A ROI Focusing Mechanism for Digital Cameras

Chu-Hui Lee, Meng-Feng Lin , Chun-Ming Huang, and Chun-Wei Hsu

Abstract—**With the development and application of digital technologies, the digital camera is more popular than other electronic products for all ages. The operations of digital cameras are facile and the user interface is similar in most brands. Auto-focusing is a common but important feature both for still image photography and video photography. Traditionally, focusing sensors are utilized for auto-focusing systems; in this study, we propose a new auto-focusing mechanism which uses the image sensor directly for measuring and estimating the focal distance. After the image is captured by Charge Coupled Device (CCD) or Complementary Metal-Oxide-Semiconductor (CMOS) image sensors, users will select Region of Interest (ROI) and the system can explore the in-focus image automatically. The experimental results show the proposed mechanism has superior precision and elasticity of auto-focusing.**

Index Terms—**Auto-focusing, Digital Camera, Discrete Cosine Transform (DCT), Standard Deviation (STD), Region of Interesting (ROI)**

I. INTRODUCTION

DUE to the development of digital technologies, the electronic products are extensively produced in our life, such as cell phones, digital cameras, GPSes, PDAs and notebooks [1]. The manipulation concept and technique of the traditional negative cameras are more complicated than those of the digital cameras. Besides, developing films is inconvenient and difficult. The digital camera is becoming one of the popular electronic products gradually, and furthermore, the digital image can be applied to medical pathologic slides image, image post-processing and growth record. The design of every brand is innovated and developed continually. The techniques of digital camera are developed speedily and applied to mobile devices including cell phones, PDAs and notebooks.

The region and number of the focusing sensors for digital camera are fix and limited for most brands. If the interesting object of users is not in the devised location of focusing sensor, the image of the interesting object will be not in the focus. Besides, the price of digital cameras is getting more expensive with the rapidly increasing number of focusing sensor which is a burden of user.

This paper proposes a new auto-focusing mechanism which uses the image sensor directly for measuring and estimating the focal distance. For a given scene, we will first fix the shutter speed and iris value, then capture series of pictures with different focal distance, after that, we select the Region of Interesting (ROI) from the picture and perform spectrum analysis on ROI region; finally, this paper can estimate the correct focal distance for each series of ROIs by selecting the ROI with highest energy for high frequency component.

The remainder of this paper is structured as follows. Section II discusses related works, including Discrete Cosine Transform (DCT) and Standard Deviation (STD). The details of the proposed mechanism are described in Section III. Section IV discusses the experimental results to test the proposed mechanism is presented in. Finally, Section V summarizes the conclusions of the paper.

II. RELATED WORKS

This section describes the related technologies for the proposed mechanism, including DCT and STD.

A. Discrete Cosine Transform

The DCT is one of multimedia processing methods which can transform the spatial or time domain of an image into spectrum domain [2]. It is similar to Discrete Fourier Transform (DFT), but the dissimilarity is that DCT only uses real numbers with DFT. In general, the transformation is applied to $n \times n$ arrays and the typically form is sized to 8×8 arrays in image processing [3]. The capability and complexity of calculations in DCT not only can be improved and reduced, but also obtains the similar effect. The DCT is a worthy method and is applied to image compression for JPEG and MPEG files [4]. Therefore, the DCT has been embedded in digital cameras.

After DCT, the frequency components which can represent the image are important for observing the detail in the picture. Several studies have presented that the image content is applied and analyzed variously to auto-focusing. Baina and Dublet proposed a mechanism which can evaluate the best setting of the lens and diaphragm by DCT [1]. Charifi et al. describes the curves of DCT are presented differently between the focusing images and blurring images [5]. Lee et al. proposed an auto-focusing method for cell phones based on intermediate frequency of DCT [6]. Present studies demonstrate the focus of digital cameras

and DCT have an inseparable relation. The number of the components after DCT is numerous and has to be further marshal.

B. Standard Deviation

The STD is a measurement concept which is used to observe the spread of one data set from the mean value. To understand this concept of STD, the larger STD shows that the most numbers are far apart form mean in the data set. Comparatively, the smaller STD describes the numbers are relatively close to mean in the data set [7]. It can be used to compare between the ranges of the data sets when the mean values of the data sets are equal. For instance, the following two sets are $\{1,5,8,10,26\}$ and $\{10,10,10,10,10\}$, respectively. The mean values of two sets are 10 identically, but the first has the larger STD which spreads out more clearly [7]. Accordingly, the STD can represent the spread degree of the observed data. It is usually accepted in some appraisal because it is easily used to calculate.

III. AUTO-FOCUSING MEASURE OF ROI

Figure 1 presents the proposed Auto-Focusing Measure of ROI (AFMR) based on DCT, which is used to analyze the features of ROI. For a given scene, we will first fix the shutter speed and iris value, then capture series of pictures with different focal distance. The ROI region is selected by users. Each ROI is cut into several $i \times j$ sub-regions. The important Focus Value (FV) of sub-regions is acquired through DCT. Finally, auto-focusing measurement will base on the maximum on FV. The following subsections introduce the details.

A. Pre-processing

There are totally m images. Each image has the same scene including ten chairs and with different focal distance. The proposed mechanism will pick out the best in-focus image for each ROI. Let $image$ denote the image database of this paper, shown as Eq. 1.

$$image = \{ image_k \; ; \text{for } k = 1 \, to \, m \} \qquad (1)$$

B. Opting the ROI

The user can select the interesting region of each image in this paper. The black rectangle marks the ROI in Fig. 2. Let M denote the interesting region of users in the image. The size of M is $b_1 \times b_2$.

C. Cutting the Several Sub-regions

M is segmented into $i \times j$ sub-regions. The size of each sub-region is $a \times a$, where i and j are $\dfrac{b_1}{a}$ and $\dfrac{b_2}{a}$ respectively. b_1 and b_2 are the size of M. The size a is an influence factor of performance of auto-focusing.
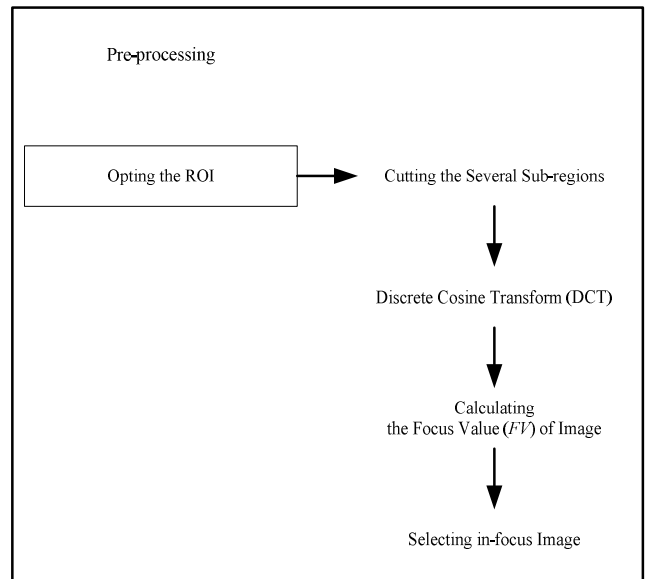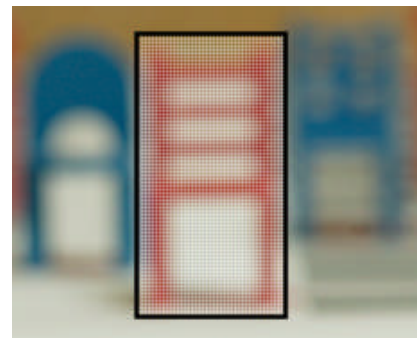

Fig. 1. The algorithms of the proposed mechanism.
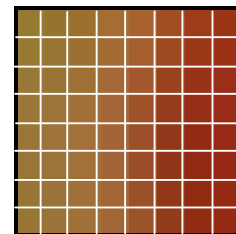

Fig. 2. The ROI of user.
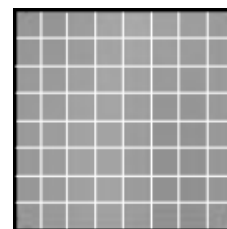

Fig. 3. An example of 8×8 sub-regions.

$$\to S_v(x,y) = \begin{bmatrix} 0.588 & 0.584 & 0.611 & 0.643 & 0.658 & 0.631 & 0.615 & 0.607 \\ 0.592 & 0.592 & 0.611 & 0.639 & 0.647 & 0.615 & 0.607 & 0.603 \\ 0.600 & 0.603 & 0.627 & 0.647 & 0.635 & 0.603 & 0.596 & 0.607 \\ 0.603 & 0.611 & 0.635 & 0.643 & 0.623 & 0.588 & 0.588 & 0.603 \\ 0.607 & 0.615 & 0.639 & 0.639 & 0.615 & 0.572 & 0.576 & 0.596 \\ 0.600 & 0.611 & 0.639 & 0.639 & 0.611 & 0.568 & 0.572 & 0.592 \\ 0.596 & 0.607 & 0.631 & 0.639 & 0.607 & 0.568 & 0.564 & 0.588 \\ 0.592 & 0.603 & 0.623 & 0.631 & 0.611 & 0.568 & 0.568 & 0.588 \end{bmatrix}$$

Fig. 4. The spatial matrix of example.

$$C(u,v)=\begin{bmatrix} 4.8701 & 0.0528 & -0.1068 & -0.0692 & 0.0554 & 0.0013 & -0.0030 & 0.0112 \\ 0.0518 & -0.0661 & -0.0201 & 0.0456 & -0.0006 & 0.0024 & -0.0007 & 0.0021 \\ -0.0080 & -0.0334 & -0.0160 & 0.0244 & 0.0015 & -0.0015 & -0.0017 & -0.0021 \\ 0.0004 & 0.0015 & -0.0034 & 0.0001 & -0.0014 & 0.0015 & 0.0012 & 0.0036 \\ 0.0025 & -0.0021 & 0.0015 & 0.0025 & -0.0005 & -0.0011 & -0.0001 & 0.0029 \\ 0.0048 & -0.0009 & -0.0032 & -0.0028 & -0.0000 & -0.0012 & 0.0035 & -0.0039 \\ 0.0049 & -0.0023 & -0.0017 & -0.0018 & -0.0001 & 0.0001 & 0.0013 & 0.0009 \\ 0.0014 & 0.0037 & -0.0004 & -0.0009 & 0.0021 & -0.0002 & 0.0018 & -0.0019 \end{bmatrix}$$
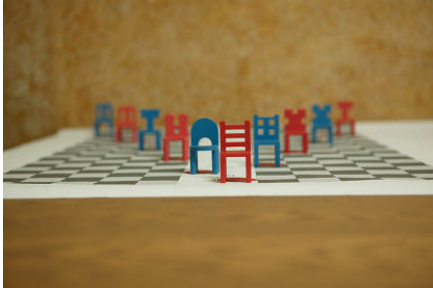
Fig. 5. The spectrum matrix of example.

Fig. 6. An image of the image database.

D. Discrete Cosine Transform

Since the DCT has a concentrate characteristic of energy. The two-dimensional DCT (2-D DCT) can transform the spatial domain of an image into spectrum domain. The 2-D DCT is presented as Eq. 2 and 3 [7]. Color is one of the most dominant and distinguished parts to human perception for CBIR. The HSV (Hue, Saturation, Value) color attribute model are used to present color features. Each sub-region is extracted the value of HSV, where $S_v(x,y)$ is an input pixel of the image. Fig. 3 displays an example of 8×8 sub-regions. The spatial matrix $S_v(x,y)$ is shown in Fig. 4. Through 2-D DCT, the spatial coefficient matrix $S_v(x,y)$ is transformed into the spectrum matrix $C(u,v)$, as shown in Fig. 5.

$$C(u,v)=\alpha(u)\alpha(v)\sum_{x=0}^{n-1}\sum_{y=0}^{n-1}S_v(x,y)\cos\left[\frac{\pi(2x+1)u}{2n}\right]\cos\left[\frac{\pi(2x+1)v}{2n}\right] \quad (2)$$

$$\alpha(u),\alpha(v)=\begin{cases}\sqrt{\dfrac{1}{n}} & \text{for } u,v \neq 0 \\ \sqrt{\dfrac{2}{n}} & \text{for } u,v = 0\end{cases} \quad \text{and} \quad \begin{cases}u=0,1,2,...,n\text{-}1 \\ v=0,1,2,...,n\text{-}1 \\ x=0,1,2,...,n\text{-}1 \\ y=0,1,2,...,n\text{-}1\end{cases} \quad (3)$$

E. Calculating the Focus Value of Image

An in-focus image contains more details than an out-focus image. This paper uses STD to observe the value of DCT coefficients. If the image has complicated content, the data of DCT coefficient is abundant. The higher value of a specific $c(u,v)$ means the image has concentrate on a specific DCT basis. This paper uses the SDT to evaluate the variation of DCT coefficient. The variation is large means detail information is more plenty. In other words, the details of region are pellucid in an in-focus image. The divergence of DCT coefficient is bigger than out-focus image. In the proposed mechanism, after each AC component is divided by DC component, the STD of all AC components is calculated. The coefficients of $C(u,v)$ can be arranged in linear, represented by $X_1, X_2, ..., X_r, ..., X_N$. X_1 is the DC value and others X_r are the AC values. The mean value $\overline{X}$ and the STD of sub-region $\sigma(i,j)$, where i and j specify the region, are defined as Eq. 4 and 5, respectively. The $FV_{i,j}$ of a sub-region $i \times j$ is defined as Eq. 5.

$$\overline{X}=\frac{1}{N-1}\sum_{r=2}^{N}\frac{X_r}{X_1} \quad (4)$$

$$FV_{i,j}=\sigma(i,j)=\sqrt{\frac{1}{N-2}\sum_{r=2}^{N}\left(\frac{X_r}{X_1}-\overline{X}\right)^2} \quad (5)$$

The spectrum matrix of all sub-regions of each ROI for each image is calculated as FV in our proposed research. The $FV_{i,j}$ of each ROI of each image is stored in the cubic matrix $T_{i,j}^k$ shown as Eq. 6, where i and j specify the region, k indicates the image. Hence $FV_{i,j}^k$ denotes the focus value for the region $i \times j$ in image k. $INO(i,j)$ defined as Eq. 7, which records the image number of largest FV in each sub-region, where the dimension is $i \times j$. The image is a best in-focus image when it is the majority in the $INO(i,j)$.

TABLE I
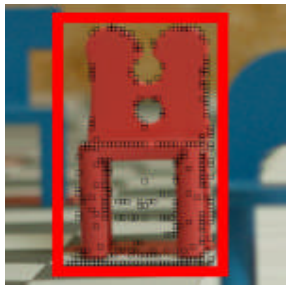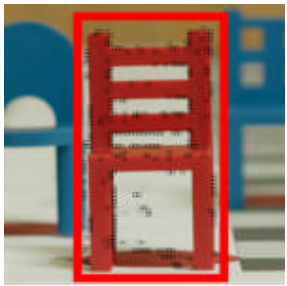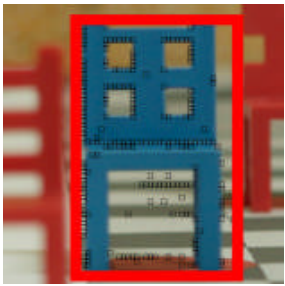### THE REALISTIC EXAMPLE OF THE PROPOSED MECHANISM

ROI-1	ROI-2	ROI-3
$I\_FV = 287$	$I\_FV = 275$	$I\_FV = 213$

TABLE II
THE EXPERIMENTAL RESULTS FOR FIVE DIFFERENT FOCUSING DEGREES OF ROI

Degree ROI	1	2	3	4	5
1	Image-2 $I\_FV = 30$	Image-12 $I\_FV = 48$	Image-22 $I\_FV = 275$	Image-32 $I\_FV = 40$	Image-42 $I\_FV = 132$
2	Image-5 $I\_FV = 28$	Image-15 $I\_FV = 56$	Image-25 $I\_FV = 170$	Image-35 $I\_FV = 23$	Image-45 $I\_FV = 36$
3	Image-11 $I\_FV = 11$	Image-21 $I\_FV = 15$	Image-31 $I\_FV = 167$	Image-41 $I\_FV = 57$	Image-51 $I\_FV = 3$
4	Image-14 $I\_FV = 11$	Image-24 $I\_FV = 13$	Image-34 $I\_FV = 105$	Image-44 $I\_FV = 23$	Image-54 $I\_FV = 14$
5	Image-19 $I\_FV = 0$	Image-29 $I\_FV = 8$	Image-39 $I\_FV = 97$	Image-49 $I\_FV = 13$	Image-57 $I\_FV = 4$

The FV of image i, denoted as I_FV, is define as the count in $\mathrm{INO}(i,j)$. If the I_FV of image i is n, then it means image i appears in $\mathrm{INO}(i,j)$ n times.

$$T_{i,j}^{k} = \left\{ FV_{i,j}^{k} \right\} \tag{6}$$

$$INO(i,j) = \left\{ \text{ image number of } (\max_{k=1}^{m}\{T_{i,j}^{k}\}) \right\} \tag{7}$$

IV. EXPERIMENTAL RESULTS

This section presents the results of the proposed mechanism. Our experiment has 57 images in the image database, which are captured by the different focal distance. The Image-1 has the shortest focal distance and the Image-57 has the longest focal distance. Each image contains ten objects with different depths. All the images are taken by Canon EOS-1Ds with dimensions 4064×2704 pixels. Fig. 6 illustrates an example image in database. Regarding experiment environment, it was as follow. The hardware was Inter Core2 Quad Q8200 CPU and 2.00GB RAM, running Microsoft Windows XP. All programs were implemented in MATLAB 2009b.

The first experiment presents a realistic example of the proposed auto-focusing mechanism in Table I with the 8×8 region size. The small black rectangles are the best focus regions. In this realistic example, the I_FV of ROI labeled 1/2/3 are 287/275/213, respectively. That means the first image has more detail information than others two. The second experiment shows the experiment results of the five different ROIs with the region size, showing the I_FV of each image in the proposed mechanism, as shown in Table II. For instance, the best focusing image of ROI-5 is Image-39 and the I_FV is 97 in the fifth row and the third column. These five ROI of chairs are arranged from near to far. The best focus image are Image-22 for the first, Image-25 for the second, Image-31 for the third, Image-34 for the fourth, and Image-39 for the fifth. The chairs of ROI are selected from near to far. Also the in-focus image number is increasing. That is comfort to the image captured condition that is image has shortest focal distance with smaller image number. Observing all the image, the selected image by our method is also the best in-focus image.

V. CONCLUSION

This paper proposed a new auto-focusing mechanism for the digital camera images. And the experimental results also demonstrated the proposed mechanism can select the best focus image. However, the size of sub-region in ROI is

an interesting topic for further research. The different scene has its proper region size. It is an important factor of process time. The future work is to explore how to reduce the time complexity of calculation base on different region size.

REFERENCE

[1] J. Baina & J. Dublet, "Automatic Focus and Iris Control for Video Cameras," *in Proc. the 5th International Conference on Image Processing and its Applications*, pp.232-235, 1995.

[2] N. Ahmed, T. Natarajan & K.R. Rao, "Discrete cosine transform," *IEEE Transactions on Computer*, Vol. 23, No. 1, pp.90-93, 1974.

[3] 94009.pdf. Available: http://www.csie.leader.edu.tw/pdf/project/.

[4] M. Tayyab, M.F. Zafar & S.S. Ali, "Performance Comparison of 2D-Discrete Cosine Transform and 2D-Discrete Wavelet Transform for Neural Network-Based Face Detection," *in Proc. the 2009 Soft Computing and Pattern Recognition*, pp.387-392, 2009.

[5] M. Charfi, A. Nyeck & A. Tosser, "Focusing Criterion," *IEE Electronics Letters*, pp.1233-1235, 1991.

[6] S.Y. Lee, S.S. Park, C.S. Kim, Y. Kumar & S.W. Kim, "Low-Power Auto Focus Algorithm Using Modified DCT for The Mobile Phones," *in Proc. the International Conference on Consumer Electronics*, pp.67-68, 2006.

[7] Standard_deviation. Available: http://en.wikipedia.org/wiki/Standard_deviation.

[8] S.A. Khayam, *The discrete cosine transform: theory and application. Technical Report ECE 802-602: Information theory and coding, Department of Electrical & Computer Engineering*, Michigan State University, Michigan, USA, 2003.