

中華大學

資訊工程學系

人臉追蹤與嘴部辨識遊戲應用

Face Tracking and Mouth Detection with the gaming applications

專題生

朱必君

CHU, PI-CHUN

中華大學

資訊工程學系

b09702180@chu.edu.tw

黃育甫

HUANG, YU-FU

中華大學

資訊工程學系

b09702063@chu.edu.tw

曾新雅

TSENG, HSIN-YA

中華大學

資訊工程學系

b09702306@chu.edu.tw

指導教授：黃雅軒

摘要

現在是一個資訊飛躍的時代，在這個資訊化的社會，人人都追求更佳的便利性，而在市面上出現越來越多不同種類的資訊化產品，如：體感互動遊戲、觸碰式遊戲、視訊辨識系統...等許多視訊類產品，這些例子說明了資訊娛樂是目前極為熱門的行業，因此我們朝著視訊類型去做研究。

本研究主要的目的為在於能使用此系統，對另一個系統或是程式做出反應，達成下指令與表達所需目的的動作，為達到此目的，我們首先利用攝影機取得影像，並且取得臉部與嘴部位置的資訊，再使用其資訊去追蹤人的臉部，最後使用一些影像處理的技術，對嘴部去做開闔的辨識，以達成在其他系統或程式中下達指令或做出動作，最後我們以一簡單的遊戲程式來進行系統的實驗，並驗證此系統可達成所需目的。

目錄

摘要	I
目錄	II
圖目錄	III
第一章	1
1-1 前言	1
1-2 系統流程	2
第二章	4
2-1 嘴部偵測	4
2-2 嘴部分析	6
一、 邊緣偵測法	6
二、 色彩空間辨識法	8
三、 特徵值計算(LBP)	9
四、 灰階運算法	10
第三章	12
3-1 嘴部辨識結果分析	12
3-2 遊戲成果展示	14
遊戲介紹:	15
遊戲架構:	17
第四章	18
一、 系統開發技術	18
參考文獻	18

圖目錄

圖 1.1 人臉追蹤與嘴部辨識流程圖.....	2
圖 2-1-1 分類器使用畫面.....	5
圖 2-1-2 臉部區域劃分示意圖.....	7
圖 2-2-1 嘴部擷取圖.....	7
圖 2-2-2 上圖嘴部切割後.....	8
圖 2-2-3 左圖邊緣化後.....	8
圖 2-2-4 測試成果較差原圖.....	9
圖 2-2-5 上圖嘴部切割後.....	9
圖 2-2-6 左圖邊緣化後.....	11
圖 2-9 良好框選可能圖.....	11
圖 3-1 系統介面.....	13
圖 3-2 閉嘴辨識圖.....	13
圖 3-3 張嘴辨識圖.....	14
圖 3-4 遊戲介面圖.....	16
圖 3-5 遊戲血量變更畫面.....	16
圖 3-5 遊戲蛋糕移動畫面.....	17

第一章

緒論

1-1 前言

在過去數年間，已有許多監視系統的發展，主要是人為的操縱或是消極的影像記錄，並不能即時的作判斷或提供資訊，但是伴隨近年來，數位影音的發展，以及影像辨識成為熱門的話題，由於電腦的技術進步，在加上為了滿足更多的需求，不論在做任何的事情上，加上即時的影像配合，都能有更好的體驗。

研究內容主要以兩部分組成，一個部分是人臉追蹤(Face Tracking)，另一部分是嘴部動態辨識(Dynamic Mouth Detection)，人臉追蹤是使用中華大學 VIP 實驗室所提供的人臉追蹤技術，再由我們將研究嘴部動態的辨識，此研究的主要目的在於希望能以人臉的追蹤與嘴部辨識，做出與遊戲、生活或是精密的系統...等其他內容的結合，讓這些系統使用上更為便利、有趣，未來更可努力達到唇形辨識，可以讓使用口語的人更容易溝通，使生活更加的美好。

1-2 系統流程

人臉追蹤與嘴部動態偵測系統包含了數位影像處理、取樣控制等，此系統可以偵測與追蹤嘴部的動作，了解目前的行為，流程架構如下

圖 1.1 所示。

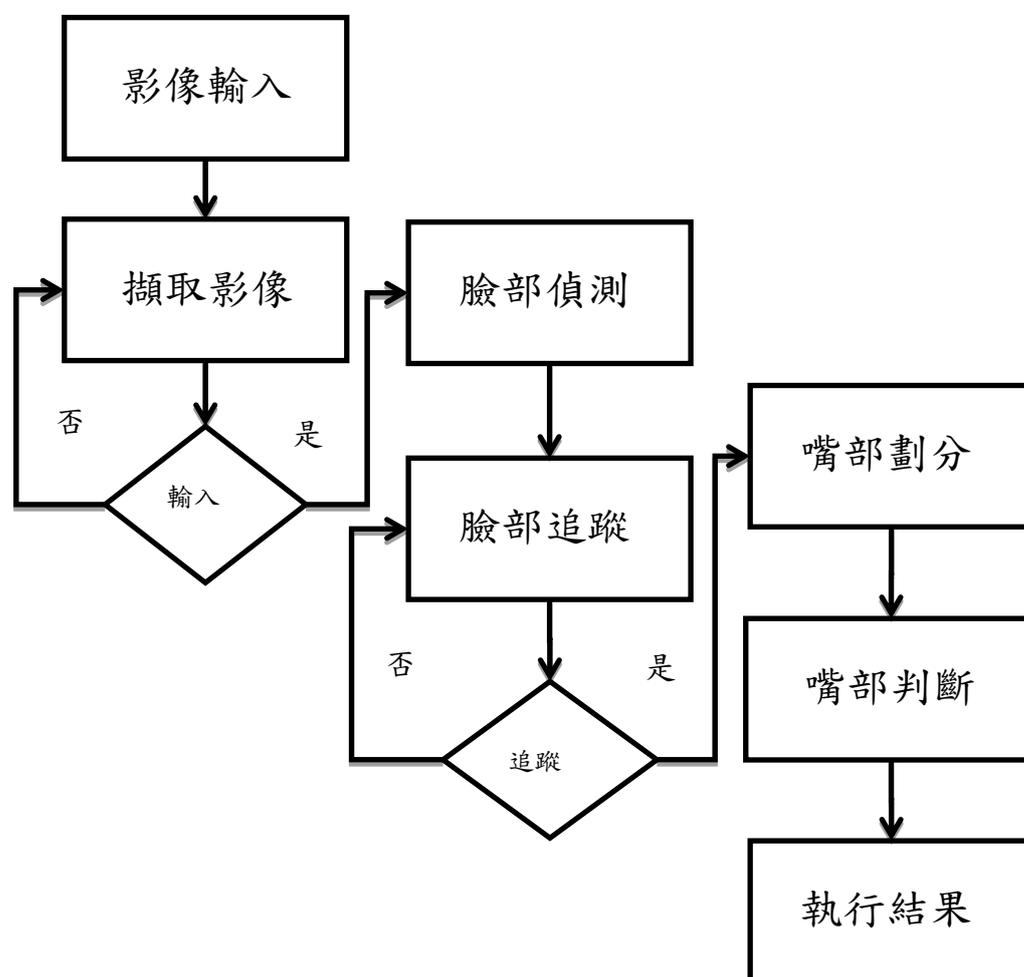


圖 1.1 人臉追蹤與嘴部辨識流程圖

系統流程分別敘述如下：

一、擷取影像：主要是從輸入裝置取得連續影像作接下來的處理，也包含影像儲存與前處理。常用的輸入裝置為網路攝影機等，裝置須固定於某處，由人進行移動。

二、臉部偵測：此研究與中華大學 VIP 實驗室合作，使用他們所提供之臉部追蹤動態連結檔(.dll)，去執行臉部偵測的部分。

三、臉部追蹤：此研究與中華大學 VIP 實驗室合作，使用他們所提供之臉部追蹤動態連結檔(.dll)，去執行臉部追蹤的部分。

四、嘴部劃分：對臉部做區域劃分。使用統計的資料對臉部做區域劃分，作為嘴部劃分的準則。

五、嘴部判斷：使用我們自行發展出的辨別方式，在嘴部偵測之後對嘴部去做張嘴或是閉嘴的判定。

六、執行結果：未來我們希望是可以使用在其他系統上，目前我們設計出一個簡易的遊戲軟體，會在成果中展示。

此遊戲使用 XNA 設計，後面將簡易介紹 XNA。

第二章

嘴部追蹤與分析

2-1 嘴部偵測

一、人臉取得

首先我們使用中華大學 VIP 實驗室所提供的動態連結檔取得臉部追蹤的位置。

二、嘴部取得

在嘴部取得部分，我們使用了二種方式去做測試，分為以下二種：1. 分類器法 2. 人臉區域劃分

(一) 分類器：

使用 OpenCv 提供的分類器去取得多個可能為嘴部的區塊，再來將取得可能區塊後再去做篩選。(圖 2-1-1)

(二) 人臉區域劃分：

使用人臉的比例，鼻尖的位置約在人臉的 $2/3$ (圖 2-1-2)，而經過我們測試，眼睛與鼻子是分類器最容易出錯的部分，以此使用這種比例分配方式能有效的排除多個可能錯誤的框選區域。

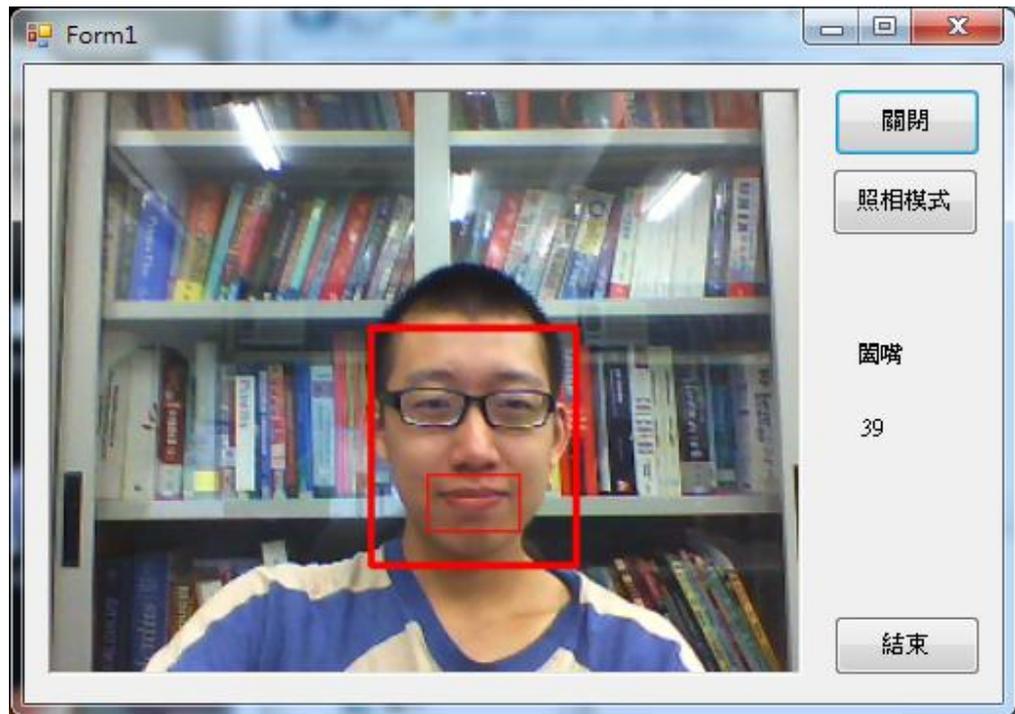


圖 2-1-1 分類器使用畫面

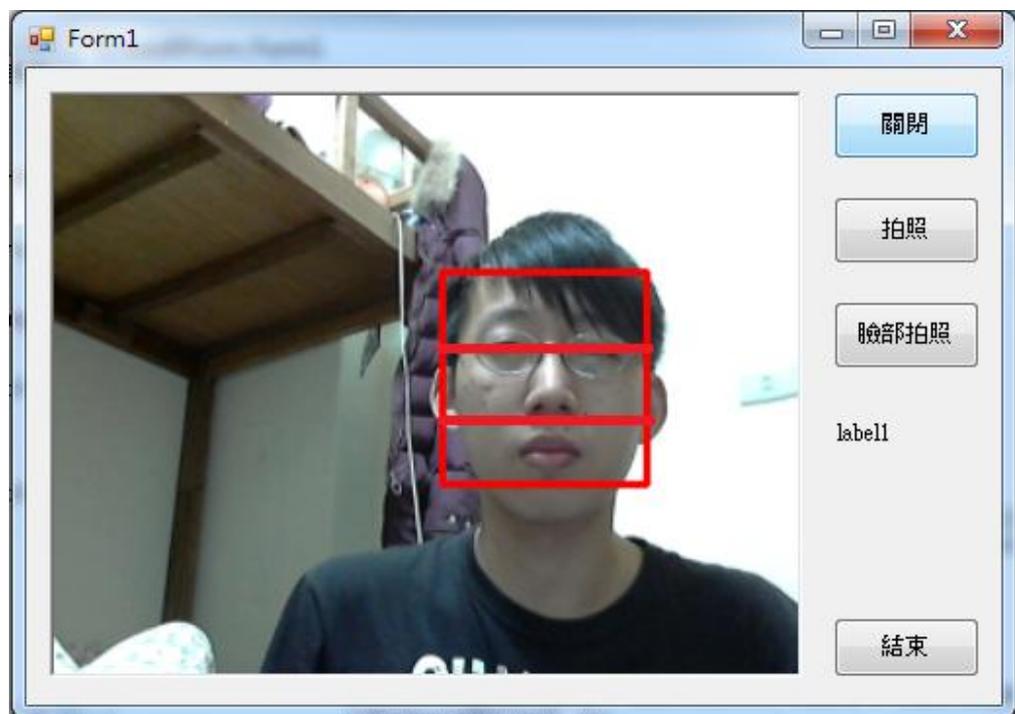


圖 2-1-2 臉部區域劃分示意圖

2-2 嘴部分析

在嘴部分析這部分，我們使用了四種方式去做測試，分為以下四種：1. 邊緣偵測法 2. 色彩空間辨識法 3. 特徵值計算 4. 灰階運算法。

一、邊緣偵測法

首先將原圖改變為灰階影像(圖 2-2-1)再擷取出嘴部區域(圖 2-2-2)，並將嘴部區域二值化後做邊緣偵測(圖 2-2-3)。

此動作是預計將唇形刻劃出來，我們在以上緣與下緣的距離做判斷的依據。但因光線不足或是每人臉部紋路深淺不同，會造成紋路過於複雜如圖(2-2-4)(2-2-5)(2-2-6)。

邊緣偵測法使用以下三步驟：

(一) 擷取嘴部區域

此處是將原圖使用到 2-1 所說明的嘴部偵測方式，因而取得嘴部區域，再將其圖像擷取出來。

(二) 二值化

二值化為圖像分割的一種方法，在二值化圖像的時候把大於某個圖像**臨界灰度值**的像素灰度最大值設為灰度最大值，將小於這個值得灰度值設為灰度最小值，從此實現二值化。

(三) 邊緣偵測

此處邊緣偵測使用 Canny 的邊緣偵測處理。

OpenCV 提供的函式庫：

```
cvCanny(pImg, pCannyImg, 50, 150, 3);
```

Canny 的處理原理為：

1. 運用高斯濾波器在灰階影像，得到平滑的影像。
2. 應用微分濾波器計算邊緣長度和方向



圖 2-2-1 測試成果良好原圖

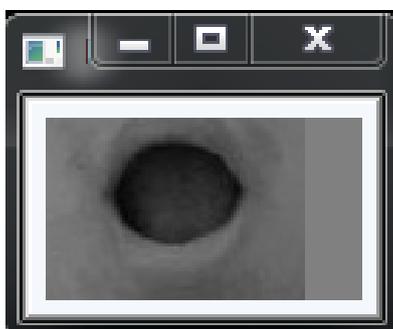


圖 2-2-2 上圖嘴部切割後

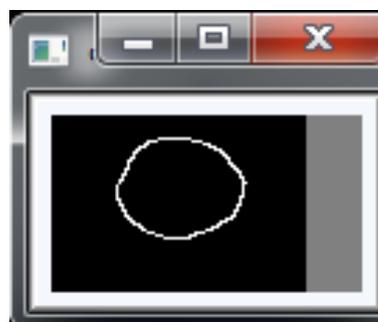


圖 2-2-3 左圖邊緣化後



圖 2-2-4 測試成果較差原圖



圖 2-2-5 上圖嘴部切割後



圖 2-2-6 左圖邊緣化後

二、色彩空間辨識法

首先我們將嘴部區塊的部分獨立區分出來，以 RGB 的方式去做擷取轉換到 YCbCr 色彩空間。

依照膚色範圍[1]

$$\begin{aligned}60 &\leq Y \leq 255 \\100 &\leq Cb \leq 125 \\135 &\leq Cr \leq 170\end{aligned}$$

在框選部分我們將不是膚色的部分判定為嘴唇內的非嘴部區域，以嘴部區域中非膚色範圍判別是否符合張嘴模型

三、特徵值計算(LBP)

LBP (Local Binary Partition)是一種用來描述區域紋理變化的特徵計算方式。

1	2	4
128		8
64	32	16

首先我們以 3X3 的遮罩運算[2]，權重方式由左到右，由上到下依照 2 的次方排列，如上圖，計算首張影像特徵值，當下使用者需為閉嘴狀態，在與後續影像特徵值比對，判斷是否為張嘴。

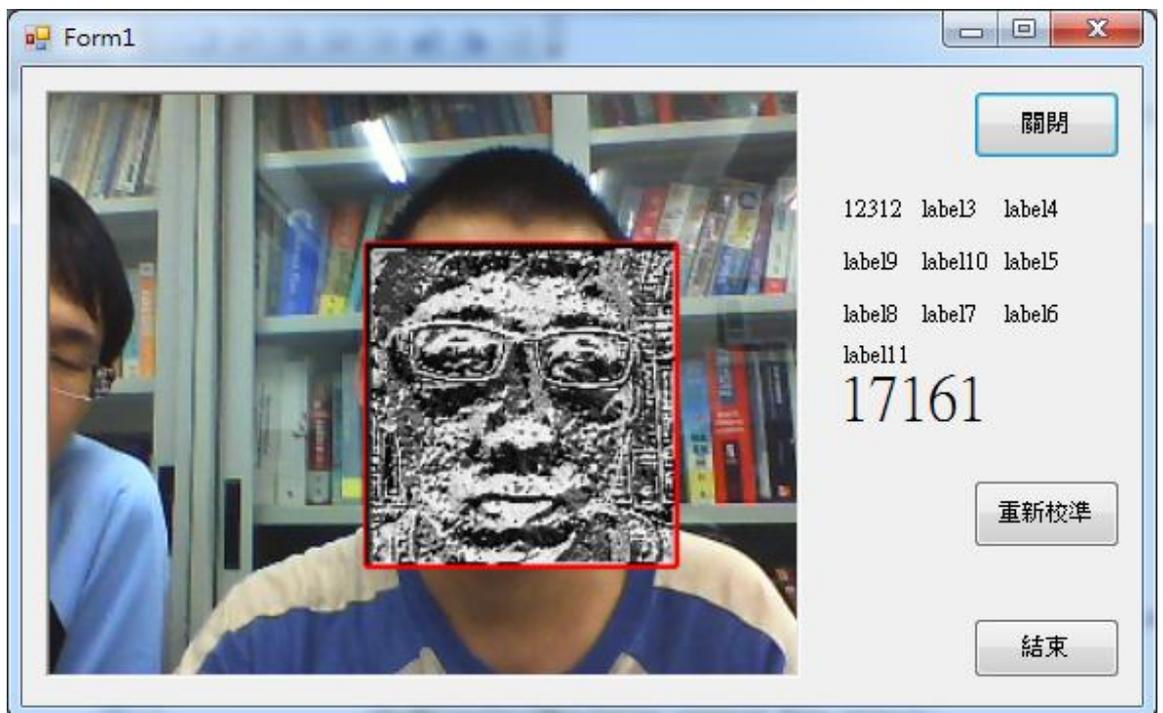


圖 2-2-7 LBP 運算後執行畫面

四、灰階運算法

此方法的目的為讓系統能在個人化的情況下執行，好處在於能更快速的符合環境與個人的狀況。首先使用者需為閉嘴狀態，以前幾張影像執行初始化，擷取嘴部區域最低灰階值，以最低灰階值與後續影像作判斷，若有大量的低數值產生則為張嘴狀態。

使用以下步驟：

(一) 灰階化

首先擷取臉部範圍中 RGB 值，使用

$$T = 0.299 * B + 0.587 * G + 0.114 * R$$

將全部範圍灰階化。

(二) 擷取嘴部區域

此處是將原圖使用到 2-1 所說明的嘴部偵測方式，因而取得嘴部區域，再將其圖像擷取出來。

(三) 結果判斷

以初始化影像取得**臨界灰度值**，在後續影像中，將小於這個值的灰度值設為 0 大於此值則為 255，方便於分辨結果(圖 2-2-8)，若為零的值大於某的判斷數量則判斷為張嘴情況。

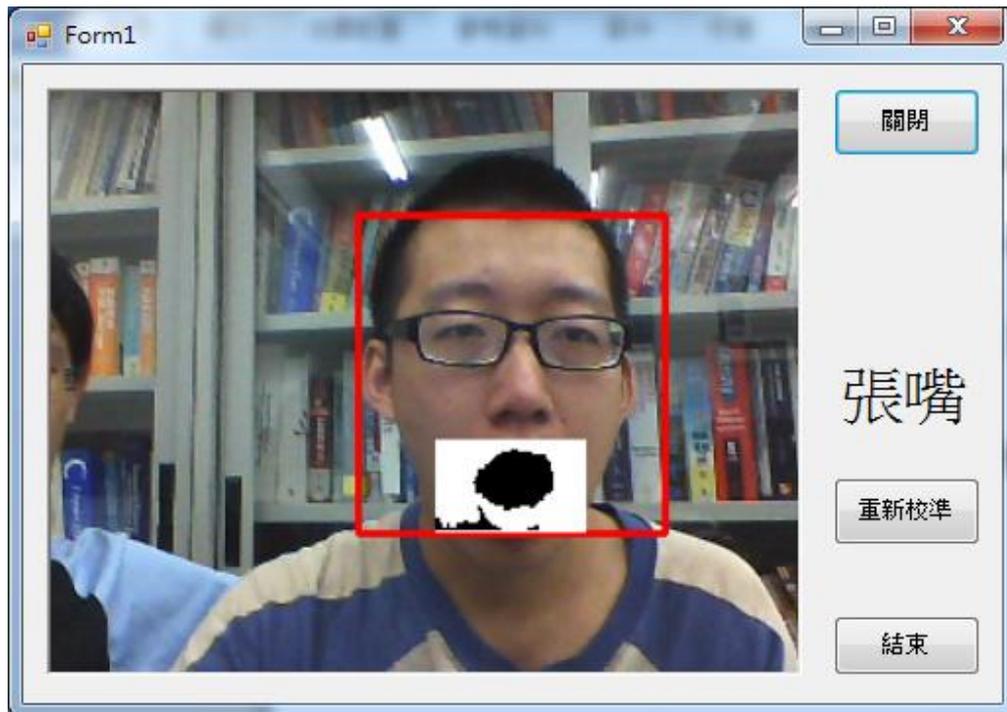


圖 2-2-8 灰階判斷張嘴情況

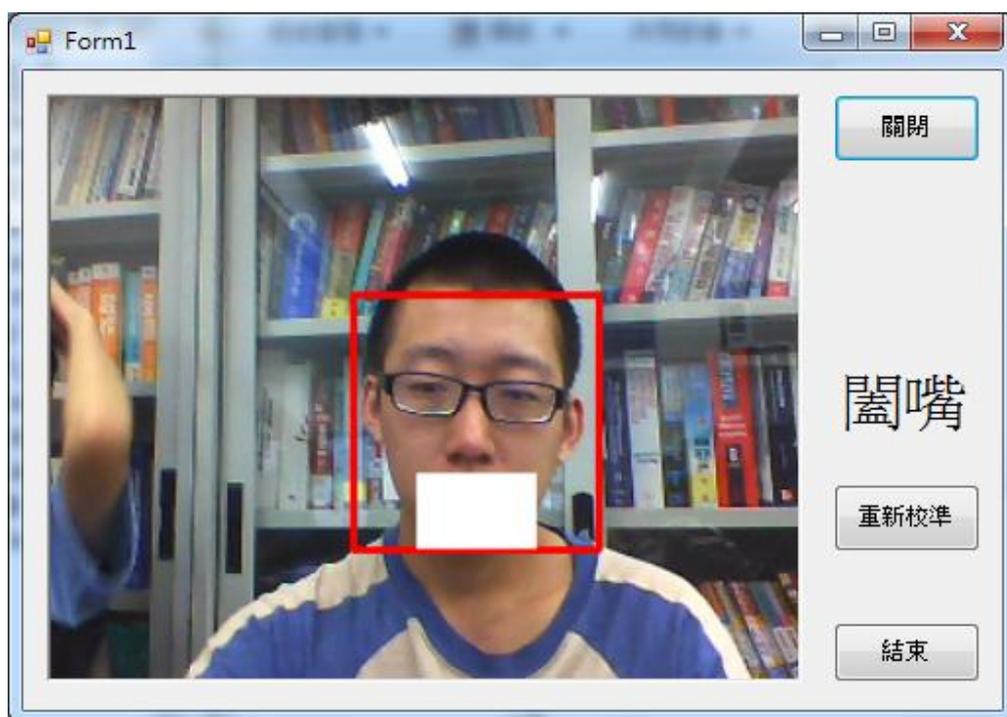


圖 2-2-9 灰階判斷闔嘴情況

第三章

結果與討論

3-1 嘴部辨識結果分析

(一) 實驗結果

臉部追蹤與嘴部動態偵測系統界面如圖 3-1 所示，紅色區域為啟動鈕，啟動後會開起攝影機的接收，在左方畫板顯示即時的圖像。藍色區域為目前的嘴部狀態，會在程式進行時即時的展示，黃色區域可關閉程式。

啟動系統後，系統會即時分析使用者的臉部與嘴部並且框選，使用者可以自行移動位置，系統會追蹤臉部並取得嘴部位置，在系統啟動時使用者先以閉嘴狀態正對螢幕，右方則會顯示目前嘴部的狀態，如圖 3-2、3-3。

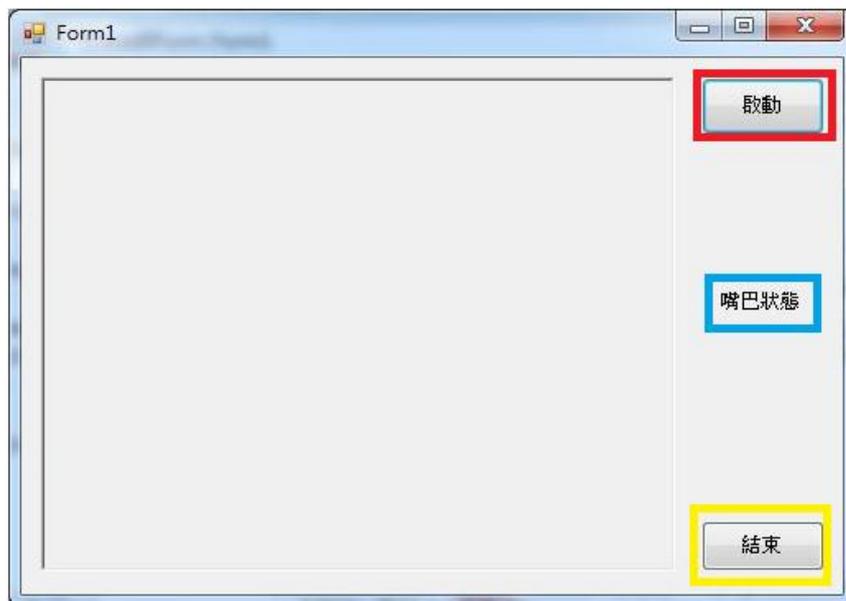


圖 3-1 系統介面

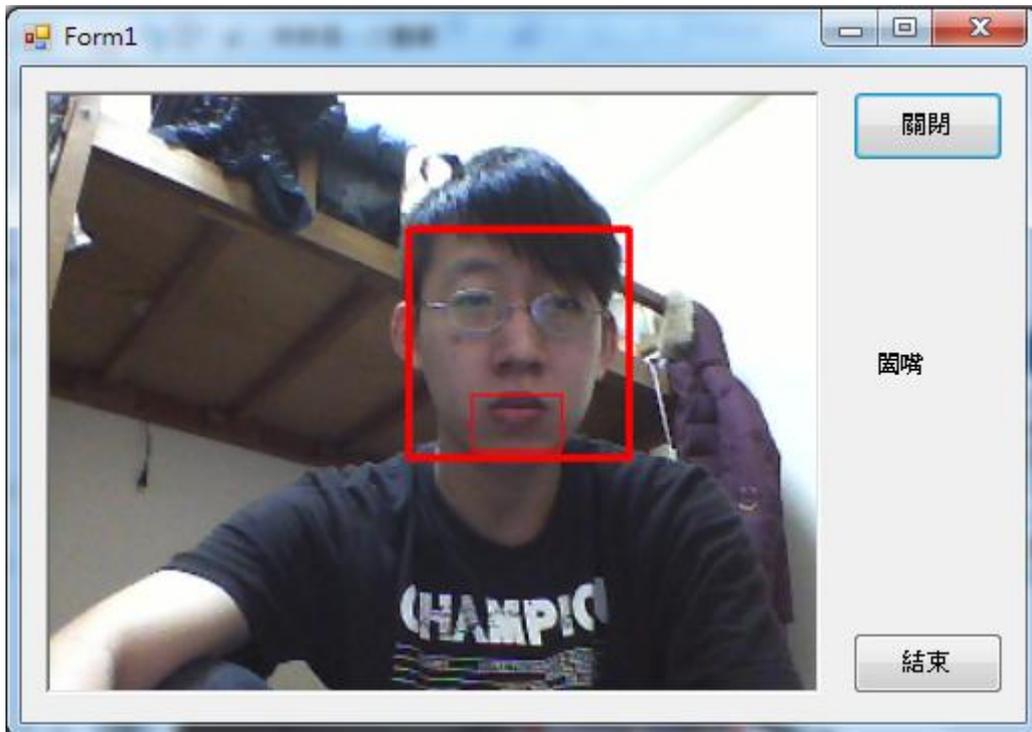


圖 3-2 閉嘴辨識圖

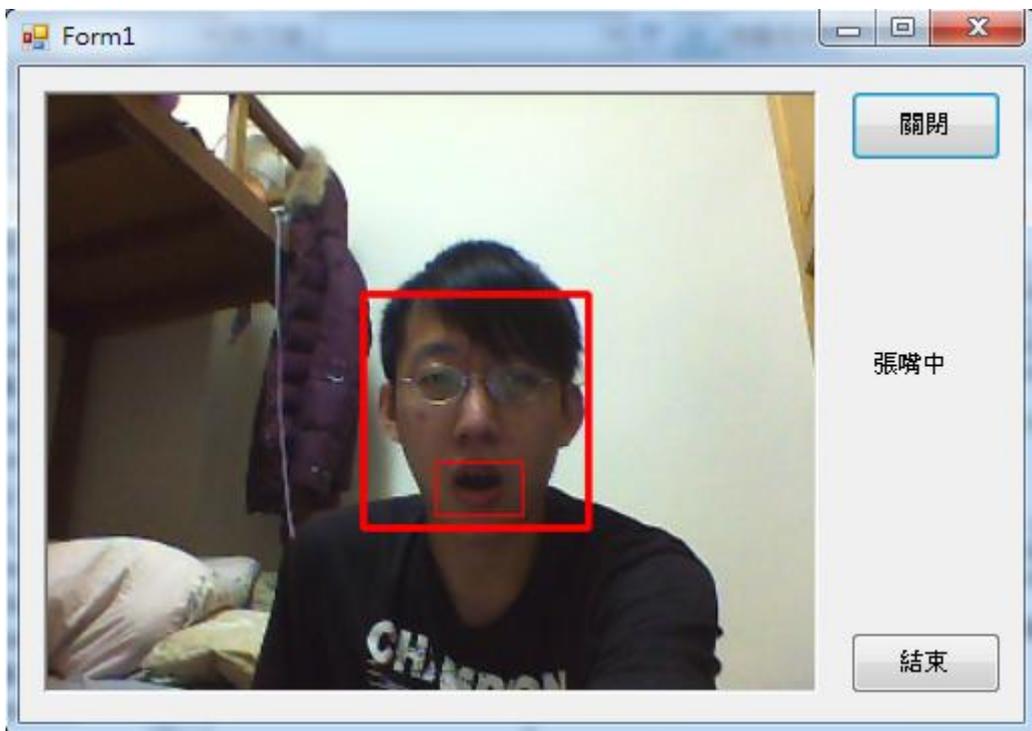


圖 3-3 張嘴辨識圖

3-2 遊戲成果展示

介紹 XNA :

由於我們的簡易遊戲是以 Visual C# Express 開發工具搭配 XNA 去做遊戲的開發，在此簡易介紹 XNA。

XNA 是微軟極力推動的次世代遊戲開發平台，搭配 Visual C# Express 開發工具。不僅開發環境完全免費，支援的遊戲平台亦可涵蓋 PC Windows、Xbox 360、以及微軟最新的 Zune 數位媒體播放器。對於想學習 2D / 3D 遊戲程式設計的初學者，是極佳的學習對象

在過去電視遊樂器的開發，是遊戲開發領域中非常專門的工作，一般只有專業的遊戲製作公司能做。直到 XNA Studio Express 版的出現，才在 360 平台上提供了業餘愛好者自己開發 360 遊戲的可能性。所以對電視遊樂器遊戲的開發，對於想要用自己的程式駕馭 Xbox 360 強大性能的程式師來說，這樣的渴望不再僅僅是夢想。

XNA 擁有幾個良好條件，跨平台開發、方便的入門套件、可提供 3D 遊戲設計，可使用 3D Max 設計模組使用。

由於以上幾個條件，所以我們選擇了 XNA3.0 版本來設計我們的遊戲，目前只是簡單的設計，為了配合良好的遊戲品質，當然也可以配合做更多的開發與使用。

遊戲介紹：

圖 3-4 為吃蛋糕遊戲界面圖，由圖我們可以看到紅色框框代表我們的使用者，它會依照我們臉部的移動去做移動，玩家可以使臉去觸碰蛋糕。藍色框框則是蛋糕，蛋糕會不停的由上方落下(圖 3-5)，若是臉部接近蛋糕且張開嘴巴，則蛋糕會被吃掉並消失，且分數將會增加，若是蛋糕掉出螢幕之外，則蛋糕會消失(圖 3-6)，綠色框框則是吃蛋糕的蟲，蟲會從螢幕下方爬上來，玩家碰到蟲分數就會減少，黃色框框代表目前分數及血量，後面會繼續應需求再設計遊戲結束情況。



圖 3-4 遊戲介面圖

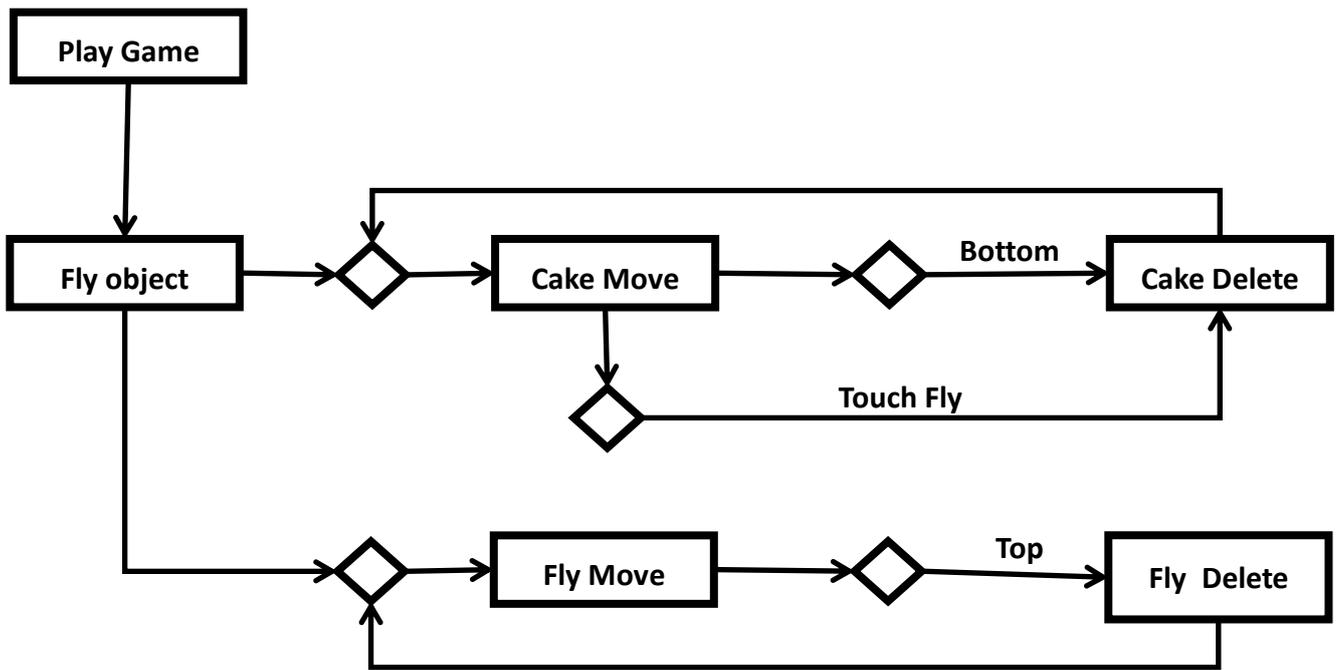
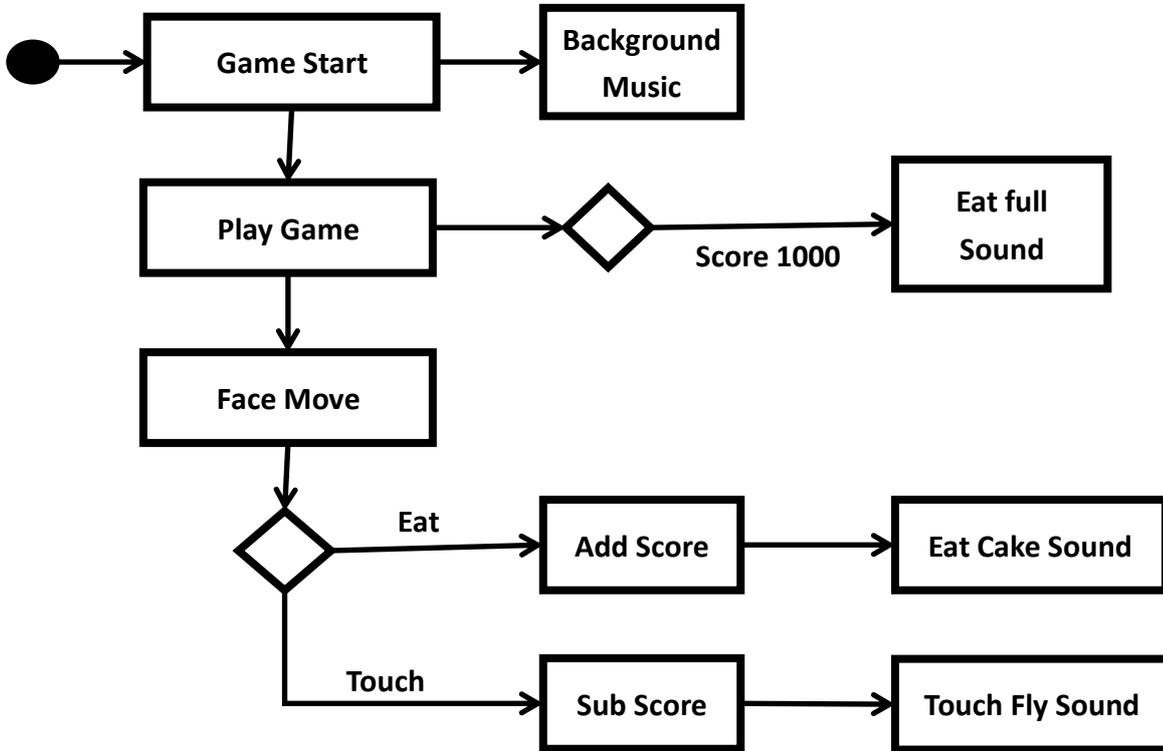


圖 3-5 遊戲開始畫面



圖 3-6 玩家吃到蛋糕會得到分數

遊戲架構：



第四章

未來工作

此研究未來改進有兩大方向，其一為多個辨識方法的結合，其二為辨識的改進。

一、系統開發技術

(一)多個辨識方法的結合

將使用的辨識方法由一種變成兩種或兩種以上一起使用，可依照個人化的要求改變，也可排除單一辨識方法錯誤的情況，以達到提升辨識能力的效果。

(二)辨識的改進

此部分會強調在邊緣偵測後的雜訊去除，由於輪廓明顯的出現原因多是因為光線不夠充足，造成臉部陰影出現過多，所以將嘗試以光線補償的方式，試圖將過暗的圖像變亮，去除大多數的輪廓雜訊，相信能大幅提升辨識成功率。

參考文獻

- [1]吳明衛，” Automatic Facial Expression Analysis System” ，國立成功大學，2003.6
- [2]劉佳諺，” A Robust Texture-Based Background Subtraction for Moving Object Detection in Video Sequences” ，義守大學，2007.6