

Intelligent Autofocus with Adaptive Depth of Field

Pu-Chuan Kang and Huei-Yung Lin
 Department of Electrical Engineering
 Advanced Institute of Manufacturing with High-Tech Innovation
 National Chung Cheng University
 168 University Rd., Chiayi 621, Taiwan

Abstract

In this paper, we propose a technique for multi-region autofocusing. The objective is to make the objects of interest at the different distance locations well focused while maintaining the shallow depth of field. Based on the image sharpness analysis of interested regions, we determine the camera's best lens position and largest aperture size such that the depth of field encompasses all objects selected by the user. Thus, in addition to emphasize the objects of interest in photography, our method can also reduce the exposure for image stabilization. Experiments with real scene images are presented.

1 Introduction

Autofocusing is one of the most important issues on modern camera design. The related principle technology exploitation has received much attention from the camera manufacturers in the past few decades. In recent years, due to the improvement of image capture and display technologies, built-in cameras have become the essential components of smartphones, tablets and other mobile devices. The users are also more and more fastidious on the image quality. However, for most of these imaging devices, their autofocus system tends to focus on the scene or object with a certain depth only. If multiple objects or the scenes with a depth range to be focused simultaneously, the camera's aperture size needs to be reduced and the obtained image would contain almost no shallow depth of field.

Some autofocus technologies utilize frequency analysis, edge detection, or join neural system [6, 4, 2] to obtain or predict the optimal camera focus distance. The regional search is usually adopted to find the object of interest [10]. If one needs the near and far objects all-in-focus, a small aperture has to be considered. However, this might cause insufficient brightness during the image acquisition process. For the research on the all-in-focus imaging such as in [11, 1, 8], a set of low-resolution sharpness measurements of the scene is analyzed while continuously varying the focus distance of the lens. From these measurements, they estimate the final lens position required to capture all objects in the scene at an acceptable focus quality[3].

In the above methods, a large number of images is captured and used for analysis or image fusion. For high-resolution images, the computation cost will become extremely high. Since the most commonly encountered autofocus situation is for faces or human-defined objects [12, 5], it is reasonable to let the users to select the objects of interest to focus on. As commonly shown in most digital cameras, the multi-point focus detects the vertical and horizontal positions of

the interested regions. When those areas are detected, they are supplied to the camera's internal AF system for focusing process. By the assistance of touch-pad panels of modern imaging devices (as such mobile phones, tablets, etc.), the objects of interest in the scene can be manually selected by the users. Thus, it is possible to analyze the content of the selected image areas for focusing evaluation to avoid the whole image computation.

To make the objects of interest well focused while maintaining the shallow depth of field, we need to derive an optimal focus range just covering the depth range associated with the objects. In this paper, we present a technique to determine the camera's best focus position and largest aperture size such that the depth of field encompasses all of the objects selected by the user. The idea is not only to emphasize the objects of interest in photography, but also to minimize the exposure time required by image acquisition. Thus, it is able to provide better image quality, especially when dealing with the scenes under low light conditions.

2 Approach

To have the acquired image contain the depth of field effect with the objects of interest well-focused, a target selection stage is first carried out. In the implementation, Canon's official EDS SDK and the camera's built-in AF system are used to develop the target selection algorithm. According to the user-selected target locations, multiple images of the same scene are acquired with different in-focus regions. This allows the user to choose the objects of interest and have them presented sharply in the captured images. Through these respective focused images, we can derive the information required in the following focus distance adjustment and sharpness analysis steps.

2.1 Focus Distance Adjustment

In the target selection stage, we first obtain multiple images with each of them focused on a single object region. To adjust the camera's lens position to have the best focus distance for all interesting object regions, we need to know the distance between the camera and each object. One way to derive the object distance is to use active devices such as on-board infrared sensors or laser rangefinders for measurement. However, they usually only provide rough distance information and might not be consistent with the camera's focus setting. Thus, we extract the focus distance information present in the image header.

It is well known that the Exif data (exchangeable image file format) contain useful image acquisition in-

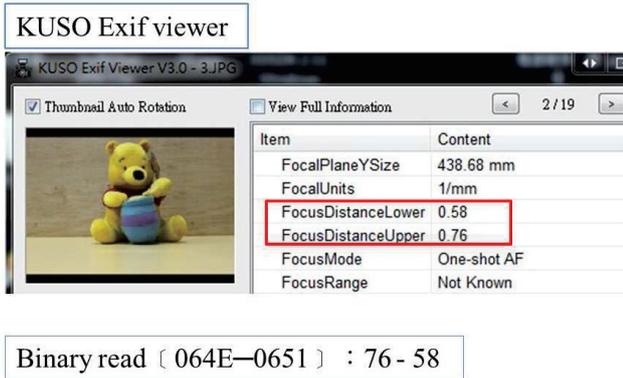


Figure 1. Depth of field tags in Exif data.

formation, including focal length, aperture, exposure time, resolution, etc. In recent years, some camera manufacturers have added two new tags which are usually termed as “FocusDistanceUpper” and “FocusDistanceLower” as shown in Fig. 1. They are used to indicate the depth of field associated with the lens focus position of the captured image. Since the scene is focused within this range, these two tags can also be used to approximate the maximum and minimum object distances. From the principle of lens imaging, we know that the depth of field will be changed due to the aperture size, focal length, and physical distance. We adopt the middle of the depth of field range as the estimated object distance, which is generally not affected too much for the inward or outward expansion due to the aperture size change. It should also be noted that, from the principle of depth of field, this range is different for a different focus distance.

For a 3-D object, the captured image must contain several small regions with different distances. Most people usually pay more attention to the details of the object regions closer to the viewer. Thus, the color information, contrast and sharpness of these regions are relatively more important than the object regions which are also viewable but spatially closer to the background. In other words, the low image focusing quality is acceptable to the human eyes for the image regions of the background scene and the object closer to the background. To take this fact into consideration, it is better to set the focus distance using the central and interior area of the target. Consequently, even if the object is not located in the range of depth of field, it can still be sharp enough as long as the difference between distances is not significant.

Although the proposed method is able to estimate the distance of the focused object based on the Exif data, the camera manufacturers do not generally release the corresponding lens focus position. To establish the relationship between the lens focus position and the depth of field range, a database for a specific camera model (Table 1) is created manually with a series of changes for the lens position and object distance. By comparing the distance information of the captured image and the database images, the difference between the target and current focus distances can be obtained. It is then used to derive the lens position for the target focusing range.

2.2 Sharpness Analysis

To make the depth of field cover all interesting object and have them well focused, an appropriate aperture size adjustment is conducted. It is based on the sharpness measurement of the selected object regions in the image. In this work, we first adopt a contrast measure based on squared Laplacian (CMSL) [14, 9] defined by

$$L(x, y) = \frac{1}{J * K} \sum_{x=1}^J \sum_{y=1}^K G(x, y)^2 \quad (1)$$

where

$$G(x, y) = \sum_{i=x-1}^{x+1} |I(x, y) - I(i, y)| + \sum_{j=y-1}^{y+1} |I(x, y) - I(x, j)|$$

and $I(x, y)$ is the intensity value. The parameters J and K are the height and width of the focusing region in the image for the contrast being evaluated.

For each image pixel, this method can be used to compute the contrast measure of the associated cross mask region. In general, the image is sharper with a high score, and the low score indicates the image is more blurred. However, the score is also affected by the image content. For example, if the selected object in the image is too bright or too dark, the measured score will be insignificant even the color and edge features are present in the region. Thus, the S_3 method [13] and Sobel edge detection are further integrated for sharpness analysis. The S_3 method utilizes the image spectral and spatial properties to measure the spectrum slope and the spatial variation. By taking the visual perception with different weighting, the generated results contain larger values to represent the sharper images.

When a series of images are taken, from blur to well-focused, more and more edge features will be detected. By combining the information of edge pixels among the images, the S_3 measures are close to a linear curve. For

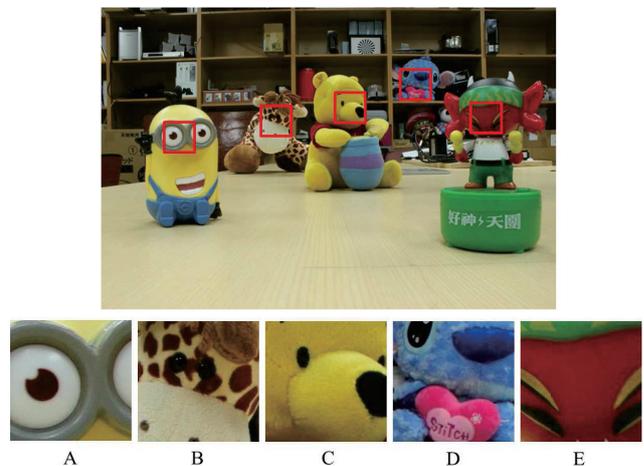


Figure 2. The image and regions of interest in the experiments.

Table 1. Lens position and focus range for Canon EOS 500D.

Number	Lower	Upper	Number	Lower	Upper
2	25	27	20	48	58
4	27	28	23	58	76
7	28	30	26	76	113
10	30	33	29	113	182
12	33	37	31	182	527
15	37	41	33	527	65535
18	41	48	35	8191	65535

different sharpness measures, the best results might appear in different image regions. If we combine various measurements directly, the results will tend to be dominated by the largest one. To comply the human vision perception, the measurements are normalized for each selected object, and the overall sharpness value is given by

$$\begin{aligned} \text{sharpness} = & \alpha \cdot \text{sharp}_{\text{contrast}} + \beta \cdot \text{sharp}_{\text{edge}} \\ & + (1 - \alpha - \beta) \cdot \text{sharp}_{S_3} \end{aligned} \quad (2)$$

where α and β are weighting parameters. The derived sharpness curve of a selected region is approximately linear under normal illumination conditions.

However, people will not simply look at a region or a few regions of the image but usually look holistic. We follow the object distance to divide the scene into the object at front, the object at middle, and the object at back. In other words, we divide the lens rotary shaft number into three regions: front, middle, and back (Table.2). Because we choose different objects, the central focus distance will be different. When focusing at front, the front objects become clear earlier than the rear ones. So that we have to choose the rear image. Similarly, focusing at back, the rear objects become clear earlier than the front ones. So we choose the front image. According to the value of our sharpness detection, we rank in ascending order. Based on the experience, we take the ranked fourth, the ranked ninth and ranked fifteen as the current best image. Then use the given weight to revise the selected number of images.

3 Experimental Results

The experiments are carried out using a Canon EOS 500D camera with a 18–55mm lens and EDSDK v2.11 for camera control. To illustrate the feasibility of our technique to capture the image of interesting objects with different distances, we choose a set of five objects in the scene as shown in Fig. 2. Using the Exif information of the image and the focus range in Table 1, the lens focus positions corresponding to the image regions A, B, C, D and E are 10, 26, 23, 31 and 4, respectively. The parameters α and β for sharpness analysis are set as 0.2 and 0.5.

In the first experiment, it is expected to have the depth of field only cover the objects A and E. The corresponding lens position is given by the average number of 7. With this lens position setting, a series of images is captured with different aperture values. The sharpness analysis is then performed using Eq. (2) and gives the best selection of aperture value F-5.0. The obtained image is shown in Fig. 3. It can be seen that,



Figure 3. The image with only objects A and E well-focused.

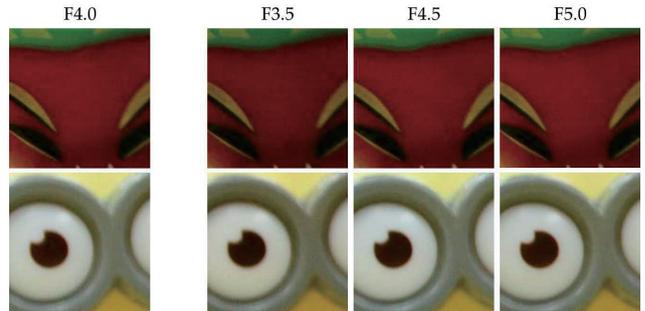


Figure 4. Regions of interest obtained with different apertures.

except for the objects A and E, the rest image regions are still blurred. Fig. 4 shows the regions of interest obtained with adjacent aperture values. F-5.0 is the largest aperture size which provides the interested image regions with sharp details.

In the second experiment, we choose C and D as the objects of interest. The corresponding lens position is given by the average number of 26. Using the same steps for sharpness analysis, the resulting image obtained with F-7.1 is shown in Fig. 5. We can see that the objects A and E are defocused while the objects B, C and D are all in the range of depth of field. These two experiments illustrate that the proposed method is able to derive the smallest depth of field to cover multiple interesting objects. If the distance range among the objects is too large, only the smallest aperture can make all objects well-focused. In this case, our technique still provides the lens focus position corresponding to the middle of the object distance range.

Table 2. System pickup table

Rotary shaft number	Front			Middle			Back		
	2	4	7	15	18	20	29	31	33
	10	12		23	26		35		
Select the ranking	No.4			No.9			No.15		
Object weight	Front	Middle	Back	Front	Middle	Back	Front	Middle	Back
	0.2	0.3	0.5	0.5	0.2	0.3	0.5	0.3	0.2

4 Conclusions

In this work, an intelligent autofocus technique with an adaptive depth of field capability is presented. It allows the user to select multiple objects of interest and make them all well-focused with the largest possible aperture size. The sharpness analysis combining various measures is used to obtain the lens focus position for each object. It is then used to derive the best focus distance for all objects. In addition to the depth of field effect to emphasize the interested regions, the image acquisition time can also be reduced. Experiments are carried out using real scene images. The results have demonstrated the feasibility of our object-selected autofocusing approach.

Acknowledgements

The support of this work in part by the National Science Council of Taiwan under Grant NSC-102-2221-E-194 -019 is gratefully acknowledged.



Figure 5. The image with only objects B, C and D well-focused.

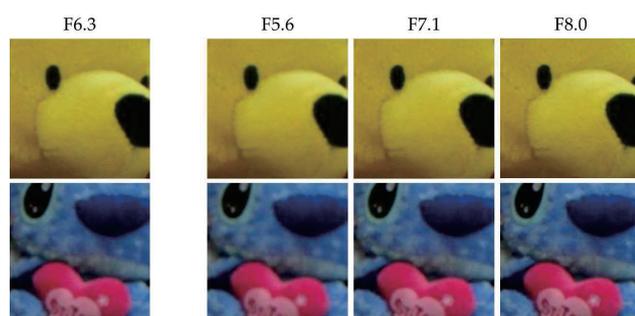


Figure 6. Regions of interest obtained with different apertures

References

- [1] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen. Interactive digital photomontage. *ACM Trans. Graph.*, 23(3):294–302, Aug. 2004.
- [2] C.-Y. Chen, R.-C. Hwang, and Y.-J. Chen. A passive auto-focus camera control system. *Applied Soft Computing*, 10(1):296 – 303, 2010.
- [3] S. Hasinoff, K. Kutulakos, F. Durand, and W. Freeman. Time-constrained photography. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 333–340, Sept 2009.
- [4] J. He, R. Zhou, and Z. Hong. Modified fast climbing search auto-focus algorithm with adaptive step size searching technique for digital camera. *Consumer Electronics, IEEE Transactions on*, 49(2):257–262, 2003.
- [5] R.-L. Hsu, M. Abdel-Mottaleb, and A. Jain. Face detection in color images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):696–706, 2002.
- [6] J. Kautsky, J. Flusser, B. Zitová, and S. Šimberová. A new wavelet-based measure of image focus. *Pattern Recognition Letters*, 23(14):1785 – 1794, 2002.
- [7] H.-Y. Lin, K.-D. Gu, and C.-H. Chang. Photo-consistent synthesis of motion blur and depth-of-field effects with a real camera model. *Image Vision Comput.*, 30(9):605–618, Sept. 2012.
- [8] T. Mertens, J. Kautz, and F. Van Reeth. Exposure fusion. In *Computer Graphics and Applications, 2007. PG '07. 15th Pacific Conference on*, pages 382–390, 2007.
- [9] L. Shih. Autofocus survey: a comparison of algorithms. In *Proc. SPIE*, volume 6502, page 65020B, 2007.
- [10] T.-H. Tsai and C.-Y. Lin. A new auto-focus method based on focal window searching and tracking approach for digital camera. In *Communications, Control and Signal Processing, 2008. 3rd International Symposium on*, pages 650–653, 2008.
- [11] D. Vaquero, N. Gelfand, M. Tico, K. Pulli, and M. Turk. Generalized autofocus. In *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pages 511–518, 2011.
- [12] P. Viola and M. J. Jones. Robust real-time face detection. *Int. J. Comput. Vision*, 57(2):137–154, May 2004.
- [13] C. Vu, T. Phan, and D. Chandler. S_3 : A spectral and spatial measure of local perceived sharpness in natural images. *Image Processing, IEEE Transactions on*, 21(3):934–945, 2012.
- [14] X. Xu, Y. Wang, J. Tang, X. Zhang, and X. Liu. Robust automatic focus algorithm for low contrast images using a new contrast measure. *Sensors*, 11(9):8281–8294, 2011.