

Autofocus with Automatic Block Selection

Wei-Sheng Liao (廖偉勝) and Chiou-Shann Fuh (傅楸善)

Department of Computer Science and Information Engineering

National Taiwan University, Taipei, Taiwan

d93002@csie.ntu.edu.tw and fuh@csie.ntu.edu.tw

Abstract

Our original autofocus algorithm segments an image into several areas and chooses three blocks for focus. This will save the calculation time. We still use the gradient operator and climbing search algorithm but with modification and improvement. Extensive experiments show that our method is fast and has comparable focus quality with other methods.

I. Introduction

The traditional autofocus algorithm uses gradient operator to calculate the focus value of the whole image. But, it takes too much time when the total pixel number in a digital camera is large. The total computational time for the gradient operator on the whole image is too large and thus impractical. Therefore, sub-sampling method seems a good solution to resolve the above problem. However, some detail information is lost in the sub-sampling method, so the lens does not get the best focus position. This may be a big problem.

To resolve this problem, sub-window method [1] is presented. This method uses the center of an image to detect the edges and leads to another problem. If the subject is present but off the center, the camera may not find the position which people want. In the human eye model, eyes always focus on the center of a scene. By this model, sub-window method focuses on the center of an image. However, different people have different composition preferences. Most

photographers usually prefer one-third rule and do not put the subject in the center of a scene. It is the challenge of autofocus to find where the subject position is. We present a method to automatically find the subject.

II. Approach

Our approach has two parts. One is the sharpness algorithm, and the other is search algorithm. The sharpness algorithm of our auto-focus method is a simple differential edge detecting operator. We first transform (R, G, B) values to Y value, and we calculate the value of the difference between $I_{(i,j)}$ pixel and $I_{(i,j+2)}$ pixel. Because of the interpolation, We think $I_{(i,j)}$ pixel may depend on $I_{(i,j+1)}$ pixel. That is R, G , or B value may be the same in $I_{(i,j)}$ and $I_{(i,j+1)}$. Finally, we check the difference along x and y axis, and if this difference is larger than a pre-defined threshold (10 in our method), we think this pixel is an edge. Equation (1) is our method.

$$S = \sum \sum |(I_{(i,j+2)} - I_{(i,j)}) + (I_{(i+2,j)} - I_{(i,j)})|, \quad (1)$$

The search method of our method is global search method. We know that when photographing, people zoom lens with a zoom lens or just do not zoom with a fixed lens. Focal length is fixed at the moment of photographing, no matter a zoom or fixed lens. What does auto-focus adjust? The auto-focus adjusts the distance between the lens and the sensor. It adjusts until acceptably sharpness is achieved according to pre-motioned sharpness algorithm.

The resolution of a DSC system is high

nowadays. A DSC system can not calculate the sharpness value for a full image, because this wastes time. Many DSC systems choose a small area and focus according to this area. This default area usually is the center of an image. Users may change the position of the area such as auto-focus lock. Choosing a small area from a pre-defined and larger area is important initially. Our method segments a pre-defined area into 4×4 blocks and chooses the three largest blocks in the first picture. We finally calculate the sharpness values of those three blocks similarly for the following pictures. If we have just one block, we may choose one complicated block. This block has many edges, but it may interfere with auto-focus. Auto-focus will make mistake in that case. After calculating, we can get three maximum values of three blocks. Each maximum presents a picture. If two blocks present the same picture, we choose that one as our focus picture. If we get three pictures, we choose the maximum of the summation of the three block values of pictures as our focus picture.

How do we choose a pre-defined area to segment? We think most people put object near center if we segment an image into 3×3 blocks. Hence, we choose a large area which has distance one-six of height from the top and bottom of the image and distance one-six of width from left and right side. We assume objects outside this area are background. Figure 1 is our method using 4×4 blocks. Figure 2 is the flowchart of our auto-focus method. We start with focus distance of infinity in our experiments, but we can start at any focus distance.

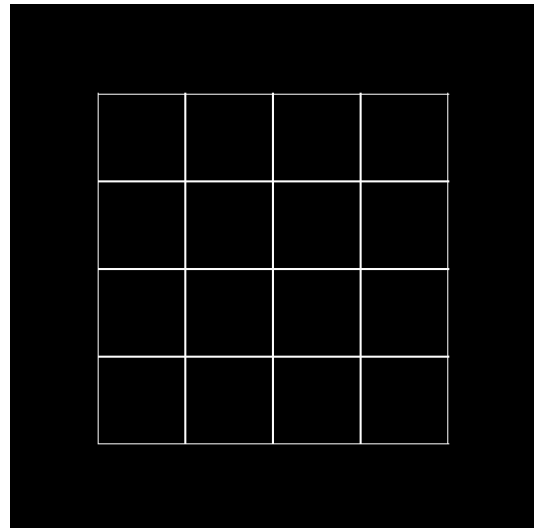


Figure 1 The blocks of our method.

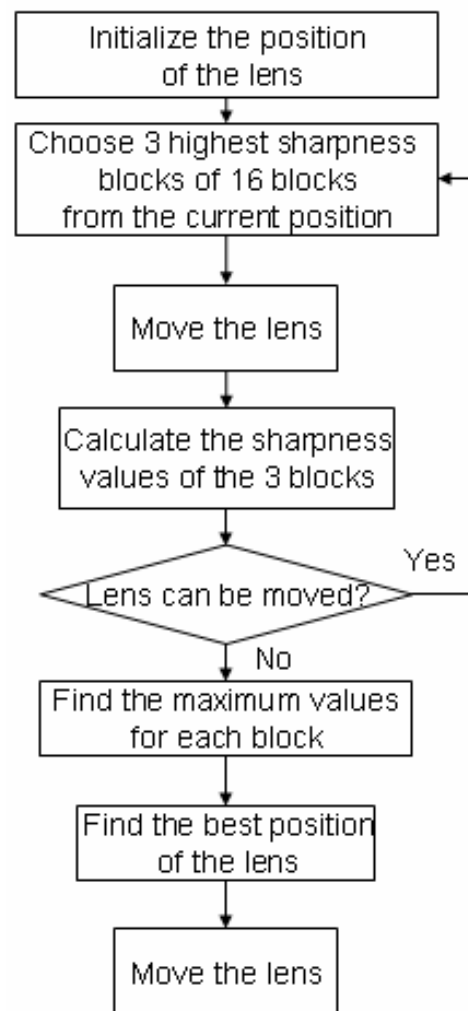


Figure 2 The flowchart of our method.

III. Experiments

We photograph a series of pictures as the positions of the lens. Choosing the focus picture automatically aims to choose the picture with the sharpest focus. We compare our method with Laplacian and YZU sharpness method [3]. YZU sharpness method is shown in Equation (2).

$$H_x = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & -6 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, H_y = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & -6 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$I_x(i, j) = \sum_{l=1}^5 I(i, j+l-3) * H_x(3, l), I_y(i, j) = \sum_{k=1}^5 I(i+k-3, j) * H_y(3, k)$$

$$S_y = \sqrt{\sum_{i=1}^m \sum_{j=1}^n (I_x^2(i, j) + I_y^2(i, j))}, \quad (2)$$

IV. Results

We photograph 50 sets of pictures. If one method focuses correctly and choose right picture, we give the method two points. If one method focuses correctly or chooses right picture, we give one point. We calculate the total points for each method. Table 1 is the result.

Method Item	Our Method	YZU	Laplacian
Score	65	65	57
Focus rate	0.58	0.66	0.54

Table 1 The result of the auto-focus.

Although the focus rate of our method is not the highest, our method is better than Laplacian. The score of our method is equal to YZU and is the highest. Although sometimes our method does not focus on subject correctly, we focus on neighboring object at the same depth and thus still make the correct choice.

V. Discussion

We test the time of the three methods with the ARM7 microprocessor. The average time of the three results is similar. That is because the time complexity of the three method is the same ($O(n)$). The main factor of taking time is the total pixel we should calculate. In our three blocks method, we just calculate the 1/12 pixels of a full image. We prove that in the following.

We assume that an image has $h*w$ pixels, w represents the width of an image, and h represents the height of an image. Each block of our method has the width $2w/12$ and the height $2h/12$. We need to calculate three blocks, and the total pixel number is P_t in Equation (3). Finally, we calculate $hw/12$ pixels totally in our three-block method.

$$P_t = \left(\frac{2*w}{12}\right) * \left(\frac{2*h}{12}\right) * 3 = \frac{w*h}{12}, \quad (3)$$

	Total pixel number	Computation time	Speedup
Exhaustive computation	480,000 (800*600)	0.028669 second	11.95
Three-block method	40,000	0.002398 second	
<hr/>			
Exhaustive computation	2,000,000	0.119333 second	12.01
Three-block method	166,666	0.009933 second	

Table 2 The computation time of different pixel numbers.

We reduce to 1/12 computation complexity. That is fast. Table 2 is the time results of different pixel numbers with the ARM7 microprocessor. We compare with the total pixel

number 480,000 and 4,000. Speedup of 11.95 ($=0.028669 / 0.002398$) times is achieved. The same comparison with 2,000,000 and 166,666 is speedup of 12.01 times. If a DSC system take a picture that has resolution of 2,000,000 pixels, a full image auto-focus may take 1.19333 seconds (if the auto-focus needs ten steps to focus). That is too long. Users can not hold a camera for so long. Adding lens motion will make it even longer. If the system uses our three-block method, the time is 0.20873 second. The speedup is 5.717 times.

We finally compare three methods with 3 blocks and with 16 blocks. We first let ten persons vote the best focused picture of a series of pictures. Second, we run three methods in 3 blocks and 16 blocks, and we get a series of focused pictures. Each focused picture has a voting score. Table 3 shows the sum of the corresponding method. The higher value represents the better result. We can see our method with 3 blocks is the best one of three methods with 3 blocks, and it is similar to three methods with 16 blocks.

Method	Our method	YZU	Laplacian
In 3 blocks	236	218	205
In 16 blocks	251	241	251

Table 3 The voting results for different methods using 3 or 16 focusing blocks.

Why do we use three blocks not other number blocks? Because we use a vote system, we let the number of blocks be odd. Because three is the smallest, positive, and odd number except one, the computation is faster.

Our method does not calculate the whole

image, and it is faster. If you want to be faster and faster, you can calculate one row and column or a small center of each block. The disadvantage of it is calculating less or no edge. That may make mistakes. That is a trade-off.

Figure 3 is a clearly discernible set of focus testing pictures. Ten persons vote the same picture Figure 3 (g). All methods choose Figure 3 (g) except Laplacian with 3 blocks (it chooses Figure 3 (h)). Figure 4 is a less discernible set of focus testing pictures. Ten persons vote different pictures, but all methods choose the same picture Figure 4 (e).



(a) Focal ring: ∞ (vote: 0).



(b) Focal ring: 10m (vote: 0).



(c) Focal ring: 7m (vote: 0).



(f) Focal ring: 1m (vote: 0).



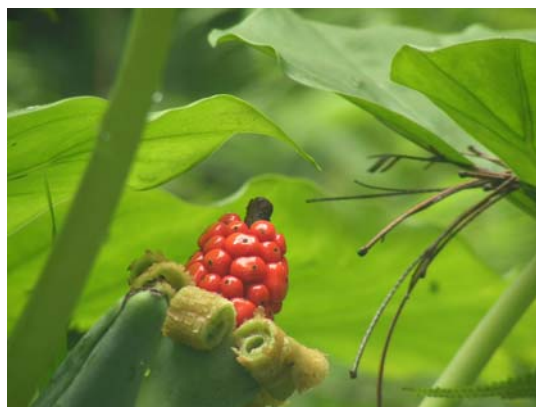
(d) Focal ring: 5m (vote: 0).



(g) Focal ring: 50cm (vote: 10).



(e) Focal ring: 3m (vote: 0).



(h) Focal ring: 30cm (vote: 0).

Figure 3 A clearly discernible set of focus testing pictures. (Shutter: 1/125 sec., F/3.5, ISO-100, focal length: 58mm)



(a) Focal ring: ∞ (vote: 0).



(d) Focal ring: 5m (vote: 3).



(b) Focal ring: 10m (vote: 0).



(e) Focal ring: 3m (vote: 2).



(c) Focal ring: 7m (vote: 2).



(f) Focal ring: 1m (vote: 1).



(g) Focal ring: 50cm (vote: 1).



(j) Focal ring: 10cm (vote: 0).



(h) Focal ring: 30cm (vote: 1).



(i) Focal ring: 20cm (vote: 0).

Figure 4 A less discernible set of focus testing pictures. (Shutter: 2 sec., F/8, ISO-50, focal length: 9mm)

Figures 5, 6, 7, 8, and 9 are examples of our method. We can see that we can find subject with our method. Then, we can focus on better position.



Figure 5 Example 1.



Figure 6 Example 2.



Figure 7 Example 3.



Figure 8 Example 4.



Figure 9 Example 5.

VI. Acknowledgement

This research was supported by the National Science Council of Taiwan, R.O.C., under Grants NSC 94-2213-E-002-032 and NSC 93-2213-E-002-073, by the EeRise Corporation, EeVision Corporation, Machvision, Tekom Technologies, IAC, ATM Electronic, Primax Electronics, Scance, Lite-on and Liteonit.

VI. Reference

- [1] K. S. Choi, J. S. Lee, S. J. Ko, "New Autofocus Technique Using the Frequency Selective Weighted Median Filter for Video Cameras," *IEEE Trans. on Consumer Electronics*, Vol. 45, No. 3, pp. 820-827, 1999.
- [2] N. K. Chern, N. P. A. Neow, and M. H. Ang, Jr., "Practical Issues in Pixel-Based Autofocusing for Machine Vision," *Proceedings of the IEEE International Conference on Robotics and Automation*, Seoul, Vol. 3, pp. 2791-2796, 2001.
- [3] C. C. Huang, "High Speed Passive Autofocus Technique," Master Thesis, Department of Institute of Mechanical Engineering, Yuan-Ze University, 2001.