Research on Auto-focusing Method Based on Pulse Coupled Neural Network

Taoyu Chen, Huawei Wang* and Jianzhong Cao
Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences,
Xi'an, Shanxi, 710119, China
*Corresponding author’s e-mail: whw@opt.ac.cn

Abstract. The key to automatic focusing method based on image processing is to construct a fast and accurate focus quality evaluation function. The method in this article introduces Pulse Coupled Neural Network processing. Using the characteristics of the network, through the enhanced image, the grayscale sum can be conveniently used to characterize clarity, thereby simplifying calculations and enhancing the real-time performance of focus control.

1. Research Background
The focus of the camera includes manual focus and automatic focus. Because manual focus has great limitations, the main research direction now focuses on automatic focus. Automatic focusing technology is widely used in military and civilian fields[1]. According to the principle, it can be divided into sensor-based automatic focusing method, self-collimating focusing method and image-based automatic focusing method. The third type is often used now[2]. After obtaining the defocus program information of the optical system, the drive motor performs focus adjustment according to the information[3][4].

In order to achieve the research purpose in the automatic focusing process, it is necessary to calculate the degree of image focus in a certain state, that is, the sharpness. The original calculation of image definition requires the use of the principle of grayscale difference. Whether it is information entropy or spatial evaluation function, it is necessary to make difference between adjacent pixels of each frame of image, and use grayscale difference information to judge, which is a huge amount of calculation. Therefore, we need to find a more efficient function. Then I get inspiration from PCNN[5].

2. Introduction and Model to Pulse Coupled Neural Network (PCNN)

2.1. Ordinary pulse-coupled neural network unit
Pulse coupled neural network is a type of spike neural network, which does not require learning or training, and can extract effective information from complex backgrounds. The structure of pulse coupled neural network is similar to ordinary neural network, with concepts such as ignition, refractory period, neuron, etc. Because PCNN is a quantitative description of the signal conduction characteristics of mammalian visual cortex neurons, its biological characteristics are consistent with the lag of human eyes perceiving brightness changes and exponential decay of visual persistence. Therefore, PCNN has rough processing of grayscale differences for parts with higher brightness, and more delicate processing of grayscale differences for darker parts[6].
2.2. Improved network structure

To make the spiking neural network closer to the real brain and eye operation principle, it is necessary to introduce the timing matrix $T$, which is based on the time domain expression of image space information, and only records the moment of the first ignition [7]. It can be used as a human eye pair for the subjective response of brightness, the input excitation $S$ is regarded as the objective brightness, and the relationship between these two pairs is logarithmic. The value of the number of ignitions recorded by the timing matrix is small, indicating that the pixels ignited at that moment are brighter, and the number of ignitions in bright areas is similar, and the processing is rough, and vice versa.

\[
F_{ij}(n) = e^{-\alpha F_{ij}(n-1)} + V_F \sum_{kl} W_{ijkl} Y_{kl}(n-1) + S_{ij} \\
L_{ij}(n) = e^{-\alpha L_{ij}(n-1)} + V_L \sum_{kl} M_{ijkl} Y_{kl}(n-1) \\
U_{ij}(n) = F_{ij}(n)(1 + \beta L_{ij}(n)) \\
E_{ij}(n) = e^{-\alpha E_{ij}(n-1)} \\
Y_{ij}(n) = \begin{cases} 1, & U_{ij}(n) > E_{ij}(n) \\ 0, & \text{other} \end{cases} \\
T_{ij}(n) = \begin{cases} n, & Y_{ij}(n) = 1, \text{Mark}_{ij}(n) = 0 \\ 0, & \text{other} \end{cases} \\
\text{Mark}_{ij}(n) = \begin{cases} \infty, & Y_{ij}(n) = 1 \\ 0, & \text{other} \end{cases} \\
E_{ij}(n) = \begin{cases} \infty, & Y_{ij}(n) = 1 \\ E_{ij}(n), & \text{other} \end{cases}
\]

Here, $\beta$ is the link coefficient of the internal activity item; $V$ and $\alpha$ are the amplification coefficient and the attenuation time constant respectively; the index $F$ represents the feedback domain, and $L$ represents the coupling connection domain; $E$ is the dynamic threshold; $M$ and $W$ are the feedback input domain and the connection matrix of the coupling connection domain.

The dynamic threshold $E$ cancels the term of raising the dynamic threshold to achieve the threshold attenuation, and uses (7) instead, because the timing matrix (6) only records the moment of the first ignition; Use Mark to assist the realization of the timing matrix $T$ to realize the output enhanced image.

In order to better highlight the edge features of the dark area, the grayscale level difference must be artificially increased, and the Laplacian filter can be used to enhance the edge.

3. Experiment and Analysis

3.1. Experimental design

Image enhancement is an image processing method that enhances the recognition and interpretation of images by enhancing useful information. When using the spiking neural network structure for image processing, it is found that it can perform image enhancement well, highlight some inconspicuous details,
and restore information that is difficult for human eyes to recognize. A simple spiking neural network can be used to construct an intensifier, which can be used to process the dynamic camera focusing process to check the effect before and after processing, especially the focus position. Note that processing pictures or videos in this way does not change the actual information, the focus position, etc., and the trend of the focus evaluation curve during focusing must also be consistent, otherwise the processing will fail, which is equivalent to adding noise to the entire process.

In the program design, there is the participation of using the Laplacian operator. The essence of making the function effective is the edge detection function of the Laplacian operator. Because the clear image has more edge information, the image after the operator has been processed, and its grayscale sum is equivalent to the sum of the edges. Only the out-of-focus image has a low function value, and the rest of the function values are difficult to distinguish. It must be strengthened by a spiking neural network to get close to real clarity. Therefore, the pulse neural network plays an auxiliary role in this program. It plays the role of feature extraction, greatly reduces the recognition complexity, and is of great help to practical engineering applications. Use the modified program to realize the system of figure 2.

In the program design, there is the participation of using the Laplacian operator. The essence of making the function effective is the edge detection function of the Laplacian operator. Because the clear image has more edge information, the image after the operator has been processed, and its grayscale sum is equivalent to the sum of the edges. Only the out-of-focus image has a low function value, and the rest of the function values are difficult to distinguish. It must be strengthened by a spiking neural network to get close to real clarity. Therefore, the pulse neural network plays an auxiliary role in this program. It plays the role of feature extraction, greatly reduces the recognition complexity, and is of great help to practical engineering applications. Use the modified program to realize the system of figure 2.

Now there are image clarity evaluation functions based on spatial domain, frequency domain, and information entropy [8]. Here I use the clarity evaluation function based on information entropy. Compared with the defocused image, the focus image has a greater degree of dispersion of grayscale value distribution and more information content, so information entropy can be used to characterize the clarity of the image. After getting the picture sequence, use the evaluation function to evaluate the degree of focus, and draw the curve as shown in the figure.

3.2. Results and Explanation

![Figure 3: Information entropy distribution of the original sequence](image)

Figure 3 Information entropy distribution of the original sequence. Here is the process of focusing using the existing autofocus program. The abscissa is the frame sequence, and the ordinate is the information entropy. On the right are a few video screenshots.
What needs to be processed is a video of the camera focusing process. The fragment is a sequence of 4 seconds and 30 frames per second, which is then decomposed into pictures frame by frame. The evaluation effect of the focus function of the image sequence strengthened by the pulse neural network has decreased. As shown in the figure above, there are three highest peaks and two secondary peaks in the original sequence of figure 3, which means that those three are the best focus points, and the positions of the two secondary peaks are very close, but they are still being adjusted.

Then use new functions to calculate clarity:

![Figure 4: The grayscale sum of the enhance sequence. The enhanced sequence trend is obvious, the highest peak and the second peak are located accurately, which can well represent the image clarity.](image)

![Figure 5: Here are Information entropy distribution of the enhance sequence and the grayscale sum of the original sequence.](image)

It can be found that the intensity of the enhanced sequence and the distribution of the information entropy of the original sequence (i.e., figures 3 and figure 4) are surprisingly consistent, and the trend can be clearly distinguished. Using this law, we can greatly optimize our calculations.

Next, verify that the pulse coupled neural network and grayscale sum work alone cannot realize full functionality.
In the left picture, the trend presented by the information entropy is correct, but the two sub-peaks become peaks of the same height, which cannot be distinguished from the true focus position. In the right picture, the grayscale sum is a trigonometric function-like trend. The maximum and minimum values are meaningless. The focus position is at the midpoint of the upper line. It is difficult to grasp, and the trend cannot be judged by the second derivative. This shows that the pulse coupled neural network and grayscale sum must be combined to achieve the function.

In addition, I used this function in different scenes (pictures about tree, tower, column and roof) and got the results shown in the figure 6. At the same time, the calculation speed is also improved (in the tree scenario, the calculation of information entropy takes 116 seconds, while the calculation of the new function only takes 7.6 seconds). The evaluation function works well in different scenes, and can get a similar function value (here above \(9 \times 10^5\)) at the focus position, which means that in the case of the same camera parameters the evaluation function can adapt to the dynamic automatic focusing process.

Figure 6: Focusing process under different scenes shot with the same camera and parameters. The videos come from the process of focusing using an existing autofocus program, and are calculated by the function I proposed.

4. Conclusion
I find a focus evaluation function, which is enhanced by pulse-coupled neural network to calculate the image grayscale sum, which can equivalently judge the focus degree of the image. I have proved this through simulation experiments. The Laplacian operator originally used for edge enhancement plays the role of edge detection here, and the neural network extracts the grayscale features, and finally can...
characterize the image clarity. So that the total calculation amount is reduced, and the overall process effect is better.

In the next research, I will optimize the structural parameters of the pulse-coupled neural network, simplify the calculation process, realize the real-time evaluation of the camera lens focus, and combine the control strategy to finally improve the camera's automatic focusing function.

Acknowledgements
This paper was completed under the guidance of teacher Wang Hua Wei so I would like to express my gratitude to him. At the same time thank the Xi'an Institute of Optics and Precision Mechanics of UCAS for its funding support.

References
[1] Lv, H.Y. (2017) Autofocus Method for Remote Sensing Cameras. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Science. Changchun.
[2] Ma, Z.L. (2017) Fast and Automatic Exposure and Focus Control Method Research for High-speed Camera. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. Changchun.
[3] Wang, H. (2015) The research of visible aviation zoom auto focus system. Changchun Institute of Optics, Fine Mechanics and Physics. Changchun.
[4] Chen, C.Y., Hwang, R.C., Chen, Y.J. (2010) A passive auto-focus camera control system. Applied soft computing,10:296-303.
[5] Lin, Z. (2020) Retina like Visual Image Reconstruction via Spiking Neural Model. In: Computer Vision and Pattern Recognition (CVPR). Seattle.pp.1438-1446.
[6] Zhang, J.Y., Lu, T. (2003) Enhancement of Image by PCNN. Computer Engineering and Applications,19:93-95,127.
[7] Ma, Y.D. (2008) Pulse coupled neural network and digital image processing. Science Press, Beijing.
[8] Tao, J.Y. (2014) Study on the Auto-focus Algorithms Based on Image Processing and its System Realization. Xi'an Institute of Optics & Precision Mechanics, Chinese Academy of Sciences. Xi'an.