. Small values of γ_g ($\gamma_g < 10^\circ$) make $o(\mathbf{x}_i)$ sensitive to partly circular structures in the image. On the other hand, by using large values of γ_g ($\gamma_g > 90^\circ$) some touching and overlapping colonies may not be detected. Any value of γ_g between 20° and 80° can be used, as the selection in this range does not significantly influence the outcome of colony detection.

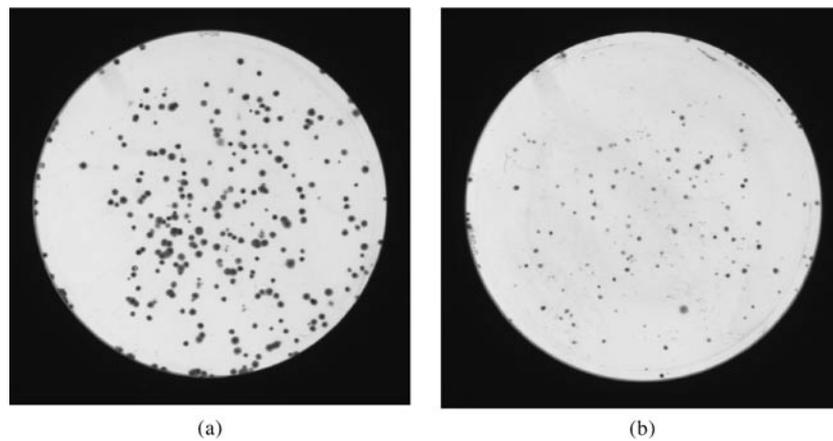


Figure 5. Cell colonies in (a) petri dish number 17 and (b) petri dish number 8.

Parameters s_{\max} and s_{\min} were set to the size of the largest and smallest expected colony in the image, respectively, and a_{\min} was set to the intensity of weakest contrast colony expected in the image. Parameters o_K and V_K depend on maximal expected overlap of colonies in the image. By selecting lower values of V_K and o_K , overlapping colonies and colonies whose shape deviates from the Gaussian shape will be detected.

With the purpose of testing the ability of the automated cell colony detection procedure to accurately detect cell colonies and to study the shape of DC3F cell colonies in detail, images of 17 petri dishes containing from 170 to 330 cell colonies (mean = 267) were analysed. The resulting model instances were manually classified into two classes: the correct model instances, corresponding to actual colonies, and the false positive model instances, corresponding to artefacts present in the image. The major sources of artefacts were: (a) colonies composed of less than 50 cells, (b) colonies with hairy outgrowths, and (c) traces of the growth medium not completely removed from petri dishes before colony fixation and staining. This was due to the established routine practice of material preparation, which paid no special attention to the subsequent automated analysis. Unmasked colonies that had not been detected by the proposed automatic method were manually located. At each manually defined location a model was fitted to the underlying image. These colonies were classified as false negative model instances. The major sources of undetected colonies were (a) isolated, small and faint colonies composed of little more than 50 cells and (b) heavily overlapped colonies in colony clusters. The consistent use of 50 cells per colony as a threshold is hard to meet in practice. Because counting cells in colonies is too labour-intensive, the threshold is applied indirectly via the appearance of a colony. However, because a small change in cell number is only slightly reflected in the colony appearance, mistakes around the threshold are inevitable, regardless of the method applied. Figure 5 shows two typical images of cell colonies in petri dishes illustrating the wide range of cell colony sizes and intensities that may appear.

Figure 6 illustrates the distributions of volume overlap, confidence measure, amplitude and size of all correct, false positive, and false negative model instances. From these distributions we extracted some global characteristics of the DC3F colonies we had analysed. The distribution of volume overlap in figure 6(a) shows that colonies were mostly isolated or overlapped less than 10%. The great majority of heavily overlapping colonies was correctly identified. The colonies and artefacts that gave rise to false negative and false positive model

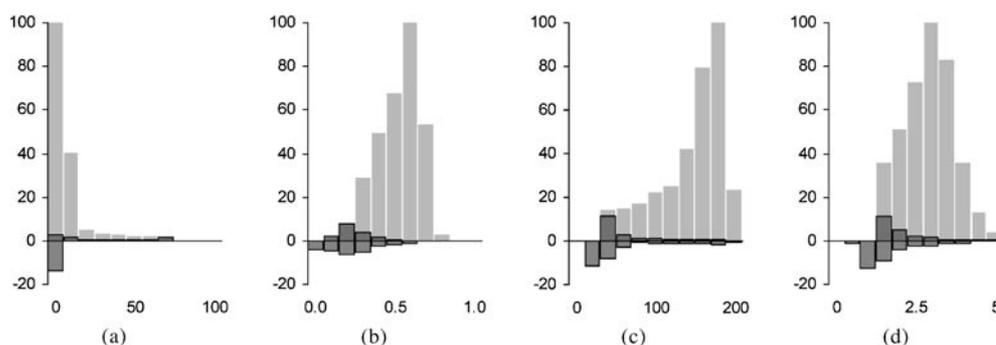


Figure 6. Distributions of over 4500 model instances classified into correct (light grey bars), false negative (dark negative bars) and false positive model instances (dark positive bars): (a) volume overlap (%), (b) confidence measure, (c) amplitude (grey levels), (d) size (pixels).

instances were mostly isolated. The distribution in figure 6(b) shows that the number of false positive and false negative model instances decreases with an increasing value of the confidence measure. All colonies represented by model instances with high values of the confidence measure were correctly detected. Similar to the distribution of the confidence measure, the distributions of amplitudes (figure 6(c)) and sizes (figure 6(d)) show a trend of false positive and false negative model instances decreasing with increasing amplitude and size. As expected, high-contrast colonies and/or colonies large in size were reliably detected. The conclusions we can make from this experiment are that the main source of false negative model instances are small and faint colonies and colonies whose appearance deviates significantly from the Gaussian model. If there were no artefacts in the image, many false negative colonies could be detected by setting lower values of s_{\min} , a_{\min} and o_K . In the presence of the artefacts the reduction of these parameters would lead to more false positive colonies.

Table 1 shows the results for the 17 petri dishes in more detail. For each petri dish it gives the number of manually defined colonies and the rate of correct, false positive and false negative model instances. It also gives the percentage of model instances selected from the sets of hypothetical model instances and the average overall time, as well as the times for the hypothesis generation and selection phases. On average 91.1% of cell colonies were correctly detected. The highest false positive and false negative rates were 9.4 and 18.2%, in petri dish 7 and 8 (figure 5(b)), respectively. The average rate of false positive and false negative was 5.0 and 8.9%, respectively. On an average 55.7% of model instances were selected in the hypothesis selection step. The average time consumption for automated colony counting was 2 min and 35 s per petri dish on a Pentium III based PC. On average 77% of the time was used by the hypothesis generation step and 23% by the hypothesis selection step. The above results show that the detection procedure is capable of identifying touching and overlapping colonies. It is very reliable and accurate in detecting colonies, which match well with the Gaussian shaped model but may fail to correctly discriminate between small, non-contrast, non-circular colonies and artefacts. However, even an experienced operator may have difficulty in counting such cell colonies.

In the second experiment the correlation with the manually counted colonies was determined and the cell survival curves obtained by automated and manual counting, were compared. To estimate the effect of masking the petri dish edges, we manually counted the colonies under the mask and added their number to the number obtained by automated counting.

Table 1. Number of model instances, classified as correct, false positive and false negative for 17 petri dishes.

Petri dish no	Colony no	Correct (%)	False ⁺ (%)	False ⁻ (%)	Selected (%)	Time (min:s)	Generation (%)	Selection (%)
1	254	93.7	4.7	6.3	36	1:56	76	24
2	258	95.0	5.4	5.0	61	1:23	77	23
3	253	90.5	5.9	9.5	57	1:34	79	21
4	225	89.3	5.8	10.7	72	0:52	82	18
5	210	84.8	4.8	15.2	78	0:37	78	22
6	278	92.8	5.0	7.2	71	1:18	80	20
7	330	87.0	9.4	13.0	41	5:14	76	24
8	176	81.8	6.3	18.2	78	0:20	85	15
9	280	92.1	5.0	7.9	48	2:54	78	22
10	312	93.9	6.7	6.1	48	4:29	80	20
11	307	96.7	2.9	3.3	39	5:08	70	30
12	170	81.8	8.2	18.2	76	0:21	80	20
13	316	92.7	2.2	7.3	46	4:21	74	26
14	271	92.6	3.3	7.4	47	4:21	81	19
15	324	93.8	2.5	6.2	50	3:28	71	29
16	325	93.8	2.5	6.2	50	3:12	71	29
17	256	96.1	3.5	3.9	50	2:30	73	27
avg.	267	91.1	5.0	8.9	55.7	2:35	77	23
max	330	96.7	9.4	18.2	78.0	5:14	85	30
min	170	81.8	2.2	3.3	36.0	0:20	70	15

The results of this semi-automated counting were given beside the results of automated and manual counting. Comparison of colony numbers obtained manually and semi-automatically for 60 petri dishes showed an almost direct linear relationship (figure 7(a)). The relationship between manual and automated counts was also linear; however, the slope of regression line deviated from the ideal case (figure 7(b)). The correlation coefficients were 0.995 and 0.994 for semi-automated and automated counting, respectively, and are comparable to the correlation coefficients obtained by Barber *et al* (2001) for cell lines HT29, A172, U118 and IN1265. The linear regression slopes were 0.964 and 0.840 and intercepts were 0.768 and -1.070 for semi-automated and automated counting, respectively. For semi-automated counting no positive or negative offset was thus revealed.

The cell survival curves are given in figures 8(a) and 8(b) for electroporated and electroporated and $C_{12}E_8$ -treated cells, respectively. Cell survivals obtained by the manual, semi-automated and automated counting are given side by side together with the standard deviations of the three repetitions. It is evident that the variability between the semi-automated, automated and manual counting is comparable to the variability between repetitions of treatments. An analysis of variance was used to determine an overall coefficient of variation (CV) of the survival fraction values for the counts made on electroporated as well as on electroporated and $C_{12}E_8$ -treated cells. The CV was based on the normalized variability averaged over voltage and gave an overall figure of merit for an experiment (smaller values indicating less error throughout the experiment). For the electroporated only cells the CVs of manual, semi-automated and automated counting were 11.35, 12.71 and 11.05, respectively. For $C_{12}E_8$ -treated cells CVs were 10.53, 10.09 and 10.61. These results indicate that the proposed method was consistent with manual counting. It is possible that to a small extent this agreement is due to the fact that false positive and false negative counts may cancel each other

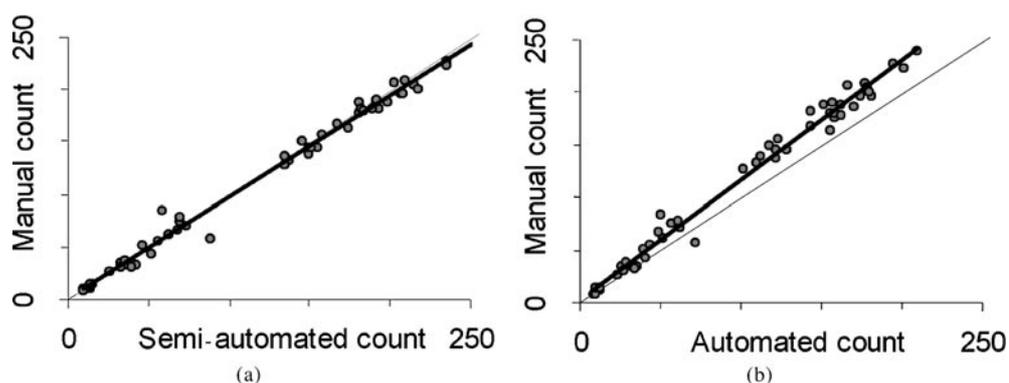


Figure 7. Correlation of (a) manual versus semi-automated and (b) manual versus automated counting. Ideal correlation and regression lines are shown by a thin and a thick line, respectively.

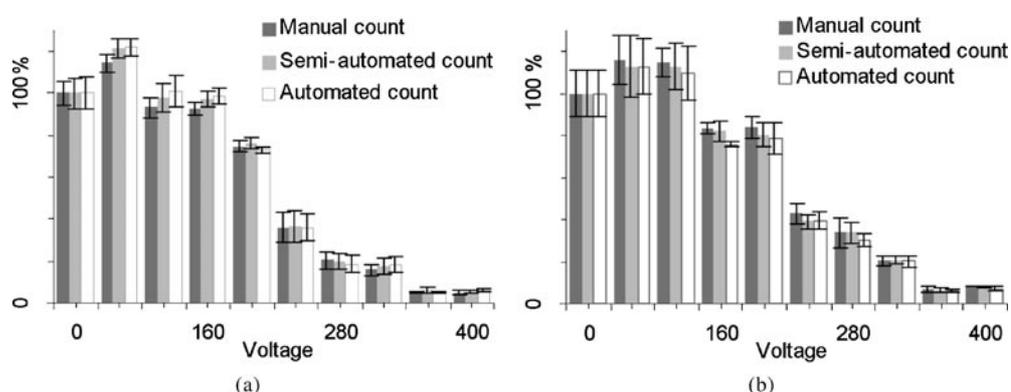


Figure 8. Cell survival curves for (a) electroporated cells and (b) electroporated and $C_{12}E_8$ -treated cells.

out. However, this may also be the case when counting cells manually. The obtained CVs are comparable to the CVs obtained by Barber *et al* (2001) for cell lines U118 and IN1265.

In summary, we have developed an automated model-based method, which we have used to analyse DC3F cell colonies. The counting accuracy is dependent on how well the colonies fit the Gaussian model. This model was selected to describe the radial intensity variation in DC3F colonies after comparing the performances of a linear model and various second-order models in the form of parabola, semi-circle, cosine and Gauss. We have also found that a circular symmetric model is better than the elliptical one when colony clusters were present. The elliptical model tended to fit to the whole colony cluster, while the circular symmetric model managed to fit to individual colonies in a cluster. While more complex models may better approximate the variety of colony appearances, the stability of the fitting process may decline and the computational cost may increase dramatically. Although we have used this method to analyse DC3F cultures, there is no reason why it should not be used to analyse different cell lines. If these lines have a different colony morphology, the Gaussian model should be replaced by the model that best describes the colony morphology.

The proposed automated method gives not only the number, but also the size and intensity of each colony. These parameters are impossible to define manually, although

they might represent valuable additional information on the effects of the treatment. The colony size is important in experiments in which cytostatic but not cytotoxic agents are tested. These agents do not reduce colony number but do reduce colony size (Freshney 1994). The proposed automated method could thus be used to study the effects of cytotoxic and cytostatic agents. Despite its complexity, the proposed method is relatively fast, as it employs optimization to precisely estimate the model parameters. Because model instances are generated independently, the hypothesis generation can be implemented in parallel. In any case, however, the detection time increases with image complexity, i.e. with increasing overlap of colonies in the image, because the selection phase is performed sequentially.

The proposed method gives comparable results to the method of Barber *et al* (2001), which, however, does not give the size and intensity of each colony. The main drawbacks of the proposed method are its inability to process colonies lying close to the petri dish edge and a number of parameters which need to be set manually. For this reason we plan to integrate a model of a petri dish edge into the method. We furthermore plan to develop an interactive training procedure to estimate optimal image processing parameters from a set of training images.

References

- Barber P R, Vojnovic B, Kelly J, Mayes C R, Boulton P, Woodcock M and Joiner M C 2001 Automated counting of mammalian cell colonies *Phys. Med. Biol.* **46** 63–76
- Biedler J L and Riehm H 1970 Cellular resistance to actinomycin D in Chinese hamster cells in vitro: cross-resistance, radioautographic, and cytogenetic studies *Cancer Res.* **30** 1174–84
- Cook J A and Mitchell J B 1989 Viability measurements in mammalian cell systems *Anal. Biochem.* **179** 1–7
- Corkidi G, Diaz-Urbe R, Folch-Mallol J L and Nieto-Sotelo J 1998 COVASIAM: an image analysis method that allows detection of confluent microbial colonies and colonies of various sizes for automated counting *Appl. Environ. Microbiol.* **64** 1400–4
- Freshney R I 1994 *Culture of Animal Cells. A Manual of Basic Technique* (New York: Wiley) pp 161–307
- Kotnik T, Macek-Lebar A, Miklavcic D and Mir L M 2000 Evaluation of cell membrane electroporation by means of a nonpermanent cytotoxic agent *Biotechniques* **28** 921–6
- Leonardis A, Gupta A and Bajcsy R 1995 Segmentation of range images as the search for geometric parametric models *Int. J. Comput. Vis.* **14** 253–77
- Malpica N, de Solorzano C O, Vaquero J J, Santos A, Vallcorba I, Garcia-Sagredo J M and del Pozo F 1997 Applying watershed algorithms to the segmentation of clustered nuclei *Cytometry* **28** 289–97
- Meer P, Mintz D and Rosenfeld A 1991 Robust regression methods for computer vision: A review *Int. J. Comput. Vis.* **6** 59–70
- Mirza M J and Boyer K L 1993 Performance evaluation of a class of M-Estimators for surface parameter estimation in noisy range images *IEEE T. Robot. Autom.* **9** 75–85
- Mukherjee D P, Pal A, Sarma S E and Majumder D D 1995 Bacterial colony counting using distance transform *Int. J. Bio.-Med. Comput.* **38** 131–40
- Noordmans H J and Smeulders A W M 1998 Detection and characterization of isolated and overlapping spots *Comput. Vis. Image Underst.* **70** 23–35
- Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in C, the art of scientific computing* (Cambridge, UK: Cambridge University Press) pp 408–12
- Rissanen J 1983 A universal prior for the integers and estimation by minimum description length *Ann. Stat.* **11** 416–31
- Rousseeuw P and Leroy A M 1987 *Robust regression and outlier detection* (New York: Wiley) p 202
- Zhang Z 1998 Parameter estimation techniques: a tutorial with application to conic fitting *Image Vis. Comput.* **15** 59–76