

互動式動作合成之模糊標籤關鍵影格橋接技術研究

劉英男、*鄧宗賢

南臺科技大學資訊工程系

*yinzx@stust.edu.tw

摘要

一般運動學習軟體無法使用虛擬角色做即時回饋的原因，在於動畫設計師無法針對所有可能互動的情境繪製。本研究提出以標籤配對的方式來銜接不同動作，透過模糊理論定義動作橋接影格的標籤，讓虛擬角色的動作與動作之間能透過動作橋接影格進行轉換，讓虛擬角色有多元的互動呈現，不會因為切換不同的影片片段而產生不連續的動作轉換。本研究開發一套 3D 虛擬角色，利用帶動唱的方式來教導小朋友跳舞，系統使用 Microsoft Kinect 鏡頭捕捉與分析玩家動作來判斷玩家動作是否正確，虛擬角色能夠重複示範動作的方式，對使用者的錯誤動作給予即時的指導回饋。

關鍵詞：角色動畫、動作合成、動作標籤、虛擬角色

Research on Fuzzy Marked Keyframe Bridging Technology in Interactive Motion Synthesis

Ying-Nan Liu, *Zong-Xian Yin

Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology

Abstract

Immediate feedback is difficult for virtual characters in fitness video games because animators cannot design keyframes for all possible interactions. This study proposes fuzzy marked keyframe bridging techniques to connect different video clips. We use the fuzzy theory to define markers for bridging keyframes. These markers can be used as switch points between two movie clips and then generate the actions of the avatar. The purpose of this research is to develop the automatic generation of interactive feedback motions of avatars without discontinuous motion caused by switching between different movie clips. This study develops a 3D virtual characters video game to teach children to dance by singing. The system uses the Microsoft Kinect to capture the player's movements and postures and analyze whether the postures are correct, and then virtual characters provide immediate coaching feedback and repeat actions until the player can perform the correct movement.

Keywords: Character animation, Motion synthesis, Marked keyframe, Avatar

壹、研究背景與動機

運動和健康觀念越來越盛行，不少業者以運動為主題設計出各種教學光碟教導觀眾有氧體操等運動，體感裝置的進步讓遊戲變得更多元，不論是使用三軸加速度感應器和紅外線感應器的任天堂公司，或者是由微軟開發的 Kinect 用於捕捉使用者動作，都推出了以健康為概念的體感遊戲(Somatosensory game)，使得玩家不需要出門即可體驗運動的樂趣。電視上的幼兒頻道也設計出了許多帶動唱的橋段，由大哥哥和大姐姐等真人偶像，或者由虛擬角色來帶領小朋友跳舞，讓小朋友不會只是坐在電視前，而是鼓勵能夠站起來運動。體適能遊戲對兒童的重要性已經有不少研究證實了，在玩樂的互動過程中兒童的語言表達能力、人際關係、想像力與創造力都經由不同經驗刺激而成長，除此之外也能強健體能幫助身體發展，陳欣茹[1]等人研究發現實施幼兒運動遊戲課程能提昇幼兒創造力，經八周的運動遊戲課程後，受到教師的引導方式與幼兒的模仿，想像力、流暢性、獨創性皆有明顯進步。郭瑜婷[2]等人研究發現創造性舞蹈教學能夠培養兒童的創造力，藉由課程的實施、教師的引導提升兒童的學習興趣。有趣的舞蹈教學能提升兒童的想像力、創造力並能夠提升學習意願。根據張鳳菊[3]的研究發現台灣的幼兒運動遊戲課程實施率以經達到 90.68%。

人與機器的互動技術日趨完備，不論是遊戲或教學等軟體，利用虛擬角色有助於使用者操作系統。虛擬角色除了需要考慮外觀，動作的真實性也會決定使用系統時的體驗感。製作虛擬角色動畫須要花費大量的時間及心力，動畫是由一連串的影格組成，當影格快速播放時，利用視覺暫留即產生動畫。目前製作動畫最常使用的方式有兩種，傳統的方式是由動畫設計師在兩張動作影格間逐一補間角色的姿勢。另一種方法是使用動作擷取系統，在被錄製的模特兒身上放置訊號發送器，而周遭的訊號接收設備能夠擷取模特兒的身體活動數據訊號，錄製成動作檔，再利用渲染產生 3D 動畫。使用動作擷取系統能夠產生逼真的動畫，不少電影及遊戲均使用此方法來製作動畫，然而動作擷取系統需要昂貴的硬體設備，並且不利於非人形的虛擬角色編排動作。

透過影片來學習運動技能是常用的學習方式，使用者跟著螢幕上的教練來學習姿勢，但因為缺乏雙向的互動，當使用者動作錯誤時，影片無法做即時的回饋，使用者本身也無法察覺到自己的錯誤，讓這種學習方式效率大打折扣。而體感遊戲通常注重於娛樂性與流暢度，較少著重於動作指導與矯正，玩家做錯動作時通常只有顏色變化或音效來提醒玩家。但動畫設計師在設計連續動作時需要考慮每個動作影格的銜接，並為每個動作設計補間動畫，而一個靈活的虛擬角色通常有不少動作，設計師需要花費不少心力去銜接動作。舞蹈教學軟體或體育教學軟體，因為姿勢與動作很多，動畫設計師若是在每個舞蹈動作之間都去設計銜接其他動作的補間動畫，將會耗費相當大的時間與精力，所以大部份的體感遊戲在遊戲進行中虛擬角色無法即時的回饋互動。

由於虛擬角色的動作關鍵影格是預先建立，當遊戲引擎接受到外部指令，系統會決定播放一組特定的動作關鍵影格，當該組特定的動作關鍵影格播放完畢後，才能再接受新的外部指令，播放下一組的動作關鍵影格。傳統的互動遊戲設計，由於角色的單一動作設計很簡單，所以單一動作關鍵影格極短，當一個動作關鍵影格播放完成後再播放下一個動作關鍵影格的方式，仍可以讓使用者誤認為動作是即時的，不會感覺到延遲。然而，當需要發展虛擬教練系統，教導舞蹈或運動姿勢時，單一動作複雜且播放時間長，動作關鍵影格數量多。如果需要等待一個動作關鍵影格播放完畢後，才能接受新的指令，很容易讓使用者感覺系統的延遲，沒有互動的感受。所以本研究希望發展一套基於模糊標籤關鍵影格橋接技術，並利用該技術建立互動式動作動畫合成系統。我們將設計一款互動軟體，由逗趣的虛擬教練與兒童來互動，教導兒童跳舞，來達到兒童心智上以及體適能的成長。虛擬教練可以在使用者做錯動作時，能即時產生的反應回饋給玩家，如同具有生命力一樣。本研究提出透過動作橋接影格，自動銜接後續新動作的設計，使得虛擬角色動作能夠更多元，而且不需要設計師為每個動作設定補間動畫。

貳、文獻探討

一、多媒體與運動學習

多媒體意指傳遞訊息的媒介，基本型態為文字、圖案、影像、聲音等。而多媒體代表多種不同媒體的整合資訊呈現方式，以教學媒體來教學已成為趨勢，可使學生增加各種學習刺激的管道，並增加學習意願，目前教材的廠商除了提供一般的教科書以外，也會提供相關的數位教材如 DVD 光碟等。相較於靜態技能的學習，運動技能學習無法只從文字或者圖像上獲取足夠的資訊，因為運動學習的基礎是建立在模仿上，要達到熟練的動作技巧唯有不斷的透過觀察與修正動作才能達成，因此透過多媒體的方式來補足動態資訊更能幫助學習者於現場操作及練習，而透過攝影機拍攝並經過數位處理後的 2D 視訊教學素材，雖然製作簡便且有真實動作的影像特點，但若事先沒有經過規劃正確的觀察點，呈現的資訊可能無法對學習者達到有效的教學[4]。Mayer[19]定義的多媒體學習為文字和圖案等方式來幫助學習，並透過兩種以上的媒體來傳播訊息，呈現訊息的方式也需要兩種以上，例如動畫搭配聲音的呈現，而學習者也需要最少兩種感官受器來接收資訊，例如眼睛接收動畫，耳朵聆聽聲音。Mayer 認為多媒體學習的設計與會直接影響學習者的學習效果。張育婷[5]歸納多位學者的看法之於多媒體的理論，認為多媒體的運用在教學上具以下幾點共通性：(1) 整合的能力；(2) 互動性；(3) 立即讀取；(4) 標準的動作示範與解釋，並可重複演練與清晰講解；(5) 益於溝通；(6) 一對一的多媒體使用，能給予立即的鼓勵與回饋。

以多媒體教學的方式除了讓學習者獲得更多資訊以外，豐富的動態資訊也能吸引學習者的注意，然而豐富的資訊與互動方式是否能增加學習者的學習意願與學習成效則是主要關鍵，在邱惠芬[6]則將多媒體學習著重於國小學童的研究中，探討國小六年級學童在接受不同多媒體介面的學習成效，分別以靜態圖像+無語音+文字模式、靜態圖像+語音+文字模式、電腦動畫+無語音+文字模式、電腦動畫+語音+文字模式的方式來探討學習成效與學習動機，而研究成果發現國小學童對於活潑多元的動畫、聲音教學模式比以往課本的靜態教學更能夠提高學習意願。傅俊德[7]探討電玩娛樂結合數位學習的方式，將傳統武術的八卦掌的招式融入對打遊戲當中，而遊戲人物與場景也針對傳統武術做設定，根據研究結果發現，系統能夠幫助使用者做出正確的單招動作，並且加深對招式動作的記憶與學習，能透過遊戲以較有樂趣的方式來學習武術，不同於以往較單一的學習模式，並能夠加強使用者的學習動機。姚玉娟[8]則是使用智慧型手機，透過三軸加速度計和 GPS 設計出體感方式的互動教學與六種體感遊戲，讓學生到訪不同的教學活動區域，並提供學習者適時、適地、適切的知識，當使用者走進設定好的教學區域，系統會自動啟動體感遊戲，並介紹區域內植物資訊，研究結果也發現直覺的體感操作和遊戲結合學習能夠增加學習者與學習環境互動，採用以手勢為基礎之遊戲式行動學習系統，將手勢動作融入遊戲的操控，藉由觸發學習者感官的動作表徵及結合學習情境，幫助學習者提取過去相關記憶及經驗並轉換成圖像、符號等有意義的內在知識，儲存於長期記憶，不易遺忘，其學習保留的效果顯著高於非遊戲式行動學習系統。陳志寶[9]互動式多媒體的教材可包含各種媒體，使用者可簡易的上載多媒體教材與共享資料，因此系統教材能夠活潑多元也較吸引使用者的興趣，甚至改變學習模式增加主動學習的意願！

二、運動學習理論

王俊傑[14]與Schmidt [18]提出運動學習指的是透過自身不斷的練習與累積經驗，達到運動技能更進一步的表現，運動學習可分為三個階段，認知階段 (cognitive phase)、定位階段 (fixation phase)、自動階段 (autonomous phase)。

- (一) 認知階段：學習者在學習新的動作技能的第一階段，此時學習者會透過課本、說明書、老師的指導、影片示範，學習者主要透過觀察與模仿的方式來了解動作技能，進而建構出學習者所需的技能原理和操作順序。
- (二) 定位階段學習者的第二個階段，由熟練的分解動作連結至完整的一連串運動技能，此時練習的程度將影響最終運動技能的熟練度。
- (三) 自動階段學習者在此時已能夠熟練的呈現運動技能，動作與經驗已經深植學習者心中，動作也不易修正。

在認知階段時正確的資訊與顯著的被模仿對象對學習者有重大的影響，而在模仿動作的同時，若能以雙向互動的方式即學習者能從模仿對象身上獲得回饋訊息，對於運動學習有顯著的成果[10]。而胡名霞[11]研究也指出教學過程中給予學習者清楚正確的資訊可幫助學習者的動作印象另外也強調回饋給學習者的重要性，學習者若是一味模仿動作可能忽略了本身的錯誤而無法察覺。Bandura [25]也指出人類的許多學習均透過了個體的觀察、模仿和由他人的示範而習得的新行為，學習者從觀察中模仿學習對象，並藉著自我回饋和自我修正，逐漸使自己的動作接近學習對象的動作，此一學習過程包含了視覺、記憶與動作結構的交互作用。由以上文獻可知訊息回饋對於運動學習的重要性，當學習者無法分析當前的資訊做錯動作，或者無法跟上教學節奏，此時的回饋提示訊息則占有相當重要的地位。

吳翊慈[12]等人研究當使用者做錯動作時能夠產生回饋，設計一套穿戴式的感應器由三軸加速規組成，在學習者身上擺設 14 組感應器來測得關節點，並將此資訊與訓練過的影像序列資訊做比對，以調整影片播放速率快慢來配合學員演練太極拳的節奏。此研究能夠給予學習者演練過程中速度的回饋，學生跟不上影片內的教練速度時，影片會自動慢速播放，然而要在使用者身上放置的位置並非任何人都相同。顏羽君[13]等人的研究使用了攝影機拍攝使用者體操動作並辨識其動作是否正確，錯誤則會出現警告提醒使用者，透過攝影機的使用者不用穿戴其他設備，然而即時的影像分析需要大量的計算，對於須要即時運算的運動教學系統，僅使用靜態的訊息提醒使用者，無法即時矯正使用者的動作。王俊傑[14]等人開發的視訊學習平台，注重於虛擬教練的動作，專業舞者透過動作捕捉裝置產生的舞蹈訊號匯入 3D 模型中，藉此產生逼真的舞步，並且使用視訊錄製使用者的影像，使用者可於事後觀察自身動作與虛擬教練動作是否有不同之處，藉此達到指導的目的。

多媒體教學當中產生的回饋方式有很多種，可透過文字、聲音、動畫等。在多媒體輔助運動學習上，學習新動作的學習者通常將注意力放在最明顯，且最容易理解的資訊上，也因此學習者通常跟著系統內的虛擬教練演練，也就是說學習者的模仿對象是虛擬教練，此時的教學回饋方式以虛擬人物的動畫呈現效果最佳。

三、電腦動畫

動畫產生的原理主要是受視覺暫留的影響，人類眼睛一秒鐘約可暫留 24 幅圖片，動畫設計師在電腦上繪製一系列不同的相關圖片經由時間軸的運行快速播放，產生了動畫的效果，而 3D 動畫與 2D 動畫相比更有臨場感，將圖像、文字等回饋資訊以 3D 動畫方式更能增加學習效率[15]。目前製作 3D 人物動畫的方式大略分成兩類。

1. 透過控制人物骨架的方式來生成動畫，動畫設計師會先創作虛擬人物的骨架，在不同影格間，經由拖拉編排設計，或者程式運算改變骨架動作進而改變與骨架相連的外層皮質。
2. 使用動作捕捉裝置，經由置放訊號發射器在被錄製者身上，周圍的接收器蒐集一系列真實的運動訊號後套用在虛擬角色上，即可產生接近真人的動畫效果，此方式雖能產生極逼真的人物動作，但是也需要花費大量成本購買硬體裝置。

不論使用何種方式動畫設計師都必須花費大量心思調整動作，也因此有不少研究針對自動產生人物動畫的方式提出研究。

(一) 基於動作實例的合成方法

將已經錄製好的各筆動作資料建立資料庫，系統能擷取資料庫的動作片段來進行合成。常見的合成方式為線性內插 (linearly Interpolating) 的方式，即將要進行合成的新動作的第一個影格與舊動作的最後一個影格進行資料比對，然後在這兩個動作間插入一個可平移過度的動作。由於每個動作可能動作差異很大，使用這種方式產生的過渡動作往往得到不自然的效果，Cameron Egbert 和 Parris K. Egbert 等人[24]提出一種即時運算的線性內插方式，將新動作與舊動作的影格進行混合重疊，從兩段影格動作中萃取出其自然平順的連續動作，解決動作不順暢的問題，但此方法的動作較無變化，利用現有兩個動作影格路徑混合搜尋產生的動作，動作雖然流暢，但往往得到的合成動作只有一種動作路徑，使得虛擬角色動

作較為單一呆板。吳亞屏等人[16]研究透過將影格標上動作的機本資訊標籤的方式讓動作在轉換時能夠找到相似動作進行轉換而達到動作順暢。

(二) 以物理學為基礎的方法

以物體上的關節點為基礎，利用物理的方式計算第一個節點到第二個節點之間的角度、位置，此方式可有效的算出兩個關節點之間的內插值，然而人體的關節是互相連接的，因父子層級關係而互相影響位置與角度，但若是套用在角色整體的關節上，可能導致電腦運算量過大，運算效率低的情形 Tainchi Lu[22]等人開發一款與海豹互動的系統，透過使用者不同的刺激方式海豹會有不同的反應，由於海豹的骨架較為單一直線，因此透過運算海豹骨架節點來產生即時動畫能產生逼真的效果。

(三) 語意、標籤分析的方法

為了簡化資料的結構，並能有效率的搜尋動作資料，有些研究人員利用語言中的詞意來標記動作資料，利用語言中的各種詞性來接合與轉換動作片段，由於此方法是利用語意的組合來合成動作，因此在文法上的解析需要大量的規則分析。Masaki Oshita[23]提出以輸入文字的方式來產生動畫，利用動詞、副詞、介詞來定義動作與物品之間的關係，並配合座標的位置與動作資料庫來計算出合成的動作。Dan Zong 等人[21]研究特別設置了兩個層別分析語意，以利使用者能夠輸入口語化的指令而不需使用艱難的動畫控制命令句。另外，賦予動作片段特定的標籤與特性，比語意標記還要明確，搜尋合成的動作資料也被分類標記，因此能快速產生合成動作的資料，並且合成的效果也比語意標記正確，故本研究採標籤的方式來實現動作合成的方法。

四、模糊理論

模糊理論是由 Lotfi Zadeh 於 1965 年提出。與傳統的明確集合 (crisp set) 不同，明確集合定義為：

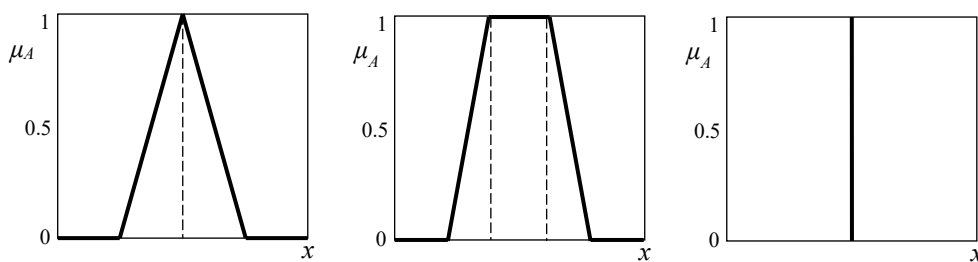
$$X_{crisp}(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

如果元素完全符合於集合 A，則為 1。相反，元素完全不屬於集合 A，則為 0。Zadeh 定義的模糊集合 (fuzzy set)，元素可以對一個集合為部分程度的屬於，而這個部分程度的屬於可以利用模擬歸屬度 (grade of membership) μ 表示，其值的範圍由 0 (元素完全不屬於集合) 到 1 (元素完全符合於集合)。模糊集合可以定義為 (U, m) ，其中的 $U = \{x_1, \dots, x_n\}$ 為有限的元素集合， $m = \mu_A$ 為歸屬度函式 (membership function)， $m(x) = \mu$ 為模擬歸屬值，當 $x \in U$ ，模糊集合定義為：

- not included : $m(x) = 0$
- fully included : $m(x) = 1$
- partially included : $0 < m(x) < 1$

模糊理論的主要目的是對現實世界進行建模，因為在現實世界中，這些現象本質上是模糊和模稜兩可的。人類對於複雜問題的知識，可以使用自然語言的不精確性來表示。而模糊集合和模糊邏輯理論，正好為此類資訊的數學表示提供有效的處理工具。模糊系統是一種基於模糊邏輯理論的資訊處理架構，主要是對於傳統明確集合或二進制邏輯無法處理的情況。系統通常為一組有明確結構的互動的元件組成，可以與外部環境區分開來。系統通過輸入和輸出與外部環境進行通訊。從外部輸入明確的資訊到模糊系統，經由模糊系統模擬人類思考模式。模糊系統它們的主要特徵以模糊條件 (if-then) 規則形式的符號知識表示。模糊系統包含四個部分，(1) 模糊化 (fuzzification)、(2) 模糊規則庫 (fuzzy rule base)、(3) 模糊推論機制 (fuzzy inference mechanism)、(4) 解模糊化 (defuzzification)。

在模糊化的過程，明確的資訊會依照開發人員設定的歸屬函數而轉化成歸屬度，歸屬函數的選擇可根據個人主觀觀點，或專家經驗來認定，一般常見的歸屬函數有高斯歸屬函數 (Gaussian membership function)、三角形歸屬函數 (triangular membership function)、梯形歸屬函數 (trapezoidal membership function)、singleton membership function 等，圖 1 為常見的歸屬函數。解模糊化的目的在於將模糊推論裡的模糊資訊以明確的數值做輸出，一般常用的解模糊化方法有最大歸屬度法、重心法、高度法及面積法。



(a) Triangular membership function (b) Trapezoidal membership function (c) singleton membership function

圖 1 常見的歸屬函數

參、研究方法

一、系統概述

本論文實作一套互動虛擬舞蹈教練系統，由虛擬教練帶領使用者執行一首舞曲，當使用者做錯動作時，虛擬教練能夠即時的矯正動作，系統使用 Kinect 鏡頭分析玩家的動作，虛擬教練則利用動作關鍵影格合成技術呈現互動的動作。系統中的虛擬教練會帶領玩家跳完一首簡單的舞蹈，並隨時偵測玩家的動作姿勢，透過動作比對分析，決定適當回饋動作，並即時產生動作關鍵影格。舉例來說，虛擬教練正做著拍手的動作，而玩家未跟著動作或者做錯動作，此時系統會將虛擬教練的動作關鍵影格切換到矯正玩家的動作關鍵影格片段，又或者是玩家太久沒有與虛擬教練互動，則虛擬教練會做出吸引玩家的動作片段。為了讓虛擬教練在轉換動作的過程能夠流暢自然，且不需要為每種動作繪製一組動作關鍵影格，本研究透過模糊理論所產生的標籤，在動畫影片內找出差異性最小的影格作為動作橋接影格，並將動作橋接影格標籤紀錄在 XML 檔。圖 2 為本系統開發之虛擬教練動作流程圖。

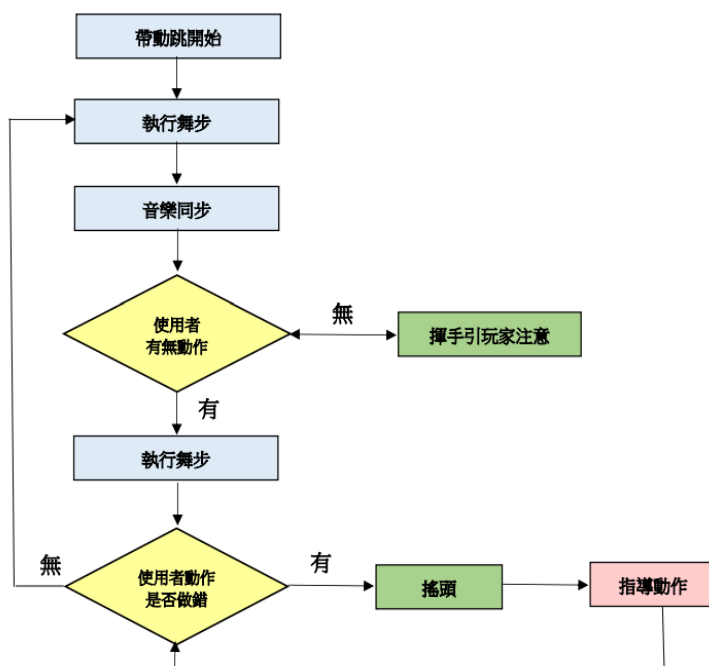


圖 2 虛擬教練動作流程圖

二、動畫合成

動作銜接的時間點除了不同的舞步連接以外，能在偵測到玩家的動作錯誤時，即時轉換至提醒使用者的動作關鍵影格，例如搖頭、揮手、放慢虛擬教練的動作等，讓虛擬教練可以與玩家回饋互動，直到使用者做對的動作後才繼續舞蹈。透過這種學習方式能夠加強玩家的學習，對於已經熟練的舞步也不用再額外從頭跳一次。圖 3 中的每一個動作，都是一組已繪製完成的動作關鍵影格，為了能讓虛擬教練在

動作轉換時，能在視覺呈現順暢動畫，我們在每動作的影格內建立動作轉換標籤，讓兩個動畫影片能在相似的影格間進行轉換。如圖 4，舞步動作與指導動作銜接時，由橋接影格來進行銜接，因此能夠避免在動作轉換時產生的閃爍、不連續缺點。

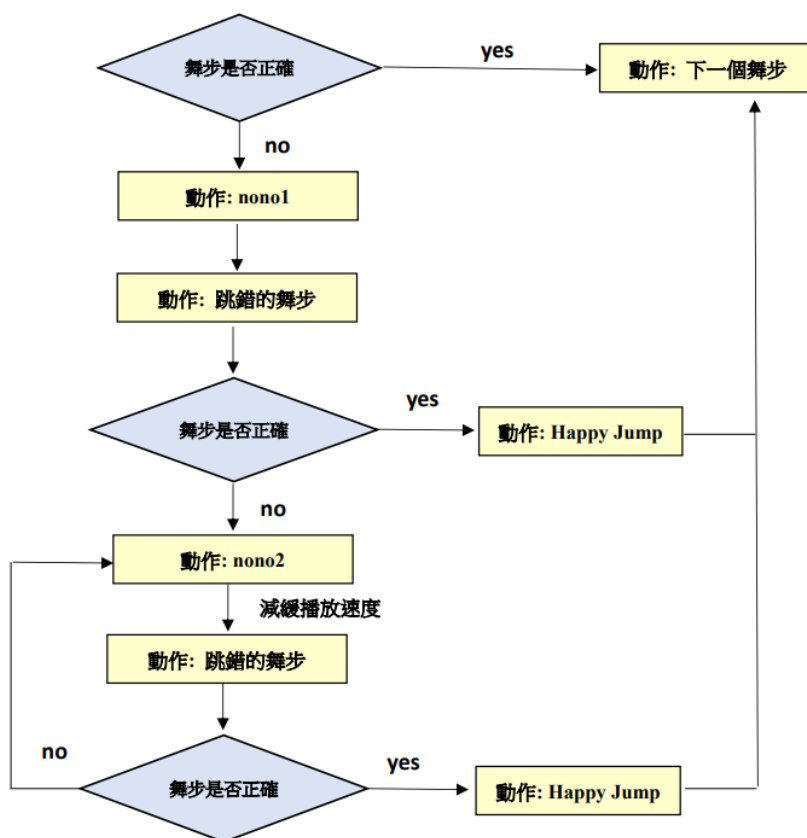


圖 3 虛擬教練動作指導動畫合成流程圖

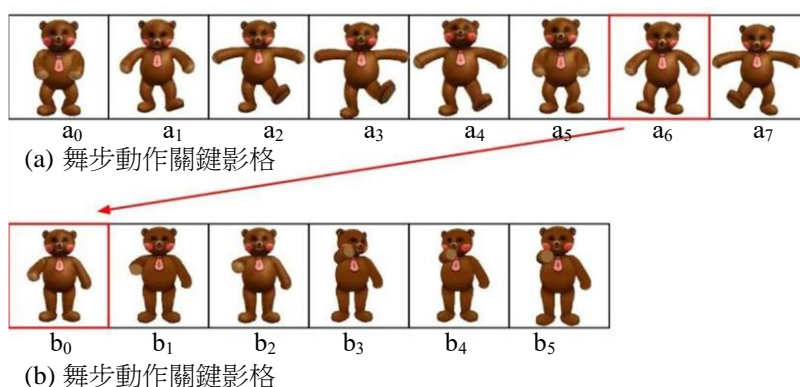


圖 4 舞步動作關鍵影格(上)與指導動作關鍵影格(下)透過橋接影格 a₆與 b₀銜接

三、建立影格標籤

虛擬人物的動作由動畫設計師繪製，虛擬人物的動作影片數量多寡決定了虛擬人物的動作流暢度。然而，像遊戲設計時所使用的動作種類繁多，且不同的遊戲關卡和遊戲任務，在不同的動作需要能夠互相接合的情況下，通常動作關鍵影格一開始有一個標準的姿勢站立，這個動作是將所有肢體固定特定姿勢，以方便不同動作關鍵影格的切換，但是這種方法不適合用於即時的互動多媒體遊戲，而舞蹈教學更是不能接受動作與動作之間有所停頓。本論文透過橋接影格標籤的方式讓不同動作能匹配到不同動作關鍵影格中相似的影格來銜接合成一動作。本研究透過三個步驟來自動產生動作橋接影格標籤。步驟一，

在所有動作關鍵影格中尋找影格動作出現頻率高，為動作橋接影格候選，原因在於定義動作橋接影格的數量越大，動作切換彈性也越大。

步驟二，本論文設置 10 個特徵值以代表 3D 模型的影格動作。依據模型四肢節點的向量座標，10 個動作特徵分別為 RH (right-hand abduction)、RHE (right-hand flexion elbow)、LH (left-hand abduction)、LHE (left-hand flexion elbow)、Posture、Trunk、RF (right-foot abduction)、RFK (right-foot knee abduction)、LF (left-foot abduction)、LFK (left-foot knee abduction)，分別代表虛擬教練的手、腳、軀幹，以及重心。在手腳的特徵主要判斷虛擬教練是否有伸出手腳或者手腳有無彎曲的情況，而軀幹的部份代表著虛擬教練的身體是否有彎腰的情況。重心主要可以表示虛擬教練的身體離地高度，用來表示跳躍、站立、坐著、躺著。使用這種方式能夠有效的分類動作，經過分類後再依照類別進行比對可以避免僅計算節點距離，不需要考慮實際動作語意上的差別。詳細的參數設定如表 1

表1 動作特徵

右手	伸展(RH)	手部位置:上(up)、中(mid)、下(down)
	關節(RHE)	彎曲與否:true、false
左手	伸展(LH)	手部位置:上(up)、中(mid)、下(down)
	關節(LHE)	彎曲與否:true、false
身體	姿勢(Posture)	姿勢狀態:躺(lie)坐(sit)站(stand)跳(jump)
	軀幹(Trunk)	彎曲與否:true、false
右腳	伸展(RF)	伸展位置:內(inward)、外(outward)
	關節(RFK)	彎曲與否:true、false
左腳	伸展(LF)	伸展位置:內(inward)、外(outward)
	關節(LFK)	彎曲與否:true、false

本研究排除頭頸特徵，因為本次研究之虛擬人物以教導小朋友跳舞為主要任務，大部份虛擬人物都直視面對操作者，且如果頭頸有大動作時，該關鍵影格不會定義為橋接影格，所以本研究排除頭頸特徵。

步驟三，本研究使用模糊理論建立動作橋接影格的動作特徵，並計算兩個動作橋接影格的相似度，如公式 (7)，其中 $k=10$ 代表 10 個動作特徵，每個動作特徵乘加權值 S ，由於舞蹈動作當中手部動作的變化非常明顯，很容易被辨識，因此我們在手部 (RF、LF、RFK、LFK) 設定的加權值最大，而身體特徵，如站立、坐立、躺的差別也非常明顯，所以其加權值同等手部；腳部動作影響整個動作並不大，因此加權值最小。若是運算結果 d 小於門檻值 0.4 則代表兩個影格的動作是相近的，即可進行兩個動作關鍵影格的切換，且可以使玩家不易察覺。

$$d = \sqrt{\sum_{k=1}^{10} (\mu_{1k} - \mu_{2k})^2 \times S_k} \quad (1)$$

四、動作動畫影格模糊函式定義

本研究利用了模糊邏輯與 10 個特徵來計算虛擬教練的動作相似度，每個動作特徵有各自的歸屬函式，本研究的歸屬函式的定義使用觀察法與經驗法，並透過重覆實驗所得的最佳結果。以下分別對 RH、RHE、Posture、Trunk、RF、RFK 等特徵作說明：RH 代表著右手伸展狀況：放下、平舉，以及抬高 (down、middle、up)，輸入值為 A 向量與 B 向量夾角 x ，代表手臂抬舉角度，歸屬函數 μ_{down} 、 μ_{mid} 、 μ_{up} 的定義如下：

$$\mu_{down} = \frac{90-x}{70}, 20 \leq x \leq 90 \quad (2)$$

$$\mu_{mid} = \begin{cases} \frac{x-20}{70}, & 20 \leq x \leq 90 \\ \frac{180-x}{90}, & 90 \leq x \leq 180 \end{cases} \quad (3)$$

$$\mu_{up} = \frac{x-90}{90}, 90 \leq x \leq 180 \quad (4)$$

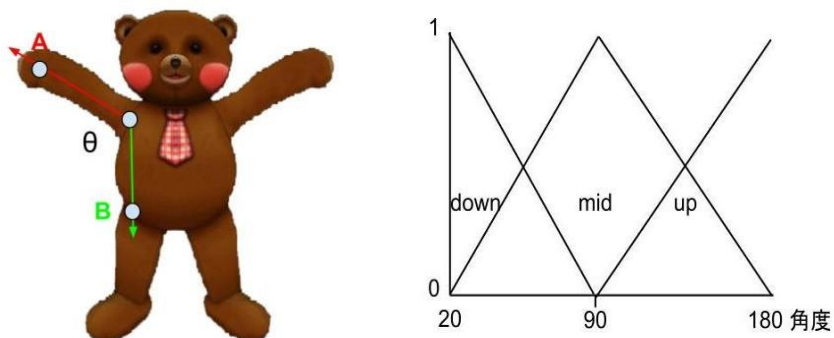


圖 5 右手伸展狀態的模糊歸屬度函

RHE 代表著右手關節是否有彎曲 (true、false)，輸入值為 C 向量與 D 向量角度 x ，歸屬函數 μ_{false} 、 μ_{true} 定義如下：

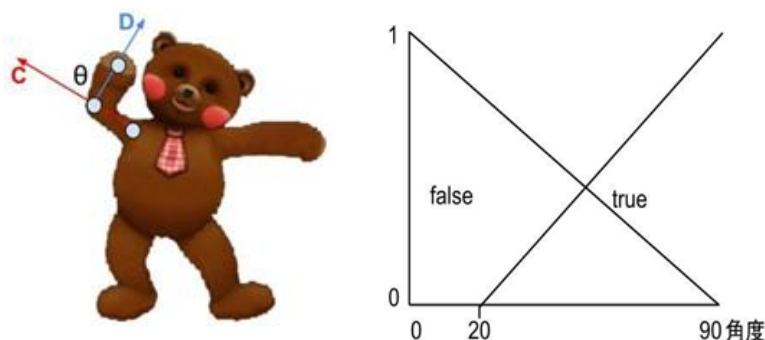


圖 6 右手的彎曲狀態的模糊歸屬度函

$$\mu_{false} = \frac{90-x}{90}, x \leq 90 \tag{5}$$

$$\mu_{true} = \frac{x-20}{70}, 20 \leq x \leq 90 \tag{6}$$

姿勢 (Posture) 是以虛擬教練的重心點位置判斷姿勢，輸入值為 Y 座標值離地高度 x 。“躺”、“坐”、“站”，以及“跳”的動作歸屬函數 μ_{lie} 、 μ_{sit} 、 μ_{stand} 、 μ_{jump} 定義如下：

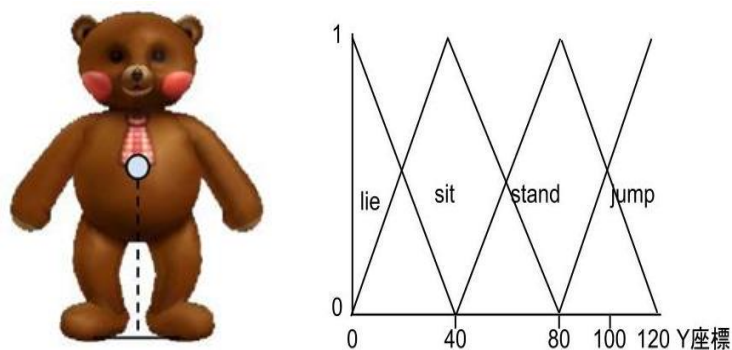


圖 7 身體姿勢的模糊歸屬度函

$$\mu_{lie} = \frac{40-x}{40}, 0 \leq x \leq 40 \tag{7}$$

$$\mu_{sit} = \begin{cases} \frac{x}{40}, & 0 \leq x \leq 40 \\ \frac{80-x}{40}, & 40 \leq x \leq 80 \end{cases} \tag{8}$$

$$\mu_{sit} = \begin{cases} \frac{x-40}{40}, & 40 \leq x \leq 80 \\ \frac{120-x}{40}, & 80 \leq x \leq 120 \end{cases} \tag{9}$$

$$\mu_{jump} = \frac{x-80}{40}, 80 \leq x \leq 120 \tag{10}$$

軀幹 (Trunk) 特徵是用來確認虛擬教練的軀幹是否有傾斜，輸入值為 E 向量與 F 向量角度 x ，軀幹有傾斜與無傾斜的歸屬函數 μ_{true} 、 μ_{false} 定義如下：

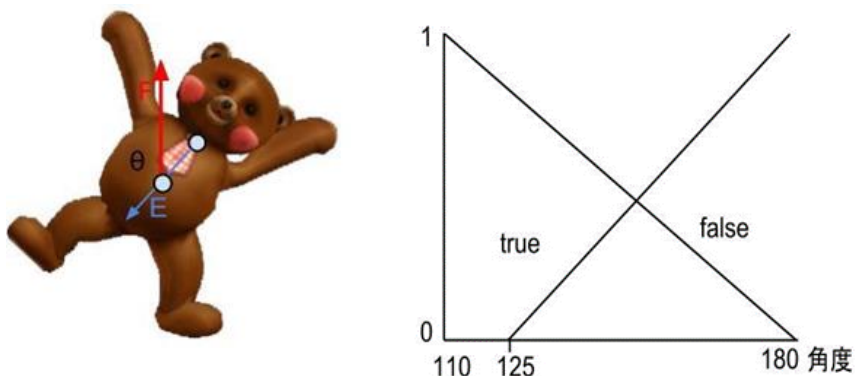


圖 8 軀幹傾斜角度的模糊歸屬度函

$$\mu_{true} = \frac{180-x}{70}, 110 \leq x \leq 180 \tag{11}$$

$$\mu_{false} = \frac{x-125}{55}, 125 \leq x \leq 180 \tag{12}$$

RF 為判斷右腳是否有伸出的情況，輸入值為 G 向量與 H 向量角度 x ，角度越大表示腳部向外展。腳部內縮與外展的歸屬函數 μ_{inward} 、 $\mu_{outward}$ 定義如下：

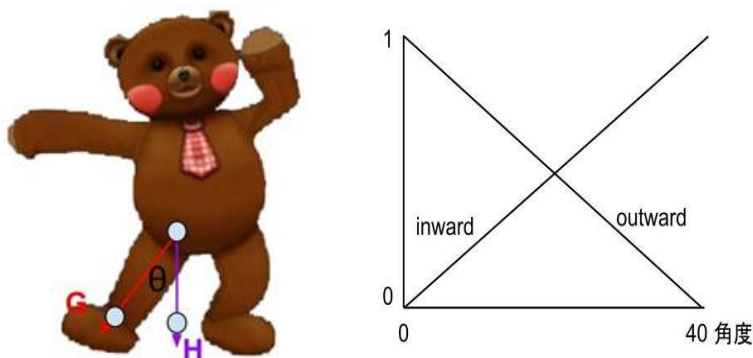


圖 9 右腳伸出的模糊歸屬度函

$$\mu_{inward} = \frac{40-x}{40}, 0 \leq x \leq 40 \tag{13}$$

$$\mu_{outward} = \frac{x}{40}, 0 \leq x \leq 40 \tag{14}$$

RFK 為判斷右腳關節是否有彎曲的情況，輸入值為 I 向量與 J 向量角度 x ，角度越小表示腳關節呈現彎曲。腳關節有彎曲與無彎曲的歸屬函數 μ_{true} 、 μ_{false} 定義如下：

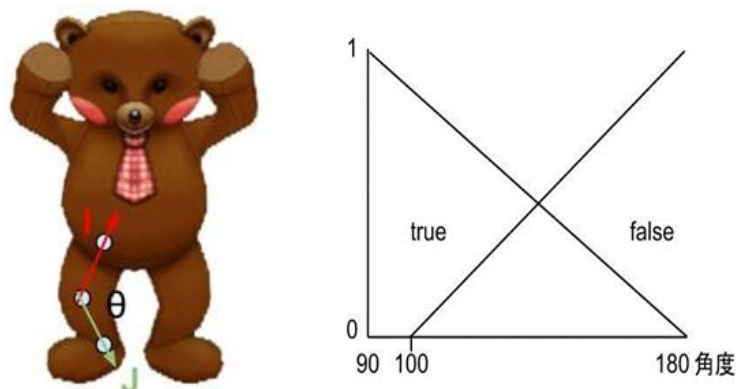


圖 10 右腳關節彎曲的模糊歸屬度函

$$\mu_{true} = \frac{180-x}{90}, 90 \leq x \leq 180 \tag{15}$$

$$\mu_{false} = \frac{x-100}{80}, 100 \leq x \leq 180 \tag{16}$$

五、動作橋接影格銜接

為了讓虛擬教練如同真人般能夠即時做出互動動作，來矯正玩家的姿勢，我們需要在每個不同的舞步之間找出至少一個能接換到“矯正玩家的動作關鍵影格”片段的動作橋接影格，在不同動作動畫影格不需要由動畫設計師為每一種可能的動作轉換，而能夠利用動作橋接影格進行切換。雖然動畫的流暢性無法像動畫設計師親自調整的完美，但考量互動系統需要繪製大量的動作動畫，其時間及經濟成本過高，使用此方法能得到相當不錯的效率。

當影格進行轉換時，系統會去查訪新動作的影格，利用動作模糊函式，從新動作關鍵影格的找出適當的影格進行切換。如下圖 11，綠色的影格代表動作橋接影格，當不同舞步進行轉換時，為了避免切換到新動作的尾端而造成不自然的動作，查訪會先由動畫影片前段開始，如果前段影格沒有找到合適的影格，則再查訪後半部，圖 12 搜尋到同樣動作的動作橋接影格，將動作關鍵影格銜接在一起完成。

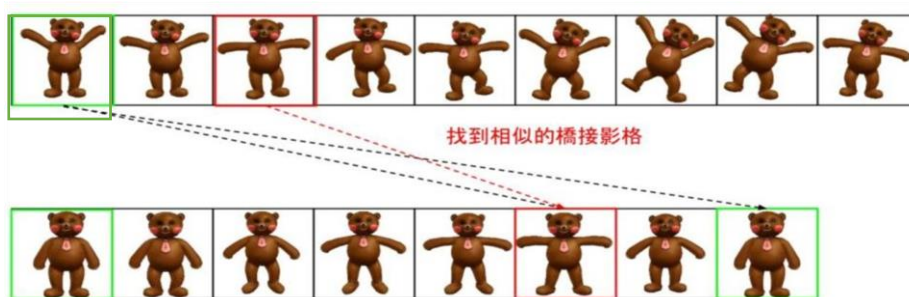


圖 11 查訪後找到合適動作橋接影格 (以紅色方框)



圖 12 動作關鍵影格銜接完成

為了避免在銜接影格時，兩個動作之間沒有合適的動作橋接影格的動作可以轉換，本研究因此設計了一組補間動作，如圖 13。當 D1 舞步與 D2 舞步銜接時，沒有合適的影格可以轉換，因此透過中間的補間動畫來銜接 D1 舞步與 D2 舞步，流程如圖 14。補間動作是一連貫由動作橋接影格所構成，因此補間動作能夠銜接在任兩個不同動作影片。

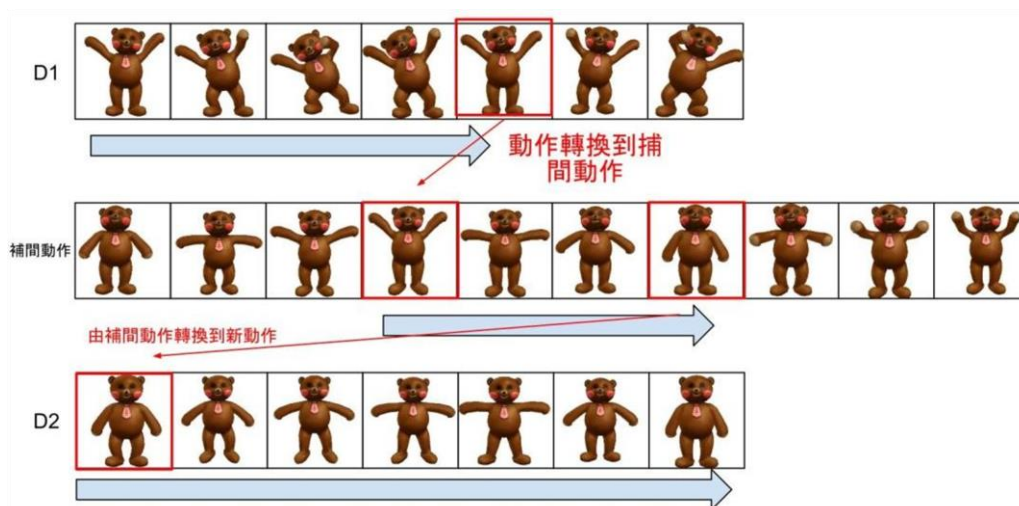


圖 13 透過補間動作銜接兩動作影片

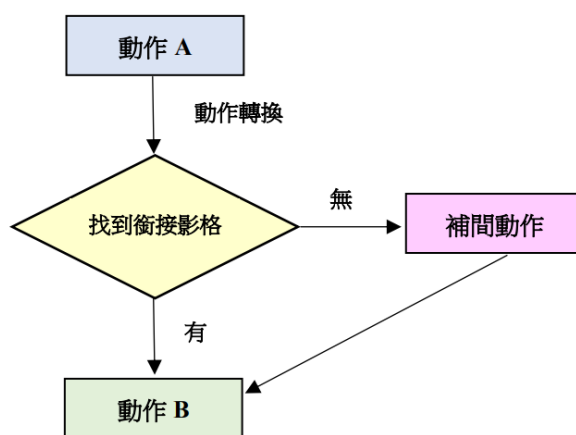
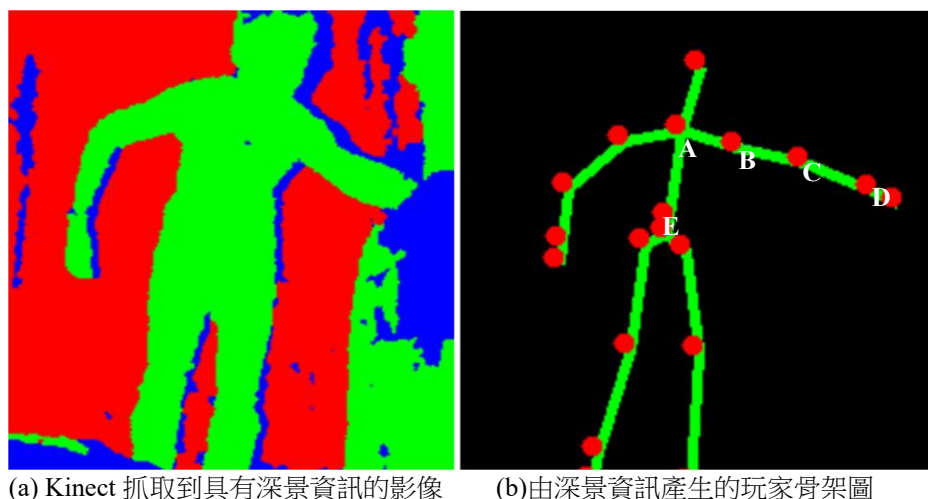


圖 14 補間動畫切換流程圖

六、動作辨識

本論文使用 Kinect 攝影鏡頭來捕捉人體動作，Kinect 為微軟開發用 XBOX360 主機上，主要可將人體動作轉換為有深度資訊的數位影像，使得玩家不需要其他裝置即可遊玩體感遊戲，Kinect 除了能捕捉彩色影像以外，透過紅外線發射器及紅外線攝影機能抓取影像深度資訊，進而將玩家影像轉換成 3D 的骨架資訊，因此我們只需要取得特定骨架節點的三維直角座標，即可透過向量運算結果來判斷玩家動作，也不需要再額外進行去背等影像分析的動作進而提升效能及即時運算的能力。

本研究的目標使用者族群為兒童，由於兒童舞蹈的編排動作較為單純，且考量兒童的動作能力，所以採用重點式的動作辨識，我們所設計的每個舞步都有一至二個明顯的重點動作，系統判定的重點為玩家的右手是否有確實彎曲，左手伸直，以及左腳跨出等，只要使用者能達到此動作的標準，系統就會判斷舞步正確。為了達到動作辨識，系統利用 Kinect 抓取到玩家骨架向量座標來分析，如圖 15(b)，在左手彎曲的地方我們用 BC 向量與 CD 向量的夾角來判斷玩家是否有達到彎曲的角度，計算 LFE 值；並由 AB 向量與 AE 向量的夾角相差判定手抬舉高低，計算 LA 值。



(a) Kinect 抓取得到具有深景資訊的影像 (b)由深景資訊產生的玩家骨架圖

圖 15 系統利用 Kinect 分析玩家動作示圖

肆、系統實作

一、人物造型

本系統設計目標族群為國小學童與學齡前幼童，因此在人物設定上以可愛的玩偶熊造型來設定，我們選擇卡通造型熊來當小朋友的虛擬教練，如圖 16，能增加使用者的好感以外，也能減少動畫設計師編排舞蹈的負擔。由於兒童舞步動作通常簡單逗趣，較不注重於技巧性，因此 3D 模型的製作面數上不需要過於龐大。



(a)使用者連續做錯動作時虛擬教練苦惱的動作 (b)使用者閒置太久，虛擬教練打瞌睡的狀態

圖 16 虛擬教練造型

兒童在學習過程犯錯時，應盡量避免給予直接的否定，當我們責備兒童時同時，也造成兒童產生負面情緒，伴隨而來的後果是打擊學童的自信產生挫折感，應該以正向鼓勵的方式來教導兒童，所以我們的虛擬教練在使用者連續犯錯後，設計出現可愛的動作，讓學童感到輕鬆沒有壓力並能夠快樂的跳舞。

二、舞蹈編排

本研究的舞蹈動作以參考幼兒帶動唱節目與體感遊戲 *Just Dance Kids 2* 來編排設計，專為兒童設計的舞步，除了有可愛的動作外，也包含了簡單的有氧運動和伸展操，一首簡單的兒童舞曲通常包含 8 到 12 個不同的舞蹈動作，透過不同的組合順序與編排組合來產生一系列舞步。

在編排舞蹈時，將相同的舞蹈動作最少兩次重覆為一組進行排序組合，使用這種編排的好處，在於重複多次的動作能讓幼童玩家容易跟隨，加深對動作的印象，不會因為舞步變換得太頻繁而感到措手不及，另一方面當幼童跳錯舞步動作時，虛擬教練進行切換動作的時間點，也能落在跳錯舞步的段落內，

避免在動作影片已經播放至下一個舞步動作，才切換成指導動作動畫片段，產生不自然的感覺。

另外，由於每台電腦的處理速度不同，系統在執行播放影格與播放音樂的速度上會有不同步的問題，使得舞蹈動作無法對上音樂節拍，為了克服這個問題，我們在特定的舞蹈段落影格中定義音樂時間標籤，當動作播放至該影格時，檢查音樂實際播放的時間。若是發現當前的音樂播放時間小於影格內標記的時間，代表音樂播放已慢於配合的舞蹈，此時系統會在下一個動作調整舞蹈的速度放慢動作，讓音樂節奏與動作影格能同步；反之亦然。

三、系統成果

系統介面，如圖 17，除了螢幕內主要的虛擬教練帶動跳以外，也以文字和圖案方式來提醒玩家舞步重點，背景主題設計成繽紛的森林幻想世界，能啟發小朋友豐富的想像力。圖 18 為本系統遊戲實際場景範例。

透過標籤式動作合成方法，能尋找動作相似影格，將兩個動作動畫影格合成在一起，因此能在每個不同的舞蹈動作切換至矯正動作，達到即時矯正的互動效果。當做錯動作時，由畫面上的小熊能指導小朋友，搭配輕快的音樂能夠讓小朋友陶醉快樂的童話世界。



圖 17 系統介面圖設計，虛擬教練與提示用的文字、圖案等資訊



(a) 場景 1






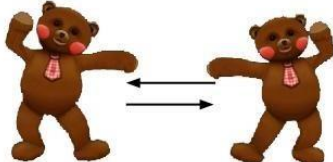






(b) 場景 2

圖 18 系統場景範例，主題為奇幻的林，風格主要呈現色彩繽紛的卡通世界

在動作部分除了基本的讓兒童流汗運動以外，也希望能增強兒童的肢體的表達能力、肌耐力、節奏感、柔軟性平衡感等身體能力。因此我們設計了一些動物動作培養兒童對舞蹈的興趣，例如模仿鳥兒飛翔的動作又或者想像力的形象動作，例如機器人的動作，而用一些基本的韻律有氣動作來增強兒童的肺活量與肌耐力，節奏感強烈的舞蹈動作來訓練節奏感，本研究之動作設計如表 2。

在音樂節奏比較快的部分我們會使用動作較敏捷、較有律動感的舞步，讓兒童增加對節奏的敏感度例如拍拍手動作與拳擊有氧，而在歌曲節奏較慢的時候我們則設定暖身的舞蹈動作，能夠有規律的收縮肌肉和放鬆，可減少運動傷害。

表2 本系統所使用的動作類型

	
(a) 簡單的伸展動作，幫助放鬆身體	(b) 有氧動作，可幫助肺活量增加
	
(c) 拳擊有氧	(d) 快速的左右搖擺動作。
	
(e) 奮力一跳，訓練爆發力	(f) 側邊踢腳，可練習平衡感
	
(g) 簡單拍手動作	(h) 左右伸展操
	
(i) 簡單的搖擺律動動作	(j) 模仿機器人

使用者操作情況如圖 19，Kinect 攝影機放至於桌面，螢幕右下角為使用者骨架投影，因此使用者能夠自行調整與 Kinect 攝影機的最佳距離，系統啟動後虛擬教練為待機狀態，當使用者準備好後可拍手喚醒虛擬教練開始舞蹈。圖 20 為使用者動作錯誤時，由虛擬教練進行校正指導，在畫面上重複示範正確動作給使用者，此時使用者只需要仔細觀察虛擬教練，不必急於跟上動作。



圖 19 使用者跟隨虛擬教練的動作跳舞

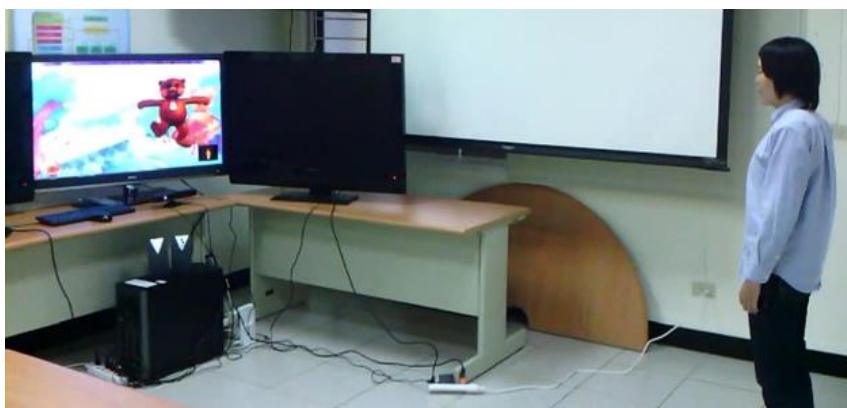


圖 20 虛擬教練指導使用者錯誤的舞步，重覆示範正確動作

伍、結論與未來展望

透過標籤式動作生成技術，並以模糊理論為基礎，對橋接影格建立標籤化，在切換影格時能夠產生自然流暢的動作轉換，動畫設計師不需要為每組舞蹈動作繪製補間動畫。當遊戲進行中，系統能即時的產生虛擬教練的反應動作，與玩家進行互動，就像是真的有位教練在教導使用者一樣，這種具互動性的虛擬角色，使用者能從回饋中學習到正確動作。

智慧電視將逐漸普及，智慧電視具備基本的辨識人體動作、聲控、網際網路，透過本研究所提出的方法，能夠為虛擬角色即時產生自然流暢的動作，未來可應用在各種運動遊戲學習軟體上，例如瑜珈、有氧體操等等。虛擬角色的互動動作也不侷限於動作指導，應用於社群連線的功能，除了讓更多親朋好友能夠連線，透過虛擬角色多元的動作的互動效果，讓社群娛樂互動更有趣。

本研究設定了 10 個特徵動作來代表虛擬教練的動作，然而，對於較複雜的流行舞蹈則顯得不足，不論是定義特徵動作，或者是動作橋接影格的數量，都必須設定的更嚴謹。而目前系統只針對使用者做錯的動作來給予指導回饋，使的系統略顯單調，未來應考量以下修改方向：

- 一、定義更精確的模糊函式，用來表示特徵動作。
- 二、使用基因演算法搭配模糊理論的方式來調整參數。
- 三、以建立模糊相似度函式量測不同動作之間的關連性，增加動作轉換的韌性。

參考文獻

- [1] 陳欣茹、黃永寬 (2009)。幼兒運動遊戲課程對幼兒創造力之影響 (碩士論文)。臺中市，國立臺灣體育大學。
- [2] 郭瑜婷、陳碧涵 (2004)。透過創造性舞蹈教學探討習舞兒童創造力之提升—以台中市三采藝術園區舞蹈班為例 (碩士論文)。臺中市，國立臺灣體育學院。
- [3] 蔡玲津、吳知賢 (2003)。電視幼兒節目中偶的運用分析與幼兒對偶的知覺之研究 (碩士論文)。臺南市，國立臺南大學。
- [4] 洪祥偉、陳五洲 (2006)。模型式虛擬實境技術於運動學習之應用 (碩士論文)。桃園市，國立體育學院。
- [5] 張育婷 (2001)。電腦多媒體輔助舞蹈學習之研究—以臺灣跳鼓為例 (碩士論文)。臺中市，國立臺灣體育學院。
- [6] 邱惠芬、林志隆 (2001)。多媒體介面對國小學童學習動機、學習成就及學習保留的影響 (碩士論文)。屏東市，屏東師範學院。
- [7] 傅俊德、陳五洲 (2009)。以3D線上電玩概念應用於程派高式八卦掌多媒體教學之研究 (碩士論文)。桃園市，國立體育學院。
- [8] 姚玉娟、區國良 (2010)。以手勢為基礎之遊戲式行動學習系統對學習成就及學習保留的影響 (碩士論文)。新竹市，國立新竹教育大學。
- [9] 陳志寶、陳登吉 (2000)。網路多媒體教材及題庫資源管理系統 (碩士論文)。新竹市，國立交通大學。
- [10] 張春興、林清山 (1981)。教育心理學。東華書局。
- [11] 胡名霞 (2006)。動作控制和動作學習。金名圖書。
- [12] 吳翎慈、石勝文 (2007)。太極拳輔助學習系統—使用加速規做師生姿態同步及比對分析 (碩士論文)。南投，國立暨南國際大學。
- [13] 顏羽君、方瓊瑤 (2007)。視覺式體操動作辨識系統 (碩士論文)。臺北市，國立台灣師範大學。
- [14] 王俊傑、方鄒昭聰 (2007)。3D/VR視訊學習平台—以Hip-Hop有氧舞蹈為例 (碩士論文)。桃園市，國立體育學院。
- [15] 林德隆 (1995)。外在回饋對運動技能學習的影響。中華體育季刊，9 (2)，67-72。
- [16] 吳亞屏、鄞宗賢 (2006)。基於標籤之動作合成方法應用於互動虛擬寵物開發研究 (碩士論文)。臺南市，南臺科技大學。
- [17] 王文俊 (2007)。認識 Fuzzy。全華圖書。
- [18] Schmidt, R.A. (1991). *Motor learning & performance: From principles to practice*. Human Kinetics Books.
- [19] Mayer, R.E. (2002). Multimedia Learning. *Psychol. Learn. Motiv.*, 42, 85-139
- [20] Elcacho, C., Schäfer, A., Dörner, R., & Luckas, A. (1998). Performing 3D scene and animation authoring tasks efficiently: An innovative approach. *Proceedings. Computer Graphics International* (pp.242- 244), Hannover, Germany,
- [21] Zong, D., Li, C., Xia, S., & Wang, Z. (2010). Instruction level motion control of virtual human for maintenance process simulation. *2010 International Conference on Audio, Language and Image Processing* (pp.1289-1294), Shanghai, China,
- [22] Lu, T., Ko, C., Lin, C., & Lin, T. (2011). Applying motion sensing technology to interact with 3D virtual

- characters. *2011 Defense Science Research Conference and Expo (DSR)*(pp.1–4), Singapore,
- [23] Oshita, M. (2008). Smart motion synthesis. *Comput Graph Forum*, 27(7), 1909–1918.
- [24] Egbert, C., Egbert, P.K., Morse, B.S. (2010). Real-Time motion transition by example. In: Perales, F.J., Fisher, R.B. (eds) *Articulated motion and deformable objects. AMDO 2010. Lecture Notes in Computer Science*, 6169 (pp.138–147), Springer, Berlin, Heidelberg.
- [25] Bandura, A. (1977). *Social learning theory*. Prentice Hall.