

人的姿勢與偵測

Detecting People and Their Pose

研究生：林建良

指導教授：李嘉晃 教授

國立交通大學資訊學院 資訊科學與工程研究所碩士班

摘要

日常生活中，人們不停的在做各式各樣的互動。在互動的過程中，頭部姿勢能提供非常重要的資訊。因此，頭部姿勢在視訊會議系統與人機互動的領域都有相關的研究文獻。很多相關研究都使用特殊的設備下或是在近距離判斷頭部姿勢，例如：紅外線或近紅外線發射器。本研究主要是在探討能否在一般的環境下擷取臉部特徵並判斷臉部姿勢。

Detecting People and Their Pose

Student : Jian Liang Lin

Advisor : Prof. Chia-Hoang Lee

Department of Computer and Information Science
National Chiao Tung University

Abstract

People communicate and interact with each other everyday. In order to interact with the computer, head pose detection is a very important task. In the past, there were many literatures proposing techniques for detection of head pose in video conference system and in high resolution video. Most of the studies used such special equipment as infrared illuminator or restricted the detection of head pose at a short distance. In our study, we remove these restrictions and use general camera in ordinary environment. We discuss how to extract facial features and use them to determine head pose.

目錄

| | |
|--|-----|
| 摘要..... | i |
| 目錄..... | iii |
| 圖目錄..... | v |
| 表目錄..... | vii |
| 第一章 緒論..... | 1 |
| 1.1 概述..... | 1 |
| 1.2 研究動機..... | 1 |
| 1.3 研究目標..... | 1 |
| 1.4 論文架構..... | 2 |
| 第二章 文獻探討..... | 3 |
| 2.1 頭部姿勢..... | 3 |
| 2.1.1 臉部特徵分析(Face Features Analysis)..... | 4 |
| 2.1.2 三維臉部模組(3D Face Model)..... | 8 |
| 2.1.3 樣本空間分析(Database Space Analysis)..... | 9 |
| 第三章 設計與實作..... | 11 |
| 3.1 整體架構..... | 11 |
| 3.2 頭部偵測..... | 12 |
| 3.3 動態顏色二分法..... | 15 |
| 3.3.1 概觀特徵..... | 18 |
| 3.4 臉部特徵偵測與追蹤..... | 18 |
| 3.5 姿勢辨識..... | 21 |
| 3.5.1 中距離姿勢辨識..... | 21 |
| 3.5.2 近距離姿勢辨識..... | 24 |
| 第四章 實作過程與結果討論..... | 27 |
| 4.1 實驗說明..... | 27 |
| 4.2 頭部偵測與追蹤結果..... | 27 |
| 4.2.1 無物體遮住人體的場景下..... | 27 |
| 4.2.2 物體遮住小部份人體的測試環境下..... | 29 |
| 4.3 臉部特徵處理效果..... | 30 |
| 4.3.1 近距離拍攝下擷取特徵..... | 31 |
| 4.3.2 中距離拍攝下擷取特徵..... | 33 |
| 4.4 姿勢辨識結果..... | 34 |
| 4.4.1 微觀頭部姿勢辨識..... | 34 |
| 4.4.2 概觀頭部姿勢辨識..... | 38 |
| 第五章 結論..... | 40 |
| 5.1 研究總結..... | 40 |

| | |
|---------------|----|
| 5.2 未來工作..... | 40 |
| 參考文獻..... | 41 |



圖目錄

| | |
|--|----|
| 圖 1：臉部姿勢..... | 3 |
| 圖 2：其他頭部姿勢下的鼻孔形狀合成圖..... | 4 |
| 圖 3：明亮瞳孔(bright pupil)與暗色瞳孔(dark pupil)..... | 5 |
| 圖 4：瞳孔突顯效果..... | 5 |
| 圖 5：臉部姿勢與雙眼中心位置位移示意圖..... | 6 |
| 圖 6：眼睛位置與上下轉動角度示意圖..... | 7 |
| 圖 7：頭部姿勢與臉部垂直特徵的關係..... | 8 |
| 圖 8：動作單元..... | 8 |
| 圖 9：三維人臉模型與影像貼合後的結果..... | 9 |
| 圖 10：人臉資料庫範例..... | 10 |
| 圖 11：系統流程圖..... | 11 |
| 圖 12：擷取人體結果..... | 12 |
| 圖 13：正面人體外圍輪廓..... | 13 |
| 圖 14：側面人體外圍輪廓..... | 14 |
| 圖 15：斜側人體外圍輪廓..... | 14 |
| 圖 16：背面人體外圍輪廓..... | 15 |
| 圖 17：頭部顏色分布範圍..... | 15 |
| 圖 18：偵測特徵方法流程..... | 16 |
| 圖 19：偵測特徵方法對臉部資料庫影像處理後的結果..... | 17 |
| 圖 20：偵測特徵方法對即時影像處理後的結果..... | 17 |
| 圖 21：臉部模組貼合人臉影像示意圖..... | 18 |
| 圖 22：人臉資料庫頭部中心位置測試結果..... | 19 |
| 圖 23：即時影像頭部中心測試結果..... | 19 |
| 圖 24：臉部特徵搜尋順序..... | 20 |
| 圖 25：鼻孔位置追蹤結果..... | 20 |
| 圖 26：眼睛位置追蹤結果..... | 20 |
| 圖 27：中距離拍攝影像經偵測特徵方法後的結果一..... | 22 |
| 圖 28：中距離拍攝影像經偵測特徵方法後的結果二..... | 22 |
| 圖 29：中距離拍攝正面影像經偵測特徵方法後的結果..... | 23 |
| 圖 30：中距離拍攝斜側影像經偵測特徵方法後的結果..... | 23 |
| 圖 31：中距離拍攝側面影像經偵測特徵方法後的結果..... | 23 |
| 圖 32：中距離姿勢辨識流程圖..... | 24 |
| 圖 33：近距離姿勢辨識流程圖..... | 25 |
| 圖 34：頭部姿勢與向量的關係..... | 26 |
| 圖 35：人體擷取與頭部偵測於簡單場景..... | 28 |
| 圖 36：人體擷取與頭部偵測於複雜場景 1..... | 28 |

| | |
|-------------------------------|----|
| 圖 37：人體擷取與頭部偵測於複雜場景 2..... | 28 |
| 圖 38：有物體遮蔽的情形下頭部偵測結果 1..... | 29 |
| 圖 39：有物體遮蔽的情形下頭部偵測結果 2..... | 29 |
| 圖 40：ORL 人臉資料庫的結果..... | 30 |
| 圖 41：UMIST 人臉資料庫的結果..... | 30 |
| 圖 42：在不同姿勢下的鼻孔位置..... | 31 |
| 圖 43：在不同姿勢下的眼睛位置..... | 32 |
| 圖 44：抬頭姿勢與偵測特徵方法的結果..... | 32 |
| 圖 45：中距離抬頭姿勢下偵測特徵方法的結果..... | 33 |
| 圖 46：中距離向右轉動姿勢下偵測特徵方法的結果..... | 33 |
| 圖 47：中距離向左轉動姿勢下偵測特徵方法的結果..... | 33 |
| 圖 48：中距離低頭姿勢下偵測特徵方法的結果..... | 34 |
| 圖 49：正面姿勢下與正面姿勢量化數值..... | 35 |
| 圖 50：正面姿勢下與正面姿勢量化數值..... | 35 |
| 圖 51：右斜側姿勢下與右斜側姿勢量化數值..... | 36 |
| 圖 52：右側姿勢下與右側姿勢量化數值..... | 36 |
| 圖 53：左斜側姿勢下與左斜側姿勢量化數值..... | 36 |
| 圖 54：左側姿勢下與左側姿勢量化數值..... | 37 |
| 圖 55：低頭姿勢下與低頭姿勢量化數值..... | 37 |
| 圖 56：抬頭姿勢下與抬頭姿勢量化數值..... | 37 |
| 圖 57：正面姿勢下與正面姿勢量化數值..... | 38 |
| 圖 58：右斜側姿勢下與右斜側正面姿勢量化數值..... | 39 |
| 圖 59：左側姿勢下與左側正面姿勢量化數值..... | 39 |

表目錄

| | |
|----------------------|----|
| 表 1：本研究使用的硬體設備：..... | 27 |
| 表 2：頭部姿勢與數值對應表..... | 35 |
| 表 3：頭部姿勢與數值對應表..... | 38 |



第一章 緒論

1.1 概述

頭部姿勢對於人機互動和電腦視覺等資訊領域是一個很重要的議題。當視訊會議系統開始普及後，很多使用者發現頭部姿勢與眼睛注視不協調時，會造成視訊會議進行與溝通上的困擾。因此，很多相關研究開始對頭部姿勢與眼睛注視的問題加以探討。早期的研究除了使用特殊的攝影裝置，例如：紅外線攝影機[1][2]或是影像處理等方式[3]外，均需近距離拍攝臉部來辨認頭部姿勢。近幾年，有很多學者開始針對遠距離拍攝環境下的頭部姿勢進行研究[4]，使得頭部姿勢可以更廣泛的被應用在各種領域上。在人機互動方面，利用頭部姿勢與臉部特徵來操作滑鼠游標[5]，或是在視訊會議中紀錄重要畫面。

本研究基於臉部的色彩分佈與特徵，提供一種在近距離與中距離皆能有效分辨頭部姿勢的方法。在近距離的環境下，此方法可以加速搜尋臉部特徵演算法。在中距離的環境下，此方法可以使得監視系統更有效的擷取監視影像，略過不重要的鏡頭。

1.2 研究動機

在一般的監視環境下，系統只提供追蹤人體的功能已經不符合需求。為了要讓追蹤系統更有效率的追蹤人體，需要更多細部的偵測與辨識。本研究主要探討如何在中距離與近距離判斷出頭部姿勢，使得監視系統能夠有效的擷取與追蹤人體。

1.3 研究目標

本研究主要是探討如何在中距離與近距離判斷頭部姿勢，使得監視系統能夠有效的擷取與追蹤移動中的人體。因此，系統需要達到臉部偵測與追蹤的效果。在穩定追蹤頭部的情況下，擷取出頭部特徵。利用人體的對稱性與脸部特徵進行整體分析與估測特徵點位置，進而判斷頭部姿勢。

1.4 論文架構

第二章文獻探討將介紹背景知識及相關研究。第三章系統設計與實作首先介紹系統整體架構。第四章系統應用實例，我們將展現系統對於頭部的追蹤表現。另外針對臉部特徵與頭部姿勢判別的穩定性做數據上的評估。第五章結論將對本研究作討論並說明其未來發展方向。



第二章 文獻探討

2.1 頭部姿勢

要決定頭部姿勢，首先要固定觀察點的位置，而頭部則是處在一個未知的姿勢。頭部姿勢就是對 x 軸、y 軸與 z 軸做旋轉動作所造成的現象，可以分作上下轉動(pitch)、左右轉動(yaw)與滾動(roll)。如圖 1 所示：

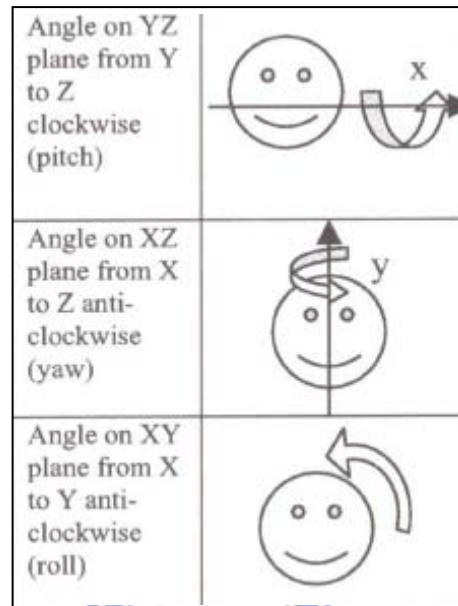


圖 1：臉部姿勢

決定頭部姿勢有一個很重要的因素就是攝影機的解析度。很多相關研究都利用高解析度的攝影機，積極的在臉部找尋穩定的特徵點位置，並利用臉部特徵點的位置決定頭部姿勢。Yoshinobu 等人[2]的研究中找尋瞳孔與鼻孔的位置，並利用幾何關係判斷出面部方向。Zitnick 等人[6]在臉部研究的文獻中建議利用鼻孔即可穩定且有效的追蹤頭部姿勢。在 2.1.1 小節內我們將詳細敘述這些方法的特點。

另外，頭部姿勢可以看做頭部模型在三維空間做轉動與平移的動作，這個觀念就發展出將三維頭部模型對應到二維影像上來判別頭部姿勢的方法。Hai Tao 等人[7]提出三維頭部模型貼合在二維影像上的演算法，而 Jilin Yu 等人[5]則利用這套演算法來達到控制螢幕游標的功能。在 2.1.2 小節內我們將詳細敘述這

些方法的特點。

2.1.1 臉部特徵分析(Face Features Analysis)

在國內外的研究中，有許多方法是透過臉部特徵位置來判斷臉部姿勢。Zitnick 等人[6]在其發表文章的附錄中，建議利用鼻孔的特徵來判斷臉部方向。這個方法首先需儲存鼻孔的初始位置與形狀，並利用此初始狀態，推算並合成出在其它頭部姿勢下的鼻孔形狀，如圖 2 所示：

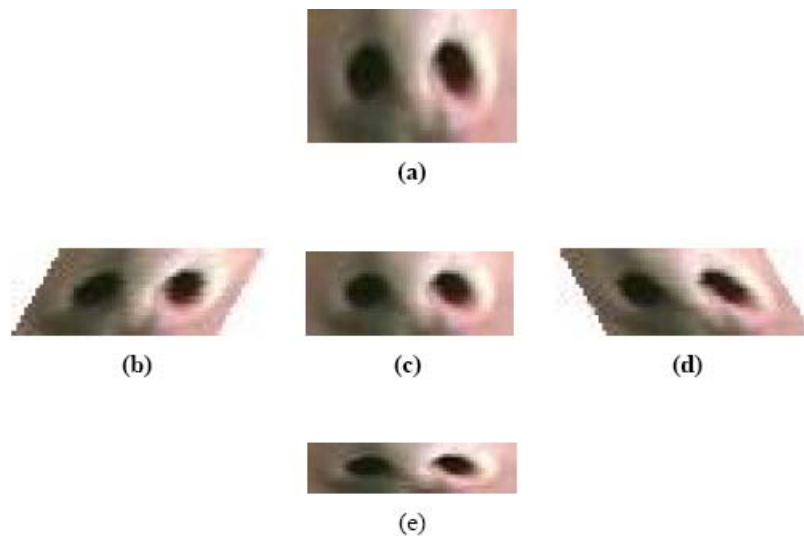


圖 2：其他頭部姿勢下的鼻孔形狀合成圖

有了鼻孔部分的位置與形狀，就可以利用其他有效的追蹤演算法追蹤鼻孔位置。並利用鼻孔形狀的比對，就可以比對出頭部姿勢。

這個方法需要很精準的取出鼻孔位置與形狀，才能計算出臉部方向。因此，需要高解析度的攝影機或是將攝影機架設在離使用者較近的距離，對於架設在離人群較遠的監控式攝影機是一件很困難的事情。另外，這個方法需要手動擷取鼻孔的初始狀態，定義出初始位置與方向。

Ji 和 Yang[1]則利用紅外線(infrared illuminator)或是近紅外線(near-infrared illuminator)發射器，照射人臉來偵測與判斷瞳孔大小、兩瞳孔間的距離與瞳孔形狀來判定簡易的頭部姿勢。其目的是要利用頭部姿勢來判斷駕駛人是否失去注意力或是過於疲憊。

因此，他們採用近紅外線照射人臉，而近紅外線照射的環境下有下列三種好處：一、減少周遭光源不足或是週遭光源不穩定的影響。二、近紅外線會有突顯瞳孔特徵的效應。三、近紅外線對於使用者而言幾乎為不可視光。並利用近紅外線光源所架設位置的不同定義出明亮瞳孔(bright pupil)與暗色瞳孔(dark pupil)的區別，如圖 3 所示。利用明亮瞳孔與暗色瞳孔的差異，更可明顯突顯出瞳孔位置，我們可以在圖 4 可以看出效果。

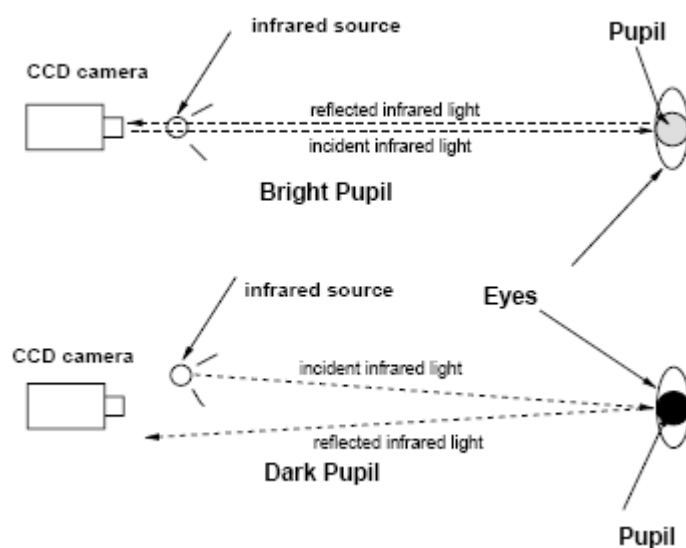


圖 3：明亮瞳孔(bright pupil)與暗色瞳孔(dark pupil)



圖 4：瞳孔突顯效果

Ji 和 Yang 並觀察瞳孔位置與形狀，來決定頭部姿勢。其觀察結果有三：一、兩瞳孔間的位置會因為頭部轉動而移動。二、瞳孔形狀會因為頭部轉動而呈現橢圓的形狀。三、瞳孔面積會因為頭部轉動而縮小。

然而，這個方法無法在普通的攝影環境中實現。另外，因為頭部姿勢需利用瞳孔位置及大小來做判斷，若瞳孔因頭部轉向過大而無法被攝影機捕捉到時，則

會產生誤判的情況。若是考慮瞳孔大小的變化，則攝影機需要較高的解析度或是擺設在離使用者較近的距離。

在 Ben Yip 和 Jess S. Jin[12]則是利用雙眼的中心位置與雙眼連線的方向等資訊，來判斷頭部是否有作 x, y 或 z 軸方向的旋轉。圖 5 為其示意圖：

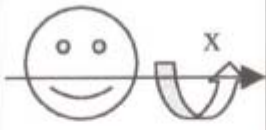











| | | -ve | 0° | +ve |
|---|--|--|---|--|
| Angle on YZ plane from Y to Z clockwise (pitch) |  |  |  |  |
| Angle on XZ plane from X to Z anti-clockwise (yaw) |  |  |  |  |
| Angle on XY plane from X to Y anti-clockwise (roll) |  |  |  |  |

圖 5：臉部姿勢與雙眼中心位置位移示意圖

這個方法先計算出在頭部正面或是某一已知頭部姿勢下，頭部中心到兩眼位置中心的向量 U ， $U=(U_x, U_y)$ 。而未知頭部姿勢下，頭部中心到兩眼位置中心的向量為 V ， $V=(V_x, V_y)$ 。其中頭部中樞為 $C=(C_x, C_y, C_z)$ ，此三維位置經過 Murat 等人研究中定義為(0 ES, -1290 ES, -290 ES)，ES 為兩眼間距除以 1024。並假設 E 為兩眼中心位置。因此，可以計算出旋轉角度，如圖 6 所示：

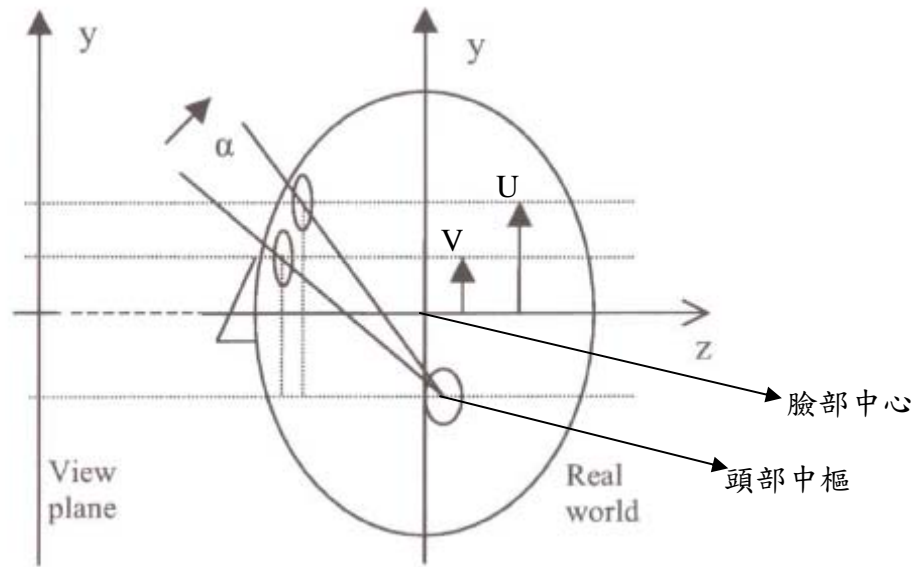


圖 6：眼睛位置與上下轉動角度示意圖

$$r = \sqrt{(U_r - C_r)^2 + (E_z - C_z)^2}$$

$$\alpha = \sin^{-1}\left(\frac{V_r - C_r}{r}\right) - \sin^{-1}\left(\frac{U_r - C_r}{r}\right)$$

這個方法需要決定出頭部的樞軸中心，這個樞軸中心並不是這麼好估測與實現的。另外，要精準算出兩眼的中心位置及向量偏轉的角度需要較機密的攝影器材才能準確推測出頭部姿勢。

另外，Yegnanaarayana 等人[3]則利用人臉在做旋轉動作的同時，會將臉部某些區域延伸，而將另外一些區域壓縮的現象，將臉部分析為垂直特徵與水平特徵，透過人臉正面垂直特徵與人臉側面垂直特徵的分佈情況，來判斷人臉姿勢。

此方法首先垂直掃描正面的人臉資訊，將垂直掃描得到的人臉資訊轉換到頻率空間上， $U = \{U_1, U_2, \dots, U_n\}$ ，其中 n 為垂直掃描線個數。另外，掃描側面人臉資訊，也轉換到頻率空間上， $V = \{V_1, V_2, \dots, V_n\}$ ， n 為垂直掃描線個數。接著，利用 DSW(Dynamic Space Analysis)分析 V_i 與 U_j 的相似性，其中 $i > 0, i < n$ 且 $j > 0, j < n$ 。

其 70×70 的人臉圖形分析結果如圖 7 所示：

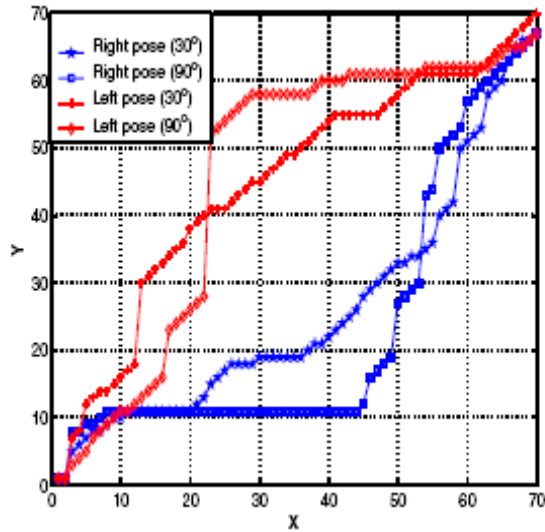


圖 7：頭部姿勢與臉部垂直特徵的關係

2.1.2 三維臉部模組(3D Face Model)

在電腦圖學的研究中，三維臉部模型已經可以透過三維掃描來建立，而臉部各種表情與動作都可以透過幾個基本的動作單元(Action Units)合成。而 Hai Tao 等人[7]將動作單元分成下列 12 種，如圖 8 所示：

| AU | Description |
|----|--|
| 1 | Vertical movement of the center of upper lip |
| 2 | Vertical movement of the center of lower lip |
| 3 | Horizontal movement of left mouth corner |
| 4 | Vertical movement of left mouth corner |
| 5 | Horizontal movement of right mouth corner |
| 6 | Vertical movement of right mouth corner |
| 7 | Vertical movement of right eyebrow |
| 8 | Vertical movement of left eyebrow |
| 9 | Lifting of right cheek |
| 10 | Lifting of left cheek |
| 11 | Blinking of right eye |
| 12 | Blinking of left eye |

圖 8：動作單元

在三維臉部模型的架構下，可以清楚看出臉部特徵的移動變化，例如：眉毛、眼睛與嘴巴的移動。之後，要將三維臉部模型穩定的貼合在二維影像上，再利用三維模型在空間上的數學關係推知二維影像上的頭部姿勢。

而 Jilin Tu 與 Thomas Huang 等人[5]則提供了一種三維模型貼合在二維影像的方法。首先需針對二維影像作人臉偵測的動作，例如：Adaboosting Algorithm。當偵測出人臉位置後，利用動態外型模組(Active Shape Model)辨認出人臉特徵位置。最後，將三維人臉模型做適度的變形來貼合這些人臉特徵。其最後的結果如圖 9 所示：

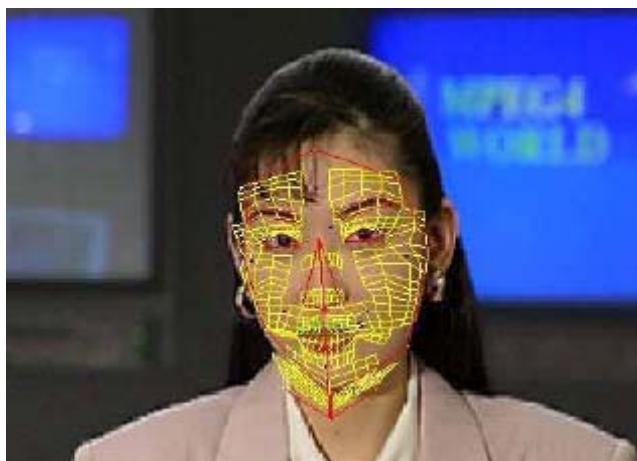


圖 9：三維人臉模型與影像貼合後的結果

2.1.3 樣本空間分析(Database Space Analysis)

頭部姿勢的問題，通常都會利用大量資料來做特徵分析，並萃取出可分別頭部姿勢的特徵後，建立出一個頭部姿勢的分類器。此方法主要流程如下：

首先是特徵點的選取，在最近的研究中，比較常見的特徵為本徵空間(Eigen Space)與赫氏小波轉換(Harr wavelet transform)。接著，利用主要要素分析法(Principal Components Analysis)或是其他機器學習演算法，例如：Adaboosting Algorithm，萃取出具有代表性的特徵向量。

因為要取得大量人臉資料，有些研究[8]是利用三維人臉建模與彩現的方式控制人臉轉向與光源照射的問題。有些研究[9]則是直接利用人臉資料庫來做機器學習的分析，圖 10 為其中一種人臉轉向資料庫。





圖 10：人臉資料庫範例

第三章 設計與實作

3.1 整體架構

本研究主要是探討如何能穩定的在中近距離判斷出人臉位置與姿勢。因此，本研究首先利用運動資訊區別動態物體與背景，之後即可追蹤與辨別頭部姿勢。頭部姿勢偵測流程為(一)頭部偵測、(二)臉部特徵偵測、(三)鼻部推測與追蹤、(四)姿勢分析與判斷。系統流程如下：

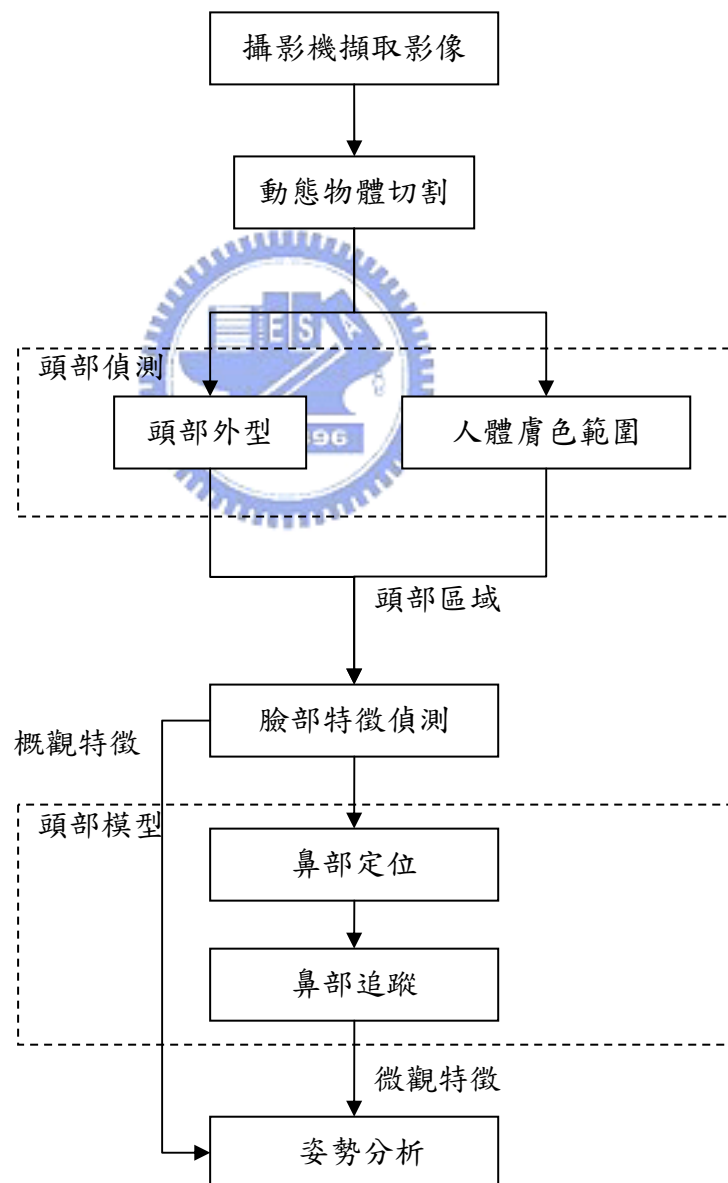
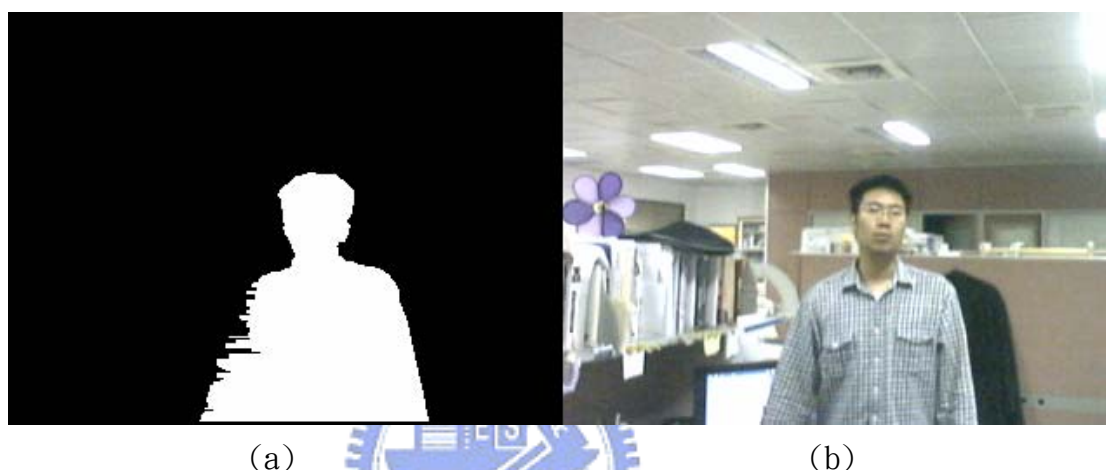


圖 11：系統流程圖

3.2 頭部偵測

我們在人體行走的狀況下，首先針對相鄰兩個串流影像做差分運算，獲得運動資訊，並利用此資訊從背景中擷取出人體。之後透過人體外圍輪廓分析，擷取出頭部位置。最後利用較寬鬆的人體膚色範圍當過濾器，並將此膚色範圍轉換到 HSI 領域上[13]，將頭部區域做小幅度的修正，達到更精確的位置。稍後我們將在下面詳細討論人體外圍輪廓分析。圖 12 為人體擷取出的狀態：



(a)

(b)

圖 12：擷取人體結果

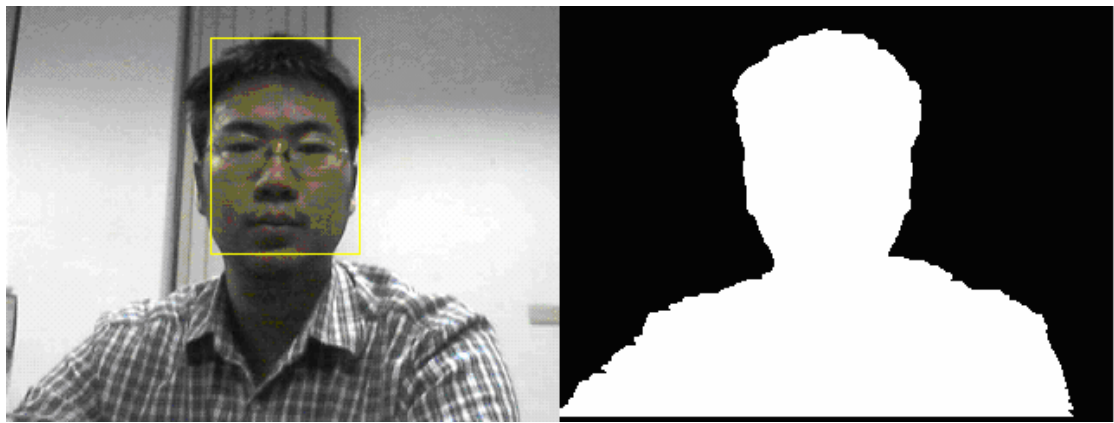
(a)從背景擷取出人體後的結果，(b)實際環境下的串流影像

我們可以從人體觀察得出，當人體只做簡單的行走動作或是站立的姿勢時，無論任何角度，頭部到頸部的外在曲線是固定的。圖 13 為正面姿勢下攝影機所擷取到的畫面與人體外圍輪廓曲線，圖 14 為側面姿勢下攝影機所擷取到的畫面與人體外圍輪廓曲線，圖 15 為斜側姿勢下攝影機所擷取到的畫面與人體外圍輪廓曲線，圖 16 為背面姿勢下攝影機所擷取到的畫面與人體外圍輪廓曲線。

因為頭部區域可視作一個剛體，因此，頭部區域的輪廓曲線在一般的情況下是固定不變的。所以，我們可以針對此剛體輪廓作曲線趨勢分析，並擷取出適當的局部最小值(Local minimum)當做頸部位置，即是圖 13 中紅色方塊所標示的位置。

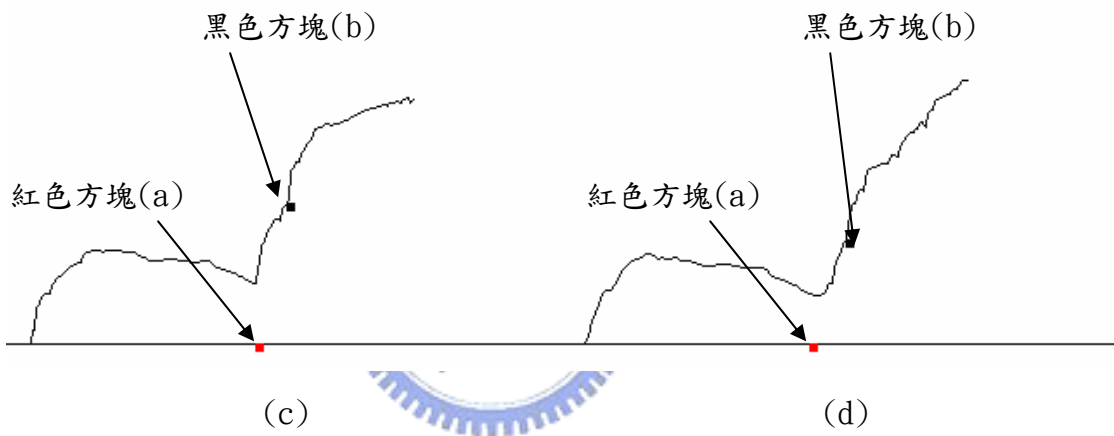
由圖 13 我們可以觀察到，紅色方塊(a)的位置為頸部位置，而黑色方塊(b)的位置大致上落在肩膀附近。經過人體外圍輪廓分析後，我們可以很容易的擷取

出頭部位置。



(a)

(b)



(c)

(d)

圖 13：正面人體外圍輪廓

(a)實際環境下的串流影像與臉部偵測的結果，(b)從背景擷取出人體的結果，(c)右邊人體輪廓曲線與頸部預測與肩部預測，(d)左邊人體輪廓曲線與頸部預測與肩部預測



(a)

(b)

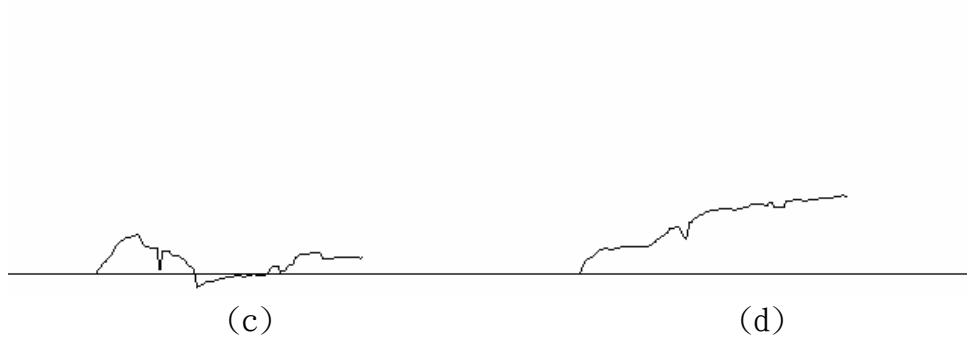


圖 14：側面人體外圍輪廓

(a)實際環境下的串流影像與臉部偵測的結果，(b)從背景擷取出人體的結果，(c)右邊人體輪廓曲線，(d)左邊人體輪廓曲線



(a)

(b)



(c)

(d)

圖 15：斜側人體外圍輪廓

(a)實際環境下的串流影像與臉部偵測的結果，(b)從背景擷取出人體的結果，(c)右邊人體輪廓曲線，(d)左邊人體輪廓曲線

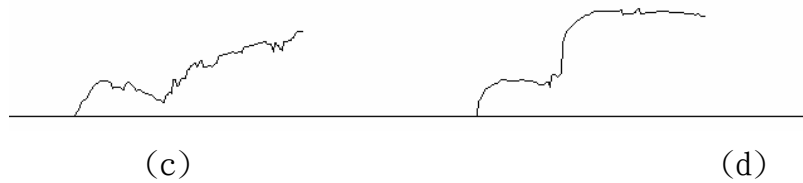


圖 16：背面人體外圍輪廓

(a)實際環境下的串流影像與臉部偵測的結果，(b)從背景擷取出人體的結果，(c)右邊人體輪廓曲線，(d)左邊人體輪廓曲線

3.3 臉部特徵偵測

取得頭部區域以後，為了做頭部特徵分析，我們設計了偵測特徵的方法。在攝影機捕捉到的影像中，我們觀察出臉部皮膚與其他如眼睛與鼻子的部位有顏色上的差距。另外，在頭部正對與側對攝影機的狀態下，膚色區域的比例較其他區域的比例高，圖 17 說明此觀察。

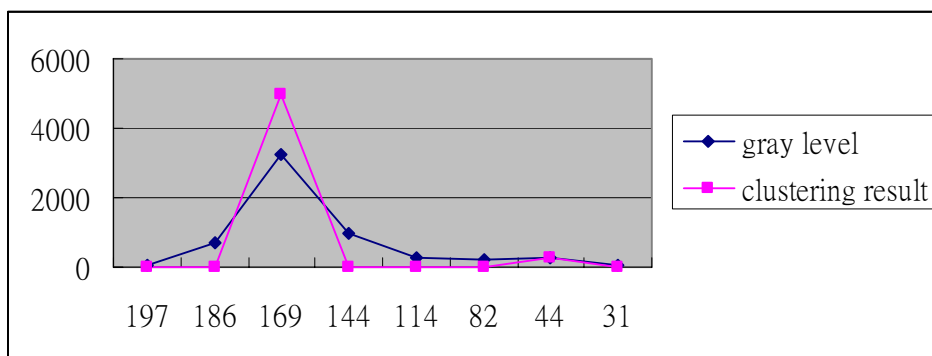


圖 17：頭部顏色分布範圍

偵測特徵方法主要步驟有三，一、萃取頭部主要顏色，二、萃取特徵主要

顏色，三、儲存分群樣板。圖 18 為偵測特徵方法的流程圖。首先，系統會做初步的顏色分群，將相近的顏色合併(merge)與分群。針對初步分群後的結果，萃取出最大群集的中心當作頭部主要顏色，並合併鄰近顏色群集。接著取出頭部區域偏暗的群集中心，並合併鄰近顏色，當作臉部特徵主要顏色。

將頭部主要顏色與特徵主要顏色儲存至分群樣板中，並針對頭部區域像素與此兩主要顏色做差值上的比對，若是與頭部主要顏色較接近，則歸類為皮膚顏色。若與特徵主要顏色較接近，則歸類為特徵顏色。

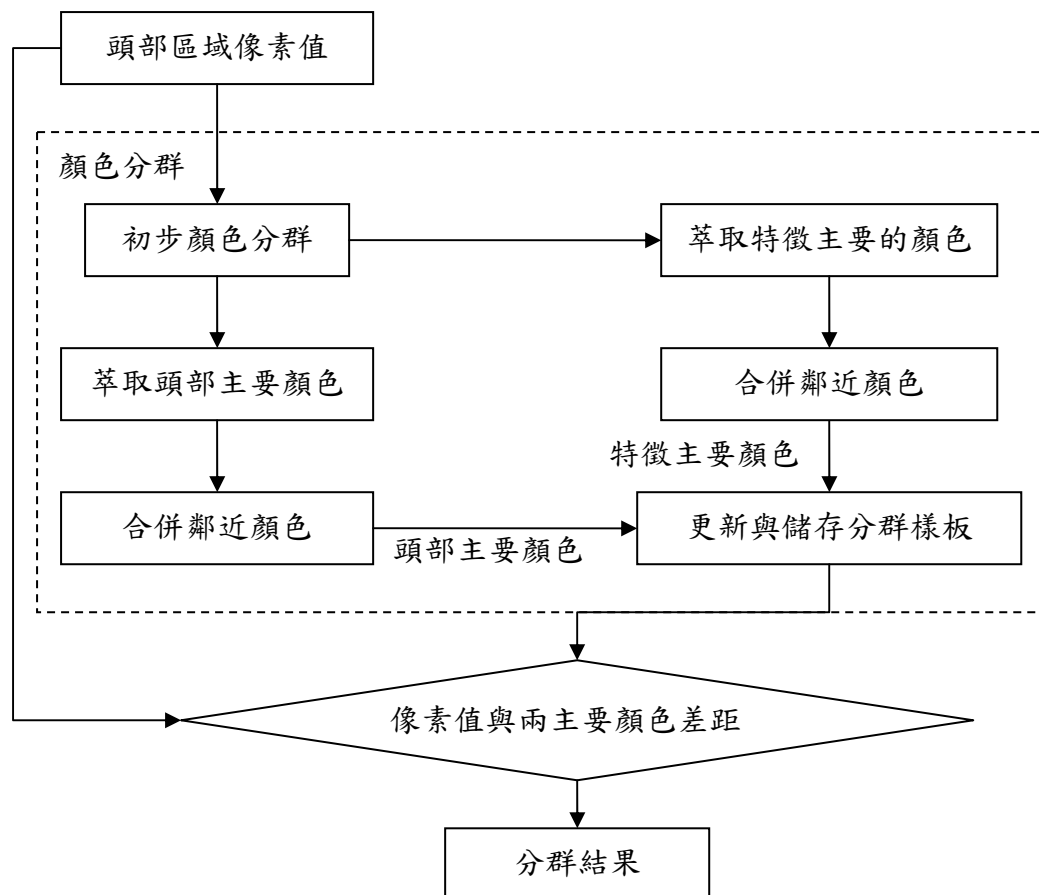


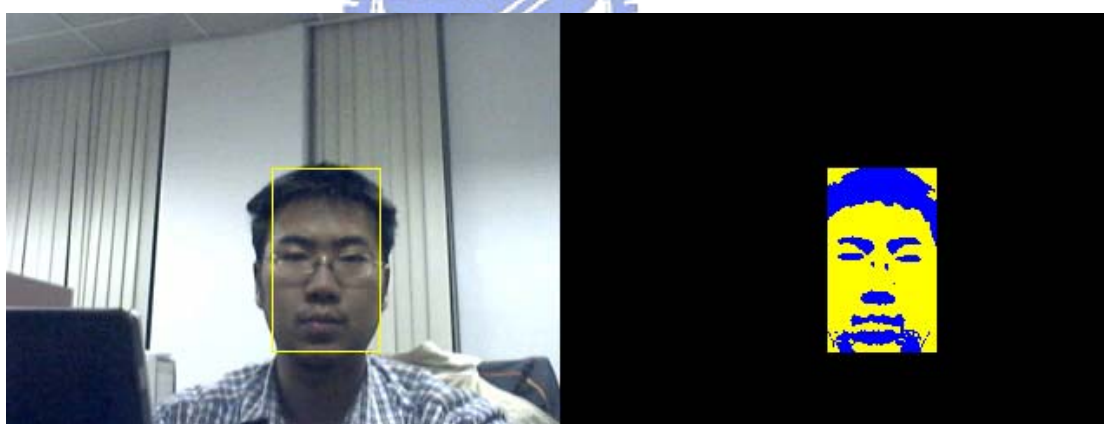
圖 18：偵測特徵方法流程

經過偵測特徵方法處理臉部資料庫影像[10]其測試結果如下：



圖 19：偵測特徵方法對臉部資料庫影像處理後的結果

經過偵測特徵方法處理即時拍攝影像結果如圖 20：



(a)

(b)

圖 20：偵測特徵方法對即時影像處理後的結果

(a)實際環境下串流影像與臉部偵測的結果，(b)針對臉部偵測特徵方法的結果

在即時拍攝的環境下，我們採用一般的室內採光。經由偵測特徵方法突顯臉部特徵後，我們觀察到偵測特徵方法可以區分出臉部膚色區域、頭髮分佈、眼睛與鼻孔位置。然而，因為室內光源無法充分照射的關係，嘴巴區域與下巴附近的顏色會有不穩定的情況。因此，在做頭部姿勢分析與判斷的動作時，我們略去這

個不穩定的因素，只對頭髮、眼睛與鼻子等區域做詳細的分析與判斷。

3.3.1 概觀特徵

頭部區域經過偵測特徵方法處理後，我們觀察到在臉部正對攝影機的情況下，偵測特徵方法所得到的結果呈現較對稱的狀態。若是頭部左右轉動，就會呈現不對稱的狀態。因此，我們可以利用此概觀頭部資訊與頭部中心位置進行比對，得到概觀的頭部姿勢：頭部向左或向右轉動。

概觀的頭部資訊提供了三種好處，一、可以快速定位眼睛和鼻孔位置，二、可以輔助推算出微觀的頭部姿勢，三、在中距離的環境下能夠穩定的偵測出頭部姿勢。

3.4 臉部特徵偵測與追蹤

經過偵測特徵方法與概觀的頭部資訊推測後，我們可以判斷出頭部是否正對攝影機或是向左、向右轉動。在此小節我們利用臉部的位置是正對或是幾乎正對攝影機的環境下，將臉部模組貼合到人臉影像上，之後快速的搜尋眼睛與鼻孔的位置，如圖 21 所示：

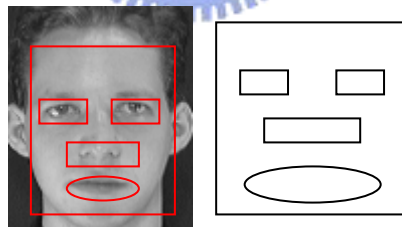


圖 21：臉部模組貼合人臉影像示意圖

首先，我們需要取得頭部區域某一特徵點當作參考點，將臉部特徵模組貼合在人臉影像上，才能偵測眼睛與鼻孔位置。在本研究所使用的人臉資料庫統計得知，在臉部的位置是正對或是幾乎正對攝影機的環境下，頭部區域的中心位置落點大致上在眼睛與鼻孔間。因此，我們取頭部區域的中心位置當做此演算法的參考點。我們使用人臉資料庫測試頭部區域中心位置的結果顯示於圖 22。我們拍攝即時影像並取出頭部區域中心位置的結果為圖 23。



圖 22：人臉資料庫頭部中心位置測試結果

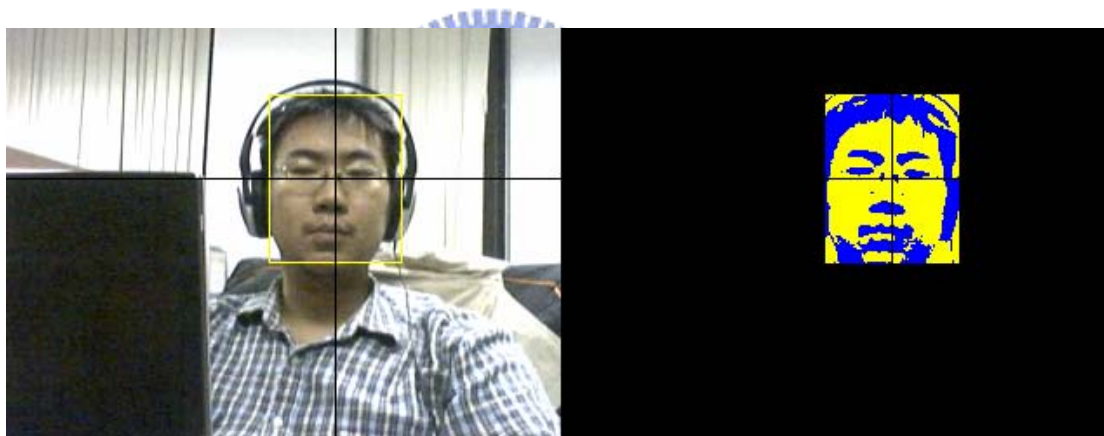


圖 23：即時影像頭部中心測試結果

(a)實際環境下串流影像與臉部偵測的結果，(b)頭部中心測試結果

因為我們利用頭部區域的中心位置當做參考點，我們便可以利用參考點的位置並配合臉部模型做小幅度的特徵搜尋。利用參考點的位置，將臉部分做四個區域，如圖 24 所示，第一區域的搜尋方式為由左到右，由下而上，定位出左眼的位置。第二區域的搜尋方式為由右到左，由下而上，定位出右眼的位置。第三區域為由右到左，由上而下，定義出鼻子偏右的區域。第四區域為由左到右，由上而下，定義出鼻子偏左的區域。之後，由第三區域與第四區域得到的位置推算出鼻孔中心的位置。

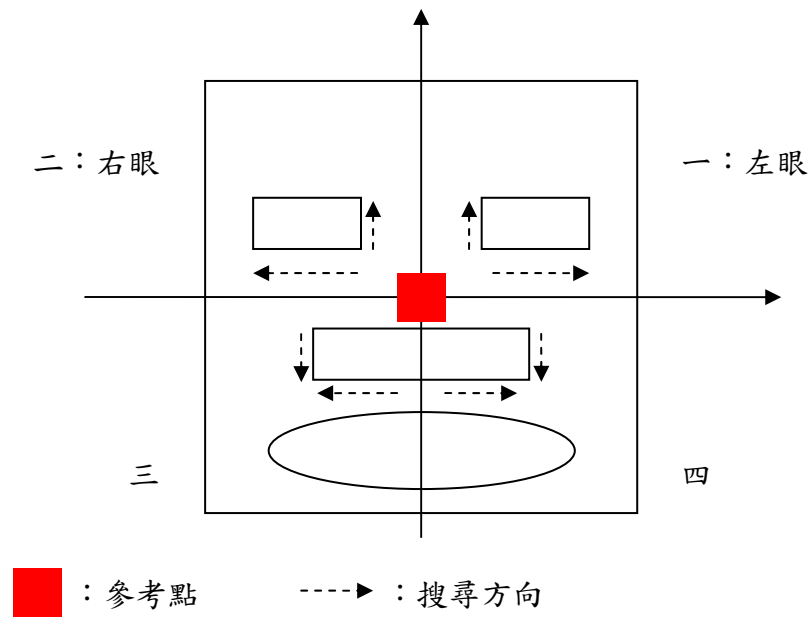


圖 24：臉部特徵搜尋順序

當我們推測出眼睛與鼻子的位置後，我們利用特徵點追蹤方法繼續追蹤眼睛與鼻孔的位置。鼻孔位置追蹤結果與雙眼位置追蹤結果如圖 25 與圖 26：



圖 25：鼻孔位置追蹤結果

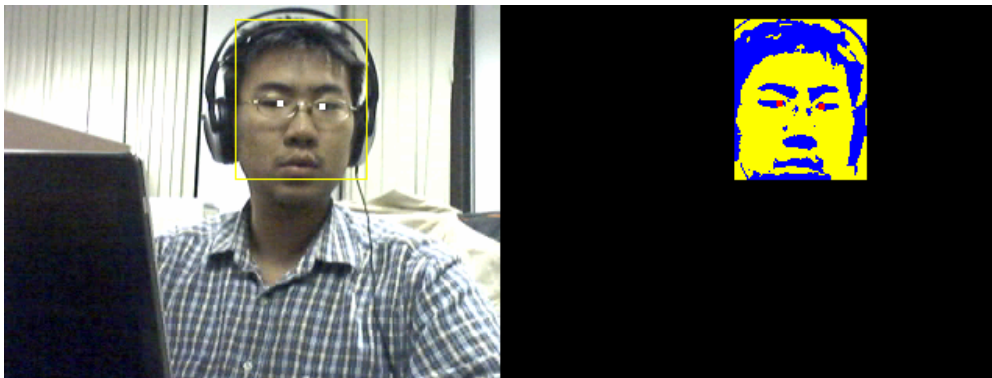


圖 26：眼睛位置追蹤結果

3.5 姿勢辨識

攝影機架設在近距離的環境下，有很多辨識頭部姿勢的相關研究[1-9]。這些相關研究擷取臉部特徵的方式，與我們系統所利用的偵測特徵方法不同，例如：利用紅外線發射器等裝置擷取瞳孔位置或是利用動態形狀模型偵測與擷取臉部特徵。在 3.5 小節裡我們會針對本系統擷取出來的臉部特徵做分析與評估，以考量利用哪些特徵來判斷頭部姿勢。

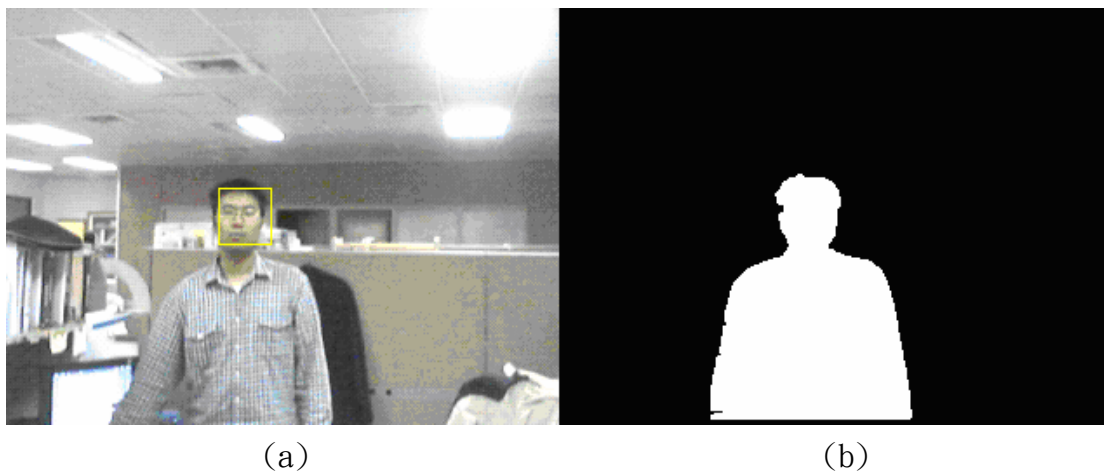
我們研究的目的是主要是判斷人臉是否朝向攝影機，因此，近距離拍攝下，我們只做臉部上下轉動及左右轉動的判斷，不做臉部滾動(Roll)的判斷。中距離拍攝的情況下會造成頭部資訊減少，目前只作左右轉動的判斷。

我們將這小節分成兩部份來討論。一部分是在中距離的環境下判斷頭部姿勢。在這個情況下，我們利用概觀的資訊判斷頭部姿勢，之後我們會做詳細的分析與探討。另一部份是在近距離的環境下，利用臉部細微特徵來判斷頭部姿勢，例如：眼睛和鼻孔位置。

3.5.1 中距離姿勢辨識

因為每台攝影機的解析度不同，因此，我們定義攝影機所擷取到的頭部區域大小為 30 像素×40 像素左右時，就視作中距離頭部拍攝。50 像素×70 像素以上時，則視作近距離頭部拍攝。

圖 27 與圖 28 為 30 像素×40 像素頭部區域經過偵測特徵方法的結果：



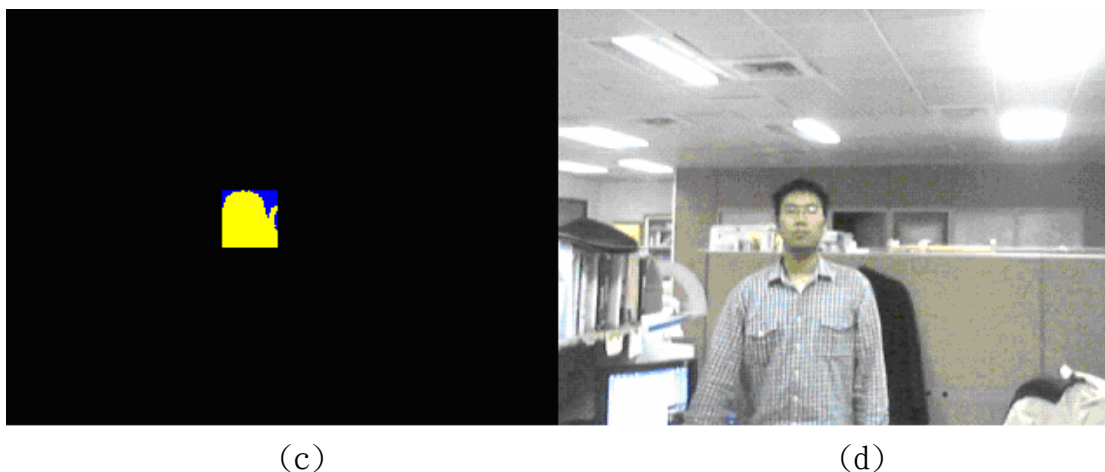


圖 27：中距離拍攝影像經偵測特徵方法後的結果一
 (a)頭部偵測的結果，(b)從背景擷取的人體，(c)針對頭部進行偵測特徵方法的結果，(d)複雜場景下的串流影像

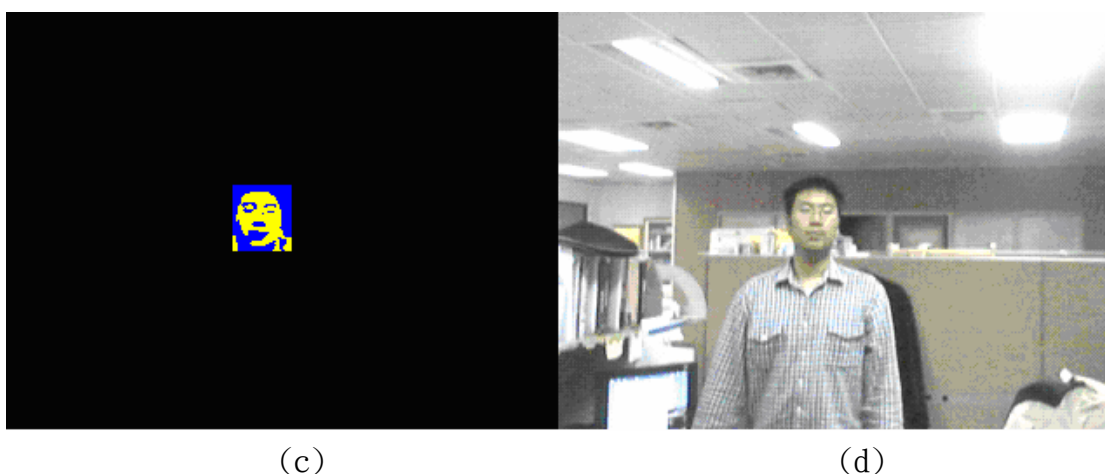
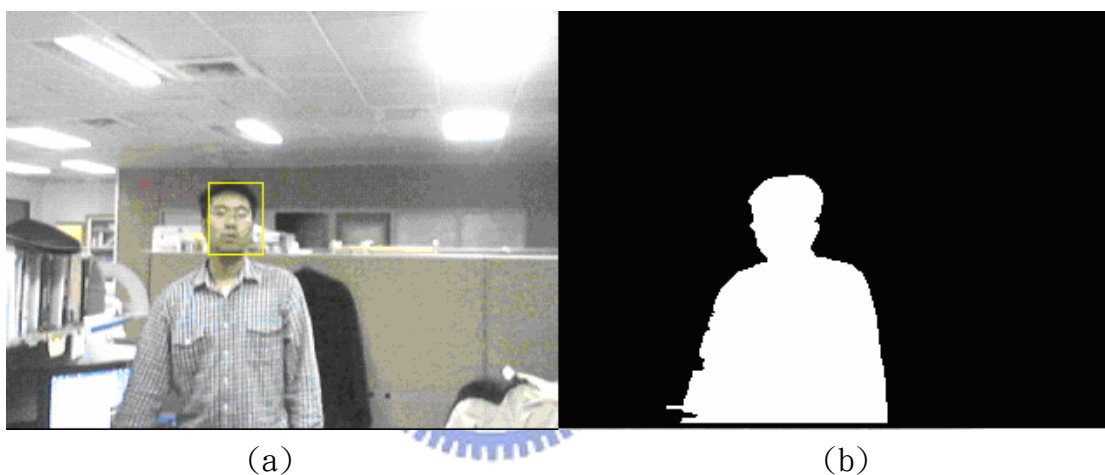


圖 28：中距離拍攝影像經偵測特徵方法後的結果二
 (a)頭部偵測的結果，(b)從背景擷取的人體，(c)針對頭部進行偵測特徵方法的結果，(d)複雜場景下的串流影像

在圖 27 與圖 28 中，因為人與攝影機的距離過遠，偵測特徵方法的效果有時

無法突顯眼睛與鼻孔的部份。只能分辨出頭髮區域與膚色區域。所以，在中距離姿勢辨識上，我們利用頭部的對稱性來做分析，依據臉部的對稱性作頭部是否向左或向右轉動的概觀性頭部姿勢推測。

我們從使用的人臉資料庫統計，在大部分的狀況下，人臉正對攝影機時，膚色區域的分佈是對稱的。當人臉從正對攝影機的姿勢作左右轉動時，膚色區域的分佈就會變得不對稱。圖 29~圖 31 顯示頭部於不同姿勢下的膚色分結果。



圖 29：中距離拍攝正面影像經偵測特徵方法後的結果



圖 30：中距離拍攝斜側影像經偵測特徵方法後的結果

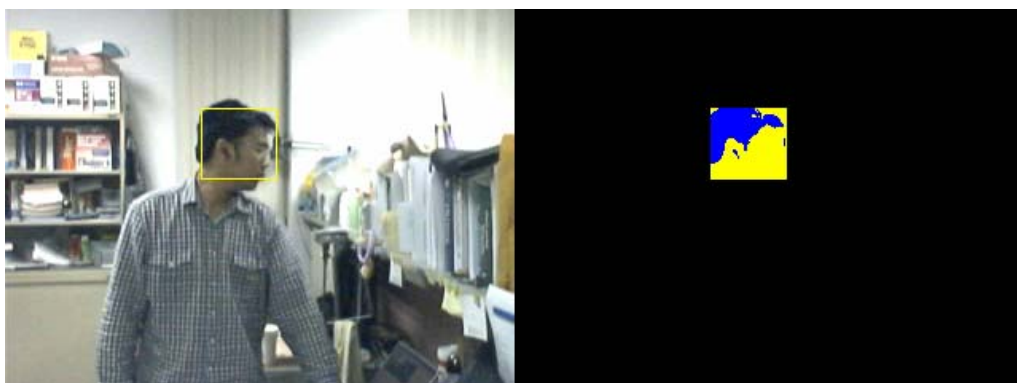


圖 31：中距離拍攝側面影像經偵測特徵方法後的結果

因此，我們可以利用膚色面積大小與正臉膚色區域的對 y 軸有對稱性，判斷臉部是否正對攝影機。我們利用算數平均數計算出膚色的中心位置(S_x, S_y)。將

膚色中心位置與之前估測出的頭部的中心位置(H_x, H_y)做相減，最後除以頭部高度則獲得量化的頭部姿勢。我們可以得到若其數值落在零附近時，則表示正對攝影機。若是正值，表示臉部向左偏轉。若是負值，表示臉部向右偏轉。

中距離姿勢辨識的流程如圖 32：

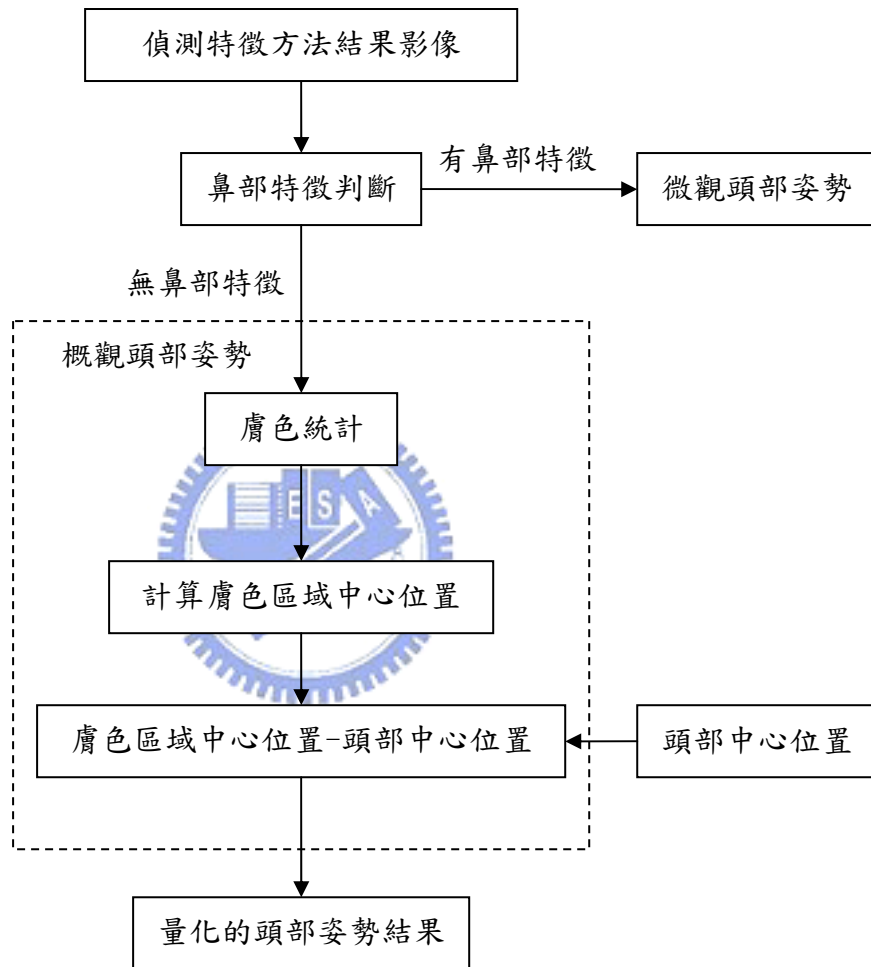


圖 32：中距離姿勢辨識流程圖

3.5.2 近距離姿勢辨識

我們在 3.4 小節提到，在近距離頭部拍攝的狀況下，偵測特徵方法可以將眼睛與鼻孔從膚色中區分出來。因此，我們可以利用更細部的臉部特徵辨別頭部姿勢。

當擷取雙眼位置與鼻孔位置後，我們可以利用雙眼位置與鼻孔位置的幾何關

係來決定頭部姿勢。然而，在頭部轉向的過程中，尤其在使用者做抬頭的動作到某種角度，攝影機無法捕捉到眼睛部位。這樣的情況下，使得偵測特徵方法在眼睛部份表現不穩定。然而鼻孔處於臉部的中央位置，在做頭部轉動時最容易被攝影機捕捉到。最後，本研究決定採用鼻孔位置作細部的頭部判斷。

近距離姿勢辨識的流程如圖 33：

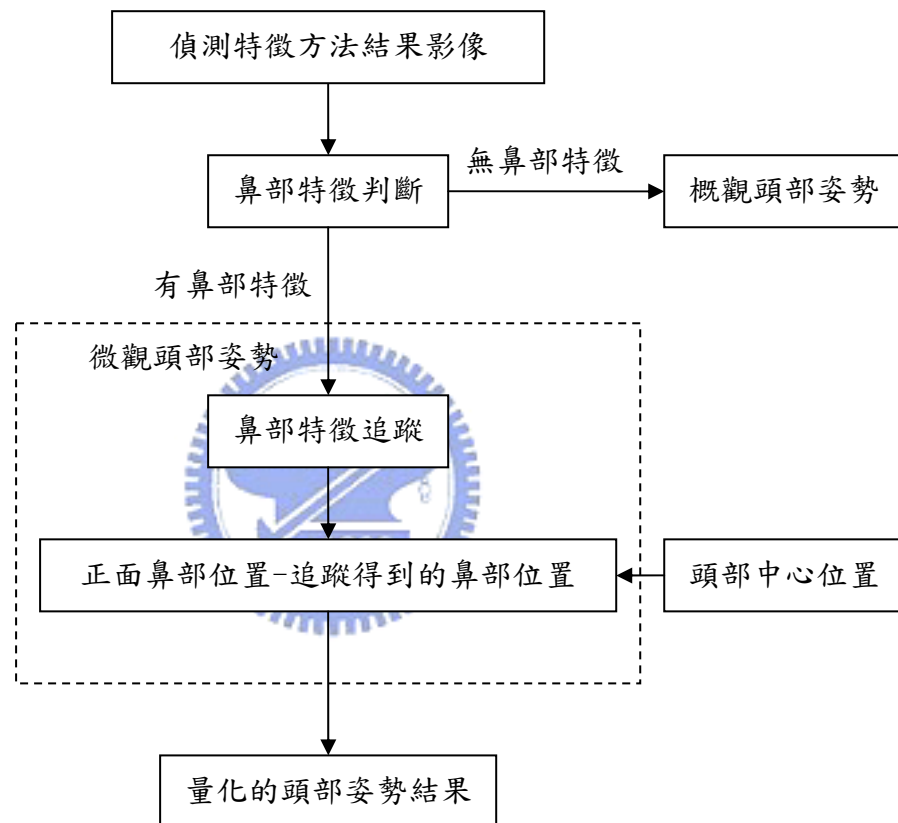
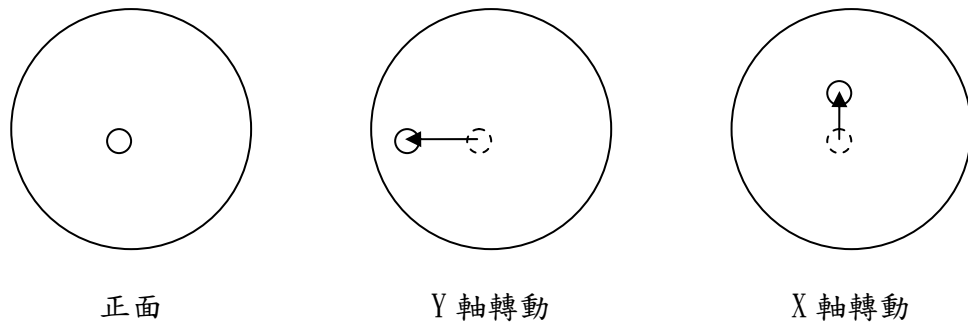


圖 33：近距離姿勢辨識流程圖

我們利用 3.4 小節所敘述的鼻部特徵追蹤演算法找尋鼻孔位置。在本研究中，正面鼻部位置可以利用頭部中心位置推算獲得，或是假設系統已經獲得正面鼻孔位置。

正面鼻部位置與追蹤到的鼻部位置相減則會形成一個方向向量，此方向向量大小除以頭部高度則得到量化的頭部姿勢。圖 34 可看出向量與頭部姿勢的關係。



⊙ : 鼻孔位置 → : 向量方向

圖 34：頭部姿勢與向量的關係



第四章 實作過程與結果討論

4.1 實驗說明

本系統首先利用移動資訊將人體與背景區隔出來，在此結果之下，再針對頭部區域做姿勢分析。整體系統方法的設計是依下列三項作為主要的考量：一、穩定且快速的追蹤頭部區域，二、偵測特徵方法的穩定性與通用性，三、姿勢辨識的正確性。底下我們會針對這些項目做實驗分析與結果討論。

表 1：本研究使用的硬體設備：

| | |
|-------|--------------------------|
| 攝影機 | 羅技 QuickCam Zoom |
| 中央處理器 | Intel Pentium M 1.50 GHz |
| 記憶體容量 | 512 MB RAM |

4.2 頭部偵測與追蹤結果

首先針對兩類測試環境下進行頭部的偵測與追蹤。一類是無物體遮住人體的測試環境下，另外一類則是物體遮住小部分人體的測試環境下。每個實驗執行時，我們皆採取固定視角的方式，並將出現在攝影機視野內的頭部標明出來。

4.2.1 無物體遮住人體的場景下

無物體遮住人體的測試環境，主要測試人體是否能從場景中區分出來。因此，測試環境會從簡單環場到複雜環場。經過頭部偵測的演算法，我們會預估人體頭部位置。

簡單環場結果如圖 35：



圖 35：人體擷取與頭部偵測於簡單場景

複雜環場結果如圖 36 與圖 37：



圖 36：人體擷取與頭部偵測於複雜場景 1

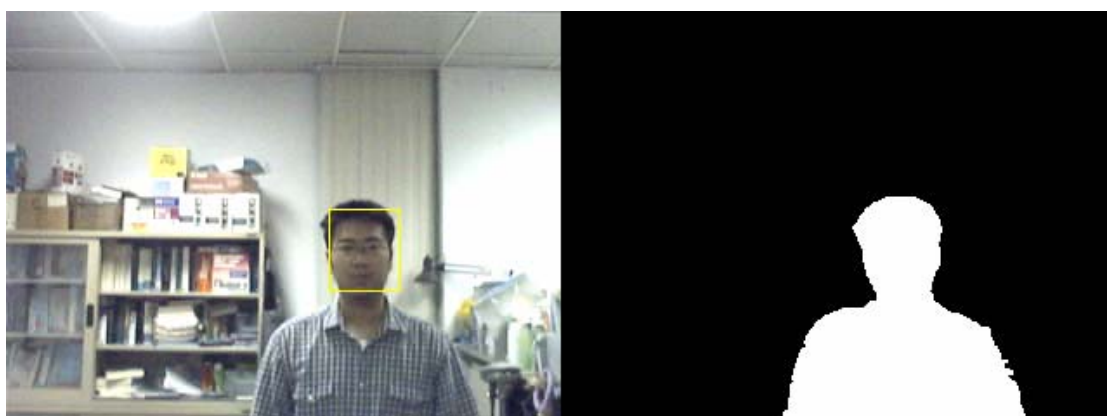


圖 37：人體擷取與頭部偵測於複雜場景 2

經過簡單場景與複雜場景的測試過後，由圖 35~圖 37 所示，本系統的頭部偵測演算法可以在無物體遮住人體的環境下，穩定且有效的偵測出頭部。之後，本系統將可以針對偵測出來的頭部區域進行一連串的頭部分析，如：偵測特徵方

法、臉部特徵分析、鼻孔位置追蹤與頭部姿勢判斷。

4.2.2 物體遮住小部份人體的測試環境下

在此小節，主要是驗證頭部偵測演算法能夠在靜態物體遮住人體的環境下，偵測出頭部位置。經過本系統的頭部偵測演算法，我們會將會將頭部標明出來或是將沒有遮住的頭部標明出來。圖 38 與圖 39 為測試的結果。



圖 38：有物體遮蔽的情形下頭部偵測結果 1



圖 39：有物體遮蔽的情形下頭部偵測結果 2

經過這樣的測試環境，本系統的頭部偵測演算法可以不受物體遮住頭部區域情況的影響。在未來我們將針對這樣的情況，對頭部姿勢進行判斷，或是進行臉部重建的工作。

4.3 臉部特徵處理效果

首先，我們先對人臉資料庫所提供的影像做處理，主要是要測試本系統所提的偵測特徵方法是否對於普遍的人臉影像有同樣的效果。

底下是使用 ORL 人臉資料庫(The ORL Database of Faces)測試的結果，如圖 40：



圖 40：ORL 人臉資料庫的結果

底下是使用 UMIST 人臉資料庫測試的結果如圖 41：



圖 41：UMIST 人臉資料庫的結果

經過人臉資料庫測試之後，顯示出偵測特徵方法可以運用在大部分的人臉影像上，並非針對特定族群的人或是場景進行臉部分析。因此，在近距離拍攝情況下，我們可以利用偵測特徵方法後的結果，進行臉部特徵追蹤，如：雙眼位置與鼻孔位置，在中距離拍攝的情況下，系統可以利用皮膚中心位置，達成頭部姿勢判斷的分析。

4.3.1 近距離拍攝下擷取特徵

近距離拍攝的測試環境下，我們主要是要測試系統是否能夠即時的反應與準確的追蹤雙眼與鼻孔位置。並測試在各種轉動過程中，雙眼位置與鼻孔位置會隨著頭部姿勢作相對應的改變。圖 42 為各種頭部姿勢下所追蹤到的鼻孔位置。圖 43 為系統所追蹤到的雙眼位置。



圖 42：在不同姿勢下的鼻孔位置



圖 43：在不同姿勢下的眼睛位置

在實作的過程中，我們發現在攝影機無法正確捕捉與追蹤臉部的情況下，系統無法利用偵測特徵方法突顯出臉部特徵，例如：眼睛。我們可以從圖 44 觀察出，當頭部抬高到某種角度後，攝影機無法正確捕捉雙眼位置，導致偵測特徵方法無法突顯眼睛位置。在經過各種頭部姿勢的測試下，我們發現鼻孔位置穩定性較高，所以在近距離拍攝情況下，本系統利用鼻孔位置來判斷頭部姿勢。



圖 44：抬頭姿勢與偵測特徵方法的結果

4.3.2 中距離拍攝下擷取特徵

在遠距離拍攝的測試環境下，通常攝影機捕捉的畫面不太能定位出雙眼位置與鼻孔位置。經過偵測特徵方法後，本系統能夠將頭髮部份與皮膚區分出來。因此，在此狀況之下，本系統利用區分出來的皮膚資訊做頭部姿勢的判斷。圖 45~圖 48 是中距離拍攝的測試結果。



圖 45：中距離抬頭姿勢下偵測特徵方法的結果



圖 46：中距離向右轉動姿勢下偵測特徵方法的結果

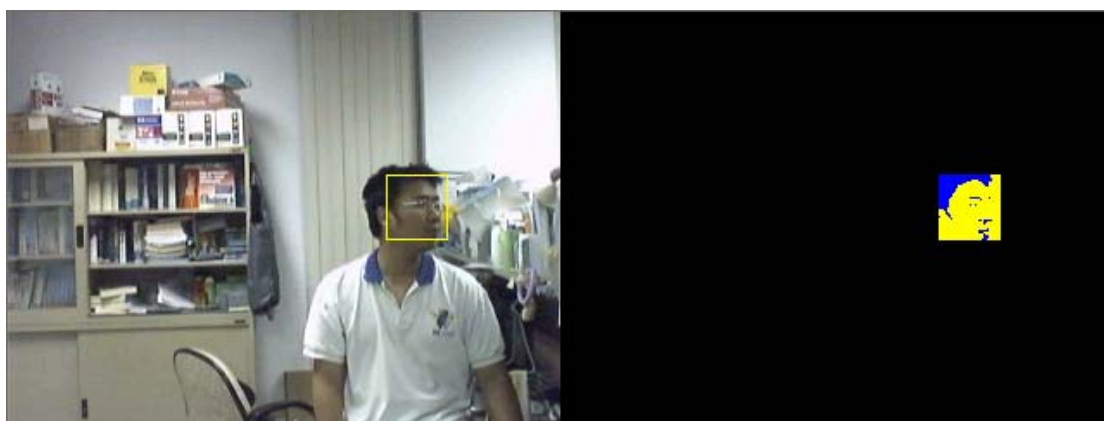


圖 47：中距離向左轉動姿勢下偵測特徵方法的結果



圖 48：中距離低頭姿勢下偵測特徵方法的結果

4.4 姿勢辨識結果

在經過偵測特徵方法後，中距離拍攝環境下，我們使用概觀的頭部資訊來判斷頭部是否向左或向右轉動。在近距離拍攝環境下，我們使用鼻孔位置來判斷頭部是否有左右轉動或是上下轉動的情形。

底下我們會分別在近距離與中距離的環境下進行測試並顯示姿勢判斷的結果。4.4.1 小節討論在固定攝影視角與近距離拍攝的環境下，微觀頭部姿勢的辨識及實驗結果。另外，4.4.2 小節討論在固定攝影視角與中距離拍攝的環境下，概觀頭部姿勢的辨識及實驗結果。

4.4.1 微觀頭部姿勢辨識

在近距離的拍攝環境下，攝影機所擷取的影像經過頭部偵測與偵測特徵方法後，本系統透過臉部特徵點偵測演算法，擷取臉部特徵位置。並利用臉部特徵位置來判斷頭部姿勢。

4.3.1 小節探討出因為雙眼特徵在某些頭部姿勢下無法被攝影機順利的捕捉，使得系統的偵測特徵方法無法順利的突顯出雙眼特徵。因此，在本研究中利用鼻孔位置當作特徵點，系統偵測並追蹤鼻孔位置來判斷頭部姿勢及頭部運動。圖 49~圖 56 清楚的顯示本系統能夠區分頭部的各類運動包括有上下與左右轉動，下方數值則表示出鼻孔位置與頭部中心位置相減並量化後的結果。另外，我們將頭部姿勢分為正面、抬頭、低頭、左側、右側、左斜側與右斜側七種姿勢，

並將其數值區間顯示於表 2。

表 2 中的數值並非連續的範圍，這是由於人為的觀感中，姿勢的判別是要超過某一定數值門檻才會呈現出來。例如：抬頭的 y 軸數值與正面的 y 軸數值並非連續的，這是因為在感知上，抬頭的動作需要超過某一定範圍才會被人認定為抬頭的姿勢。

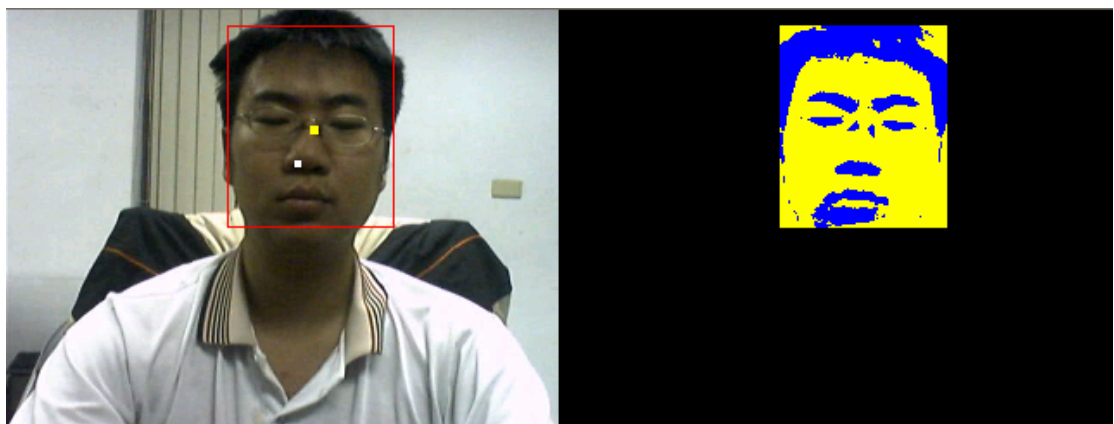
| 姿勢 | 正面 | 抬頭 | 低頭 | 右側 | 右斜測 | 左側 | 左斜側 |
|-----|--------|--------|--------|---------|---------|-------|-------|
| X 軸 | -10~10 | -10~10 | -10~10 | -25~-30 | -15~-20 | 25~30 | 15~20 |
| Y 軸 | 15~25 | -5~10 | 30~35 | 15~25 | 15~25 | 15~25 | 15~25 |

表 2：頭部姿勢與數值對應表



-4 18

圖 49：正面姿勢下與正面姿勢量化數值



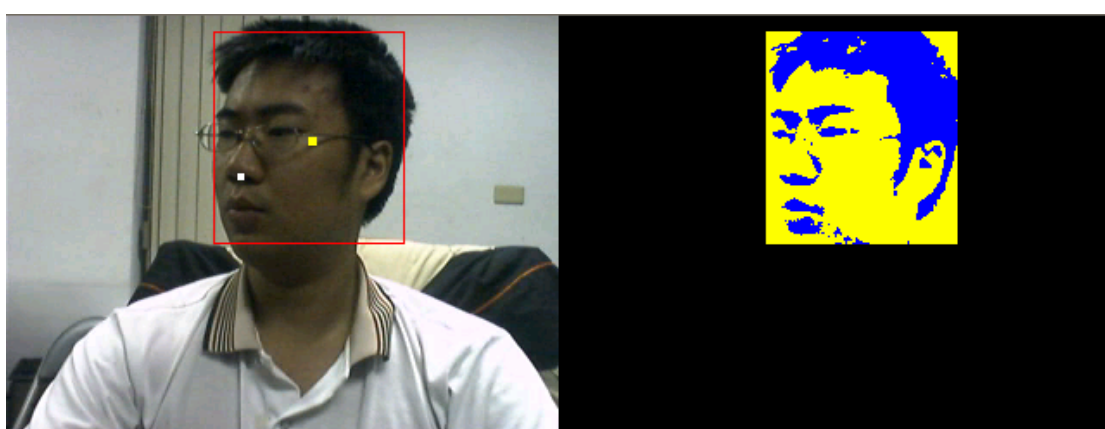
-7 22

圖 50：正面姿勢下與正面姿勢量化數值



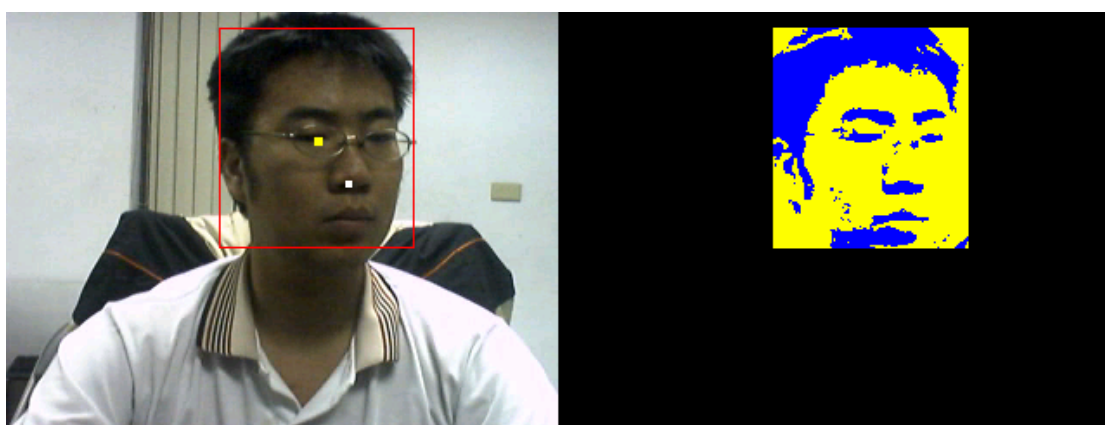
-18 18

圖 51：右斜側姿勢下與右斜側姿勢量化數值



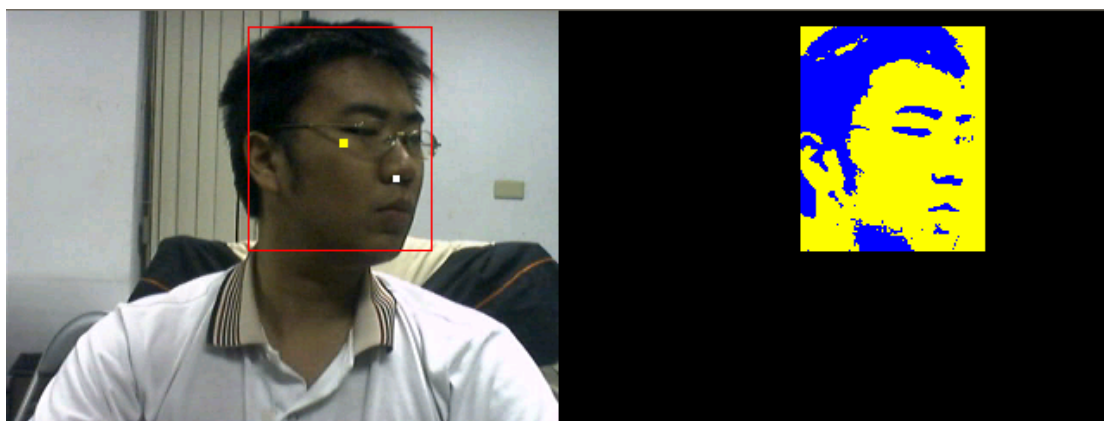
-31 19

圖 52：右側姿勢下與右側姿勢量化數值



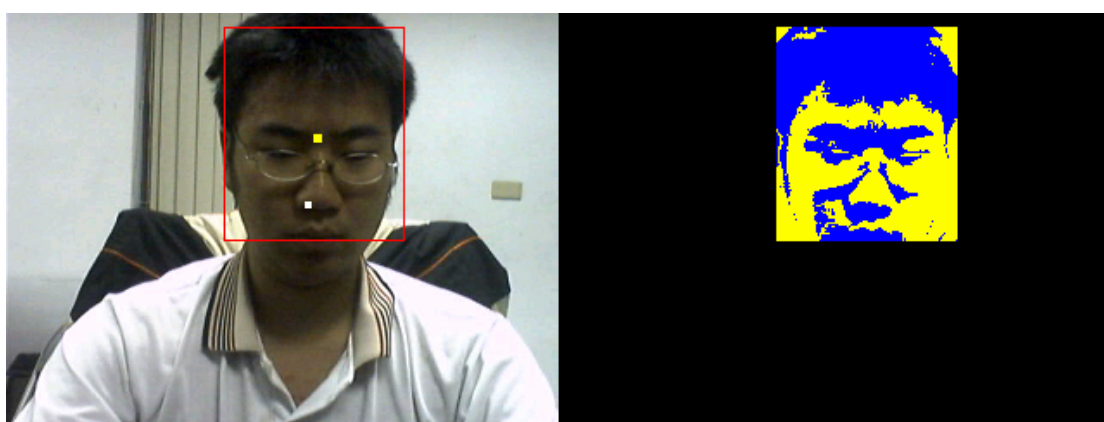
17 16

圖 53：左斜側姿勢下與左斜側姿勢量化數值



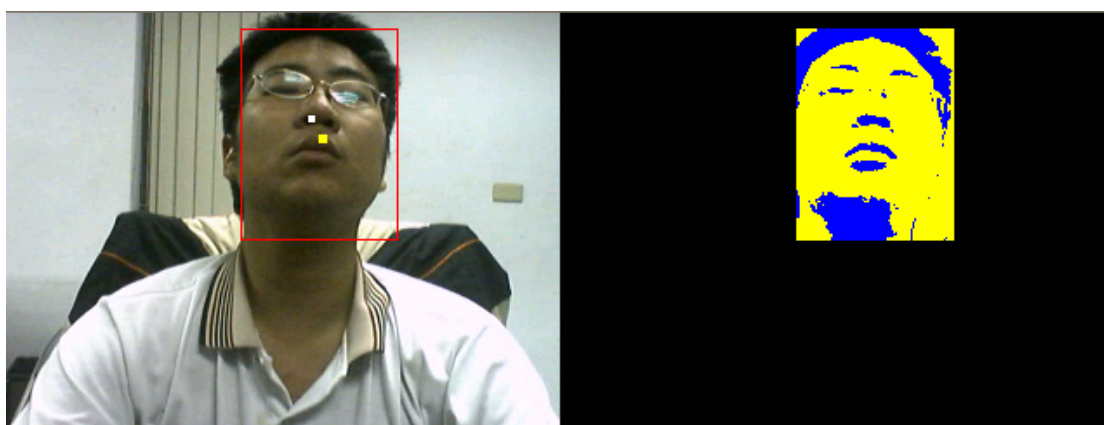
33 23

圖 54：左側姿勢下與左側姿勢量化數值



-5 34

圖 55：低頭姿勢下與低頭姿勢量化數值



-5 -5

圖 56：抬頭姿勢下與抬頭姿勢量化數值

4.4.2 概觀頭部姿勢辨識

在中距離的拍攝環境下，攝影機所擷取的影像經過頭部偵測與偵測特徵方法後，本系統利用概觀的頭部資訊來判斷頭部姿勢。

首先計算出膚色中心位置，之後與頭部中心位置相減，所得到的數值即可用來判斷概觀的頭部姿勢。圖 57~圖 59 顯示出各種頭部姿勢與數值結果。因為中距離拍攝的情況下會造成頭部資訊減少，目前只作左右轉動的判斷，而 X 軸的數值即代表轉動趨勢。另外，我們將頭部姿勢分為正面、左側、右側、左斜側與右斜側五種姿勢，並將頭部姿勢相對應的量化的結果做成表 3。

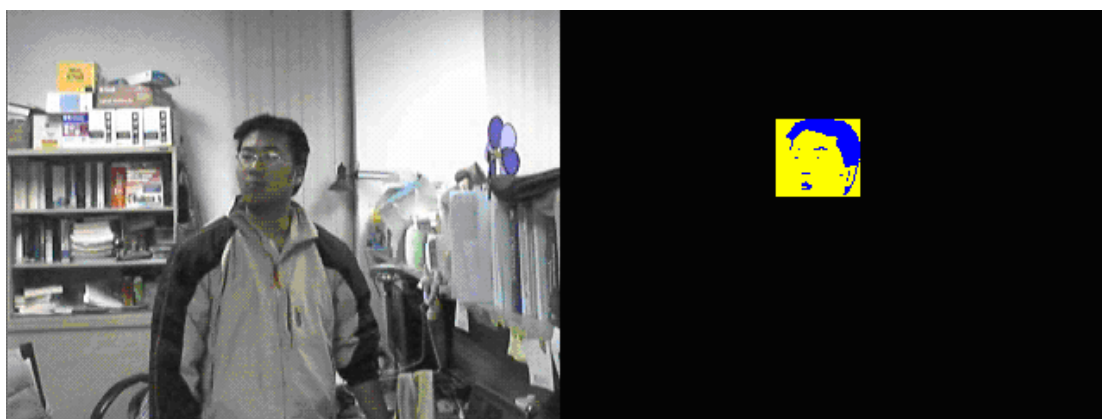
| 姿勢 | 正面 | 右側 | 右斜測 | 左側 | 左斜側 |
|-----|------|---------|--------|-------|------|
| X 軸 | -3~3 | -10~-15 | -4~-10 | 10~15 | 4~10 |

表 3：頭部姿勢與數值對應表



2 69

圖 57：正面姿勢下與正面姿勢量化數值



-5 67

圖 58：右斜側姿勢下與右斜側正面姿勢量化數值



8 67

圖 59：左側姿勢下與左側正面姿勢量化數值



第五章 結論

5.1 研究總結

在本研究中，我們提出下列三種特點，一、人體外圍輪廓，二、概觀頭部姿勢判斷，三、微觀頭部姿勢判斷。

我們利用動態資訊捕捉人體外圍輪廓，再利用人體外圍輪廓減少搜尋範圍並快速的找尋頭部區域。在站立或行走的姿勢下，經過人體外圍曲線分析後，我們可以獲得頭部位置與肩膀位置的推測，我們可以進一步利用此資訊來進行驅體與四肢的運動分析。

概觀頭部姿勢的判斷，可以加速臉部特徵搜尋演算法。在獲得概觀頭部姿勢後，可以減少頭部模型搜尋的範圍。另外，此項資訊可以改善監視系統的效率，使得監視系統可以更精確的推算出人體行徑方向，並可有效的擷取畫面與節省記憶空間。微觀頭部姿勢判斷可提供更仔細的頭部動作，可讓有變焦功能的監視系統獲得更細部的資訊。



5.2 未來工作

本研究提出以人體外圍輪廓減少搜尋範圍與快速的找尋頭部區域，在這各方法下，要如何更準確的將人體外圍輪廓取出，讓之後的分析更精確就是一個需要更深度研究的課題。除此之外，中距離或是遠距離的頭部姿勢推測是否更應該擷取運動資訊在配合偵測特徵方法所擷取出的特徵。利用這兩個資訊更準確推測人體行徑方向澤是需要再做更進一步的研究。

參考文獻

- [1] Ji, Q., Yang, X, “*Real-time eye, gaze, and face pose tracking for monitoring driver vigilance,*” *Real-Time Image*, Volume 8, Issue 5, pages 357-377, October 2002.
- [2] Q. Ji and R. Hu, “*3D face pose estimation and tracking from a monocular camera,*” *Image and Vision Computing*, Volume 20, Issue 7, pages 499-511, 2002.
- [3] B. Yegnanarayana, Anil Kumar sao, B. V. K. Vijaya Kumar, and Marios Savvides, “*Determination of Pose Angle of Face Using Dynamic Space Warping (DSW),*” *International Conference on Information Technology: Coding and Computing*, Volume 1, page 661, 2004.
- [4] Jilin Tu, Thomas Huang, and Hai Tao, “*Accurate Head Pose Tracking in Low Resolution Video,*” In *International Conference, FGR*, 2006
- [5] Jilin Tu, Thomas Huang, Hai Tao, “*Face as Mouse Through Visual Face Tracking,*” In *2nd Canadian Conference on Computer and Robot Vision*, pages 339-346, 2005.
- [6] Zitnick, C.L., Gemmell, J., and Toyama, J, “*Manipulation of video eye gaze and head orientation for video teleconferencing,*” *Microsoft Research, Technical Report, MSR-TR-99-46*, 1999.
- [7] Hai Tao and Thomas Huang, “*Explanation-based Facial Motion Tracking Using a Piecewise Bezier Volume Deformation Model,*” In *IEEE Conference on Computer Vision and Pattern Recognition*, Volume 1, pages 611-617, 1999.
- [8] Hideo Saito, Akihiro Watanabe, Shinji Ozawa, “*Face Pose Estimating System Based on Eigen Space Analysis,*” In *Proc. ICIP*, pages 638-642, 1999.
- [9] Shumeet Baluja, Mehran Sahami, and Henry A. Rowley, “*Efficient Face Orientation Discrimination,*” *International Conference on Image Processing*, pages 589-592, 2004.
- [10] <http://www.cl.cam.ac.uk/Research/DTG/attarchive/facedatabase.html>
- [11] <http://images.ee.umist.ac.uk/danny/database.html>

- [12] Ben Yip, Jesse S. Jin, "Pose determination and viewpoint determination of human head in video conferencing based on head movement," In Proc. MMM, page 130, 2004.
- [13] Kabakli, T. Celik, T. Dernirel, H. Ozkaramanli, H. Uyquroqlu M. Kondoz, A.M, "*Robust Face Tracking Using Color Segmentation Aided With Connected Components Labeling and Morphology*," In Proc. Signal Processing and Communications Applications Conference, 2004.

