

應用電腦視覺技術於電玩遊戲之設計與探討

Research and Development of Games Based on Computer Vision

林鴻斌* 陳禮揚**

國立台北科技大學電機工程系

* t5318080@ntut.edu.tw **t5318082@ntut.edu.tw

論文摘要

電玩遊戲廣受大眾的喜愛，是人們最方便的休閒娛樂之一，自從電玩遊戲面世以來，各種型態的遊戲不斷的推陳出新，這篇論文將討論把電腦視覺技術應用在電玩遊戲的方法，我們會討論電腦視覺的技術、系統的硬體架構和軟體架構，期望能開創一種全新的電玩遊戲型態。

關鍵詞：電腦視覺、體感遊戲、視訊攝影機、形狀分析、移動追蹤

Abstract

Many people like game. Game is convenient for people to recreation and relaxation. Since game appears in the world. Many type games are to weed through the old to bring forth the new. We discuss the method of using computer vision technique in games. We will discuss technique of computer vision, hardware structure and software structure in our system, and hope to create one new type game.

Keywords: Computer Vision, Physical Game, Webcam, Sharp Recognition, Motion Analysis

1. 前言

本研究的主要精神為「解放你的手，讓身體動起來」，大約 20 年前日本任天堂公司俗稱「紅白機」的家用型遊戲主機面世後，遊戲便走入了家庭，隨著硬體設備的進步和網路通訊的發達，遊戲畫面的精緻、透過網路全世界的人同樂，已經成為基本要求，近日任天堂公司的 Wii 主機上市，開創了一種新的遊戲型態，我們稱之為「體感遊戲」，玩家透過 Wii 搖桿操控角色時，可以用揮動搖桿的方式玩遊戲(例如棒球遊戲)，因為「體感遊戲」讓玩家和遊戲情境可以更進一步融合，所以帶起新一波的遊戲熱潮，但 Wii 還是需要搖桿才能操作遊戲，這樣的方式依然無法讓玩家跟遊戲完美的融合。

對現代人而言，遊戲不單單只是玩樂而已，它已經成為一種文化，變成生活裡密不可分的一部分，是許多人休閒、娛樂和解放壓力的選擇，有別於傳統遊戲「坐著玩」這種比較靜態，只動動手指頭玩遊戲的方式，本研究期望以一種用身體玩遊戲的方式，讓玩家與遊戲情境能完美的融合，結合運動、休閒、娛樂等元素，這比「體感遊戲」更進一步所以本研究稱之為「運動遊戲」。

許多人在玩遊戲的時候往往無法控制時間，常常一玩就是一整天，這樣長時間玩遊戲，對身體健康其實是很大的負擔，電視新聞也常常報導，有人因為廢寢忘食的玩網路遊戲，甚至連性命都送掉了。我們希望能改善這種情況，讓遊戲和運動融合一體，帶給大家一種全新的遊戲，或者也可以稱為一種新的運動方式。電腦視覺和遊戲結合有以下四種優點：

- ◆ **操控性：**本研究運用 **webcam** 視訊攝影機搭配電腦視覺、圖形識別和影像處理技術代替傳統的輸入控制設備，解放你的雙手，達成一種以人體為本的新操作方式，如圖 1 所示：



圖 1 以人體為本的操控方式

- ◆ **互動性：**以人體為本的控制方式，提昇了人和機器的互動，不再是用遙控器上面的幾顆按鈕操作機器，而是以自己的姿勢或動作來操作。
- ◆ **融合性：**讓「體感遊戲」的精神發揮到極致，以身體操控遊戲人物，讓遊戲跟

著你一起動，玩家在玩遊戲的時候更有帶入感，和遊戲世界完美融合。

- ◆ **教育性：**「運動遊戲」的精神可以延伸到教育方面，例如手語教學、空手道教學、幼兒教學等方面。

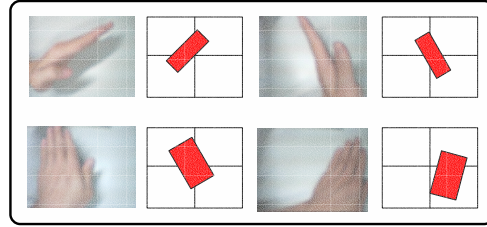


圖 2 圖形轉換為圖形矩

2. 電腦視覺

在這小節我們將討論電腦視覺的技術，因為並沒有任何一種電腦視覺技術可以適用於全部類型的遊戲，所以本研究使用了圖形矩、形狀識別和移動分析三種技術，我們將討論這三種技術的原理和它們適用的遊戲。

2.1 圖形矩(Image Moments)

圖形矩技術可以用數學式算出一張圖片的特性，將圖片中的物體近似成矩形，這可以讓我們得到圖片中物體的形心和傾斜角，因為圖形矩技術需要對圖片中的每個像素點(pixel)作運算，且光線對顏色的 RGB 值影響非常大，所以本研究計算圖形時將使用明度(Intensity)代替 RGB 值，明度的定義為像素點的 $(R+G+B)/3$ 。

假設 $I(x,y)$ 是圖片中位置於座標 (x,y) 像素點的明度，圖形矩定義以下特性項目：

$$M_{00} = \sum_x \sum_y I(x, y) \quad M_{11} = \sum_x \sum_y xyI(x, y)$$

$$M_{10} = \sum_x \sum_y xI(x, y) \quad M_{01} = \sum_x \sum_y yI(x, y)$$

$$M_{20} = \sum_x \sum_y x^2 I(x, y) \quad M_{02} = \sum_x \sum_y y^2 I(x, y)$$

由這些特性項目我們可以得到圖形的形心(圖形中心)位置 X_c Y_c 、傾斜角 θ 以及圖形矩的長 L_1 和寬 L_2 ，計算公式如下：

$$x_c = \frac{M_{10}}{M_{00}} \quad y_c = \frac{M_{01}}{M_{00}}$$

$$a = \frac{M_{20}}{M_{00}} - x_c^2 \quad b = 2\left(\frac{M_{11}}{M_{00}} - x_c y_c\right) \quad c = \frac{M_{02}}{M_{00}} - y_c^2$$

$$\theta = \frac{\tan^{-1}(b, (a-c))}{2}$$

$$L_1 = \sqrt{6(a+c + \sqrt{b^2 + (a-c)^2})}$$

$$L_2 = \sqrt{6(a+c - \sqrt{b^2 + (a-c)^2})}$$

取圖片的每一個像素點經過數學運算後，我們可以得到圖形的圖形矩，如下圖 2 所示：

2.2 形狀識別(Sharp Recognition)

形狀識別技術用於辨識圖形的形狀，例如分辨照片中的圖形是圓形、方形或三角形，在本研究中我們用一些預先設定的手勢當作遊戲命令，玩家可以透過預定的手勢來操作遊戲，我們選擇的形狀識別演算法為傾向統計(Orientation Histograms)，使用統計的方式來辨識玩家得手勢。

使用傾向統計演算法取圖片的每一個像數點，經過演算後我們可以得到像數點的傾向角，統計所有的傾向角，可以得到圖形的代表統計圖表，傾向角的計算方式如下圖 3 所示：

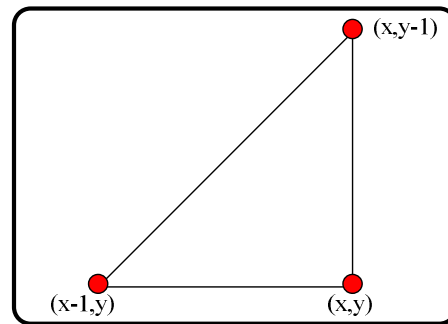


圖 3 傾向角三角表示法

使用基本的三角公式對像素點取明度計算：

$$\theta(x, y) = \tan^{-1} [I(x, y) - I(x-1, y), I(x, y) - I(x, y-1)]$$

統計圖形的全部傾向角後，我們可以得到類似圖 4 的曲線統計圖：

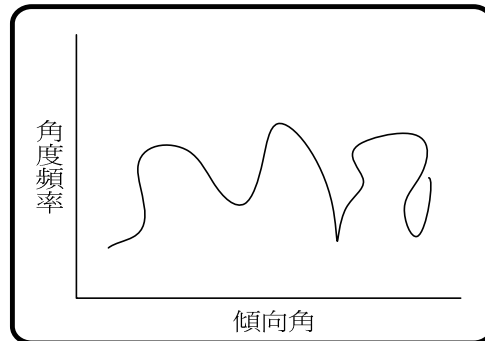


圖 4 傾向角統計圖

我們可以用傾向統計演算法，統計每個預定的手勢，為了計算方便我們會將曲線圖轉換成極座標的形式，如下圖 5 所示：

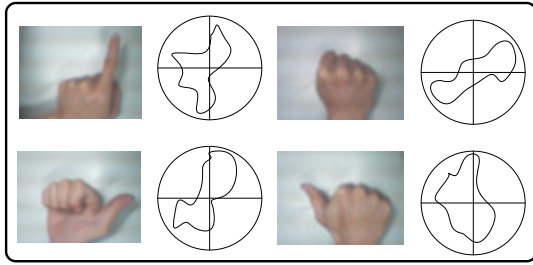


圖 5 預定手勢極座標統計圖

2.3 移動分析(Motion analysis)

移動分析用於追蹤畫面中玩家的移動並加以分析和判斷其動作，這需要從視訊攝影機中取得物體的連續移動資訊，在本研究裡視訊拍攝的背景必須為靜止不動之背景，且限制畫面中除了玩家之外不能有其它會移動的物體，將兩張連續的視訊畫面相減，畫面的背景會被消去，我們以下圖 6 為例說明：

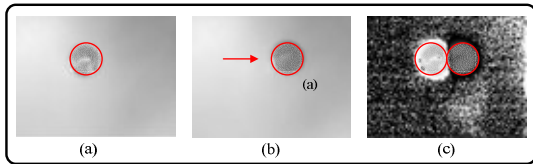


圖 6 物體的移動分析

圖 6(a)中是由視訊攝影機拍攝的第一張圖，圖中有一個靜止不動的一元硬幣，在圖 6(b)中這個硬幣往右邊稍微移動，將這兩張連續的圖相減去除背景得到 6(C)，由 6(C)中可以看出一元硬幣稍微往右邊移動。

使用這種移動分析的方法，我們可以對玩家的身體動作或移動方向作分析，將分析的結果當成遊戲命令控制遊戲，下圖 7 是我們用視訊攝影機追蹤手掌移動(從畫面左邊移動到右邊)的結果，畫面中的綠色框線會追蹤手掌移動並將手掌輪廓用框線畫出。



圖 7 視訊追蹤手掌移動

3. 系統架構

在本小節將說明我們規劃的系統架構，本研究分為兩個部份設計架構，分別為硬體和軟體架構，硬體架構包括一部個人電腦和兩部視訊攝影機，軟體架構分為動作識別模組和遊戲模組。

3.1 硬體架構

本系統的硬體架構極為簡單，以一部 PC 和兩

部視訊攝影機建構系統操作環境，如下圖 8 所示：

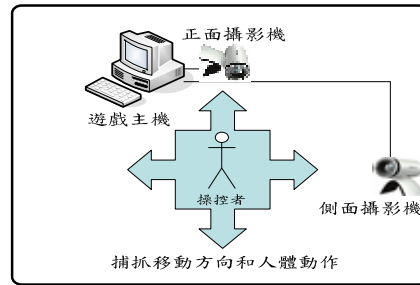


圖 8 系統硬體設置架構

系統以一部遊戲主機連接兩部視訊攝影機，一部位於遊戲主機上方拍攝玩家的正面動作和前後移動，另一部位於玩家的側邊拍攝玩家的側面動作和水平移動軌跡，因為單部視訊攝影機只能取得平面訊息，如果需要三維的立體資訊必須要多部攝影機拍攝才能做到，所以我們設置了兩部視訊攝影機，但如果玩家遊玩的遊戲只需要平面資訊，則只需要一部視訊攝影機即可。攝影機拍攝玩家動作後，接下來由動作識別模組分析畫面，判斷玩家的動作。

3.2 軟體架構

本系統的軟體架構分為兩個部份，分別為動作識別模組和遊戲模組，動作識別模組主要功能為抓取操控者的移動方向和身體動作，經過影像處理後再進一步分析操控者目前的動作，輸出給遊戲模組當作操作命令，軟體模組溝通圖如下圖 9 所示：

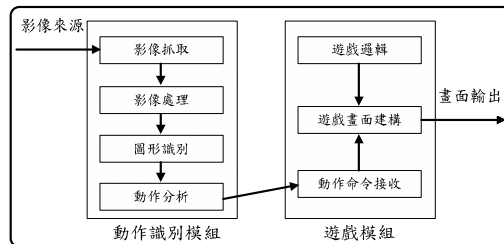


圖 9 系統軟體架構

以視訊攝影機為影像來源，輸入的影像經過動作識別模組和遊戲模組的處理後，才輸出為玩家看到的遊戲畫面：

- ◆ 影像抓取：我們使用 JAVA 的 JMF(Java Media Framework)API，來撰寫影像抓取的功能，擷取視訊攝影機的影像。
- ◆ 影像處理：為了減少雜訊必須對影像作去背景、二值化和破洞修補等處理。
- ◆ 圖形識別：依據遊戲類型，挑選使用圖形矩、形狀識別和移動分析等技術。
- ◆ 動作分析：分析玩家的手勢或身體動作，將這些連續動作轉換成遊戲命令。
- ◆ 命令接收：遊戲模組接收識別模組的遊戲命令，將之轉換為遊戲角色的動作。
- ◆ 遊戲邏輯：包括遊戲規則、人工智慧、

- 資料結構這些處於遊戲背景面的設計。
- ◆ 畫面建構：我們的系統設計使用 MVC 架構，畫面建構是表現給玩家的 View，對應的 model 就是遊戲邏輯。

4. 遊戲設計實例

我們於第 2 小節討論電腦視覺技術和第 3 小節討論系統的設計架構，討論完這些研究知識和背景之後，這個小節將會討論如何將電腦視覺應用在遊戲設計上，我們會以兩個不同的遊戲類型說明如何運用電腦視覺的技術，分別是射擊類遊戲和動作類遊戲。

4.1 射擊類遊戲

一般射擊類遊戲的操作都非常簡單，只需要瞄準敵人、射擊、角色移動和特殊攻擊四種命令，就可以組合射擊類遊戲的動作，在第 2 小節有提到一個遊戲可能需要搭配多種電腦視覺技術，而射擊類遊戲我們會以圖形矩搭配形狀識別兩種技術來說明遊戲的動作識別模組如何運作。

在 2.1 節說明圖形矩技術的時候，提到圖形矩有一個特性項目「傾斜角」 θ ，這個項目代表圖形的傾斜方向和角度，射擊類遊戲瞄準敵人的命令可以用傾斜角來達成，我們用圖 10 說明這個概念：

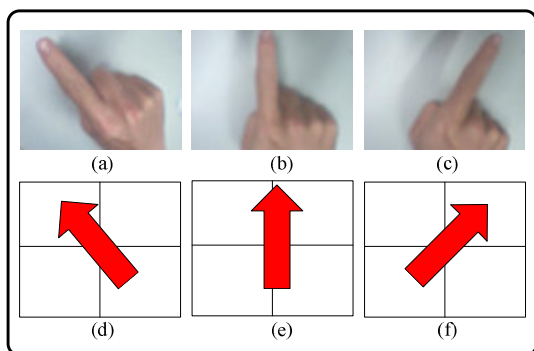


圖 10 圖形矩的傾斜角

圖 10(a)代表玩家瞄準螢幕左邊時的手勢，它的圖形矩會像 10(d)示一個傾向左上角的矩形(為了強調圖形矩的傾斜角方向，我們以箭頭符號表示圖形)，在設計遊戲時只要在遊戲畫面上，給予玩家即時的方向回饋(例如遊戲畫面裡的十字瞄準符號)，玩家可以很輕鬆的用手指的方向來瞄準敵人。

遊戲需要的動作命令，例如射擊、角色移動和特殊攻擊我們使用形狀識別技術來達成，這需要預先設定的手勢，因此我們為射擊遊戲設定了一些

手勢，圖 11 說明這些手勢代表的命令：

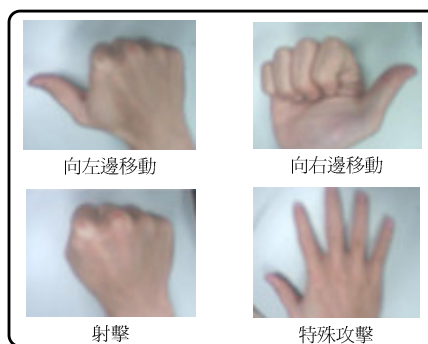


圖 11 遊戲預先設定的手勢

這些預先設定得手勢，經過動作識別模組分析後，會轉換成遊戲的命令，玩家可以透過這些單純的手勢來控制遊戲模組玩遊戲，比起用遊戲搖桿操作瞄準、移動和射擊等動作，這樣的操作方式對玩家而言是一種更直覺的方式。這種操作方式並不限定於射擊遊戲，只要是需要控制角色方向的遊戲都可以適用這個設計，例如賽車遊戲我們可以用傾斜角來算出車子轉彎的角度，圖 11 手勢代表的意義則改成加速、煞車和使用道具等命令，這套手勢就可以適用於賽車遊戲，最後我們以圖 12 表示以預設手勢操作遊戲的概念。

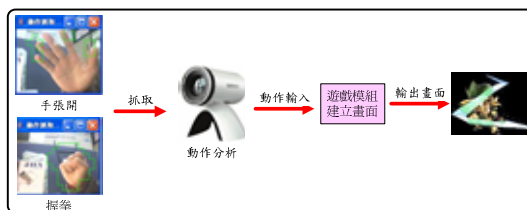


圖 12 以預先設定的手勢操作遊戲

4.2 動作類遊戲

動作類遊戲的種類很多，並沒有一個嚴格的分類限制，例如格鬥遊戲、動作闖關遊戲和跳舞機等都可以歸類為動作遊戲，在這篇論文裡動作類遊戲泛指，「遊戲裡人物的手、腳、身體」可以被玩家用命令控制的遊戲。動作類遊戲的操控，並不像射擊類遊戲可以用手勢來操作，我們期望玩家可以用自己的身體動作來操控遊戲人物，例如玩家舉起右手，遊戲裡的人物也舉起右手，又或者玩家跳起來，遊戲裡的人物也跳起來，這需要用到移動分析的技術，視訊攝影機拍攝玩家的連續動作，系統的

動作識別模組根據這些連續拍攝畫面，分析判斷玩家目前的行為，我們會以「揮動右手」來說明，動作識別模組如何分析「揮動右手」這個動作。

首先將視訊攝影機拍攝到的畫面，分成九宮格，如圖 13 所示：

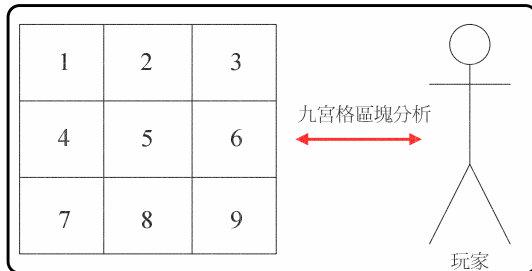


圖 13 用九宮格分析人體動作

我們將視訊畫面分成九個區塊，使用區塊變動量的大小來分析人體的行為，一般而言玩家有動作的部位視訊畫面的變動量會較大，例如玩家的身體上半身有動作，則區塊 1~6 的變動會比較大，反之若玩家的下半身有動作，則區塊 4~9 的變動量會比較大，我們分析視訊攝影機拍攝的連續畫面，判斷出哪些區塊是變動量比較大的，再根據變動量比較大的區塊數字來決定玩家的動作，下圖 14 標示「揮動右手」時九宮格變動量比較大的區塊：

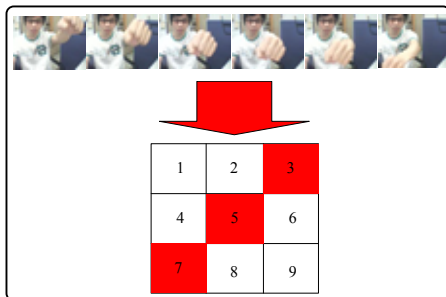


圖 14 揮動右手的區塊變動

可以看出「揮動右手」時區塊 3、5、7 的變動量會比其他區塊大很多，動作識別模組便可以根據這些變動量比較大的區塊判斷出玩家現在的行為是「揮動右手」，接下來我們在舉玩家「原地跳起」和身體「往右移動」兩個例子，圖 15 是這兩個例子的區塊變動圖：

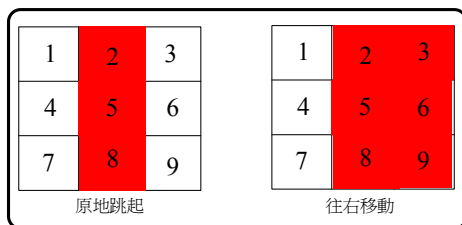


圖 15 人體動作區塊變動圖

使用這個方法，我們可以很容易的分析玩家的行為和動作，動作識別模組則會將分析結果當成遊戲命

令，傳送給遊戲模組，玩家便可以透過自己的身體操控遊戲人物，和遊戲互動。

5. 結論與討論

在本篇論文，提供了將電腦視覺技術應用在電玩遊戲的方法和概念，我們討論了合適用於遊戲的電腦視覺技術、系統架構的雛型概念，還有這些技術實際運用到遊戲設計的方法。電腦視覺技術和電玩遊戲結合後，將產生一種全新型態的遊戲，人們的休閒、娛樂和運動因此而結合為一體，使玩家在玩遊戲的時候更有帶入感與遊戲更進一步融合。

電腦視覺的應用可以無限寬廣，其和遊戲的結合，讓我們看到電腦視覺在「控制」這方面的潛力，以此為基礎，未來我們期望將此技術推廣運用到更多方面，例如電視和遊樂器遙控器的替代、幼教軟體、體育教學、簡報控制和虛擬控制面板。

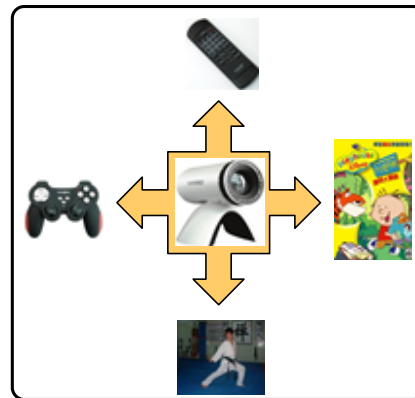


圖 16 更寬廣的應用

參考文獻

- [1] Colombo C., Del Bimbo, A. and Valli A., "real-time tracking and reproduction of 3d human body motion" Image Analysis and Processing, 2001. Proceedings. 11th International Conference on 26-28 Sept. 2001 Page(s):108 – 112
- [2] Godil A., Grother P., Ressler S., "Human identification from body shape" 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings. Fourth International Conference on 2003 Page(s):386 – 392
- [3] Green R.D., Ling Guan, "Quantifying and recognizing human movement patterns from monocular video Images-part I: a new framework for modeling human motion" Circuits and Systems for Video Technology, IEEE Transactions on Volume 14, Issue 2, Feb. 2004 Page(s):179 – 190
- [4] Takahashi K., Uemura T., Ohya J., "Neural-network-based real-time human body posture estimation" Neural Networks for Signal Processing X, 2000. Proceedings of the 2000 IEEE Signal Processing Society Workshop Volume 2, 11-13 Dec. 2000 Page(s):477 – 486

- [5] Ramanan D., Forsyth D.A., Zisserman A.,
“Tracking People by Learning Their
Appearance” Pattern Analysis and Machine
Intelligence, IEEE Transactions on Volume 29,
Issue 1, Jan. 2007 Page(s):65 – 81
- [6] Jonathan Knudsen 著，張晏誠 譯，JAVA 2D
圖學技術，歐萊禮，2000。
- [7] Brackeen David 著，許承強 發行，大師談
JAVA 遊戲設計，上奇科技，2003。
- [8] Andrew Rollings, Dave Morris 著，史萊姆工
作室 譯，大師談 遊戲架構與設計理論，
2004 。