Optics Communications 410 (2018) 269–276

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Improving the performances of autofocus based on adaptive retina-like sampling model

Qun Hao [a], Yuqing Xiao [a], Jie Cao [a,b,*], Yang Cheng [a], Ce Sun [a]

[a] Beijing Key Lab. for Precision Optoelectronic Measurement Instrument and Technology, School of Optoelectronics, Beijing Institute of Technology, Beijing, 100081, China

[b] NUS Suzhou Research Institute (NUSRI), Suzhou, Industrial Park, Suzhou, 215123, China

ARTICLE INFO

Keywords:
Retina-like sampling
Accuracy
FWHM
Autofocus functions
Real-time

ABSTRACT

An adaptive retina-like sampling model (ARSM) is proposed to balance autofocusing accuracy and efficiency. Based on the model, we carry out comparative experiments between the proposed method and the traditional method in terms of accuracy, the full width of the half maxima (FWHM) and time consumption. Results show that the performances of our method are better than that of the traditional method. Meanwhile, typical autofocus functions, including sum-modified-Laplacian (SML), Laplacian (LAP), Midfrequency-DCT (MDCT) and Absolute Tenengrad (ATEN) are compared through comparative experiments. The smallest FWHM is obtained by the use of LAP, which is more suitable for evaluating accuracy than other autofocus functions. The autofocus function of MDCT is most suitable to evaluate the real-time ability.

1. Introduction

Autofocusing optical system has received much attention in recent years due to the advantages of adjustable focal length, simple structure and high accuracy [1–3]. It plays an important role in imaging system and has been applied in various fields, such as augmented reality [4,5], microscopic imaging [6,7] and target tracking [8,9]. To the best of our knowledge, autofocusing method can be divided into two categories, namely active and passive autofocusing methods. Active autofocusing method employs a launch equipment to actively emit laser light or waves (infrared or ultrasound) [10–12]. Although high autofocusing efficiency is obtained, the remarkable drawback of this method is low autofocusing accuracy [13]. Unlike the active autofocusing method, the passive autofocusing method does not require external equipments, therefore it is more concise than active autofocusing method. The passive autofocusing method adjusts the focal length according to the sharpness of the digital image acquired by the imaging sensor, and the edge information is obtained by algorithms to estimate the image quality [14]. Unfortunately, the current passive autofocusing method uses uniform sampling method, i.e., target and background are sampled in uniform resolution. Actually, for practical use, not whole scene in the whole field of view (FOV) is interested. For target, it should be sampled with high sample resolution, because those sampling areas contain more useful information than background. But for background, it is not necessary to sample in high resolution as the target, because it can result in low autofocusing efficiency.

Inspired by sampling structure of the human eye, retina-like sampling model provides an available and effective solution for the aforementioned issue. It is well known that photosensitive cells in the human retinas is non-uniform distribution [15,16], i.e., high resolution in the fovea and low resolution in the periphery. The retina-like sampling structure has a remarkable feature compressing redundant information in the periphery while keep high resolution in fovea [17,18], and the feature has been employed to improve the sampling efficiency [19,20]. Our group has studied retina-like sampling model in autofocusing application and developed basic theory [21], but the sampling window and adaptability of typical autofocus functions are not studied.

Therefore, the main contribution of this paper is proposing an adaptive retina-like sampling model (ARSM) and carry out experiments to study which autofocus function is more suitable for the ARSM. The rest of paper is organized as follows: the principle and theoretical analysis are illustrated in Section 2, including adaptive retina-like model and flow chart of adaptive sampling window. In Section 3, we carry out the comparative experiments between the adaptive retina-like sampling and the uniform sampling, then analyze the effect of different autofocus

* Corresponding author at: Beijing Key Lab. for Precision Optoelectronic Measurement Instrument and Technology, School of Optoelectronics, Beijing Institute of Technology, Beijing, 100081, China.

E-mail addresses: qhao@bit.edu.cn (Q. Hao), ajieanyyn@163.com (J. Cao).

https://doi.org/10.1016/j.optcom.2017.10.017
Received 11 July 2017; Received in revised form 29 September 2017; Accepted 8 October 2017
0030-4018/

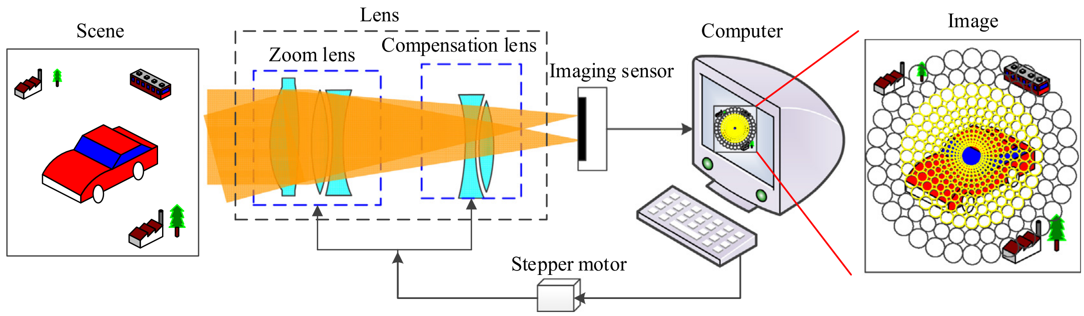

Fig. 1. The components of the typical autofocusing system.

functions in the ARSM. The optimal retina-like sampling structure is discussed in Section 4. Finally, the conclusions are listed in Section 5.

2. Methods

2.1. Principle

The components of the typical autofocusing system is shown in Fig. 1. The system includes lens (zoom lens and compensation lens), imaging sensor, computer and stepper motor. The zoom lens and compensation lens are used to achieve zoom and focus. The scene is imaged on the imaging sensor. Then the image information is sampled and sent to computer. The computer uses autofocus function to estimate sharpness of the image and control stepper motor to adjust the position of compensation lens based on focus value, until the highest sharpness of image is obtained.

The comparison between traditional method and proposed method is shown in Fig. 2. The contents of the black dashed box are traditional method, the scene including target (car) and background (trees, factories and bus) are uniformly sampled in the Cartesian plane, and focusing value is estimated by all pixels in the scene, which results in low autofocusing efficiency. Actually, only the target is used to calculate focused rather than background. Different from the uniform sampling method, the proposed method employs ARSM to sample the scene, shown in red dashed box. Firstly, the proposed method employs retina-like model to sample image, and the area of the sampling unit increases with the increasing of the ring radius. Secondly, the proposed method optimizes sampling window to ensure that background does not participate in calculating focus value, which is helpful in eliminating background information interference. Thirdly, target in optimal sampling window (red circle) proceeds Log-polar transformation (LPT) based on the retina-like structure, which is beneficial of compressing volume data significantly. Then autofocus function is employed to calculate focus value, and the sharpness of each image is estimated by the corresponding focus value, i.e., F_1, F_2, …, F_k. The focus position is obtained when focus value reaches maximum. From Fig. 2 we can see the difference between two methods is that the proposed method uses optimal sampling window and LPT to perform autofocus calculation and obtains focus value. That is the remarkable feature of proposed method. Although the process of retina-like structure seems more complicated than traditional method, such feature can obtain better performances than traditional method through adaptive retina-like sampling model and LPT. Besides that, with the development of retina-like sensor based on hardware [22], LPT image is obtained directly by sensor. Therefore, the efficiency is improved further.

2.2. Adaptive retina-like model

The human retina is non-uniform distribution which is over-sampling in the fovea (keep high resolution of target) and under-sampling in the periphery (provide situational awareness and context). Such feature is capable of possessing a high resolution of the target while compressing redundant background information. Therefore, it is more efficient due to using limited data and can be implemented as an image processing technique for information transmission [23]. It is worth noting that the relationship between retina and visual cortex in human visual system meets the log-polar law [21], shown in Fig. 3, which is written as

$$\begin{cases} q = \dfrac{1+\sin(\pi/N)}{1-\sin(\pi/N)} \\ r_1 = \dfrac{r_0}{1-\sin(\pi/N)} \\ D_1 = \dfrac{2r_0\sin(\pi/N)}{1+\sin(\pi/N)} \\ r_i = r_1 \cdot q^{i-1} \\ r_{\max} = r_0 \cdot q^M \\ D_i = q^{i-1}\cdot D_1 \\ \xi_i = \log_q(r_i) = \log_q(r_1) + i - 1 \quad (i = 1,2,3\ldots M) \\ \theta_j = \alpha_j = \dfrac{2\pi}{N}\cdot j \quad (j = 1,2,3\cdots N), \end{cases} \tag{1}$$

where q is the increasing coefficient between adjacent rings, r_0 is the radius of the blind area, $r_{\max}$ is the maximum radius of the adaptive sampling window, r_i is the radius of the ith ring and D_i is the diameter of the ith pixel of the LPT, M is rings and N is sectors in each ring. The angle between neighbor pixels is $2\pi/N$.

In terms of the adaptive sampling window of the proposed method, the size of the adaptive sampling window is determined by the image intensity distribution. Calculating the intensity distribution of the image and determining the size of the adaptive sampling window are performed in the Logarithmic polar coordinates. For arbitrary point $(x, y)((x, y)\in(X,Y))$ in Cartesian coordinates, it is represented by Logarithmic polar coordinates $(\upsilon, v)(\upsilon = 1,2,3,\ldots, N;\ v = 1,2,3,\ldots, M.)$. Under the Logarithmic polar coordinates, statistical dispersion (SD) is computed in initial adaptive sampling window, which is written as

$$SD = \sum_{\upsilon=0.7\sigma M}^{M} \sum_{v=1}^{N} |U(\upsilon, v) - \mu|, \tag{2}$$

where σ is the ratio between the ring number occupied by the target and the total ring number of the image, and it can be obtained by prior knowledge. Too large initial ring number decreases autofocusing accuracy and too small initial ring number increases the calculation consumption and influences the real-time ability of autofocus. It is reasonable to select the initial ring number between 60% and 80%. Here, we choose 70% of the ring number occupied by the target as the initial ring number to calculate the SD in our experiments. SD is the absolute difference between the intensity value $U(\upsilon, v)$ and its intensity median value μ. The specific process is comparing SD with a threshold T [24], which is the intensity of the target. Small SD value indicates low texture area and few high-frequency information, it cannot to achieve precise focus value. When the SD is less than the threshold T, the vertical coordinate v (ring number) in the Logarithmic polar coordinates system