HIGH SPEED AUTOFOCUS FOR MICROSCOPIC IMAGES

J. E. Fischer, D. Homeister, J. Lehmler, G. Roos†

Institute for Physical Electronics, University of Stuttgart
Pfaffenwaldring 47, 7000 Stuttgart 80, Germany

† Institute for Microelectronics Stuttgart
Allmandring 30a, 7000 Stuttgart 80, Germany

e-mail: fischer@ipesun.e-technik.uni-stuttgart.de

Abstract

The processing of biological specimens or the examination of material surfaces for quality control requires a fast and precise autofocus system. In order to process several fields of view per second, a high speed autofocus was developed and integrated in an ASIC-chip. The system is able to process a specimen without any supervision. The system operates on the well known dual-out-of-focus principle by using three images in three different focal planes, where multiple standard CCD-cameras are used for autofocus, each one with an 8 bit digitizer running at a pixel rate of 14.2MHz. By taking every half image of the interlaced TV-frame, 50 measurements per second are made. With the results from three different focal planes, the determination whether the current scene is already in focus or not, can be made within 20 milliseconds. The system is controlled by a microprocessor which is moving the scanning stage from field to field and which focusses the scenes. The software is made intelligent by the way that it learns, how to use the statistical properties of the investigated surface to predict the position of the next focal plane.

1 Introduction

For the automatic investigation of samples under a microscope a motorized scanning stage is needed to move the specimen in the lateral and axial direction. One critical task is to focus the scenes after the scanning stage was moved from one field of view to the next one. For this purpose a standard microscope with an automated scanning stage was taken and an autofocus system was developed and added to the microscope. The autofocus uses the well known dual-out-of-focus principle by taking three images in three different focal planes simultanously. The three image planes are generated by using mirrors to project the object onto the CCD camera targets with a different length of the light path. Figure 1 shows the basic principle of the optical paths through the microscope. The signals of the three focus cameras are then analyzed by special hardware to determine the sharpness of the current scene. A color camera is connected to the host of the main image processing system, which is grabbing and analyzing the focussed images. With this equipment a fully unsupervised analysis of a microscopic sample can be done.

2 Computational Theory

2.1 Overview

The determination, whether an image is in focus or not is dependent on various facts. An absolute statement about the quality of the focussing is possible, only if a model of the scene and the image pickup system is available. This is also true for our eye-brain-system with its given resolution and the knowledge about the world. If there is no a priori information about the scene, a statement about the quality of the focussing can only be made with respect to another image by changing the parameters of the image pick-up, e.g. the focal length of the camera. This means that several images in different focal planes are needed to find the optimal image. In addition to this, the objects are normally three-dimensional and are projected onto a two-dimensional image plane. In general, all parts of the image are not in focus at the same time. With the given aperture of the camera or the microscope and the size of a pixel or element of a CCD-camera, the depth of focus can be determined.

For the development of our autofocus system the following requirements were taken into account:

Quality: The focus algorithm has to detect the optimal focus position precisely.

Unambiguity: The computed focus value must increase continuously with the degree of focussing and must have its maximum at the optimal focus.

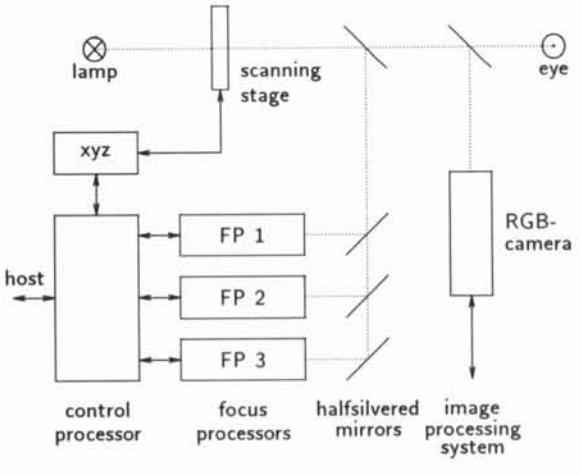

Fig.1: Block diagram of the microscope.

Reproducibility: The focus value must be independent of the previous history and must be the same no matter from which direction the approach was made.
Insensitivity: The focus algorithm must be insensitive to noise and changing illumination.
Dynamic Range: Within a given range, the system must be able to operate quickly and reliably (tracking). The system must also be able to find the optimal plane of focus beginning from any state (lock-in).
Computational Costs: The time to compute the focus value for a given image should be as small as possible.

Many of the proposed algorithms in the literature were tested with respect to the above criteria. The methods can be divided into three groups. The first one only compares single pixels in subsequent images and the quality of the focus is computed for every pixel [Haeu82], [Piep83]. These methods are not suitable for our application, because a large number of images at different focal planes are needed to compensate for noise, as our experiments showed. The second group uses the local neighborhood of a pixel to compute the contrast or the local variance [Ligt82], [Sugi85], [Krot86], [Krot89]. This group of algorithms was the most widely used one and gave the best results. There are many different suggestions for a suitable operator but they all use the same principle: the local variance is computed by determing differences between adjacent pixels. Some of them are using nonlinear elements to enhance the differentiation process, to combine operators working in different spatial directions or to smooth the results in the perpendicular direction. A third group uses global operators like the Fourier-transform to compute the amount of the higher frequencies in an image. Although these algorithms can yield the optimal focal plane, the computational costs are so high, that they were not considered.

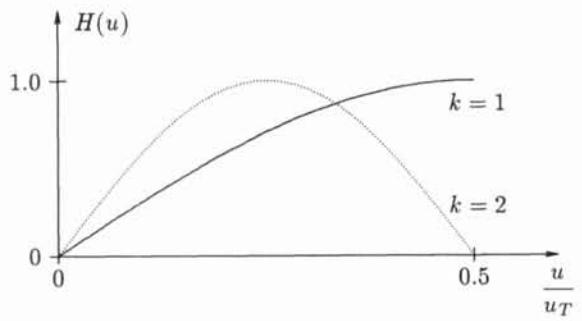

Fig.2: Transfer function H(u) of the difference operator $I(x,y) - I(x-k,y)$ with $k = 1$ and $k = 2$. u is the spacial frequency and u_T the sampling frequency.

2.2 Implementation

To select the best algorithm for our implementation, experiments with all procedures listed above were made and the final choice was made with respect to the accuracy and the computational costs. The experiments showed that the quality of some simple operators were as good as the quality of more sophisticated ones. Therefore we choose the one-dimensional operator

$$F = \sum_{y=1}^{Y} \sum_{x=1}^{X} [I(x,y) - I(x-k,y)]^2 \qquad (1)$$

where X and Y are the number of pixels in x- and y-direction of the image intensity function I within the selected image area. k denotes the distance between two pixels which are used for the differentiation. Finally, F is the focus value obtained for the whole area of interest. Fig. 2 shows the transfer function of the differentiation operator $I(x,y) - I(x-k,y)$ with $k = 1$ and $k = 2$.

This operator is a highpass filter for $k = 1$ and a bandpass filter for $k = 2$. The following nonlinear square enhances the bigger pixel differences more than the smaller ones and thus enhances steeper gradients in the images more than smaller ones. All these locally computed values are then added within the given region of interest and the value F is obtained to indicate the degree of focussing.

This operator is one of the best operators we tested. In addition, the computational costs are very small, because it works only in the x-direction and uses only two pixels to compute the local contrast. The fact that this operator is one-dimensional is not a problem for our application, because it is very unlikely to have only horizontal structures in our microscopic images. This is especially true for biological specimens.

3 ASIC-Design

3.1 Specifications

In order to have a fast autofocus system with low costs, standard TV-technology was used in combination with a stepper motor attached to a conventional microscope. One TV-frame consists of two half-images of each 20ms duration. The pixel rate was choosen to be 14.2MHz to get square pixels. Both the vertical and horizontal synchnonization pulses and the pixel clock are taken from the CCD camera as input to the focus processor. The host processor has access to several control registers on the chip. The area of interest within a TV-frame can be defined as well as an upper and lower gray value threshold to be able to eliminate image regions without objects. Within these given spatial and gray value borders the focus value is computed according to equation (1) and integrated over the whole region. The parameter k can be selected to be 1 or 2 depending on the lowpass characteristics of the output amplifier of the used CCD-camera. In addition the gray value sum within the region of interest is also computed. This is needed to calibrate the autofocus system during the initialization phase.

All the operations described above, the spatial window, the gray value window, the pixel difference,

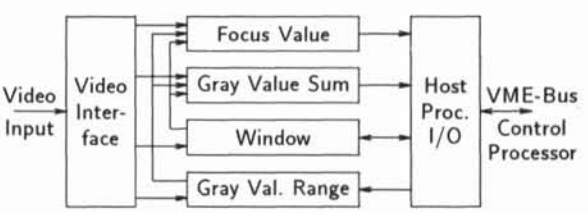

Fig.3: Block diagram of the focus processor.

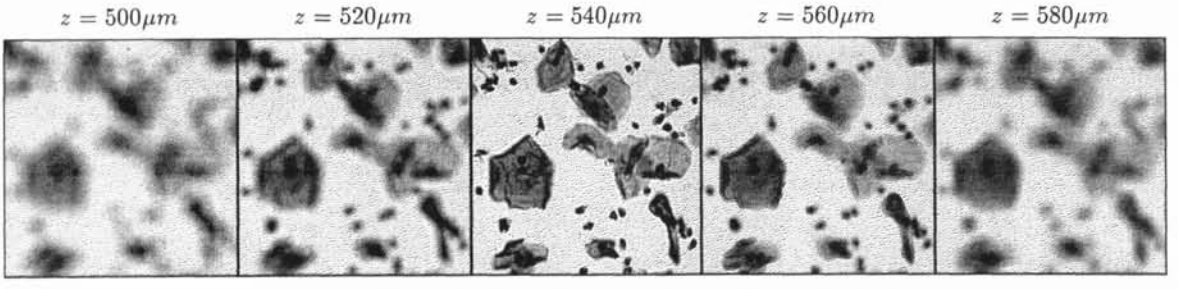

Fig. 4: Scene from a biological specimen in 5 different focal planes

the square of differences and the summation are done at the pixel rate of 14.2MHz. Then after 20 milliseconds (one TV-half-frame) a value for the degree of focussing is available and can be read from the host processor. In this way the focus processor computes 50 focus values per second. Figure 3 shows the block diagram of the focus processor ASIC.

3.2 Chip Design

For the realization of the focus processor an ASIC (**A**pplication **S**pecific **I**ntegrated **C**ircuit) was chosen. This was considered to be a good compromise between a full custom design and the use of programmable logic device (PLD). ASICs are made of premanufactured wafers containing an array of transistors which are connected by a 2-layer metal back-end process, according to the customers logic specifications. The design and development time, the speed of the used technology and the price per chip unit were taken into account. The advantages and disadvantages of the ASIC-design are outlined below:

+ fast design time (due to the use of existing circuit libraries)
+ fast production time (especially in conjunction with E-Beam direct writing)
+ suitable for low volume production (especially in conjunction with E-Beam direct writing)
+ low overall costs
+ short training time
− design constraints with a given master
− not optimal use of chip space

The way from the specifications listed above to a ready chip lead through many design stages that are the same no matter if a full custom design, an ASIC or a PLD is used. First a set of testvectors was defined to be able to test the whole circuit. Then the blocks shown in fig. 3 were designed on a workstation. Compared to PLDs it is neccessary to integrate a testpath on the chip to be able to reach every node on the chip in a testmode (boundary scan test). The blocks were tested separately and then put together and tested on a logical level. After the automatic placement and routing all switching times were known and the design verification showed that the chip was expected to work at a frequency of at least 25MHz.

The chip master is a GFG4 from the *GATE FOREST* or *sea-of-gates* family of the *Institute for Microelectronics Stuttgart*, using a 1.2μm CMOS technology with two metal layers. The use of E-Beam direct writing in the back-end processing for the metallization (contact, metal 1, via, metal 2) lead to a fast turnaround time and to low cost. 25,442 transistors are used in our application, the 84 pins of the chip are TTL and CMOS compatible and are able to drive up to 12mA. The chips are tested on the wafer with the vectors from the design phase. The finished integrated circuit then is tested in our application environment described in chapter 4.

4 Control Hardware

The hardware uses a VME-bus concept (half-height) with a 68000 control processor. There are three focus processor ASICs on one printed circuit board, which is directly connected to the board with the analog/digital converters for the image signal. A 3-axis motor driver board is used to move the scanning stage of the microscope in three dimensions.

A scene from a biologial specimen in 5 different focal planes is showed in figure 4. The selection of a focal plane is done by moving the scanning stage with a stepper motor up and down in the z-direction. The resolution of the z-axis is 0.5μm, the calculated depth of focus is about 5μm in the example of fig.4 with a 20× objective. The images there were taken with a distance of $z = 20\mu$m, which is also the distance between the images which are projected onto the three CCD-cameras. So, a series of 3 succeeding images in fig. 4 are seen by the cameras of the autofocus systems.

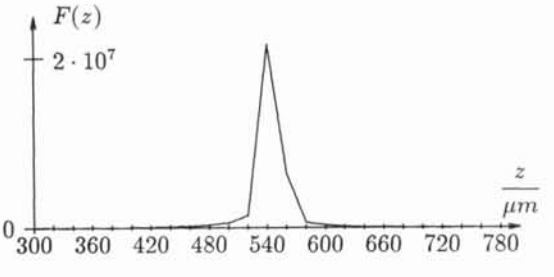

Fig.5: Measurements of the focus value F with the coarse focus algorithm as a function of the relative z-position of the scanning stage to the 3 cameras. 3 subsequent samples are each taken at once.

5 Experimental Results

5.1 Coarse Focus

During the initalization phase, the system has to focus on the first presented scene after a slide is loaded. Due to variation between slides the microscope is usually too far from the point where the dual-out-of-focus priciple could succeed. In this case a coarse focus algorithm is started, which seaches for the best focussed image within a wide range and which uses

increments of 20μm for the motor in the z-direction. Figure 5 shows a typical curve of such a coarse focus run. With 25 different positions we can cover the relevant range of half a millimeter depth. These values work fine for biological specimens and can be modified for different applications.

5.2 Fine Focus

The fine focus algorithm is used when the current scene is nearly in focus. This is the case when the scanning stage moves from one field of view to the next one. The function of the focus value F is shown with a step size of 1μm for the camera in the central location. The middle image in figure 4 belongs to the position z_2 with the focus value $F_2 = F(z_2)$. The two other cameras give nearly the same curves. They are not exactly the same because they are at different positions in the light path of the microscope. This effect can be neglected. A bigger issue is that the light intensity is not the same on the three cameras, because of the imperfection of the half silvered mirrors in the optical path (see fig. 1). With this knowledge, the focus values $F_1 = F(z_1)$ and $F_3 = F(z_3)$ for the two other cameras can be added to figure 6 at the positions z_1 and z_3 assuming that the middle camera would give these values when it is moved to z_1 or z_3 respectively.

The correction for the different light light intensities is made by dividing the focus value by the square of the mean gray value within an empty field of view. The mean gray value is obtained by the summation of all pixels within the region of interest, thus

$$F_i' = F_i / M_i^2 \tag{2}$$

where M_i is the mean and F_i' is the corrected value of the camera i with $i = 1, 2, 3$. From these values F_i' a correction Δz for the z-position can be obtained by

$$\Delta z = \lambda \frac{F_3' - F_1'}{F_2'} \tag{3}$$

This is a robust estimation with respect to noise. λ is a constant determined by the optical system, and suppresses oscillations of the focussing process. This correction step is repeated until Δz is smaller than a given threshold ϵ which also depends on the optical parameters of the system and the value of λ.

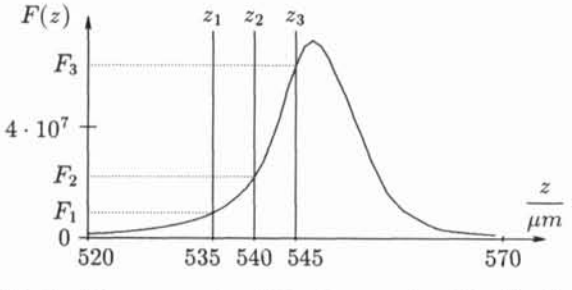

Fig.6: Measurements of the focus value F with the fine focus algorithm as a function of the relative z-position. The focus values F_1, F_2 and F_3 for the three cameras are determined by mapping the different focal planes into one curve to the positions z_1, z_2 and z_3.

5.3 Focal Plane Estimation

To make the automated microscope as fast as possible, we implemented a predictive autofocus, which adapts itself to different surfaces to investigate. The z-position of the next focal plane is estimated in advance and adjusted together with the x- and y-position of the scanning stage to minimize the number of fields which have to be refocussed. Three different algorithms are available:

1. The next position is equal to the previous one, that means no adjustment is done at all.
2. The last $n = 2...10$ predecessors are taken to compute the next z-position.
3. All previous positions are taken to compute the best fitting plane for z which is a function of x, y.

At the beginning of the scanning process the first algorithm is selected to estimate the z-position of the next focal plane. With an increasing number of scanned fields the statistical models for the second and third algorithms are produced. Then an estimate is made with all three models and the quality of the estimate is gained from the next focussing process. The algorithm with the highest hit rate is selected for the next estimate. Note that the accuracy for the focus adjustment is only as good as the selected minimal ϵ for the correction Δz in z-direction.

With this prediction the system checks to see if a scene is in focus before going through the unnecessary work of refocussing. In a parallel process, image acquisition by the main computer for scene evaluation can be done. Only if the image is not in focus, a correction will be made and another image is grabbed. The time to move the scanning stage to the next field of view is 150 milliseconds. The time to grab a full image in the host computer is 200 milliseconds for one scene. The video timeframe therefore is 40 milliseconds per frame. To refocus a scene we need a multiple of 80 milliseconds depending on the number of correction steps. The number of scenes that have to be refocussed could be cut down to about 10% on the biological samples we used, so an average speed of more than 4 frames per second could be achieved using a standard microscope and standard TV technology. Therefore a fast and cost efficient device could be designed.

References

[Haeu82] G. Häusler, E. Körner: *Expansion of Depth of Focus by "Image De-Puzzling"* Proc 6 Int Conf Pattern Recognition, Munich, 1982

[Krot86] E. P. Krotkov, J. Summers, F. Fuma: *Computing Range with an Active Camera System,* Proc 8 Int Conf Pattern Recognition, Paris 1996

[Krot89] E. P. Krotkov: *Active Computer Vision by Cooperative Focus and Stereo,* 1989

[Ligt82] G. Ligthart, F. C. A. Groen: *A Comparison of Different Autofocus Algorithms,* Proc 6 Int Conf Pattern Recognition, Munich, 1982

[Piep83] R. J. Pieper, A. Korpel: *Image Processing for Extended Depth of Field,* Applied Optics, Vol 22, No. 10, 1983

[Sugi85] S. A. Sugimoto, Y Ichioka: *Digital Composition of Images with Increased Depth of Focus Considering Depth Information,* Applied Optics, Vol 24, No. 14, 1985