

## CONTENT-BASED AUTOFOCUSING IN AUTOMATED MICROSCOPY

PETER HAMM<sup>1</sup>, JANINA SCHULZ<sup>2</sup> AND KARL-HANS ENGLMEIER<sup>1</sup>

<sup>1</sup>Institute for Biological and Medical Imaging, Helmholtz Zentrum München German Research Center for Environmental Health, Ingolstädter Landstr. 1, D-85764 Neuherberg, Germany; <sup>2</sup>Carl Zeiss MicroImaging GmbH, Kistlerhofstr. 75, D-81379 Munich, Germany  
e-mail: peter.hamm@helmholtz-muenchen.de, schulz@zeiss.de, englmeier@helmholtz-muenchen.de  
(Accepted September 7, 2010)

### ABSTRACT

Autofocusing is the fundamental step when it comes to image acquisition and analysis with automated microscopy devices. Despite all efforts that have been put into developing a reliable autofocus system, recent methods still lack robustness towards different microscope modes and distracting artefacts. This paper presents a novel automated focusing approach that is generally applicable to different microscope modes (bright-field, phase contrast, Differential Interference Contrast (DIC) and fluorescence microscopy). The main innovation consists in a Content-based focus search that makes use of *a priori* knowledge about the observed objects by employing local object features and Boosted Learning. Hence, this method turns away from common autofocus approaches that apply solely whole image frequency measurements to obtain the focus plane. Thus, it is possible to exclude artefacts from being brought into focus calculation as well as locating the in-focus layer of specific microscopic objects.

Keywords: autofocus, classification, fluorescence microscopy, object detection, phase contrast.

### INTRODUCTION

Being the most important imaging tool in biomedical research, microscopy takes a significant role in the research advances in proteomics, genomics, bio-chemistry, and molecular biology. The achievements in these fields are the driving force for continuous improvements in medical diagnostics and drug development as well as drug targeting. Hereby, the technical progression in the recent years, namely computer and imaging sensor technologies, provides the opportunity to automate microscopy tasks to a large degree. In this context, the task of automated focusing is a fundamental problem to be solved to enable High-Throughput and High-Content Screenings. The common basic approach to face the automated focus task is to look for the optical plane along the *z*-axis of the specimen that contains the highest image contrast. In order to obtain such contrast measurements previous works applied mainly image frequency calculations such as simple gradient filter (Santos *et al.*, 1997) or, more sophisticated, wavelet transform (Forster *et al.*, 2004; Widjaja and Jutamulia, 1998) and discrete cosine transform (DCT) (Feng *et al.*, 2007). Also statistical analysis via image variance (Groen *et al.*, 1985) and autocorrelation (Vollath, 1987) calculation successfully proved to be working well to distinguish between out-of-focus and in-focus images. Additionally, improvements have been accomplished by selecting only areas of interest within an image and further enhancements

on existing methods, *e.g.*, modelling the focus curve (Brázdilová and Kozubek, 2009). Although numerous image-based autofocus methods have been proposed and some found their ways into commercial products, they all suffer from limited applicability when either the observation of specific objects is wanted or when objects in different optical layers need to be in focus. A patent pending method that handles these difficulties by assuming a completely new approach to the problem is described in this paper. The main idea is to refrain from using the hitherto existing whole image analysis and to take the actual image content into account. Our approach uses preceded class training to gain knowledge of the research object. With this *a priori* information we are able to direct closer focus measurements only to regions of interest and, thus, exclude unwanted objects and artefacts from the observation.

### METHODS

The herein presented novelty to approach the focussing problem is not due to the development of just another focus measurement function but to explicitly use the knowledge of the examined object to constrict the focus search solely to regions of interest. The general proceeding can be formulated as a task of three steps (see Fig. 1). In the first step, a captured image is scanned for regions holding strong object hypotheses while assuming that the microscope is in some initial

position within the sample and the captured image does contain a most likely blurred state of the observed specimen. In the second step, focus measurements are processed only upon the detected regions along the  $z$ -axis to yield the respective in-focus layer. In the event of multiple objects being at different  $z$ -layers – as it is the case for thick and non-monolayer preparations – consequentially multiple layers will be given out as result. In the final third step, all found layers are closer verified to ensure that the focused images do factually hold the desired object. Thereby, layers are rejected that turn out to contain unwanted objects (e.g., artefacts) when surveyed in focused state.

In order to conduct the above described steps, it is necessary to feed the focusing system with information on the subject to study. By means of classification tasks known from the field of pattern recognition - that is feature extraction from known sample images and subsequent classifier training - a decision function is built to locate potentially interesting image regions. For this, we apply a well-known object detection framework that makes use of Haar Features and boosted classifier learning (Viola and Jones, 2001). The important key elements of this system and its application to our issue is explained in the following section.

### FEATURE EXTRACTION AND CLASSIFIER TRAINING

In terms of preparing for the classifier training an image set is needed. This sample set is generated as the microscopist is asked to manually mark representative samples in the in-focus layer. Hereby, the number of samples does not need to exceed a dozen to keep user interaction within reasonable limits. Though the selection has to cover the variety of the considered object category. The sample set is furthermore complemented by blurred and rotated instances to add defocused states and ensure rotation invariance which is not supported by the selected features that will be described in the following. In our experiment we use simple Gaussian filtering for blurring and linear interpolated image rotation transform that produces instances each with an angle offset of  $30^\circ$  while mirrored background padding inhibits negative border effects. The step of feeding the training set with samples is supposed to happen only once while it is possible to add samples in further observations to refine the classifier.

The feature extraction uses the Haar wavelet set (Fig. 3) introduced by Papageorgiou (Papageorgiou *et al.*, 1998) which has been successfully used for a

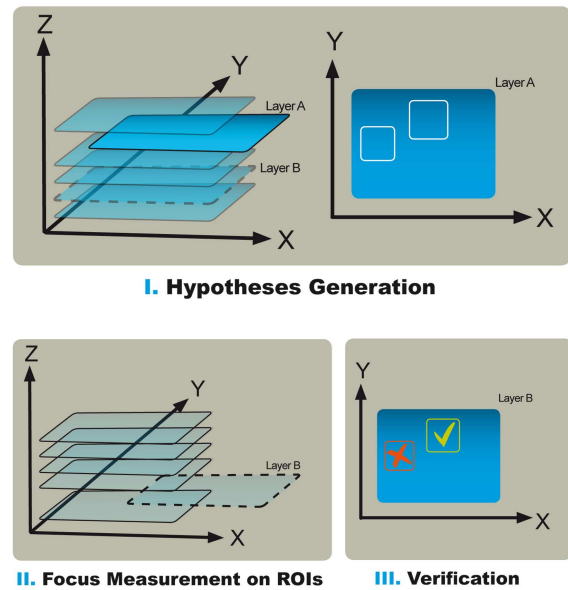


Fig. 1. Schematic focus working steps. I. From a random  $z$ -position (Layer A) within the specimen, a layer is captured upon which object hypotheses are found. II: Focus search is limited to the determined areas yielding one or multiple in-focus layer(s). III: Within the in-focus layer (Layer B) a more detailed classification is processed to verify the focused areas.

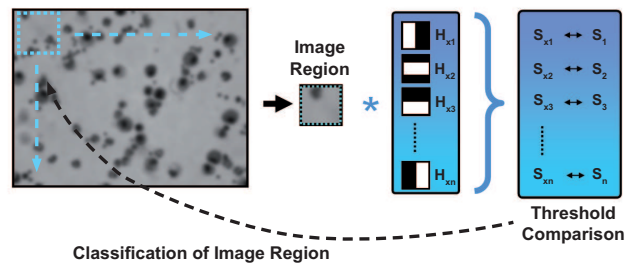


Fig. 2. The user marks positive and negative areas in an initial process to provide a training image database for the boosted training with Haar Features.

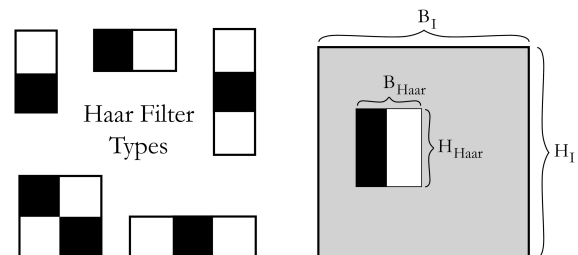


Fig. 3. Haar Wavelet Types and their exemplarily depicted positions and scales within an image.

variety of detection tasks like face (Viola and Jones, 2001; Lienhart and Maydt, 2002), object and even for cell recognition (Smith *et al.*, 2009). This choice of features have been made for two reasons: on one hand Haar features are generally applicable to any kind of probe and, since we have to consider unsharp representation of any object when operating on heavily blurred images, features of, *e.g.*, morphological kind cannot be taken into account.

On the other hand the detection routine with Haar features can be unbeatably fast when compared to for example scale-invariant keypoint features (Lowe, 2004). An extracted image feature corresponds to the response of convoluting a particular Haar filter with an image area. At this, each Haar filters does vary in type, position within an image and size as displayed in Fig. 3. The number of features  $N$  per Haar filter type can be computed from the size of the image with the width  $W_I$  and the height  $H_I$  and the initial size of the specific Filter (width  $W_{\text{Haar}}$  and height  $H_{\text{Haar}}$ ) as follows:

$$N = \frac{W_I}{W_{\text{Haar}}} * \frac{H_I}{H_{\text{Haar}}} \left( W_I + 1 - W_{\text{Haar}} * \frac{W_I}{W_{\text{Haar}}} + 1 \right) * \left( H_I + 1 - H_{\text{Haar}} * \frac{H_I}{H_{\text{Haar}}} + 1 \right). \quad (1)$$

The amount of possible Haar features is by far over-complete and of great number even for small images. Keeping Fig. 3 in sight and using the above equation, a rather small image (region) of  $30 \times 30$  pixels ( $W_I = H_I = 30$ ) would produce almost  $N = 105000$  features for only one Haar filter type with an initial size of  $2 \times 1$  pixels ( $W_{\text{Haar}} = 2; H_{\text{Haar}} = 1$ ).

Only a few of these many Haar filters can contribute to classification and, moreover, a single Haar filter would obviously represent a weak decision-maker. Thus, classifier training is needed to select those of the thousand features that discriminate best between the object classes. As for Viola and Jones' recognition approach (Viola and Jones, 2001), the method of choice is adaptive boosting (AdaBoost) (Freund and Schapire, 1995) in combination with cascading the classifiers as introduced by Viola and Jones (Viola and Jones, 2001).

AdaBoost selects and concatenates the Haar filter by weighting training images according to the difficulty to be classified correctly. Thereby, each Haar filter (*i.e.*, each possible classifier) that passes through the iterative training process has to focus on hard to classify samples. The training result is a combination of several weak decision-makers that together form a

strong classifier. Additionally, trained classifiers are arranged into cascade of stages, in a way that each stage fulfils minimum detection and maximum false positive requirements.

Stages are added subsequently until a specified overall classification performance is reached. This cascading of classifiers achieves a tremendous speed-up during the detection process: a positive image has to pass all cascades to be classified as such. Any rejection by one classifier within any cascade will instantly reject the examined image (see Fig. 4) without passing the complete cascade.

## FINDING THE FOCUS LAYER

In the routine of finding the focus we assume an initial position in  $z$ -direction (image I in Fig. 5) that lies within the upper and lower bounds of the specimen, though improbable representing the in-focus layer. Based on the afore trained classifier, the detection process scans the entire image by means of sliding a search window across the image and returning the regions with an object hypothesis (red squares in image II of Fig. 5) if classified as such. Hereby, the minimum and maximum size of the search window is chosen accordingly to the size variation of the observed object class. Then, the system applies first-order Gaussian derivative (Geusebroek *et al.*, 2000) focus measurement with a sigma value of 1.5 solely to the areas of hypotheses along the  $z$ -direction in parallel.

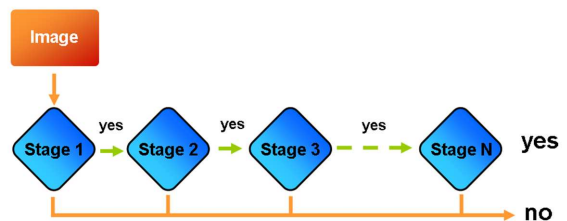


Fig. 4. Cascade of stages containing varying number of weak classifiers.

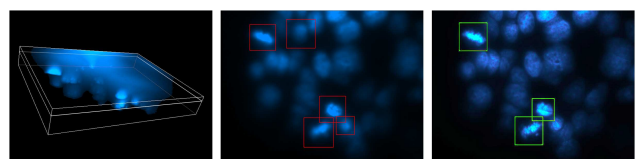


Fig. 5. Focus detection of mitotic cells.

As the first step of hypotheses generation is performed with a rather coarse detection, it assures a high rate of true positive detection while deliberately accepting a high rate of false positive identifications, too. In order to reject these false positives, all found regions are finally verified in their respective in-focus layer (green squares in image III of Fig.5). This is achieved by a second classifier, which has been trained only on non-blurred samples.

## RESULTS

The proposed method was tested on data sets captured with the inverted microscope Carl Zeiss Axio Imager.Z1 equipped with an AxioCam HR3 and a motorised stage in  $z$ -direction. The autofocus method was implemented as macro plugin to work online within the AxioVision software environment. Though, for better reproducibility of results and with fading fluorophores in fluorescent samples in mind, the probes were captured as  $z$ -stacks and the autofocus routine was conducted upon these. All testing have been processed on a notebook with 1.80 GHz Intel Core2Duo and 3 GByte of RAM. The dimensions of the captured stacks have been 1388 by 1040 pixel per image with 16 bits per pixel and up to 75 layers per stack.

The test examples contained several challenges regarding the task of autofocus such as non-monolayer specimens and bright artefacts that usually bias conventional autofocus techniques. As starting point of observations of new specimen types, the user marks image regions within an in-focused sample image with respect to the class affiliation.

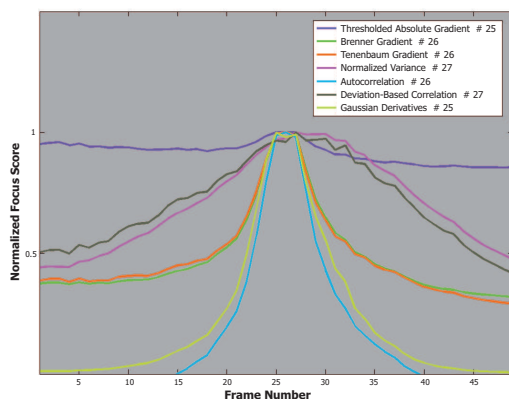


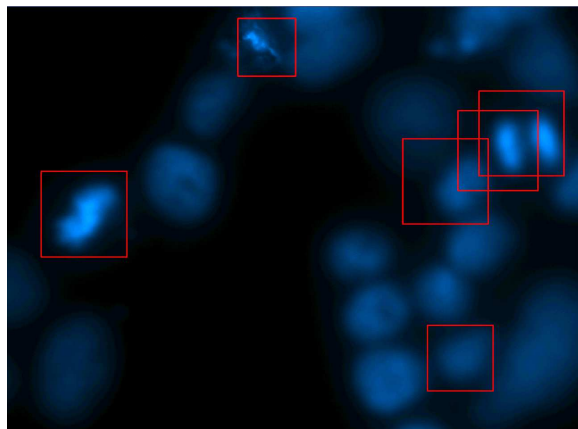
Fig. 6. Focus curves obtained by applying the most common focus measurements to the sample shown in Fig. 7.

We again make a point of having only as few as possible samples for the training input (maximum one to three dozens each class) to keep user interaction at minimum. Finally, we compared the results of the Content-based Autofocus system with the most known autofocus routines (Geusebroek *et al.*, 2000; Sun *et al.*, 2004) (see complete list in label of Figs. 11a and 6) found in literature. Furthermore, a biologist with profound knowledge in cell biology was conducted to give his expert opinion regarding the subjective focus position for every probe.

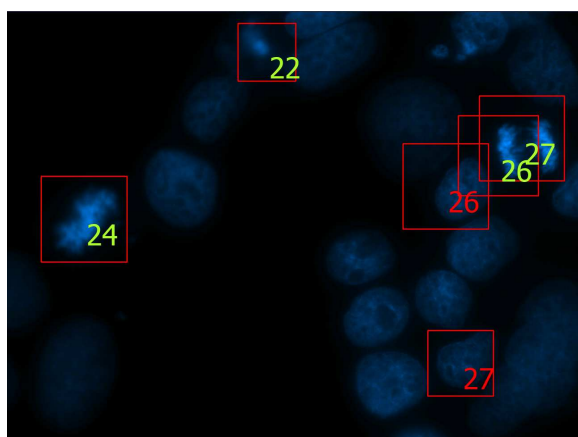
The majority of evaluation tests was performed on CACO-2 preparations (Human colon adenocarcinoma) nuclear stained with Hoechst 33258 and showing cells in different stages of cellular division. Cell proliferation and cycle analysis of CACO-2 cell lines are widely used in clinical cancer research and pharmaceutical industry. In our test case with altogether 22 stacks, we trained the classifier to detect cells in metaphase and anaphase when the spindle apparatus is notably expressed. The  $z$ -stacks were captured with a step size of 30  $\mu\text{m}$  since an oil immersion objective with a magnification of 63 $\times$  and a numerical aperture of 1.4 was used. The number of layers between the lowest and highest in-focus layer averaged at 2.53 ( $\sim 76 \mu\text{m}$ ) which points out the existence of more than only one focus layer. In fact, some stacks contained cells of interest in up to three different focal layers being spatially separated by 150  $\mu\text{m}$ .

All in all 16 positive and 34 negative samples plus their blurred and rotated instances have been used as sample set for the preceded classifier training. The results shown in Fig. 7a stand exemplary for our tests with the CACO-2 preparations which hold cells in different focal planes. While the mainly unimodal focus curves of the best-performing known methods AutoCorrelation, First-Order Gaussian Derivatives and Normalized Variance (see Fig. 6) point to Layer #26, #25 and Layer #27, Content-based Autofocus approach is able to bring out the focus of all planes at  $z$ -position 22, #24 and #26 respectively #27. From Fig. 7b it can also understandingly seen why common methods aim at the Layer around #26. Apart from the rightmost dividing cell, most of the non-mitotic cells are located at this layer and thus giving the greatest overall structural input to standard focus functions. While this example shows an excellent result, the current performance of the presented method can be summarised in this way: on one hand it is truly able to get desired objects into focus while excluding the vast majority of unwanted objects – which is confirmed by a sensitivity of 0.836 – on the other hand, the performance is yet limited by some False-Positive-Detections, too.





(a)



(b)

Fig. 7. Focus evaluation of mitotic cells in three different focal planes. (a) Result of the hypotheses generation in a defocused layer; (b) Identification of focal planes at each ROI (denoted with the respective layer number) and evaluation that successfully throws out two hypotheses containing non-mitotic cells (red numbers). The displayed image shows layer #26 in which only the most right detection is in focus.

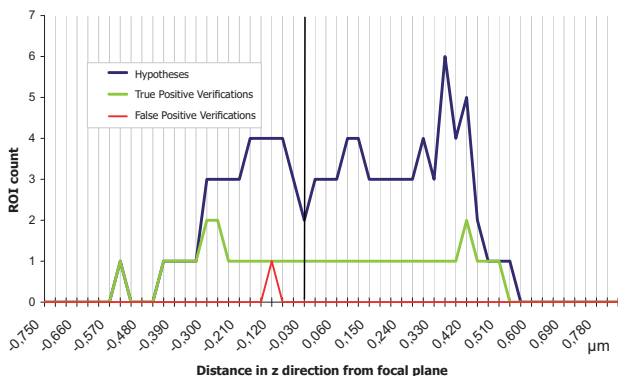


Fig. 8. The number of found hypotheses rapidly decreases with increasing distance from the focal plane (marked with the vertical black line at 0 µm).

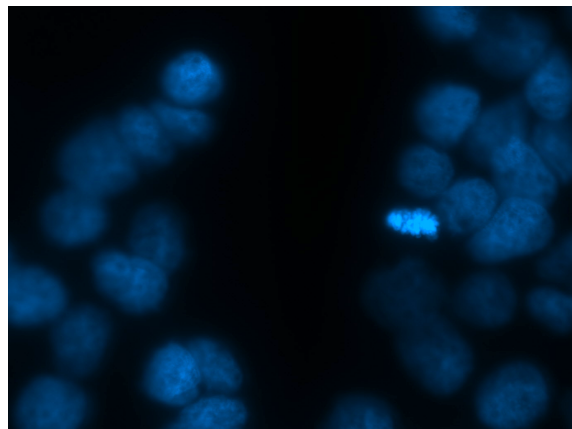


Fig. 9. Focus layer of the observation that gives the chart in Fig. 8.

The initial z-position wherefrom the hypothesis generation starts does of course have a direct relation to the amount of detections because the image information rapidly decreases with the blurring. As it can be seen in Fig. 8, it is meaningless to set the initial position beyond a certain range. We found that there is not enough image information left for a reasonable hypotheses search if we start farther away than around 400 µm from the focal layer in our case.

We standardised our testing in such a way that we carried out two runs for every image stack each with different initial z-position: one starting 300 µm below and the second 150 µm above the focal plane respectively the middle of all focal planes. The overall results (see Table 1) achieved with the Content-based approach show that 71.31% of all possible focal planes containing mitotic cells were found with the Content-based approach while there were on average 0.84 layers per stack erroneous detections caused by false positive verifications. The first-order Gaussian derivative method reached on the same sample set a rate of 56.43% whereas the autocorrelation algorithm gives a rate of 46.21% as result and normalized variance definitely fails with a rate of 12.80%.

Table 1. Comparison of focusing results achieved by the Content-based approach and classical methods.

CACO-2 SERIES	Detection rate
Content-based autofocusing w/ Gaussian derivatives	71.31%
First-order Gaussian derivatives	56.43%
AutoCorrelation	46.21%
Normalized Variance	12.80%

In another small test series cytoskeleton preparations were observed in the phase contrast channel. Partially detached round cellular structures (see Fig. 11) make these samples a difficult challenge because they show a typical phase transition effect of bright halos with gradients and intensities above average. Accordingly, most autofocus methods are not able to find the correct focal plane and their focus curves show multiple local maxima (Fig. 11a). In the cytoskeleton sample displayed in Fig. 11c, the Gaussian derivatives approach lead to a – from human standpoint – acceptable layer #20. Though, only the autocorrelation method brings out the correct result with layer #21. Likewise does the content-based autofocus approach which succeeded by finding layer #21, too.

However, we can see a considerable performance differences in our implementation: while the standard Gaussian Derivatives method with an exhaustive search strategy required 19.07 seconds for calculation in the above case, the Content-based Autofocus with the Gaussian Derivatives function as focus measure took 13.14 seconds and only 5.53 seconds when the focus function is substituted with the Normalized Variance function - and still achieving the correct focal plane. This performance boost originates mainly from the exclusion of large areas also with artefacts from the outset. While the speed up is true for the example in Fig. 7, the increasing number of object hypotheses will inevitably increase the foci search time.

## DISCUSSION

Our main intention was to present a new way of automated focusing that explicitly deals with non-monolayer preparations and the occurrence of artefacts and additionally, is per se not limited to the nature of the objects of study. In summary, the presented method puts a region-of-interest detection in front of a common autofocus measurement function. By this means our focus approach is not only able to narrow the focus search down to certain areas but is also capable of detecting multiple in-focus layers while excluding uninteresting objects. We have developed a fully functional software plug-in to work with a motorised microscope where the user furthermore has the possibility to generate sample sets and train a classifier that is used during the detection step during focus search. The results of our testing show demonstrative improvements compared to common autofocus methods. The system performance means a considerable speed up of the focus search in case of numerable hypotheses. This is not the case for tissue

samples that show large connected structural areas of cells rather than spatially distinct objects. Here, a trained classifier will lead to a large quantity of hypothesis regions and the subsequent focus search slows down the overall focusing procedure. But the main advance of the presented method is the ability to detect the correct focal plane respectively planes largely due to the elimination of the influence of artefacts and uninteresting objects.

It has to be emphasised that, what have been described in this paper, is a first realisation of a promising concept. We are certain that further development and testing will release the full potential of the Content-based Autofocusing. First of all, it is desirable to extend the test onto a larger variety of different preparations; also with industrial sample to prove the method's universal applicability. The usage of Haar features in combination with AdaBoost learning seems reasonable in terms of detection speed and quality of hypotheses generation. Nevertheless, the not negligible number of false positive detections in the verification step motivates for studies on different classification methods. However, they will still have to consider a procedure that is not limited to only some specimen types.

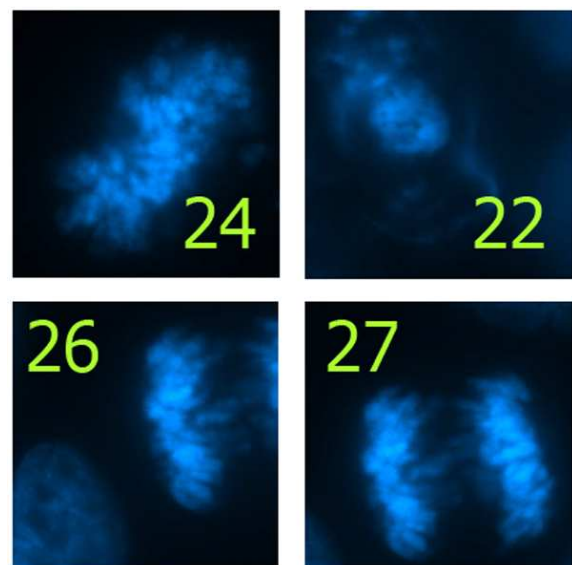
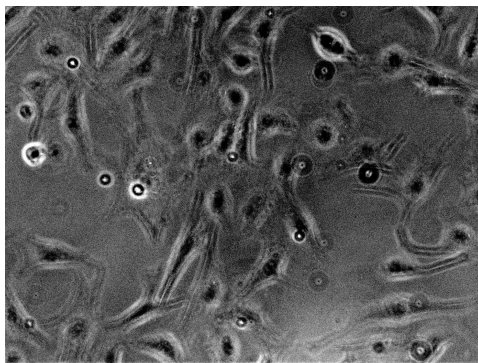
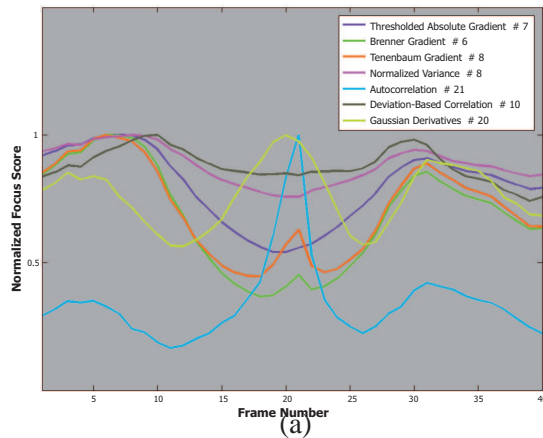
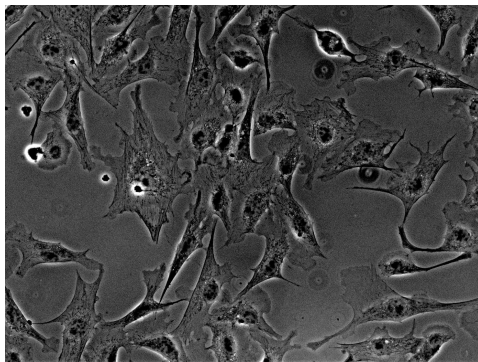


Fig. 10. *Magnifications of areas within their focus layer (numbers show layer index) found by the Content-based autofocus application as displayed in Fig. 7.*



(b)



(c)

Fig. 11. Cytoskeleton specimen in microscope phase contrast mode and autofocus measurement results by classic approaches and the respective in-focus layer found by the Content-Based Method and as well as Autocorrelation at layer #21. (a) Focus curves obtained from common focus functions. Only Autocorrelation has a comparatively unimodal and steep curve while all others are highly affected by distracting cellular structures around Layer #9 and #30. Only Autocorrelation led to the correct in-focus layers #21; (b) The wrong finding of Normalized Variance Algorithm; (c) Focal plane of cytoskeleton specimen at layer #21.

## ACKNOWLEDGEMENT

The authors highly appreciate Dr. Andreas Murr's assistance in giving his expertise for the evaluation.

## REFERENCES

- Brázdilová SL, Kozubek M (2009). Information content analysis in automated microscopy imaging using an adaptive autofocus algorithm for multimodal functions. *J Microsc* **236**:194–202.
- Feng Q, Han K, Zhu X (2007). A new auto-focusing method based on the center blocking DCT. In: Zhang YJ, ed., *Proc 4th Int Conf Image Graph (ICIG)*. Chengdu, China. Aug 22-24. IEEE Computer Society.
- Forster B, Van de Ville D, Berent J, Sage D, Unser M (2004). Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images. *Microsc Res Techniq* **65**:33–42.
- Freund Y, Schapire RE (1995). A decision-theoretic generalization of on-line learning and an application to boosting. In: Vitányi PMB, ed., *Proc 2nd Eur Conf Comput Learn Theory (EuroCOLT)*. Barcelona, Spain. Mar 13–15., no. 904 in *Lect Notes Comp Sci*. Berlin, Heidelberg: Springer.
- Geusebroek J, Cornelissen F, Smeulders A, Geerts H (2000). Robust focusing in microscopy. *Cytometry* **39**:1–9.
- Groen F, Young IT, Ligthart G (1985). A comparison of different focus functions for use in autofocus algorithms. *Cytometry* **12**:81–91.
- Lienhart R, Maydt J (2002). An extended set of haar-like features for rapid object detection. In: *Proc Int Conf Image Proc (ICIP)*. Rochester NY, USA. Jun 24–28. IEEE Computer Society.
- Lowé DG (2004). Distinctive image features from scale-invariant keypoints. *Int J Comput Vision* **60**:91–110.
- Papageorgiou CP, Oren M, Poggio T (1998). A general framework for object detection. In: *Proc 6th Int Conf Comput Vision (ICCV)*. Bombay, India. Jan 4–7. IEEE Computer Society.
- Santos A, Ortiz de Solórzano C, Vaquero JJ, Pena JM, Malpica N, Del Pozo F (1997). Evaluation of autofocus functions in molecular cytogenetic analysis. *J Microsc* **188**:264–72.
- Smith K, Carleton A, Lepetit V (2009). Fast ray features for learning irregular shapes. In: *Proc 12th Int Conf Comput Vision (ICCV)*. Kyoto, Japan. Sep 29–Oct 2. IEEE Computer Society.
- Sun Y, Duthaler S, Nelson BJ (2004). Autofocusing in computer microscopy: Selecting the optimal focus algorithm. *Microsc Res Techniq* **65**:139–49.
- Viola P, Jones M (2001). Rapid object detection using

- boosted cascade of simple features. In: Proc Conf Comput Vision Pattern Recogn (CVPR). Kauai HI, USA. Dec 8–14. IEEE Computer Society.
- Vollath D (1987). Automatic focusing by correlative methods. *J Microsc* 151:279–88.
- Widjaja J, Jutamulia S (1998). Use of wavelet analysis for improving autofocusing capability. *Opt Commun* 151:12–4.