Automatic Coregistration, Segmentation and Classification for Multimodal Cytopathology

Thomas Würflinger[a], Jens Stockhausen[a], Dietrich Meyer-Ebrecht[a], Alfred Böcking[b]

[a]Institute for Measurement Techniques and Image Processing, RWTH Aachen, Germany
[b]Institute for Cytopathology, Heinrich-Heine-University, Düsseldorf, Germany

Abstract

The paper describes the key component of the Multimodal Cell Analysis approach, a novel cytologic evaluation method for early cancer detection. The approach is based on a repeated staining of a cell smear. The correlation of features and data extracted from different stainings, and related to relocated individual cells, will yield a dramatic increase of diagnostic reliability.

The necessary fully automatic preprocessing steps are presented: coregistration of multimodal images, segmentation, and classification of cell nuclei. Both efficiency and robustness of all steps reached at the current stage of research, are high regarding medical image material, and strongly support clinical application.

Keywords:

Cytopathology; Cancer Diagnosis; Cell Microscopy; Multimodal Coregistration; Image Segmentation; Active Contour Models; Classification.

1. Introduction

The widely applied histopathological method for cancer diagnosis requires tissues for microscopic investigation. These are removed from the patient by bloody biopsies or operation. In contrast, cytopathologic investigations only require cells, obtained non-invasively and without discomfort. Sediments from diverse body fluids, smears from mucosal surfaces or fine needle aspirates from different organs contain enough cells to establish valid cancer diagnoses. Thus the compliance of patients to have lesions investigated, if they are suspicious for cancer, is higher for cytopathology than for histopathology. Furthermore, investigation of cells, as opposed to tissues, often allows an earlier diagnosis of cancer and therefore higher rates of cure. Yet, there are at least two obstacles to a broad application: the subjective microscopic evaluation of cytologic specimens is too time-consuming, and mere morphologic inspection of cells does often not offer enough information to achieve sufficient diagnostic accuracy.

Yet, the combination of different adjuvant diagnostic methods, like DNA-measurements, immunocytochemical marker demonstrations or AgNOR-analysis would allow accurate cancer diagnosis on a few cells [2]. In our interdisciplinary research we introduced the new approach of *Multimodal Cell Analysis*, which circumvents the limitations of current evaluation methods. Fixation of cells on microscope slides allows for repetitive staining, with intermediate application of a solvent. Interesting positions on the slides are retrieved under the microscope. Images from identical cells but different stainings are then acquired. Once exactly aligned, cell nuclei can be multimodally segmented and their cell types can be classified, after which the information retrieved from each staining can be evaluated and combined for each cell in order to gain a multivariate statistic. This leads to a considerable

increase in diagnostic accuracy and allows for valid diagnoses based on a very few cells.

At the reported stage, our objective was to prove that the automatic image processing algorithm could be made robust and efficient enough for the clinical routine. The algorithm was tested on a representative set of image series successfully.
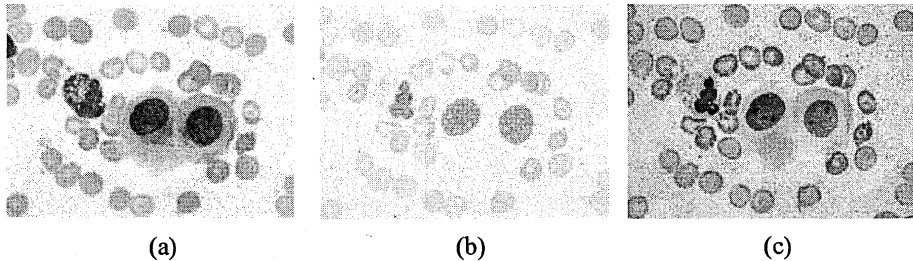

(a)   (b)   (c)

Figure 1 - Multimodal image set of three cells: the intensity of cytoplasm and of erythrocytes around the cells differs, (a) MGG, (b) FEU, (c) AGN.

2. Experimental Design

So far three cytological stainings have been incorporated into the test environment (see fig. 1): The May-Grünwald-Giemsa staining (MGG) shows both nuclei and cytoplasm. The Feulgen staining (FEU) stoichiometrically dyes DNA and hence serves for ploidy measurement [1]. The silver staining (AGN) is less specific, but it stains active nucleolar organizing regions (AgNORs). Their number and area allows conclusions on the intensity of protein synthesis and thus cell proliferation [11].

Seven series, each of 11 to 20 images per staining, were captured. Preparations were derived from serous effusions. Two slides contained cells from pleurisy due to chronic lymphocytic inflammation, two due to malignant pleural mesothelioma and three slides due to metastatic adenocarcinoma. Altogether 108 multimodal colour image triples showing a total of 693 nuclei were digitized using a Sony DXC 3-chip CCD camera mounted on a Zeiss Axioscope. Image size was 768 * 568 pixels with a pixel resolution of 0.165 μm. The software was designed to run unsupervised from the first to the last preprocessing step.

3. Adaptive Colour Space Transformation

All stainings show different colours but are nearly monochromatic, since information is carried merely in the staining intensity. The colour space conversion is carried out for each staining by a cluster analysis [4] of the RGB histogram and subsequent application of the Fisher transform [5]. Given initial estimates, the first step determines foreground and background colour distributions, with which the second performs the transformation yielding optimal foreground-background contrasts. This combination of methods makes the colour space transformation adaptive to differences in individual staining results while keeping it robust by the addition of a-priori knowledge to the start estimates.

4. Coregistration

If only mechanically retrieved, the misalignment of images is still up to 30 μm of translation and up to 3° of rotation, whereas subpixel accuracy should be reached. Alignment is a precondition to multimodal segmentation. Therefore knowledge about the image scene (e.g. positions of the cell nuclei) cannot be employed for coregistration. A common method for coregistering two images [12] is to construct a function F from the

following two components: one image is chosen as the reference image **R** (here: FEU), the other as the object image **O** (MGG resp. AGN), which is passed through an affine transformation **T(p, O)** with parameter vector **p** in the first step. Second **R** and **T(p, O)** are compared by means of a so called cost function **C** that measures their dissimilarity. It should be sensitive to misalignments while insensitive to other differences. Here a scaled variant of the pixelwise L2 norm is utilized. The overall function is

$$F(p) = C(R, T(p, O)) \qquad (1)$$

where **T**, in our case, is a 2D rigid body transformation. **F** is minimized according to **p** to gain the smallest dissimilarity between **R** and **T(p, O)**. The Powell algorithm [10] is used for numerical minimization.

The problem of local minima of the cost function was eliminated by use of a multiscale environment, the Gaussian pyramid [7]. The quality of **C** could be improved by masking foreground pixels in FEU using the colour distributions from section 3 and adapting the numbers of nonzero voxels in the other images appropriately. For each image triple two coregistrations were needed (FEU-MGG and FEU-AGN). 215 of 216 coregistrations succeeded, allowing multimodal processing of 107 of the 108 images triples (99.1 %).

5. Segmentation

A segmentation method that lets a membrane-like contour representation iteratively converge toward an underlying image feature is called active contour or snake [8]. Its movement is driven by external image forces, usually derived from image edges, and internal forces that model features like stiffness or tensibility. Because of these simulated physical features the method is known to be robust against noisy or partially missing image information. Forces act on and between points distributed along the snake, which are here called knots. Both their number and the fraction of image forces versus the inner forces make for the degree of freedom of the snake movements. They need to be balanced in order to make the snake both flexible enough to enclose the image contour and stiff enough not to be partially distracted by other image structures, which is not trivial, particularly if the variance in object appearance is high, as with cell nuclei.

These characteristics of a snake may also be intrinsic to the contour representation: When using B-Snakes [3][9], based on B-splines [6], the stiffness of the contour is solely induced by the dependency of the knot positions due to the spline polynomials, rendering explicit inner forces unnecessary. Stiffness is modeled by the density of knots and the number of sampling points between them. The latter assure a fine sampling of the image; they influence movements of their neighbouring knots but do not increase the degree of freedom of contour movements. This method results in a much smaller number of knots and reduction in computational cost.

Furthermore, the adjustment of the stiffness parameter becomes more intuitive. In our implementation the contour representation is initialized with only four knots. After the snake has converged, further knots are added *only where necessary* (fig. 2(a)). This behaviour can, for example, be driven by the average quadratic approximation error along spline segments between two knots, retaining the snake noise-resistant.

In every iteration the knots are moved within defined neighbourhoods, here contour normals. External forces of different kinds and here, from different modalities, are incorporated by a weighted summation. Edge forces are usually implemented by means of common image edge filters. The thereby induced loss of information on the edge direction disables snakes to differentiate between neighbouring objects, a major cause to the regularly criticized high dependency of snakes on initialization. In this work contour normals are convolved by a derivated Gaussian, preserving the edge direction and thus such

differentiation (fig. 3). As a side-effect, image normals can be made longer without the risk of attraction by other objects, facilitating the segmentation of concave objects (fig. 2(b)).
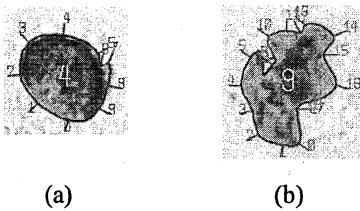

(a)   (b)

Figure 2 – Two examples of B-Snakes and variable knot density: (a) only ten knots describe the contour, with a higher density around the notch, (b) irregular nucleus.

MGG

FEU

AGN

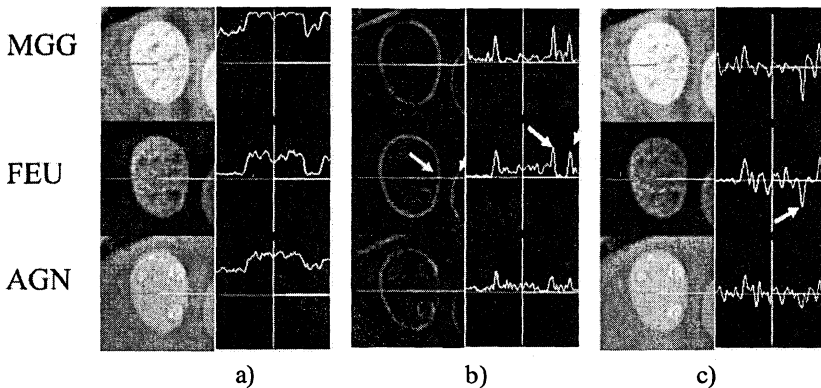a)   b)   c)

Figure 3 – The impact of directed edge filtering; multimodal, directed profiles of two neighbouring nuclei (rear grey, front white): (a) identity function, (b) same on an edge-filtered image; all edges are maxima, (c) directed edge filtering discriminates between foreground-background edges (minima) and background-foreground edges (maxima).

The snakes are initialized using a rough segmentation on binary images generated by means of the colour distributions from section 3. After the initialization they are superimposed on the grey scale image triple and the segmentation is finished.

Occasionally a snake encloses two or more nuclei after initialization. This problem is solved by two approaches: First, if the nuclei are sufficiently far apart, the snake will move into their interspace allowing for automatic separation by means of a distance criterion. If the nuclei nearly touch each other, the snake will have two opposite concavities after convergence, resulting in the second criterion for splits.

FEU   edge   identitiy   both

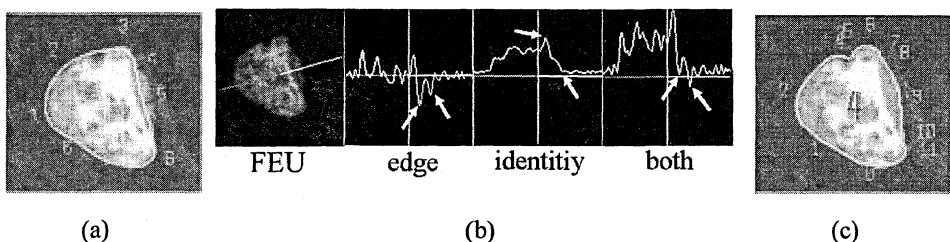(a)   (b)   (c)

Figure 4 - Image forces and their impact on segmentation: (a) incorrect segmentation due to inner edges, (b) combination of edge force and identity force; the position of the minimum changes from inner edge ("edge") to outer edge ("both"), (c) correct segmentation according to (b).

FEU proved the most suitable staining for segmentation. The exclusive use of edge forces, however, does not allow differentiation between the object border and other edges. Thus the snakes broke into objects (fig. 4(a)). The addition of the identity function, i.e. the grey level, nearly eliminated this effect, because contour edges lie on a deeper grey level than structures within an object (fig. 4(b),(c)). The overall success rate yielded 92.6 % (642 of 693 nuclei). An image triple with overlaid segmentations is shown in figure 5.
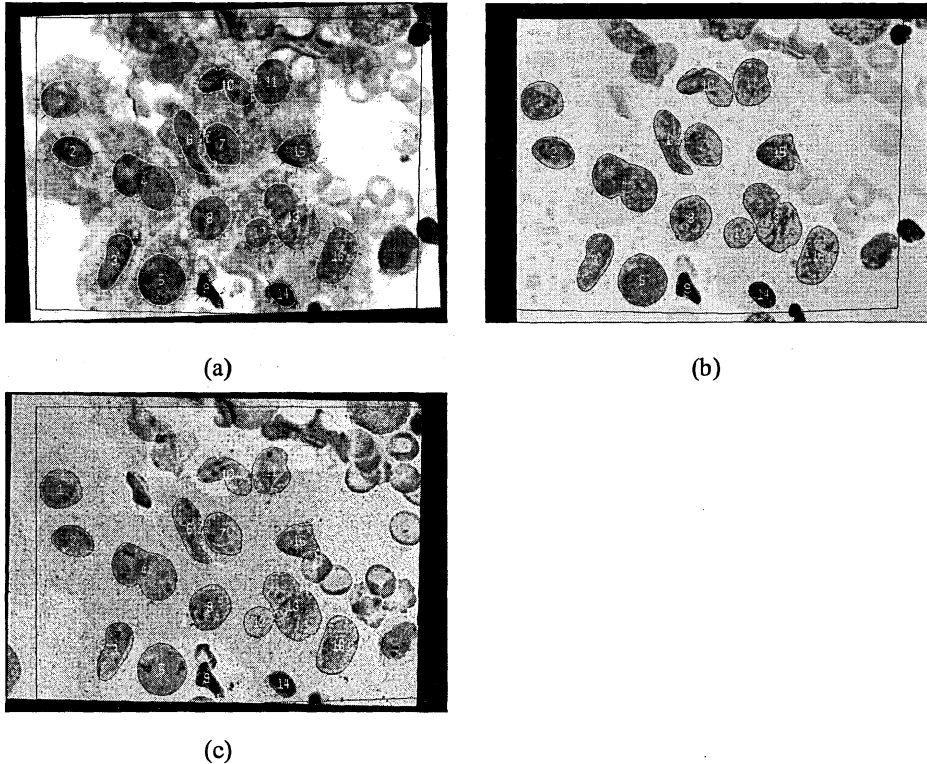

(a)

(b)

(c)

Figure 5 - Image triple with overlayed segmentation: a) MGG, b) FEU, c) AGN. All nuclei are correctly segmented using FEU, despite partially strong inner edges and some nuclei nearly touching each other. Only objects within the common area (black frame) were processed.

6. Classification

For the automatic analysis the nuclei need to be classified regarding to their cell type. There are four different classes of interest in serous effusions: one class for analysis, which contains mesothelial cells (normal or neoplastic) and carcinoma cells; and lymphocytes, granulocytes, and macrophages, which are all used for calibration. So far a Bayesian and a knn-classifier [4] have been tested on a basic set of eight features gained from different modalities. The optimal error rate is 12.5 %.

7. Overall Results, Summary and Perspective

Disregarding possible correlations between error sources, the product of the single success rates may be regarded as a "worst-case" estimation for the overall efficiency:

$$99.1\% \cdot 92.6\% \cdot 87.5\% = 80.3\%$$

   The reported project was driven by two major objectives: the first one was to develop a fully automatic preprocessing algorithm for multimodal sets of cell images acquired by means of a fixation technique in order to explore its potential for clinical routine use. The outcome clearly supports this application. The second objective was to find out whether multimodality could support the steps of segmentation and classification. As for classification this is the case because object features were extracted from different stainings. Segmentation results were best using the FEU staining only, but improvements through multimodality are still likely.

   The segmentation process will be further improved by more general approaches of object separation. The success rate of the classification process will be increased by the addition of more features, and further research on the classifier will be performed.

   In order to assure user compliance and usability, the graphical user interface has been a major component of research from the beginning. It facilitates the validation of the diagnostic advantage of Multimodal Cell Analysis, which is currently being undertaken in three pilot studies on preparations from serous effusions, oral mucosa, and thyroid.

References

[1] Böcking A. DNA Measurements – When and Why? Wied GL, Keebler CM, Rosenthal DL (eds): *Compendium on Quality Assurance, Proficiency Testing and Workload Limitations in Clinical Cytology,* in *Tutorials of Cytology,* Chicago, 1995

[2] Böcking A. Towards a Single-Cell Cancer Diagnostics. *The distinguished Ploem Lecture, 7th congress of the European Society for Analytical Cellular Pathology.* April 2001, Caen, France.

[3] Brigger P, Hoeg J, and Unser M. B-Spline Snakes: A Flexible Tool for Parametric Contour Detection. *IP* 2000: 9 pp 1484-1496.

[4] Duda R and Hart P. *Pattern Recognition and Scene Analysis.* John Wiley and Sons, 1973.

[5] Fisher RA. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 1936: 7 (2 Suppl.) pp 179-188.

[6] Foley J, van Dam A, Feiner S, and Hughes J *Computer Graphics Principles and Practice.* Addison-Wesley, Reading, Mass., 1990.

[7] Jähne B. *Digitale Bildverarbeitung.* Springer Verlag, 1991.

[8] Kass M, Witkin A, and Terzopoulos D. Snakes: Active Contour Models. *Proc. of IEEE Conference on Computer Vision.* 1987: 8-11, London, England, pp 259-268.

[9] Menet S, Saint-Marc P, and Medioni G. B-snakes: implementation and application to stereo. *Proceedings of Image Understanding Workshop.* 1990.

[10] Press WH, Teukolsky SA, Vetterling WT, and Flannery BP. *Numerical Recipes.* 2nd ed. Cambridge University Press, Cambridge, 1992.

[11] Rüschoff J. *Nukleolus Organisierende Regionen in der Pathomorphologischen Tumordiagnostik.* 1st ed. Gustav Fischer Verlag: Veröffentlichungen aus der Pathologie, 1992: 139.

[12] Woods R, Grafton S, Holmes C, Cherry S, and Mazziotta J. Automated image registration: I. general methods and intrasubject, intramodality validation. *J Comput Assist Tomogr*, 1998: 22(1), pp 139-152.

Address for correspondence

Thomas Würflinger

LfM Institute for Measurement Techniques and Image Processing
RWTH University of Technology Aachen
D-52056 Aachen, Germany
phone: +49 241 80 27974, fax: +49 241 80 22200
Wuerflinger@LfM.RWTH-Aachen.De, http://www.lfm.rwth-aachen.de