# Sirindhorn International Institute of Technology

# Thammasat University


**Thesis xxxxxxxx**

---

**DEVELOPMENT OF AN EYE TRACKING SYSTEM FOR OPHTHALMIC DIAGNOSIS, TREATMENT AND SURGERY**


**Rawikarn Suppharachyothin**

# DEVELOPMENT OF AN EYE TRACKING SYSTEM FOR OPHTHALMIC DIAGNOSIS, TREATMENT AND SURGERY

A Thesis Presented

by

**Rawikarn Suppharachyothin**

Master of Engineering

Information, Computer, and Communication Program Program

Sirindhorn International Institute of Technology

Thammasat University

April 2010

# DEVELOPMENT OF AN EYE TRACKING SYSTEM FOR OPHTHALMIC DIAGNOSIS, TREATMENT AND SURGERY

A Thesis Presented

By

Rawikarn Suppharachyothin

Submitted to

Sirindhorn International Institute of Technology

Thammasat University

In partial fulfillment of the requirement for the degree of

MASTER OF ENGINEERING

Approved as to style and content by

| | |
|---|---|
| Advisor and<br>Chairperson of Thesis Committee | ________________________<br>Asst. Prof. Toshiaki Kondo, Ph.D. |
| Co-Advisor | ________________________<br>Prof. Yasuharu Koike, Ph.D. |
| Committee Member and<br>Chairperson of Examination Committee | ________________________<br>Itthisek Nilkhamhang, Ph.D. |
| Committee Member | ________________________<br>Kanokvate Tungpimolrut, Ph.D. |

External Examiner : Assoc. Prof. Chedsada Chinrungrueng, Ph.D.

April 2010

# Acknowledgement

The author would like to express her profound gratitude to her advisor, Asst. Prof. Dr. Toshiaki Kondo, whose encouragement, guidance and support from the initial to the final of this research. She is also grateful to all of the committee members, Dr. Itthisek Nilkhamhang, Dr. Yasuharu Koike, Dr. Kanokvate Tungpimolrut, and Dr. Waree Kongprawechnon, for their suggestion and information. Moreover, this research is financially supported by Thailand Advanced Institute of Science and Technology - Tokyo Institute of Technology (TAIST-Tokyo Tech), National Science and Technology Development Agency (NSTDA), Tokyo Institute of Technology (Tokyo Tech), and Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU).

# Abstract


Development of an Eye Tracking System for Ophthalmic Diagnosis, Treatment and Surgery

March 2010

by

Rawikarn Suppharachyothin

B.Eng, Sirindhorn International Institute of Technology

Thammasat University, 2007

This thesis presents the development of a real time eye tracking system that helps ophthalmologists to perform more convenient diagnosis, treatment, and surgery. This research starts with a comparative study between three real-time image processing techniques for an eye tracking system. The techniques tested in this research are a projection-data based method, the circular Hough transform, and the template matching method. The strength and weakness of each technique is revealed through computer simulations. This research also proposes a new technique named a scale-variable template matching technique for real-time human eye tracking. The proposed technique is a combination of a conventional template matching technique and the circular Hough transform. Firstly, the template matching technique is used for tracking the patients eye (the pupil and iris). The circular Hough transform is then applied to the best matching region to acquire the size of the eye. When the scale of the eye varies, the size of the template is equally varied. In this way, a scale-variable template matching technique is realized and implemented.

# Table of Contents

# List of Figures

| Figure | Title | Page |
|---|---|---|

# List of Tables

# Chapter 1

# Introduction

Structure of this thesis is start with an introduction about eye tracking and eye tracking system. Chapter 2 is describes the details of conventional eye tracking technique and the proposed technique. Chapter 3 reveal the results of each technique and discussion. Last chapter, chapter 4, is summary of this research.

## 1.1 Background

Eye tracking is the process of measuring the eye movements. An eye tracker is a device for measuring eye positions and eye movements. Eye tracking technique can be applied to many applications such as cognitive study, medical research, communication systems for disabled, improved image and video communications, and equipment for Ophthalmic Diagnosis. Nowadays, the most widely used design for eye tracking is video-based eye trackers. There are a number of methods for measuring eye movements. The most popular variant uses video images from which the eye position is extracted. This thesis is aimed to improve the stabilization of the eye tracker by using Visual C and OpenCV with the eye of the patient in operating video file which is converted to AVI format. The position of the patients pupil is propose to detect in real time by applying image processing techniques to the video sequence of eye operating process and use the positional information to control the microscope position to track the eye. This study is based on 3 algorithms, projection-data based method, Circular Hough Transform method, and template matching method. In this way, stabilized views of the moving eye are obtained and provided to ophthalmologists and their assisting nurses. The challenge of this work is the patients eye is detected harder than normal eye due to some disorder of ciliary body, lens, retina, sclera, cornea, and iris.

## 1.2 Open Source Computer Vision Library (OpenCV)

OpenCV stands for Intel Open Source Computer Vision Library. It is an open source computer vision library that including a C functions and some C++ classes that implement many popular image processing and computer vision algorithms mainly designed for computational efficiency and with a strong focus on real time application. It is free for both non-commercial, commercial use and research use under a Berkeley Software Distribution (BSD) license. The library is written in C and C++ can run under Linux, Windows and Mac OS X. There is active development on interfaces for Python, Ruby, Matlab, and Other languages.

OpenCVs application areas include

- Image Analysis

- Structural Analysis

- Motion Analysis and Object Tracking

- Object Recognition/Identification

- 3D Reconstruction

- Human-Computer Interaction (HCI)

- Mobile Robotics

One of OpenCVs goals is to provide a simple-to-use computer version infrastructure that helps people build fairly sophisticated vision applications quickly but not all video formats can be used by OpenCV. To be able to read and process the video files, they must be converted into types that OpenCV can be supported. The OpenCV library contains over 500 functions that span many areas in vision, including factory product inspection, medical imaging, security, user interface, camera calibration, stereo vision and robotics.

### 1.3 Structure of Human Eye

Human eye consists of many parts such as sense organ that control the vision, rod and cone cells in retina that control conscious light perception and vision including color differentiation and the perception of depth. The protective outer layer of the eye, sometimes referred to as the white of the eye is called the sclera and it maintains the shape of the eye. The front portion of the sclera, called the cornea, is transparent and allows light to enter the eye. The cornea is a powerful refracting surface, providing much of the eye's focusing power [1]. Attached to the sclera are six extra ocular muscles responsible for movement of the eyes. The choroid is the second layer of the eye and lies between the sclera and the retina. It contains the blood vessels that provide nourishment to the outer layers of the retina. The iris is the part of the eye that gives it color. It consists of muscular tissue that responds to surrounding light and constricting the pupil to allow more or less light into the eye. The pupil or circular opening in the center of the iris is larger or smaller depending on the brightness of the light, see Fig. 1.1.

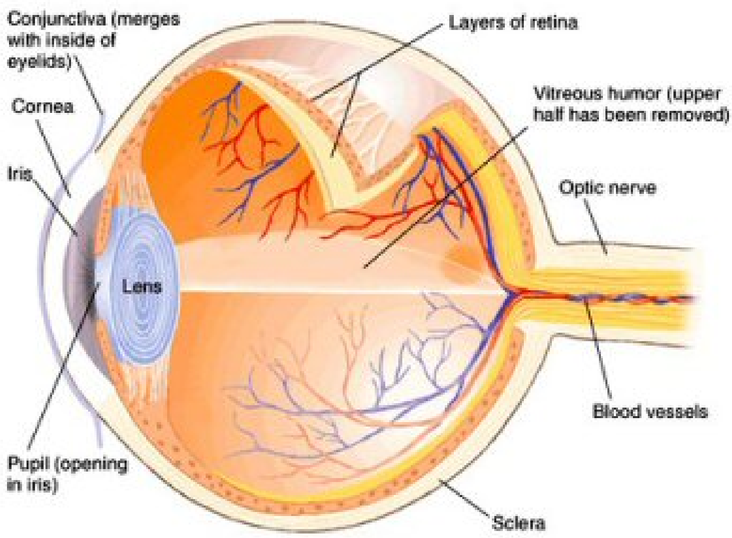



Figure 1.1 Structure of Human Eye.

The dimensions of human eye are differing only one or two millimeters among adults. The vertical measure is generally less than the horizontal distance, is about 24 mm among adults, at birth about 1617 mm. (about 0.65 inch) The eyeball grows rapidly, increasing to 22.523 mm (approximate 0.89 in) by the age of three years and the eye attains its full size when age 13 years old. The volume is 6.5 ml (0.4 cu. in.) and the weight is 7.5 g. (0.25 oz.).

## 1.4 Eye Tracking System

We are concerned with a real-time human eye tracking system for assisting ophthalmologists who perform medical operations. At the moment, the ophthalmologists we work with need to move a microscope to follow the motion of patients eye using a foot pedal while performing eye treatment and surgery. A smooth and adequate positioning of the microscope requires a highly sophisticated skill. Consequently, a long-term experience is necessary. The objective of this research is to automate the positioning of the microscope. Fig. 1.1 shows an ophthalmic operating microscope. The ophthalmologist needs to control the position of the microscope following the patients eye movement using a foot pedal while performing eye treatment and surgery. It is apparent that this manipulation requires a highly sophisticated skill. The ultimate goal of the project is to realize a fully automatic eye tracking system to assist the ophthalmologist to concentrate on surgical operations. For this purpose, we have been currently developing a real-time image processing technique to locate the position of the eye (pupil and iris). Subsequently, the detected eye position will be sent to a motor that controls the position of the microscope.

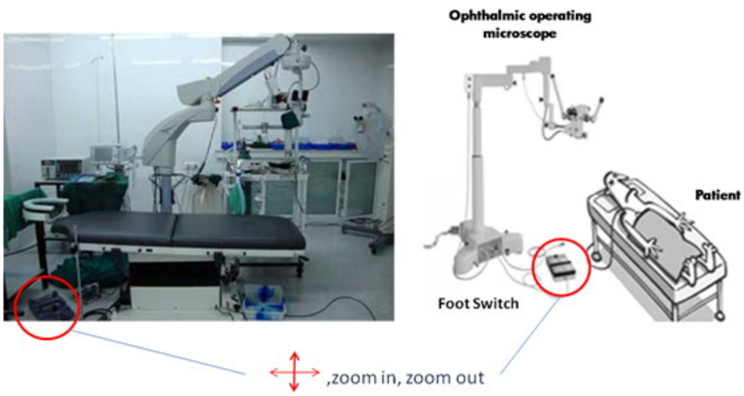



Figure 1.2 Overview system of operation equipment.

# Chapter 2

# Eye Tracking Technique

There are many existing methods for eye tracking such as eye tracking based on grey prediction [2], video-oculograph [3], the circular Hough transform (CHT) [4], ellipse fitting [5], thresholding [6] $-$ [9], and edge based technique [10], [11]. Tang and Zhang [2] use gray prediction model to predict the position of moving eye in the next frame, and use that position as the reference point for the searching region of eye. Kim et al. [3] proposed pupil/iris detection algorithm for video-oculograph -based eye movement tracking, the center of pupil is hard to detect when it has a reflection and more than half of the pupil/iris is hidden by something. Toennies et al. [4] demonstrated Hough transform-based following the assumption that an average ratio between iris size and distance is known and the distance between iris and camera is limited. T. Funahashi et al. [5] show that ellipse fitting technique can detect an oval object, but not suitable for image of low quality (low contrast, noisy, blurred) as a precise edge information is necessary. Thresholding techniques are a well suited to real-time applications, but sensitive to the change of image intensities. These approaches can be used only under a well-controlled lighting condition [6] $-$ [9] by setting the threshold value. For instance, if the intensity value of a pixel is above the specified threshold value, the intensity value in that pixels turns to 255 or white color. If the intensity value of a pixel is lower than the specified threshold value, the intensity value in that pixels turns to 0 or black color. Edge-based techniques are heavily dependent on the quality of input image sequences, and thus can be used in combination with other techniques [10], [11]. The existing image processing techniques for detecting the pupil that we studied are template matching, Hough transform, ellipse fitting, active contour model (ACM), thresholding, and edge based technique.

Template matching is a technique that uses to find small part in image by searching and match the image, see Fig. 2.1. After the function comparison, the best matches can be found as global minimums or maximums depends on each method. The computational cost of this technique is growing with the number of images [12]. There are many methods to match the template with the image such as cross-correlation, Sum-of-absolute differences matching (SAD)and Sum-of-squared differences matching (SSD):

$$R(i,j) = \sum_{k=1}^{N} \sum_{l=1}^{M} [I_1(k,l) \cdot I_2(i+k, j+l)] \tag{2.1}$$

where R store the cross-correlation coefficients, $I_1(k,l)$ is template, $I_2(i+k,j+l)$ is an input image, M and N are the i and j dimension of the image, respectively.

$$SAD(i,j) = \sum_{k=1}^{N}\sum_{l=1}^{M} |I_1(k,l) - I_2(i+k,j+l)| \tag{2.2}$$

$$SSD(i,j) = \sum_{k=1}^{N}\sum_{l=1}^{M} |I_1(k,l) - I_2(i+k,j+l)|^2 \tag{2.3}$$

where $I_1(k,l)$ is the template that must be not greater than source image (in this research, the template is image of pupil), (k,l) is coordinate of each pixel in the template, $I_2(i+k,j+l)$ is the image that we use to search (in this research, it is whole eye image),and (i,j) is the coordinate of each pixel in that image

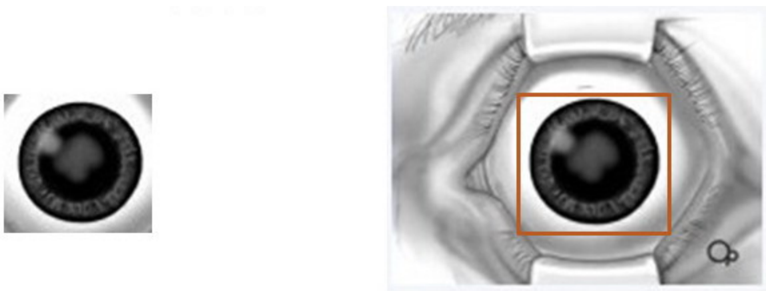


Figure 2.1 Template matching a) Template, b) image after matched with template.

Hough Transform is one of the classical approaches. Hough Transforms will make decision depending on vote in parameter space [13]. It can detect any curve as

$$f(X,C) = 0 \tag{2.4}$$

where, X is a vector of image, C is a vector of parameter in Hough transform, Equation to describe a circle is

$$(x-a)^2 + (y-b)^2 = r^2 \tag{2.5}$$

where, a, b, r is the center and radius of circles, the parameter space should be three dimension. In this parameter space A, every element is A(a,b,r), its value is the point sum of satisfied (2.4). For every edge point, A(a,b,r) is calculated as:

$$A(x, y, R) = A(x, y, R) + 1 \tag{2.6}$$

Then depending on vote the value of A(a,b,r), a circle is being decided mostly

$$A(x_0, y_0, R) = max(UA(x, y, R)) \tag{2.7}$$

The result is that the circle is one with center ($x_0$,$y_0$)and the radius R. In Hough space, the parameters of max(A) will be responding to the best circle of image space, see Fig. 2.2.

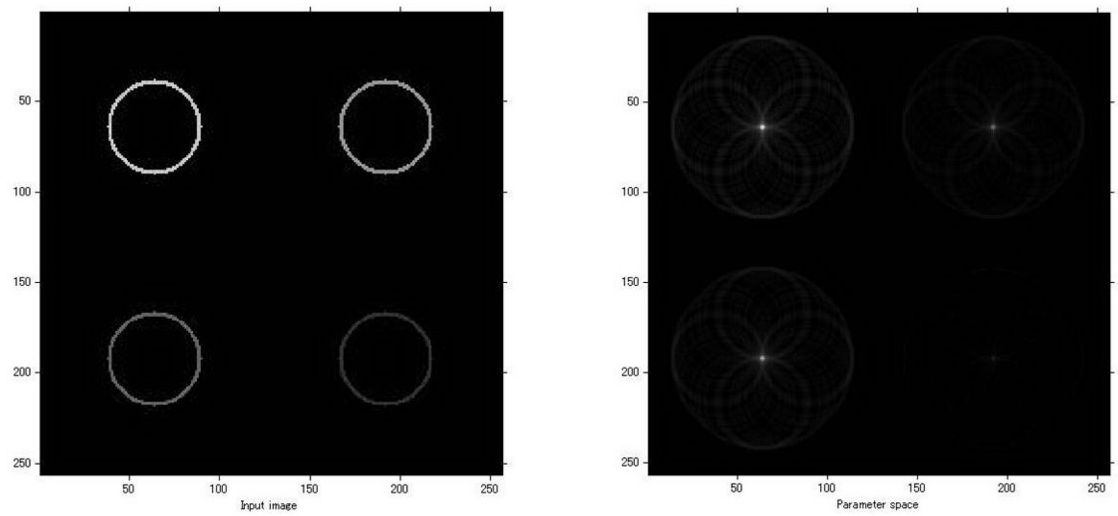



Figure 2.2 Hough transforms a) Input image, b) Parameter space.

Following is detail description of Hough transforms Algorithm:

1) Pre-processing image, get edge image by using edge extracting arithmetic operators.

2) Initialize Hough transforms matrix, let A(x,y,R)=0.

3) For every edge points, let radius r varies from $r_{min}$ to $r_{max}$, calculating every points satisfied (2.4), let A(x,y,R)=A(x,y,R)+1.

4) Make a decision of seeking for circle depending on vote from A.

Ellipse Fitting technique can detect an oval object with high accuracy rate for high-resolution image. This algorithm finds the center of the iris from the detected iris edge point (use Sobel edge operators to find the edges) and fit them with a dynamic ellipse [5]. This technique is not require the assumption of the size of an iris. The drawback of this technique is it not suitable for low-resolution, low quality, low contrast, noisy, blurred image because it need precisely edge information and high contrast between sclera and iris areas.

Active contour model (ACM) is robust to many conditions which are camera defocusing, image blurring, scales and fairly robust against occlusions changes of illumination. This method based on particle filtering and the Expectation Maximization (EM) algorithm [14]. Anyway, the robustness of the algorithm needs to maintain a set of the hypothesis and there is a trade-off between tracking accuracy and speed. The time consuming algorithm is not well suit to real-time processing application.

Thresholding technique is easy and convenient way to perform segmentation based on the basis of different intensities in the foreground and background region. This technique use to convert the gray-scale image to the binary image by setting the threshold value and the condition when value of pixels is greater than or lower than the threshold,

$$dst(x, y) = 255, if src(x, y) > threshold \tag{2.8}$$

$$dst(x, y) = 0, otherwise \tag{2.9}$$

Thresholding provides an easy and convenient way to perform segmentation based on the basis of different intensities in the foreground and background region. Due to the average area of pupil is bigger than glints and smaller than average background [8], an estimated location of the pupil is got from this technique. Thresholding is less computation time, so it is suitable for real-time application. The drawback of Thresholding is it sensitive to light changing condition, sensitive to change of image intensity. It suited for only well-controlled lighting condition.The example of thresholding technique for gray-scale image is shown in Fig. 2.3.

Figure 2.3 Thresholding a) gray-sacle image of eye, b) image after applied thresholding.

Edge Based technique is very useful to detect the edge of the pupil but it is heavily dependent on the quality of input image sequences. It is suited for only high quality image but it can combine with other techniques to get better result [10]. Edge Based technique has many methods such as Sobel, Laplacian, Canny edge method. The process of sobel filter is a 2-D filter that computes the approximation of the gradient of the image intensity at each point in both horizontal and vertical edges depending on the arrangement of the values of the masks coefficient. Typically, it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. The smallest size of sobel edge detector is a pair of 3x3 convolution masks, one estimating the gradient in the x-direction (columns) and the other estimating the gradient in the y-direction (rows). The size of convolution mask need to smaller than the size of actual image. As a result, the mask is slid over the image, manipulating a square of pixels at a time. The sobel masks size 3x3 are shown in Fig.2.4 and the example of sobel edge detection is shown in Fig.2.5. It also give the direction of the largest possible increase from light to dark and the rate of change in that direction.

| | | |
|---|---|---|
| -1 | 0 | +1 |
| -2 | 0 | +2 |
| -1 | 0 | +1 |

Gx

| | | |
|---|---|---|
| +1 | +2 | +1 |
| 0 | 0 | 0 |
| -1 | -2 | -1 |

Gy

Figure 2.4 Sobel mask size 3x3, Gx for gradient x direction, Gy for gradient y direction.

The process of laplacian filter is a second derivative function that designed to measure the changes in intensity. Canny edge detection is first-derivative based edge detection method, the example of canny edge detection is shown in Fig.2.6. Process of Canny edge detection is

-Smooth an image with a circular 2-D Gaussian filter, the larger the width of the Gaussian mask, the lower is the detector's sensitivity to noise

-Compute the gradient magnitude and direction, follow by apply Nonmaxima suppression to the gradient magnitude to trace along the edge in the edge direction

-Suppress any pixel value that is not considered to be an edge

-Apply double thresholding and connectivity analysis

In this research, many techniques that use to detect the pupil are combined together to get better result for real-time input.

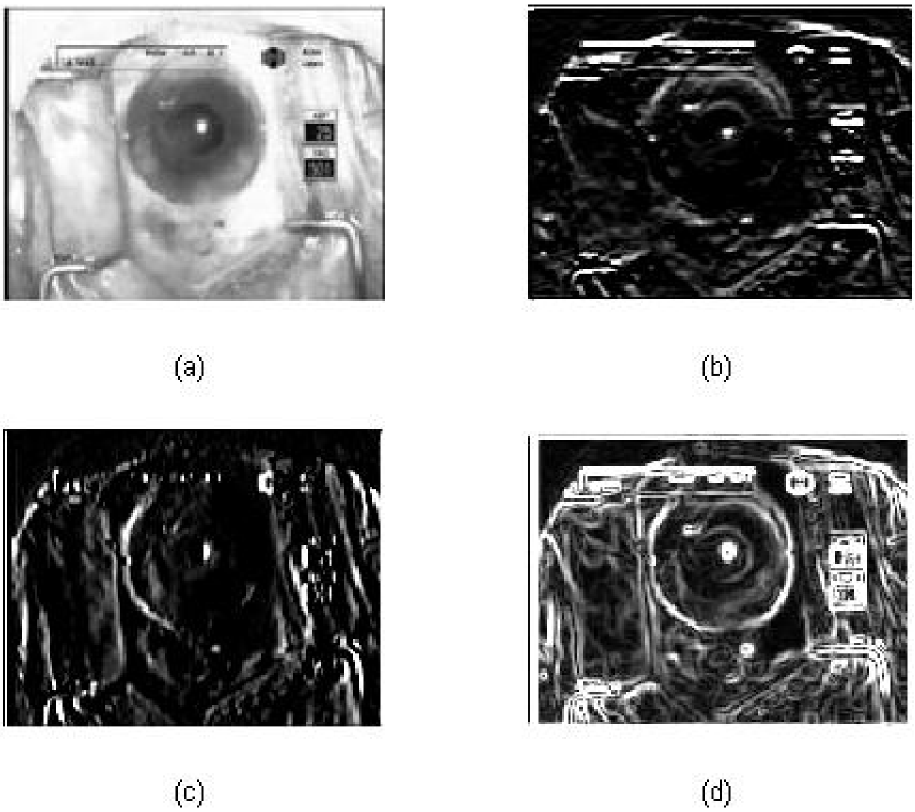



Figure 2.5 Result of sobel edge detection a) original gray-scale image, b) image after applied sobel operators in horizontal direction, c)image after applied sobel operators in vertical direction, d) image after combine edge result from both direction

## 2.1 Conventional Eye Tracking Techniques

This section presents a comparison between three real-time image processing techniques for an eye tracking system. The techniques tested in this research are a projection-data based method, the circular Hough transform (CHT), and the template matching method (TM). We reveal the strength and weakness of each technique. The position of the eye and pupil may be detected using both vertical and horizontal projections of image intensities in the projection-data based method. The computation of the method is simple and works well when the contrast of the pupil/iris is high. The CHT is based on the boundary edges of the pupil/iris. Thus, the CHT is well-suited to sharp images. Both the projection-data based and CHT methods can be applied to scale-varying images. In the TM, a target region to track is firstly cropped from a reference frame in an image sequence as a template. Thus, the technique can handle various types of eyes of different colors and shapes. Under the constant scaling condition, the TM outperforms the other two methods.

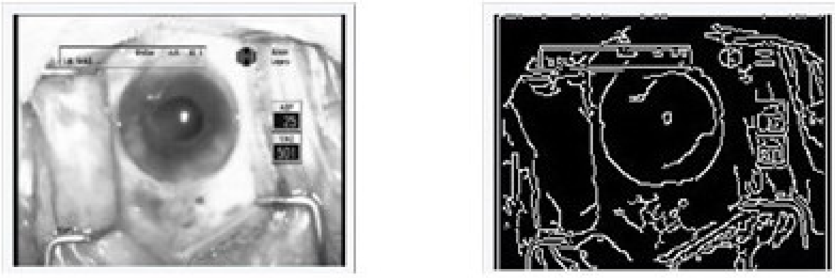
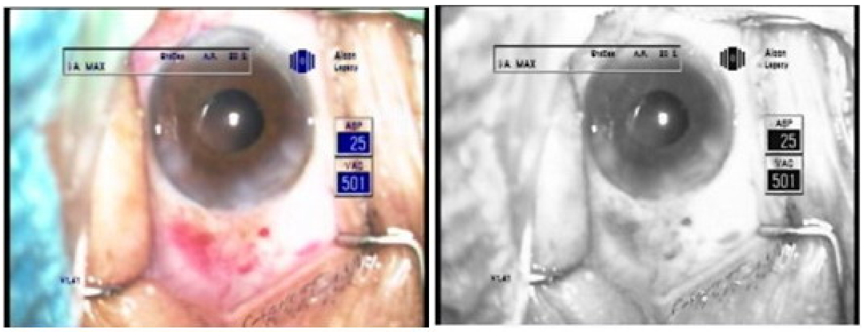


Figure 2.6 Result of canny edge detection a) original gray-scale image, b) image after applied canny operators

### 2.1.1 Projection-data Based

The projection data based method locates the center of the eye using the average of both vertical and horizontal projections of image intensities, sum the intensity value of each row and column, then divided by the number of pixels in each row and column respectively [15]. A color input image is converted to a gray-scale image first, See Fig. 2.7.

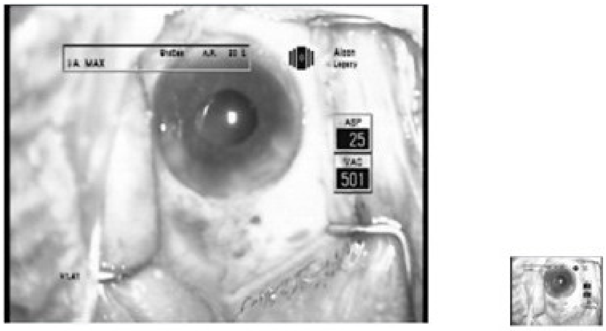



(a) (b)

Figure 2.7 Gray-scale transformation (a) An input image, (b) a gray-scale image of (a).

The image is then down-sampled to the one fifth of the original image, see Fig. 2.8. This step is effective for removing small unimportant details such as bright reflection spots as well as reducing the computational cost. With this spot, the algorithm will have difficulty to find the center of iris. Due to the size of the iris is large, when the image size is reduced, it will not be affected to the process to find the center of the iris.

Next, we apply Median filtering next to further remove noise. Median filtering is use to perform high degree of noise reduction in an image. The median filter considers the



(a) (b)

Figure 2.8 Downsampling (a) A gray-scale image of eye, (b) the one fifth of image (a).

nearby neighbors pixel and calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value, see Fig. 2.9. After applied median filtering, the small bright spots are completely removed, see Fig. 2.10.

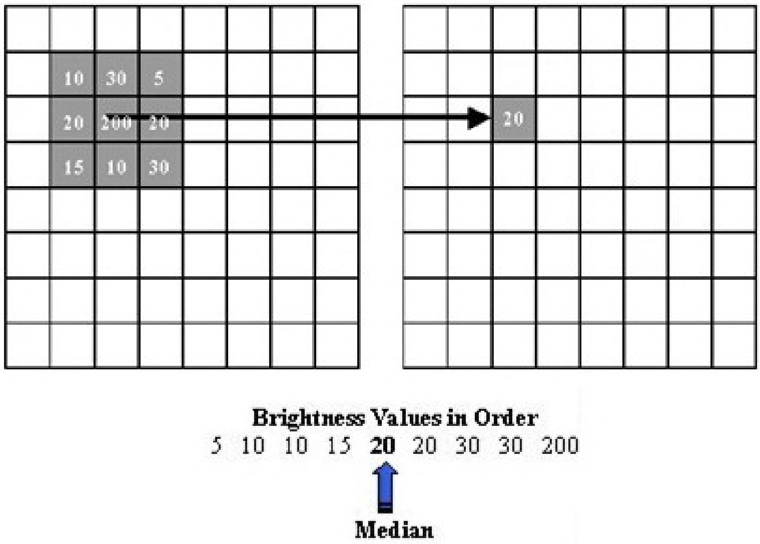
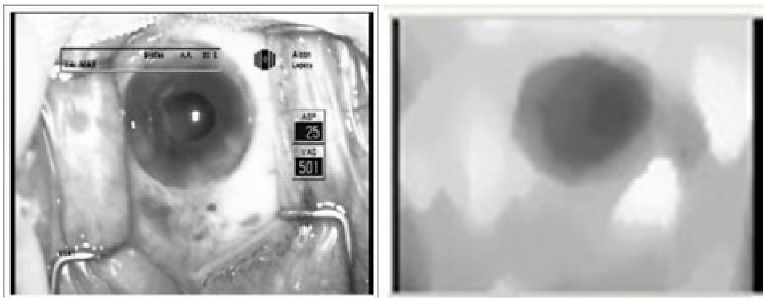



Figure 2.9 Calculating the median value of pixel neighborhood.

Then vertical and horizontal projections of image intensities are computed. Vertical projections of image intensities is computed by measure intensity of every pixel along the column, sum up all intensity in the column, and find the average value in each column. For horizontal projections, we applied similar calculation to every row of the image, see Fig. 2.11. Since the eye region of interest (pupil/iris) is darker than its surrounding tissues, those projections data always show valleys that correspond to the eye, see Fig. 2.12. In this



(a) (b)

Figure 2.10 (a) eye image in gray-scale, (b) eye image after applied Median Filtering.

way, the detection of the eye is reduced from 2-D search to the valley detection in two 1-D projection data graphs in this scheme. The process of the projection-data based method is summarized in the flow chart in Fig. 2.13.

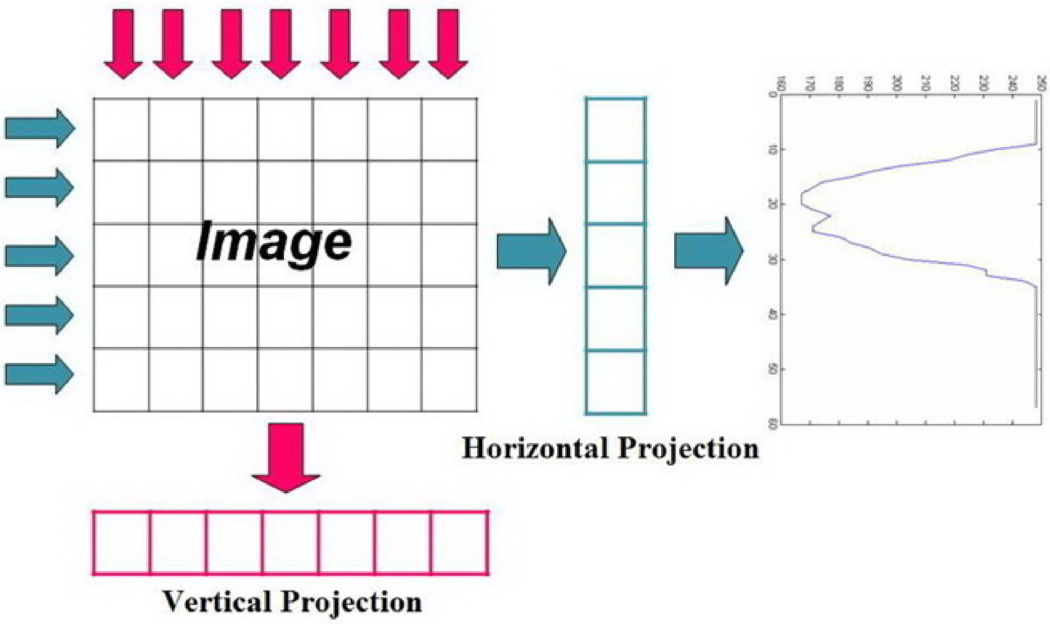



Figure 2.11 Horizontal and vertical projections.

### 2.1.2 Circular Hough Transform

The process of the CHT is based on the boundary edges of the pupil/iris region in the image. Thus the method starts with edge detection. The CHT is then applied to the detected

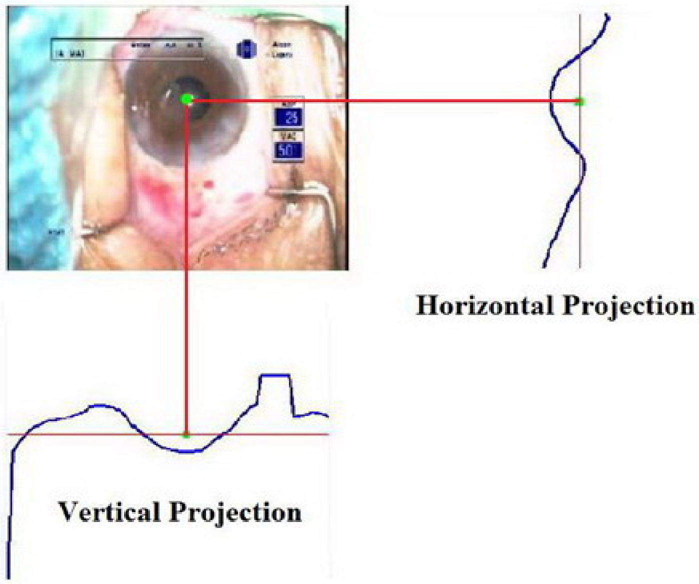



Figure 2.12 Graph of horizontal and vertical projections of eye image.

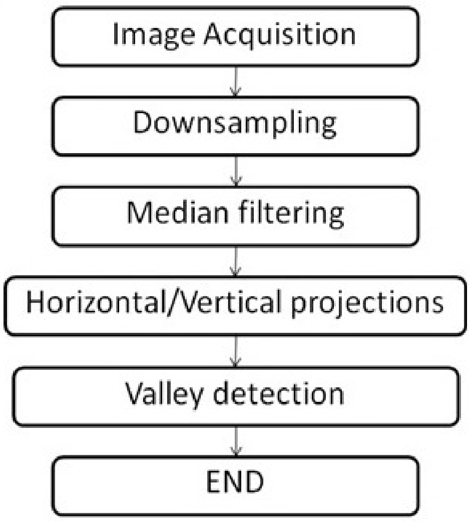



Figure 2.13 Steps of the projection-data based method.

circumference. The procedure of CHT is shown in Fig. 2.14.

First step, downsampling is used to get a sharper image by compressing the resolution to 50% of the original image. The size of the pupil/iris is large, even if the image resolution is reduced, the diameter (or radius) of the pupil/iris can be determined reliably, see Fig. 2.15.

In addition, downsampling helps to speed up the eye detection step. Edge detection heavily depends on the quality of input images. Thus, it is suited for only high quality image but it can combine with other techniques to get better results [10]. In this study, we extract

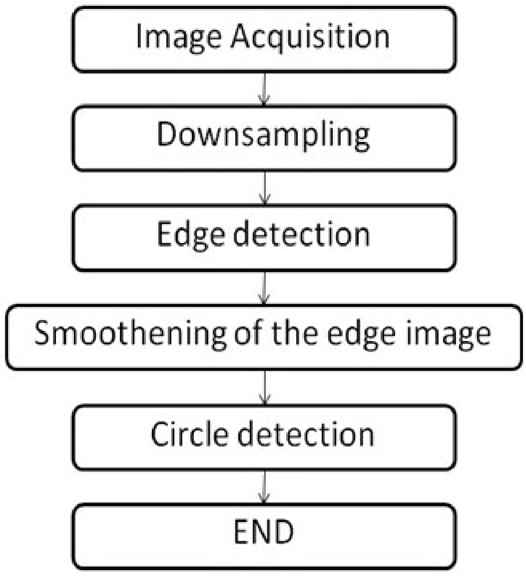



Figure 2.14 Steps of the circular Hough transform method.

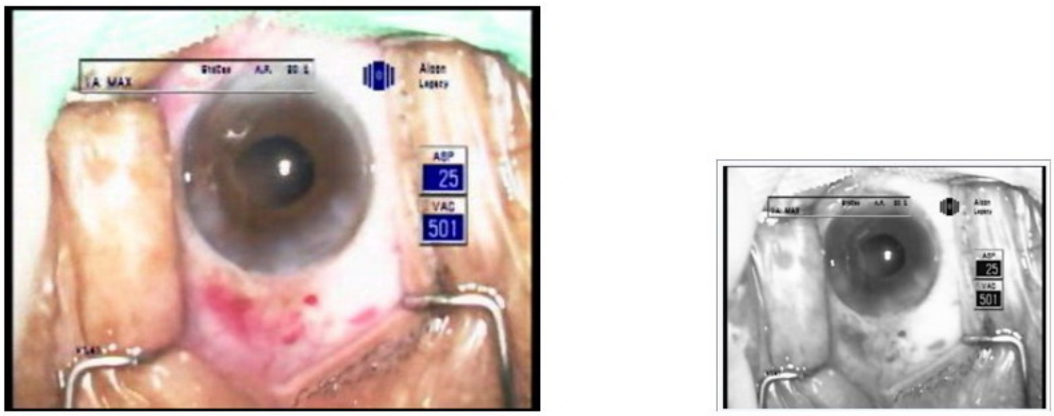
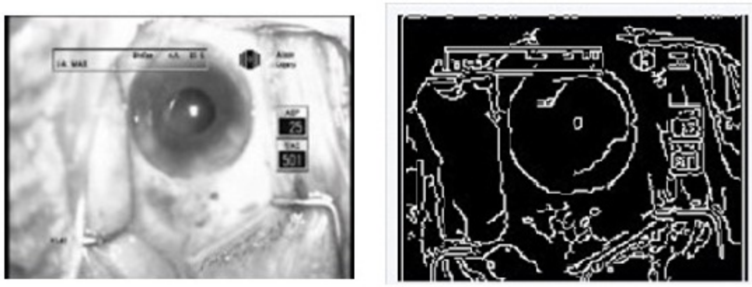



Figure 2.15 Compressing the resolution to 50% a)Original image, b)Image after downsampling.

edge information of the images by using the Canny edge detector, see Fig. 2.16. Canny edge detector is a first-derivative based edge detection method [16]. The procedure comprises smoothening an image with a circular 2-D Gaussian filter, computation of the gradient magnitude and direction, nonmaxima suppression, double thresholding, and connectivity analysis. Then, the edge image needs to be smoothened to remove noise and also widen the edges. A large mask can reduce more noise, but if it is too large, the small edge information will be lost, see Fig. 2.17. In this research, the image is convolved with a Gaussian mask of size 5x5 pixels, see Fig. 2.18.



(a) (b)

Figure 2.16 Extract edge information a) A gray scale image, b) Edge information of a).

Figure 2.17 Image that convolved with too large mask.

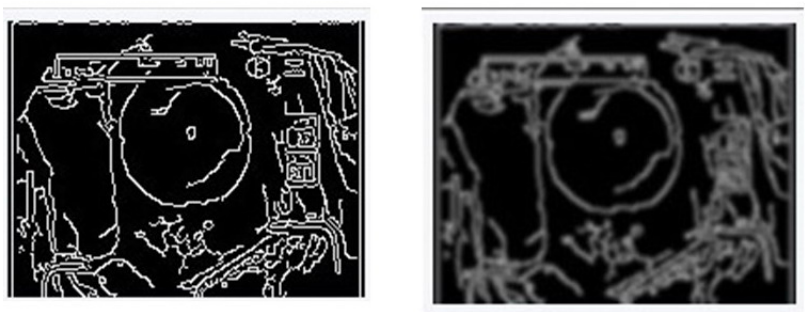


Figure 2.18 Smoothening the image a)Edge image, b)Image after convolved with a Gaussian mask of size 5x5 pixels.

Lastly, the circumference and the center of the pupil/iris can be obtained by using the CHT technique. It finds the best position through voting in the parameter space [13]. In our approach, the minimum and maximum radii of the circle are set to reduce the chance of having false circle detection.

### 2.1.3 Template Matching

Template matching (TM) is widely used in many applications of image and signal processing. It is a technique for determining where the image contains a particular object or region in which we are interested in [17]. There are many methods to match the template with the image such as sum of absolute differences and cross correlation method. In this research, the cross-correlation method is used. It finds the best matching position by determining the global maximum after vector product is performed as in (2.1). The template used in this research is defined in a reference frame by a user manually. Traditional TM methods employ a fixed size of template.

## 2.2 Proposed Technique

The proposed technique, called a scale-variable template matching technique for real-time human eye tracking, is a combination of a conventional template matching technique and the circular Hough transform. Firstly, the template matching technique is used for tracking the patients eye (the pupil and iris). The circular Hough transform is then applied to the best matching region to acquire the size of the eye. When the scale of the eye varies, the size of the template is equally varied. In this way, a scale-variable template matching technique is realized and implemented. Results show that the proposed technique can track a scale-changing eye successfully.

Traditional template matching methods employ a fixed size of template. However, in our application, the ophthalmologists move the micro scope up and down to obtain an appropriate visibility of the eye during eye surgery. As a result, the scale of the eye varies, which causes a problem for template matching. Therefore, we monitor the size of the eye every frame and if the scale of the eye is changed, we adjust the size of the template adaptively. In real operation, the ophthalmic operating camera needs to zoom. When the camera zooms in and out, the size of pupil/iris in the video is bigger and smaller respectively. Therefore, the scale-variable template matching method is constructed.

Scale-variable template matching is based on template matching and the size of template can vary according to pupil/iris diameter. It consists of 3 main steps, see Fig. 2.19, first step is find the diameter of pupil/iris by using CHT. The process of the CHT is used to locate the pupil/iris region of the image. This method starts after template matching in the first frame is finished. Then, CHT is applied to the area where the template matchings result is. This method starts from extract each frames from the video file and then downsampling it follows by detecting the edge and smoothing the image. After that, CHT is applied to detect the circumference and the center of the pupil/iris. From this point, the diameter of the pupil/iris is used for determining and resizing the template for current frame. Second step is resize the template according to the size of pupil/iriss diameter and set the limit follow (2.10) to stabilize the view and to avoid track extremely large change

$$0.7 \times D_O < D_N < 1.3 \times D_O \tag{2.10}$$

where $D_O$ and $D_N$ is previous and new diameter respectively. If value of new diameter exceeds this range, it becomes the value of diameter of previous frame. Third step, after the template is resized, we do template matching by using cross correlation method to find the best matching position using (2.10).

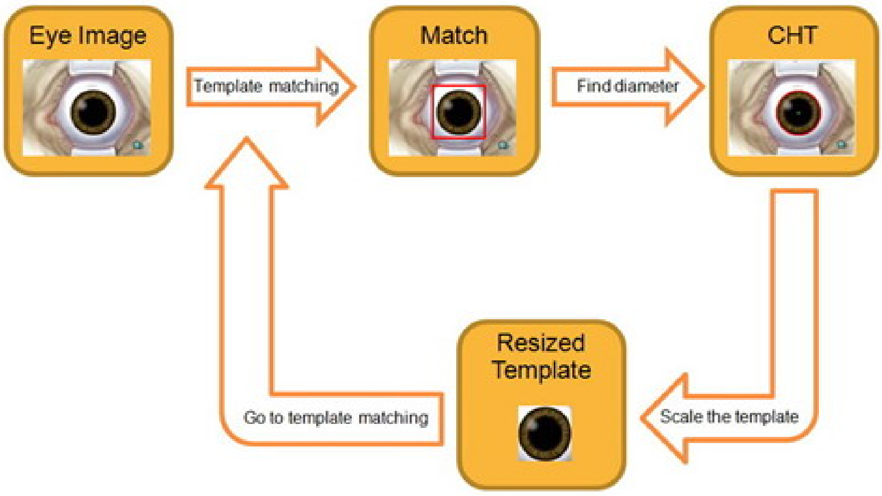



Figure 2.19 Process of scale-variable template matching.

# Chapter 3

# Results and Discussion

## 3.1 Comparative Study of Eye Tracking Technique

The results of the projection-data based, the circular Hough transform (CHT), and template matching technique (TM) are shown in Figs. 3.1 $-$ 3.3, respectively. We evaluate experimental results by means of the ratio of the estimation error and the diameter of the pupil/iris as shown in (3.1). The estimation error is a deviation between the manually selected center of the pupil/iris ($C_{tx}$,$C_{ty}$)and that of the estimated by the proposed method ($C_{px}$,$C_{py}$). The video sequence used in this experiment contains 676 frames. The size of each input image is 320 $\times$ 240 pixels. The template is cropped only the area of the pupil/iris in square shape. In order to simplify the computation, we randomly select 100 frames out of 676 frames to evaluate the result.

$$Error(\%) = \frac{\sqrt{(C_{tx} - C_{px})^2 + (C_{ty} - C_{py})^2}}{PupilDiameter} \times 100 \tag{3.1}$$

where Ct and Cp is center of template and center of pupil respectively.

### 3.1.1 Projection-data Based

The projection-data based method can detect the position in every frame but it has the lowest accuracy among these three methods. It also obtains the least success detection. The limitation of this method is its heavy dependence on the quality of input image sequences. In general, the applicable cases of this approach may be limited because it is based on the assumption that the pupil/iris is darker than its surrounding tissues. This may not be always the case, for example, if a patient has cataract or eyes of a light color. However, this method can handle scale-varying images. In Fig. 3.1, a diamond shaped marker denotes where the center of the pupil is, which is calculated by using the projection-data based method.

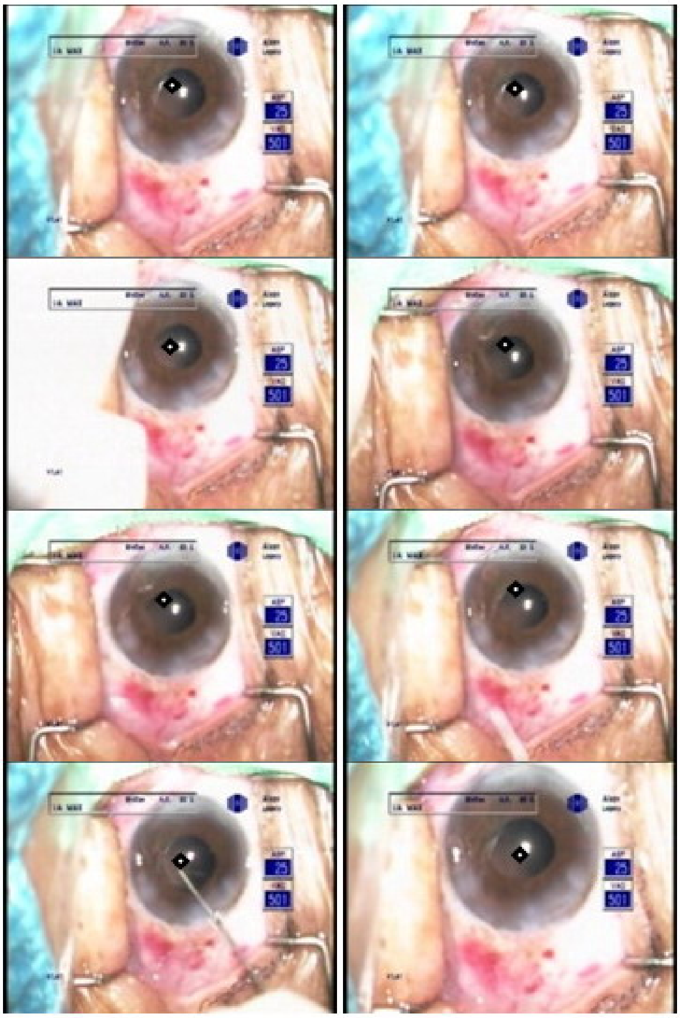


Figure 3.1 Results of the projection-data based method.

### 3.1.2 Circular Hough Transform

The circular Hough Transform (CHT) seems to work well in most frames. However, when input images are not very sharp, the CHT fails to detect a circle. In Fig. 3.2, a dark point indicates the center of the pupil/iris detected by the CHT. The CHT is dependent on the sharpness of input image sequences because the technique is dependent on edge information. When an input image is blurred, the threshold values for edge detection need to be lowered adaptively, which is not always easy to perform. An advantage of the CHT is that it can also handle scale-varying images.

### 3.1.3 Template Matching

For template matching (TM), the template image cropped out from the pupil in the reference frame of the video sequences. If the scale of the image is does not vary much, this

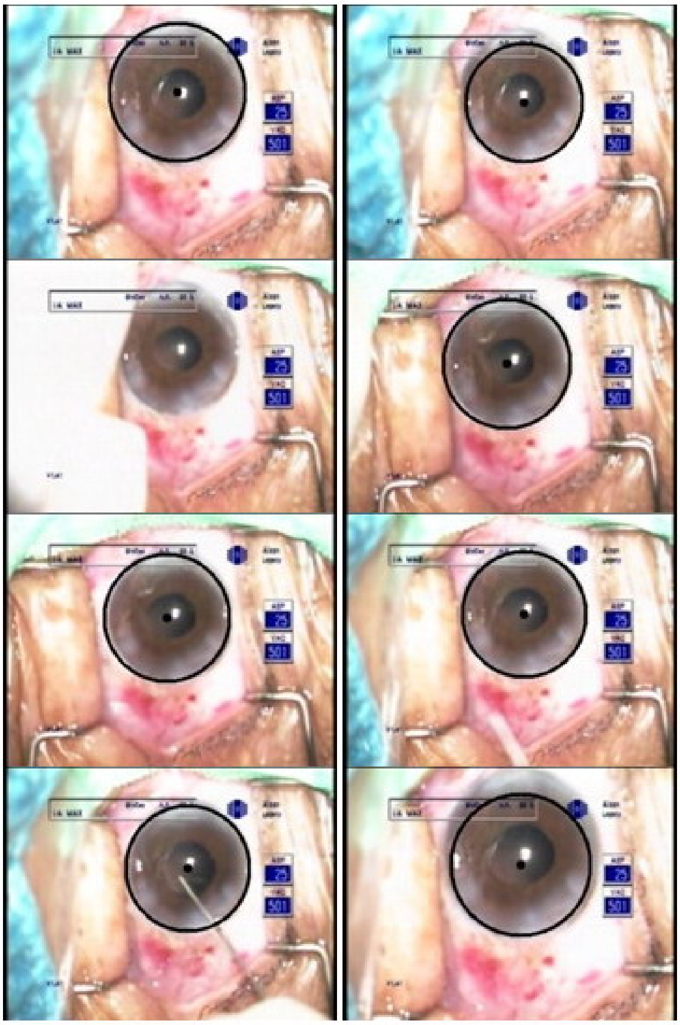


Figure 3.2 Results of the circular Hough transform method.

method works best. Fig. 3.3 shows the best matching position determined by the TM.

The threshold of the error in each frame is set to determine the accuracy. If the error is less than 10%, it will consider as successful detection. The success rates and mean error of the projection-data based, the CHT, and the TM method is shown in Table 3.3.

The simulation results show that the template matching technique is the best among the three, whereas the projection-data based method is the worst. The projection-data based method works well when the pupil/iris region is much darker than its surrounding tissues. This implies that the technique is weak if the eye region is not dark enough, for example, with cataract. In addition, the technique will not be suitable for the patients with eyes of light colors. The circular Hough transform (CHT) works well when input images are on-focus and sharp because the method is based on edge information. This means that the CHT will not suitable for the eyes with some anomaly, such as cataract, that obscures the eye region. Finally, the TM technique works the best in all circumstances, excluding variations

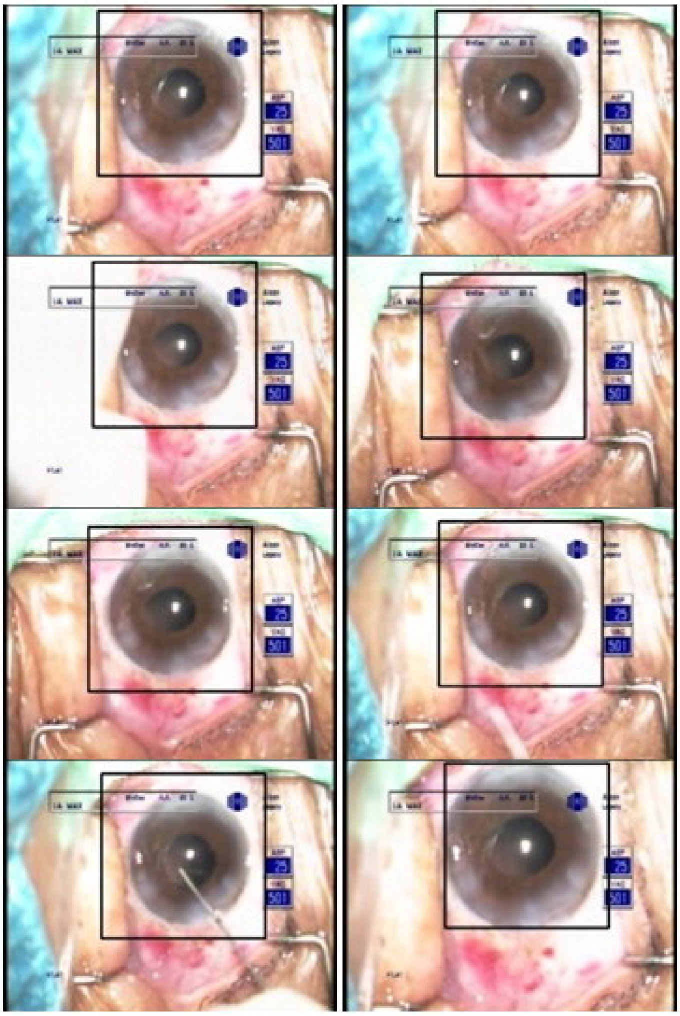


Figure 3.3 Results of the template matching method.

in scale. Since our objective is to assist the ophthalmologist to perform surgery on the eye with some anomaly. The TM should be the most suitable technique for the purpose. One limitation of the TM, however, is that the technique is sensitive to variations in scale.

## 3.2 Scale-variable Template Matching for Eye Tracking Technique

For circular Hough transform, the circles in Fig. 3.5 are circumferences of pupiliris and the point is the center of that circle. The CHT cannot detect every frame due to changing lighting conditions and image quality. If an image is too blurred, a fixed threshold will fail to detect edges that are necessary for the CHT. Therefore, some frames cannot find the suitable circular edge. For template matching, pupil/iris area cropped from the first frame is used as the template for this video sequence, shown in Fig. 3.4. In this method, the size of template is fixed. In case the size of pupil/iris from video sequences is varied too much, the result will cause an error. Most existing object recognition systems suffer from poor scaling with

Table 3.1 Mean error and success rates of each method (%).

| Methods | Mean Error | Success Rates |
|---|---|---|
| Projection-data based | 9.62% | 62% |
| Circular Hough Transform | 6.94% | 72% |
| Template Matching | 4.12% | 98% |

the number of recognizable objects [18]. For scale-variable template matching, the square is also the best matching position and the size of template can vary according to the pupil/iris diameter which can be derived from the CHT. The result is shown in Fig. 3.5.

Figure 3.4 Pupil and iris are used as a template.

We evaluate experimental results by means of the ratio of the estimation error and the diameter of the pupil/iris as shown in (3.1). The video sequence used in this experiment is in the length of 500 frames. The size of each input image is 320 $\times$ 240 pixels. The template is cropped only the area of pupil/iris in square shape. In order to simplify the computation, we randomly select 200 frames out of 500 frames in a video sequence to evaluate the result. The threshold of the error in each frame is set to determine the accuracy. If the error is less than 10%, it will consider as successful detection. The success rates, mean error and standard deviation of the error is shown in Table 3.2. In addition, we simulate normal Template Matching and scale-variable Template Matching on the same video file to compare the accuracy. The additional result of Template Matching method is shown in Fig. 3.6 and the comparison of the accuracy of these two methods are shown in Table 3.3.

Table 3.2 Mean error and standard deviation (%).

| | |
|---|---|
| Success Rates | 83% |
| Mean Error | 8.28% |
| Standard Deviation | 9.37% |

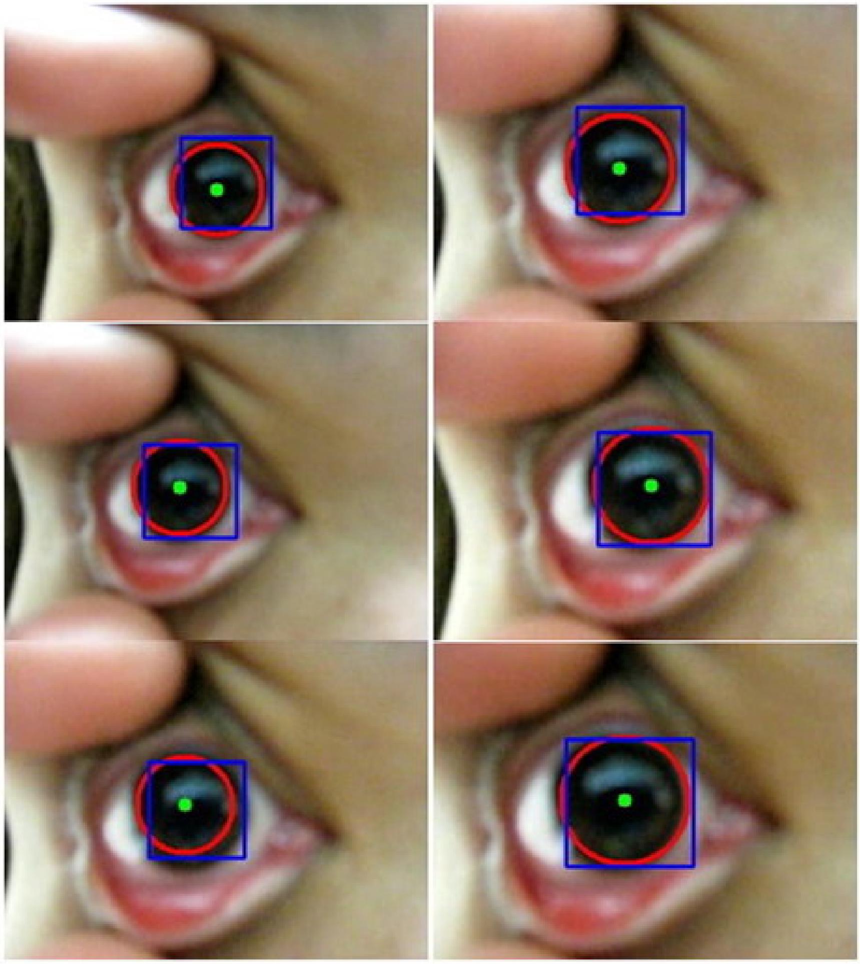


Figure 3.5 Results of scale-variable template matching method.

Table 3.3 Comparison between normal Template Matching and scale-variable Template Matching (%).

| Methods | Mean Error | Success Rates |
|---|---|---|
| Template Matching | 9.88% | 79% |
| Scale-variable TM | 8.28% | 83% |

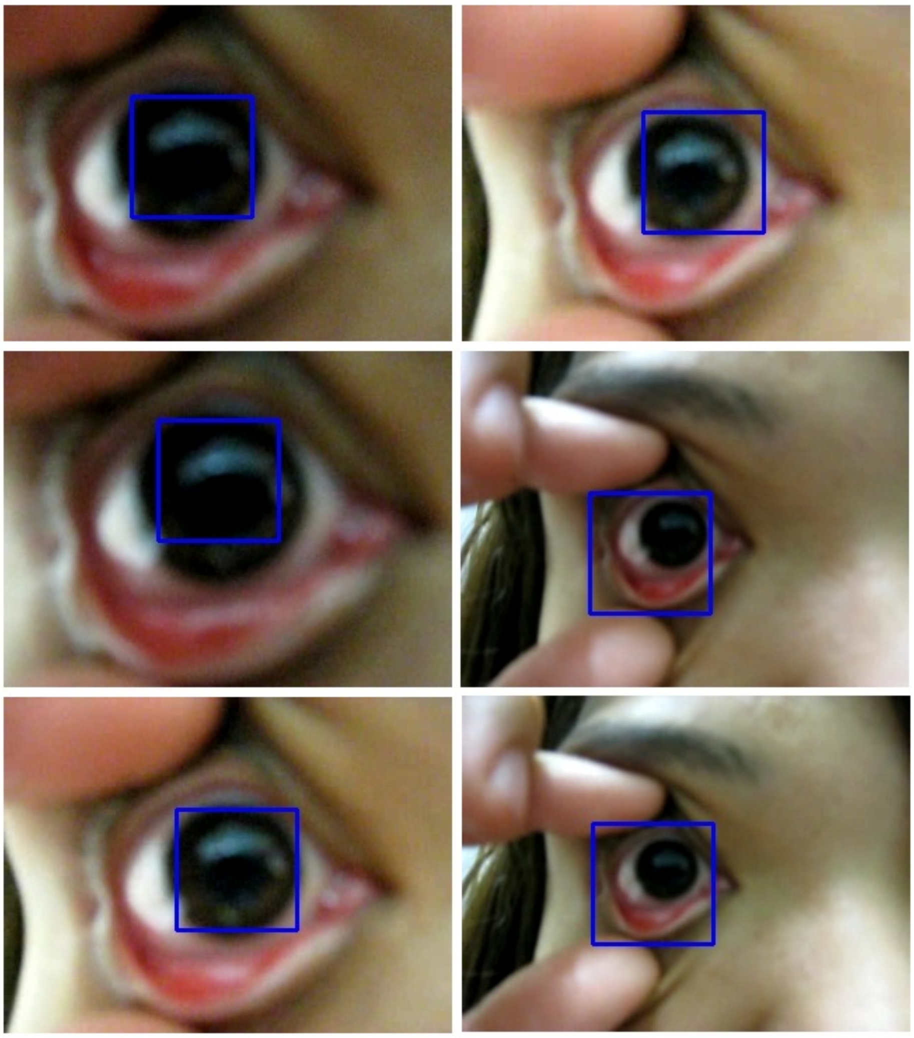


Figure 3.6 Additional results of template matching method.

# Chapter 4

# Conclusions

The study of Eye Tracking System for Ophthalmic Diagnosis, Treatment and Surgery is motivated by the inconvenience of the ophthalmologists during a surgical operation because the ophthalmologists need to control many types of equipment at the same time. Moreover, the ophthalmic operating microscope which has an eye tracking system is high cost, so the hospital that has low budget cannot buy it. To give an opportunity to those hospitals, the low cost ophthalmic operating microscope with an eye tracking system is needed. This research proposes a novel approach for eye tracking. The proposed technique is based on a template matching technique as it is well-suited to the detection of a fixed pattern form a frame to another. One limitation of the template matching technique is that it cannot detect a scale-varying object because of the fixed size of the template. For this, we introduced the CHT to obtain the size of the pupil/iris. The CHT is applied only to a limited small region determined by the proceeding template matching technique. In this way, the computational cost of the CHT is reduced, and more importantly, false circle detection may be avoided. With the aid of the CHT, we can monitor the size of the eye. Therefore, it is possible to adjust the size of the template immediately when a scale change occurs. Simulation results show a satisfactory performance when the size of the pupil/iris or iris varies.

# References


[1] Dictionary of Eye Terminology, B. Cassin and S. A. B. Solomon, 1997

[2] Jianxiong Tang and Jianxin Zhang, Eye Tracking Based on Grey Prediction, *1st International Workshop on Education Technology and Computer Science*, pp.861-864, 2009.

[3] S. I. Kim, J. M. Cho, J. Y. Jung, S. H. Kim, J. H. Lim, T. W. Nam, J. H. Kim, A Fast Center of Pupil Detection Algorithm for VOG-based Eye Movement Tracking, *IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp.3188-3191, September 2005.

[4] Klaus Toennies, Frank Behrens, and Melanie Aurnhamer, Feasibility of Hough-transform-based Iris Localisation for Real-time-application, *16th International Conference on Pattern Recognition*, pp.1053-1056, 2002.

[5] J. Zhu and J. Yang, Subpixel Eye Gaze Tracking, *5th IEEE International Conference on Automatic Face and Gesture Recognition*, 2002.

[6] JK. S. Jin, S. G. Cho, J. S. Lee, and J. J. Hwang, Real-time Pupil Detection Based on Three-step Hierarchy, *7th International Conference on Signal Processing*, 2004.

[7] R. Jain, T. Kasturi, and B.G. Schunck, *Machine Vision*, McGraw-Hill, Inc., 1995.

[8] S. Goni, J. Echeto, A. Villanueva, and R. Cabeza, Robust Algorithm for Pupil-glint Vector Detection in A Video-oculography Eyetracking System, *17th IEEE International Conference in Pattern Recognition*, 2004.

[9] A. Villanueva, R. Cabeza, and S. Porta, Eye tracking: Pupil Orientation Geometrical Modeling, *Image and Vision Computing*, pp.663-679, 2006.

[10] S. Dey and D. Samanta, An Efficient Approach for Pupil Detection in Iris Images, *15th International Conference in Advanced Computing and Communications*, Wiley/Interscience, 2007.

[11] A. Discant and S. Demea, An Algorithm for Extraction of Iris Information, *International Conference in Automation, Quality, and Testing, Robotics*, vol. 2, pp. 445-450, May 2006.

[12] Luke Cole, David Austin, Lance Cole, Visual Object Recognition using Template Matching, National ICT Australia, Locked Bag 8001, Canberra, ACT 2601.

[13] Qi-Chuan Tian, Quan Pan, Yong-Mei Cheng, and Quan-Xue Gao, Fast Algorithm and Application of Hough Transform in Iris Segmentation, *Proc. 3th Int. Conf. on Machine Learning and Cybernetics, Shanghai*, pp. 3977-3980, Aug. 2004.

[14] D. W. Hansen and A. E. C. Pece, Iris tracking with feature free contours, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1625-1630, December 2003.

[15] S. Keerativittayanun, K. Rakjaeng, T. Kondo, W. Kong-prawechnon, K. Tungpimolrut, and T. Leelasawassuk, Eye Tracking System for Ophthalmic Operating Microscope, *ICROS-SICE International Joint Conference 2009*, pp 653-656.

[16] Paul Bao, Lei Zhang, and Xiaolin Wu, Canny Edge Detection Enhancement by Scale Multiplication, *IEEE Transactions on Pattern analysis and machine intelligence*, Vol. 27, No. 9, September 2005, pp.1485-1490.

[17] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Second edition, pp 206.

[18] Ardeshir Goshtasby, Template Matching in Rotated Images, *IEEE Transactions on Pattern analysis and machine intelligence*, Vol. PAMI-7, No. 3, pp. 338-344, May 1985.

# Appendix A

# OpenCV Installation

The step to install OpenCV are:

1. Double click icon OpenCV1.0.exe to start installation process to your computer.

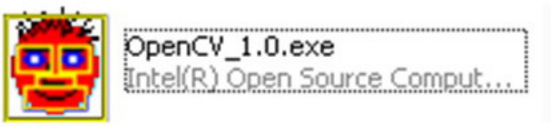



Figure A.1 Set up icon for OpenCV version 1.0.

2. The setup wizard window will appear on the screen. Then click Next button.

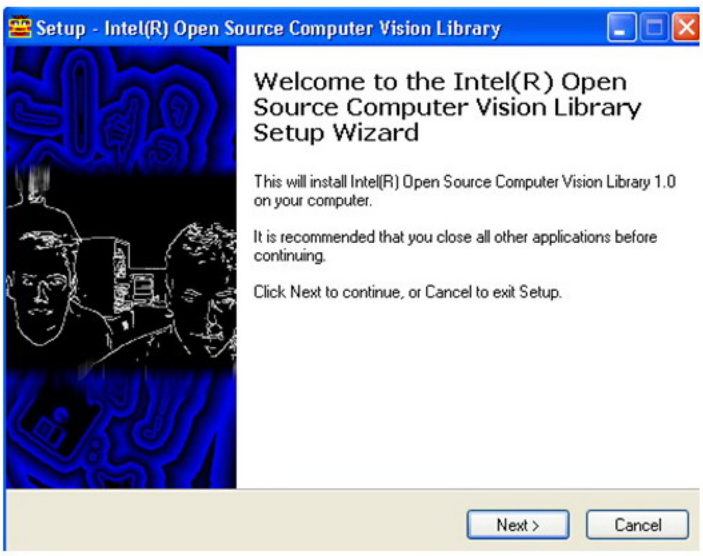



Figure A.2 Set up window (1)

3. Press Install and wait until installation process complete.

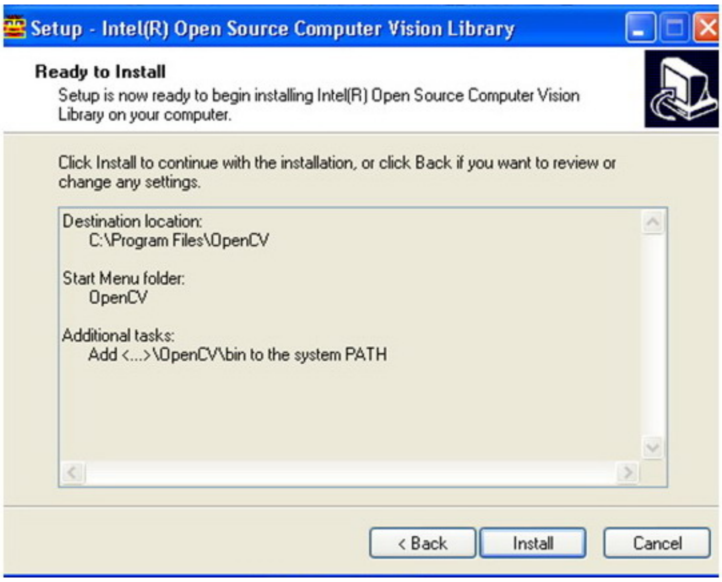



Figure A.3 Set up window (2)

## A.1 Getting Started with OpenCV by Using Microsoft Visual C++

Before using OpenCV with Microsoft Visual C++, Microsoft Visual Studio must be installed in computer. To create OpenCV-based project in MS-Visual C++ do the following steps:

### A.1.1 Customize Global Options

Perform the following steps only on the first time using OpenCV with MS Studio.Net

- Open the MS-Visual C++ Application. In the menu bar, select Tools − > Options

- In the listing, choose Projects and Solutions − > VC++ Directories.

- First, select Library files from the ”Show Directories for” List Box.

- Click the Insert New icon, and locate the folder where you have installed OpenCV.

- Consider that it is installed in ”C:/Program Files/OpenCVVS2005”.

- In the Library files list in the list box, locate and add ”C:/ Program Files/ OpenCVVS2005/ lib” , see Fig.A.4.

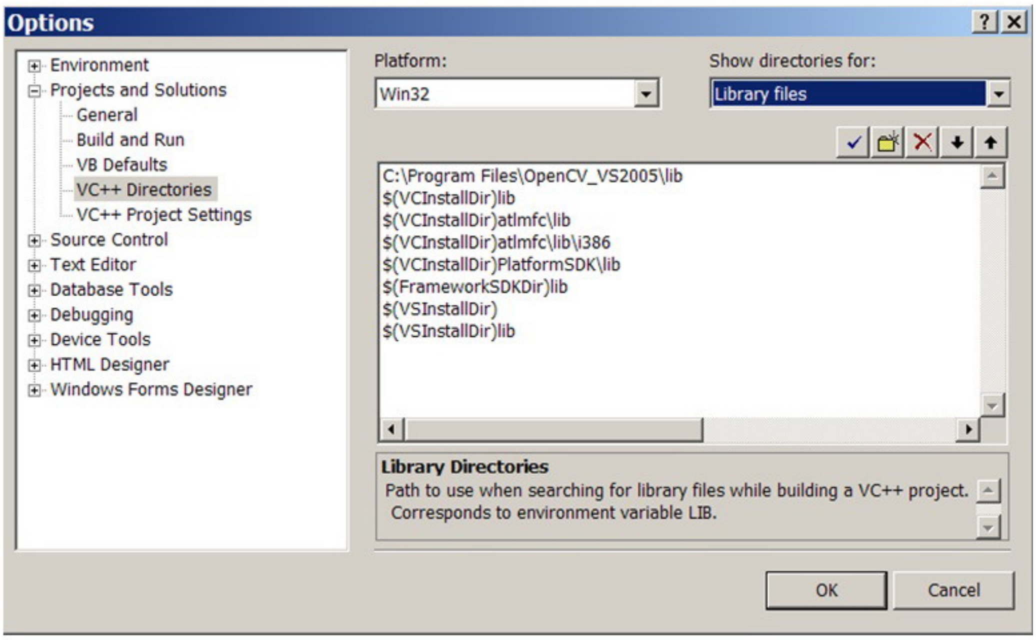



Figure A.4 Show directories for Library files

- Now choose Include files in the list box, locate and add the directories follow Fig.A.5.

- Next, choose source files in the list box, locate and add the directories follow Fig.A.6.

- Now click OK in the Options dialog. - Successfully configured the global settings.

### A.1.2 Create New Project

Perform the following steps every time when creating new project

- First, select from menu "File" − > "New..." − > "Projects". - Choose project types as Visual C++ − > Win32 and select "Win32 console application". - Create project name and location then OK, see Fig.A.7.

- Click Finish button on Win32 Application Wizard window, see Fig.A.8.

- Go to menu bar and select Project − > test1 properties, see Fig.A.9.

After the above steps done Developer Studio will create the project folder (by default it has the same name as the project), <project name>.vcproj file, Solution <project name>.sln and, Three Source files: <project name>.cpp, stdafx.cpp and stdafx.h. Stdafx files are pre-compiled header files, which can be very useful if you want to reduce the compilation time.

- Choose "Linker" − > "General" − > "Additional Library Directories", see Fig.A.10.

```
"C:\Program Files\OpenCV_VS2005\cv\include"
"C:\Program Files\OpenCV_VS2005\cxcore\include"
"C:\Program Files\OpenCV_VS2005\otherlibs\highgui"
"C:\Program Files\OpenCV_VS2005\cvaux\include"
"C:\Program Files\OpenCV_VS2005\otherlibs\cvcam\include"
```

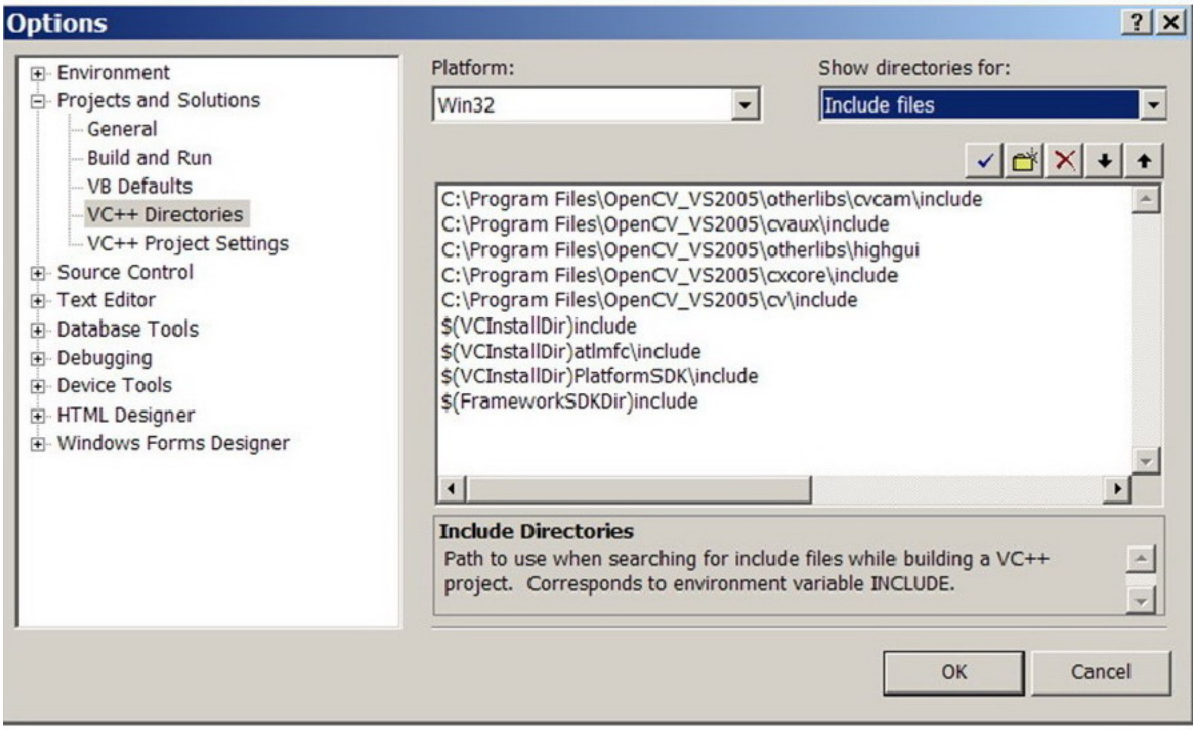



Figure A.5 Show directories for Include files

- Choose directories of OpenCV. Then click OK button, see Fig.A.11.

- Choose ”Linker” – > ”Input”– > ”Additional Dependencies. Add the paths to all necessary import libraries (cxcore.lib, cv.lib, highgui.lib), see Fig.A.12.

After created a new ”test1” Project. Open the test1.cpp file, and include the OpenCV-related include directives:

include <cv.h>

include <cxcore.h>

include <highgui.h>

- Press F7 to compile the project

```
"C:\Program Files\OpenCV_VS2005\cv\src"
"C:\Program Files\OpenCV_VS2005\cxcore\src"
"C:\Program Files\OpenCV_VS2005\cvaux\src"
"C:\Program Files\OpenCV_VS2005\otherlibs\highgui"
"C:\Program Files\OpenCV_VS2005\otherlibs\cvcam\src\windows"
```

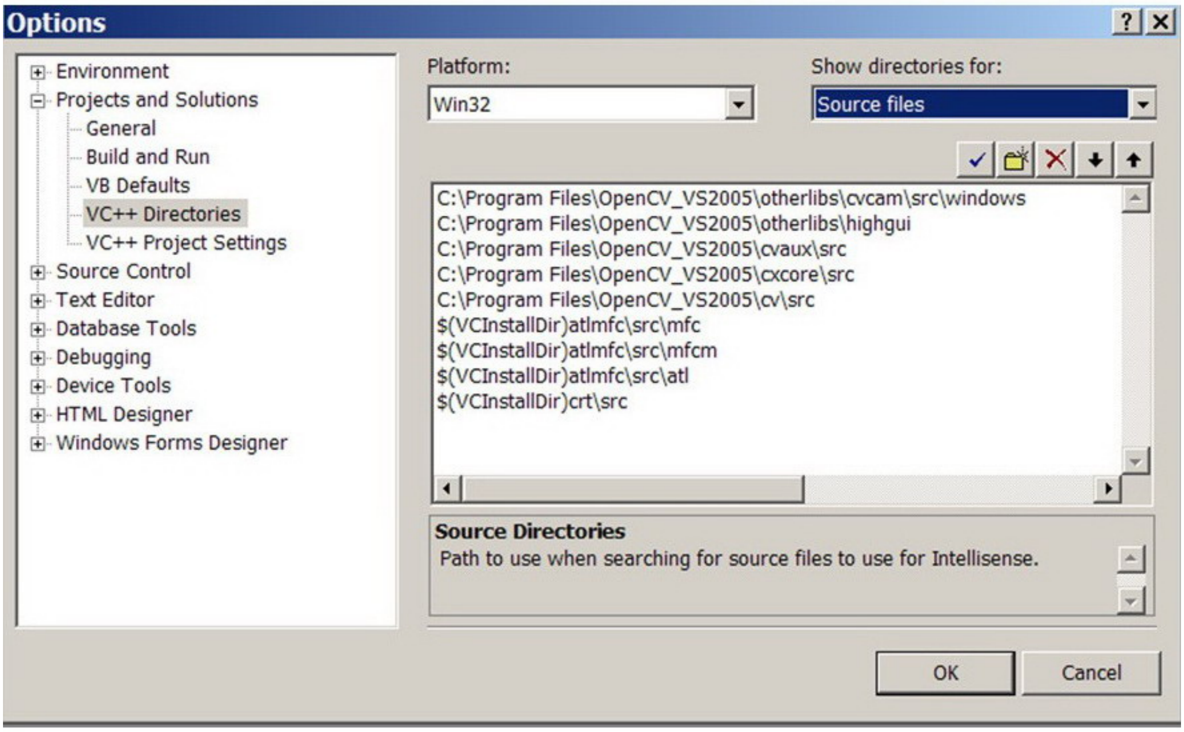



Figure A.6 Show directories for Source files

- Go to the directories C:/Program Files/OpenCVVS2005/bin and copy all .dll files that consist of 8 files:

cv100.dll

cvaux100.dll

cvcam100.dll

cxcore100.dll

cxts001.dll

highgui100.dll

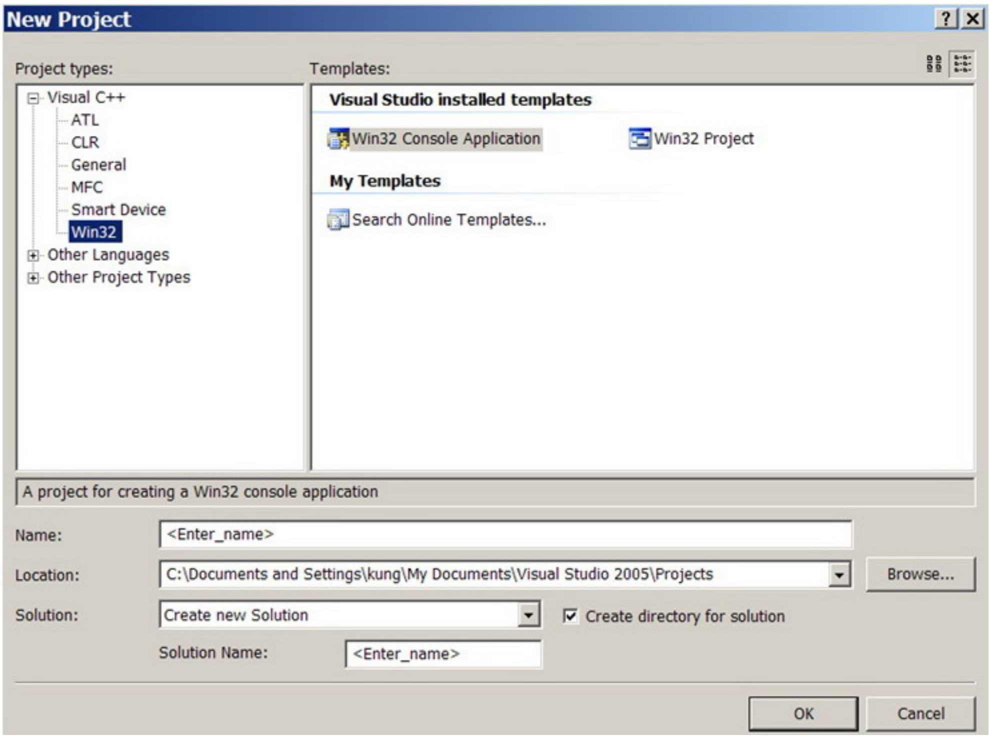



Figure A.7 New Project window

libguide40.dll

ml100.dll

- Paste all .dll files on C:/Documents and Settings/...../My Documents/Visual Studio 2005/Projects/test1/debug, see Fig.A.13.

- Successfully creating a new project.

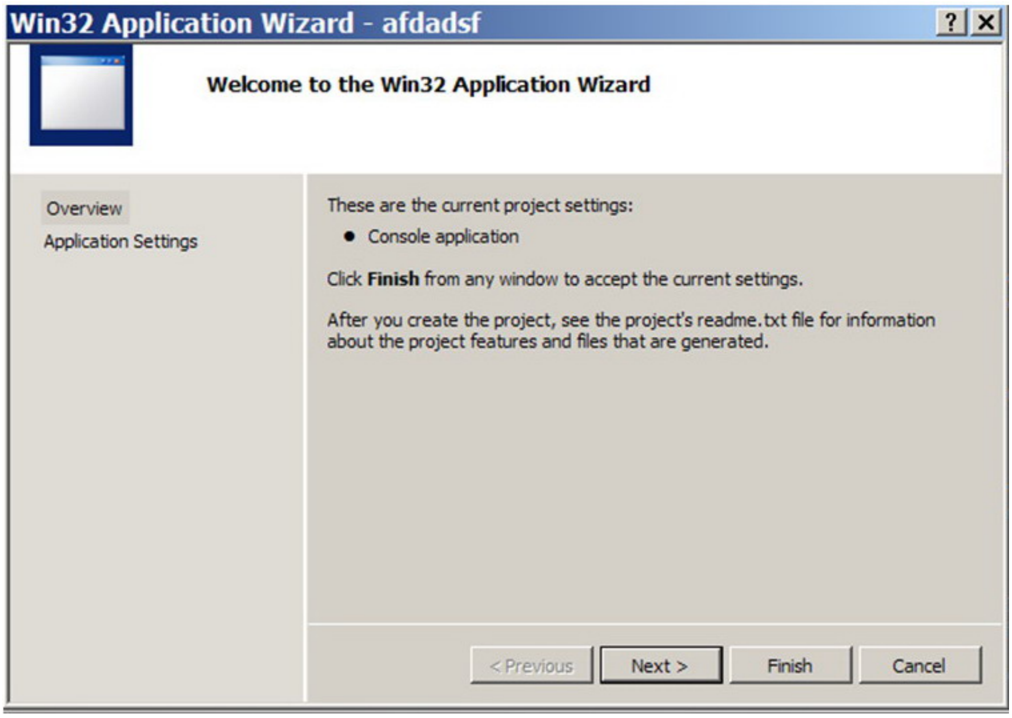



Figure A.8 Win32 Application Wizard window

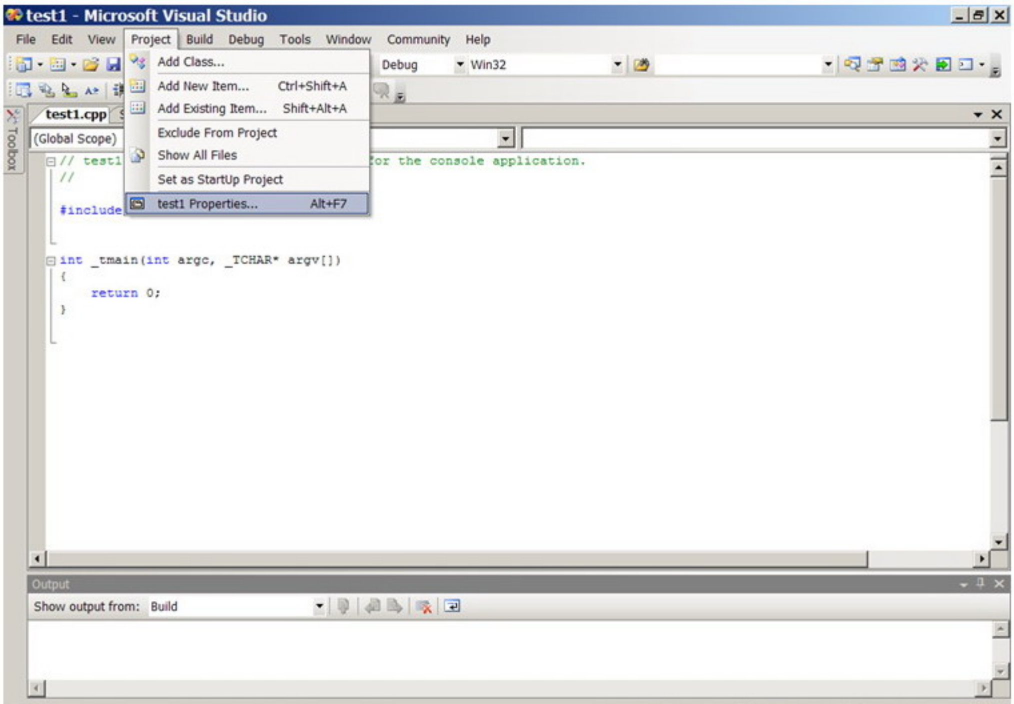



Figure A.9 test1 properties

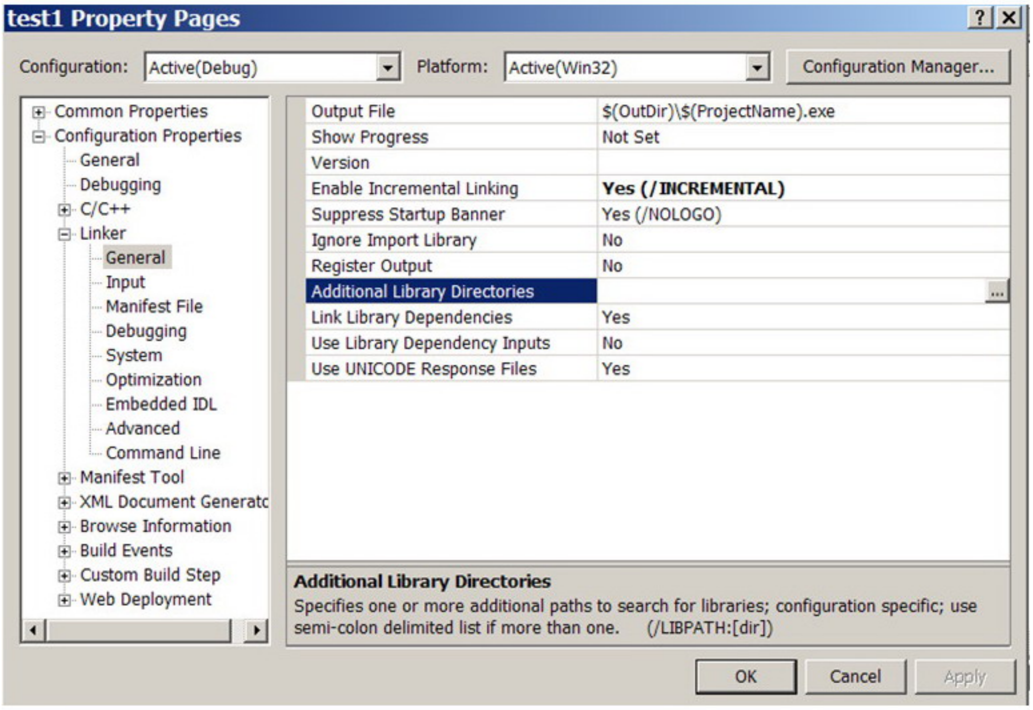



Figure A.10 Property Pages

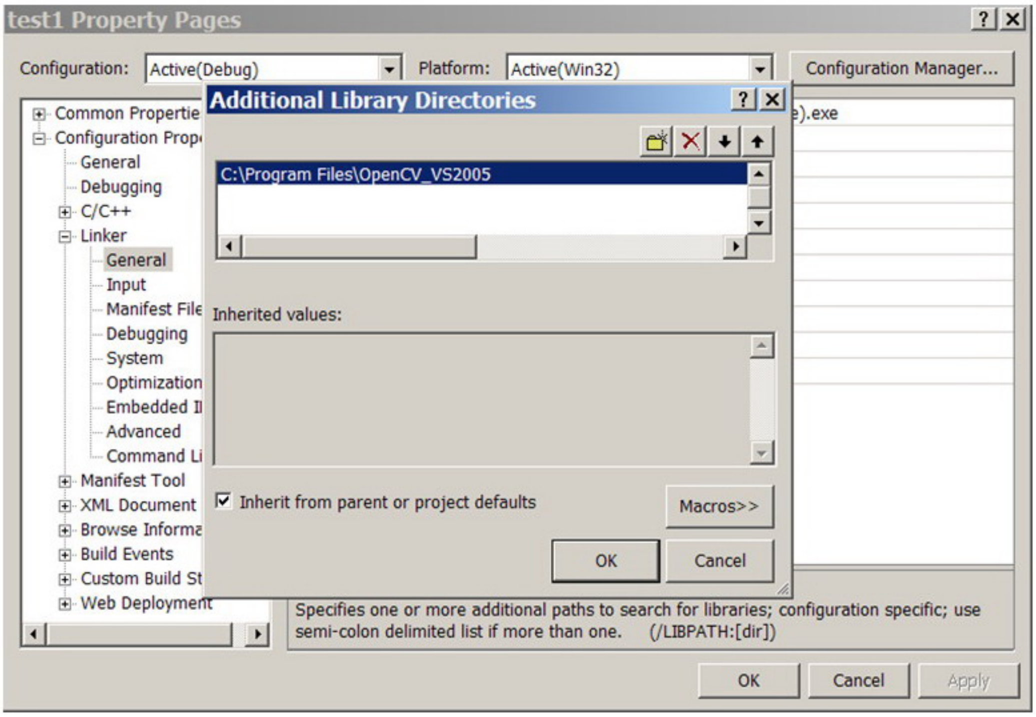



Figure A.11 Additional Library Directories

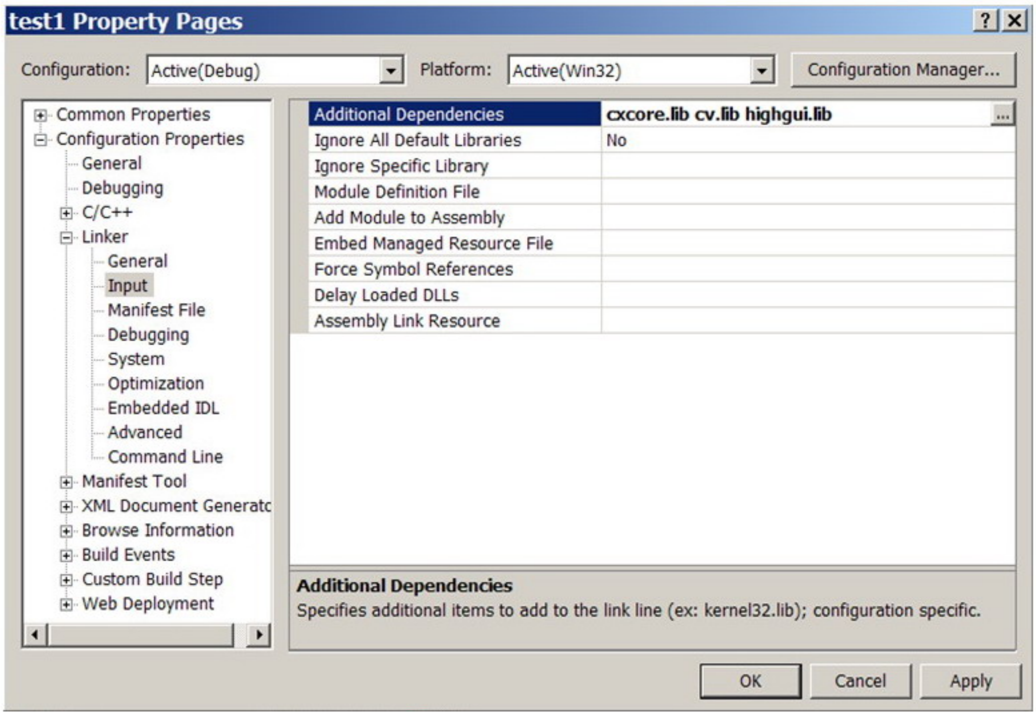



Figure A.12 Additional Dependencies

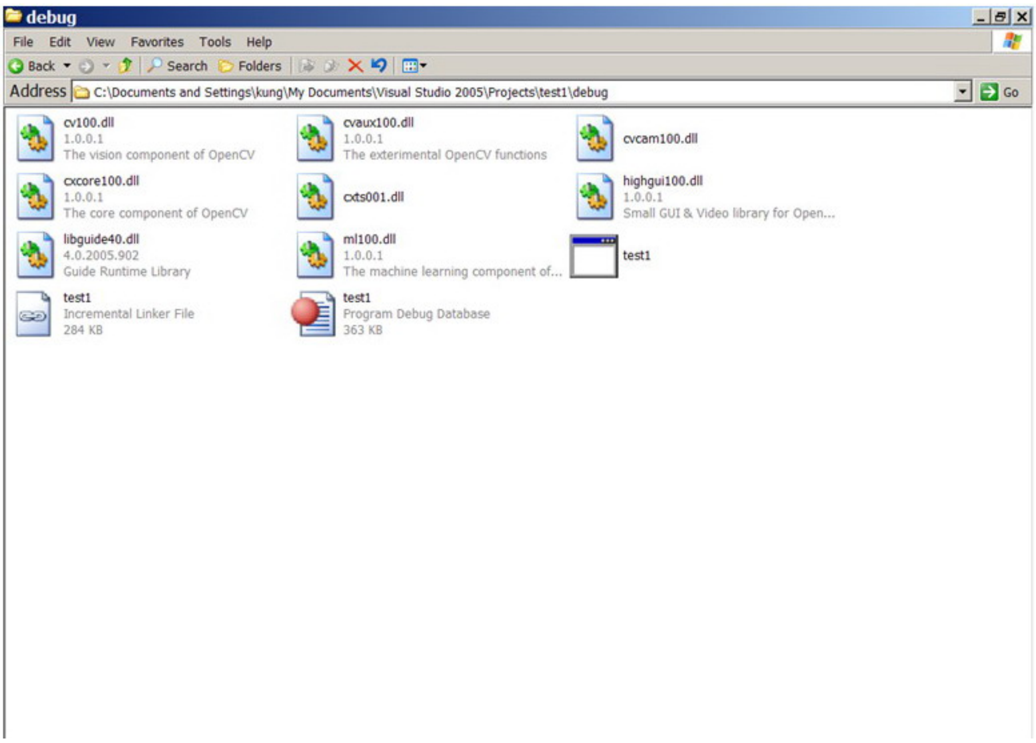



Figure A.13 Debug window

# List of Publications

[1] R. Suppharachyothin, J. Nittayasak, T. Kondo, I. Nilkhamhang, Y. Koike, K. Tungpimolrut, T. Leelasawassuk, "A Comparative Study of Eye Tracking Techniques ," *Proceedings of the International Conference on Information and Communication Technology for Embedded Systems*, Pathumthani, Thailand, 2010.

[2] R. Suppharachyothin, J. Nittayasak, T. Kondo, I. Nilkhamhang, Y. Koike, K. Tungpimolrut, T. Leelasawassuk, "A Scale-variable Template Matching Technique for Eye Tracking," *Proceedings of the International Conference on Embedded Systems and Intelligent Technology* , Chiang Mai, Thailand, 2010.