IMAGE STABILIZATION WITH BEST SHOT SELECTOR AND SUPER RESOLUTION RECONSTRUCTION

Jing-Fu Chen (陳景富) and Chiou-Shann Fuh (傅楸善)

Department of Computer Science and Information Engineering
National Taiwan University, Taipei, Taiwan, R.O.C.
R92091@csie.ntu.edu.tw and fuh@csie.ntu.edu.tw

ABSTRACT

In this paper, we present an image stabilization algorithm based on Nikon's Best Shot Selector [11]. We implement Best Shot Selector and use super resolution concept to reconstruct blurred images. Image stabilization includes optical and digital methods. Optical stabilization applies gyroscopic sensors and corrective lens to compensate camera shake optically. Digital stabilization applies image processing algorithms to restore motion blurs. Best Shot Selector avoids the risk of capturing blurred images by sequential capturing and selection. Furthermore, super resolution reconstruction combines images into a sharper image even if the result image of Best Shot Selector is blurred. The experiment results show that our image stabilization algorithm indeed selects and enhances the sharpest image.

1. INTRODUCTION

1.1. Motivation

As the Digital Still Camera (DSC) becomes more and more popular, it develops more and more features. Some of the features provide convenience, such as portability and lightweight. Some of the features provide good image quality, such as telephoto and wide-angle capabilities and high resolution. The two aspects sometimes collide with each other. Hand-held and lightweight device often brings about vibrations which result in blurred images. Hand-shake ruins more photos than anything else. Therefore, Image Stabilization becomes an important issue when we use the compact camera.

1.2. Image Stabilization

Image Stabilization, also known as vibration reduction, is a DSC technology which helps prevent images from becoming blurred. It reduces vibration caused by camera shake, slow shutter speed, or when using a long telephoto lens without a tripod. If you have ever photographed in dim light or tried to hand-hold a long telephoto lens, you know how easy it is to get blur in your images from camera shake. In most cases, we resort to tripods or other camera supports. However, Image Stabilization provides a new way to solve the problems. Image Stabilization is divided into optical stabilization and digital stabilization according to different method.

The DSC with optical stabilization typically has vertical and horizontal built-in gyroscopic sensors. The optical stabilization uses gyroscopic sensors to detect hand-shake, then shifts a corrective lens inside the lens system to compensate for the shake. Although the optical stabilization results in better images, the cost of the compensator is expensive. Another way to get better images is the digital stabilization. The digital stabilization uses image processing techniques to compensate for vibrations.

1.3. Best Shot Selector

Best Shot Selector [11] is a Nikon exclusive technology that takes up to 10 shots as long as the shutter release is pressed. The sharpest single shot is saved and the rest is discarded. In this thesis we present a novel method of image stabilization for DSC. We use the burst capture mode of DSC to get a sequence of images. Best Shot Selector picks up a sharp and clear image among the sequence of images. We can easily select the sharpest image by using gradient edge detector. For different applications, calculation of edge response is flexible. Total calculation or fixed region of interest calculation depends on different applications. Best Shot Selector avoids the risk of capturing blurred images.

1.4. Super Resolution Reconstruction

Although we have a high probability of obtaining the sharpest image after the Best Shot Selector, the sharpest image could still be blurred and not clear enough. We use the other blurred images to enhance the sharpest image by applying super resolution concept. Super Resolution Reconstruction has three principal steps which are Harris Corner Detection [4], Feature Matching, and Image Combination. Harris Corner Detector finds the distinct corners to serve as the features of an image. Feature Matching uses the minimum sum of squared difference error to find the corresponding features correctly. Then Image Combination reconstructs the sharpest image according to the corresponding features.

2. IMAGE STABILIZATION

Image stabilization provides two major methods which are optical and digital. Before explaining the two methods, we should know how the blur occurs and why we need image stabilization. Section 2.1 describes the causes of image blur. Section 2.2 explains optical stabilization in detail. Section 2.3 explains digital stabilization.

2.1. The Causes of Image Blur

One of the biggest problems for a photographer is when a picture is not sharp enough. The reason is either bad focus or hand-shake. Bad focus is usually caused by wrong focal distance. Hand-shake happens almost anytime. Even the steadiest hands are never perfectly still. Even slight hand-shake is transmitted to the camera, where it causes a tiny shift in the optical axis. This effect, called jitter, makes blurred images. There are two major factors for taking blurred images. They are slow shutter speed and telephoto shots.

2.1.1. Slow Shutter Speed

Hand-shake does not always produce blurs. That is because the shutter speed is usually fast enough to prevent jitter. The DSC receives light for just an instant in bright environment. At that fast shutter speed, the photo is captured before jitter appears in Charge-Coupled Device (CCD) sensor. In most ordinary situations, a fast shutter speed prevents jitter. On the contrary, a slow shutter speed probably results in blurred image. The slower the shutter speed, the greater the hand-shake. Figure 2.1 shows how the shutter speed influences image quality in the same camera shake [9]. The left image recorded camera shake motion for only 1/250 second and the hand-shake is virtually unnoticeable. The right image recorded camera shake motion for 1/8 second and the hand-shake ruins the whole image.

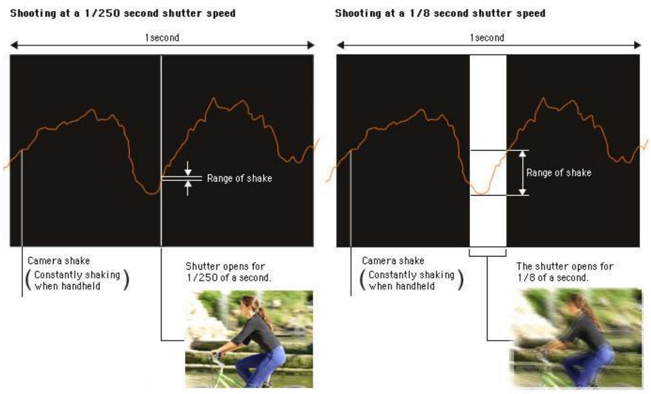


Figure 2.1 Slow shutter speed causes blurred image.

The key to reduce blur is to minimize the length of time that camera shake can enter the image. However, since most DSCs have a zoom function and a small lens diameter, the brightness at the maximum aperture tends to be low. This often calls for slow shutter speeds, so the camera is more susceptible to hand-shake.

2.1.2. Telephoto Shots

When the field of view is narrow, such as close-ups and telephoto shots, the effect of jitter is magnified. During zoom-in, an extremely small area is fitted into the limited space of the frame, and the field of view is narrow. That magnifies the effect of any hand-shake. In addition, the flash in telephoto shots does not reach the subject. Therefore, it is the best to use natural light rather than a flash when taking telephoto shots or close-ups.

Telephoto shots are more vulnerable to hand-shake. Because the focal length is short in a wide-angle shot and the captured area is large, a little hand-shake will not blur the image obviously as shown in Figure 2.2 (a). The focal length is long in a telephoto shot and the captured area is small, so even a small amount of hand-shake will blur the image seriously as shown in Figure 2.2 (b). Therefore, a narrower field of view increases the relative effect of hand-shake. Almost all compact DSCs today come with a zoom function. That makes hand-shake more noticeable in the resulting photos.

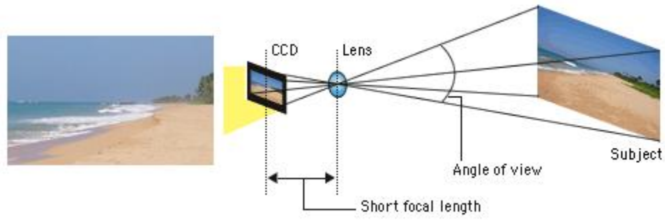


Figure 2.2 (a) Hand-shake effect is small with wide field of view.

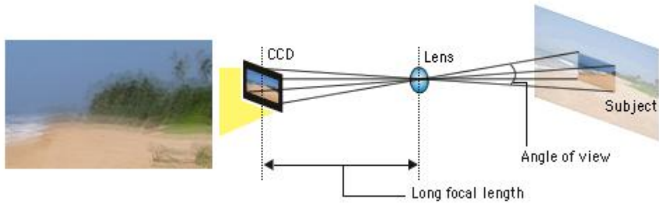


Figure 2.2 (b) Hand-shake effect is great with narrow field of view.

2.2. Optical Stabilization

2.2.1. Vibration Reduction System

Nikon makes an optical stabilization system named Vibration Reduction (VR) System as shown in Figure 2.3 [8]. The VR System contains two gyroscopic sensors, two Voice Coil Motors (VCM), and a corrective lens. The three components are essential for optical stabilization. To compensate for image blur, the angular velocity of camera shake should be detected accurately. The two gyroscopic sensors detect angular velocities in horizontal and vertical directions respectively. Then these angular velocity data is sent to a microcomputer to calculate the amount of lens movement. VCM moves the VR lens in a particular direction. This is controlled by the electronic current inside the VCM’s magnetic field. VR lens corrects the slanted optical axis by vibration.

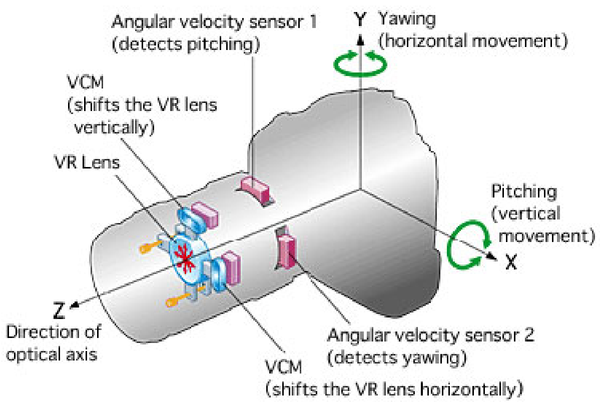


Figure 2.3 Nikon VR system.

2.2.2. Gyroscope Principle

Gyroscopes are well known as sensors to detect the revolution movements or angular velocities of objects. Gyroscopes using piezoelectric vibrators are called piezoelectric vibrating gyroscopes [7]. The gyroscope utilizes a physical phenomenon called Coriolis force. Foucault pendulum is frequently used to explain this force as shown in Figure 2.4. When a mass (*m*) is vibrating with a velocity ($\vec{v}$) and given a revolution (angular velocity $\vec{\omega}$), Coriolis force (*Fc*) operates in a direction perpendicular to $\vec{v}$ as expressed by $Fc = -2m\vec{\omega} \times \vec{v}$ [10].

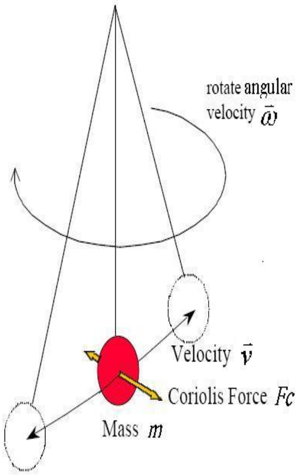


Figure 2.4 Foucault pendulum.

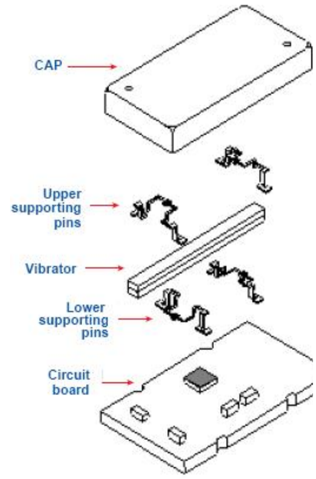


Figure 2.5 ENC-03M structure.

Since this gyroscope uses a piezoelectric ceramic, Coriolis force can be detected and transformed to electric signals by the basic principle of piezoelectric ceramic. Figure 2.5 shows the structure of piezoelectric vibrating gyroscope made by Murata Manufacturing Co., Ltd. [1].

Murata's Gyroscope Application Guide describes exactly how it works [6]. The structure of a ceramic bimorph vibrator is shown in Figure 2.6. The piezoelectric ceramic vibrates at the resonant frequency in a vertical direction. When a rotational velocity is introduced into the system, the Coriolis force causes the ceramic to vibrate in the horizontal direction. This horizontal vibration causes the ceramic material to distort from the left to the right of the material. This distortion causes the piezoelectric material to change the phase angle of the input voltage from the left side to the right side. The sensors on the top of the bimorph material read the analog output voltages as vibrating amount.

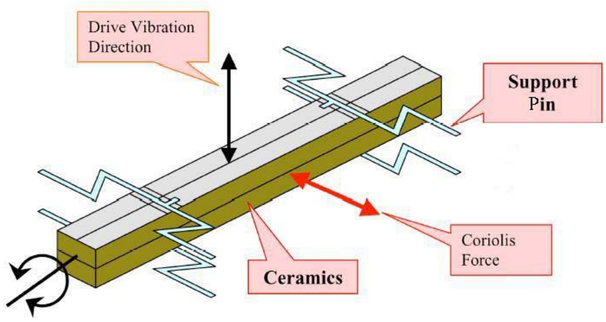
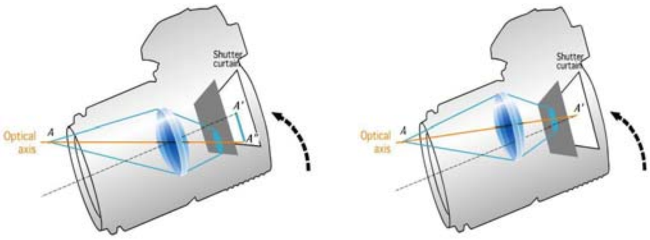


Figure 2.6 Ceramic bimorph vibrator.

2.2.3. Compensation Effect

Figure 2.7 shows how the light from subject (*A*) reaches focal plane (*A'*) through the lens. Without camera shake, point *A* on optical axis should be captured at position *A'* on the CCD. However, if the camera vibrates in vertical direction, the light from *A* will converge on *A"* as shown in Figure 2.7 (a). This shift from *A'* to *A''* results in image blur. The VR System works by adjusting the lens to make sure light from *A* converges on *A'* as shown in Figure 2.7 (b). When the gyroscopic sensors detect θ angles slant, the lens shifts $focal\_length \times \tan\theta$ displacement to compensate the slant. This special mechanism ensures clear and sharp pictures even when hand-shake occurs.



(a) VR off. (b) VR on.

Figure 2.7 Compensation effect of VR system.

We analyze the compensation effect as shown in Figure 2.8. Without vibration, point *A* should converge on *A'*. With θ slant, point *A* converges on *A''*. To correct the shift from *A'* to *A''*, the lens center moves from *C* to *D* perpendicular to the optical axis. The optical axis moves from $\overleftrightarrow{PC}$ to $\overleftrightarrow{QD}$. Line segments $\overline{DF}$ and $\overline{CE}$ are the same focal length. We want to know why the lens movement $\overline{CD}$ is $focal\_length \times \tan\theta$. The following calculations prove it.

$$\Delta ADC \sim \Delta AA'A'' \qquad \Delta AGD \sim \Delta ABA'$$

$$\therefore \frac{\overline{CD}}{\overline{A''A'}} = \frac{\overline{AD}}{\overline{AA'}} \qquad \therefore \frac{\overline{AD}}{\overline{AA'}} = \frac{\overline{GD}}{\overline{BA'}} \qquad \exists \overline{BA} = \overline{GC}$$

$$\Delta GDF \sim \Delta GCA'$$

$$\therefore \frac{\overline{GD}}{\overline{GC}} = \frac{\overline{DF}}{\overline{CA'}} \qquad \Rightarrow \frac{\overline{CD}}{\overline{A''A'}} = \frac{\overline{DF}}{\overline{CA'}}$$

$$\overline{A''A'} = \overline{CA'} \times \tan\theta$$

$$\overline{DF} = focal\_length$$

$$\therefore \overline{CD} = \frac{focal\_length}{\overline{CA'}} \times \overline{CA'} \times \tan\theta$$

$$= focal\_length \times \tan\theta$$

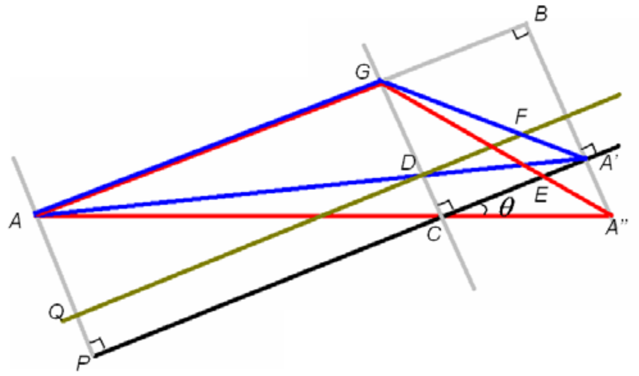


Figure 2.8 Compensation analysis.

The gyroscopic sensors detect vertical and horizontal camera slants, and then the corrective lens moves according to the angle of inclination. Therefore, the optical stabilization system successfully avoids the image blur by little camera shake.

Here is an example how well the VR system works. The Coolpix 8800 is the first Nikon digital camera with VR technology [5]. Figure 2.9 (a) and (b) are parts of test images by Coolpix 8800. The two images are both taken at 1/15 second shutter speed.

(a) VR off. (b) VR on.

Figure 2.9 Examples of VR system.

2.3. Digital Stabilization

Digital stabilization uses software to adjust images after capturing. At first, digital stabilization applies to DVCs (Digital Video Camcorders) for reducing unwanted camcorder motion. For example, digital stabilization crops the captured image and attempts to correct for shake by moving the area of the crop within the frame as shown in Figure 2.10. The area of CCD sensor is larger than recorded area. The algorithm finds the most similar crop with minimum mean square error to compensate the displacement by vibration. Besides, there are various digital stabilization algorithms for DVCs and DSCs. In the thesis, Best Shot Selector and super resolution reconstruction are also digital stabilization algorithms for DSCs and described at chapter 3 and 4.

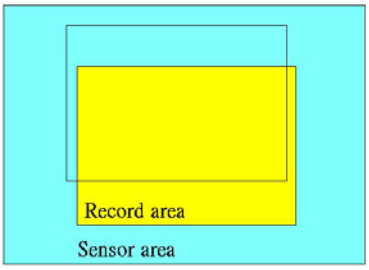


Figure 2.10 Digital stabilization.

Digital stabilization often results in an artificial looking image. Although digital stabilization is usually considered inferior to optical stabilization, digital stabilization saves camera manufacturers additional hardware and takes up less space inside the camera.

3. BEST SHOT SELECTOR

Nikon's BSS identifies and saves the sharpest image automatically from ten sequential shots. In this chapter, we implement Best Shot Selector by gradient edge detector. Section 3.1 describes the principle of gradient edge detector. Section 3.2 shows total calculation and region of interest calculation. Section 3.3 has the flowchart and experimental results.

3.1. Gradient Edge Detector

How do we determine an image is sharp or blurred? The key point is the edge. Edges are placed in the image with strong intensity contrast. Since edges consist of mainly high frequencies, we can detect edges by applying a high-pass frequency filter in Fourier domain or by convolving the image with an appropriate kernel in the spatial domain. In practice, edge detection is performed in the spatial domain, because it is computationally less expensive and often yields better results. Edges correspond to illumination gradients, and therefore we can measure them by calculating the derivatives of the image.

We use the gradient magnitude g to determine whether the pixel is an edge or not. Convolution masks K_v and K_h are Sobel Edge Detector kernels [3]. They come from a smooth mask and a differential mask convolved together. The smooth mask suppresses noise. The differential mask calculates the vertical and horizontal gradients. For an input image $I(x,y)$, the magnitude $G_v(x,y)=\sum_{i=-1}^{1}\sum_{j=-1}^{1} I(x+i,y+j)\times K_v(i,j)$ represents the vertical edge response and the magnitude $G_h(x,y)=\sum_{i=-1}^{1}\sum_{j=-1}^{1} I(x+i,y+j)\times K_h(i,j)$ represents the horizontal edge response of the pixel (x,y). Therefore, the magnitude $g=\sqrt{G_v^2+G_h^2}$ indicates the edge response of the pixel (x,y). To suppress the influence of noise, we only sum up the values over threshold (*grad_threshold*=50).

3.2. Region of Interest (ROI) Calculation

There are two ways to calculate the edge responses. The straightforward way is to calculate responses for every pixel of the whole image. Perhaps the total calculation takes much time, but it includes the details of the whole image and is more accurate. Another way is to calculate responses for the pixels in ROI. We divide an image into 8x8 grids and select the center area as the ROI as shown in Figure 3.1. Assume that the subject is generally in the center. We choose the central 12 grids and the corner 4 grids from total 64 grids. ROI calculation economizes on computing time. We only need to compute a quarter of total pixels and spend quarter time.


Figure 3.1 Region of interest.

3.3. BSS Flowchart and Experimental Results

The procedure of our program is shown in Figure 3.2. The beginning initializes a criterion value $max=0$. Then we decide the calculation region is ROI or total. Compute Sobel edge response and compare with *max* to keep sharper image in buffer. The sharper image has larger edge response. We only keep two images in memory at the same time to compare which is sharper. After checking all images, the image in buffer is the sharpest one.

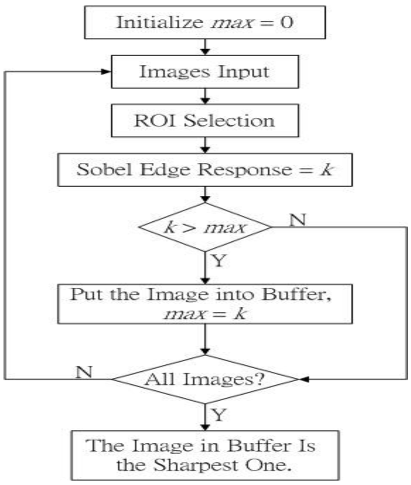


Figure 3.2 Best Shot Selector flowchart.

We have 30 groups of testing images where every group has three images captured by sequential shots of OKWAP M216. The burst capture mode of M216 can take three shots in about two seconds. Figure 3.3 is an example that program result is the sharpest image with the largest response.

(a) 10933863 (b) 24195702 (c) 18536121

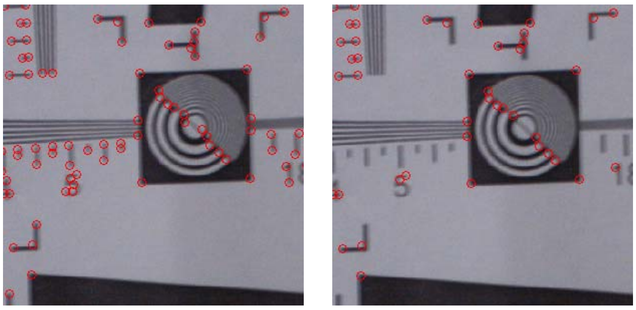
Figure 3.3 Responses of BSS.

4. SUPER RESOLUTION RECONSTRUCTION

The original goal of super resolution is to combine images with similar scene into a high resolution image. However, our goal is not to get a high resolution image but to get a much sharper image. We follow and modify the method of super resolution to achieve our goal.

4.1. Harris Corner Detector

To combine images accurately, we first find out the characteristic parts of these images. The distinct corners serve as the features of an image. Harris and Stephens provide a simple algorithm to detect corners [4]. There are four major steps to find out corners. First, compute horizontal and vertical derivatives for each pixel. Second, define autocorrelation matrix $M=\sum_{x,y} w(x,y)\begin{pmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{pmatrix}$ and calculate corner response at each pixel. Then, set a threshold to suppress noises and find local maximum as the corner point. Figure 4.1 shows the results of Harris corner detector.



(a) *threshold*=400 748 corner. (b) *threshold*=1300 298 corner.

Figure 4.1 Harris corner detector examples.

4.2. Feature Matching

4.2.1. Minimum Sum of Squared Error

Minimum sum of squared error is a mathematical technique to find the similarity. To determine if two images are similar, we use the sum of squared difference between two images. The smaller the sum of squared error is, the more similar the two images are.

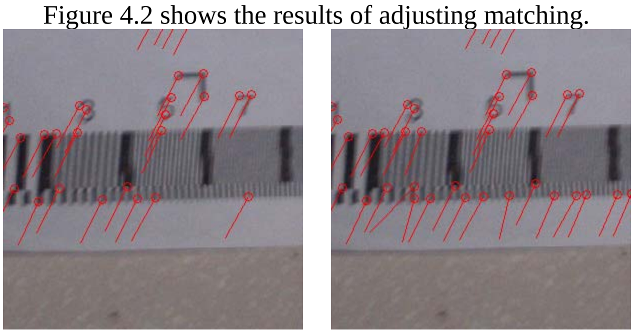
We assume that the scene does not change too much during burst capturing. For each feature, we take the 15x15 neighborhoods as the feature descriptor which has 225 pixels (*diff_scale*=15). Then we calculate the sums of squared difference error with the features only in 200x200 neighborhoods (*match_range*=200). Eventually, decide the corresponding feature with the minimum sum of squared difference error.

4.2.2. Two-way Check and Adjusting Matching

The purpose of two-way check is to guarantee that the corresponding features both have the minimum sums of squared error with each other. We have to calculate sum of squared error both sides and ensure the corresponding features with minimum value.

To match more accurately, adjusting matching is required. We assume that the scene has a consistent moving direction. Therefore, discard the moving vectors which diverge from the median. We filter out the vectors which are not centralized enough. First, find out the median value of matched vectors. Second, cancel the vectors which have values out of 10 pixels range of median value (*adjust_vector*=10). Finally, discard the vectors whose angles exceed 5 degrees error (*adjust_angle*=5).

Figure 4.2 shows the results of adjusting matching.



(a) 284 features matched. *adjust_vector*=10 *adjust_angle*=5 (b) 327 features matched. *adjust_vector*=40 *adjust_angle*=20

Figure 4.2 Adjusting matching examples.

4.3. Image Combination

4.3.1. Grid Vector Calculation

Grid vector calculation is the preprocessing of image combination. There are only some pixels which have feature vectors. We use grid-based combination and divide an image into many grids of 128x128 pixels (*comb_scale*=128). For each grid, find the average of feature vectors falling in this grid. When there is no feature falling in the grid, use the nearest neighbor as the grid vector. Although we assume that the scene has a consistent moving direction, the feature vectors still have little difference in fact. The object displacement in

3D scene varies with object distance, so we do not use a uniform vector for image matching.

4.3.2. Proportional Combination

Proportional combination gets the sharper parts of two images. The combination is proportional to edge responses of two grids. For example, if the edge responses of grids A and B are 4:6, we combine the two grids in a proportion of 40% and 60%. However, when the part of much blurred 40% A influences the part of sharper 60% B, we can not get the best result as shown in Figure 4.3 (a).

(a) Original proportion. (b) Enhanced proportion. (c) Winner takes all.

Figure 4.3 Examples of proportional combination.

Why not use the rule of winner-takes-all? If we only compare the edge responses to decide which grid takes up all proportion of combination, block effect frequently occurs in the result image. There are obviously unnatural seams between grids as shown in Figure 4.3 (c). Therefore, we use a proportional factor (*comb_factor*=0.4) to raise the ratio but avoid the block effect. For example, the ratio 4:6 will be raised to $(40\times 0.4):(100-40\times 0.4)=16:84$. We use the enhanced ratio as the proportion of combination as shown in Figure 4.3 (b).

The advantage of our combination method is to integrate the sharper region of the two original images. Therefore, if the two original images blur in different directions, this method produces an excellent result image. On the contrary, if one image of the two original images blurs seriously, there is almost no chance to sharpen the other one. Figure 4.4 shows that our method integrates two images with different blurred directions into a better image.

(a) Image A. (b) Image B. (c) Combined Image.

Figure 4.4 Examples of combination.

4.4. Super Resolution Reconstruction Flowchart and Experimental Results

We summarize super resolution reconstruction. First, detect the corners as features for two input images A and B. Second, match the features from two images and output the corresponding vectors. Finally, combine the two images according to the corresponding vectors and edge responses. Figure 4.5 shows the flowchart of super resolution reconstruction.

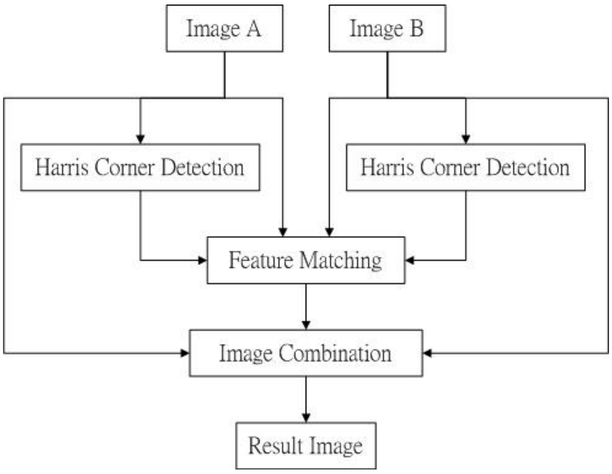


Figure 4.5 Flowchart of super resolution reconstruction.

Figure 4.6 shows the distribution of edge responses. The edge responses of combined images are all higher than two input images. Figure 4.7 is an example of super resolution reconstruction.

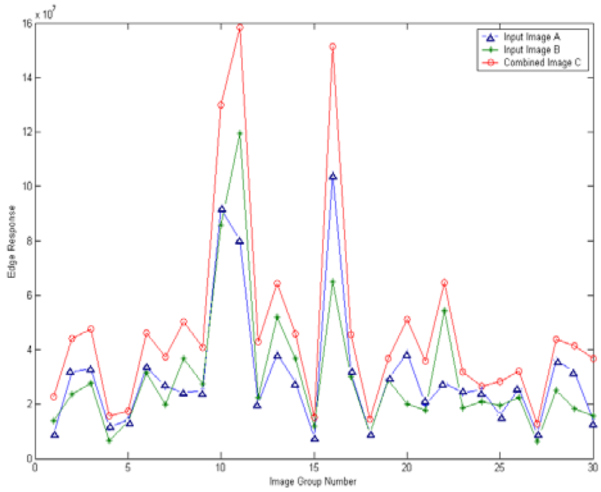


Figure 4.6 Edge responses of combined image.

(a) 80115989 (b) 158345703 (c) 119508425

Figure 4.7 Examples of super resolution reconstruction.

5. CONCLUSIONS AND FUTURE WORKS

5.1. Conclusions

We use gradient edge detector and region of interest calculation to implement Best Shot Selector. Best Shot Selector aims to avoid the risk of capturing blurred images by sequential capturing and to save the sharpest image. When capturing in a dark environment or using a long telephoto lens, Best Shot Selector works well and is worth to be one of capturing mode.

However, we have to notice that Best Shot Selector saves the sharpest image rather than the best image possible with the scene. If the subject is moving during sequential shooting, the saved image may not be what the image user wants. A good image always records the best moment. The problem is also the general problem of photography and is not the issue of Best Shot Selector. For the purpose of sharp and clear image, Best Shot Selector still achieves the goal.

Moreover, we assume that the scene has a consistent moving and nearly uniform motion. When the subject moves randomly and causes non-uniform motion blurs, super resolution reconstruction has difficulties integrating the images.

5.2. Future Works

To improve the digital stabilization algorithms, the key point is the tradeoff of fast algorithm and high accuracy. The computation time should be in a tolerable range. Fast and accurate algorithm is necessary for implementing on DSC.

Generally speaking, optical stabilization can obtain more image details during capture. Digital stabilization only compensates and reconstructs images in limited image details. Even optical stabilization can not avoid motion blur entirely. Therefore, if we integrate optical and digital stabilization methods, the result image could be better. Digital stabilization can get the moving information from gyroscopic sensors of optical stabilization to restore image precisely. Optical stabilization needs the assistance of digital stabilization to eliminate motion blur completely. Image stabilization mixed with optical and digital methods is a much better system.

6. ACKNOWLEDGEMENT

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

7. REFERENCES

[1] H. Abe, “Applications Expand for Downsized Piezoelectric Vibrating Gyroscopes,” Asia Electronics Industry, Vol. 8, No. 79, pp. 49-50, 2003.

[2] D. Capel and A. Zisserman, “Computer Vision Applied to Super Resolution,” IEEE Signal Processing Magazine, Vol. 20, No. 3, pp. 75-86, 2003.

[3] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 2nd Ed., Prentice-Hall, Upper Saddle River, New Jersey, 2002.

[4] C. Harris and M. Stephens, “A Combined Corner and Edge Detector,” Proceedings of Alvey Vision Conference, Manchester, UK, pp. 147-151, 1988.

[5] J. Keller, “Digital Camera Resource Page Review: Nikon Coolpix 8800,” http://www.dcresource.com/reviews/nikon/coolpix8800-review/index.shtml, 2004.

[6] Murata Manufacturing, “Application Guide for Gyrostar,” http://www.murata-northamerica.com/, 2005.

[7] Murata Manufacturing, “Piezoelectric Vibrating Gyroscopes,” http://www.murata.com/catalog/s42e3.pdf, 2005.

[8] Nikon, “Precise Camera-Shake Compensation at Every Angle VR (Vibration Reduction) System,” http://www.nikon.co.jp/main/eng/portfolio/about/technology/nikon_technology/vr_e/index.htm, 2005.

[9] Panasonic, “Technology That LUMIX Takes the Shake out,” http://panasonic.co.jp/pavc/global/lumix/popup/qp_guide/basic/basic04.html, 2005.

[10] D. Van Domelen, “Getting around the Coriolis Force,” http://www.physics.ohio-state.edu/~dvandom/Edu/newcor.html, 2000.

[11] P. Yuen, “Nikon Best Shot SelectorTM,” http://digitalimaging.patyuen.com/shootout/cp5kvg2/bss.htm, 2003.