Bioimaging **4** (1996) 93–106. Printed in the UK

Fluorescent dot counting in interphase cell nuclei

Hans Netten†§, Lucas J van Vliet†, Hans Vrolijk‡, Willem C R Sloos‡, Hans J Tanke‡ and Ian T Young†

† Faculty of Applied Physics, Lorentzweg 1, Delft University of Technology, NL-2628 CJ Delft, The Netherlands
‡ Laboratory for Cytochemistry and Cytometry, Wassenaarseweg 72, University of Leiden, NL-2333 AL Leiden, The Netherlands

Submitted 27 June 1996, accepted 19 September 1996

Abstract. Fluorescence *in situ* hybridization allows the enumeration of chromosomal abnormalities in interphase cell nuclei. This process is called dot counting. To estimate the distribution of chromosomes per cell, a large number of cells have to be analysed, particularly when the frequency of aberrant cells is low. Automation of dot counting is desirable because manual counting is tedious, fatiguing, and time consuming.
We have developed a completely automated fluorescence microscope system that counts fluorescent hybridization dots for one probe in interphase cell nuclei. This system works with two fluorescent dyes—one for the DNA hybridization dots and one for the cell nucleus. A fully automated scanning procedure has been used for the image acquisition. After an image is acquired it has to be analysed in order to find the nuclei and to detect the dots. This article focuses upon the dot detection procedure. Three different algorithms are presented. The problems of 'overlapping' dots and split dots are discussed. The automated dot counter has been tested on a number of normal specimens where DAPI was used for the nucleus counter stain and a centromeric probe was used to mark the chromosome 12. The slides contained lymphocytes from cultured blood. The performance of the different algorithms has been evaluated and compared with manually obtained results. The automated counting results approximate the results of manual counting.

Keywords: fluorescence *in situ* hybridization, chromosome enumeration, dot counting, overlapping dots, image analysis

1. Introduction

Fluorescence *in situ* hybridization (FISH) techniques in interphase cell nuclei have great potential both in research and in clinical applications such as minimal residual disease and early relapse detection in leukemias. FISH has made it possible to selectively stain various DNA sequences in interphase cell nuclei. The sequences may be chosen so as to detect specific abnormalities or to facilitate the process of identification and quantification of numerical and structural chromosomal abnormalities (Eastmond and Pinkel 1989, Hopman *et al* 1991, Nederlof *et al* 1989).

Fluorescence has a high sensitivity and allows the use of multiple colors to detect multiple targets simultaneously. The target signals in interphase cells become visible as colored dots. Enumeration of these signals is called dot counting. Chromosome enumeration requires the analysis of a large number of cells to determine the distribution of chromosomes per cell and to be able to detect small aberrant sub-populations. The number of cells that must be analysed depends on the frequency of aberrant cells and the count accuracy. In practical situations this can vary from only a few cells to more then 10 000 cells (Carothers 1994, Castleman and White 1995, Kibbelaar *et al* 1993). Current manual counting procedures leave much to be desired including the need to work in a dark environment and the fatiguing nature of the work. An automated dot counting system is a practical requirement.

We have developed a completely automated microscope system that counts fluorescent hybridization dots for one probe in interphase cell nuclei. Only two colors can be used—one for the counter stain and one to make the chromosome visible. A complete description is presented in (Netten *et al* 1997). A critical part of the system is

§ To whom correspondence should be addressed. E-mail address: hansn@ph.tn.tudelft.nl

0966-9051/96/020093+14$19.50

the dot detection algorithm. Previous results showed that an average of 11% of the cells are counted incorrectly (Netten *et al* 1997). About 6% of the cells that are counted incorrectly are caused by the dot detection algorithm. Missed dots, false dots, touching dots and split dots have a significant influence on the count accuracy of the system. This paper presents a more detailed discussion about the dot detection algorithm. Different image processing algorithms have been developed and evaluated.

The paper is organized as follows. Section 2 gives an overview of the system and hardware and scanning procedures are described. Section 3 describes the image processing and image analysis procedures used in the dot counter. The problem of overlapping FISH dots in 2-D images of interphase nuclei is discussed in section 4. We have evaluated the dot detection algorithms using a set of 352 images acquired with the automated scanning system. The experimental results are given in section 5. Finally, in section 6, we draw some conclusions and discuss the results.

2. Materials and methods

We have developed a completely automated microscope system that counts FISH dots of a single probe in counter-stained interphase cell nuclei. This section describes the different components that are necessary for the image acquisition. An overview of the instrumentation is given, along with how the actual screening is implemented and what kind of specimens are used to test the system.

2.1. Biological material

The automated dot counter has been tested on a number of slides where DAPI is used for the nucleus counter stain and a centromeric 12 probe (CEP 12, Vysis, Downer's Grove, Il, USA) is used to mark the desired chromosomes. The probe is labeled with Spectrum Orange†. The samples have been provided by Vysis and contain lymphocytes from cultured blood. All samples are normal specimens.

2.2. Screening procedure

The system is built around a Zeiss Axioskop microscope. The microscope is fully automated. Focus, scanning stage, excitation filter wheels and shutter (Ludl Electronic Products Ltd) are controlled by computer. A Macintosh Quadra 840 AV computer controls all hardware and takes care of the image processing. A KAF 1400 Photometrics Series 200 camera (Aikens *et al* 1990) is used to digitize the microscope images. It is a slow scan (500 kHz), cooled ($-42\,^{\circ}$C) CCD camera. The CCD chip contains 1317×1035 pixels with a size of $6.8 \times 6.8\,\mu\text{m}^2$. The performance of the camera in terms of the signal-to-noise ratio (SNR) is excellent. Due to a slow readout rate and cooling, the camera is photon limited. The SNR of a single pixel is limited only by photon statistics. Other noise sources are not significant (Mullikin *et al* 1994). The image SNR is limited by the variation of the gain and offset of each pixel. The maximum image SNR is 30 dB (Netten *et al* 1994). It is possible to correct for the pixel variation using flat field correction. Because flat field correction is a time consuming operation and the SNR is adequate, we have not used this technique.

Screening a slide consists of a number of steps that are repeated until a preset number of nuclei is analysed. This is called the scanning cycle. Each field of view (FOV) is focused automatically after which an image is acquired. When the acquisition is finished, the image is processed and the stage is moved to the next FOV.

The auto-focusing algorithm is based on a derivative filter (Boddeke *et al* 1994). A digital filter [1,0,−1] is applied in the x direction of the image. The 'energy' in the resulting image is computed. This is called the focus function and it has to be maximized with respect to the z position for the image to be in focus. A sequence of images is acquired at different z positions to find that maximum.

We have used a $40\times/1.3$ objective (Plan NEOFLUAR, Zeiss) in combination with 2×2 binning for the acquisition of images after the focusing process. Binning is the process of combining adjacent pixels on the CCD into one larger pixel. This is used to increase the signal amplitude and to reduce the spatial sampling frequency. The imaging setup results in a spatial sampling frequency of 2.9 pixels/μm. The choice of the objective is mainly determined by the high numerical aperture. A high numerical aperture has the advantage that the integration time can be relatively short. The image brightness is commonly assumed to be proportional to the fourth power of the NA (Inoué 1986). We use an integration time of 1.0 s.

The two dyes are acquired simultaneously into one *monochrome* image. We have used a dual band-pass filter block, both emission and excitation, that is suitable for the two dyes, DAPI counter stain and Spectrum Orange. The dual band-pass filter block maps the blue fluorescence (DAPI, peak emission spectrum $\lambda = 452$ nm) into the middle gray values ($\sim$ 20% at full dynamic range) and the red signal (Spectrum Orange, peak emission spectrum $\lambda = 615$ nm) into the light gray values ($\sim$ 60% at full dynamic range). Through the judicious use of the color filters, the image processing is based on brightness contrast and not on color information. A typical image is shown in figure 1.

The whole system is embedded in a 'user-friendly' user-interface to control the microscope, to adjust parameters, and to evaluate the results. Before the system starts screening, an operator has to define a scan area on the slide. The area is scanned following a meander pattern. After a preset number of nuclei have been analysed, the

† Spectrum Orange is a trademark of Vysis Corporation, Downer's Grove, Illinois, USA.

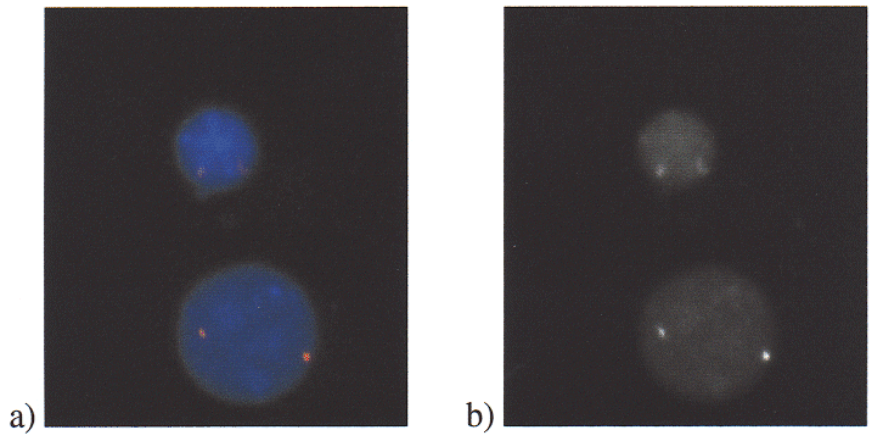

Figure 1. Double stained cells (DAPI + Spectrum Orange). These images are 143 by 165 pixels sampled at 2.9 pixels/μm. (a) Two-color FISH acquired in a true, multi-filter color image. (b) Same FOV acquired in one monochrome image. The bright dots are easy to distinguish from the gray nuclei.

system stops, and the result can be interactively verified and corrected. The individual nuclei are automatically relocated under the microscope. Visual inspection can be done either using the monitor display or through the microscope. The output of the dot counter can be a dot histogram, a confusion matrix, and/or a gallery of images of every cell that has been analysed (Netten *et al* 1997).

3. Image analysis

After the image acquisition, each image is processed to determine the number of dots per nucleus. The image analysis consists of finding the nuclei in the image, detecting the dots within each nuclei, counting the number of dots per nucleus and updating the results including a dot histogram for the entire specimen. The algorithm must be accurate. The number of false positives and false negatives must be as low as possible. On the other hand the algorithm must be fast. In a practical situation more then 70 MBytes of data have to be processed to analyse 500 nuclei.

The algorithm can be divided into four steps: (1) find a region that contains a nucleus, (2) find the nucleus in the region, (3) find the dots in nucleus, and (4) count the dots and update the results. This paper focuses on the third step, the dot detection procedure. Three different approaches are presented. The nucleus segmentation (steps 1 and 2) is discussed only briefly. A more detailed discussion is presented in Netten *et al* (1997).

3.1. Cell detection

3.1.1. Find region of interest (ROI). The goal of this first step is to find those regions that contain cell nuclei. To speed up the algorithm the original image is first sub-sampled by a factor of 8. The reduced image is then filtered to suppress the noise and to correct for shading. The entire resulting image is segmented by an automatically-chosen threshold (Zack *et al* 1977). An enclosing rectangle for each object in the segmented image defines the region of interest and is used for the next step.

3.1.2. Detect nuclei in ROI. For each ROI the original image is processed again at full resolution to define a mask for the nucleus. A gray-value opening is applied to remove the dots. The high intensity of the dots can influence the threshold level calculated in the next step. The iso-data thresholding algorithm (Ridler and Calvard 1978) is used to segment the ROI into object and background. The resulting object mask is then further processed using binary morphological operations to remove small objects and to separate slightly touching nuclei (Haralick *et al* 1987).

After segmentation, size, shape, and intensity features are measured for each object (Young and Roos 1988). The features are used to select single nuclei and to reject touching nuclei, debris, etc. The selected single nuclei are used in the next step where the dots are to be detected.

3.2. Dot detection

To count the dots, the original image is segmented again within the mask of the nucleus. Three different techniques are presented. The first algorithm is the most simple one and is based on a tophat transform followed by constant threshold. This technique is called *tophat threshold*. The second technique is called *Laplacian threshold* (nL threshold) and is an extended version of the tophat threshold. After the tophat transform a nonlinear Laplacian (nL) filter is applied to separate touching dots. The third technique is again based on the tophat transform but instead of a constant threshold a variable threshold is used to label the dots. This third technique is named *dot label*.

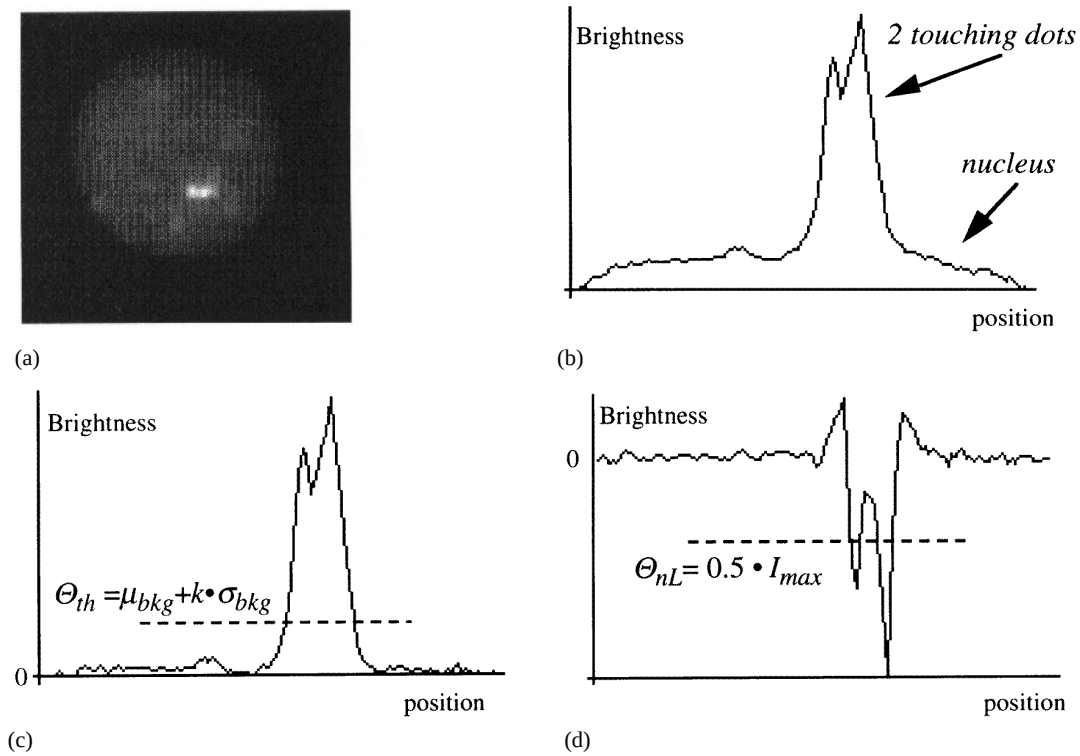

Figure 2. An example of tophat threshold and nL threshold. (a) Original image. (b) Intensity plot of a line crossing the two dots. (c) Tophat transform, tophat[Im] = Im−Max_5[Min_5[Im]]. (d) Laplacian, nL[tophat]=Max_3[tophat]+Min_3[tophat]− 2tophat. The index of the Max/Min operator defines the size of the square kernel.

3.2.1. Tophat threshold. A tophat transform (Meyer 1979) is performed on the original image to remove the DAPI counter stain. For light objects, on a darker background, the tophat transform is given by

$$Tophat(A, B) = A - \max_B \left(\min_B (A) \right) \qquad (1)$$

where max() and min() are maximum and minimum filters, respectively, over a region B. We typically choose B to be a 5×5 window. This is illustrated in figure 2. The tophat transform is only applied within the mask of the nucleus. The resulting image only contains the Spectrum Orange dots on a noisy background. The size of the tophat kernel must be slightly larger then the dots. A constant threshold is performed on the tophat transform to find the dots. The threshold level is given by $\Theta_{th} = \mu_{bkg} + k\sigma_{bkg}$ where μ_{bkg} and σ_{bkg} are the mean and standard deviation of the background inside the mask of the nucleus. The mean and standard deviation are estimated using the pixels below the 90% percentile intensity of the tophat image. The parameter k has to be determined on the basis of a limited number of nuclei used as a training set.

3.2.2. nL threshold. Most dots are detected properly with the tophat threshold, but some dots appear merged. Therefore, an extra step is included after the tophat threshold. A nonlinear Laplacian (Van Vliet *et al* 1989) is performed on the tophat image. This step is only applied within the mask of the tophat threshold. A threshold on a negative level will separate touching dots (see figure 2). The threshold level Θ_{nL} is determined by half the minimum intensity of the Laplacian image. The mask of the tophat transform and nL are combined into a mask of detected dots.

3.2.3. Dot label. The tophat threshold has the disadvantage that it is difficult to define a proper threshold level. Figure 2 shows that if the threshold level is too high a dot will be missed, and if it is too low, two dots will be merged. The nL filter is used to separate touching dots but has the disadvantage that it is sensitive to noise. False dots could be the result. To overcome these problems a different approach is presented.

Again a tophat transform is performed on the original image. In contrast with the tophat threshold, the dot label algorithm uses a variable threshold level. The basic idea of the algorithm is that pixels with an intensity that is equal to a threshold level are assigned to a dot if they are connected to that dot. If they are not connected to an existing dot, a new dot is created. The threshold level

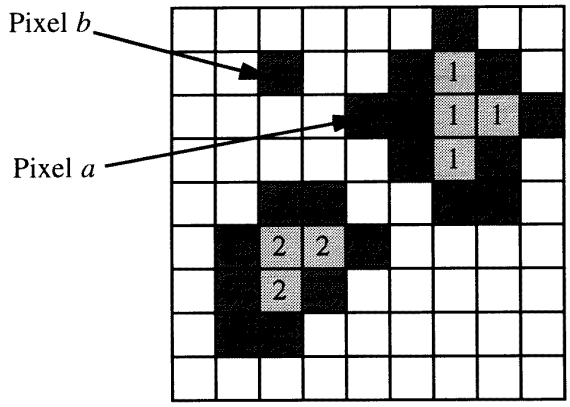

Figure 3. An example of the definition of connectivity. The light gray pixels are already labeled with 1 or 2. The dark gray pixels have an intensity within the threshold-band. Pixel a will be assigned to dot 1. Pixel b will create a new dot if the intensity is equal to Θ_{seed}.

Θ_{seed} starts at the maximum intensity of the image I_{max} and runs down until it is just above the background level $\Theta_{seed} = \mu_{bkg} + k\sigma_{bkg}$. A second threshold level Θ_{conn} is introduced to avoid false dots. Small variations in the image due to noise can create false dots. Instead of a threshold level, a threshold-band is used. Pixels within a threshold-band ($\Theta_{seed} \geq I(x, y) > \Theta_{conn}$) are assigned to a dot if they are connected to that dot. Pixels that are equal to the threshold level Θ_{seed} and not connected to a dot will create a new dot. The width of the threshold-band is related to the SNR of the images, specifically the σ associated with the noise. The second threshold level is defined as $\Theta_{conn} = \Theta_{seed} - 3\sigma_{camera}$ where σ_{camera} is the standard deviation of the camera at the intensity Θ_{seed}.

A pixel is connected to a dot if there is a path from that pixel to one pixel of the dot, considering all the pixels within the threshold-band. An example is given in figure 3. Two dots have been created and the pixels are labeled with 1 or 2. The dark gray pixels have an intensity within the threshold-band. Considering pixel a, there is a path, in a 4-connected neighborhood, along the gray pixels to a pixel of dot 1. So this pixel would be labeled as 1. Considering pixel b, there is no path to one of the dots. If the intensity of pixel b is greater than or equal to Θ_{seed} then this pixel is the seed of a new dot with label 3. Figure 4 shows three intermediate steps of the algorithm. The threshold level Θ_{seed} starts at the maximum intensity of the image. This pixel is the seed of the first dot. Pixels within the threshold-band are labeled if they are connected to that dot. The second dot is created when Θ_{seed} is equal to the maximum intensity of that dot. The threshold level Θ_{seed} decreases until it is just above the background level.

3.3. Feature extraction

To refine the result of the dot detection algorithm a number of features are measured. Those features can be used to verify if a detected dot is a real hybridization dot, to detect split dots, or to distinguish 'overlapping' dots from single dots. The problem of overlapping dots is discussed in the next section. This section presents a number of features that can be measured and how they can be used to improve the result.

Often features are based on the resulting mask of the segmentation procedure. For example the area is commonly estimated by counting the number of pixels in the mask. Because the dots are relatively small (area of dot $\approx$ 11 pixels), the area strongly depends on the threshold level. To make the features independent of the segmentation procedure, the measurements are not based on the dot mask. The features are measured using the pixels that have an intensity larger than a fraction γ of the maximum intensity of the dot. The maximum intensity I_{max} of a dot is the maximum intensity within the dot mask. As an example, the total intensity is defined as the sum of the intensities that are larger than γI_{max} and is given by:

$$I_{tot} = \sum_{x,y \in dot\ mask} I(x, y) clip\,(I(x, y) - \gamma I_{\max}) \quad (2)$$

where

$$clip(q) = \begin{cases} 1 & q \geq 0 \\ 0 & q < 0 \end{cases}$$

and the set (x, y) is chosen within the dot mask. The fraction γ is an arbitrary value between 0 and 1. If γ is close to zero, background noise will influence the measurements. If γ is too high the measurements will be based on only a few pixels. We have used a value $\gamma = 0.33$. The input image $I(x, y)$ is assumed to be corrected for background variation. The tophat image is used as the input image. The tophat transform subtracts the estimated background signal from the original image.

The following features are measured: maximum intensity, area, total intensity, average intensity, relative intensity, and eccentricity.

- Maximum intensity I_{max} is the maximum intensity within the dot mask.
- Area A_{dot} is the number of pixels with $I(x, y) > \gamma I_{max}$.
- Total intensity I_{tot} is the sum of the intensities with $I(x, y) > \gamma I_{max}$.
- Average intensity I_{avg} is the total intensity divided by the area.
- Relative intensity I_{rel} is the total intensity relative to the maximum total intensity within a nucleus and is defined as:

$$I_{rel}(d_i) = \frac{I_{tot}(d_i)}{\max\,(I_{tot}(d_1), \ldots, I_{tot}(d_n))} \quad (3)$$

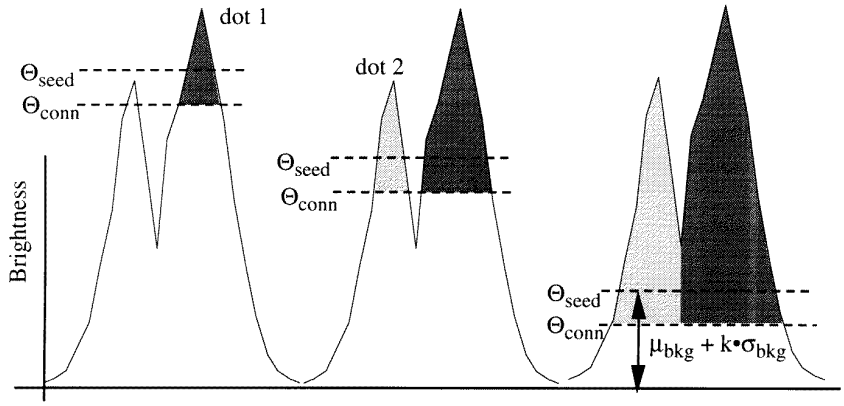

Figure 4. Three steps of the dot label algorithm. Θ_{seed} starts at the maximum intensity of the image. Dot 2 is created when Θ_{seed} is equal to the maximum intensity of that dot. The algorithm stops when θ_{seed} is just above the background level.

where $I_{tot}(d_i)$ is the total intensity of dot i. The relative intensity of a dot with the maximum total intensity will always be one.

- Eccentricity. The definition for the eccentricity is:

$$E_{dot} = \sqrt{\frac{\eta_{20}}{\eta_{02}}} \tag{4}$$

where η_{20} and η_{02} are the invariant second order normalized moments (Gonzales and Woods 1990). For a circular dot with a brightness distribution, that is circularly symmetric, the eccentricity will be one.

There are several factors that can influence the measurements. The number of pixels per dot, which is related to the sample density, affects the accuracy of the feature estimators. A number of publications (Van Vliet 1993, Young 1988) have reported on the relationship between the sample density and the coefficients of variation (CV) of several estimators. Considering the area, the theory predicts a percentage error below 10% (Young 1988), even though the dots are relatively small. However, the measured CV of the area is 35% (the percentage error and the CV are roughly equivalent measures). Thus the variation caused by a limited number of pixels is not a complete explanation of the difference between theory and experiments. A second source that can effect the measurements is the image brightness noise. In our case using a Photometrics camera and an average dot intensity above 500 ADU (camera range is 4095 ADU) the SNR will be about 30 dB (Netten *et al* 1994). Again a variation in the intensity of less then 4% due to noise is not significant in comparison to the measured variation. CVs of 21% and 42% for the intra- and internuclear distribution, respectively, of the total intensity have been reported (Nederlof *et al* 1992). The non-uniform illumination of the fluorescence microscope is another source that can contribute to the large variation of the intensity features. We have measured a shading of 20% . The shading is defined as the $(I_{max} - I_{min})/I_{avg}$ of a homogenous FOV. Finally the effect of the focus position of a dot must be considered.

Because nuclei are not flat, the dots are not always in the same focal plane, and it is possible that a dot is out of focus. It is also possible that the auto-focusing routine fails because it focuses on debris instead of the dots. Figure 5 shows the relative error for each feature as a function of the focus position. The relative intensity is not included. The features are measured at different focus positions. The relative error of a feature f for a dot at position z is defined as:

$$\varepsilon_f(z) = \frac{|f(z_0) - f(z)|}{f(z_0)} \tag{5}$$

where z_0 is the in-focus position of the dot. The average relative error for 12 dots is plotted in figure 5. All features strongly depend on the focus position. A focusing error of 0.5 μm yields a relative error of more then 15% for the area, average intensity and maximum intensity. Only the total intensity and the eccentricity have a relative error less than 10% if the focusing error is on the order of 1 μm. Each source that influences the measurements will contribute to the large variation of the features. We have measured CVs from 22% for the relative intensity to 51% for the average intensity.

3.4. Dot classification

The total intensity and the relative intensity are used to verify if a detected dot is a real hybridization dot. If the value of a feature is not within a certain interval the detected dot will be rejected. The interval is defined from the minimum value to the maximum value of that feature as observed in a training set. The training set contains only hybridization dots and no false dots.

The relative intensity is also used to detect split dots. As a result of cell replication, one target chromosome can appear as two dots that are close together. This is called

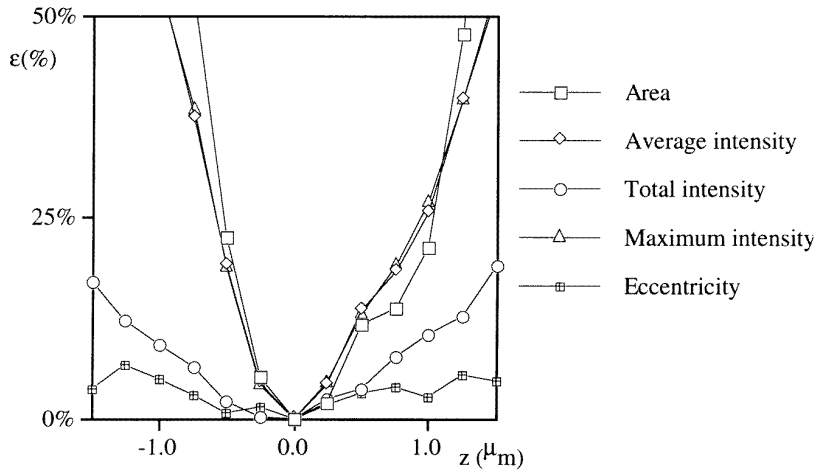

Figure 5. The relative error for five different features as a function of the focus position. Each data point is an average of the relative error measured for 12 dots. At $z = 0$ the dots are in focus. The error intervals of each plotted point are not included because they are significantly smaller than the estimated relative error.

a split dot. Those two dots should be counted as one dot. Mis-interpretation of a split dot can lead to an apparent high percentage of trisomy cells in what is actually a normal specimen. Figure 6 shows a gallery of cell nuclei that contain split dots. All these cells should be interpreted as cells containing two dots. The two dots of a split dot are always close together and the intensities of both dots are in general weaker then the third dot. Two dots d_i and d_j will be classified as a split dot if the following rule is true:

$$D(d_i, d_j) < D_{split}$$

and

$$\bar{I}_{rel}(\ldots, d_i + d_j, \ldots) > \bar{I}_{rel}(\ldots, d_i, d_j, \ldots) \tag{6}$$

where $D(d_i, d_j)$ is the projected distance between the two dots (d_i and d_j) and $\bar{I}_{rel}$ is the average relative intensity within a nucleus. Two dots are combined ($d_i + d_j$) if the average relative intensity is larger than if they are not combined (d_i, d_j). In other words the relative intensity of the combined dots must be closer to one. This rule is only applied to a cell with more than two detected dots.

4. Overlapping FISH dots in 2D images

It is important to realize that we are observing a three-dimensional nucleus through a two-dimensional projection, and thus one dot can hide 'behind' another dot. Normal interphase cell nuclei contain two copies of all autosomes. Microscope imaging projects the three dimensional nucleus onto a two dimensional sensor. Two dots can be distinguished if the projected dots do not severely overlap. Overlapping dots means in this article that the projected distance between two dots is too small to separate the dots properly.

4.1. Probability of an overlap

We have calculated the probability that in a nucleus with two dots the projected dots overlap by modeling the nucleus as an oblate spheroid. Cell nuclei in suspension have a spherical shape. In the process of slide preparation they turn into bodies of revolution called oblate spheroids, i.e. ellipsoids with principal axes R_c, R_c, R_h where R_c is the radius of a nucleus in the xy-plane and R_h is the principal axis parallel to the z-axis. Given an oblate spheroid, the height $h(x, y)$ of the body in the xy-plane is then given by:

$$h(x, y) = 2R_h\sqrt{1 - (x^2 + y^2)/R_c^2}. \tag{7}$$

If the position (x, y, z) of a dot is equally likely to be anywhere within the volume of the nucleus, the probability density function to find a dot inside an oblate spheroid at position (x, y) is:

$$p(x, y) = \frac{h(x, y)}{\frac{4}{3}\pi R_c^2 \cdot R_h} = \frac{3\sqrt{1 - (x^2 + y^2)/R_c^2}}{2\pi R_c^2} \tag{8}$$

where the denominator is the volume of an oblate spheroid. The two dots overlap when the second dot lies inside a cylinder with radius r_0 that is centered at the x, y position of the first dot. This is illustrated in figure 7. The radius r_0 is the smallest distance between two projected dots that can still be separated. The probability of an overlap is then given by:

$$P_0(D(d_1, d_2) < r_0) = \int_{x,y} p(x, y)(p(x, y)\pi r_0^2)\,\mathrm{d}x\mathrm{d}y$$
$$= \frac{9}{8}\left(\frac{r_0}{R_c}\right)^2 \tag{9}$$

with $D(d_1, d_2)$ the lateral distance between the two dots. The above calculation assumes that: (1) the probability

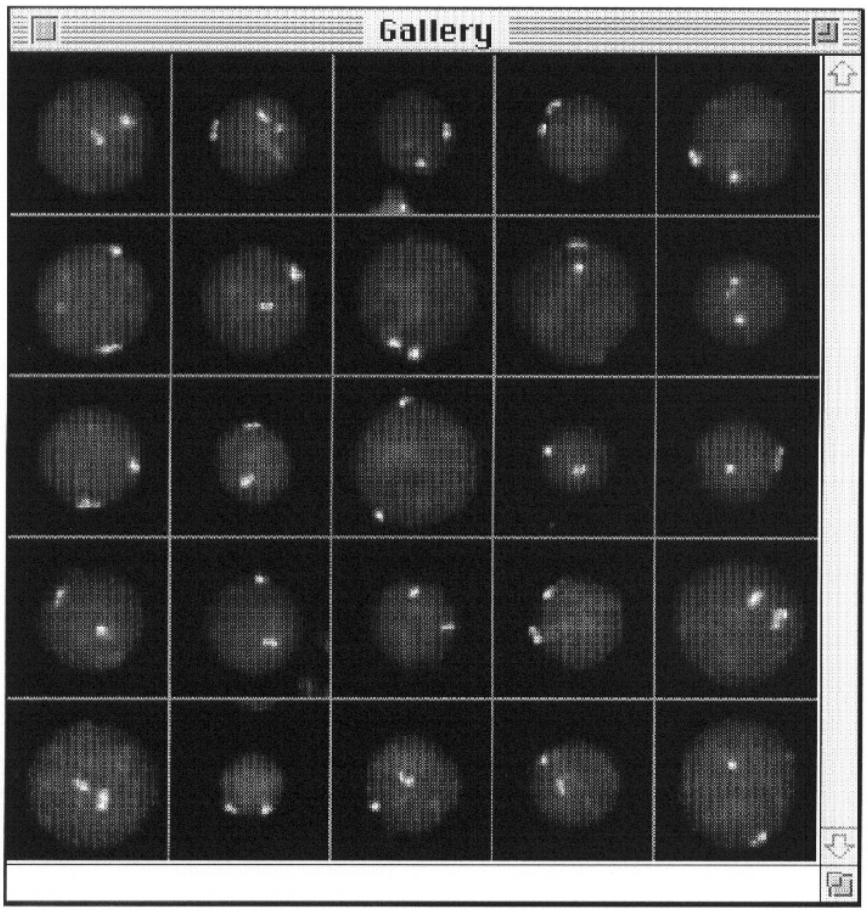

Figure 6. A gallery of images of cells that contain split dots.

density function of the position (x, y, z) of a dot inside a nucleus is uniformly distributed; (2) the positions of the dots are independent; (3) $r_0 \ll R_c$. Frequency distribution curves of observed distances between two targets have been compared with a model that assumes uniform and independent distribution of point-like targets. There is evidence that chromosomes occupy distinct territories in cell nuclei. Dietzel *et al* have shown a significant difference between observed data and this model (Dietzel *et al* 1995). But the differences are small, especially when the distance between the targets is small. Although the above assumptions are not completely correct, equation (9) yields a good approximation of the probability of an overlap.

It is interesting to see that the probability of an overlap is independent of the spheroid's eccentricity and is inversely proportional to the projected area of the nucleus. In other words, flattening the nuclei on a slide, which increases the area, will reduce the probability of an overlap.

The distance r_0 depends on the radius of an observed dot R_d and on the capability of the image processing algorithm to separate the dots. R_d is determined by the physical size of a dot and the point-spread-function of the optical system which depends on the NA of the objective and the emission wavelength λ of the fluorescent dye. If

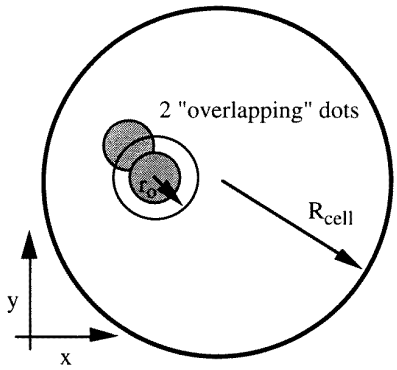

Figure 7. Two 'overlapping' dots when we are observing a three-dimensional nucleus through a two-dimensional projection. The two dots overlap when the second dot lies inside a cylinder with radius r_0 that is centered at the x, y position of the first dot.

the physical dots are smaller than a wavelength of light, each observed dot is approximately an Airy disc with $R_d = 0.61\lambda/\mathrm{NA}$. The probability of an overlap becomes:

$$P_0 = 0.42(\alpha\lambda/\mathrm{NA}R_c)^2 \tag{10}$$

where α is a parameter that represents our ability to segment two adjacent dots. In this case r_0 is defined as $r_0 = \alpha R_d$. If $\alpha = 1$, r_0 is equal to the Rayleigh criteria. The Rayleigh criterion defines a distance at which two superimposed Bessel functions can still be separated based on the maximum intensity. As an example for $\alpha = 1$, $\lambda = 0.615\,\mu$m (peak emission spectrum), NA = 1.3, and $R_c = 5\,\mu$m, the probability of an overlap $P_0 = 0.4\%$. In our case, using a centromeric probe, the radius of an observed dot is generally larger than the radius of an Airy disc. The physical size of a dot is significantly larger than the wavelength of light. In this case, the radius r_0 has to be determined experimentally.

4.2. Overlapping dots detection

The problem of 'overlapping' dots is especially important for the detection of a monosomy. Monosomy means that a proportion of the nuclei has only one copy of the target chromosome instead of two. 'Overlapping' dots mainly affect the estimated proportion of cells that contain one dot. In practice most cells have two copies of a chromosome and only a small sub-population of cells have an aberrant number of chromosomes. Due to overlapping, some of the cells with two chromosomes are counted as one. Because the proportion of cells with two dots is much larger than the proportion with one dot, the error will be significant for the estimated proportion of cells with one dot. Although overlapping also occurs with cells with three dots, the effect on the estimated proportion of cells with two dots is much smaller. The proportion of cells with three dots is in practice much smaller than those with two dots.

To improve the results of the image processing algorithm we want to include an extra step that classifies detected dots into single dots or 'overlapping' dots based on these features. We may expect that two 'overlapping' dots will have twice the total intensity of a single dot or that the eccentricity of two touching dots is larger then the eccentricity of a single dot. A nearest neighbor classifier (Fukanaga 1990) has been used to see if it is possible to discriminate 'overlapping' dots from single dots based on the total intensity and eccentricity. The classification is only applied to the cells with one detected dot.

5. Results

The performance of the dot counter has been evaluated. A set of images has been used to test the different algorithms. Two slides have been scanned automatically. From each FOV an image has been acquired and has been saved on computer disk. Afterwards the images have been checked manually to see if they were properly focused. Because the purpose of the experiments was to investigate the performance of the image analysis algorithms, focusing errors have been excluded. A total of 352 images have been acquired, containing 1014 nuclei. A subset of 113 images with 200 nuclei has been used as an independent training set to adjust the parameters of the algorithms. The other 239 images have been used to test the system.

The result of the dot counter is a dot distribution that gives the proportion of cells containing 0, 1, 2, 3 or > 3 dots. The proportion p_i is estimated by

$$p_i = \frac{n_i}{N} \tag{11}$$

where n_i is the number of cells with i dots and N is the total number of cells. Assuming that the probabilities of the counting errors are constant and the selection of the cells is random, the proportion p_i is a multinomial distribution. The standard deviation s_i is then given by:

$$s_i^2 = \frac{p_i(1-p_i)}{N}. \tag{12}$$

For large N the multinomial distribution is approximately Gaussian. In that case a 95% confidence interval of the estimated proportion is approximately between $p_i \pm 1.96 s_i$ (Castleman and White 1995).

5.1. Counting results

The dot distribution of the test set has been estimated with the three different algorithms; tophat threshold, nL threshold and dot label (table 1). Together with these fully automated counting results, manually obtained results are also given. Each nucleus from the test set has been counted manually using the monitor display. The first row of table 1 shows the dot distribution for a normal specimen, given by the product specification of the CEP 12 Spectrum Orange probe (Vysis). This dot distribution will be used as the 'ground truth'. The manually obtained results of the test set are comparable with the specifications. Using a monitor display instead of the microscope does not influence the counting result. The automated results differ from the manually obtained dot distribution. The best result has been obtained with the dot label algorithm. About 2% of the nuclei are counted incorrectly. Using the tophat threshold the error percentage increases to 8%.

5.2. Overlapping dots

The percentage of nuclei containing one dot is significantly larger for the automated results obtained with the tophat threshold and nL threshold than with manual counting. This is especially true when the tophat threshold has been used. Most errors occur because 'overlapping' dots cannot be properly separated. The proportion ϵ_0 of nuclei that are

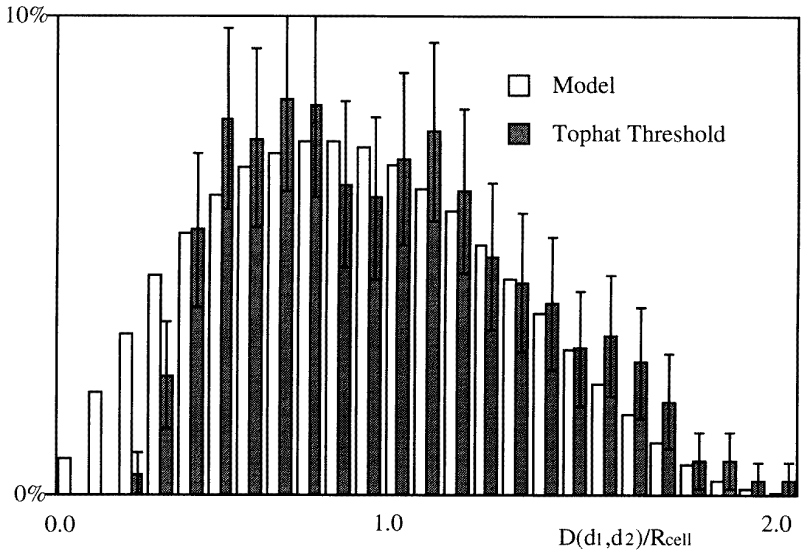

Figure 8. The frequency distribution histogram of the projected distance between two dots within a nucleus. The distance is normalized by the radius of the nucleus. The measured distribution curves are given for all nuclei with two detected dots obtained (1) with the tophat threshold and (2) with the distribution of a model which assumes uniform and independent distribution of point-like dots.

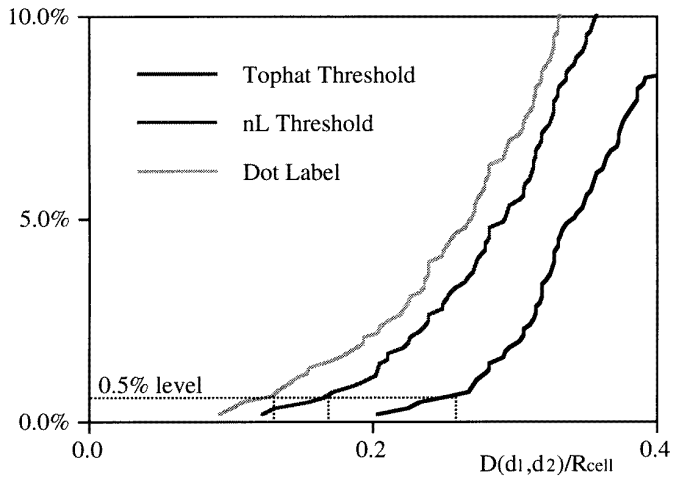

Figure 9. The cumulative frequency distribution curves of the projected distance between two dots as observed in cells with two detected dots. The distance is normalized by the radius of the nucleus. The curves have been obtained with the three algorithms. The 0.5% level defines r_0. Only a small range near zero is plotted.

counted as one instead of two has been obtained. The results of the three algorithms are examined visually using the monitor display. Each nucleus that has an automated scoring of one dot has been classified manually as a single dot or an 'overlapping' dot. Table 2 shows the number of 'overlapping' dots for each algorithm as a percentage of the number of nuclei containing two dots. The error interval is defined as two times the standard deviation given by equation (12). The smallest distance r_0 that can be separated is estimated using the frequency distribution of the distances between two dots. The distance between two dots is normalized by the radius of the nucleus R_c. The result is also given in table 2, together with the estimated probability of an overlap P_0 using equation (9). For each algorithm the distances between two dots have been measured for all nuclei with two correctly detected dots. Because the frequency distribution of distances does not contain those cells for which the dots are not properly separated, we may expect that the smallest distance that occurs in the distribution is the smallest distance for which the dots can be separated. But it is possible that the smallest distance of the distribution is an outlier caused by noise. In other words, two touching dots, at a certain distance, are sometimes properly separated and sometimes

Table 1. Percentage of cells containing various numbers of dots for the test set as identified manually using a monitor display and with full automation using three different algorithms. The dot distribution for normal specimens is also given (Vysis). The number of cells that have been counted $N = 814$.

Chromosome 12	0	1	2	3	> 4
Manual					
Spectrum Orange™	0.2%	1.5%	97.1%	1.2%	0.1%
Monitor display	0.6%	1.4%	97.2%	0.6%	0.3%
Automated					
Tophat threshold	0.6%	8.7%	89.3%	1.1%	0.3%
nL threshold	0.5%	3.9%	94.2%	1.0%	0.4%
Dot label	0.5%	1.8%	95.1%	2.1%	0.5%

Table 2. Because 'overlapping' dots cannot always be properly separated, some nuclei are counted as having one dot instead of two. The proportions of 'overlapping' dots, ε_0 as a percentage of the total number of cells containing two dots, observed with the three different algorithms are given, together with the measured ratio r_0/R_c and the estimated probability of an overlap given by equation (9). The number of cells with two dots is $N \approx 800$.

Algorithm	ε_0	r_0/R_c	$P_0(r_0/R_c)$
Tophat threshold	$7.1 \pm 1.8\%$	0.26 ± 0.02	$7.6 \pm 1.2\%$
nL threshold	$2.4 \pm 1.0\%$	0.17 ± 0.01	$3.1 \pm 0.4\%$
Label dot	$0.9 \pm 0.6\%$	0.13 ± 0.01	$1.9 \pm 0.3\%$

not. Therefore r_0 is defined as the distance at the 0.5% level of the cumulative frequency distribution. In our case with $N \approx 800$, r_0 becomes equal to the fourth smallest distance (sample) in the distribution. The error interval is defined as the difference between the third smallest distance and the fifth smallest distance. Figure 8 shows the measured frequency distribution as a result of the tophat threshold in combination with the distribution of a model which assumes the uniform and independent distribution of point-like dots. The model frequency distribution is obtained from simulations. Figure 9 shows the cumulative distributions for the three algorithms with the resulting r_0 distance at the 0.5% level. The dot label algorithm has the best performance in segmenting 'overlapping' dots followed by the nL threshold. The nL filter significantly improves the result of the tophat threshold. There is no significant difference between the estimated probability of an overlap using equation (9) and the measured proportion of 'overlapping' dots ε_0, but the error intervals are large.

Including an extra classification step to distinguish 'overlapping' dots from single dots has been evaluated. The result of the tophat threshold has been used to see if it is possible to improve the results based on the total intensity and eccentricity. Table 3 gives the mean and standard deviation of the total intensity and the eccentricity for 'overlapping' dots and single dots. The number of single dots is much larger than the number of 'overlapping' dots because we have used all dots from cells with two detected dots. As we expected, the average total intensity of 'overlapping' dots is twice the total intensity of single dots. Also the eccentricity is larger for 'overlapping' dots. The variation, however, is also larger. A scatter plot (figure 10) illustrates the overlap between the two classes. A nearest neighbor (NN) classifier (Fukanaga 1990) has been applied to these data. The set of 1452 single dots and 57 'overlapping' dots has been used as a training set and as a test set. Because the number of 'overlapping' dots is small, we did not want to split the set into a separate training set and test set. Table 4 gives the result of 6-NN classifier. Because the training set has been used as the test set these results are optimistic. The classifier has been trained in such a way that the number of false negatives (single dots that are classified as 'overlapping' dots) is close to zero. Only 54% of the 'overlapping' dots are classified correctly. This result shows that an extra classification step only slightly improves the result. The results of the nL threshold algorithm and dot label algorithm are still better than the results of the tophat threshold with the extra classification step. The large variation of the features makes it difficult to distinguish 'overlapping' dots from single dots.

5.3. Split dots

The ability to separate touching dots means that a split dot could be detected as two dots. Mis-interpretation of a split dot can lead to a high percentage of trisomy cells in normal specimens. A simple rule has been used to detect split dots based on the distance between the two dots and the relative intensity. This rule is only applied on nuclei with more than two dots. If two dots are classified as a split dot, these dots are combined into one dot. To test this rule the test set has been analysed again using the three algorithms but without split dot detection. The results given in table 1 were obtained *with* the use of split dot detection. The results of the two experiments are now compared. Table 5 gives the proportion of nuclei that have been counted incorrectly because of a split dot, with and without the use of split dot detection. Because the tophat threshold does not have the ability to separate touching dots, no split dots have been detected as two dots and the tophat threshold is therefore not included. If the ability to separate dots improves, the number of split dots that have been detected as two dots increases. Using the split dot detection, most of the split dots have been combined into one. Four nuclei have been counted incorrectly due to split dots using the dot label algorithm and cause a slightly higher percentage of trisomies (table 1).

6. Discussion

We have developed a completely automated fluorescence microscope system that can examine interphase cell nuclei in order to determine the proportions of cells containing

Table 3. The mean and standard deviation of the total intensity and the eccentricity for single dots and 'overlapping' dots. The result of the tophat threshold is used to calculate these values.

Features	Total intensity (ADU)		Eccentricity	
	μ	σ	μ	σ
Single dots ($N = 1452$)	7458	4079	1.5	0.5
Overlapping dots ($N = 57$)	14823	7677	3.1	1.3

Table 4. A confusion matrix for a 6-NN classifier using the total intensity and the eccentricity to distinguish 'overlapping' dots from single dots. The learning set was also used as test set. More than 99% of the single dots are classified as single dots but only 54% of the 'overlapping' dots are classified correctly.

Classification \ Test set	Single dots ($N = 1452$)	Overlapping dots ($N = 57$)
Single dots	99.7%	45.6%
Overlapping dots	0.3%	54.4%

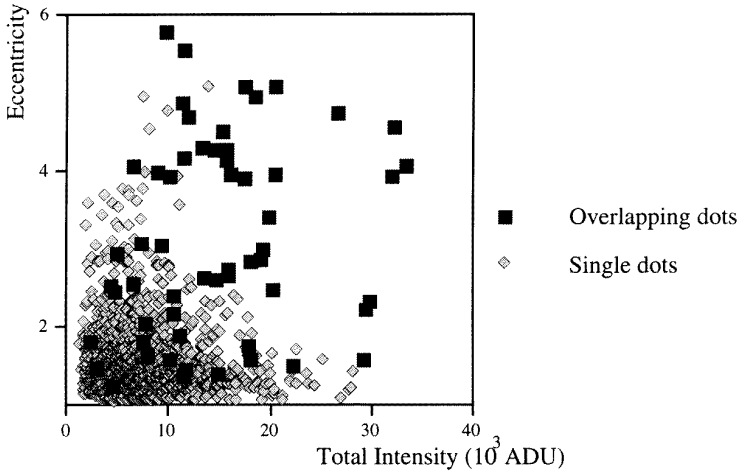

Figure 10. Scatter plot of the total intensity and the eccentricity for single dots (light gray) and 'overlapping' dots (dark gray).

Table 5. The proportion of cells that have been counted incorrectly due to a mis-interpretation of a split dot. The percentages are given for nL threshold algorithm and dot label algorithm with and without split dot detection. The split dot detection combines most of the split dots. The total number of cells is $N = 814$.

Algorithm	No split dot detection	Split dot detection
nL threshold	3.2%	0.4%
Dot label	9.1%	1.2%

0, 1, 2, 3, or $>$ 3 dots. This paper describes the image processing algorithms that are responsible for the dot detection. Three techniques have been presented: tophat threshold, nL threshold, and dot label. The different techniques have been evaluated and the results have been compared with manual counting. All experiments have been done with a set of images stored on computer disk. Because we wanted to evaluate the image processing algorithms, the images have been checked manually in order to avoid errors caused by the image acquisition.

The results of the dot label algorithm approximate the manually obtained results. Only 2% of the cells are counted incorrectly. Using nL threshold or tophat threshold the error rates increase to 3% and 8%, respectively. Because the images of the test set have been checked manually the error rate of the complete system, automated scanning and image analysis, will be higher. Previously presented results (Netten *et al* 1997) showed an error rate of 3% (out of the total 11%) to be caused by focusing errors. These results showed also that the variance of the dot distribution, using full automation, is significantly larger than the expected variance of a multinomial distribution. Our experience is that the large variation is related to the slide quality. Clean slides will make the error rate of the dot counter more predictable.

'Overlapping' dots are the main problem of the dot segmentation. The overall error rate is strongly related to the ability to separate touching dots. The estimated distance r_0 gives an indication of how well the dots can be separated. The smallest distance has been measured with the dot label algorithm. The estimated error percentage due to an overlap is comparable with the calculated probability of an overlap using the distance r_0 (equation (9)). Of course this result depends on how we have defined r_0. The distance r_0 has been defined as a fixed distance. In other words two dots at a distance smaller than r_0 will not be separated, and if the distance is larger they will be separated. The frequency distribution of the distances between two dots, as a result of the tophat threshold in comparison with the results of the model, shows that it is not a fixed distance but an interval where some dots are separated and some are not. If the distance between the dots becomes larger, the distribution of the model, which assumes uniform and independent distribution of point-like dots, is in agreement with the result of the tophat threshold. The underlying assumption of equation (9) seems to be reasonable. Although the definition of r_0 is ambiguous, equation (9) yields a good approximation to the probability of an overlap.

Including an extra step, that distinguishes 'overlapping' dots from single dots based on the eccentricity and total intensity, has been tested. Although the total intensity of 'overlapping' dots is twice the total intensity of a single dot, and the eccentricity is larger for 'overlapping' dots, the large variation of the features yields only a small improvement of the results.

If the ability to separate touching dots increases, the number of split dots that are detected as two dots will increase. A simple rule is applied to detect split dots. If two dots are classified as a split dot these two dots are combined. Without the split dot detection, the results of nL threshold and dot label would have a higher estimated percentage of trisomy cells and the results would be unreliable. The split dot detection combines most of the spit dots. We have only tested the split dot detection on a normal specimen. Therefore we could not measure the false negative rate of the split dot detection. The false negative rate has been defined as the percentage of real trisomy cells that are counted as two dots because of a false detection of a split dot. None of the few trisomy cells in the test set has been detected as a split dot. The number of trisomy cells is not sufficient to give an estimation of the false negative rate.

Acknowledgments

This work was partially supported by Vysis, Downer's Grove, Il, USA and by the European Concerted Action on Automation of Molecular Cytogenetic Analysis (CA AMCA).

References

Aikens R S, Agard D A and Sedat J W 1990 Solid-state imagers for microscopy *Fluorescence Microscopy of Living cells in Culture Part A* ed L Taylor and Y Wang (New York: Academic) pp 291–313

Boddeke F R, van Vliet L J, Netten H and Young I T 1994 Autofocusing in microscopy based on the OTF and sampling *Bioimaging* **2** 193–203

Carothers A C 1994 Counting, measuring and mapping in FISH-labelled cells: sample size considerations and implications for automation *Cytometry* **16** 298–304

Castleman K R and White B S 1995 Dot count proportion estimation in FISH specimens *Bioimaging* **3** 88–93

Dietzel S, Weiland E, Eils R, Múnkel C, Cremer C and Cramer T 1995 Three-dimensional distribution of centromeric or paracentromeric hetrochromatin of chromosomes 1, 7, 15 and 17 in human lymphocyte nuclei studied with light microscopic axial tomography *Bioimaging* **3** 121–33

Eastmond D A and Pinkel D 1989 Aneuploidy detection by analysis of interphase nuclei using fluorescence *in situ* hybridization with chromosome-specific probes *Prog. Clin. Biol. Res.* **318** 277–84

Fukanaga K 1990 *Introduction to Statistical Pattern Recognition* (New York: Academic)

Gonzales R C and Woods R E 1990 *Digital Image Processing* (Reading, MA: Addison-Wesley)

Haralick R M, Sternberg S R and Zhuang X 1987 Image analysis using mathematical morphology *IEEE Trans. Pattern Anal. Machine Intell.* **9** 532–49

Hopman A H N, Moesker O, Smeets A W G B, Pauwels R P E, Vooijs G P and Ramaekers F C S 1991 Numerical chromosome 1, 7, 9 and 11 aberrations in bladder cancer detected by *in situ* hybridization *Cancer Res.* **51** 644–51

Inoué S 1986 *Video Microscopy* (New York: Plenum)

Kibbelaar R E, Kok F, Dreef E J, Kleiverda J K, Cornelisse C J and Raap A K 1993 Statistical methods in interphase cytogenetics: an experimental approch *Cytometry* **14** 716–24

Meyer F 1979 Iterative image transformation for an automatic screening of cervical cancer *J. Histochem. Cytochem.* **27** 128–35

Mullikin J C, van Vliet L J, Netten H, Boddeke F R, van der Feltz G and Young I T 1994 Methods for CCD camera characterization *Proc. SPIE Image Acquisition and Scientific Imaging Systems (San Jose)* **2173** pp 73–84

Nederlof P M, Van der Flier S, Raap A K and Tanke H J 1992 Quantification of inter- and intra-nuclear variation of fluorescence *in situ* hybridization signals *Cytometry* **13** 831–6

Nederlof P M, Van der Flier S, Raap A K, Van der Ploeg M, Kornips F and Geraedts J P 1989 Detection of chromosome aberrations in interphase tumor nuclei by non-radioactive *in situ* hybridization *Cancer Genet. Cytogenet.* **42** 87–98

Netten H, Van Vliet L J, Boddeke F R, De Jong P and Young I T 1994 A fast scanner for fluorescence microscopy using a 2-D CCD and time delayed integration *Bioimaging* **2** 184–192

Netten H, Young I T, van Vliet L J, Vrolijk H, Sloos W C R and Tanke H J 1997 FISH 'n Chips: automation of fluorescent dot counting in interphase cell nuclei *Cytometry* in press

Ridler T W and Calvard S 1978 Picture thresholding using an iterative selection method *IEEE Trans. Syst. Man Cybernet.* **SMC-8** 630–2

Van Vliet L J 1993 Grey-scale measurements in multi-dimensional digitized images *PhD Thesis* Delft

University of Technology
Van Vliet L J, Young I T and Beckers A L D 1989 A non-linear Laplace operator as edge detector in noisy images *Comput. Vision, Graphics Image Process.* **45** 167–95
Vysis *Chromosome 12 product specifications* Downer's Grove, Illinois, USA
Young I T 1988 Sampling density and quantitative microscopy *Anal. Quant. Cytol. Histol.* **10** 269–75
Young I T and Roos R 1988 Acuity: image analysis for the personal computer *Pattern Recognition and Artificial Intelligence* ed E S Gelsema and L N Kanal (Amsterdam: Elsevier) pp 5–16
Zack G W, Rogers W E and Latt S A 1977 Automated measurement of sister chromatid exchange frequency *J. Histochem. Cytochem.* **25** 741–53