IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 4, 2020, accepted June 27, 2020, date of publication July 1, 2020, date of current version July 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3006356

Passive Autofocusing System for a Thermal Camera

RASHID ALI[1], PENG YUNFENG[1], (Member, IEEE), AHMAD ALI[2], HAIDER ALI[3], NAEEM AKHTER[3], JAVED AHMED[4], (Member, IEEE), AND ABDUL JALIL[5]

[1]School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), Beijing 100083, China
[2]Department of Software Engineering, Bahria University, Islamabad 44000, Pakistan
[3]Department of Computer and Information Sciences, PIEAS, Islamabad 44000, Pakistan
[4]Military College of Signals, NUST, Islamabad 44000, Pakistan
[5]Department of Electrical Engineering, FET, International Islamic University, Islamabad 44000, Pakistan

Corresponding author: Rashid Ali (b20150558@xs.ustb.edu.cn)

This work was supported by the GIT Key Laboratory Foundation under Grant 2003008002.

ABSTRACT Most of the Thermal (Infrared) cameras nowadays are equipped with a motorized lens for focusing a scene manually. The subjective nature of manual focusing makes it an inefficient and cumbersome process. In contrast, Autofocusing (AF) obtains the best focused image based on a quantitative measure with the benefits of convenience and intelligence. Various AF systems for visual cameras have been developed, but relatively less amount of work has been done for thermal imaging systems. This paper presents a Vision and Control based Autofocusing System (VCAFS) comprising: (1) an uncooled thermal camera with motorized lens, (2) a passive contrast-based focus measure, (3) a smoothing operator to avoid local extrema, and (4) two different lens motion controllers. Experimental results show the efficacy of the proposed system on live videos even when the scene and its depth are continuously changing.

INDEX TERMS Contrast detection, focus measures, passive autofocusing, thermal imaging system.

I. INTRODUCTION

Image acquisition is an important task in a wide variety of applications such as security, medical science, law enforcement, agriculture, entertainment and power industry, resulting in an increase in popularity of digital cameras. Effectiveness of the image acquisition directly relates to the quality of acquired image. Objects tend to appear sharper in a well-focused image in contrast to being blurry in a defocused image. The degree of sharpness of an image is a function of the distance between camera lens and its imaging sensor [1]. This distance [2] can be controlled automatically by lens' movement performed by a special system of digital camera called the Autofocus System (AFS). A typical AFS consists of a hardware part responsible for lens movement acting on the commands provided by the software part. In this paper we propose a framework for AF of a thermal camera.

Modern day thermal cameras include a motorized lens which can be controlled manually, or automatically to focus on a certain scene. Numerous AFSs have been developed for visual cameras but only few exist for thermal imaging systems. The most recent work on autofocus for thermal imagery combines IR and visible light for microscope focus correction [3]. Pinkard *et al.* [4] proposed CNN based autofocus architecture for microscope; it learns the physics of illumination to find the best focus. Juočas *et al.* [5] developed an array of low cost cameras multi-focus fusion of microscopic imagery. Najibi *et al.* [6] proposed a deep learning based autofocus framework to accelerate multi-scale inference for generic object detection. Sumon and Ratnakirti [7] provided a comparison of different autofocus techniques for digital techniques. The most recent work related to our research was done by Chunping *et al.* [8]. It provides comparison of different image estimation methods for daylight dataset of images of different scenes; lacks control part of autofocus.

We have proposed a Vision and Control based Autofocus System (VCAFS) for thermal cameras. Thermal cameras for long-range surveillance usually do not come with autofocusing, because it further increases their cost. Furthermore, autofocusing algorithms are heavily dependent on the nature of the scene. VCAFS is comprised of (1) an uncooled thermal camera with motorized lens, (2) a passive contrast-based focus measure, (3) a smoothing operator to avoid local extrema, and (4) two different lens motion controllers.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaoqing Pan.

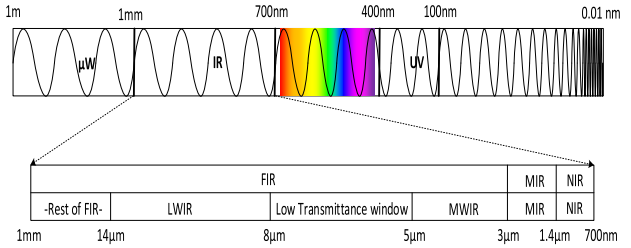

FIGURE 1. Electromagnetic spectrum [9].

A. VISIBLE VS. THERMAL IMAGING

Visual cameras capture images by sensing visible light reflected from an object's surface and are therefore dependent on scene's illumination, direction, color balance etc. providing unsatisfactory information, especially in dark environments. These limitations can be overcome by using sensors which respond to the energy radiated by a body instead of reflected by it. Thermal cameras are equipped with such sensors to capture images in IR range of electromagnetic (EM) spectrum which contains most of the thermal energy emitted by a body. This thermal energy is in accordance to the black body radiation law and is a direct function of object's temperature; the hotter the object, the more the thermal radiation it emits. IR spectrum spanning the wavelengths from 700 nm to 1 mm is divided into three main bands: Near Infrared (NIR), Medium Infrared (MIR) and Far Infrared (FIR). Figure 1 [9] illustrates these bands and their sub-bands.

Thermographic cameras are classified into two types based on the type of infrared detector incorporated: uncooled or cooled. More commonly used uncooled thermal cameras operate at ambient temperature and have lower image quality. However, they are quieter, smaller and less expensive. On the contrary, cooled thermal cameras are more expensive and larger due to the inclusion of cryogenic coolers and offer the advantages of better image quality and improved sensitivity.

B. AUTOFOCUSING

Focusing is achieved by moving focal plane of the lens to a position dictated by the thin lens formula [10] so that the object appears the sharpest in the image. The lens can be moved either manually (called Manual Focusing (MF)) or automatically by the camera system (called Autofocusing (AF)). MF requires user intervention, sufficient skills, is tedious, complicated and time consuming and has higher probability of error thereby resulting in blurry images. AF, on the other hand, provides the benefits of minimal user involvement and faster and easier implementation, leading to the best quality image. An AFS mainly consists of a sensor, a motorized lens and a control system to rotate the motor so that the distance between sensor and lens can be adjusted to focus a certain scene. AF techniques are broadly classified into two categories depending on the type of input provided to the lens control system. These are explained below.

1) ACTIVE AUTOFOCUSING

An Active Autofocus System [10] makes use of an energy transmitting device to decide whether an object is in focus or not. The transmitter (IR or ultrasonic) located on the camera emits a beam which gets reflected by the object of interest and reaches the receiver situated on the camera. The time difference is measured which is used to calculate the distance between camera lens and object of interest.

2) PASSIVE AUTOFOCUSING

No active sensors are required in Passive Autofocus System because they depend upon information obtained from the image itself. This information is in the form of several metrics (or measures) calculated by image processing algorithms. Passive AF is achieved by either phase detection or contrast detection. Phase detection makes use of a secondary mirror and micro-lens to split the incoming light and directs it onto an autofocus (AF) sensor producing a couple of images which are then compared to find similar intensity patterns. The comparison enables us to calculate both the direction and amount of lens movement. Contrast-based AFS does not require the use of AF sensor, secondary mirror or micro-lens because the image captured by the main sensor is used directly. It consists of three steps:

i. Selection of focusing region.
ii. Computation of focus measure.
iii. Search for extremum.

Value of the focus measure, corresponding to image captured at each focus position, is obtained which should ideally produce a Gaussian like focus function.

C. LITERATURE OVERVIEW

A lot of work has been done on the comparison, evaluation and selection of different focusing operators. The earliest comparison was done by Groen in [1] in which 11 different focusing operators were evaluated. A noticeable work from implementation point of view was done by Krotkov in [11]. Autofocusing in the field of computer microscopy was investigated in [12] by Sun. Some novel and robust focusing operators were proposed by Lee in [13] and Joen in [14]. Work on autofocusing operators, specifically for thermal images, has been done in [9] and [15]. Chen and van Beek [16] employs the technique of supervised machine learning to address the problem of finding optimal lens position. A recent literature overview comparing and evaluating various focus measures and search methodologies for both visible and thermal cameras was done in [17]. In [18], design and implementation of a real time, fixed step size autofocus system for thermal imagers is presented.

Autofocusing operators, reported in literature, can mainly be divided into two categories: spatial and frequency domain operators. Gradient [11], histogram [12] and statistics based measures [19] fall in the category of spatial domain operators while Discrete Cosine Transform (DCT) [20], Fast Fourier Transform (FFT) [21] and Discrete Wavelet

Transform (DWT) [22] are the frequency domain operators. Spatial domain operators are generally faster owing to the reason that they operate directly on gray values of an image. On the other hand, frequency domain operators are more robust to noise and perform better than spatial counter parts but because of their high computational complexity, they are rarely used in real time applications.

D. CONTRIBUTIONS OF THE PAPER

As evident from literature review in the previous section, most of the autofocusing work has been based on the comparison and evaluation of focus measures using databases, with little attention to their implementation in real time systems. Moreover, relatively less work has been done in autofocusing of IR camera due to less features in its imagery.

In this paper, we compare different focusing methods and find a suitable focusing measure for thermal imagery; we propose a control framework for autofocusing a thermal camera (which does not provide feedback about its lens position) on live video feed with continuously changing scene and depth. A bang-bang controller in two different configurations based on fixed and adaptive step size, along with a moving average filter to avoid local extrema, is devised and implemented to control the motorized lens of an uncooled thermal camera.

II. FOCUSING OPERATORS

In this section, we will first describe the desired characteristics of an ideal focusing function and then explain some of the selected focusing operators. These focusing operators are selected from different categories established from literature. A lot of focusing operators were tested specifically for thermal images and only the best performing from different categories are presented here. The selected operators include one blob detector called Laplacian Blob Detector (LBD) [23], one feature detector called Fast Feature Detector (FFD), two statistical measures including variance and entropy and a gradient based operator called Tenengrad.

A. REQUIRED PROPERTIES OF FOCUSING OPERATOR

To determine the best focus position of a certain scene, we must define a certain measure which can exhibit the sharpness content of that scene to a certain approximation. In essence, we should have a focus value corresponding to each focus position producing a focus function as shown in Figure 2 [1]. Following criteria are used in the selection and evaluation of a suitable focus function.

a. Unimodality: The focus function should have only one extremum which helps to avoid any local extremum.
b. Focus resolution (ε): The extremum must correspond to the best focused image, so ε should be ideally zero.
c. Range (η): The tail of the extremum must be as broad as possible so that in-focus image can be obtained over a wide range. In other words, η (width at a low percentage l of the maximum) should be large.

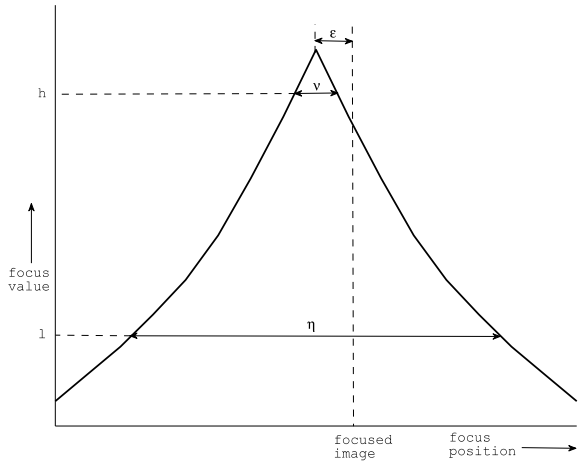

FIGURE 2. Focus function [1].

d. Width (υ): Extremum should be sharp meaning that υ (width at a high percentage h of the maximum) should have a small value.
e. Reproducibility: Good reproducibility is ensured by a sharp extremum.
f. Robustness: The algorithm used to obtain the focus function should be robust to noise, camera jitter, etc.
g. General applicability: The algorithm should be generic enough to work on a wide variety of images.
h. Implementation: The focus function generating algorithm should not be computation intensive.

B. FOCUSING OPERATORS UNDER CONSIDERATION

1) LAPLACIAN BLOB DETECTOR

Laplacian Blob Detector (LBD) [23], one of the most common blob detectors, is based on Laplacian of Gaussian (LoG). In this technique, the input image is first convolved with a Gaussian kernel at a particular scale to get a state space representation [23] which when passed through a Laplacian operator produces strong response for blobs. This procedure is repeated for different scales of image to detect blobs of all sizes. Total numbers of blobs, representing our focus measure, are obtained by adding no. of blobs at all scales.

2) FAST FEATURE DETECTOR

Features from Accelerated Segment Test (FAST) [24], [25] is a corner detection method. It extracts such features which are useful for the purpose of tracking and mapping objects in various computer vision tasks. It is faster than any of many other renowned feature extraction methods, such as Difference of Gaussians (DoG) used by SIFT, Harris and SUSAN and is therefore appropriate for real time video processing.

3) TENENGRAD

Tenengrad [11], [26], [27] focus measure is based on calculating the magnitude of the 2D spatial gradients using the Sobel operator. This measure convolves the image $I(x, y)$ with the

following 3 × 3 Sobel masks:

$$i_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad i_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \tag{1}$$

These convolution operations result in horizontal and vertical gradient images. Focus measure value for a single pixel is then given by the following formula (5):

$$S(x, y) = [i_x * I(x, y)]^2 + [i_y * I(x, y)]^2 \tag{2}$$

A single value for whole image is obtained by:

$$f_{tenengrad} = \sum_{x=1}^{M} \sum_{y=1}^{N} S(x, y) \tag{3}$$

where M and N represent the rows and columns of image. Measures based on mean and standard deviation of matrix $S(x, y)$ have also been reported in the literature.

4) VARIANCE

Variance of gray levels of an image is a statistical measure given by:

$$f_{variance} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} [i(x, y) - \mu]^2 \tag{4}$$

where $i(x, y)$ is the gray value and μ represents the mean of the image. Gray level local variance [28] has also been reported in the literature.

5) ENTROPY

This operator bases itself on the assumption that a focused image contains more information and lesser randomness thereby producing a minimum at the point of optimal focal position. Its formula is given by [29]:

$$f_{entropy} = -\sum_{i} p(i) \log_2 [p(i)], \quad p(i) \neq 0 \tag{5}$$

where $p(i)$ is normalized histogram count value.

III. PROPOSED FRAMEWORK

In this section, framework of proposed Vision and Control based Autofocusing System (VCAFS) is first described with the help of a block diagram. A suitable focusing operator is then selected after careful evaluation of the selected operators on a thermal image database. The later part of the section explains the effects of running average on focusing function profile. Controller configurations for lens motion are provided at the end.

A. SYSTEM DESCRIPTION

Figure 3 shows our implementation framework using a block diagram. There are four main blocks in the diagram: scene, image acquisition, image processing and lens motion control. These are explained below.

Scene is any object emitting thermal energy in the form of IR radiations. These radiations are captured by the lens

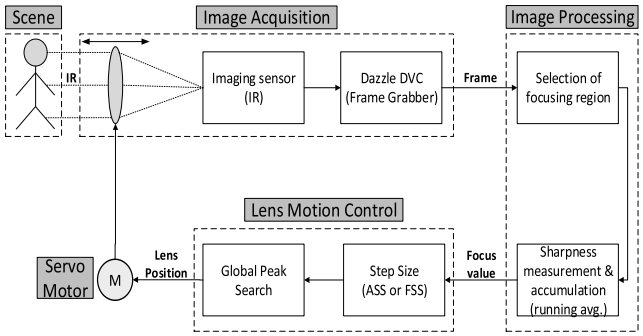

FIGURE 3. Block diagram of proposed VCAFS.

of camera and converged onto the IR imaging sensor which generates an electronic image/video depicting the thermal energy of the scene. A frame grabber is employed to feed digital frames to the personal computer (PC) where image processing and decision making regarding lens movement takes place.

In the image processing block, the first task is to select an appropriate focusing region in the captured frame which is usually central part of the image. In our case, the resolution of the camera is not very high and Tenengrad can process the frames at an appropriate speed, so there is no need for cropping. Next, we apply our focusing operator on the image to obtain its focus value which, along with the focus values of previously captured frames, contributes in calculating the running average. This is done to avoid the possible local extrema in our focusing function profile and helps in the subsequent step i.e., locating the global peak. The controller keeps on rotating the servo motor attached to the lens in either direction, based on the trend of running average, until global peak is reached eventually resulting in the best focused image. The thermal camera used for experimentation of the devised autofocusing system is an uncooled microbolometer type working in wavelength range of 8-14 μm with focal length and thermal sensitivity of 100 mm and 80 mK, respectively. Its focal plane array is 320 × 240, its frame rate is 25 fps and its pitch (i.e., pixel size) is 25 μm.

B. SELECTION OF A SUITABLE FOCUSING OPERATOR

The operators discussed above are tested using a thermal database developed by [9]. This database is developed using an uncooled thermographic camera (TESTO 880-3) with wavelength in the range of 8-14 μm, FPA resolution of 160 × 120 pixels and minimum focus distance of 100 mm. It contains 10 different image sets of scenes with varying amount of information and focus depth. Each set contains 96 images taken at different focus steps 1 mm apart. The database contains 4 sets of images of telematic equipment at different distances (TE1, TE2, TE3, TE4), a set of images of an electronic circuit (EC), a set of images of a corridor illuminated by fluorescents (CF), a set of images of hand (H), a set of images of a person's face (F), a set of images of a heater (H) and a set of images of a laptop transformer (LT).

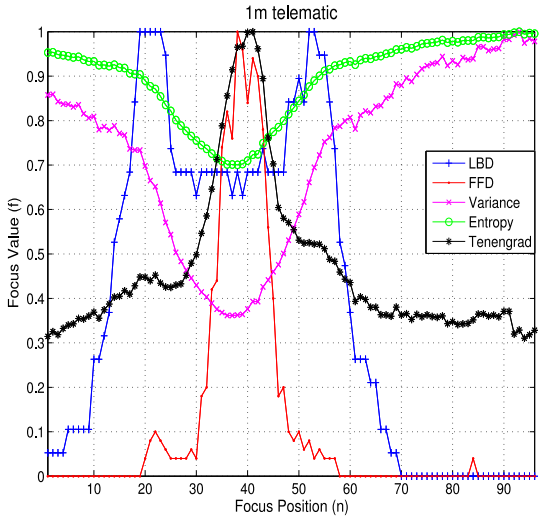

FIGURE 4. Focus functions for 1m telematic (TE1).

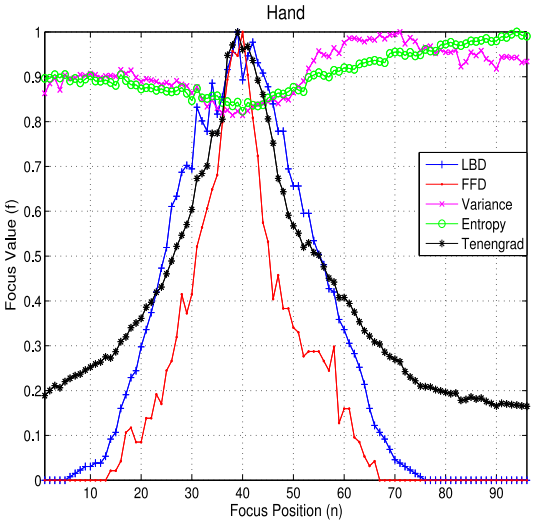

FIGURE 6. Focus functions for Hand (Ha).

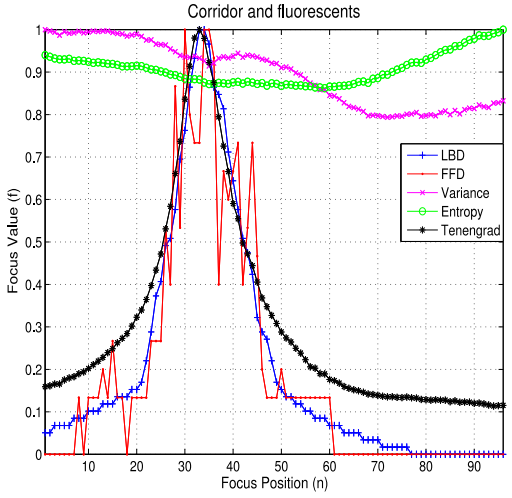

FIGURE 5. Focus functions for Corridor and Fluorescents (CF).

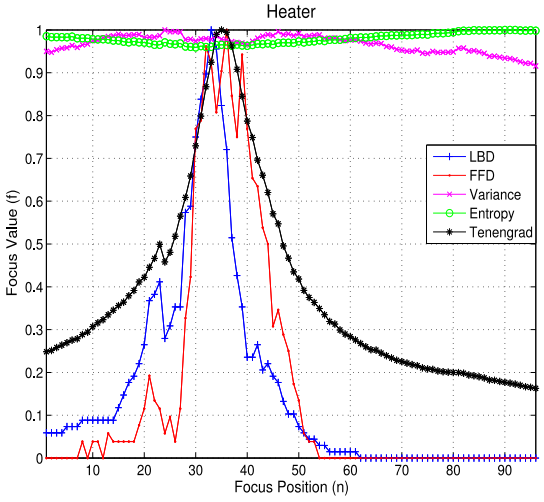

FIGURE 7. Focus functions for Heater (H).

Due to space constraints, results of all the five operators on all 10 image sets could not be presented. Therefore, results of only four databases (TE1, CF, Ha and H) are shown. For each of these four databases, these selected five operators are applied one by one, normalized by their respective maximum value and then plotted on the same figure for comparison purposes. An ideal operator should produce a Gaussian like focus function with extremum in the middle and increasing/decreasing on the sides.

Figure 4 shows the results of applying the selected five focusing operators on TE1 (telematic equipment) database. Focus functions produced by LBD and FFD are not desirable because of multimodality and narrow tail respectively. Tenengrad suffers from the problem of local extrema but still produces a favorable focus function having a sharp peak as compared to that of variance and entropy.

The results of CF (corridor and fluorescents) database, shown in Figure 5, reveal that Tenengrad produces a perfect focusing function while statistical measures fail because of scene's low contrast and large difference between the sizes of object (fluorescents) and background (corridor). FFD also fails due to the absence of definite features while LBD gives a fairly acceptable focus function. Similarly, Tenengrad gives the most desirable results when applied on Ha (hand) and H (heater) databases as shown in Figures 6 and 7 respectively. In both the databases, variance and entropy generate an unsuitable flat focus function while LBD and FFD face the issues of local extrema and narrow tail.

In terms of computational time per image calculated on a Core i3, 2.40GHz Windows computer running MATLAB R2013a, LBD suffers the most because of its multi-scale nature. FFD provides a significant improvement but still not suitable for our application. Variance and entropy perform the best regarding the computational time while Tenengrad takes slightly more time as shown in TABLE 1. Considering all the characteristics of an ideal focus function, we select Tenengrad

TABLE 1. Comparison of computational time.

OPERATOR	TIME (ms)
LBD	2400
FFD	300
Variance	25
Entropy	27
Tenengrad	30

as our operator of choice for calculating the sharpness of thermal images.

C. RUNNING AVERAGE

As evident from Figures 4-7, the profiles of the focusing functions are not always smooth due to the presence of local extrema which can prevent us from achieving the best focused position. This problem can be solved using the technique of moving average which helps to smooth out short-term variations while emphasizing the general trend of our focus function. It works by calculating the simple or weighted average of some r data points in the vicinity of a particular data point and then replacing that data point with the calculated average. Figure 8 shows the effect of applying running average on the focusing function of Electronic Circuit (EC) database [9]. It is clear that the running average profiles are devoid of small fluctuations previously present in Tenengrad's focus function profile, specifically in the tail regions. Another worth noting difference is the forward shift in peak position of running average profiles. This is because of the reason that only previous r data samples are used instead of taking equal samples on either side. The amount of this lag is equal to one half of r as confirmed by Figure 8 where peak positions of smoothed profiles for r = 4 and r = 6 are 36 (2 point forward shift) and 37 (3 point forward shift) instead of 34, respectively. Note that the running average is only applied on the focus measure values and not on the acquired images themselves.

D. LENS MOTION CONTROL

Two types of controllers are devised based upon the step size provided to the servo motor responsible for lens movement. The step size is either same (fixed) for all steps or variable (adaptive) depending upon its position from the global peak.

1) FIXED STEP SIZE (FSS)

In this configuration, servo motor moves the lens in constant steps until the peak is reached. This simple approach suffers from the drawbacks of vulnerability to local extrema and not producing the optimal result if focus position lies between two focus steps.

2) ADAPTIVE STEP SIZE (ASS)

Tenengrad produces Gaussian like focus function with small variance meaning that consecutive focus values in tail region have small difference as compared to their difference in the peak region. An ASS controller makes use of this fact by making the step size inversely proportional to the difference between two consecutive focus values as follows:

$$adaptive\ step\ size \propto \frac{1}{difference\ value} \tag{6}$$

In this way, we are able to traverse the tail region quickly thus avoiding possible local extrema. In addition, it generally produces sharper image than FSS approach.

Figure 9 shows the pseudo code of our proposed controllers. The difference between the two controllers is the step size applied at lines 8 and 16.

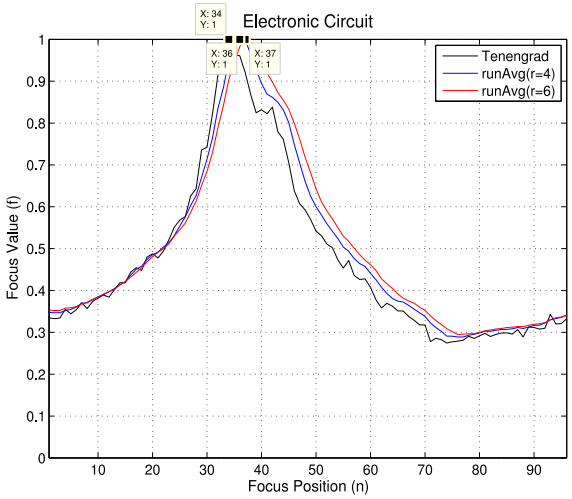

FIGURE 8. Effect of running average on EC database.

IV. RESULTS

Results of applying both the proposed controllers on live video feed of a human face are presented in this section. Comparison is done by analyzing the series of frames captured at each focus step by each controller. Running average with past four focus values (r = 4) is employed in these experiments. Images have been cropped to show only the relevant portion.

Figure 10 shows the results for FSS controller on the scene of a human face. As shown in the first two rows, there is a constant and steady increase in image quality as the lens moves with a fixed step. The third row exhibits similar trend until the sharpest image is obtained at focus step 11. The autofocusing algorithm continues moving the lens further thereby decreasing the sharpness because the lens has now moved past its optimal focus position. This movement of 4 additional focus steps is a consequence of two reasons:

i. Lag of $\frac{r}{2} = 2$ focus steps introduced due to running average
ii. Capture 2 more frames to ensure that the global peak is reached.

The controller then issues a single command to move the lens 4 steps back to its calculated optimal position,

Algorithm Passive Autofocusing Algorithm

Require: IR video/frames
1: **Procedure** Autofocusing
2: **loop**
3: rotate motor clockwise 3-4 times with certain step size
4: capture frames at each focus step
5: calculate focus value for each frame and running average
6: **if** running average is increasing **then**
7: **while** global peak not reached **do**
8: rotate motor clockwise with fixed/adaptive step
9: capture frame
10: calculate focus value and running average
11: **end while**
12: pause(5)
13: **goto** 23
14: **else**
15: **while** global peak not reached **do**
16: rotate motor counter-clockwise with fixed/adaptive step
17: capture frame
18: calculate focus value and running average
19: **end while**
20: pause(5)
21: **goto** 23
22: **end if**
23: **while** scene not changed **do**
24: pause(5)
25: **end while**
26: **end loop**
27: **end procedure**

FIGURE 9. Pseudo code of passive autofocusing algorithm.

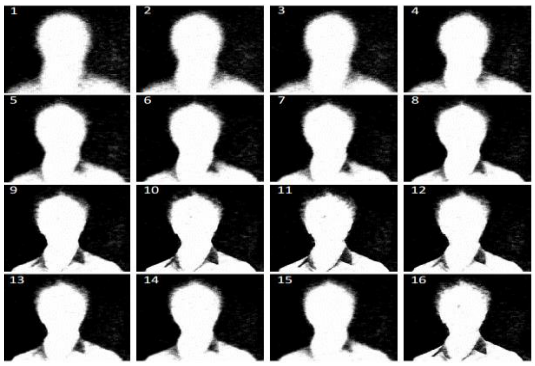

FIGURE 10. FSS controller results - Face (Image 11 is the sharpest).

eventually producing the sharpest image. This explains the sudden difference between the last two images.

Results of implementing ASS controller on the same scene as before are given in Figure 11. Note that the initial images in Figures 10 and 11 are similar implying that both the controllers are started at the same focus position. The second image in Figure 11 is taken by moving the lens by a fixed small step. Both images look similar with only a small difference d in their focus values. A significant improvement in sharpness is observed in the third image of this series, when compared to that of Figure 10 due to the reason that a large focus step is taken in adaptive controller based on its inverse relation with the previously calculated small

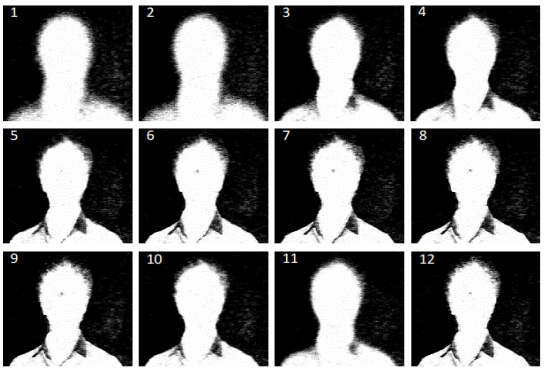

FIGURE 11. ASS controller results - Face (Image 7 is the sharpest).

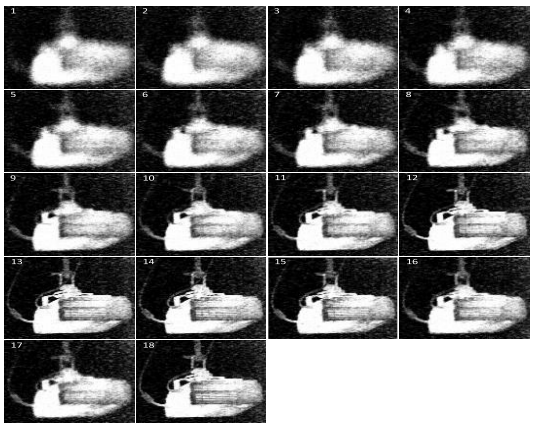

FIGURE 12. FSS controller results - Projector (Image 13 is the sharpest).

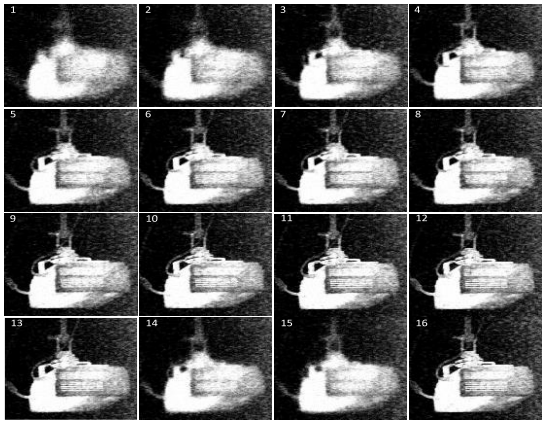

FIGURE 13. ASS controller results - Projector (Image 11 is the sharpest).

difference d. In other words, an adaptive controller is able to *skip* some of the focus steps in tail region. But at the same time, it requires more steps in the vicinity of peak region for precise refinement. This can be verified by Figure 10 where three steps (10, 11 and 12) are taken near the global peak as compared to five steps (5, 6, 7, 8 and 9) of Figure 11.

Similar results are obtained for the second scene of a ceiling mounted projector. Figure 12 shows the results of applying the FSS controller on the said scene. Note that in this case image 13 is the sharpest as the starting position of lens is

TABLE 2. Comparison of controllers' sharpness values.

SCENE	SHARPNESS VALUE (TENENGRAD)		
	FSS	ASS	% increase
Face	11827	13684	15%
Projector	20128	24040	19%

TABLE 3. Comparison of proposed system with recent study.

PARAMETER	PROPOSED SOLUTION	RECENT STUDY [8]
Thermal Images	✓	
Adaptive Controller	✓	
Tenengrad Focusing Measure	✓	
Dynamic Adaptive Focusing Window		✓

farther away from the optimal position when compared with Figure 10. When ASS controller is applied on the projector scene, results in Figure 13 are obtained.

After reaching the optimal focus position, the focus value of the next frame is calculated continuously to check if either the scene or its depth has changed. If the new focus value is within a certain bound i.e., 90% of the focus value of the last optimal position, the lens is kept at the same position; otherwise the focusing process is done for the new scene.

As mentioned before, ASS controller produces sharper optimal focus position image as compared to the FSS controller as evident by values in TABLE 2.

Once the best focus position of the lens is achieved, it is kept at the same position for some time (for example, 5 seconds) in order to avoid overheating of lens motor.

A. COMPARISON WITH A RECENT STUDY

TABLE 3 provides a comparison of our proposed solution with a similar recent autofocus study [8] done on telescopic images.

The study [8] focuses on visible spectrum images and uses the measure of Tenengrad as the focusing measure. Moreover a single controller (hill-climbing method) is employed with dynamic adaptive focusing window to help find regions of interest and reducing computation time.

V. CONCLUSION AND FUTURE WORK

A passive autofocus system for a thermal camera is developed using a sharpness-based focus measure. Five different focus operators, namely Laplacian Blob Detector (LBD), FAST Feature Detector (FFD), Tenengrad, variance and entropy are tested using a thermal image database and the gradient based operator i.e., Tenengrad's performance is found out to be the best among them. The technique of moving average is used to get rid of local extrema present in a typical focus function profile. A hardware implementation consisting of an uncooled thermographic camera and a personal computer (running MATLAB) is done to verify and compare the applicability of the two proposed controller configurations called the fixed and adaptive step size controllers. It is found out that both the controllers require approximately the same time and number of frames to reach the optimal focus position with ASS producing the sharper image than the FSS does. The developed system is tested on live video with human subject by varying both the scene and its depth. The proposed system presents a cost effective autofocus system for an uncooled thermal camera.

In current system configuration, once optimal focal position for a particular scene is achieved, scene change is detected by continuously calculating focus value every 5 seconds resulting in wastage of computational resources. Scene change detection algorithm will be incorporated in future to automatically detect when a scene or its depth has changed. The proposed algorithm will also be designed to be implemented on FPGA to reduce the focus value computation time thereby resulting in faster autofocusing.

REFERENCES

[1] F. C. A. Groen, I. T. Young, and G. Ligthart, "A comparison of different focus functions for use in autofocus algorithms," *Cytometry*, vol. 6, no. 2, pp. 81–91, Mar. 1985.

[2] M. Gamadia, N. Kehtarnavaz, and K. Roberts-Hoffman, "Low-light autofocus enhancement for digital and cell-phone camera image pipelines," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 249–257, 2007.

[3] X. Zhang, F. Fan, M. Gheisari, and G. Srivastava, "A novel auto-focus method for image processing using laser triangulation," *IEEE Access*, vol. 7, pp. 64837–64843, 2019.

[4] H. Pinkard, Z. Phillips, A. Babakhani, D. A. Fletcher, and L. Waller, "Deep learning for single-shot autofocus microscopy," *Optica*, vol. 6, no. 6, p. 794, 2019.

[5] L. Juočas, V. Raudonis, R. Maskeliūnas, R. Damaševičius, and M. Woźniak, "Multi-focusing algorithm for microscopy imagery in assembly line using low-cost camera," *Int. J. Adv. Manuf. Technol.*, vol. 102, nos. 9–12, pp. 3217–3227, Jun. 2019.

[6] M. Najibi, B. Singh, and L. Davis, "AutoFocus: Efficient multi-scale inference," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2019, pp. 9745–9755.

[7] S. Day and R. Roy, "Autofocus technologies in digital cameras: Comparative study," *Int. J. Res. Eng. Appl. Manage.*, pp. 43–46, Nov. 2018, doi: 10.18231/2454-9150.2018.0924.

[8] C. Yang, M. Chen, F. Zhou, W. Li, and Z. Peng, "Accurate and rapid autofocus methods based on image quality assessment for telescope observation," *Appl. Sci.*, vol. 10, no. 2, p. 658, Jan. 2020.

[9] M. Faundez-Zanuy, J. Mekyska, and V. Espinosa-Duró, "On the focusing of thermal images," *Pattern Recognit. Lett.*, vol. 32, no. 11, pp. 1548–1557, Aug. 2011.

[10] C.-Y. Chen, R.-C. Hwang, and Y.-J. Chen, "A passive auto-focus camera control system," *Appl. Soft Comput.*, vol. 10, no. 1, pp. 296–303, Jan. 2010.

[11] E. Krotkov, "Focusing," *Int. J. Comput. Vis.*, vol. 1, no. 3, pp. 223–237, 1988.

[12] Y. Sun, S. Duthaler, and B. J. Nelson, "Autofocusing in computer microscopy: Selecting the optimal focus algorithm," *Microsc. Res. Technique*, vol. 65, no. 3, pp. 139–149, 2004.

[13] S.-Y. Lee, J.-T. Yoo, Y. Kumar, and S.-W. Kim, "Reduced energy-ratio measure for robust autofocusing in digital camera," *IEEE Signal Process. Lett.*, vol. 16, no. 2, pp. 133–136, Feb. 2009.

[14] J. Jeon, J. Lee, and J. Paik, "Robust focus measure for unsupervised autofocusing based on optimum discrete cosine transform coefficients," *IEEE Trans. Consum. Electron.*, vol. 57, no. 1, pp. 1–5, Feb. 2011.

[15] M. G. Chun and S. G. Kong, "Focusing in thermal imagery using morphological gradient operator," *Pattern Recognit. Lett.*, vol. 38, pp. 20–25, Mar. 2014.

[16] R. Chen and P. van Beek, "Improving the accuracy and low-light performance of contrast-based autofocus using supervised machine learning," *Pattern Recognit. Lett.*, vol. 56, pp. 30–37, Apr. 2015.

[17] Y. Zhang, L. Liu, W. Gong, H. Yu, W. Wang, C. Zhao, P. Wang, and T. Ueda, "Autofocus system and evaluation methodologies: A literature review," *Sens. Mater.*, vol. 30, no. 5, pp. 1165–1174, Jan. 2018.

[18] S. A. Srivastava and N. Kandpal, "Design and implementation of a real-time autofocus algorithm for thermal imagers," in *Proc. Int. Conf. Comput. Vis. Image Process.* Singapore: Springer, 2017, pp. 377–387.
[19] A. Santos, C. O. De Solórzano, J. J. Vaquero, J. M. Peña, N. Malpica, and F. Del Pozo, "Evaluation of autofocus functions in molecular cytogenetic analysis," *J. Microsc.*, vol. 188, no. 3, pp. 264–272, Dec. 1997.
[20] C.-H. Shen and H. H. Chen, "Robust focus measure for low-contrast images," in *Int. Conf. Consum. Electron. (ICCE) Dig. Tech. Papers*, 2006, pp. 69–70.
[21] E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*. Amsterdam, The Netherlands: Elsevier, 2004.
[22] S.-Y. Lee, Y. Kumar, J.-M. Cho, S.-W. Lee, and S.-W. Kim, "Enhanced autofocus algorithm using robust focus measure and fuzzy reasoning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1237–1246, Sep. 2008.
[23] T. Lindeberg, "Scale selection properties of generalized scale-space interest point detectors," *J. Math. Imag. Vis.*, vol. 46, no. 2, pp. 177–210, Jun. 2013.
[24] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2, Oct. 2005, pp. 1508–1511.
[25] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Eur. Conf. Comput. Vis.*, vol. 1, pp. 430–443, May 2006.
[26] J. F. Schlag, A. C. Sanderson, C. P. Neumann, and F. C. Wimberly, "Implementation of automatic focusing algorithms for a computer vision system with camera control," Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-RI-TR-83-14, Aug. 1983.
[27] J. M. Tenenbaum, "Accommodation in computer vision," Ph.D. dissertation, Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, 1970.
[28] J. L. Pech-Pacheco, G. Cristóbal, J. Chamorro-Martinez, and J. Fernández-Valdivia, "Diatom autofocusing in brightfield microscopy: A comparative study," in *Proc. 15th Int. Conf. Pattern Recognit.*, vol. 3, 2000, pp. 314–317.
[29] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. Upper Saddle River, NJ, USA: Prentice-Hall, 2003, ch. 11.
[30] X. Xu, X. Zhang, H. Fu, L. Chen, H. Zhang, and X. Fu, "Robust passive autofocus system for mobile phone camera applications," *Comput. Electr. Eng.*, vol. 40, no. 4, pp. 1353–1362, May 2014.

RASHID ALI received the B.S. degree in electronics and telecommunication, in 2013, and the master's degree in signal and information processing from Northwestern Polytechnical University, Xi'an, China, in 2015. He is currently pursuing the Ph.D. degree with the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, China. He has published one book in Lambert Academic Publishing, Germany. He has several publications in international journals and international conference indexed by EI. His research interests include image processing, image denoising, image enhancement, image compression, and control systems. He was a recipient of the Chinese Government Scholarship (CSC).

PENG YUNFENG (Member, IEEE) received the Ph.D. degree from Shanghai Jiaotong University, in 2007. Since 2011, he has been a full-time Professor with the Department of Communication and Information Engineering, University of Science and Technology Beijing (USTB), Beijing, China. His research interests include image processing and network communications.

AHMAD ALI received the bachelor's degree in computer science from U.E.T., Lahore, Pakistan, in 2003, and the M.S. degree in systems engineering and the Ph.D. degree from PIEAS, Islamabad, Pakistan, in 2015. He is the author of a number of publications in image and speech processing. He is currently doing research in the field of computer vision and artificial intelligence.

HAIDER ALI received the bachelor's degree in electrical engineering from the National University of Sciences and Technology (NUST), Pakistan, in 2013, and the master's degree in systems engineering from the Pakistan Institute of Engineering and Applied Sciences, Pakistan, in 2015. Since then, he has been working with a Research and Development organization in the fields of image processing, computer vision, embedded systems, and digital design.

NAEEM AKHTER received the bachelor's degree in agricultural engineering from the University of Agriculture, Faisalabad, Pakistan, in 2001, the master's degree in information technology from the Pakistan Institute of Engineering and Applied Sciences, Islamabad, Pakistan, in 2003, and the Ph.D. degree in computer vision from the Vienna University of Technology, Vienna, Austria, in 2012. He is currently an Assistant Professor with the Pakistan Institute of Engineering and Applied Sciences. His current research interests include data-driven image processing, image synthesis, and 3D modeling.

JAVED AHMED (Member, IEEE) received the bachelor's degree in electronics engineering from the NED University of Engineering and Technology, Karachi, Pakistan, in 1994, the M.Sc. degree in systems engineering, securing 2nd position under the Fellowship Program, from the Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan, in 1997, and the Ph.D. degree in electrical (telecom) engineering from the National University of Sciences and Technology (NUST), Islamabad, in May 2008. His current research interests include image processing, computer vision, artificial intelligence, and industrial automation. His publications include 15 international conference papers and six international journal articles. He received the Certificate of Achievement from the Computer Vision Laboratory, University of Central Florida, USA, in February 2007, for conducting an eight-month outstanding joint research with them. He is listed in Who's Who in America 2011 (65th Edition), Who's Who in Asia 2012 (2nd Edition), and Who's Who in the World 2015 (32nd Edition). He has been recently nominated for the President's Medal for Technology.

ABDUL JALIL received the degrees in electronics from Quaid-i-Azam University, Islamabad, Pakistan, in 1986 and 2000, respectively, and the Ph.D. degree in image and signal processing from Mohammad Ali Jinnah University, Pakistan, in 2006. Then, he remained affiliated with the University of Sussex for Postdoctorate of nine months, from 2008 to 2009. He worked in promising chemical and electronic industries in Pakistan, from 1987 to 2002. From 2002 to 2016, he has served as an Academician with the Pakistan Institute of Engineering and Applied Sciences (PIEAS), as a Professor. He has been working as a Professor with International Islamic University, Islamabad, since February 2016. His research interests include machine vision and signal, and image processing.