

BUILDING A GOOD IMAGE FROM TWO COLOR IMAGES OF CAMERA MODEL USING GA AND DISCRETE WAVELET TRANSFORM

Pham The Bao, Pham Thanh Trung

University of Natural Sciences, VNU-HCM

(Manuscript Received on November 18th, 2007, Manuscript Revised December 2nd, 2007)

ABSTRACT: *We set up a system to build a good image from two-camera model. Our system is imitated a human visual system with two cameras referring to two eyes; each eye captures its own view and two separate images are sent to brain for processing to unit into one image. Similarly, in our system, we present a process to find out the relationship of two images captured from two cameras, and then these images are synthesized to build a good image. This image can be applied in other processes such as object recognition, detection or tracking...And our system can be used as robot vision.*

Keywords: *2-cameras model, GA, discrete wavelet transform, robot, fusion.*

1.INTRODUCTION

The human visual system includes two eyes having the same structure and the same function, but there is small difference of two captured images due to the different positions. Human brain takes advantage of the small differences for processing to build a single image that contains better information; that is human vision. Each eye is referred to a camera being able to capture its view and form an image. In the real world, since the camera is affected by external factors, for example: illuminations, environment..., the taken image is not good quality. Therefore, if there are two cameras co-operated a system like human eyes, we could have a better image from this system. The problem is to find out the relationship of two images based on the relationship of two cameras, then synthesizes these images to form final image having better quality.

Two-camera model is also called stereo vision has a classical problem, or correspondence problem; the objective is to determine pairs of points that correspondence to the same scene point [1]. There are two main approaches to solve this problem: correlation-based and feature-based [2, 15]. Recently data fusion, especially image fusion [4, 5] is one of research areas concerned by many people. Image fusion is and will be an integral part of many applications such as: intelligent robot or remote sensing [17].

We build the system including two cameras that work together simultaneously. The objective of this system is to build a single image having good content or good information. The images captured from camera system are processed to find out the disparity or the common regions, and then are synthesized to produce a single image better than two first images. We used GA approach in solving correspondence problem [9] and Discrete Wavelet Transform (DWT) in image fusion [5, 10] to be the basic of the whole system processing.

2.GENERAL SYSTEM

2.1.System Framework

Our camera system is described in detail via diagram 1.

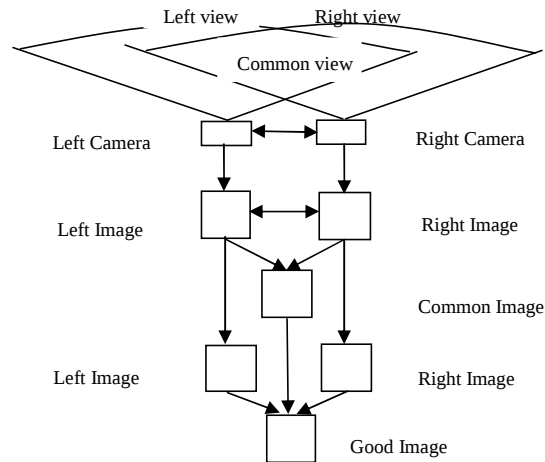


Diagram 1. Two-camera system

2.2.The Positional Relationships of Two Cameras

Human visual system includes two eyes having the same structure, the same view direction, and the same field of view; the distance between them is approximately 6.5cm. Similarly, we set up the system including two cameras having the same internal parameters. If the first camera is set at origin of coordinate system, then the second camera is translated along z-axis; in the other word, the angle between two view directions is zero. For our eyes, the crystalline lens is changeable to enable view near or far objects. Nevertheless, the focal length of camera is fixed or unchangeable and the field of view is much smaller than human eye's. Therefore, the distance between two cameras in our system is able to change from 4cm to 7cm depending on the distance between objects and camera system.

3.BUILDING A GOOD IMAGE

3.1.The Relationship of Two Images

Two images captured from camera system are slightly different due to the difference of camera's position. The distance between two cameras is set closely enough to able to preserve the order relationship of two images, fig. 3; this is an important condition when we set up the camera system. Furthermore, the information in the common view (diagram 1) on two images is not completely similar due to the external factors, fig. 1.



Figure 1. The corresponding relationship of two images

3.2.Determining Disparity and Common Regions of Two Images

In this step, we need to determine the disparity of two images, and then find out the common regions of the same view. Since the camera system is set specially, there is just the horizontal disparity of two images, no angle disparity. Moreover, there is the order relation between two images, fig. 3, so we just need to determine the disparity of some corresponding pairs of points on two images. Consider $A(x, y)$ is a sample pixel in left image, and then the corresponding pixel of A is $A'(x, y')$ in right image. At this time the horizontal disparity between two point is $d = |y' - y|$, then we get $y' = |y - d|$. Now to be able to find out point A' , we just change the value of d in its range until we can get the most suitable one. The degree of correlation between A and A' can be determined by using sum squared differences (SSD), equation 1, based on the pixel intensity [3]. The most suitable value of d will give us to minimum value of function SSD in equation 1.

$$SSD_d = \sum_w (I_L(x, y) - I_R(x, y + d))^2 \quad (1)$$

- $I(x, y)$ is intensity at pixel (x, y) .
- W is the region containing the neighbors around pixel (x, y) .

However, in practice due to the errors when we set up the camera system, two images usually have both horizontal and vertical disparity. Therefore, we use equation 2 to determine the degree of correlation instead of equation 1. Yet it takes a mount of cost to compute, so we applied GA [14] to reduce the cost of computing [9, 13].

$$SSD_d = \sum_w (I_L(x, y) - I_R(x + d_x, y + d_y))^2 \quad (2)$$

The disparity $d = (d_x, d_y)$ of two images will be encoded in a 12 bit binary string, fig. 2, with the first 4 bits referred to vertical disparity and 12 bits rest referred to horizontal disparity.

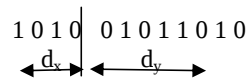


Figure 2. The disparity d is encoded in 12 bits binary string

At the beginning, the population includes 50 solutions or chromosomes generated randomly, and then new chromosomes will be created by genetic operators such as mutation and crossover.

Assumption, A_1 and A_2 are two chromosomes in the population:

$A_1 = 1 0 1 1 1 0 1 1 0 0 0 0$

$A_2 = 1 0 1 0 1 0 1 1 0 1 1 0$

The crossover of these two chromosomes is performed by exchanging 6 last bits. As the result, we will have 2 new chromosomes:

$P1 = 1 0 1 1 1 0 1 1 0 1 1 0$

$P2 = 1 0 1 0 1 0 1 1 0 0 0 0$

The mutation is performed on single chromosome by change the value of bit 6th from 1 to 0 or from 0 to 1. Suppose $B_1 = 1 0 1 1 1 0 1 1 0 0 0 0$ is chromosome need to mutated, and then $Q_1 = 1 0 1 1 1 1 1 1 0 0 0 0$ is the result after mutation.

After generating new chromosomes via crossover and mutation, we just choose 50 best chromosomes with high accommodation. We can evaluate the accommodation of these chromosomes according to fitness function, equation 3.

$$F(d) = \sum_w (I_L(x, y) - I_R(x + d_x, y + d_y))^2 \quad (3)$$

At each step in the process, 50 new chromosomes is created by crossover, and then 3 worst ones will be mutated; this can generate better chromosomes. This process will be stopped when there is a new chromosome that has accommodation or fitness value greater than a threshold $t=15$, the value of fitness function. However in some case, that special chromosome never occurs, so we should force to stop this process after 7 generations, and then choose the best one.

After determining the disparity between two images, we have a region in the left image (CI_L) and a region in the right image (CI_R) being corresponding or having the same view, fig. 4. These regions then are combined to build a better image.

3.3.Fusing Two Common Images

Although the field of view of our eyes is very wide, we just concern the objects in front of our face. That means we just enable to see clearly objects in the common region of two views. Similarly, the only information in the common view can be fused to make them better, diagram 1.



Figure 3. The order relation is preserved.



Figure 4. (a) Left region; (b) Right region

After we have CI_L and CI_R , we can consider that each is a set of data. Then, these two set of data are combined to form a new better set of data. In practice, there are many factors affecting the quality of images. Therefore, multi-resolution analysis (MRA) is the best way to decompose two sets of data before fusing and forming a new set of data.

DWT is one of popular tools used in MRA by using low-pass and high-pass filter [6, 7, 8]. Consider $\mathcal{W}$ is DWT performance, the analysis and synthesis is described via equation 4 below.

$$SI = \omega^{-1}(\min(\omega(CI_L), \omega(CI_R))) \quad (4)$$

We use ω to analyze CI_L and CI_R into two collections D_L and D_R include wavelet coefficients as in equation 5 and 6, diagram 2. Then, D_L and D_R are combined by choosing the minimum wavelet coefficient between corresponding pairs of coefficient, equation 7. Finally, synthesized image (SI) will be perceived by taking inverse DWT, diagram 3.

$$D_L = \omega(CI_L) = \{LL, LH, HL, HH, \dots\}_L \quad (5)$$

$$D_R = \omega(CI_R) = \{LL, LH, HL, HH, \dots\}_R \quad (6)$$

$$\{LL, LH, HL, HH, \dots\}_{New} = \min(D_L, D_R) \quad (7)$$

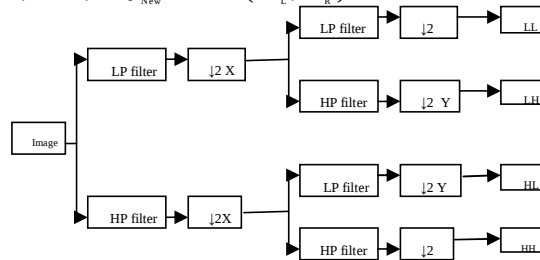


Diagram 2. Analysis process ω .

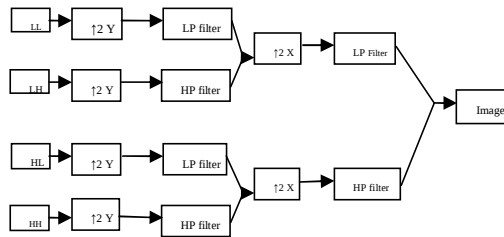


Diagram 3. Synthesis process ω^{-1}



Figure 5. Two images in the common view: (a) CIL; (b) CIR.

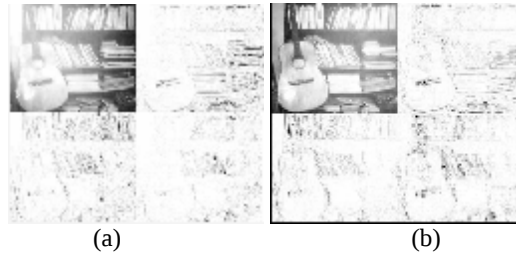


Figure 6. Analyzed result using Daubechies Wavelets level 1: (a) got from 5a; (b) got from 5b.

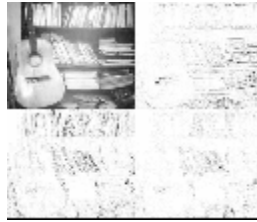


Figure 7. Synthesis data.

Each color image combines three separate colors: red, green and blue. These are 3 basic colors from that we can create lots of different colors. Therefore, we need to perform the analysis and synthesis on each color separately to build a good color image. Fig. 5 and 6 are two red color images perceived from two original color image 4a and 4b. After passing these images through the filter system of Daubechies Wavelets at level 1 [12], we get four images from each one of size equal to quarter the size of original image, fig. 6. Next, we fuse in succession corresponding pairs of image for those by minimum operator to get four new better ones, fig. 7. Finally, we can synthesize from these images into a red image with complete information, fig. 8. For other colors, we perform the same operator.



Figure 8. Synthesized image with red color.

4.CONCLUSION

We performed capture images from our camera system with many different positions. Fig. 9a and 9b are the taken results when the distance between objects and camera system is 2m and the distance between two cameras is 6cm. Due to the external factors, the quantity or resolution is not good, some regions in the images is blurry. So we need to increase the

resolution of image to have complete information. Fig. 9c is the result after processing via our system; the final image is better than two original images.

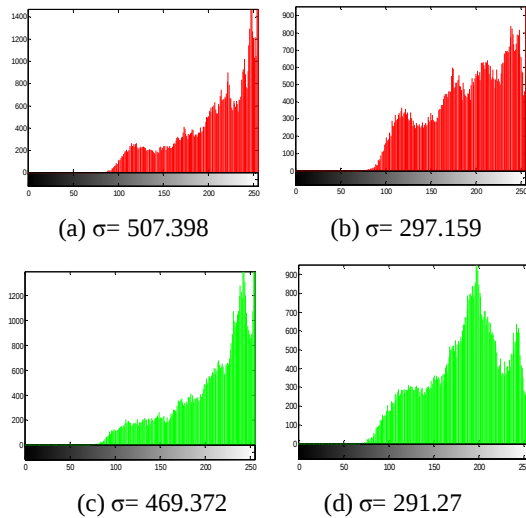
The taking image position plays a very important role for our system. If we take images near to our camera system, then the information of two images may be the same and enable to build a better image. If we take images far to our system, the difference between two images is much and cannot complete each other. These cases still happen with our visual system normally.



Figure 9. (a) Taken image by left camera,
(b) Taken image by right camera, (c) Synthesized image.

After experiencing many times with many different positions and various light condition, our system can give good result with range from 1.5m to 5m for capturing position; the average time for running the process is 2.9 seconds performed with Matlab language on PC Pentium 4 and 320x240 resolution cameras.

The objective of this system is to build a good image with more meaningful information than two original images. Now the quantity of synthesized image can be evaluated via standard deviation σ of histogram. The smaller the standard deviation is, the better the image is. Fig. 10 shows the comparison about histogram and standard deviation in RGB separately of left image (fig. 9a) and synthesized image (fig. 9c). Similarly, fig. 11 is for right image (fig. 9b) and synthesized image (fig. 9c). It is shown clearly that the σ value of synthesized image is smaller than two original ones. In other words, the quantity of final image is better.



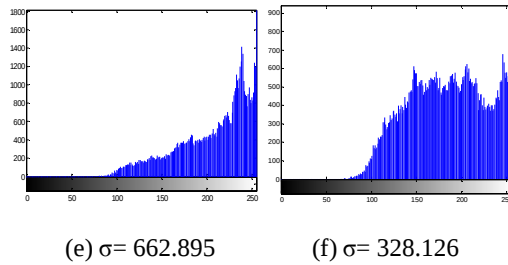


Figure 10. Histogram and standard deviation comparisons. (a), (c), (e) are histograms and standard deviation corresponding to R, G, B color of image 9a; (b), (d), (f) are histograms and standard deviation corresponding to R, G, B color of image 9c

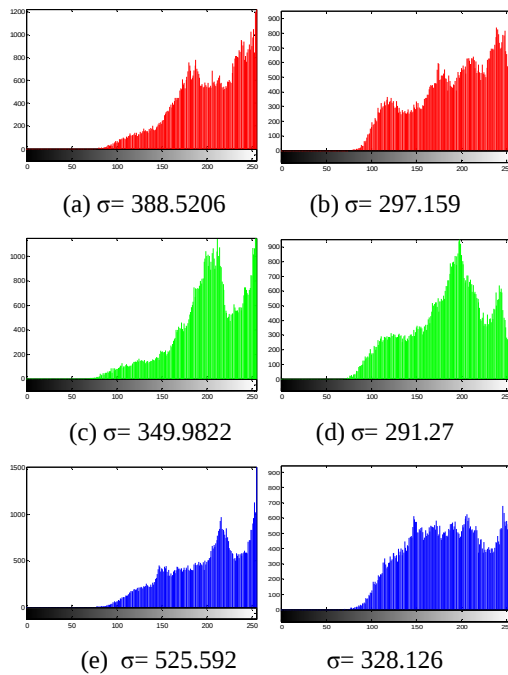


Figure 11. Histogram and standard deviation comparisons. (a), (c), (e) are histograms and standard deviation corresponding to R, G, B color of image 9b; (b), (d), (f) are histograms and standard deviation corresponding to R, G, B color of image 9c

In the future, we will try with various focal lengths, various positions of cameras and different fields of view for our camera system. We also set up the fuzzy system for these relationships to be able to choose the best position of cameras giving best quality image as changing these values.

XÂY DỰNG ẢNH CHẤT LƯỢNG TỐT TỪ HAI ẢNH ĐƯỢC LẤY TỪ MÔ HÌNH CAMERA BẰNG PHƯƠNG PHÁP GA VÀ BIẾN ĐỔI WAVELET RỜI RẠC

Phạm Thế Bảo, Phạm Thành Trung
Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM

TÓM TẮT: Chúng tôi xây dựng một hệ thống cho phép nhận được một ảnh có chất lượng tốt từ mô hình hai camera. Mô hình của chúng tôi dựa trên hệ thống thị giác của con người, với hai camera như hai mắt người; mỗi mắt sẽ nhận một ảnh rồi truyền về não bộ để tổng hợp thành một ảnh mà chất lượng sẽ tốt hơn từng hình ảnh của từng mắt. Chúng tôi trình bày một quá trình tìm mối tương quan của hai ảnh được lấy từ hai camera, dựa vào mối tương quan này chúng tôi tổng hợp một ảnh mới có chất lượng hơn. Ảnh tổng hợp này có thể được dùng trong các hệ thống nhận dạng, xác định và theo vết chuyển động, ... và hệ thống này có thể được xem như thị giác của robot.

REFERENCES

- [1]. R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, second edition, (2004).
- [2]. Håkan Bjurström and Jon Svensson, *Assessment of Grapevine Vigour Using Image Processing*, Master Thesis, Linköping University, Sweden, (2002).
- [3]. Hill P.R., Bull D.R., and Canagarajah C.N., *Image Fusion Using A New Framework For Complex Wavelet Transforms*, Image Processing, IEEE International Conference, (2005).
- [4]. Gema Piella Fenoy, *Adaptive Wavelets and their Applications to Image Fusion and Compression*, Ph.D Thesis, Mathematics and Computer Science (CWI), Amsterdam, (2003).
- [5]. Laure J. Chipman and Timothy, *Wavelet and Image Fusion*, Proceedings of international conference on image processing, pp. 248–251, (1995).
- [6]. William F. Herrington, Jr Berthold K.P. Horn, and Lchiro Masaki, *Application of the Discrete Haar Wavelet Transform to Image Fusion for Nighttime Driving*, Proceedings of Intelligent Vehicles Symposium, IEEE, (2005).
- [7]. M. A. Berbar, S. F. Gahe, and N. A. Ismaill, *Image Fusion Using Multi Decomposition Levels Of Discrete Wavelet Transform*, Visual Information Engineering, International Conference on Page(s):294 – 297, VIE (2003).
- [8]. Lee A. Ray and Reza R. Adhami, *Dual Tree Discrete Wavelet Transform with Application to Image Fusion*, Proceeding of the Thirty-Eighth Southeastern Symposium, Page(s):430 – 433, (2006).
- [9]. Pengcheng Zhan, Dah-Jye Lee, and Randal Beard, *Solving Correspondence Problem With 1D Signal Matching*, Intelligent Robots and Computer Vision XXII, Proceedings of the SPIE, Volume 5608, pp. 207-217, (2004).
- [10]. Zhang, Z and Blum, R.S., *Image Fusion for A Digital Camera Application*, Signals, Systems & Computers. Conference Record of the Thirty-Second Asilomar, Page(s):603 – 607, (1998).

- [11]. Bloch I and Maitre H, *Data Fusion In 2D And 3D Image Processing: An Overview*, Computer Graphics and Image Processing, Proceedings of X Brazilian Symposium, (1997).
- [12]. William k. Pratt, *Digital Image Processing*, PIKS Inside, Third Edition, (2001).
- [13]. P. Chalermsart and T. El-Ghazawi, *Multi-resolution image registration using genetics*, Proceedings of International Conference on Image Processing, vol. 2, pp452-456, Japan, (1999).
- [14]. By Michael D. Vose, *The Simple Genetic Algorithm: Foundations and Theory*, MIT Press, (1999)
- [15]. Internet, <http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html>.
- [16]. Internet, <http://perso.orange.fr/polyvalens/clemens/wavelets/wavelets.html>
- [17]. Internet, <http://www.carlpisaturo.com/STR/StereoCamera.html>