

Automated microscopy system for detection and genetic characterization of fetal nucleated red blood cells on slides.

Ilya Ravkin and Vladimir Temov *

Applied Imaging Corp., 2380 Walsh Avenue, Bldg. B, Santa Clara, CA 95051

ABSTRACT

The detection and genetic analysis of fetal cells in maternal blood will permit noninvasive prenatal screening for genetic defects. Applied Imaging has developed and is currently evaluating a system for semiautomatic detection of fetal nucleated red blood cells on slides and acquisition of their DNA probe FISH images.

The specimens are blood smears from pregnant women (9-16 weeks gestation) enriched for nucleated red blood cells (NRBC). The cells are identified by using labeled monoclonal antibodies directed to different types of hemoglobin chains (gamma, epsilon); the nuclei are stained with DAPI.

The Applied Imaging system has been implemented with both Olympus BX and Nikon Eclipse series microscopes which were equipped with transmission and fluorescence optics. The system includes the following motorized components: stage, focus, transmission, and fluorescence filter wheels. A video camera with light integration (COHU 4910) permits low light imaging. The software capabilities include scanning, relocation, autofocusing, feature extraction, facilities for operator review, and data analysis.

Detection of fetal NRBCs is achieved by employing a combination of brightfield and fluorescence images of nuclear and cytoplasmic markers. The brightfield and fluorescence images are all obtained with a single multi-bandpass dichroic mirror. A Z-stack of DNA probe FISH images is acquired by moving focus and switching excitation filters. This stack is combined to produce an enhanced image for presentation and spot counting.

Keywords: fetal cells, prenatal diagnostics, automated microscopy, FISH, DNA probes, image analysis

1. INTRODUCTION

Current procedures for prenatal genetic testing (chorionic villus sampling and amniocentesis) are invasive and carry a risk of miscarriage. These procedures are performed only when the risk of a genetic defect is higher than the risk of the procedure itself, e.g. in women older than 35 years. Development of a noninvasive procedure would open prenatal genetic testing to women of all ages. Isolation of fetal cells from the peripheral blood during pregnancies has been shown to be a feasible procedure that can be available to all women without the danger of triggering a miscarriage^{1,2,3}.

The fetal cell project at Applied Imaging Corporation has 3 main components: enrichment of fetal NRBCs from maternal blood, positive identification of fetal NRBCs, and genetic analysis of fetal NRBCs. The instrument to be described, plays a major role in the second and third components.

The enrichment procedure starts with 20mL of EDTA anticoagulated whole blood from pregnant women at 9 to 16 weeks post conception. It is calculated from PCR studies⁴, that this amount of blood contains approximately from 20 to 100 fetal cells. The goal is to achieve a 10^4 reduction of maternal blood cells without significant loss of target cells. This will permit the deposition of the fetal NRBC on a slide. The first step of enrichment is the separation of the blood sample into density fractions by centrifugation of the whole blood in a specially designed plastic tube⁵. The NRBCs are harvested from the interface layer between white blood cells (WBC) and red blood cells (RBC). The second step is a selective lysis of the remaining maternal RBCs. The third and final step separates NRBCs from the remaining WBCs in a three-layer density

* I.R.(correspondence): Email: ir@aicorp.com; Telephone: 408-562-0250.
V.T.: Email: vt@aicorp.com

gradient formed of a silica colloid Percoll (Pharmacia, Uppsala, Sweden) suspended in gelatin under hypertonic conditions. After centrifugation, the NRBCs are harvested from the bottom of the gradient and deposited on a slide.

The resulting slide contains NRBCs, RBCs and WBCs, all of them of possible fetal or maternal origin. In order to identify the fetal NRBCs we must create a set of features that would distinguish them from other types of cells. This is done by creating one type of contrast in cells containing fetal hemoglobin, and another type of contrast in cells having nucleus. The slide is first reacted with the primary antibody - mouse anti fetal hemoglobin (HbF), then with the secondary antibody, - goat anti mouse conjugated to biotin, - lastly streptavidin conjugated with alkaline phosphatase is added followed by Vector Blue substrate. The result is a blue precipitate on the cytoplasm of cells, which contain fetal hemoglobin. A DNA intercalating agent (DAPI) gives all nuclei a fluorescent blue stain; the presence of both these contrasts determines a fetal NRBC.

In order to achieve the desired goals, the instrument must have the following capabilities:

1. Automated finding of cells; storage of cell images and their slide coordinates.
2. Interactive review, classification, and selection by operator.
3. Relocation of cells.
4. Acquisition, enhancement and presentation of probe images, and counting of FISH spots.

The main design challenge is the high variation among specimens, which complicates both the design and use of the system. The instrument has to adapt to difference in staining intensity and probe brightness, type of cell deposition on slides (smear or cytospin), cell density, different kinds of samples and prevalence of different cell types.

2. HARDWARE CONFIGURATION

The system includes a microscope with transmission and fluorescence capabilities, a trinocular head, and 10X, 20X, and 40X objectives. Typically an Olympus BX-60 microscope (Olympus America, Inc. Melville, NY) is used. A single or a multi-slide scanning stage (Maerzhauser Co., Upper Saddle River, NJ) is mounted on the microscope with a 7 position transmission filter wheel, a 12 position fluorescence filter wheel, and a focus drive (TOFRA, Palo Alto, CA). All of these devices are based on stepping motors and are controlled by microstepping motor controllers (Intelligent Motion Systems, Marlborough, CT). Images are acquired through a video camera with light integration capability (COHU 4910, Cohu, Inc., San Diego, CA) and a custom built frame grabber board, which includes a 10 bit ADC and frame averaging. Computer workstation running Windows 95 operating system is used to control all motion, perform image acquisition and processing, and user interface functions.

Fig. 1 shows the optical configuration for the sequential or simultaneous detection of DAPI fluorescence and Vector Blue absorbance positive cells, and subsequently also for FISH imaging. Its main feature is the permanent presence of the multiband mirror and emission filter for the detection of both absorption and fluorescence. The epiillumination starts with a mercury arc, traverses the DAPI excitation filter and is reflected down by the multiband mirror; excites the blue fluorescence of the DAPI stained cells, the emitted light returns through the objective, and passes through both the mirror and the emission filter to the camera. The transillumination starts with a halogen lamp, passes through a long-pass (red) filter, is absorbed by the cells stained with Vector Blue, and passes through the objective, the multiband mirror and emission filter to the camera. Depending on DNA probes, we use Chroma 83000 triple band filter set (Chroma Technology Corp., Brattleboro, VT) or Vysis quad DAPI/Aqua/Green/Orange filter set (Vysis, Inc., Downers Grove, IL).

Custom slides were developed and are employed for all work done on the instrument. These slides have four painted squares in the corners with crosses laser etched in each square. Prior to scanning, coordinates of the reference points are recorded in the scan data file. At any time later, the reference points can be easily found and centered in the camera field of view. The offset is used for accurate relocation to all other objects in the scan file.

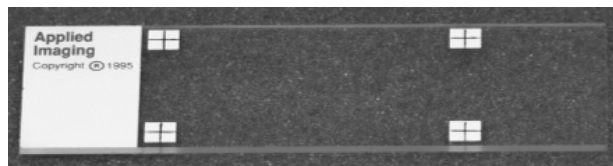


Figure 2. Slide with reference marks.

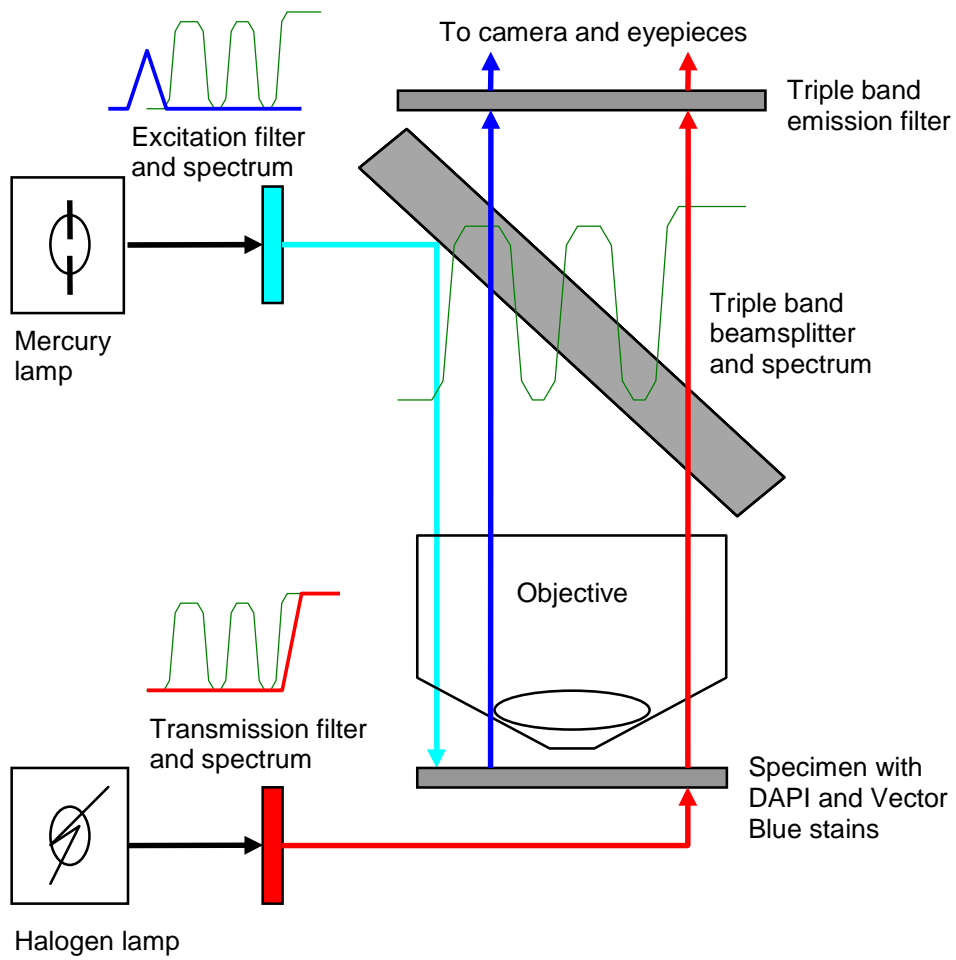


Figure 1. Optical configuration of the system. Spectra are for illustration purposes only.

3. IMAGING MODES

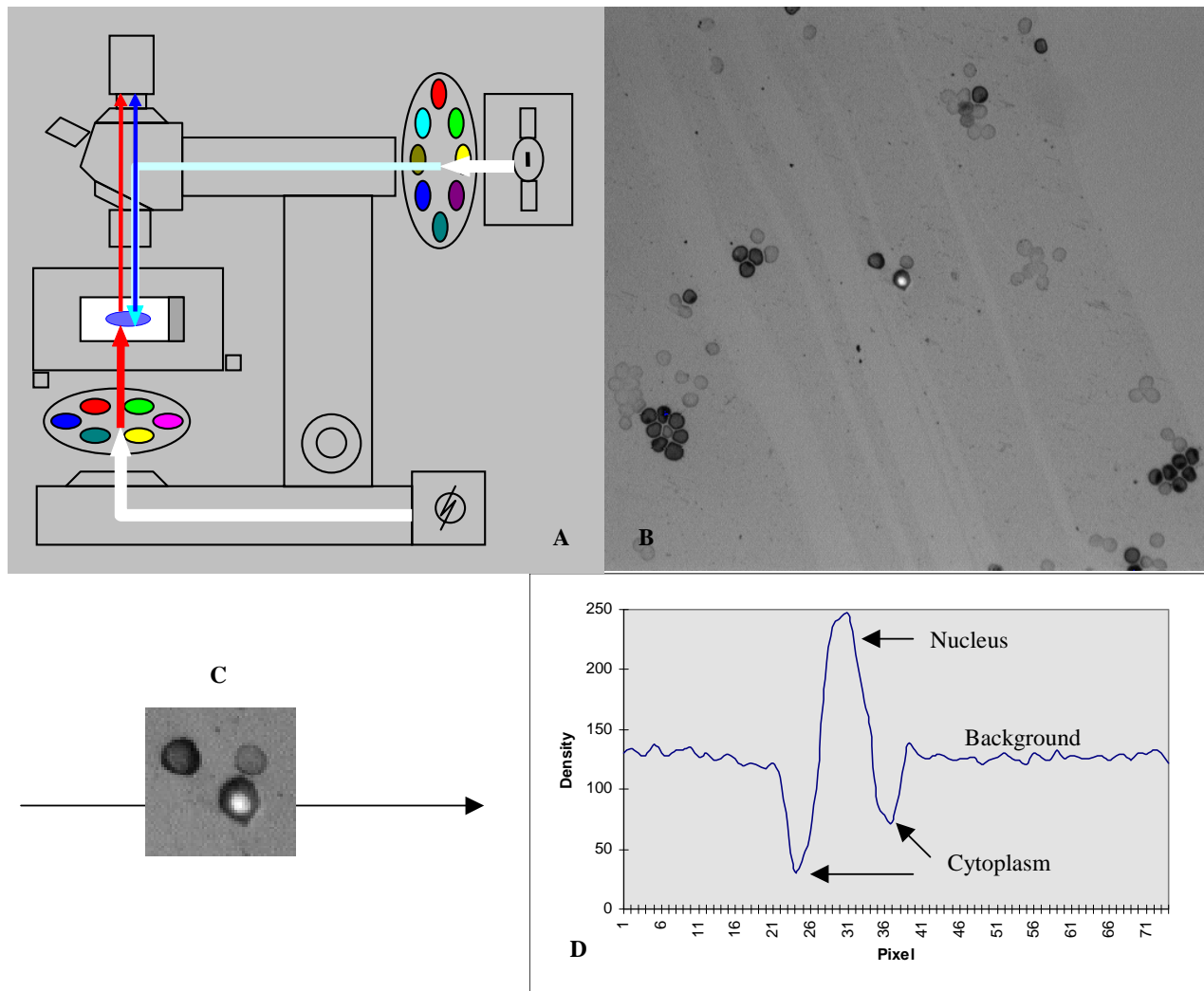


Figure 3.

A - Optical configuration for simultaneous imaging of brightfield and fluorescent components of stained specimens.

B - Image of a field containing RBCs and an NRBC. Antibody-stained objects appear darker than the background and DAPI-stained objects appear brighter than the background.

C and D - A nucleated red cell and a density profile through it.

In the simultaneous acquisition mode (Fig. 3) the sample is transilluminated with red light and epiilluminated with UV light at the same time. The intensities of both illumination systems light are balanced so that the empty background corresponds to a mid-level gray. Antibody-stained objects appear darker than the background and DAPI-stained objects appear brighter than the background. A combined light intensity profile through a fetal NRBC shows a significant drop in the cytoplasm, rise in the nucleus and again drop in the cytoplasm.

In the sequential acquisition mode, two images are taken on every field. For the first image, epiillumination is blocked by the filter wheel, and the red light goes through the specimen to the camera. The absorption image shows objects positive for antibody staining. For the second image, transillumination is blocked, and the specimen is excited by UV light going down through the objective. The fluorescence image shows all nuclei in the field. Fetal NRBCs are the objects having dark cytoplasms in the first image and bright nuclei in the second.

Simultaneous illumination has the advantages of speed (because there is only one image to acquire and no filter switch for every field), and easy human recognition of NRBCs under the microscope. On the other hand, the disadvantage is a tighter requirement for cell fixation and sharpness of DAPI image. There are conditions when the nuclear counterstain image is fuzzy and overlaps with the cytoplasm. In this case, simultaneous illumination cancels out both contrasts and only the sequential method can be used.

4. RECOGNITION OF FETAL CELLS

Recognition of fetal cells in the simultaneous acquisition method (Fig. 4) involves two additional steps compared with sequential acquisition since the images, which have been combined optically, now have to be separated digitally. First, the background value is determined as the mode of density distribution. Second, the combined images (Fig. 4, Col. 1) are split into the below-the-background component, characterizing antibody absorption (Fig. 4, Col. 3), and the above-the-background component, characterizing DAPI fluorescence (Fig. 4, Col. 4). In the sequential acquisition, the background is removed separately for brightfield and fluorescence images, and it is estimated as the brightest (or darkest) pixel in the area around the potential object. The rest of the algorithm is the same for both methods.

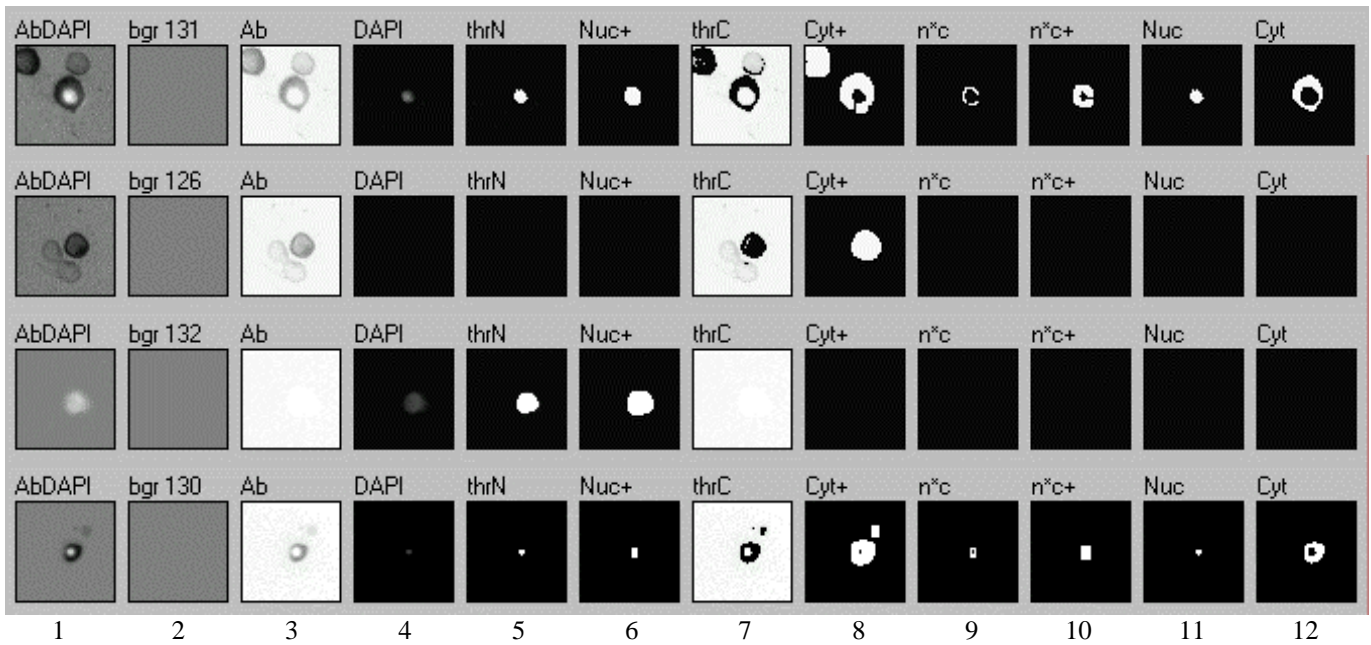


Figure 4. Recognition of fetal NRBCs.

Columns:

- 1 – optically combined brightfield and fluorescence images
- 2 – background images estimated as modes of density distribution
- 3 – below-the-background component, characterizing HbF antibody absorption
- 4 – above-the-background component, characterizing DAPI fluorescence
- 5 – threshold applied to the nuclear image produces a binary nuclear mask
- 6 – dilated nuclear mask
- 7 – threshold applied to the cytoplasm image produces a binary cytoplasm mask
- 8 – dilated cytoplasm mask
- 9 – intersection of dilated masks of the nucleus and cytoplasm
- 10 – dilated intersection
- 11 – reconstruction of dilated intersection within nucleus
- 12 – reconstruction of dilated intersection within cytoplasm

Rows:

- 1 - HbF antibody positive, DAPI positive cell and two HbF antibody positive, DAPI negative cells (magnification 20X)
- 2 - HbF antibody positive, DAPI negative cell (magnification 20X)
- 3 - HbF antibody negative, DAPI positive cell (magnification 20X)
- 4 - HbF antibody positive, DAPI positive cell (magnification 10X)

The finding algorithm is based on the following general assumptions: 1) the objects of interest are different in their intensity in one or more spectral components from the background; 2) the objects are compact and occupy a small fraction of the total image area. The first assumption allows the use of thresholding as the basis of discrimination of objects from the background. The second assumption in conjunction with the first determines the overall strategy, namely: 1) reduce as fast as possible the whole image to the areas containing potential objects; 2) analyze each of these areas as carefully as needed. Since the total of these areas is a small fraction of the whole image, the processing time will not increase significantly.

The specific feature of the cells of interest is the overlap or close proximity of antibody-positive and DAPI-positive areas. Figure 4 shows the main steps of analysis of areas, which contain potential NRBCs. Two separate thresholds are applied to the nuclear image (Fig. 4, Col. 4) and to the cytoplasmic image (Col. 3) producing corresponding masks (Cols. 5, 7). These masks are dilated¹⁵ (Cols. 6, 8) and intersected (Col. 9). The intersection itself is dilated (Col. 10) and reconstructed within the original masks of nucleus and cytoplasm, giving a nonempty result if the areas of DAPI and HbF antibody activity intersected or were close. The rows in Figure 4 show how the algorithm works on fetal NRBCs, RBCs and WBCs.

As described so far, the algorithm looks for one structural feature: overlap of DAPI fluorescence and HbF antibody absorption. This proved to be a robust feature missing few cells (<3%). The false positives are the overlaps of RBCs, WBCs, and debris. To separate NRBCs from false positives, the following parameters are used: average and integrated intensity in nucleus and cytoplasm, and area and shape of the nucleus and of the union of nucleus and cytoplasm. The shape parameters are: the ratio of principal axes, which measures elongation of an object, and compactness (defined as the ratio of the moment of inertia of a disk with the same area to the actual moment of inertia of the object). This gives a measure of irregularity of the object's shape.

The speed for scanning a 4 cm² (typical smear) is 1/2 hour (at magnification 10X). The average relocation error is 2 μm (4 μm maximum). The program finds 97.2% of the NRBCs found by a trained operator.

5. WORKFLOW FOR SLIDE ANALYSIS

The workflow for analysis of slides consists of three major steps: scan, review, and probe acquisition. The first step starts with a slide and produces a data file, which contains images and measurements of detected objects. In the second step, the operator reviews results of the scan, classifies and annotates the cells, and selects cells of interest for probe acquisition. The review is done primarily from the screen images, but the slide is available for immediate relocation and inspection under the microscope. Finally, probe images of the selected cells are acquired and the probe spots are counted.

To set up a scan the user performs the following actions:

1. Define the scan area to correspond to the actual smear location on the slide.
2. Select a representative field, and adjust the image digitization parameters for each of the filters used in scanning. (Sometimes the digitizer adjustment is done on calibration slides).
3. Run the finding algorithm on the selected field and adjusts the parameters if needed.
4. Set position of the reference point to provide for accurate relocation in the future.
5. Give a name to the scan, and start the scan.

The scan can be interrupted at any time to look at the objects and readjust parameters before resuming the scan.

After the scan is complete, the user returns for review. The main form of presenting information for review is in a data grid, which contains images, measured features, classification, and annotation. The objects can be revisited any number of times and each time a group of several images can be acquired for each object (for example, original images from scanning, same filter images at higher magnification, and probe images as in Fig. 5). Each group of images may be shown differently (e.g. in monochrome, in color, or both). The review process is aided by 2-D distribution plots of any two of the measured features. Each class in the distribution is shown in a different color and the display resembles that in flow cytometry. Switching of views among stored cell images, distribution of cells in the feature space, and microscope relocation is instantaneous. Another useful facility is sorting by linear combination of the measured features, which presents cells for review in the order of most likely candidates first. The result of the review is a classification and selection of cells for probe acquisition. At this time, the finding algorithms are designed to look for objects of one class (defined by the algorithm and its parameters), and there is no further automatic classification in the feature space. Classification is done manually by the user.

To set up probe acquisition the following must be defined:

1. Sequence of filters for the used fluorochromes.
2. Filter for focusing.
3. For each filter definition, whether to acquire a single plane, or a Z stack.
4. Z stack parameters (if used): number of planes, step between planes, background removal parameters.
5. Size of the image acquired for each cell.
6. Digitizer parameters for each filter.
7. Set of cells for probe acquisition.

After the set up, the process is automatic. The cells of interest are relocated and focused. For every cell, a set of monochrome images corresponding to the used fluorochromes is acquired according to the above parameters (the acquisition can also be done manually).

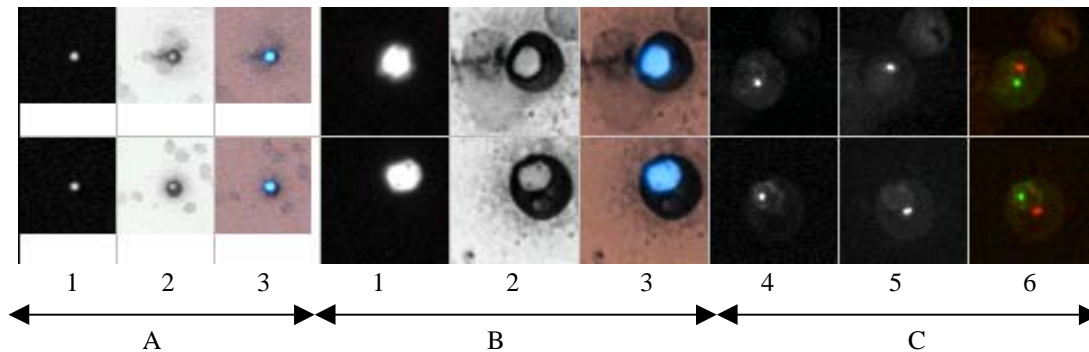


Figure 5. Images collected during scanning and subsequent multifilter acquisition on a maternal blood sample spiked with fetal cord blood.

A – Images of DAPI and HbF antibody collected in scanning. Magnification 10X.

B – Images of DAPI and HbF antibody collected later. Magnification 40X. This is usually not required and done for documentation purposes.

1 – Images of DAPI stained nucleus.

2 – Images of HbF antibody stained cytoplasm.

3 – Color composite of 1 and 2.

C – Images of X and Y probes. Magnification 40X.

4 – Y probe (Vysis, SpectrumGreen)

5 – X probe (Vysis, SpectrumOrange)

6 – Color composite of 4 and 5.

6. AUTOFOCUSING

Autofocusing is used in the system in two distinct situations: 1) During scanning, the whole field must be considered in calculation of the focus function. Typically, in scanning the starting point of focusing would be already close to focus – this is how the frequency of focusing is adjusted, but on sparse slides the algorithm may skip empty fields and drift far out of focus. 2) Autofocusing while revisiting a set of previously recorded objects. In this case the object is relocated in the middle of the frame and only a small part of the frame containing the object, can be used for calculation the focus function. The order of objects may not be according to their slide coordinates, and focal difference between the objects may be substantial. In addition, this is done at higher magnification than during scanning and the depth of field is shallower.

The second characteristic of focusing in this project is high noise. Much of the focusing is done in fluorescence on low light images. Special consideration was given to the development of a focusing algorithm, which can handle large defocusing and noisy images.

1. Field rejection.

Fields, which do not have enough contrast over background, are rejected. Background is estimated as a low frequency component in the image. After background removal, the image content is characterized by the difference between high and low quantiles of the histogram. Typically low quantile is at 1% and high quantile is at 99% of the image brightness distribution.

2. Contrast functions.

Use of two focus functions with better performance near focus and far from focus has been suggested previously⁶. The near range focus function is similar to difference criterion of Vollath⁹, squared gradient of Groen, Young and Ligthart⁶, and focus function of Shippey¹⁰. Summation is done on a subsampled (in X and Y) image in the X direction only, and only on pixels

$$F_1 = \sum_{i=1}^N \left\{ E(|g_i - g_{i-1}| - L) \right\}^2 / N, (g_i > T, g_{i-1} > T) \quad (1)$$

above the given threshold T. Optionally, the summation may be limited to the region of interest within the whole image. To avoid random noise, differences smaller than contrast limit L are forced to zero. $E(x) = x$ if $x \geq 0$; $E(x) = 0$ if $x < 0$. Subsampling in Y is determined only by speed considerations. Subsampling in X determines the frequencies to which the focus function is sensitive⁸. Threshold level T separates data from background and may be adjusted on every focusing field so that no fewer, than a set number of pixels, are available for calculation of focus function.

The far range contrast function is similar to the Mendelsohn and Mayall's focus function⁷. It is the average pixel value calculated on pixels above a given threshold: Summation follows the same rules, as in (1).

$$F_2 = \sum_{i=1}^N g_i / N, (g_i > T) \quad (2)$$

3. The following steps provide a simplified description of the focusing algorithm:

- Make a trial step. If the near-range focus function is below a predetermined limit, start using the far-range focus function. If the step was in the wrong direction, return to the original point.
- Set focus drive into motion in the direction determined by the trial step. The speed of focus motion is scaled down by magnification and camera exposure time. The speed of focus move is selected so that samples are separated in Z direction by steps equal or smaller than theoretical depth of field (for the example in Fig. 6, step is 0.5 μm).
- In the loop: read current focus position, acquire the image, and calculate both focus functions. If the near-range function exceeds the limit, switch to it. Keep moving according to the stop rule below.
- Return to the point of maximum or to the original point.

4. Stop rule

Presence of noise affects the rule, which determines when the focus search should stop. The whole range of focus contrast for the near-range function is divided into three intervals (Fig 6). Tests of over-travel limit and timeout limit are applied in all three intervals. [0-C1] is the interval where the near-range function is unreliable and focusing is guided by the far-range function. [C2- ∞] is the interval where the search stops when the near-range function drops by a given percent of its accumulated maximum. In the middle interval [C1-C2] there are two additional conditions for end of search. The purpose of these conditions is to eliminate stops on spurious local peaks of the focus function. The first condition is that no fewer, than a predetermined number of points must be analyzed. The second condition characterizes the shape of the focus curve. If a measure of oscillations, defined as the number of reversals of direction of the focus function to the number of focusing points, is high, it is a high noise situation and the search should continue.

5. Return rule.

After the search has been stopped, a decision must be made where to return and whether to consider autofocusing finished. If the maximum was reached, as defined in the stop rule, the algorithm returns to the point of this maximum and the procedure is finished. In case of timeout, it returns to the original point. This is an indication of some malfunctioning. In case of over-travel we distinguish two situations. If the achieved maximum is significantly higher than the starting point, return to the point of maximum. This typically means that the algorithm was on the right track, but the over-travel limit was smaller than the actual defocusing. If the achieved maximum is not significantly higher than the starting point, return to original point and try focusing in the opposite direction. This typically means that the direction, established by the first step, was wrong probably due to noise.

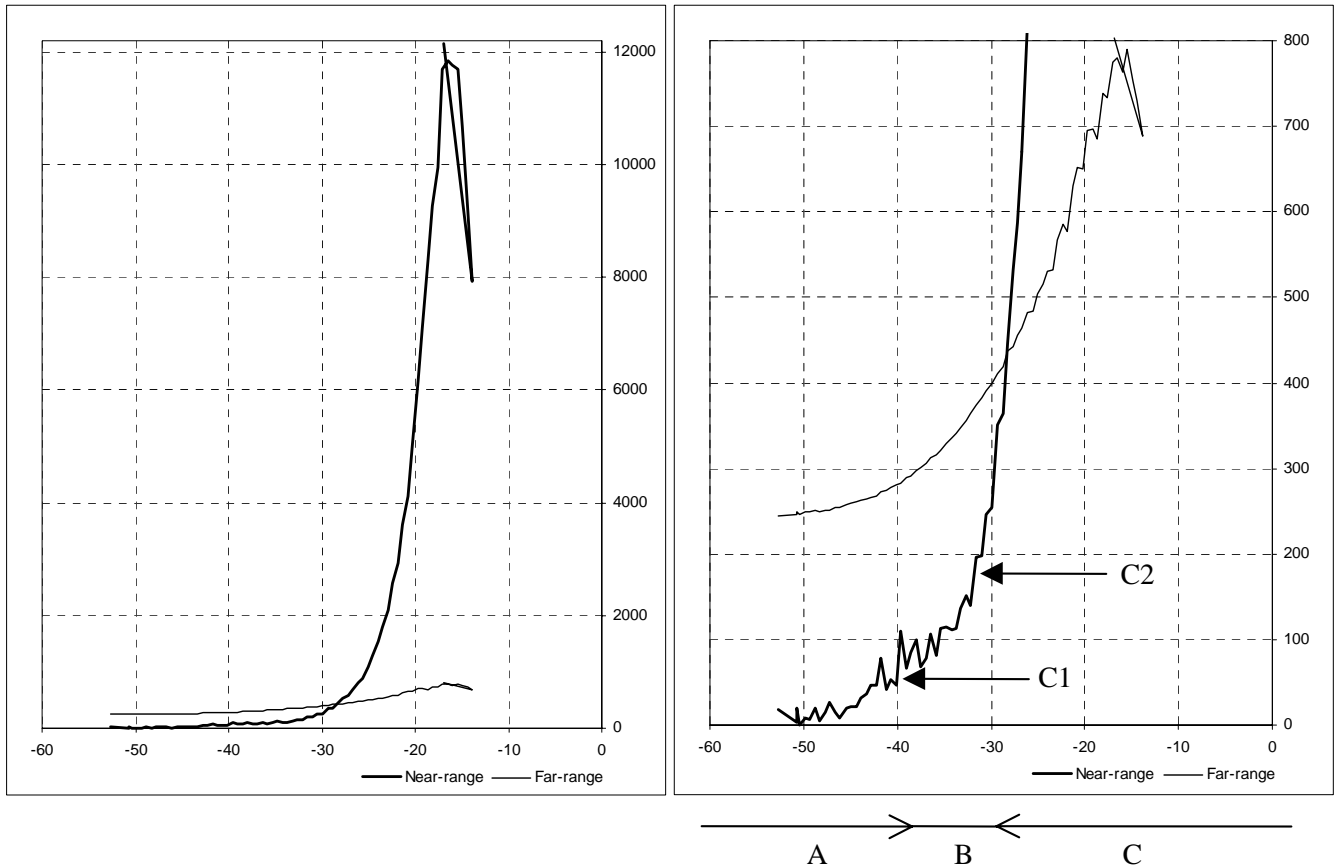


Figure 6. Example of focus functions estimated during autofocus. DAPI stained specimen of human blood. Olympus BX60 microscope, objective UplanFl 20X/0.50. Pixel size 0.65 μm . X - focus position in microns. Y - focus contrast. Sampling frequency in Z is 0.5 μm . Last segment shows return to optimal focus position. (Point of return is slightly different from maximum of contrast due to focus backlash compensation built into the algorithm) Left and right graphs show the same data in different scale for contrast. The graph shows noise of near-range function far out of focus, and noise of far-range function close to focus.

We analyzed surfaces formed by the autofocus positions on slides and found that they usually are a superposition of a smooth low order surface and random noise (Fig. 7). Main contributors are tilt of the stage, wedging of the slide and cushioning of the cover slip due to mounting media. While care is taken to reduce each of these factors, it is useful to reduce their effect computationally by dynamically estimating the focus surface and predicting focus position during scanning.

The focus prediction is based on a second order polynomial function:

$$z = p_{00} + p_{10}x + p_{01}y + p_{20}x^2 + p_{11}xy + p_{02}y^2$$

Coefficients p_{ij} are calculated by the least squares method from the points on which focusing was already done and are used to predict focus for the new stage position. Previous focusing points are weighted in reverse proportion to their distance

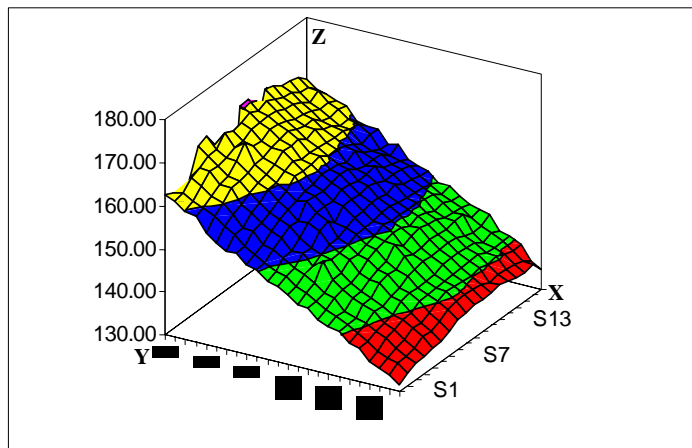


Figure 7. Focus Map. Best focus position was estimated in a grid of 24 (X) by 19 (Y) positions. Step in both directions 950 μm .

from the point of prediction. Coefficients p_{ij} are the solution of the following system of linear equations:

$$\mathbf{W} \cdot \mathbf{M} \cdot \mathbf{p} = \mathbf{W} \cdot \mathbf{Z} \quad (3)$$

Where: N is the number of previous focusing points; $\mathbf{p} = (p_{00}, p_{10}, p_{01}, p_{20}, p_{11}, p_{02})$ is a 6×1 vector of the polynomial coefficients; \mathbf{M} is a $N \times 6$ matrix with rows: $(1, x_i, y_i, x_i^2, x_i y_i, y_i^2)$, $i = 1, \dots, N$; \mathbf{W} is a $N \times N$ diagonal matrix of weights: $w_{ii} = 1/(d_0 + \text{distance}((x_i, y_i), (x_{\text{new}}, y_{\text{new}})))$. We used the Euclidean distance function, parameter d_0 was adjusted empirically. The overdetermined system of linear equations (3) was solved through the normal equations¹¹.

7. ACQUISITION OF FISH IMAGES

Smears of blood used in this study have a thickness of 2-3 μm . Typically for FISH imaging we use 40X/0.75 objectives. Depth of field is about 1 μm . Under these conditions, probe spots that differ in the Z plane but are at the same XY position in the cell are probably unresolvable. On the other hand, a single plane is often not enough for adequate presentation of all spots in a cell (Fig. 8). Based on this, our approach is to create a single image in which each spot would be represented as if it were in the best focus, in other words, computationally increase depth of field¹² or form a synthetic projection¹³. The probe acquisition algorithm has the following steps:

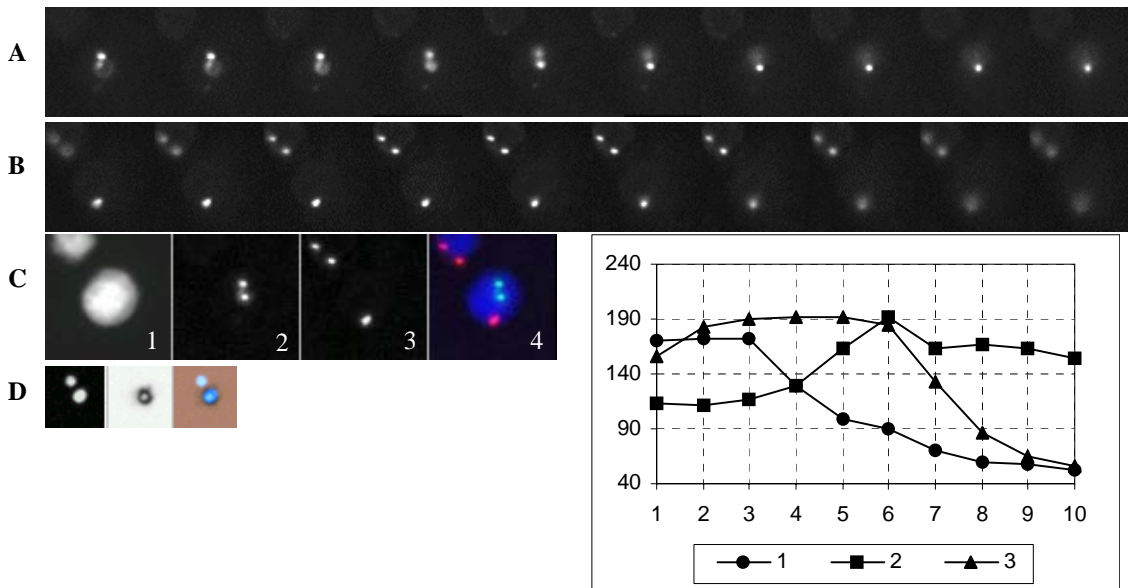


Figure 8. Z stack acquisition and processing.

A,B,C - Images of an NRBC, taken with Olympus UPlanFl 40X/0.75 objective.

A,B - series of images at 0.5 μm separation between Z planes.

A - Y probe (Vysis, SpectrumGreen), B - X probe (Vysis, SpectrumOrange).

C: 1 - DAPI stain; 2, 3 - result of processing of Z stacks A and B; 4 - composite image.

D - images of DAPI and HbF antibody of the same cell at 10X.

Plots of maximum intensity of probe spots in the same cell, according to Z position. Points correspond to image stacks: 1 - top spot in A stack, 2 - bottom spot in A stack, 3 - bottom spot in B stack.

1. Autofocus on nuclear counterstain
2. For each probe fluorochrome, acquire a Z stack with a small (4-7) number of planes
3. For each plane, remove the low frequency background
4. Combine all planes in a stack after background removal by maximum operation.

The role of a stack of images is not only to account for spots possibly lying in different planes, but it also accounts for differences in focal position between DAPI and probes.

8. COUNTING OF DNA PROBE SPOTS

The algorithm counts the number of FISH signals in a nucleus. Currently, counting of spots is done independently for each fluorochrome. This will be expanded to consider in one algorithm spots in all acquired spectral bands. Combining all of the information in one step will eliminate from the count nonspecifically fluorescing debris. It is assumed that the nucleus of the cell is known, either from a separate image of nuclear counterstain, or can be extracted from the probe image.

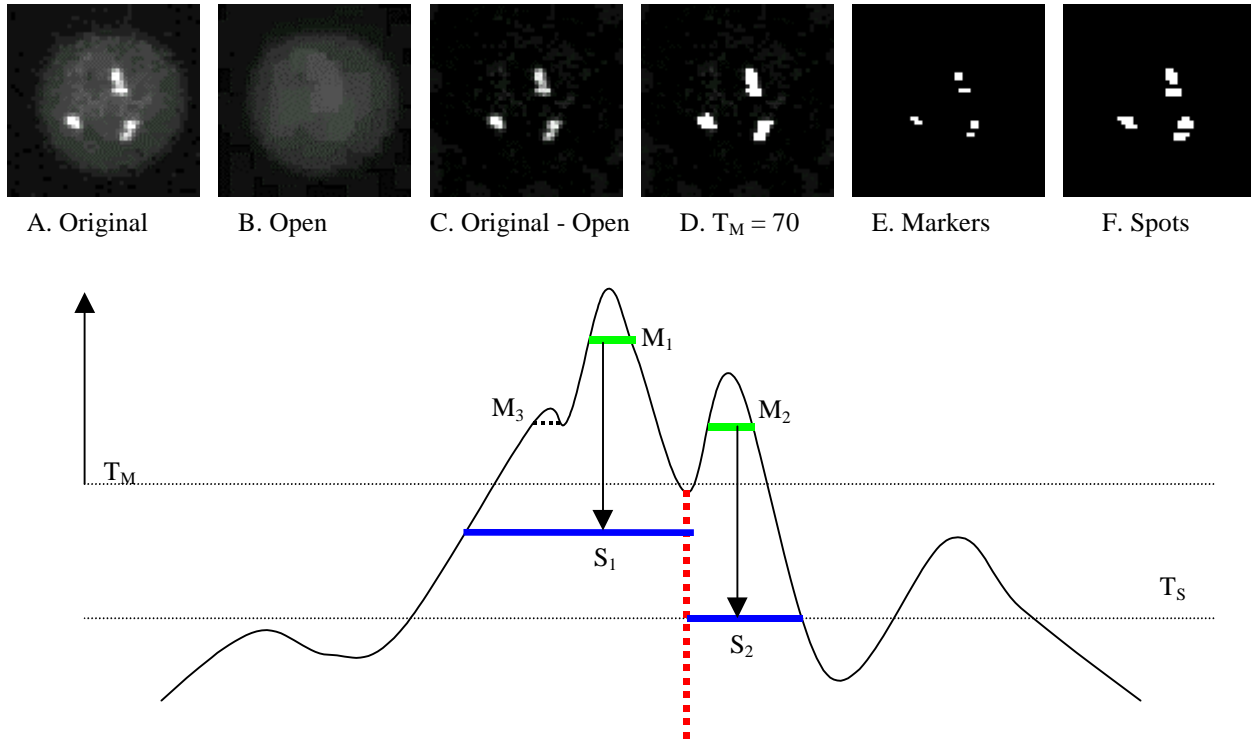


Figure 9. Counting of fluorescent probe spots.

A - Original image of probe spots of X chromosome in cultured lymphocytes, slides from Vysis. Objective - Olympus UPlanFl 40X/0.75.

B - Opening of A removes spots smaller than the kernel size of opening operation. Here, kernel radius is 2.

C - Subtraction of open image B from original A gives spots of interest on a uniform background.

D - Mask at marker threshold T_M . Segmentation tree is constructed starting with this threshold up.

E - Markers - leaves of the segmentation tree not smaller, than A_{Min} .

F - Spots - grow down from markers as watershed, constrained by spot threshold T_S , and A_{Max} .

The algorithm consists of the following steps:

1. Preprocessing
2. Formation of the segmentation tree and selection of markers
3. Growing of spot markers
4. Postprocessing to apply rules for split and merged spots and for rejection

1. Preprocessing (see Fig. 9) is essentially the same as described by Netten et al.¹⁴ Opening¹⁵ of the original image with an appropriate size structuring element removes spots of interest. The size of the structuring element is a parameter of the algorithm and is scaled by magnification. This background image is subtracted from the original.

2. The segmentation tree starts to grow up from the marker threshold T_M which is set high enough to exclude undesired areas but low enough to include at least a few pixels of all potential spots. For each connected set, the threshold is raised until the set splits into several new sets or disappears. If a new set has area smaller than A_{Min} it is not considered as a potential marker and its branch of the tree is stopped (Fig. 9, M_3). A_{Min} controls the size of the detail of the brightness distribution, which is considered "noise". These areas will be considered only as part of stronger spots. At the end of the procedure, leaves of the

tree are spot markers. To speed up the next step, the actual markers are made by a section of the tree, which gives them maximal area without merging and not exceeding A_{Max} .

3. From the markers, the spots are grown down in a procedure, which is essentially watershed¹⁶ constrained by two parameters: A_{Max} and T_S . A_{Max} is maximum allowed area for a spot; T_S is the threshold until which a spot can grow. In Fig. 9 growth of S_1 is stopped by area and growth of S_2 - by threshold.

4. A postprocessing step is needed to account for possible artifacts of cell hybridization. This part of the algorithm is not finished and will follow the approach of Netten et al.¹⁴ Area and intensity features will be computed and rules will be formulated for split and merged spots. We have produced a gallery of difficult to score cells and have given it as images to investigators at different laboratories. The results of counting are so different that we will not try to pursue "correct counting", but rather formulate rules for rejection of cells. These cells, however, will be available for manual scoring.

9. REFERENCES

1. D.W. Bianchi, "Prenatal diagnosis by analysis of fetal cells in maternal blood", *The Journal of Pediatrics*, **127**, pp. 847-858, 1995
2. D.E. Lewis, W. Schober, S. Murrell, D. Nguyen, J. Scott, J. Boinoff, J.L. Simpson, F.Z. Bischoff, S. Elias, "Rare event selection of fetal nucleated erythrocytes in maternal blood by flow cytometry", *Cytometry*, **23**, pp. 218-227, 1996
3. M.C. Cheung, J.D. Goldberg, Y.W. Kan, "Prenatal diagnosis of sickle cell anemia and thalassaemia by analysis of fetal cells in maternal blood", *Nature Genetics*, **14**, pp. 264-268, 1996
4. D.W. Bianchi, J.M. Williams, L.M. Sullivan, F.W. Hanson, K.W. Klinger, A.P. Shuber, "PCR quantitation of fetal cells in maternal blood in normal and aneuploid pregnancies", *Am. J. Hum. Genet.* **61**, pp. 822-829, 1997
5. A.M. Saunders, M.A. Zarowitz, P.J. Baldwin, "Method for separating rare cells from a population of cells", US Patent 5,432,054
6. F.C.A. Groen, I.T. Young, G. Ligthart, "A comparison of different focus functions for use in autofocus algorithms", *Cytometry* **6**: pp. 81-91, 1985
7. M.L. Mendelsohn, B.H. Mayall, "Computer-oriented analysis of human chromosomes - III. Focus", *Comput. Biol. Med.* **2**: pp. 137-150, 1972.
8. F.R. Boddeke, L.J. van Vliet, H. Netten, I.T. Young, "Autofocusing in microscopy based on the OTF and sampling", *Bioimaging* **2**: pp. 193-203, 1994
9. D. Vollath, "The influence of the scene parameters and of noise on the behavior of automatic focusing algorithms" *J. Microsc.* **151**: pp. 133-146, 1988
10. G. Shippey, "Automatic focussing device", US Patent 4636051
11. W.H. Press, S.A. Teukolsky, B.P. Flannery, *Numerical recipes in C: the art of scientific computing*, 2nd Ed., Cambridge University Press, 1992
12. W.J. Hannan, "Imaging system with enlarged depth of field", US Patent 4404594
13. D.A. Agard, Y. Hiraoka, P. Shaw, J.W. Sedat, "Fluorescence microscopy in three dimensions", *Methods in Cell Biology*, **30**, pp. 353-377, 1989
14. H. Netten, L.J. van Vliet, H. Vrolijk, W.C.R. Sloos, H.J. Tanke, I.T. Young, "Fluorescent dot counting in interphase cell nuclei", *Bioimaging* **4**: pp. 93-106, 1996
15. J. Serra, *Image Analysis and Mathematical Morphology*, Vol. 1. Academic Press, London, 1989
16. S. Beucher and F. Meyer, "The Morphological Approach to Segmentation: The Watershed Transformation" in: *Mathematical Morphology in Image Processing*, E.R. Dougherty - Ed., pp. 433 - 481, Marcel Dekker, New York, 1993