Image Processing for Automated Analysis of the Fluorescence In-Situ Hybridization (FISH) Microscopic Images

Jan Schier[1,*], Bohumil Kovář[1], Eduard Kočárek[2], and Michal Kuneš[1]

[1] Institute of Information Theory and Automation of the ASCR,
Pod vodárenskou věží 4
182 08 Prague 8, Czech Republic
`{schier,kovar,xkunes}@utia.cas.cz`
[2] Department of Biology and Medical Genetics
2nd Faculty of Medicine
Charles University Prague
V Úvalu 84
150 06 Prague 5, Czech Republic
`eduard.kocarek@lfmotol.cuni.cz`

Abstract. The paper describes automated segmentation and analysis of the microscopic images resulting from fluorescence in-situ hybridization (FISH) analysis. FISH is a popular molecular cytogenetic method. The output of a single FISH analysis is a set of several tens or hundreds microscopic images — a single evaluated sample is of roughly 20mm diameter. The goal of an automated evaluation is to replace the subjective evaluation of images by the laboratory technician to achieve higher uniformity of results. Following explanation of the principle of the method and the typical contents of images, the processing flow of image segmentation is outlined and the results are presented on several example images. With emphasis on a low-cost solution, the ITK library is used for implementation.

Keywords: fluorescence in-situ hybridization (FISH), image processing, image segmentation.

1 Introduction

The fluorescence in situ hybridization (FISH) method enables us to detect genetic changes both in mitotic (dividing) and interphase (preparing to divide) cells (see Figure 2). It belongs to the most common of molecular cytogenetic methods and is widely applied in routine clinical genetic diagnostics.

The FISH technique concerns application of a fluorescently labeled DNA or RNA probe onto a biological sample containing the target DNA or RNA from the patient (e.g. as tissue sections, separated cells, isolated interphase nuclei

---

* Corresponding author.

or chromosomal metaphase spreads). In the process of hybridization, the probe attaches to its complementary sequences on the target DNA or RNA molecule. Under the fluorescence microscope we observe the sequence with the attached probe as a strikingly luminous hybridization signal (a fluorescent dot or an area within the biological structure). If the corresponding sequence is missing then no hybridization signal is present (for a detailed description of the FISH protocol see [4]).

Currently FISH is not only widely applied in clinical genetic diagnostics, but it is a useful diagnostic tool in reproductive medicine, oncology, preventive medicine, and other medical branches ([4], [18]). Formerly FISH significantly contributed to the successful mapping of the human genome [1].

We are paying attention to FISH analysis of chromosomal aneuploidies – i.e., deviations from chromosomal number. These abnormalities are responsible for well-known clinically significant syndromes such as Down's syndrome (trisomy 21), Turner's syndrome (monosomy X), Klinefeter's syndrome (XXY constitution) – see [8], [15], [3], and [19]. They are also frequently associated with cases of miscarriages (monosomy X and trisomy 16 belong to the most frequent genetic findings in spontaneously aborted fetuses – see [6]) and tumor diseases (as a result of non-disjunction during rapid and unregulated mitotic division of a cancer cell).

In these cases FISH findings are enabling us to conclude or to verify the genetic diagnosis (in postnatal examinations) or prognosis of pregnancy (in prenatal examinations). In oncology it contributes to an estimation of the patient's prognosis and to the selection of an appropriate therapy (some tumors with specific chromosomal aberrations are susceptible to certain cytostatic drugs or to a special combination of them and radiation therapy – see [2]).

1.1   FISH and Fluorescence Microscopy

In the aforementioned, the samples with hybridized probes are analyzed using an epifluorescence microscope. Good summary of the technical principles of fluorescence microscopy can be found in [14]. Its principle is the fluorescence of the dye, present in the sample: the UV illumination excites the electrons in the molecules of the fluorescent dye. When the electron returns back to the ground state, then a photon with lower energy, one with longer wavelength, is emitted. For practical reasons, the fluorophores are selected so as to emit light in the visible spectrum. An assembly of *filters* combined with a *dichroic mirror* (referred to as *filter cube*) allows both the UV illumination and the fluorescence signal to pass through the objective, while only the fluorescence signal emitted from the dye is allowed to pass through the eyepiece. A schematic drawing of the epifluorescence microscope is given in Figure 1 (as adapted from Henry Mühlpfordt, Wikipedia.org)

After passing through the optical system, the image is typically captured by a digital camera. For final evaluation of hybridization, an objective with $100\times$ magnification is used (see [10] for more details). A single image then covers only a
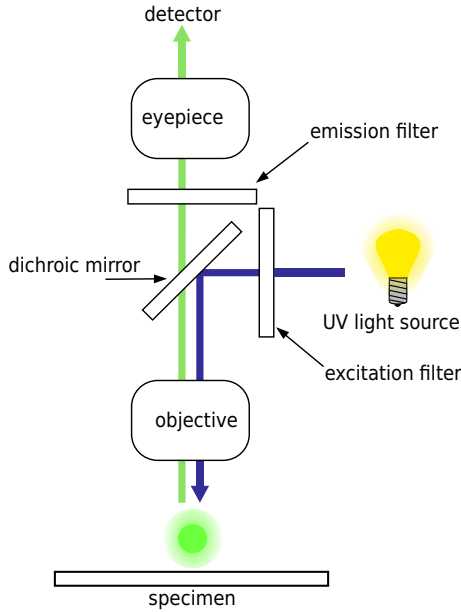
Fig. 1. Epifluorescence microscope scheme

small fraction of the sample (which has approx. 20mm in diameter — dependent on the preparation process), resulting in a set of tens or hundreds of images that have to be evaluated to be able to conclude examination of a single patient.

1.2   Motivation for the Automated Processing of Images

An image set of this size requires automated processing to achieve results of acceptable accuracy and reproducibility. While commercial systems for evaluation of FISH analyses exist, we believe that there is still space for further research. The reasons, in our opinion, are of both an academic research and technology insertion nature (rapid development of new cytogenetic methods, rapid development of segmentation and data mining methods in the past decade). Also, research in this field is motivated by the economics of open source software — on the one hand, the high cost of commercial tools is a barrier to their adoption in a non-commercial laboratory, on the other hand, high-quality open-source image processing tools are available for free (examples are the ITK (`http://www.itk.org`) or the OpenCV (`http://opencv.willowgarage.com/wiki/`) libraries).

   This paper is focused on the segmentation of the images, that is, on finding the objects of interest — the cell nuclei and the hybridization signals. Section 2 describes the typical features of the images resulting from the FISH analysis. Section 3 outlines the image processing flow and the segmentation methods. Section 4.1 describes the edge-based segmentation, section 4.2 the level-set based segmentation. Section 5 concludes the paper.

2   FISH Image: Composition and Contents

Example images from the FISH analysis are shown in Figure 2. They show the interphase and metaphase nucleus — two basic forms of the cell nucleus, which correspond with two basic phases of the so-called cell cycle.
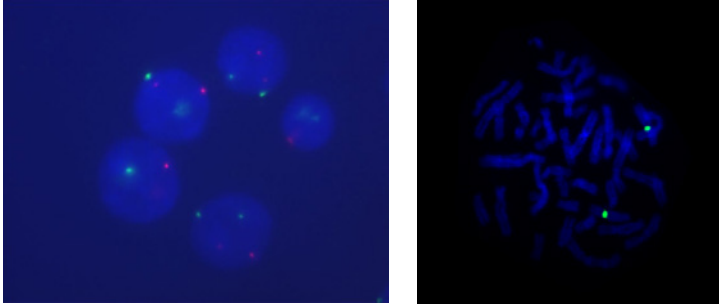
(a) Interphase cells   (b) Metaphase cell

Note: the images are not in the same scale. The metaphase cell image has been preprocessed to suppress the background.

Fig. 2. Example FISH images

The image of the interphase cells shows the cell nuclei, each of them containing the hybridization signals, in this case in red and green colour. The green dots represent the X-chromosome, red ones the Y-chromosome. Blue ares are the nuclear structures counterstained with the DAPI dye. The image of the metaphase cell shows dividing cell nucleus — the spread of chromosomes can be seen. Let us note that very often, both interphase and metaphase cells are contained in the same image.

The image is, in fact, displayed in artificial colours. — it is composed of several image layers (greyscale images). Each layer image is taken with a different filter set (see the scheme of microscope, Figure 1), so its actual colour depends on the particular emission filter used. Each filter set is tuned for different emitted light wavelength; one is used to take the image of the cell nuclei (and the image background), the others to take the images of the hybridization signals of different colors (for more detailed discussion of the fluorescent dyes and filters, see e.g. [14], Chapter 2).

The layers composing the image in Figure 2a are, for illustration, shown in Figure 3. For the metaphase cell, the layers would look alike.

2.1   Image Contents

It can be seen from the above examples that the objects in the image are either solid objects of roughly circular shape (the case of the interphase nuclei) or roughly circular "clouds" — the spread of the ribbon-like chromosomes

(a) nuclei   (b) signals (green filter)   (c) signals (red filter)
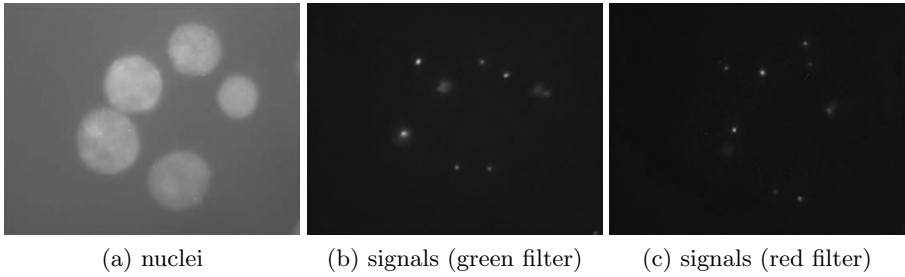
Fig. 3. Interphase cells: image layers

(metaphase nuclei). The hybridization signals are contained within the nuclei or on the individual 'ribbons'.

This description represents an ideal image. Most often, however, there are defects present. While the classification of these defects depends on the protocol used in the particular laboratory, the general guidelines may be found in[4]. They include especially:

- irregularities of the background intensity
- irregularities of the nuclei shapes
- fluorescent signals out of nuclei (in the background)
- hybridization signals in the area of nuclei overlap (impossible to decide what nucleus the signal belongs to)
- hybridization signals too close to each other

The system for FISH evaluation has to be able to detect these defects to either eliminate them, or where possible, to compensate for them.

3   System Processing Flow

A basic scheme of data processing flow in the system for FISH analysis is given in Figure 4. In the image processing block, the objects of interest, as introduced in the previous section, are located and checked for the defects. For the objects that satisfy the criteria, the parameters such as roundness, mean intensity, size, etc. are determined and passed to the next block, where statistics are computed and data mining is performed. The final operation is the visualization of the output of both image and statistical processing, as needed. For simplicity and without loss of generality, the scheme does not include blocks of the kind: microscope control, image database, etc.

4   Image Segmentation

Image segmentation has gained importance as key to the field of automatic analysis of medical images. Segmentation algorithms can be historically clustered as belonging to one of these three development periods [20]:
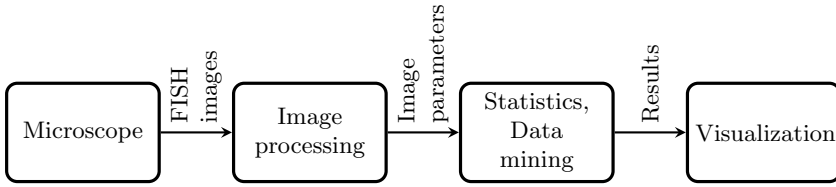
Fig. 4. Basic block scheme of a system for FISH analysis

The earliest methods use only low-level processing techniques without any apriori knowledge. These methods include particularly thresholding, region-growing and edge tracing algorithms. The accuracy of segmentation is mainly influenced by heuristics that help to initialize segmentation. While the segmentation results using these methods are insufficient for complex medical images, they are often used to initialize the more advanced methods.

The second generation of algorithms is based on image models, optimization methods and uncertainty models. A characteristic feature is the absence of the aforementioned heuristics of the earliest methods. From our perspective, deformable models are one of the most important techniques of the second generation. Energy minimization models [9] (active contours) or front propagation methods based on level-sets [16] are common types of deformable models. Although the second generation of segmentation techniques is very well applicable to medical image segmentation, these techniques are not always suitable for accurate and automatic segmentation.

The third generation of algorithms extends previous methods by using additional knowledge about the segmented objects (e.g. statistical shape models) [7],[5]. We assume that for the segmentation of cell nuclei we will use segmentation techniques from first and second generation, namely edge tracing segmentation and segmentation based on deformable models.

4.1   Segmentation Algorithm Based on Edge Tracing

Cell nuclei detection

The first segmentation algorithm that we used in this project is based on edge tracing. The main part of the method uses a border following algorithm introduced by Suzuki [17]. Border following is one of the fundamental techniques in the processing of binary images. It derives a sequence of the coordinates or the chain codes from the border between a connected component of pixels that correspond to objects (pixel value = 1) and a connected component of pixels that belong to background or hole (pixel value = 0). The binary image is first converted into the border description and then topological structures of the image are extracted. These structures are usually described by contours (Fig. 5), which represent the segmented objects.

From the description above, it is clear that the algorithm used for image binarization will have a major impact on the quality of segmentation. Among
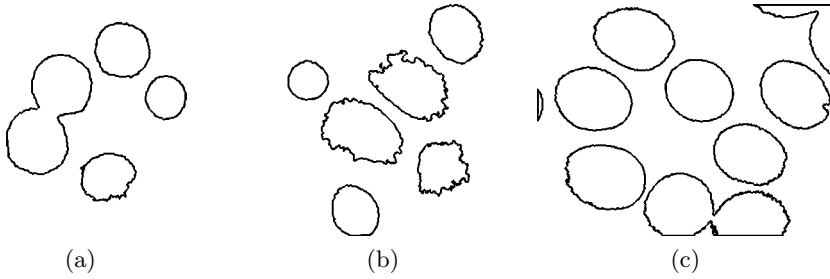
Fig. 5. The results of the border following algorithm described by contours

the many existing algorithms, we have chosen the well-known Otsu algorithm, introduced in 1979 by Nobuyuki Otsu [13]. Using discriminant analysis, Otsu defined the between-class variance of the thresholded image as

$$\sigma_\omega^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t), \tag{1}$$

where ω_i are the probabilities of the two classes separated by a threshold t and σ_i^2 are variances of these classes. For bi-level thresholding, Otsu verified that the optimal threshold t^* is chosen so that the between-class variance $\sigma_\omega^2(t)$ is maximized; that is,

$$t^* = \mathrm{argmax}(\sigma_\omega^2(t)), \quad 1 \le t \le L, \tag{2}$$

where L is the highest gray-level intensity value in the image. This nonparametric and unsupervised method is well suited for the reduction of a graylevel image to a binary image due to an automatic threshold selection.
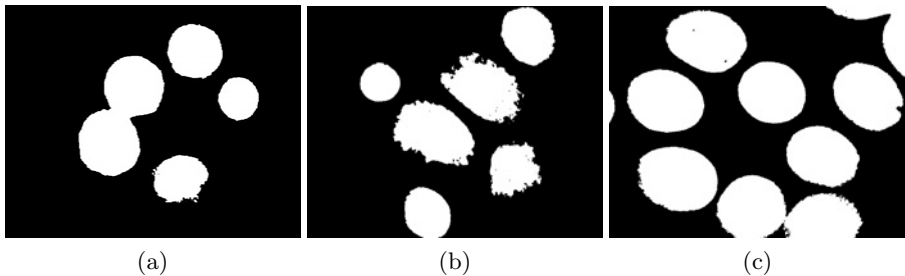

Fig. 6. The results of image binarization based on the Otsu thresholding

Description of the segmented objects using closed contours only is not convenient for further processing. We approximate the detected closed contours by the ellipse or the convex-hull (Fig. 7). The detected cell nuclei, their contours, convex hull and elliptical approximation are then used to calculate the relevant statistical/shape characteristics such as centroid, average intensity, standard deviation of intensity, eccentricity, roundness etc. Subsequently, we label each detected nucleus with all the necessary parameters for further statistical processing. Only

the nuclei with circular or slightly elliptical shape are suitable for further processing, so we do not evaluate those with eccentricity $e > 0.6$. The eccentricity e (3) is defined as the relative difference in magnitude of the eigenvalues λ_i (4) of the covariance matrix of the image I.

$$e = \sqrt{1 - \frac{\lambda_2}{\lambda_1}}, \tag{3}$$

$$\lambda_i = \frac{\mu'_{20} + \mu'_{02}}{2} \pm \frac{\sqrt{4\mu'^2_{11} + (\mu'_{20} - \mu'_{02})^2}}{2} \tag{4}$$

where μ'_{pq} are second order central moments.
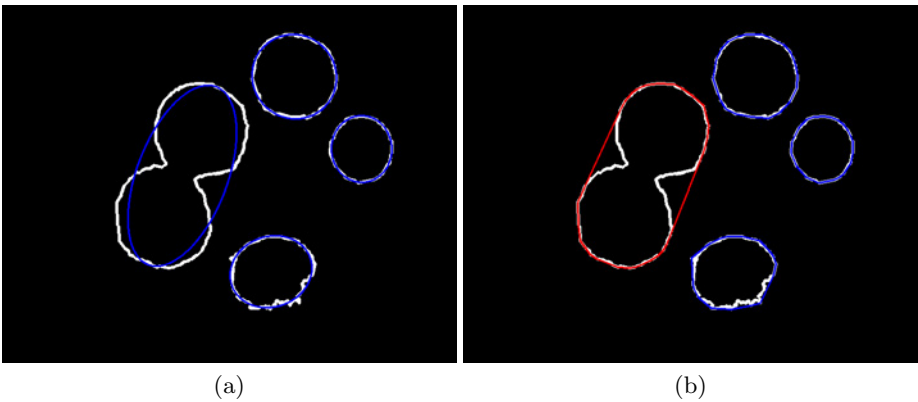

(a)   (b)

Fig. 7. Cell nuclei approximated by (a) ellipse and (b) convex hull. The two nuclei marked in red are touching each other — the eccentricity of the ellipsoid covering both of them is greater than 0.6.

Hybridization signals detection

Hybridization signals are detected by the same algorithm as the cell nuclei, but in other layers of the input image (Fig. 3). Thus, there is no need to change the segmentation algorithm or its parameters for the processing of signals. This can be considered as the main advantage of this approach. Illustrative results are shown in the Figure 8.

4.2   Level-Set Based Segmentation

Image segmentation based on the technique described in the previous part is very simple and suitable only for "ideal" samples. Given that we are trying to create a robust system for processing and analysis of FISH images, we know that

it will be necessary to use advanced segmentation techniques. For this reason, we began to test the level-set segmentation algorithms. The method described in the next section is based on an active contour formulation with shape priors developed by Leventon *et al.* [11] for nuclei segmentation. The implementation of this method is available in the ITK library for medical imaging.

Segmentation methods based on boundary finding (Section 4.1) are generally local algorithms that require some feature (the edge, for example) to be present along the boundary of the object. These methods may be sensitive to the initialization and may "leak" through the nucleus boundary if the edge feature is not salient enough in a certain region in the image.

Leventon stated the evolution by the following pde [12]:

$$\frac{d\Psi}{dt} = f(I)(\alpha c + \beta \kappa)|\nabla \psi| + \gamma \nabla f . \nabla \Psi + \delta(\Psi^* - \Psi) \tag{5}$$

where Ψ is the level-set function and the zero level-set corresponds to the segmentation curve/surface, α is an expansion weight, β is a smoothness parameter and γ is a boundary attraction term. The first terms include propagation, curvature and advection terms. The final term $(\Psi^* - \Psi)$ is a recent addition by Leventon *et al.* [11] wherein they incorporated principal components of the segmentation shape model to drive the update equation. The surface Ψ^* is the maximum aposteriori shape given the current segmentation and image information. The parameters α, β, γ are user-defined settings for the relative scaling of the three speeds. We typically want the level-set to fit our object and not escape out, hence, we specify higher values of γ and β relative to α.

The level-set evaluation separates the foreground into the individual cell nuclei. Their evolution is controlled by the speed function. In order to resolve the affiliation of small fragments Mosaliganti *et al* [12] assign pixels in a given connected component to the nearest cell available in that component.
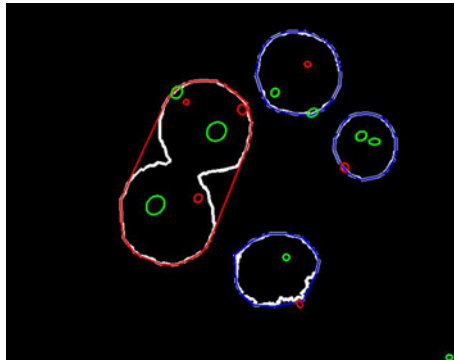

Fig. 8. Detected cell nuclei with hybridization signals. Note that some of the detected signals do not correspond to the real signals in the sample.
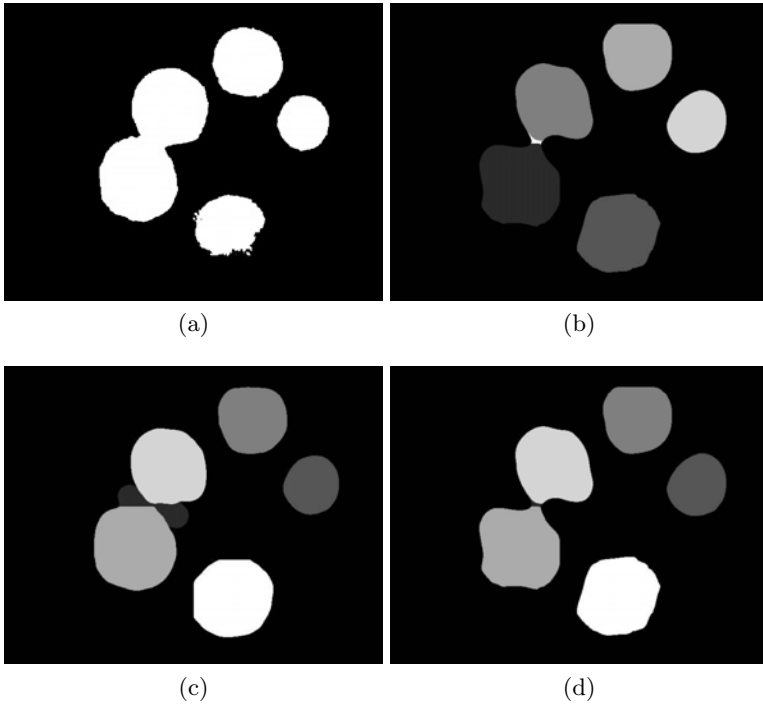
Fig. 9. Segmentation comparation: (a) Otsu thresholding, (b) multi-threshold binarization, (c) level-sets, (d) Voronoi level-set refinement

5   Conclusions

In the paper, we have discussed the evaluation of the microscopic images resulting from the fluorescence in-situ hybridization (FISH) analysis. The typical objects of interest contained in these images have been described. Several image segmentation methods of various levels of complexity have been reviewed, with their output demonstrated on example images.

Two points should be emphasized here: our project on automatic FISH image processing and evaluation is in a very early stage of development at this moment in time. It is why only few images are presented, instead of statistical evaluation of the methods. The shape deviations of the level set segmentation output need to be analyzed and explained. There are also many topics in the FISH image processing that we have not discussed at all, such as the recognition of interphase and metaphase nuclei, quality evaluation of the objects in the image, etc. Nevertheless, at least one important conclusions can be reached, namely, that automatic segmentation and evaluation of FISH images is possible with modern segmentation methods.

In future work, we want to address the accuracy of segmentation and the detection of nuclei of irregular shape, of overlapping nuclei, etc., required to achieve acceptable robustness of the system. Last but not least, the number of image samples in evaluations has to be increased substantially.

References

1. Adinolfi, M., Crolla, J.: Nonisotopic in situ hybridization. clinical cytogenetics and gene mapping applications. Adv. Hum. Genet. 22, 187–255 (1994)
2. Albertson, D., Collins, C., McCormick, F., Gray, J.: Chromosome aberrations in solid tumors. Nature Genetics 34(4), 369–376 (2003)
3. Annerén, G.: Turner syndrome. The American Journal of Human Genetics 68(3), 2–9 (2007)
4. Beatty, B., Mai, S., Squire, J.: FISH. Oxford University Press, Oxford (2002)
5. Beichel, R., Bischof, H., Leberl, F., Sonka, M.: Robust active appearance models and their application to medical image analysis. IEEE Transactions on Medical Imaging 24(9), 1151–1169 (2005)
6. Benn, P.: Trisomy 16 and trisomy 16 mosaicism: a review. American Journal of Medical Genetics 79(2), 121–133 (1998)
7. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - their training and application. Comput. Vis. Image Underst. 61, 38–59 (1995)
8. Harper, P.S.: Practical genetic counselling, 6th edn. Arnold, London (2004)
9. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 1(4), 321–331 (1988)
10. Kumar, G.L., Zucker, R.M.: Fluorescence In Situ Hybridization (FISH) Imaging, 5th edn., ch. 13, pp. 81–95. Dako (2009)
11. Leventon, M.E., Grimson, W.E.L., Faugeras, O.D.: Statistical shape influence in geodesic active contours. Computer Vision and Pattern Recognition, 1316–1323 (2000)
12. Mosaliganti, K., Gelas, A., Gouaillard, A., Megason, S.: Microscopy image analysis: Blob segmentation using geodesic active contours. The Insight Journal (June 2009), paper Id=321
13. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9, 62–66 (1979)
14. Rittscher, J., Machiraju, R., Wong, S.T.C. (eds.): Microscopic Image Analysis for Life Science Applications. Artech House, Inc., Boston (2008)
15. Roizen, N., Patterson, D.: Down's syndrome. The Lancet 361(9365), 1281–1289 (2003)
16. Sethian, J.: Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge (1999)

17. Suzuki, S., Abe, K.: Topological structural analysis of digitized binary images by border following. Computer Vision, Graphics, and Image Processing 30(1), 32–46 (1985)
18. Varella-Garcia, M.: Molecular cytogenetics in solid tumors: Laboratorial tool for diagnosis, prognosis, and therapy. The Oncologist 8(3), 45–48 (2003)
19. Visootsak, J., Graham, J.: Klinefelter syndrome and other sex chromosomal aneuploidies. Orphanet Journal of Rare Diseases 1(42) (2006), biomed Central
20. Withey, D., Koles, Z.: Medical image segmentation: Methods and software. In: Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, NFSI-ICFBI 2007, pp. 140–143 (October 2007)