JPMTR 107 | 1809
DOI 10.14622/JPMTR-1809
UDC 681.6-004.93|62.5

Research paper
Received: 2018-03-20
Accepted: 2018-06-23

Robust Fourier-based focusing method for post-press inspection

Julian Schäfer, Felipe Fernandes and Edgar Dörsam

Institute of Printing Science and Technology,
Darmstadt University of Technology,
Magdalenenstr. 2, 64289 Darmstadt, Germany

schaefer@idd.tu-darmstadt.de
fernandes@idd.tu-darmstadt.de
doersam@idd.tu-darmstadt.de

Abstract

Post-print optical inspection of printed sheets becomes more and more substantial in job printing in order to reduce both errors and setup time in the cutting process. This paper provides a new method for on-sensor based focusing for cameras using a two-dimensional FFT (Fast Fourier Transform). Deriving from contrast based focusing and two-dimensional FFT an approach is shown which uses the amplitude spectrum provided by FFT. Emphasizing high frequencies and therefore masking the power spectrum results in a method which presents sufficiently accurate focusing on stacks of paper sheets. A variety of tests have been performed to verify the method provided in this paper. Four different types of paper were tested using four different test charts in order to cover a broad span of print layouts found in job printing.

Keywords: power spectrum, two dimensional Fast Fourier Transform, job printing, inline inspection

1. Introduction and background

1.1 Optical inspection in post processing

With the aim to increase efficiency in the printing process, post processing comes into focus as the level of automation is low compared to pre-press and press. Cutting represents a bottleneck in the printing process as succeeding processes such as hemming, gathering and binding rely on the results. As the size of printed sheets in job printing increases to include more jobs on a single printing plate and therefore reduce cost and setup time, minimizing errors during the cutting process becomes even more crucial. Even though job data is available, disruptive factors of the printing process influence the printed sheets resulting in deformation thus the job data does not match the physical data anymore. Consequently, quality inspection is a main target of manufacturers of cutting systems.

The cutting process allows stacks of sheets of various dimensions to run through the process. The maximum dimensions are given by the size of the machine table of the cutting machine, while the minimum sliceable dimensions are usually smaller than 10 cm. While the change of width and depth between jobs means that the measuring range of an inspection system has to be adaptable, the change of the height of the paper stack means that the inspection system has to adjust its focus for every job. In this respect, the inspection task differs from most tasks seen in inline inspection systems where objects and distances usually stay the same as shown for example in Pawlowski (2011).

In this paper we provide a method for focusing with an industrial standard 5 MP area scan camera with no built-in focusing functions by adjusting the distance between the surface of a stack of sheets and the camera using the power spectrum of a two-dimensional Fast Fourier Transform (FFT). The method is tested on different types of paper with different printed test patterns. To optimize the significance of the method and reduce calculation time the frequency range is adjusted.

1.2 On-sensor focusing methods

In order to focus on a subject just by using data from the imaging sensor, a difference in brightness is needed as the imaging sensor only detects brightness. Figure 1 shows the two borderlines. On the left side, a blurred digitalized edge resulting in a brightness gradient

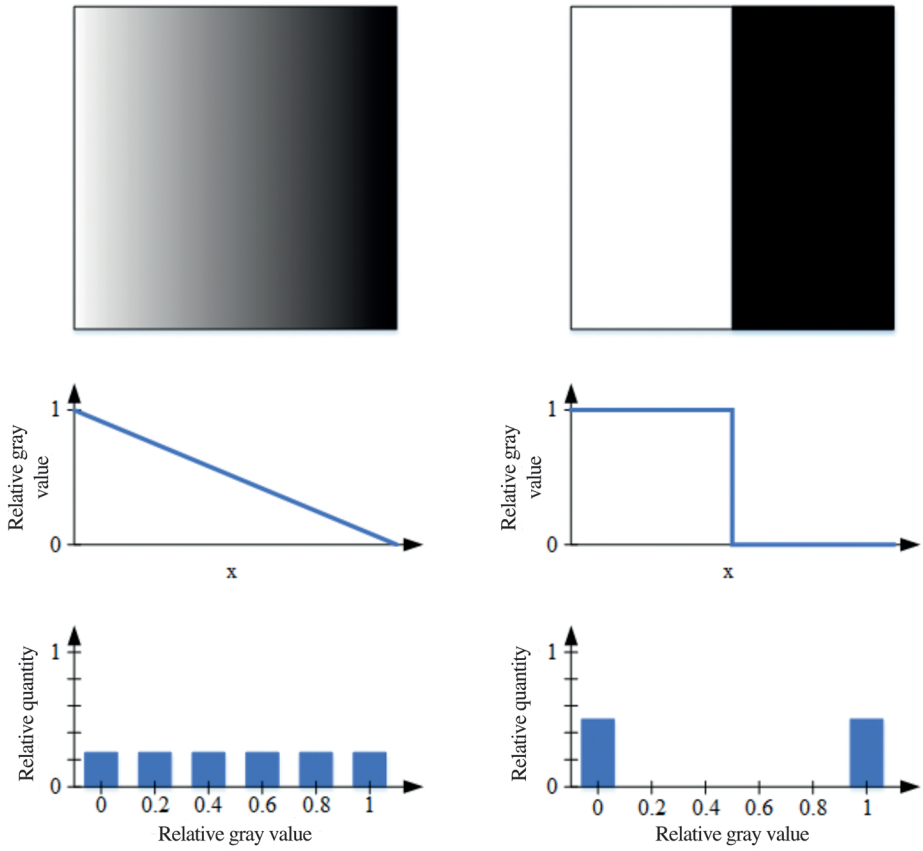

Figure 1: Blurred and sharp edges (left and right, respectively) and their corresponding brightness distribution

is shown with uniformly distributed gray values. On the right side the same edge is digitalized with maximum sharpness. The difference in gray values at the edge and thereby the contrast is at the maximum what allows the assumption that the image is focused.

As seen in Figure 1, the sharpest image is associated with maximum gap in gray values histogram. Evaluation of the gap of an image histogram performs well for simple printing jobs, where the amount of different gray values is at a minimum. However, evaluation gets complicated for an unknown distribution of gray values where the histogram tends to be more uniform. Without a priori knowledge of the printed motive another method is required for focusing a camera on the large variety of printing products.

In order to improve accuracy and robustness, a focusing method based on FFT is provided. Low-pass filters reduce contrast on edges by adapting gray values of pixels according to their neighborhood. As a result, the image gets blurred (Chaudhuri and Rajagopalan, 1999; Utcke and Burkhardt, 1999; Vision & Control, 2007). By implication this means that high frequencies occur most if the image is focused at its best. A popular method in image processing therefore is the two-dimensional FFT (Burger and Burge, 2005).

1.3 Aim of research

The aim of this research is to provide a robust method, which allows adjusting the focus plane independent of the print layout on the sheets by varying the distance between camera and surface of the stack of sheets, when only data from the imaging sensor is available and no further sensor should be used. First, we introduce the method in section 2.1 before the test rig and the test charts which are used are presented in section 2.2. After discussing the results of our research in section 3, we conclude and provide an outlook in section 4.

2. Materials and methods

2.1 FFT-based focusing method

Image processing using the frequency domain offers a variety of opportunities as it allows operating with features which are not detectable in spatial domain (Bredies and Lorenz, 2011). Furthermore, processing effort decreases due to different mathematical approaches such as the FFT (Erhardt, 2008). Figure 2 shows the fitting of an ideal step as seen in Figure 1 on the right side with Fourier series with different parameters.

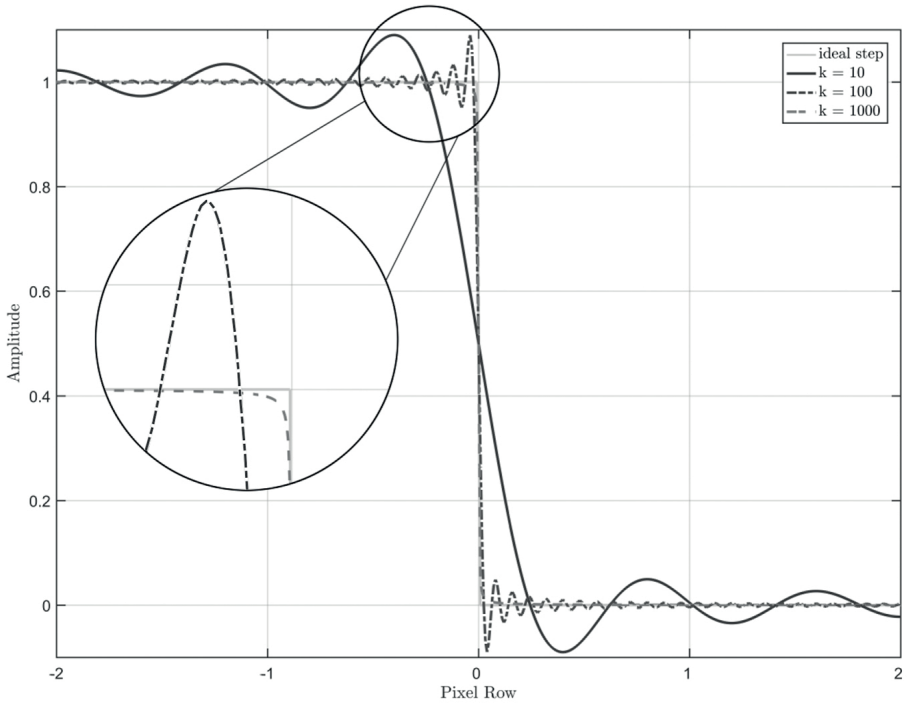

Figure 2: Fitting an ideal step at zero with Fourier series with k = {10,100,1000} terms; crop not true to scale

After capturing an image, the data is available in the spatial domain. Each pixel is usually encoded with 8 bit resulting in 256 gray values. As a result, a gray image $\boldsymbol{G}(x,y)$ captured with an imaging sensor with m pixels in height and n pixels in width is represented by a $m \times n$ matrix. Because of further processing steps, the 256 gray values are in the interval from 0 to 1. Each pixel implies a gray value $g_{m,n}$ representing the brightness at the position specified by m and n of the image $\boldsymbol{G}(x,y)$, as shown in Equation [1].

$$\boldsymbol{G}(x,y) = \begin{pmatrix} g_{1,1} & \cdots & g_{1,n} \\ \vdots & \ddots & \vdots \\ g_{m,1} & \cdots & g_{m,n} \end{pmatrix} \qquad [1]$$

As an image in general is composed by sine and cosine waves, any image is describable as a combination of sine and cosine waves with customized coefficients (Jähne, 2005, p. 43ff.), where u and v represent the frequency of the wave. Every pixel $g_{m,n}$ of the image $\boldsymbol{G}(x,y)$ with $1 \le x \le m$ and $1 \le y \le n$ of an input image is transformed into freqency domain by a discrete two dimensional FFT described by Equation [2]. The resulting complex valued Fourier image $\boldsymbol{F}(u,v)$ has the same size as the image $\boldsymbol{G}(x,y)$.

$$\boldsymbol{F}(u,v) = \frac{1}{\sqrt{m \cdot n}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \boldsymbol{G}(x,y) \cdot \left\{ exp\left(-i2\pi \cdot \left(\frac{u \cdot x}{m} + \frac{v \cdot y}{n} \right) \right) \right\} \qquad [2]$$

Image $\boldsymbol{G}(x,y)$ is multiplied with the exponential function elementwise and added up over rows and columns of the image matrix $\boldsymbol{G}(x,y)$ and then scaled. Sine and cosine as initial functions for FFT are represented through the exponential function in the complex range of numbers. The Fourier image $\boldsymbol{F}(u,v)$ contains the frequency and phase information of the spatial image $\boldsymbol{G}(x,y)$. An example of a spatial image and the corresponding amplitude spectrum is shown in Figure 3. A high gray value in the amplitude spectrum originates by high amplitude of detected FFT frequency $|\boldsymbol{F}(u,v)|$ in image $\boldsymbol{G}(x,y)$ with phase $\arctan(\boldsymbol{F}_{\mathrm{Im}}(u,v)/\boldsymbol{F}_{\mathrm{Re}}(u,v))$, where $\boldsymbol{F}_{\mathrm{Re}}(u,v)$ and $\boldsymbol{F}_{\mathrm{Im}}(u,v)$ are the real and imaginary components, respectively.

The square of the value of the Fourier transformation presented in Equation [3] is called power spectrum according to Süße and Rodner (2014). It describes the contribution of single frequencies u and v to the image $\boldsymbol{G}$.

$$\boldsymbol{P}(u,v) = |\boldsymbol{F}(u,v)|^2 \qquad [3]$$

As can be derived from section 1.2 and Equation [3], the power spectrum at high frequencies is most pronounced in a perfectly focused image. To cut down processing time herein a customized section of the power spectrum is discussed. For this purpose, only the high frequency range $\Psi(b)$ of the spectrum with the low limit frequency b is considered (see Figure 4) which we further call the frame frequencies.

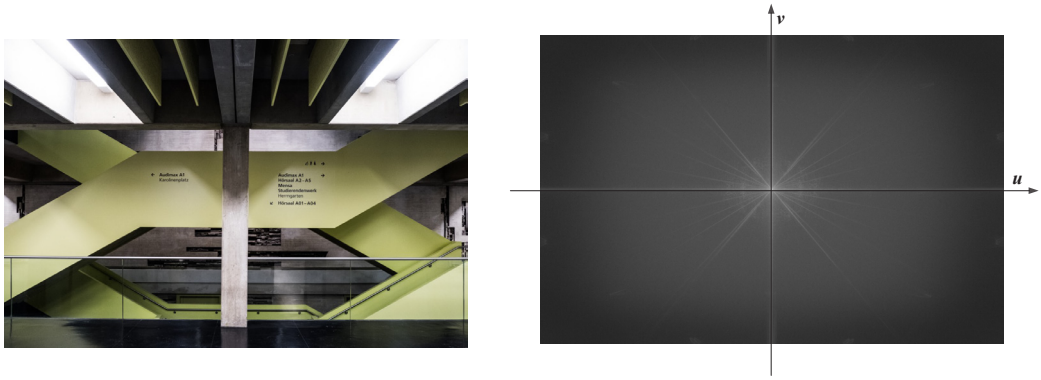

Figure 3: Input image (left) and resulting amplitude spectrum of a two dimensional FFT (right)

For the perfectly focused image three low limit frequencies with different frequency ranges are defined in the image processing, which are $b_{\text{high range}} = 3.7075\ \text{mm}^{-1}$, $b_{\text{medium range}} = 5.0950\ \text{mm}^{-1}$ and $b_{\text{low range}} = 6.4825\ \text{mm}^{-1}$.

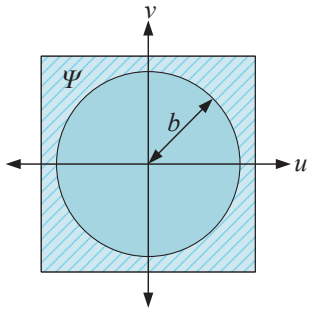

Figure 4: Frame Ψ (hatched area) with low limit frequency b, which is used in the FFT in order to concentrate on high frequencies

By adjustment of Equation [3], the power spectrum of the frame $\Psi(b)$ – further called frame power $\widehat{\boldsymbol{P}}_{\text{b}}$ – is calculated and shown in Equation [4]. Herein u and v represent the frequency of the wave in the Fourier image $\boldsymbol{F}(u,v)$, whereby u, v, x and y are elements of integer numbers $\mathbb{Z}$.

$$\widehat{\boldsymbol{P}}_{\text{b}} = \frac{1}{n_{\Psi(b)}} \sum_{u,v \in \Psi(b)} |\boldsymbol{F}(u,v)|^2 \ \text{with}\ u,v,x,y \in \mathbb{Z} \quad [4]$$

Frame power $\widehat{\boldsymbol{P}}_{\text{b}}$ therefore can be described as a high-pass filtering with additional summation of the elements in the remaining frame $\Psi(b)$ with low limit frequency b of the power spectrum normalized by the number of elements $n_{\Psi(b)}$ in the frame $\Psi(b)$.

2.2 Experimental setup

For the experimental setup the 5 MP industrial monochrome area scan camera P83M-GigE-AS from PicSight (Leutron Vision, 2010) was used in combination with the Fujinon 25 mm f/1.4–22 lens HF25SA-1 (Fujifilm, 2010). As camera and lens are not equipped with an automatic focus or aperture, the focal plane of the optical system can be assumed to be in a constant distance from the imaging sensor. As a compromise between light intensity and accuracy of the setup an aperture of f/4 was used, resulting in a focal depth of ± 4 mm around the actual point of focus (equivalent to a focal plane distance of f = 0 mm). The left side of Figure 5 illustrates the geometry used for identifying the stack's height using the best focus.

In the test rig (see Figure 5 on the right) the camera is mounted to a room gantry which allows us to move and vary the height of the camera above the machine table on which the stack of sheets will be inspected once the best focus for its height is found. For the investigation, the camera is once focused manually at a specific height on a single test-chart the height of which is assumed negligible. The resulting spatial resolution of the optical setup in this camera position is 72 µm per pixel. Subsequently a series of 75 images is taken of every test chart, while the camera is moved 1 mm upwards after the acquisition of each image, beginning from the lowermost camera position. By this, the focal plane is moved through the pre-defined point of focus. Thereby the spatial resolution is inevitably changed for every image. For testing the developed method of focusing through varying the cameras height above the stack and calculating the frame power from the power spectrum, test charts were designed for two categories (see Figure 6). The first category encloses a text sample as well as a color and monochrome picture, which are closely related patterns for print jobs. In the second category, image patterns with generic frequency distributions are used. A stroke pattern essentially contains one frequency with a clear orientation of the phase vector (u,v). In contrast, a broad spectrum extending to high frequencies, but without any preferential orientation is represented by the random binary noise pattern. In addition, it is assumed that the spectral characteristics of a real printing product lie in between of those of the generic frequency test charts. All charts are shown in Figure 6.

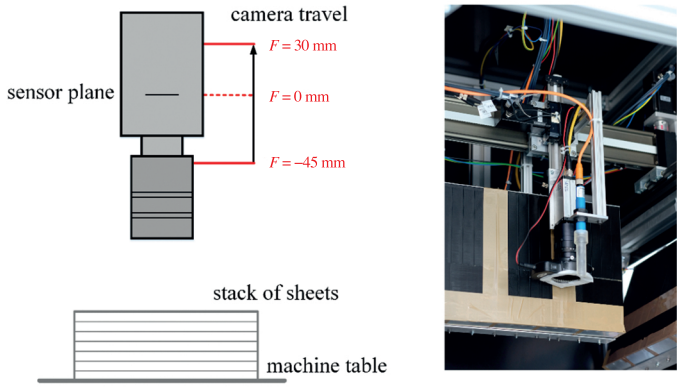

Figure 5: Geometry for identifying the height of the stack of sheets (left) and test rig (right)

Figure 6: Test charts used in the experiment, where the first category is closely related to printing products and encloses an image in monochrome as well as colored and a text pattern, whereas the second category is build up by generic frequency distributions with a stroke pattern and a binary noise pattern

The text pattern is used as a special case of a stroke pattern, and to study the effect of a preferential direction (PD). Therefore, four different orientations were defined with angles 0°, 15°, 45° and 75° to a reference direction at which the text was aligned during the experiment. For investigating the robustness of the algorithm, the test charts were printed on four different types of paper with a laser printer. These include one glossy (Luxo Magic 150 g/m^2), one matte glossy (Zanders Ikono Matt 150 g/m^2), one matte (Tauro Offset 100 g/m^2 SB) and one slightly rough paper (Lorsatzpaper 100 g/m^2 SB). A constant lighting setting was established by using a circular LED camera light.

3. Results and discussion

The frame power $\widehat{\boldsymbol{P}}_{\mathrm{b}}$ is calculated for each image of the series and scaled to the maximum frame power of the series $\widehat{\boldsymbol{P}}_{\mathrm{b,\,max}}$. By implication, we assume that the test chart is focused when the maximum frame power $\widehat{\boldsymbol{P}}_{\mathrm{b}}$ is achieved and that it is in between the borders of focal depth.

All results shown are based on the matte Tauro Offset paper, but for the comparison between matte/glossy, the glossy Luxo Magic paper is used in addition. Results for the three other papers are very similar to the results on Tauro Offset paper.

When comparing different low limit frequencies b = {3.7075 mm^{-1}; 5.0950 mm^{-1}; 6.4825 mm^{-1}} (listed from high range to low range) from two different test charts (binary noise and text aligned at 0°), the results show a very similar appearance qualitatively before entering the focal depth (see Figure 7).

The magnitude increases with a larger frame range for both test patterns in the same way. For a low limit frequency of b = 6.4825 mm^{-1} resulting in the smallest frame the lowest frame power is observed as well as for a low limit frequency of b = 3.7075 mm^{-1} resulting in the largest frame the largest overall frame power is observed. Therefore, a lower low limit frequency also considers lower frequencies, which on the one side are detectable earlier even when the test pattern is not in focus and secondly the percentage of lower frequencies in image data is higher by implication. We can confirm that the focus is found with each of the considered low limit frequencies, but because of the most distinctive results and reasons of computational speed upcoming experiments are only done with a high frequency range, considering frequencies from b = 3.7075 mm^{-1}.

Results of frame power obtained from color and monochrome pictures show a similar appearance before entering the focal depth, but the absolute magnitude of the frame power $\widehat{\boldsymbol{P}}_{\mathrm{high\ range}}$ is larger for the colored picture (see Figure 8).

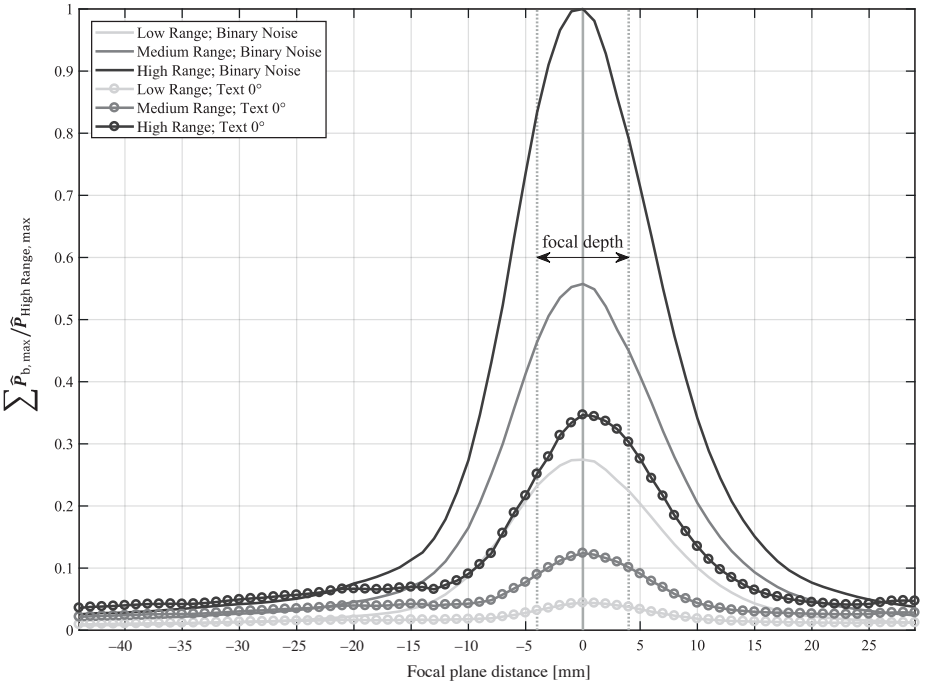

Figure 7: Power spectra of binary noise (solid lines) and text aligned at 0° (circled lines) test chart using Low, Medium and High Ranges of frequency bandwidths

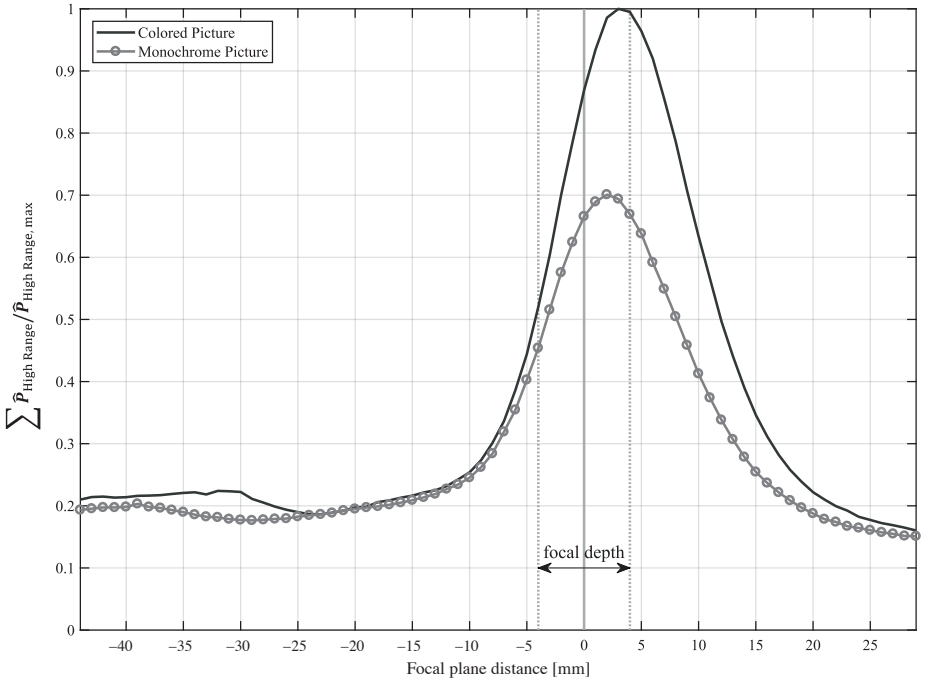

Figure 8: Power spectra of colored picture and monochrome picture test chart using the high frequency range

Digitalized image of the colored picture is noisier than the monochrome picture as cutouts show in Figure 9. Also differences in the printing screens of colored and monochrome picture printouts can be detected in visual control. For a distinct association further investigations are required.

Continuing, in the experiments a major influence due to different angles of text cannot be detected. Figure 10 shows the results of text aligned at 0°, 15°, 45° and 75° using the large frame range with a low limit frequency of b = 3.7075 mm^{-1}. These results hereby match the statement that frequency spectra are invariant of input

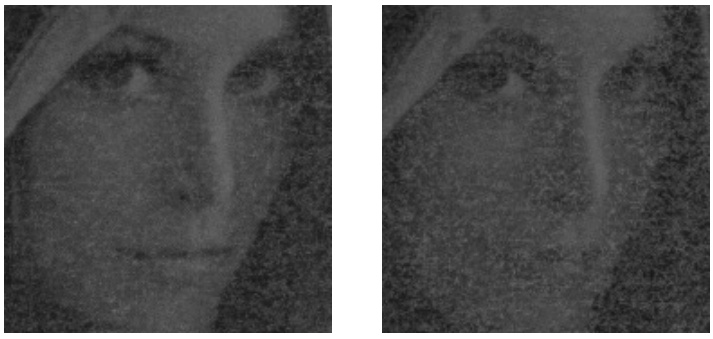

Figure 9: Cutouts of the digitalized monochrome (left) and colored (right) picture test pattern using a monochrome camera

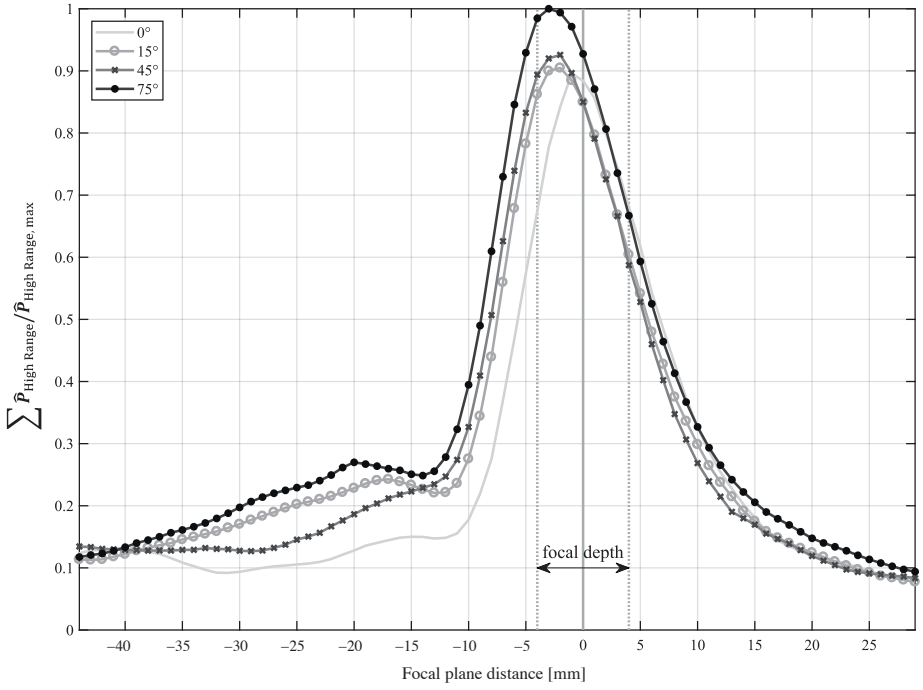

Figure 10: Power spectra of text test chart aligned at 0°, 15°, 45° and 75°, using the large frequency range

image rotation (Ó Ruanaidh and Pun, 1998). Beside a global peak of the frame power $\hat{\boldsymbol{P}}_{\text{high range}}$, small local peaks are existent (see Figure 9), when using the sample text pattern.

Figure 11 shows the results of the sensor filling stroke pattern and a windowed stroke pattern. The behavior of local peaks is even more remarkable and invariable which is noticeable for all investigations using the stroke pattern, when it fills the camera's sensor completely (see solid line in Figure 11). The deviation between the maximum frame power $\hat{\boldsymbol{P}}_{\text{high range, max}}$ and the actual point of focus f = 0 mm, which is defined by $|\Delta f(\hat{\boldsymbol{P}}_{\text{b}})|$, goes up to $|\Delta f(\hat{\boldsymbol{P}}_{\text{high range, max}})|$ = 5 mm. This behavior occurs because of strokes entering the image due to upwards camera travel resulting in a larger acquisition area (see Figure 12 on the left side). Reducing the printing size of the test chart so that the complete pattern is detected in the lowermost camera position (see Figure 12 on the right side) prevents local peaks from occurring. The magnitude of the power spectrum of the windowed stroke pattern is much lower because the region of interest, or more precisely the area that contains a frequency pattern is much smaller. Even though the frequency and amplitude of the windowed stroke pattern is the same, less harmonics are found by our FFT algorithm in the reduced area, what decreases the power spectrum by more than 80 %.

Nevertheless, we can confirm that windowing the stroke pattern and hereby preventing strokes from entering the acquisition area eliminates local peaks. Furthermore the accuracy increases as the deviation between maximum frame power $\hat{\boldsymbol{P}}_{\text{high range, max}}$ and the actual point of focus f is reduced by 1 mm to $|\Delta f(\hat{\boldsymbol{P}}_{\text{high range, max}})|$ = 4 mm.

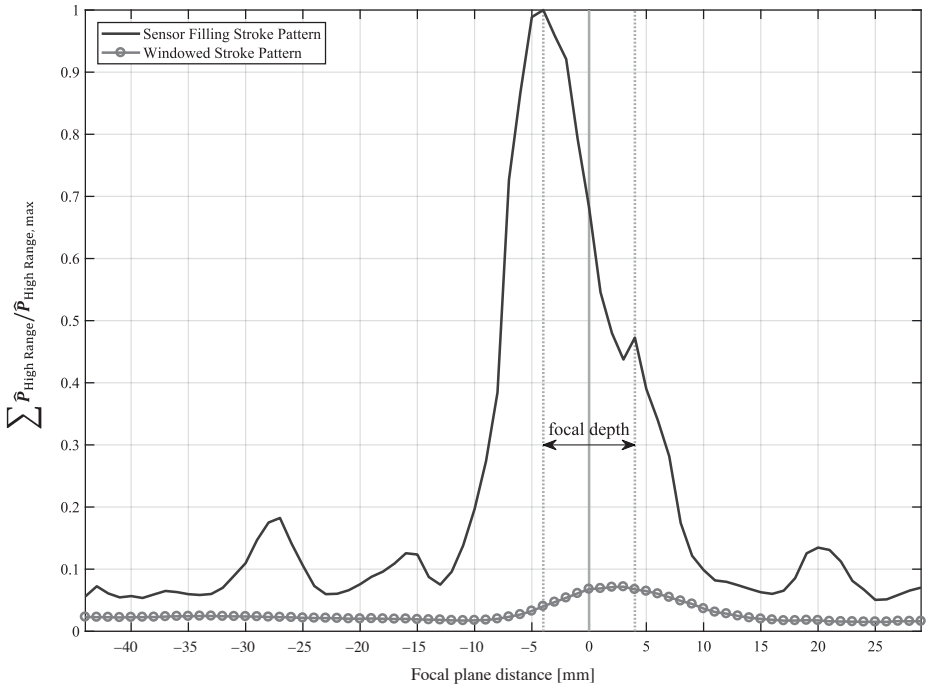

Figure 11: Power spectra of sensor filling stroke pattern and windowed stroke pattern showing local peaks besides a global maximum occur with the sensor filling stroke pattern, while the windowed stroke pattern is preventing local peaks in power spectrum; a large frequency range is used

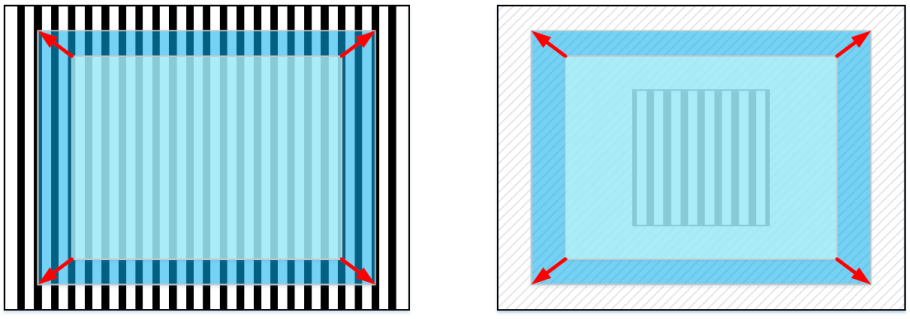

Figure 12: Sensor filling stroke pattern (left) and windowed stroke pattern (right); the acquisition area in the lowermost camera position is highlighted in light blue, in the uppermost position in blue, the red arrows mark the enlargement of the acquisition area

However, we assume that an image sensor filling stroke pattern is not found frequently in real printing products, therefore the suitability of the introduced method is not less applicable because of these results.

Contrary to the indistinct results of the sensor filling stroke pattern, outcomes of the binary noise pattern are highly convergent (see Figure 13). In visual control, we detect a major influence caused by the circular light though. Because of parallelism of lighting direction and the cameras optical axis, partial reflections of the light source are captured in the images when using glossy papers. This biases the results of FFT and thus the power spectrum. For correction a light diffusor is installed which prevents sharp reflections on glossy surfaces. The reflectance of the paper sheet influences the absolute values of the power spectrum as well. The frame power $\hat{\boldsymbol{P}}_{\text{high range}}$ is raised by usage of glossy papers without exception (see Figure 13), what possibly is explained through higher contrast. Hence a resulting higher contrast of the image leads to higher amplitudes of the corresponding FFT functions, what implies larger absolute values of the FFT and thereby the power spectrum (see Figure 11). Deviations of $\hat{\boldsymbol{P}}_{\text{high range max}}$ to the actual point of focus, when using the binary noise test chart, lie in a span of $\Delta f(\hat{\boldsymbol{P}}_{\text{high range max}}) = \pm 1$ mm. Out of focus, small structures of the noise pattern become indistinct, hence the edge contrast is reduced dras-

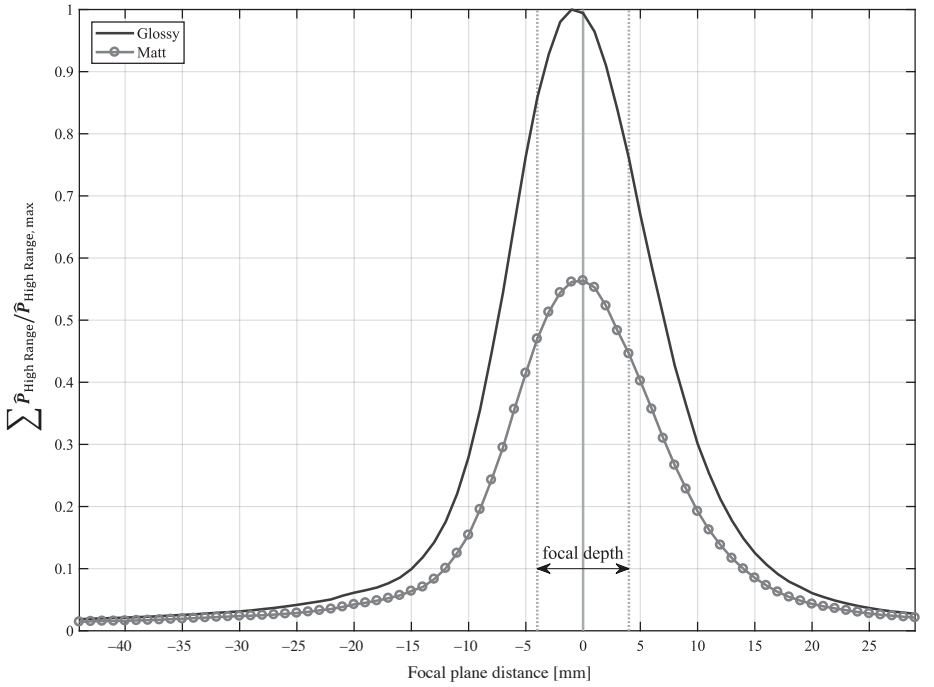

Figure 13: Power spectra of random binary noise pattern for glossy and matt paper using a large frequency range

tically. Inside the focal depth, structures are distinguishable and the high spatial resolution of the pattern results in a sudden detection of high frequencies. Out of all experiments, the results of the noise pattern show the best results. Nevertheless scaling the frame power to the maximum of the series of a single test pattern is essential for obtaining a distinct statement of which image is focused best. Yet we can confirm that the presented algorithm found a maximum frame power for each series of images inside the focal depth of the acquisition assembly for all the used test patterns.

4. Conclusions

The method shown in this research serves as a robust alternative for focusing tasks where only data from the imaging sensor is available. No further sensor is needed within the system. The method was evaluated by using four different test charts representing a large variety of print layouts found in printing jobs. Deviations of $\widehat{\boldsymbol{P}}_{\text{b, max}}$ to the actual point of focus for all experiments done lie in a span of $\Delta f(\widehat{\boldsymbol{P}}_{\text{b, max}}) = \pm 4\text{mm}$. While a decrease in low limit frequency and thus considering also lower frequencies results in a higher frame power $\widehat{\boldsymbol{P}}_{\text{b}}$, even a low limit frequency of just $b = 6.4825\ \text{mm}^{-1}$ results in precise prediction of the actual focus. However, for the most distinctive results and because of computational speed a low limit frequency of $b = 3.7075\ \text{mm}^{-1}$ is considered the best.

A variety of glossy and matte papers were tested and even though glossy paper caused difficulties with partial reflections due to circular lighting and power spectrum performed on matte sheets showed lower maxima, the method still worked sufficiently.

Acknowledgments

The authors gratefully acknowledge the financial support by the AiF (of the German Federal Ministry of Economic Affairs and Energy) within the framework of the ZIM (Zentrales Innovationsprogramm Mittelstand), grant # KF2012459PT3.

References

Bredies, K. and Lorenz, D., 2011. *Mathematische Bildverarbeitung: Einführung in Grundlagen und moderne Theorie*. Wiesbaden: Vieweg+Teubner Verlag / Springer Fachmedien Wiesbaden GmbH.

Burger, W. and Burge, M.J., 2005. *Digitale Bildverarbeitung: Eine Einführung mit Java und ImageJ*. Berlin, Heidelberg: Springer-Verlag.

Chaudhuri, S. and Rajagopalan, A.N., 1999. *Depth from defocus: a real aperture imaging approach*. New York, NY: Springer-Verlag.

Erhardt, A., 2008. *Einführung in die Digitale Bildverarbeitung: Grundlagen, Systeme und Anwendungen*. Wiesbaden: Vieweg+Teubner / GWV Fachverlage GmbH.

Fujifilm, 2010. *Fujinon CCTV lens: for FA/machine vision*. [pdf] Fujifilm Corporation. Available at: <http://www.fujifilmusa.com/shared/bin/FUJINON-FA-MV-CCTV-LENS-Brochure.pdf> [Accessed 15 February 2017].

Jähne, B., 2005. *Digitale Bildverarbeitung: 6., überarbeitete und erweiterte Auflage*. Berlin, Heidelberg: Springer-Verlag.

Leutron Vision, 2010, *PicSight P83M-GigE-AS GigE vision camera*. [pdf] Leutron Vision AG. Available at: <http://www.leutron.com/english/product/picsight/gige/P83M-GigE.pdf> [Accessed 15 February 2017].

Pawlowski, K., 2011. *Untersuchungen zur Bestimmung der Farbdichte druckfrischer Offsetfarben mit einem Druckinspektionssystem*. Göttingen: Sierke Verlag.

Ó Ruanaidh, J.J.K. and Pun, T., 1998. Rotation, scale and translation invariant spread spectrum digital image watermarking. *Signal Processing*, 66(3), pp. 303–317.

Süße, H. and Rodner E., 2014. *Bildverarbeitung und Objekterkennung: Computer Vision in Industrie und Medizin*. Wiesbaden: Springer Vieweg.

Utcke, S. and Burkhardt H., 1999. *Praktikumsversuch: Autofokus*. [pdf] Albert-Ludwigs-Universität Freiburg. Available at: <http://lmb.informatik.uni-freiburg.de/lectures/praktika/BVPraktikum-I/autofokus/autofokus.ps.gz> [Accessed 29 May 2016].

Vision & Control, 2007. *Tutorial: Filter*. [pdf] Vision and Control GmbH. Available at: <http://www.vision-control.com/uploads/media/TU_Softwarefilter_de_2.0.0_02.pdf> [Accessed 15 February 2017].