行政院國家科學委員會專題研究計畫 期末報告

一個自動平衡清晰度和視覺效果的影像融合方案

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 101-2221-E-040-010-
執 行 期 間 ：101 年 08 月 01 日至 102 年 07 月 31 日
執 行 單 位 ：中山醫學大學醫學資訊學系

計畫主持人：張炎清

計畫參與人員：大專生-兼任助理人員：闕董辰
大專生-兼任助理人員：謝政學
大專生-兼任助理人員：徐永軒
大專生-兼任助理人員：鍾月華

報 告 附 件 ：出席國際會議研究心得報告及發表論文

公 開 資 訊 ：本計畫可公開查詢

中 華 民 國 102 年 10 月 25 日

中文摘要： 多聚焦(multi-focus)影像融合(image fusion)主要是結合具有不同焦點的多個影像使其成為一個單一且銳利(sharp)的影像。其基本原理是先比較在不同的輸入影像中的每個畫素或區塊的局部內容，接著選擇它們中具有最大的資訊成分。直覺上，較大的對比(contrast)被視為影像中較清楚的畫素或區塊。很不幸地，這個觀點並不總是對的，除非一個完美的銳利度測量(sharpness measure)被採用，因此，影像清晰度的測量扮演一個非常重要的角色。對某些情況而言，例如在影像中平滑且樸素的(smooth and plain)的區塊，一個較高的測量值並非總是來自於較聚焦的區域。由於一個錯誤的選擇會導致區塊效應(blocking effects)，因此，多聚焦影像融合法通常需要額外的資訊來修正先前選擇的結果使得合成(composite)或融合的(fused)影像更加宜人。此外，適合所有種類的影像的最佳區塊大小選擇也是一項挑戰，但目前尚未存在這樣的技巧。在這篇論文中，我們提出一個具有兩個程序的自動影像融合方案來解決區塊效應的問題：一個是像素對像素處理，另一個是區塊對區塊處理。實驗結果證實結合我們所提出的方案到銳利度測量中，在僅犧牲一點點銳利度的情況下就能夠有效達到更加滿意的視覺效果。

中文關鍵詞： 多聚焦、影像融合、聚焦測量、銳利度測量。

英文摘要： Multi-focus image fusion aims to combine multiple images with different focuses to form a single, sharp image. The basic principle is to first compare the local content information of every pixel or block on distinct input images, and then choose the maximum among them. Intuitively, a larger contrast is viewed as a clearer pixel or block of an image. Unfortunately, this is not always the case, except that a perfect sharpness measure is adopted. Hence, the measures of image clarity play a very important role. For some cases, especially for smooth and plain areas on images, a higher measure does not always come from a more focused region. A wrong selection will result in blocking effects, and thus a multi-focus image fusion method generally needs extra information to modify the previously chosen results to make the composite or fused image pleasing. In addition, the selection of the optimal block size suitable for all sorts of images is also a challenge. No existing method has provided the skill yet. In this paper, we propose an automatic image fusion scheme with two procedures to solve the blocking problem in an effective way. One is for pixel-by-pixel processing and the other for block-by-block processing. Experimental results show that a fusion method combining any sharpness measure with our proposed scheme can achieve more satisfactory visual quality than one without considering the scheme, at the cost of mild sharpness.

英文關鍵詞： Multi-focus, image fusion, focus measure, sharpness measure.

AN AUTOMATIC IMAGE FUSION SCHEME FOR BALANCING CLARITY AND VISUAL EFFECTS

ABSTRACT

Multi-focus image fusion aims to combine multiple images with different focuses to form a single, sharp image. The basic principle is to first compare the local content information of every pixel or block on distinct input images, and then choose the maximum among them. Intuitively, a larger contrast is viewed as a clearer pixel or block of an image. Unfortunately, this is not always the case, except that a perfect sharpness measure is adopted. Hence, the measures of image clarity play a very important role. For some cases, especially for smooth and plain areas on images, a higher measure does not always come from a more focused region. A wrong selection will result in blocking effects, and thus a multi-focus image fusion method generally needs extra information to modify the previously chosen results to make the composite or fused image pleasing. In addition, the selection of the optimal block size suitable for all sorts of images is also a challenge. No existing method has provided the skill yet. In this paper, we propose an automatic image fusion scheme with two procedures to solve the blocking problem in an effective way. One is for pixel-by-pixel processing and the other for block-by-block processing. Experimental results show that a fusion method combining any sharpness measure with our proposed scheme can achieve more satisfactory visual quality than one without considering the scheme, at the cost of mild sharpness.

Keywords: Multi-focus, image fusion, focus measure, sharpness measure.

I. INTRODUCTION

Image fusion is usually used in many fields, including remote sensing, computer vision, defense systems, medical imaging, and microscopic imaging. Its aim is to produce a detailed description of a scene than any of source images by integrating complementary information. Among image fusion methods, multi-focus image fusion is an important, interesting, and inviting research topic.

A lens can precisely focus on objects within a limited distance at a time, whereas the sharpness gradually decreases as other objects are away from both sides of the focused distance because of a limited depth of field (DOF), which is the distance between the nearest and farthest objects in a scene that appear acceptably sharp in an image. The limited depth of field usually makes cameras impossible to acquire an image containing all objects of interest in focus. In order to capture a pleasing image, one can focus on different objects using distinct camera settings, and finally fuse them into a single, clear image using a multi-focus image fusion technique [4].

Since the out-of-focus regions are often more blurry than the in-focus regions, an intuitive idea of constructing a fused image is to choose the clearer image pixels or blocks from source images. Therefore, how to evaluate the local content information of the input images is an enormous challenge. In order to effectively differentiate between the out-of-focus and in-focus regions, various measures have been developed, including spatial frequency [3], energy of gradient [7], phase coherence [14], bilateral gradient-based sharpness [12].

On the other hand, other families of fusing techniques includes multi-resolution analysis [10], frequency selective weighted median filter [1], pulse coupled neural networks [13], multiscale directional bilateral filter [5], and image matting [8]. The multi-resolution analysis applies a forward multi-resolution decomposition to each input image, integrate all the decompositions to form a composite representation according to certain measures, and finally reconstruct the fused image via an inverse multi-resolution transform. One of the drawbacks of this type of methods is time-consuming.

An important step in multi-focus image fusion is to choose the sharper or more informative pixels or blocks by comparing measures of source images. For example, the gradient information of images is considered to be an effective and informative measure. Unfortunately, the gradient strength of the in-focus regions is not always larger than that of the out-of-focus regions. In order to raise the correction rate of determining in-focus pixels or blocks, an extra consistency-verification procedure is often indispensable, which generally needs further information on the source images to make correct decisions. Because of the continuity of imaging, users can remove isolated pixels or blocks by exploiting this useful information to manually adjust parameters. Without this useful information, users can also use mathematical morphology to remove these annoying regions by trial and error, or manually delete them by perceptual inspection.

In this paper, we propose a reliable image fusion scheme which can automatically perform the fusion procedure to obtain a pleasing fused image without extra consistency-verification procedure. The scheme can be implemented by two processing approaches: pixel-by-pixel and block-by-block processing, and applied to any sharpness measure, such as the energy of gradient (EOG), the energy of Laplacian (EOL).

The rest of this paper is organized as follows. Section II formulates the problem of multi-focus image fusion. Section III briefly introduces six sharpness measures. An automatic image fusion scheme is presented in Section IV. Experimental results are discussed in Section V. Finally, the conclusions are drawn in Section VI.

II. PROBLEM FORMULATION

A set of images were acquired by taking different imaging settings and aligned well. The aim of multi-focus image fusion is to integrate the most information or sharpest content among source images into a single fused image. The simplest multi-focus image fusion method is to take the average of all source images, which is expressed as

$$F(x, y) = \frac{1}{n}\sum_{i=1}^{n} f_i(x, y). \qquad (1)$$

Because more detailed and informative regions have the same weightings as the blurry, it is difficult to obtain a good fused image. A more reasonable approach should adopt a weighted average of all images according to the importance of each pixel or block, which can be expressed as

$$F(x, y) = \frac{\sum_{i=1}^{n} w_i(x, y) f_i(x, y)}{\sum_{i=1}^{n} w_i(x, y)}, \qquad (2)$$

where $w_i(x, y)$ is the weighting assigned to pixel (x, y) on the ith image.

The selection of weightings plays a very important role in the performance of the fused image. The weightings should be able to reflect the important information content of each pixel in the image. One feasible approach is to give a larger weighting to the pixel with a sharper neighborhood. A common way of choosing weightings is to take the weighting with the maximum measure as one, the other weightings as zeros, which easily leads to blocking effects.

In order to overcome these annoying phenomena, an automatic and effective selection method will be adopted in this paper. According to the local characteristics of each pixel with different sizes of neighborhoods, one can choose the occurrence rate of the maximum sharpness at each pixel as the weighting. Then the weighting relation between two adjacent pixels will be changed gradually, not abruptly. Therefore, blocking effects will be greatly improved.

III. SHARPNESS MEASURES

The measurement of image clarity mainly depends on the sharpness measure of an image, also called the “focus measure” [6]. A good measure should be consistent with image clarity, independent of image content—the clearer an image is, the larger the measure. Since focused images usually have sharper edges and more high-frequency content, sharpness measures are often measured using the magnitude of gradients.

In this section, a number of sharpness functions are reviewed. Let f be an image of gray level, and a block of size $M \times N$ is considered.

1. Variance [6]

$$S_{VAR} = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left(f(x, y) - \mu \right)^2 , \tag{3}$$

where

$$\mu = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y). \tag{4}$$

2. Energy of Image Gradient (EOG) [6]

$$S_{EOG} = \sum_{x=0}^{M-2} \sum_{y=0}^{N-1} f_x^2 + \sum_{x=0}^{M-1} \sum_{y=0}^{N-2} f_y^2 , \tag{5}$$

where

$$f_x = f(x+1, y) - f(x, y), \tag{6}$$

for $x = 0, 1, \ldots, M-2$ and $y = 0, 1, \ldots, N-1$ and

$$f_y = f(x, y+1) - f(x, y), \tag{7}$$

for $x = 0, 1, \ldots, M-1$ and $y = 0, 1, \ldots, N-2$.

3. Tenengrad (TEN) [2, 6]

$$S_{TEN} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \left|\nabla f(x, y)\right|^2 \quad \text{for} \quad \nabla f(x, y) > T, \tag{8}$$

where T is a threshold and $\left|\nabla f(x, y)\right| = \sqrt{f_x^2 + f_y^2}$ is the magnitude of the image gradient performed by the Sobel operators

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \tag{9}$$

along the *x*-axis and *y*-axis, respectively. They can be expressed as

$$\begin{aligned} f_x = & -f(x-1, y-1) - 2f(x-1, y) - f(x-1, y+1) \\ & + f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1) \end{aligned} \tag{10}$$

and

$$\begin{aligned} f_y = & -f(x-1, y-1) - 2f(x, y-1) - f(x+1, y-1) \\ & + f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1). \end{aligned} \tag{11}$$

Some other high-pass filters can be also used to replace the function of the Sobel operator.

4. Energy of Laplacian (EOL) [6]

$$S_{EOL} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \left[\nabla^2 f(x, y)\right]^2, \tag{12}$$

where $\nabla^2 f(x, y) = f_{xx}(x, y) + f_{yy}(x, y)$ denotes the image gradient performed by the Laplacian operator

$$\begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix}. \tag{13}$$

The detailed operations are expressed as

$$\begin{aligned} \nabla^2 f(x, y) = & -f(x-1, y-1) - 4f(x-1, y) - f(x-1, y+1) \\ & -4f(x, y-1) + 20f(x, y) - 4f(x, y+1) \\ & -f(x+1, y-1) - 4f(x+1, y) - f(x+1, y+1) \end{aligned} \tag{14}$$

5. Sum-Modified-Laplacian (SML) [6, 9]

$$S_{SML} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} S(x, y), \tag{15}$$

where

$$S(x, y) = \sum_{i=x-w}^{x+w}\ \sum_{j=y-w}^{y+w} \nabla_{ML}^2 f(i, j) \quad \text{for} \quad \nabla_{ML}^2 f(i, j) \geq T, \tag{16}$$

and

$$\begin{aligned} \nabla_{ML}^2 f(i, j) = & \left|2f(i, j) - f(i-\text{step}, j) - f(i+\text{step}, j)\right| \\ & + \left|2f(i, j) - f(i, j-\text{step}) - f(i, j+\text{step})\right|, \end{aligned} \tag{17}$$

where T is a threshold, the window size of computing the sharpness measure is $W = 2w+1$, and the step denotes a variable space between the pixels used to compute the derivatives.

6. Spatial Frequency (SF) [3]

$$S_{SF} = \sqrt{(RF)^2 + (CF)^2}, \tag{18}$$

where RF and CF are the row and column frequencies:

$$RF = \sqrt{\frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=1}^{N-1} \left[f(x, y) - f(x, y-1) \right]^2} \tag{19}$$

and

$$CF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^{M-1} \sum_{y=0}^{N-1} \left[f(x, y) - f(x-1, y) \right]^2} . \tag{20}$$

Since $f_y = f(x, y) - f(x, y-1)$ for $x = 0, 1, \ldots, M-1$ and $y = 1, 2, \ldots, N-1$ and $f_x = f(x, y) - f(x-1, y)$ for $x = 1, 2, \ldots, M-1$ and $y = 0, 1, \ldots, N-1$. Thus,

$$RF = \sqrt{\frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=1}^{N-1} f_y^2} \tag{21}$$

and

$$CF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^{M-1} \sum_{y=0}^{N-1} f_x^2} . \tag{22}$$

Thus,

$$S_{SF} = \sqrt{(RF)^2 + (CF)^2} = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=1}^{N-1} f_y^2 + \frac{1}{M \times N} \sum_{x=1}^{M-1} \sum_{y=0}^{N-1} f_x^2 . \tag{23}$$

In fact, when we use the following definitions

$$f_x = f(x+1, y) - f(x, y) \quad \text{and} \quad f_y = f(x, y+1) - f(x, y) \tag{24}$$

to replace the original ones

$$f_x = f(x, y) - f(x-1, y) \quad \text{and} \quad f_y = f(x, y) - f(x, y-1), \tag{25}$$

we can obtain

$$S_{SF} = \frac{1}{M \times N} \sum_{x=0}^{M-2} \sum_{y=0}^{N-1} f_x^2 + \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-2} f_y^2 = \frac{1}{M \times N} S_{EOG} . \tag{26}$$

The spatial frequency is a scaled version of the energy of image gradient (EOG) compared to (5). Since it was usually used as a sharpness measure, it is introduced here for a comparison with other sharpness measures.

IV. AN AUTOMATIC IMAGE FUSION SCHEME

Fig. 1 shows a schematic diagram for the multi-focus image fusion in a block-by-block processing way. For simplicity, we only consider fusing two source images into a single image. The technique can be easily extended to more than two source images.

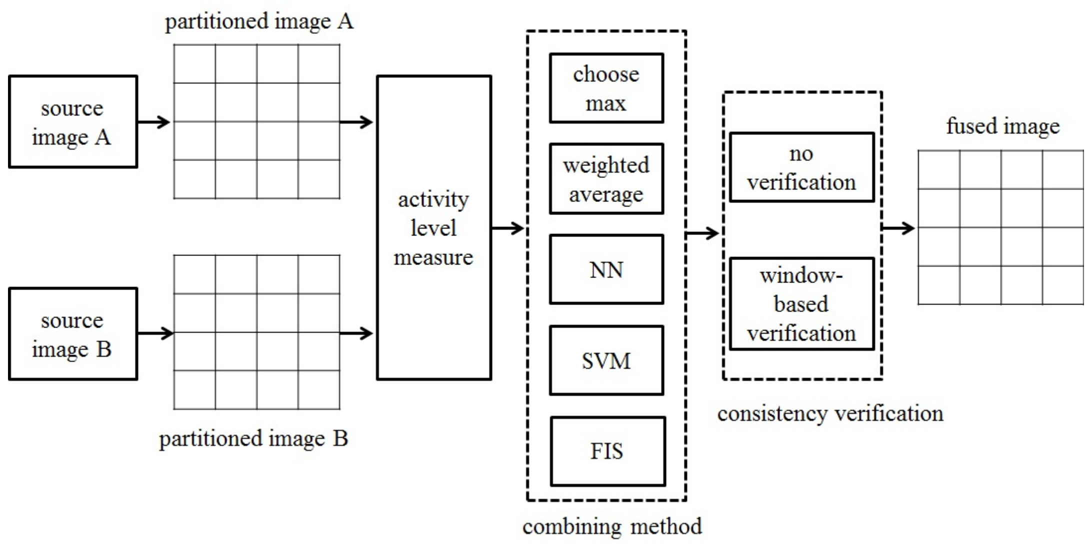

Fig. 1. Schematic diagram for multi-focus image fusion.

The standard procedure for multi-focus image fusion:

1. Divide two source images *A* and *B* into rectangular blocks of same size $M \times N$, and denote the *i*th blocks of *A* and *B* by A_i and B_i, respectively.
2. Compute the sharpness measure of each block on A_i and B_i, and the corresponding values are denoted as S_{A_i} and S_{B_i}, respectively.

3. Construct an intermediate fused image according to the sharpness values S_{A_i} and S_{B_i}, with the block C_i

$$C_i = \begin{cases} A_i, & S_{A_i} \geq S_{B_i} \\ B_i, & \text{otherwise} \end{cases}. \tag{27}$$

4. Repeat the Steps 2 and 3 for all divided blocks of the source images.

5. Use, if necessary, a majority filter to correct the intermediate fused image. The corrected composite image is referred to as the fused image, F.

The majority filter aims to eliminate isolated blocks so that the appearances of the final fused image are consistent with visual effects. The operation of the majority filter is as follows: If the central block comes from image A, but the majority of the block and its surrounding comes from image B, then this block will be replaced with the corresponding block from image B and vice versa. Generally, the size of the majority filter is chosen as 3×3.

According to the principle of imaging, the in-focus or out-of-focus regions should be continuous, and thus if isolated blocks happen they are not reasonable. In addition, perfect discrimination between in-focus and out-of-focus regions cannot be obtained by the existing image fusion methods. In order to remove isolated blocks, users appeal to a majority filter even though it usually causes side effects. For example, the sharper corners may be taken away because of the innate characteristic of the majority filter. Therefore, using a majority filter is not a kind of absolute positive operation. It needs extra information to help improve the processing, such as the shapes of focused objects.

Although some flaws occur in the fused image, the whole performance of the fused image is usually sharper than every one of the source images, and its visual effects are also pleasing as long as an appropriate block size is chosen. Unfortunately, for some uniform and plain images, a fused image constructed by block-by-block processing might result in unnatural appearances, especially as the block size is larger. Even a majority filter is exploited to remove isolated blocks, the visual quality of the fused image is still not guaranteed except for further information being provided.

On the other hand, some sharpness criteria (e.g., Tenengrad or SML) must adjust the parameters in advance to compute sharpness measures, and the determination of parameters is generally case by case. Consequently, the decision mechanism of this kind of image fusion lacks automation and generality. In order to provide a universally applicable fusion method, we propose an automatic image fusion scheme implemented by two ways: one is for block-by-block processing and the other for pixel-by-pixel processing.

The main idea of our proposed scheme is how to choose a meaningful weighting for each pixel in an automatic way, and the scheme is appropriate for all sorts of images, regardless of the content of an image. An effective approach is to compute the occurrence rate of the maximum sharpness for each pixel which belongs to a few of different block sizes. For example, we consider m different block sizes and thus each pixel will have m sharpness measurements for each source image. The detailed procedures for block-by-block and pixel-by-pixel processing are shown below:

A. An automatic image fusion scheme: block-by-block processing

1. Choose m different window or block sizes, $W_i \times W_i$, $i = 1, 2, \ldots, m$, where W_i is any number larger than 2, and set the initial i as 1.
2. Set the indicator matrix $M_i = \mathbf{0}$ for each block size, with the size of the matrix being the same as the size of any source image, and set the window size as $W = W_i$.
3. Divide two source images A and B into rectangular blocks of equal size $W \times W$, and denote the corresponding blocks of A and B by A_i and B_i, respectively.
4. Compute the sharpness measure of each block on A_i and B_i, and denote the corresponding values as S_{A_i} and S_{B_i}, respectively.
5. Compare the sharpness values S_{A_i} and S_{B_i}, choose the sharper block, and set the indicator matrix M_i as 1 in the corresponding block, denoted by C_b, by the following formula:

$$M_i\left(C_b\right) = \begin{cases} \mathbf{1}, & S_{A_i} \geq S_{B_i} \\ \mathbf{0}, & \text{otherwise} \end{cases}. \tag{28}$$

6. Repeat Steps 4 to 5 for all divided blocks of the source images.

7. Perform $i = i + 1$; if $i <= m$, then go to Step 2, otherwise go to the next step.

8. Establish the fused image by the following two formulas:

$$R(x, y) = \frac{1}{m} \sum_{i=1}^{m} M_i(x, y) \tag{29}$$

and

$$F(x, y) = R(x, y)A(x, y) + (1 - R(x, y))B(x, y), \tag{30}$$

for all pixels or coordinates.

Compared to the average method constructed by (1), our proposed method can be considered to be the method constructed by (2), where the weighting matrices are $w_1(x, y) = R(x, y)$ and $w_2(x, y) = 1 - R(x, y)$, respectively. Thus, it is simply called the "ratio method" because the weighting of each pixel depends on the occurrence rate of the maximum sharpness from different block sizes.

In the above procedure, the only two parameters chosen by users are the number of block sizes and the size of each block. Setting these parameters is an easy work. In this paper, we choose the number as 12, and the sizes range from 4 to 26 with increment 2. Block-by-block processing will be proven to be an effective approach by experiment. More importantly, it is more efficient than pixel-by-pixel processing. For completeness, the pixel-by-pixel processing is also provided here.

B. An automatic image fusion scheme: pixel-by-pixel processing

1. Choose m windows or blocks of size $W_i \times W_i$, where W_i is an odd number, and set the initial i as 1.
2. Set the counter matrix $M = \mathbf{0}$, with the size of the matrix being the same as the size of any source image, and set the window size as $W = W_i$.
3. Extract the blocks with the center being at each pixel (x, y) from two source images A and B, and denote the corresponding blocks of A and B by $A_{(x, y)}$ and $B_{(x, y)}$, respectively.
4. Compute the sharpness measure of each block on $A_{(x, y)}$ and $B_{(x, y)}$, and denote the corresponding values as $S_{A_{(x, y)}}$ and $S_{B_{(x, y)}}$, respectively.
5. Compare the sharpness values $S_{A_{(x, y)}}$ and $S_{B_{(x, y)}}$, choose the pixel (x, y) with the sharper block, and accumulate the counter matrix by the following formula:

$$M(x, y) = M(x, y) + 1 \text{ for } S_{A_{(x, y)}} \geq S_{B_{(x, y)}}. \tag{31}$$

6. Repeat Steps 4 to 5 for all pixels of the source images.
7. Perform $i = i + 1$; if $i <= m$, then go to Step 2, otherwise go to the next step.
8. Establish the fused image by the following two formulas:

$$R(x, y) = M(x, y)/m \tag{32}$$

and

$$F(x, y) = R(x, y)A(x, y) + (1 - R(x, y))B(x, y), \tag{33}$$

for all pixels or coordinates.

V. EXPERIMENTAL RESULTS

In this experiment, we adopt five types of sharpness measures to compare their performance, including variance (V), energy of image gradient (EOG), Tenengrad (TEN), energy of Laplacian of the image (EOL), sum-modified-Laplacian (SML). For efficiency, we apply the procedure of block-by-block processing to each measure, and consider two types of implementation: one is for implementation with a majority filter and the other for implementation without a majority filter.

Figure 2 shows two source images including one focused on the right and the other on the left. Table 1 is the results measured by the objective image fusion performance measure, proposed by Xydeas and Petrović [15] for five sharpness measures and each measure has two types of considerations on a majority filter, with and without the filter. Mutual information proposed by Qu et al. [11] is another common performance measure. The notation R is denoted as our proposed ratio method.

Fig. 2. Source images: focus on the right (left); focus on the left (right).

For simple comparison, we chose the fused image of the maximum performance measure implemented by different block sizes and the five sharpness measures as the compared subject; another subject was the ratio method for the corresponding sharpness measure. Figures 3 and 4 show the fused images constructed by the EOG because it has the maximum performance measure among the five sharpness measures.

Table 1. Various measures with and without a majority filter for different block sizes and our proposed ratio method.

	4	6	8	10	12	14	16	18	20	22	24	26	R	max	min
V	0.6557	0.6669	0.6741	0.6678	0.6758	0.6741	0.6756	0.6752	0.6704	0.6770	0.6724	0.6697	**0.6726**	0.6770	0.6557
V*	**0.6361**	0.6452	0.6655	0.6609	0.6708	0.6650	0.6683	0.6732	0.6704	0.6687	0.6670	0.6721	**0.6717**	0.6732	**0.6361**
EOG	0.6697	0.6768	0.6789	0.6804	0.6790	0.6809	0.6804	0.6808	0.6813	**0.6820**	0.6817	0.6801	**0.6743**	**0.6820**	0.6697
EOG*	0.6682	0.6719	0.6763	0.6788	0.6800	0.6770	0.6802	0.6817	0.6795	0.6791	0.6789	0.6792	**0.6796**	0.6817	0.6682
TEN	0.6645	0.6673	0.6745	0.6711	0.6743	0.6788	0.6788	0.6743	0.6789	0.6814	0.6798	0.6774	**0.6729**	0.6814	0.6645
TEN*	0.6657	0.6590	0.6696	0.6722	0.6742	0.6671	0.6784	0.6772	0.6792	0.6704	0.6760	0.6760	**0.6766**	0.6792	0.6590
EOL	0.6682	0.6752	0.6791	0.6789	0.6795	0.6800	0.6810	0.6803	0.6804	0.6807	0.6808	0.6797	**0.6750**	0.6810	0.6682
EOL*	0.6785	0.6789	0.6805	0.6805	0.6803	0.6813	0.6818	0.6810	0.6801	0.6790	0.6788	0.6790	**0.6787**	0.6818	0.6785
SML	0.6692	0.6755	0.6777	0.6776	0.6795	0.6796	0.6805	0.6801	0.6797	0.6804	0.6808	0.6795	**0.6762**	0.6808	0.6692
SML*	0.6781	0.6791	0.6803	0.6788	0.6802	0.6810	0.6813	0.6809	0.6797	0.6797	0.6788	0.6785	**0.6789**	0.6813	0.6781

* denotes a measure using a majority filter.

Fig. 3. Energy of image gradient (EOG): with a block of size 22 × 22 (left); for the ratio method (right).

Fig. 4. Energy of image gradient (EOG) through a majority filter: with a block of size 18 × 18 (left); for the ratio method (right).

Figures 3 and 4 obviously show that the fused images using EOG with and without a majority filter bring about blocking effects in several places. The experimental results tell us that the maximum performance cannot guarantee great visual effects on its fused image, unless we can find a perfect sharpness measure. However, it's almost impossible for all sorts of images. Even if a majority filter was adopted to remove the isolated blocks, the majority filter cannot clear up all isolated objects of different shapes, except for manual adjustments or further knowledge of information on in-focus objects. On the other hand, the majority filter might incur an extra problem of unduly eliminating the borders of in-focus objects.

On the contrary, our proposed ratio method of image fusion will provide satisfactory visual effects at the expense of mild sharpness values. Particularly, it is simple yet effective, and can easily be applied to any method based on sharpness measures.

VI. CONCLUSIONS

Most of the existing multi-focus image fusion methods lack a consistent way of achieving satisfactory results. They heavily depend on the block size chosen, but the optimal block size cannot usually be applied to other images. Moreover, objective performance measures also play a very important role in image fusion because different performance criteria will easily give distinct outcomes. Whether an area is in focus or not is a very complicated issue. Even a good sharpness measure cannot guarantee a clear and sharp result because a higher measure doesn’t always denote a clearer appearance. A decision according to the magnitude of a measure will possibly result in a wrong selection, thereby producing isolated blocks. Correcting these isolated blocks is not an easy work, even if a majority filter is used. It needs the knowledge and shapes of in-focus and out-of-focus objects, as well as a great sharpness measure.

In this paper, we propose an automatic image fusion scheme for balancing clarity and visual effects, without the cost of how to choose the optimal block size and further information of source images. The proposed method, on one hand, improves the drawback of the average method, which treats all weightings the same. On the other hand, it can be directly and extensively applied to any existing method of multi-focus image fusion to help alleviate blocking effects in order to raise visual effects. The experimental results show that all fused images with considering our proposed automatic scheme are more natural and pleasing than ones fused by other existing methods; no visible blocking effects appear in the fused images at the expense of mild clarity.

ACKNOWLEDGMENTS

This work was supported by the National Science Council, Republic of China, under Grant NSC 101-2221-E-040 -010.

REFERENCES

1. Aslantas, V. and Kurban, R., “A comparison of criterion functions for fusion of multi-focus noisy images,” *Optics Communications*, Vol. 282, pp. 3231-3242 (2009).
2. Buerkle, A., Schmoeckel, F., Kiefer, M., Amavasai, B. P., Caparrelli, F., Selvan, A. N., and Travis, J. R., “Vision-based closed-loop control of mobile microrobots for micro handling tasks,” *Proceedings SPIE*, Vol. 4568, Microrobotics and Microassembly III, pp. 187-198 (2001).
3. Eskicioglu, A. M. and Fisher, P. S., “Image quality measures and their performance,” *IEEE Transactions on Communications*, Vol. 43, No. 12, pp. 2959-2965 (1995).
4. Goshtasby, A. A. and Nikolov, S., “Image fusion: Advances in the state of the art,” *Information Fusion*, Vol. 8, pp. 114-118 (2007).

5. Hu, J. and Li, S., “The multiscale directional bilateral filter and its application to multisensor image fusion,” *Information Fusion*, Vol. 13, pp. 196-206 (2012).

6. Huang, W. and Jing, Z., “Evaluation of focus measures in multi-focus image fusion,” *Pattern Recognition Letters*, Vol. 28, pp. 493-500 (2007).

7. Li, H., Manjunath, B. S., and Mitra, S. K., “Multisensor image fusion using the wavelet transform,” *Graphical Models and Image Processing*, Vol. 57, No. 3, pp. 235-245 (1995).

8. Li, S., Kang, X., Hu, J., and Yang, B., “Image matting for fusion of multi-focus images in dynamic scenes,” *Information Fusion*, Vol. 14, pp. 147-162 (2013).

9. Nayar, S. K. and Nakagawa, Y., “Shape from Focus,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 8, pp. 824-831 (1994).

10. Pajares, G. and de la Cruz, J. M., “A wavelet-based image fusion tutorial,” *Pattern Recognition*, Vol. 37, pp. 1855-1872 (2004).

11. Qu, G., Zhang, D., and Tan, P., “Information measure for performance of image fusion,” *Electronics Letters*, Vol. 38, No. 7, pp. 313-315 (2002).

12. Tian, J., Chen, L., Ma, L., and Yu, W. “Multi-focus image fusion using a bilateral gradient-based sharpness criterion,” *Optics Communications*, Vol. 284, pp. 80-87 (2011).

13. Wang, Z., Ma, Y., and Gu, J., “Multi-focus image fusion using PCNN,” *Pattern*

Recognition, Vol. 43, pp. 2003-2016 (2010).

14. Wong, A. and Bishop, W., "Efficient least squares fusion of MRI and CT images using a phase congruency model," *Pattern Recognition Letters*, Vol. 29, pp. 173-180 (2008).

15. Xydeas, C. S. and Petrović, V., "Objective image fusion performance measure," *Electronics Letters*, Vol. 36, No. 4, pp. 308-309 (2000).

國科會補助專題研究計畫出席國際學術會議心得報告

日期：102年 10 月 21 日

計畫編號	NSC 101－2221－E－040－010－		
計畫名稱	一個自動平衡清晰度和視覺效果的影像融合方案		
出國人員姓名	張炎清、許淮均、闕董辰	服務機構及職稱	中山醫學大學醫學資訊學系 副教授
會議時間	102 年 5 月 23 日 至 102 年 5 月 24 日	會議地點	中國北京
會議名稱	(中文) 2013 電腦、網路和通訊工程研討會 (英文) 2013 International Conference on Computer, Networks and Communication Engineering (ICCNCE 2013)		
發表題目	(中文) 使用區塊二元像素和當特徵的車牌字元辨識 (英文) License Plate Character Recognition Using Block-Binary-Pixel-Sum Features		

一、參加會議經過

現今網路資訊發達，各種訊息經常透過電子郵件傳送，也因此經常可以收到來自國科會、國科會相關學門或其他研討會的訊息。在得知各類研討會後，首先篩選適合的主題，其次審視當時是否有適合發表的題材，當然地點也是一項考量，在綜合三個條件後，發現此研討會符合個人研究背景和興趣，且有適合的題材可以發揮，最重要的也符合當初申請國科會所提出的需求說明，因此便著手規畫此一行程。

為達到更大的效果，個人希望可以藉此機會提升學生的國際觀和視野，因此也與當時參與此題材的專題生討論是否有人願意出席這個盛會。出乎意料之外，原以

為學生會恐懼，沒想到有兩個人表達出高度的興趣並期待能參與，因此便著手寫此次研討會的論文。在論文被接受為口頭報告(oral presentation)後，一方面請旅行社辦理相關的事項，另一方面開始續練學生如何口頭報告，以及請學生規畫相關的參觀行程。

二、與會心得

這次是我第二次參加國際研討會，第一次是 2008 年底在新加坡，儘管沒有第一次的完全新鮮感，但因睽違如此久的時間，還是有某種程度的新鮮度，也多了一層濃濃的比較性，因此收穫良多，尤其是參與的兩位學生，他們都是第一次參加，因此格外興奮，甚至還包括他們的家長，當學生得知他們的老師是博士畢業後才登上國際研討會時，他們更加雀躍；學生家長也覺得不可思議，小朋友不但出國還參加甚麼研討會，聽起來挺嚇人的，非常值得驕傲的一件事。這次的收穫歸納如下：

1. 增加個人並培養學生的國際觀：與來自不同國家的研究人員，利用英語交換意見和心得，可以知道自己領域外的相關知識。
2. 訓練個人和學生的英語聽說能力並增加英文的需求感：在自己的國度接觸英文大都是以讀寫為主，透過此次研討會，參與者必須使用英語聽說能力來和不同國家的研究人員溝通，因此刺激和拓展英文的需求度。
3. 提升個人的研究視野和學生的知識廣度：透過聆聽別人的研究報告可以擴展自己的研究領域，以達到跨領域的碰撞機會；參與的學生也由此了解自己所知相當有限，了解到知識的浩瀚無崖，每個人只是地球進化的一顆小螺絲釘而已。
4. 學習在謙卑中增進人類的福祉：平日在自己的領域奔馳經常誤以為世界就是如此

而已，透過國際研討會的交流，除了吸收不同的專業知識外，還能了解到世界各地的不同面貌。不同的文化造就出不同的國家風格和經濟環境，也間接成就人的不同，以及對世界的看法的差異性。由此讓自己可以學習到世界本是一家，沒有誰優誰劣的問題，每個人每個民族都應該依此來增進共同的福祉。

三、發表論文全文或摘要

發表論文全文。

四、建議

國科會在經費允許的情況下，可以提供更多的機會給研究者來增進國際交流以拓展視野。

五、攜回資料名稱及內容

研討會行程表一份、光碟一片(包含所有被接受的論文集)、參與名牌、現場拍照等。

六、其他

這次借花獻佛，多增加一天的行程讓自己和學生可以有充分的時間參觀北京的風土民情，其中包括參觀北京交通大學、北京大學、清華大學三大名校，由此可以了解到大學竟然可以如此的大、如此的宏偉、如此的美麗。其次也參觀北京紫禁城，以零距離的眼光了解時代變遷和中國的歷史遺跡，還有透過北京的地鐵網了解大陸的現代化和風格。

此外，全聚德烤鴨只有烤鴨還可以，其他餐點普通；北京俏江南的菜還不錯，但記得要告訴服務人員不要辣或者小辣就好，否則還算貴的水果汁會喝不停；需要 lady 幫忙的時候，記得要放下身段，無論美醜都要先稱呼美女，千萬不要稱呼小姐，

否則不但不理你還會招白眼，因為小姐指的是特種行業的女人。

綜合以上觀點，如果台灣如果能善用自己一流的美食、一流的醫學環境、一流的科技水準，還有特殊的地形和人文風貌，很適合發展 7-10 日的美食觀光科技島，以及 10-14 日的健診美食觀光科技島。

國科會補助計畫衍生研發成果推廣資料表

日期:2013/10/21

<table>
<tr><td rowspan="3">國科會補助計畫</td><td>計畫名稱: 一個自動平衡清晰度和視覺效果的影像融合方案</td></tr>
<tr><td>計畫主持人: 張炎清</td></tr>
<tr><td>計畫編號: 101-2221-E-040-010-　　　　學門領域: 視訊與影像分析</td></tr>
<tr><td colspan="2">無研發成果推廣資料</td></tr>
</table>

101 年度專題研究計畫研究成果彙整表

計畫主持人：張炎清　計畫編號：101-2221-E-040-010-

計畫名稱：一個自動平衡清晰度和視覺效果的影像融合方案

成果項目			量化：實際已達成數（被接受或已發表）	量化：預期總達成數(含實際已達成數)	量化：本計畫實際貢獻百分比	單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
國內	論文著作	期刊論文	0	1	100%	篇	
		研究報告/技術報告	0	1	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	為求計畫內容、觀點或技術可以快速在網路上流通，以便進行下一個相關的研究題材，本計畫成果已投稿至國內某 SCI 期刊審稿中。

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：□已發表 ■未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

由於鏡頭潛在的景深(depth of field, DOF)問題，鏡頭在一個時間上只可以聚焦在一個距離上，因而造成銳利度會隨著有限的景深而逐漸衰減。使得相機無法在一個時間點上獲取多個感興趣的物體都聚焦的影像。為此，多聚焦(multi-focus)影像融合(image fusion)是一項非常必要的技術。

由於既存的透過銳利度比較法所得到的融合影像，無法解決因錯誤判斷所造成區塊效應(blocking effects)，導致融合影像只有提高銳利度或對比，但其所產生的視覺效果實在難以令人接受。此時影像融合通常需要額外的資訊來修正初步的融合結果，使得合成(composite)或中間的融合影像更加宜人。一般情況下，待融合的影像都沒有先驗資訊(prior information)可茲利用，因此，引用一個多數決濾波器(majority filter)是一個直覺且廣泛被使用的修正工具。另一方面，應用數學形態學(mathematical morphology)也是這個階段的另一項選擇。只可惜，前者容易造成邊緣部分的錯誤置換，而後者須借助於試誤法(trial and error)，無法達到自動化的效果。

針對上述的問題，本計畫提出一個自動平衡清晰度和視覺效果的影像融合方案，此方案可以有效解決區塊效應，同時具有很高的清晰度和很好的視覺效果。此外，此法可應用在其他新提出的銳利度測量法中，因此，此法不僅具有學術價值，也深具應用價值，且由於它的自動性，也深具有市場價值。