A New Sampling Method of Auto Focus for Voice Coil Motor in Camera Modules

Wei Hsu and Chiou-Shann Fuh

Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan, R.O.C
{r93063, fuh}@csie.ntu.edu.tw

Abstract. We present an effective and efficient sampling table for auto focus searching algorithms to phone camera modules. Many searching algorithms base on focus value curve, and they do not consider the correlation between sampling and image characterization. Proposed table is calibrated by the camera images to meet results of optical lens, sensor response, noise, image pipeline, compressed method, and focus value function. We verify the algorithms in an image signal processor with mega-pixel level cameras. The table reduces total searching steps from 412 to 16, and images are still sharp. Our searching algorithm with the table can focus a scene within 0.6 seconds from infinity to 8 cm. For further progress, they can easily apply to mobile phones.

Keywords: Auto focusing, focus value, phone camera, voice coil motor.

1 Introduction

Auto Focus (AF), Auto Exposure (AE), and Auto White Balance (AWB) algorithms are the most important controls in image quality of digital cameras, and they are key differentiations to other cameras. Image sensors of CCD (Charged-Coupled Device) and CMOS (Complementary Metal-Oxide Semiconductor) are popular today. CCD sensors have better image quality, but prime cost and power consumption are disadvantageous. CMOS sensors with build-in processor are applied to mobile phones and webcam widely, but image quality is just acceptable. A new architecture, ISP (Image Signal Processor) with a raw sensor, has been proposed. Many people believe the architecture can bring a balance between image quality and cost. Actually, a famous company had promoted the design for a time. The sensor provider will not integrate ISP to CMOS sensors of high resolution to reduce the production cost recently.

In a typical ISP, AF finds a focus. AE controls exposure value, and AWB balances color temperatures. Color calibration maps device's colors into standard (or preferred) colors. Image pipeline processes Bayer [1] images to full-color images. JPEG [2] (Joint Photographic Experts Group) encoder compresses images into data streams for communication or storage. In the future, ISP will support H.264 to meet video applications of high compressed rate and quality.

Small dimension and high resolution of phone camera modules are noticeable features. Especially in height of sizes, thickness of mobile phone is a very alluring

condition in marketing, but the thickness is usually constrained by camera's height. The height decreases, but optical advantage becomes weakly. In beyond applications, fixed-focus cameras apply to lower resolution (under 2 mega-pixels), and their images are always blurred. An AF function can compensate the weakness, and it effectively upgrades image sharpness, especially at near distance. If high-pixel sensors bond with fixed-focus lens, the output images are still blurred. Therefore, AF is a very important feature in high resolution age.

Normally in AF algorithms, they can be divided into sharpness functions and searching loops. Sharp evaluation and searching servo are related as closely as each in reliability and speed. The computation of sharpness functions is usually greater than searching one. Sharpness functions give statistics to recognize blurred degrees of images. Simple evaluated functions are always influenced by noise. Complex functions can provide more complete information, but their computations are huge against embedded systems. The evaluation of suitable functions is necessary for the purpose.

The optical lens for phone camera has larger depth of filed (DOF), but AF still spends many times to search from infinity to macro distance. Our new sampling table by calibration can reduce searching times and preserve reliability, and proposed searching algorithm performs effectively and efficiently for mobile phones.

2 Focus Value

Focus value is sharp information of images, and auto-focus algorithms measure image quality by focus value functions. Authors present different focus value functions, and we classify functions into gray level, spatial-filter, frequency, and probability domains.

Gray level functions often base on amplitude, energy, gradient, absolute value, and variance [3], [8]. In the spatial-filter domain, high-pass or edge filters evaluate statistics. Tenegrad and Laplacian are famous filter processes [3]. Conveniently, one directional filters in horizontal are developed in real platforms. The unilateral sharp information is incomplete, but it is enough to focus. DCT (Discrete Cosine Transform), FFT (Fast Fourier Transform), and wavelet are used in frequency analyses. Human coding length can be focus value [7] in probability. In JPEG compression, frequency data are encoded by Huffman coding. In-focus images have more complex distribution in frequency lines, and stream length is larger.

Sum-Modulus-Difference (SMD), Laplacina, FFT, or coding length is the representation in each domain. Focus value curves of Fig. 1 show a sequence of JPEG from infinity to macro range by five focus value functions. Gray level amplitude often has many local maximum values, and many searching methods lose. SMD algorithm is nice in the images, because it have a sharp and clear curve. FFT fits to analyze images, and Laplacina is a robust method in many objective conclusions. Images are very complex results from captured devices, and they are consisted of the scenes, noise, sensor response, optical lens, image pipeline, compressed specification, and captured condition. After quantization, JPEG images have lost some information, and it is added some extra noise.
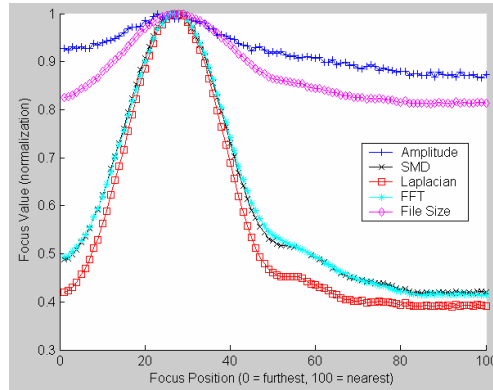
Fig. 1. Focus value curves are by five different functions. Sequential JPEG images are captured from the furthest end to the nearest end step by step. Gray level amplitude, SMD, Laplacian, FFT, and file size algorithms have different image features by the curves.

Different functions perform different image features. Many methods normalize results, and we can ignore the operation in search for computational reduction. For the fastest method, we only read file size as focus value, and it costs nothing in time and space. Frequency domain methods have the most complete information, but too complex. Algorithms in gray level domain are the fastest, but they are sensitive to slight variation of images. Methods in the spatial-filter domain are applicable in embedded systems, but a hardware engine should be implemented for real-time requirements. We can combine, modify or transfer focus value functions to improve statistical reliability [4]. Two or more focus value functions can be computed at the same time, and there is more information for search.

In different applications, we capture images by many data formats. Focus value functions usually measure sharpness by luminance and green color. We also can consider file sizes or stream lengths from compressed data. If there is a DCT hardware engine of an encoder, we can use frequency transformation for the most complete and accurate statistics conveniently. Focus value functions are difficult to say what is the best, but we can choose a suitable one to meet specific requirements under constraints of system ability and computational time.

3   Focus Sampling

AF sampling is scenic capture from different focus distances. A focus value curve records focus values at all of focus positions, and we usually evaluate AF accuracy by the curve. Window segment and adaptive step size have flexible applications to improve AF performance. A sampling table by calibration can speed up focus time, and we discuss with them as follows.

3.1 Focus Value Curve

Stepping motor moves lens by gear rotation, and VCM (Voice Coil Motor) drives lens by coil displacement to change different focusing. Input images of captured devices are discrete signals, and fixed focus positions must be repeated under mechanical and electrical controls. Generally, total steps can cover all focuses. In searching rang, searching algorithms find a focus to decrease unnecessary lens movement. Global search scans all steps, and it records focus value every step in Fig. 2.
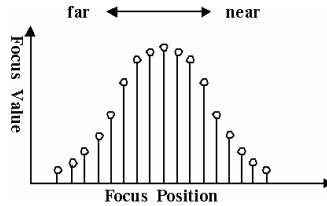
Fig. 2. Discrete signals of a focus value curve. Vertical axis is amplitude of focus value, and focuses at different position are on horizontal axis. Global search scans all steps to plot the curve, and we usually analyze AF methods by the curve.

Noise, camera shaking, and object movement often make vibration in a curve. The situation is always confusing searching algorithms. The maximum statistic has the sharpest image in the most definition of sharpness functions, and focus value of in-focus range is greater than one of out-focus range [6]. After a search, AF moves lens to the position with the maximum statistic to finish the action. Other searching algorithms can be imagined that they jump on a focus value curve by global search to find their best focus.

3.2 Window Segment

The purpose of window segment can reduce computation and improve 3A (AF, AE, and AWB) performance as Fig. 3. Statistics by a full image is an average result, and we are hard to see scenic details. We segment an image into many blocks, and compute their information. Statistics of blocks usually include focus value, color average, or color summation.

A small number of blocks lose details, and majority has huge computation. Basically, bigger image sizes and more blocks are better for precise evaluation. If there is no constraint in computation, the best preciseness is by one pixel per block. The block-average operation can reduce noisy interference, and it also reduces computation to meet the real-time requirement. In our experiments, we propose 80 to 200 blocks in 640 by 480 pixels.

3A algorithms could give different weights to blocks or select areas for precise controls. Central, spot, and matrix of AE metering are typical applications. Many AWB algorithms analyze statistical distribution in specific domains to decide samples for color correction. A central area is important in AF as Fig. 3 (a), and we can give suitable weights to correspond to human vision by location. Relatively, peripheral blocks are unconcerned even though we can ignore them. If we know the subject in
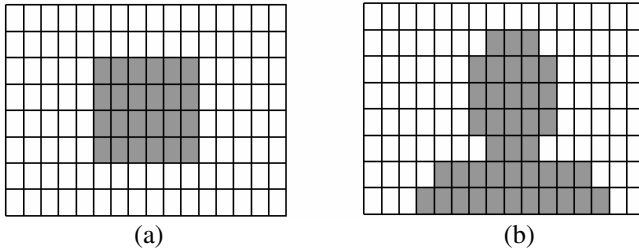
(a)   (b)

Fig. 3. Window segment. The blocks in dark color are interesting features. We can give weights on the blocks to enhance reliability. (a) A central area. (b) An adaptive area.

which blocks as Fig. 3 (b), AF can refer to the blocks for the most precise search. Such as object detection and face tracking are the applications for adaptive areas. Window segment is very flexible and effective for algorithmic development, and it is tendency towards ISP.

A focus value curve by a small number of blocks makes a steep mountain, and it easily differentiates from a peak or hillside. The disadvantage is sensitive to noise, and a local maximum often appears. Smooth curves are accumulated by more blocks, but searching algorithms decide the peak slow. It is trade-off between computational blocks and focus value reliability. Many AF methods choose different areas between out-focus and in-focus range for accurate and fast purposes, and the areas are different central sizes [4], [5], [6]. In out-focus range, a smooth curve avoid local maximum to converge to in-focus range fast. AF can accurately find a peak in a steep curve of in-focus range through area-size exchange.

3.3 Step Size

Searching algorithms change step sizes to speed up convergence. Global search uses the minimum step to sample all images in searching range. It is the slowest method, but the most accurate. Binary search compares focus value to use decrement of half steps to find the maximum value back and forth. Fibonacci search is similar to binary search, but the step sizes are arranged by Fibonacci sequence. Binary search and Fibonacci search are faster than global search, but their accuracy is doubtful.

Many searching methods consider amplitude [4], slope [5], or ratio [9] of focus value to change step size during hill-climbing. Methods of Adaptive step size can decrease focus time efficiently. Adaptive methods are very difficult to decide step sizes in different searching stages, and they do not concern with scenes, optical lens, and focus value functions. Focus value cures are very different in many cameras, and controlled parameters in the AF must be fined tuning to improve performance and reliability for a specific platform.

3.4 Sampling Table

Searching algorithms become efficiently and effectively by a sampling table. The table records focus steps, and a searching algorithm tracks the table to move lens position. Fig. 4 illustrates non-linear relation between object-distance and focus steps by a phone camera module.
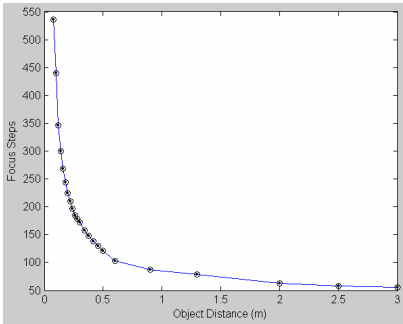
Fig. 4. An object-distance and focus steps curve. In the camera module, the optical effective steps are around 490 steps. DOF is narrow at near end, but it is wide at far distance.

Optical lens has the DOF feature, and we can utilize lesser steps to represent all steps. Lens movement should be involved in the table as a result of non-linear relation for efficient searches. A sampling table can be built by DOF or fixed distances with probability. By fixed distances, the purpose is focus at some fixed distance, but it is discontinuous in optics. By DOF solutions, sampling steps have optical difference, and they can reduce probability of local maximum. Searching algorithms with the table also can avoid noisy interference to improve accuracy. Because of the optical nonlinear relation, lens movement should not consider logic judgement only. We should track specific focuses in sampling to improve AF performance.

In adaptive step controls, a big step could jump a peak, and a small one is slow. Adaptive controls read focus value to adjust steps. Normally, high-contrast scenes have a sharp focus value curve, and low-contrast scenes have a gradual one. Searching algorithms often confuse in-focus with out-focus, because they never know the curvature. Noise also is a major cause to interfere in focus value curves. Preview mode with high frame rate needs huge image data, and it accompanies more noise. If searching algorithms can base on optical continuous features to search, it would not lose necessary sampling steps. An object at the far end with large DOF could be use small steps, and we can use big steps at the near end respectively. It is the reason to keep reliability and improve speed. Searching steps are reduced, and searching times are saved without quality loss. Images are complex. If we only consider the DOF feature, it is insufficient. We present a new sampling method to calibrate by a camera.

4   The Algorithms

Our proposed method has a sampling table and searching algorithm for VCM lens in camera modules. The table establishment is described in section 4.1, and a fast and robust searching method is stated in section 4.2.

4.1   The Sampling Table

Fig. 5 draws a focus value curve, and DOF makes clear images in a range. An acceptable threshold differentiates between blurred and clear image. Clear images have over acceptable threshold in the curve, and blurred images are opposite.
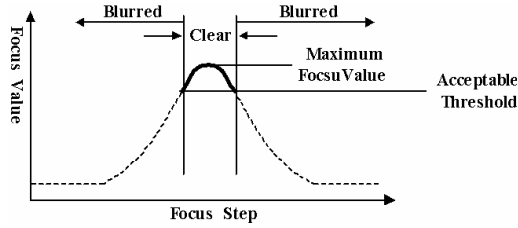
Fig. 5. A focus value curve corresponds to clear and blurred images by an acceptable threshold. Clear images are at the hilltop, and they are in a nusmbers of focus steps.

The curve has the maximum focus value (*MFV*), and the blurred ratio (*BR*) is calculated by

$$BR = Focus\ Value/MFV \qquad\qquad (1)$$

For an efficient search, we build the sampling table from a blurred end (infinity) to the other blurred end (macro). At the first, we need give an acceptable threshold ratio (*ABR*) to decide clear or blurred images. A testing chart is set at infinity distance, and global search finds the focusing step. Focusing is moved toward near direction step by step until *BR* is smaller than *ABR*. At the condition, we record the focus step into the proposed table, and set the testing chart at the focus step with a new *MFV* by global search. After the iterative processes from the furthest to nearest end, we can build the sampling table. The table by the current focus value function can cover all steps in searching range. Searching algorithms refer to the specific steps which are peaks of curves efficiently in Fig. 6, and all scenes have sharpness over *ABR* by the sampling.

ABR should be fined tuning by a target camera. Focus value curves of a fixed scene are very different by different cameras and focus value functions. Noise and artifact interfere in adjustable degrees of edge enhancement. Therefore, we should estimate the image quality to decide *ABR* to meet the customization.
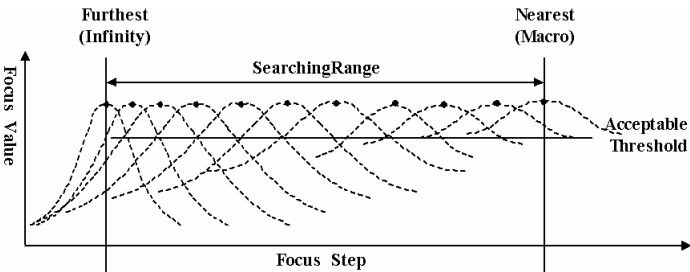

Fig. 6. Sampling steps for auto focus. A focus value curve is calibrated one by one in searching range. Every curve is different gradient after normalization, and the sampling steps are at the peaks.

4.2 Searching Algorithm

Our proposed method has an initial and searching loop stage, and we consider the central window segment for focus value evaluation. Fig. 7 draws the search flow chart.
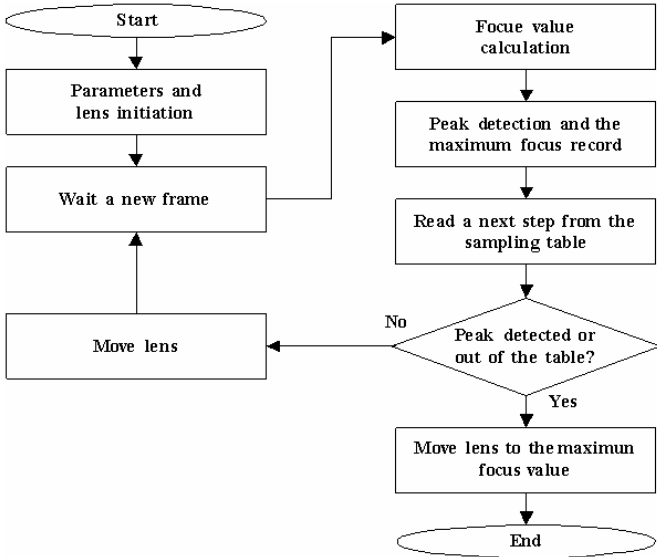

Fig. 7. Proposed auto focus searching. Initial and searching loop stages are the backbone in our auto focus servo. A new frame drives the loop progress. Peak detection decides a peak of a focus value curve to reduce searching times. Finally, lens is moved to a focusing.

In the initial stage, we move lens to focus infinity and reset controlled parameters. The *basic focus value* is recorded for reference of the current scene. The loop progress refers to the sampling table frame by frame, and focus values renew synchronously. The maximum focus value is also updated every frame.

When the table runs out, it terminates the loop. Peak detection also can break the loop, and it bases on the *basics focus value*. We assume the focus value of in-focus range is over 1.1 times than one of out-focus range [9] in high-contrast scenes. If the focus value decreases on going and accumulates to seven frames under 0.9 times condition of *basics focus value*, lens has been over the peak. If peak detection does not break the loop, all of sampling steps are be scanned. After the loop termination, lens is moved to the position with the maximum focus value, and we finish AF.

5 Experimental Results

In our experiments, we use a VCM camera module which is 10 x 10 x 7.3 mm^3, and the effective searching range is from infinity to 8 cm. The actuator driver outputs the voltage to make lens displacement for VCM. The driver has 10 bits resolution for

voltage control, and we can arrange 1024 steps. Effective steps (infinity to 8 cm) are 412 steps in searching range, and others cannot be focused. Popular CMOS sensors of 2 and 3 mega-pixels are verified by our AF algorithm. The sensors can provide 10 bits Bayer data of 640 by 480 pixels at 30 fps (frame per second) in preview mode, but we control in 28.5 fps for better image quality. Our experimental ISP has ability to process the image size in real-time. The 3A hardware engine in ISP calculates red, green, blue, and focus value statistics with 12 by 8 window segment. The computational time and lens displacement can be finished within a frame flush, and the software is developed in language C.

Optical DOF of VCM camera modules is larger, and we reduce 4 steps to 1 step for global search evaluation. Global search spend 3.61 seconds in a round by 103 steps. In our proposed table, the acceptable threshold ratio is set to 0.94 to calibrate our camera module, and we reduce 412 steps to 16 steps. Our method spends 0.56 seconds in the worst case, and the least time is 0.32 seconds. Peak detection can reduce searching steps in high-contrast scenes, and we spend less time for a search.

In accuracy, we base on the sampling table to search in searching range. Peak detection is not work in low-contrast scenes for reliability. If the object is at the near end, the peak detection also looks ineffectively. In other words, our search scans all of steps in the sampling table. The maximum focus value of global search indeed is different from ours, but the differences are within acceptable range. Therefore, images have sharpness over acceptable vision. Actually, it is fast even though we do not care the peak detection. Our AF scans all of focuses within 0.6 seconds in our platform.

6   Conclusion

Our proposed algorithm performs not only fast but also accuracy by camera modules and platform. In the sampling table, focus values of neighbor values are very different, and the sampling method can avoid local maximum situations possibly. The table has the minimum and most efficient sampling steps for other searching algorithms. For more accurate, we also can search the maximum focus value among three largest sampling steps.

Images are consisted of the scenes, sensor response, optical lens, image pipeline, and compressed method. Especially in image pipeline, edge enhancement functions can compensate sharpness, but too strong operation is not naturally. Noise often emphasizes at the same time. In the current stage, we estimate the acceptable blurred ratio by vision. PSNR (Peak Signal-to-Noise Ratio) is a good method to evaluate a difference of two images. When we change focal lengths, the scenes are different. We are hard to evaluate the blurred degree by PSNR. The combination of aperture sizes and optical zooms has different DOF, and tables must be calibrated by conditions.

In the future works, we will focus on the acceptable blurred ratio for objective evaluation. The developed software can be transferred to the family chips easily, and we will apply them to ISP chip for further progress. Zoom tracking has the similar concept, and we will transfer the idea.

References

1. Bayer, B.: Color imaging array. In: U.S. Patent No. 3,971,065. (1976)
2. Wallace K.: The JPEG still picture compression standard. IEEE Trans. on Consumer Electronics.(1991)
3. Chern N. K., Neow P. A. and Jr M. H. A.: Practical issues in pixel-based autofocusing for machine vision. Int. Conf. On Robotics and Automation. (2001) 2791- 2796
4. He J. et al: Modified fast climbing search auto-focus algorithm with adaptive step size searching technique for digital camera. IEEE. Trans. on Consumer Electronics, Vol.49, No.2. (2003) 257-262
5. Choi K. and Ko S.: New autofocusing technique using the frequency selective weighted median filter for video cameras. IEEE Trans. on Consumer Electronics, Vol. 45, No. 3. (1999) 820-827
6. Lee J. S., Jung Y. Y., Kim B. S., and Ko S. J.:An advanced video camera system with robust AF, AE, and AWB control. IEEE Trans. on Consumer Electronics, Vol. 47, No. 3. (2001) 694-699
7. Weintroub J., Aronson M. D., and Cargill E. B.: Method and apparatus for detecting optimum lens focus position. In: U.S. Patent No. 0,117,514. (2003)
8. Lee, J.H et al: Implementation of a passive automatic focusing algorithm for digital still camera. IEEE Trans. on Consumer Electronics, Vol. 41, No. 3. (1995) 449-454
9. Kehtarnavaz N. and Oh H. J.: Development and real-time implementation of a rule-based auto-focus algorithm. Journal of Real-Time Imaging, Vol. 9, Issue 3. (2003) 197-203