

數位家庭中智慧型互動電視之實驗平台的設計

An Experimental Platform for Smart Interactive TV in Digital Home

劉炳億

國立政治大學資訊科學系
g9527@cs.nccu.edu.tw

許書瑋

國立政治大學資訊科學系
g9204@cs.nccu.edu.tw

李蔡彥

國立政治大學資訊科學系
li@nccu.edu.tw

李宏偉

玄奘大學應用心理學系
spoon@hcu.edu.tw

黃淑惠

國立政治大學心理科學系
slh@nccu.edu.tw

摘要

視訊節目的觀看，是數位生活空間中不可或缺的一項娛樂。近年來，智慧型互動電視的應用是不少數位生活空間研究的焦點。然而，過去的研究多著重在軟硬體系統的設計及如何滿足使用者對娛樂節目的個人偏好上。我們認為，要設計出好的互動式數位電視系統，必須將使用者在觀看過程中的情緒與注意力納入考慮，方能設計出合用的智慧型人機介面。因此，我們的研究著重在藉由實驗設計輔助我們找出使用者在瀏覽互動電視時的行為特徵及主觀需求，未來透過適當的使用者塑模及行為偵測，我們希望能理解使用者的需求及情緒反應，進而解決使用者在當時情境下所遭遇到的問題。本篇論文提出一個進行中研究的階段性成果，介紹我們所設計的實驗方法和系統。實驗初期將會分成兩個階段，我們已自行設計完成實驗記錄與檢視的工具來輔助觀察受測者的行為，也將規劃不同瀏覽模式下的實驗平台，確保未來所得的實驗結論與智慧型系統設計能在各種不同的瀏覽情境下使用。

1. 前言

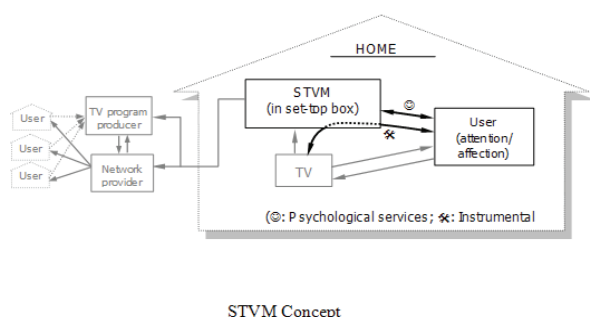
自從微軟和英特爾等 IT 巨人提出數位家庭 (Digital Home) 的概念後，這幾年來無論是產業界或學術界在相關研究上均有蓬勃的發展。隨著的科技發展，透過網路、行動通訊、晶片設計、電子、數位內容等整合技術，許多為數位家庭概念而設計的

產品也如雨後春筍般冒出。這股風潮未來將帶領使用者到達另外一種不同層次的生活方式與習慣。其中數位家庭的概念涵括了許多層面的應用，例如客廳的電視、電話，臥房的電腦、冷氣機，廚房的冰箱、洗衣機等眾多的電子產品，皆屬數位家庭的應用範疇；其中數位電視 (Digital TV) 的應用更是扮演著十分重要的角色。長久以來，電視一直在客廳中扮演舉足輕重的角色，幾乎每個家庭都會配置一台電視機；若電視機能夠數位化，再加上網路科技的普及及媒體內容的創新突破，那麼數位電視將會逐漸成為家中個人 PC 以外的第二個媒體娛樂中心。因此，像是目前次世代遊戲主機的兩大龍頭廠商—微軟和索尼 (SONY)，都相繼喊出以自家遊戲主機當做數位機上盒 (DSTB)，希望以電視結合網路的方式使其產品成為家庭客廳的數位娛樂中心。

隨著數位電視的誕生，「互動電視」 (Interactive TV, iTV) 的觀念在未來將逐漸取代傳統電視的瀏覽習慣。在互動電視的概念裡，希望使用者從以往「被動」的觀賞電視方式，轉變成「主動」提供自己的需求，向電視索取資訊，而產生互動的作用。舉例來說，像是股票查詢、即時路況、新聞、天氣、遊戲下載、數位教學等，都可以依照個人需求透過電視輕鬆取得這些服務。

因此，在未來數位生活的趨勢下，許多技術上

和內容上的需求也逐漸浮出檯面上；例如該怎麼感測人們的需求來主動提供服務，像是看到恐怖、厭惡的新聞畫面，電視感應到你的情緒反應而立即建議轉台或者產生一些緩和情緒的動畫效果；或是當你輾轉難眠時，收音機能夠自動撥放柔和的音樂來幫助你較好入眠等。我們希望能透過情意計算，讓科技更貼心的融入到家庭空間之中，進而創造更優質的居家生活環境。本論文的研究主軸在提出一個「智慧型電視管理員(Smart TV Manager, STVM)」的觀念，並建立所需要的互動電視實驗平台。我們的研究目標有二：一方面，我們將著重在建立一個以電視為基礎，且兼備情意計算服務的優質娛樂環境。另一方面，我們將互動電視的定義範疇更進一步延伸，藉由強化觀賞者與電視之間互動的能力，讓觀賞者在看電視時能有更好的體驗。如圖一所示，我們將 STVM 整合在數位機上盒中，使用者透過遙控器(Remote Control)和機上盒溝通，來幫我們管理電視互動的部分。



圖一、Smart TV Manager 的觀念

為了達到 STVM 的目標，我們有兩項前置工作必須先做：(1)首先需要觀察使用者在觀看電視時，有哪些行為特徵(Behavioral patterns)和觀看電視時主觀上的個人需求(Subjective needs)。(2)我們需藉由多模(Multimodal)的方式[2]來建立生理上和行為上的辨識系統；這個系統包括眼動儀、肢體和特定姿勢的辨識系統、臉部表情辨識、腦電波偵測儀器(EEG)等。

在第一項工作中，我們將探討使用者在看電視以及對電視的個人需求，以規劃互動電視之功能，

並有助於未來互動電視使用者塑模(User Modeling)的工作。傳統上，過去相關研究所提議建立於互動電視上的使用者模組，通常都是應用在電子節目選單(EPG)上，為了推薦給使用者喜愛的媒體內容而去紀錄使用者的觀賞偏好行為，而不是為了了解使用者觀看電視時的行為、動機以及相關生心理狀態關係，對於情緒反應的探測更是少見。因此，我們將收集使用者的習慣行為和情感資訊，來推斷使用者的個人特質和情緒反應。使用者模組未來該收集哪些個人特質的行為資料和如何對應使用者的情緒及注意力，將會是我們研究關心的焦點。長久以來，資訊系統多由軟體設計師負責規劃設計，較缺乏心理學的專業評估，因此本研究也希望透過跨領域的合作，使實驗設計及分析上能夠更加完善，讓科技和人性更為貼近。

接下來在這篇論文中，我們會在第二節開始探討相關背景及研究，審視過去這幾年來互動電視的應用，並簡介未來實驗主要平台為何挑選 Windows Media Center(MCE)來做為基礎平台，以及傳統使用者模型在互動電視上的研究和扮演的角色。而第三節裡，我們會介紹實驗系統的設計動機與目的，以觀測使用者在瀏覽電視時有哪些行為特徵和主觀上的個人需求，並對於情緒上的認知能更深入探討。在第四節裡，我們將會規劃兩階段的實驗來觀察量測，並會設計不同情境下的電視觀看模式來觀測使用者行為與情緒反應。第五節中，將介紹我們所開發的 MOD 觀看模式的實驗平台。另外，我們也將設計一套實驗觀察工具，以增進觀察上的效率和使用模型偵測的輸入來源。最後，在第六節中，我們將會為目前成果做個簡單的結論，並說明往後實驗的規劃和待努力的方向。

2. 相關背景與研究

2.1 互動電視平台之現況

以傳統的電視觀看經驗而言，使用者只能被動的依照頻道商所提供的節目表來定時收看，無法掌

握時間上的主動權。因此在數位化後的互動電視上，不只把媒體內容傳播的過程數位化，而是更進一步希望使用者能夠擁有時間上的主動權，主動索取媒體內容。例如中華電信的多媒體隨選系統(MOD)[9]即是互動電視初期的應用之一，MOD 是利用網路串流技術將豐富的高畫質數位影音內容送進電視，拋棄一般傳統第四台需要「定時收看」的束縛，取而代之的則是「隨選即看」的新體驗。此外 MOD 中的影片也可以輕易的透過遙控器快轉或者暫停等，想看哪裡就看哪裡而不受限制；MOD 除了觀看影片外。也可提供家庭金融或者氣象查詢等功能，方便取得生活資訊。

除了 MOD 之外，近年在美國相當風行的 TiVo 也已經在台灣上市[8]，TiVo 有別以往其他的錄放影機，能透過網路取得當地有線電視節目表，而使用者可以由簡單直覺的方式錄製節目，不需要記得節目時間，往後即使節目時間更動，TiVo 也將會自動更改錄製時間。在觀看現場節目(Live TV)的過程中，也可自動進行錄製的動作。並在邊看邊錄中，隨時可以暫停倒轉。而這樣的操作方式也與互動電視的基礎觀念相當契合，主動掌握看電視的媒體內容。因此，TiVo 在美國是相當受人歡迎的一項產品。

另外，在 2006 年 MIT Media Lab 的 Lee 發表了一項 HiTV 的研究[3]，允許觀眾直接對電視傳達情緒上的反應；例如使用者看電視時，看到不喜歡的角色或者劇情時，可以用他們所設計的一顆軟球丟向電視，螢幕上的畫面就會有類似變形爆炸的效果。像這樣可以直接將情緒宣洩於電視上的互動方式，產生的動畫特效具有調節情緒的作用。有別於以往和電視互動的方式，僅限制於遙控器的操作上，而引用多模的觀念方式，允許多個輸入來源來和電視進行溝通，此類多重介面在互動電視上的研究具有相當高的創意啟發性。

2.2 MCE 介紹

一般來說，互動電視方面的研究多數都是以個

人電腦(PC)模擬互動電視的介面，許多媒體操作功能需另外開發撰寫，未來還需要移植到目標平台（如機上盒）上，因此整個互動電視的實驗環境，在架設上需要花費許多心力。這樣的平台若自行獨立開發，所需耗費的時間與人力成本將會十分多。因此，在本研究中，我們選用了微軟 Microsoft Windows Vista 中的 Windows Media Center Edition (MCE) [10]來做為實驗平台的解決方案，並以 MCE SDK 5.0 來開發相關的應用程式。

MCE 提供了強大的家庭影音整合平台，它整合上網、觀賞 DVD、看電視、聽音樂、燒錄光碟等各種多媒體應用，並且可使用遙控器簡化這所有的操作。目前市面上也有 MCE 機上盒的產品，連接液晶電視和網路，即可成為一個互動電視的作業平台。與 Tivo 之類的系統不同的是，MCE 可以提供開發者一個方便的開發環境，允許開發者建立額外的程式在 MCE 的平台上，包括啟動時的背景應用程式(Background Application)、另外點選執行的附加應用程式以及網頁應用程式。程式語言可使用 C# 或者網頁語法來撰寫，而附加應用程式的介面(UI)則是利用 Media Center Markup Language (MCML)設計的。因此，MCE 平台在環境架設上以及往後的功能整合上，都提供相當有效率的解決方案。

2.3 使用者模組於電視上的相關研究

近年來，藉由使用者模組來達到個人化電視的相關研究越來越多，如 O'SULLIVAN 等人[5]在 2002 年曾指出，當使用者瀏覽行為等資訊較少時，其推薦節目系統時所面臨的問題。因此，他們建議利用資料探勘(Data Mining)等相關技術來輔助改善 EPG 中推薦節目的成效。另外，在 2003 年 Hara 等人[1]則是分析了日本人在看電視時的類型，並加以分類。另外也有些研究不只探索個人興趣，而更考量到看電視通常是一種社群行為，往往電視機前面的觀眾不只一個；如 2003 年 Masthoff 等人[4]在推薦節目時考慮到團體關係。2005 年 Yu 等人[7]也是

提出利用多個使用者紀錄檔(User Profile)合併的技術，來針對多人推薦節目。上述許多使用者模組的相關研究，大都探討使用者觀看電視的個人偏好或者家庭群組共同喜愛的媒體內容。目前塑造使用者模組於電視上的應用也大多是如此。

但人們有些時候並不是單純喜好某類型的影片或節目，而是因為當下的情緒產生特殊的需求。例如心情不好時，有些人喜歡看綜藝節目紓解壓力，平時喜歡看的文藝劇情影片則絲毫不想碰觸。加上未來互動電視中，人們對於媒體內容將會更有掌控權，隨之而來的多媒體應用將會日漸豐富，像是可以參與節目的互動過程，用電視遙控器即可購物或者投票等，使用者所面對的不再只是單純的收看電視節目，而是數以百計的相關多媒體應用服務。因此，如何能夠掌握使用者的情緒將是未來互動電視中使用者建模的一個重要課題。

3. 實驗動機與目標

我們實驗系統的建置目標，在於觀察並量測使用者瀏覽互動電視時的行為狀態和個人需求。行為狀態通常又可分成長期行為和短期行為；長期行為比較偏向於使用者的個性特質及興趣偏好，這方面需要較長時間的統計觀察來才能作出推論，屬於較為隱性的行為；而短期行為則是看電視時當下的情緒反應，外顯行為比較明顯，但行為的生命周期較短，過一段時間可能就消失了；例如，看影片時受到劇情驚嚇而引發的恐懼反映。通常來說，多數使用者模型於互動電視中的相關研究較多屬於長期行為的量測，而且偏向研究使用者對於媒體內容的興趣程度，對於個人特質的探討研究相對比較少。另一方面，量測使用者在觀看電視時的情緒反應的相關研究則更少，可能因為這樣的研究通常需要情意計算技術的支援。情意計算自 1997 年由 MIT Media Lab 中的 Picard 教授提出以來，目前仍算是一個新興的領域，雖然這幾年來投入建立情緒模型的相關研究從未停過，但由於人類的情緒包羅萬象，其複雜程度相當難以探測，即使是人類相處之

間有時也難以了解對方目前情緒為何，因此目前進展仍是有限，也尚未被廣泛的應用。我們希望除了能在使用者長期行為的量測上更為深入之外，主要對於短期行為上的辨認，像是情緒和情感狀態能夠有一定程度上的了解。所以，我們未來對於 STVM 的使用者模型會結合長期行為和短期行為的量測，適時給予適當的輔助回饋，提供給使用者觀看互動電視時更佳的體驗。接下來，我們會設計實驗來解決這樣的問題，並取得未來使用者模型分類行為資料的依據和驗證其效用性。

目前我們規畫兩個階段的實驗。在第一階段中，主要是觀察被實驗者在看電視的行為上有哪些特殊的行為訊息可以提供給我們，所以我們會設計一些觀察工具來輔助紀錄。這些觀察到的資訊量將會非常的多，因此我們會整理出其中較為重要顯著的行為特色出來，來做為我們之後第二階段實驗和 STVM 中使用，這也是我們第一階段實驗主要的目標。

在第二階段的實驗裡，我們會建置使用者模型，而使用者模型需要取得哪些行為特徵資訊則是依賴我們第一階段的結果。有別於第一階段實驗，我們不再只是單純的觀察行為，而是會根據受測者目前的狀態給予一些回饋，觀察使用者情緒或者操作上是否有獲得改善，因此被實驗者和電視的互動過程也會較頻繁。在這個階段的實驗中，我們希望能夠評估第一階段的成果是否有效之外，也能夠對之後 STVM 系統的互動功能設計上有些啟發上的提示。這將是我們第二階段實驗主要的目標。

4. 實驗系統設計

4.1 第一階段實驗

第一階段實驗中，我們希望實驗環境除了使用者以自然狀態下觀看傳統電視的行為外，也能進一步了解特定主題影片對使用者行為所產生的差異。因此，我們的實驗環境將會分成傳統電視的瀏

覽環境和可控制內容的電影播放環境兩種：

• **瀏覽一般電視之環境 (Nature TV Viewing)**：受測者觀看傳統電視時，除了選台外，不會去影響頻道媒體內容。此環境主要的目的在於觀察受試者在瀏覽傳統電視時的行為特徵，其情境因為較輕鬆自然，長期行為的表現較容易出現。

• **瀏覽特定電影之環境 (Selective Movie Viewing)**：將會從過濾篩選過的幾十部特定主題的電影中，選出一定數量的電影（或片段）提供給受測者觀看。我們希望在觀看這些電影時，能營造出激發特定情緒的情境，其行為特徵也比較為顯著突出，容易引發短期行為的情緒性反應，以方便我們深入觀察。

不論是一般電視的收視或特定電影的觀看，我們都需要一個整合的電視瀏覽系統以做為實驗的平台。理想中的電視觀看實驗平台應能支援上述兩種觀看情境，並且能方便的安裝在使用者家中，讓使用者能經由網路點選收看統一提供測試的影片。若是我們想要更換測試主題，就只需要在伺服器端更改即可，增進我們第一階段觀察實驗設計上的效率。除了電視觀看的環境外，我們還需要設計通用的實驗觀察工具，能記錄使用者在觀看電視的過程中曾發生過的事件（包含使用者控制遙控器的事件），以利後續的行為分析。在未來第二階段和 STVM 中，這些觀察工具也都會繼續扮演量測使用者狀態的輸入來源，並且我們也會盡量以非入侵式的方式觀測受測者的行為，確保實驗的公正性。

在第一階段的實驗裡，我們將首先進行傳統電視環境的觀察，讓受測者在觀察時間內自由瀏覽電視頻道，我們則利用觀察工具記錄下外在行為、頻道內容和遙控器按鈕行為等，由這些外在的行為和當時媒體內容來進行判斷。實驗後也會針對受測者進行更深入的訪談，對於一些較為特殊的行為訊息來詢問他們當下的想法和需求，並藉由心理學的認知專業進行綜合判斷，分類出使用者在觀賞電視時有哪些重要的特徵和其觸發對應的事件。有了這些

初步的實驗結果後，我們會再以 MOD 系統播放特定的主題電影給受測者，引發特定的情緒來觀察使用者的行為表現，這些主題將會涵括 Ekman 等所提出的六大基本情緒：快樂(happy)、憤怒(anger)、害怕(fear)、驚訝(surprise)、哀傷(sad)及厭惡(disgust)。以這些基本情緒狀態主題作為基礎，我們將瞭解這些基本情緒的分類是否適用於電視觀看過程中的行為。除此之外，這些實驗結果也可幫助了解人們在看 MOD 和傳統電視這兩者不一樣的收視模式下是否會有行為上的不同。

4.2 第二階段實驗

在第二階段的實驗中，第一階段實驗後的瀏覽行為將做為我們使用者模型分類行為情緒的來源依據，觀察受測者在看電視的過程中，是否有特定的行為或情緒反應出現，系統再給予適當的回饋來觀察使用者的反應。這些回饋一開始將由我們多數討論的結果來設計，之後則依據使用者的反應來調整，因此第二階段實驗將會是反覆性的實驗過程。而我們觀測的重點也將分成三個主要層次：**(1)環境層次**；探討使用者目前的環境因素使得使用者產生哪些行為，或者哪些特殊事件造成使用者狀態的轉移。**(2)行為層次**；觀察環境因素或者事件所引發情緒而造成的肢體行為動作，希望能歸納出是否有多數相似的行為，從異中求同裡來定義出某類情緒可能會出現的行為動作。**(3)生理指標層次**；我們將會偵測使用者特定情緒產生時，當下的生理指標變化，例如心跳、腦波、血壓、體溫、呼吸頻率等，並將這些資訊和情緒狀態之間進行對應。

在量測工具和互動介面上，我們將以多模方式設計。傳統上，觀測使用者的行為多以遙控器的按鈕選擇和對應的媒體內容進行交叉分析，而這樣的方式通常只用於觀測於長期行為的興趣偏好。若是想要更進一步了解使用者狀態的轉移以及情緒的探測，則比較不足。而在互動介面上，傳統人機介面也是以鍵盤滑鼠或者遙控器來溝通，其他型態的互動界面較為缺乏。因此我們將以下面三個角度來

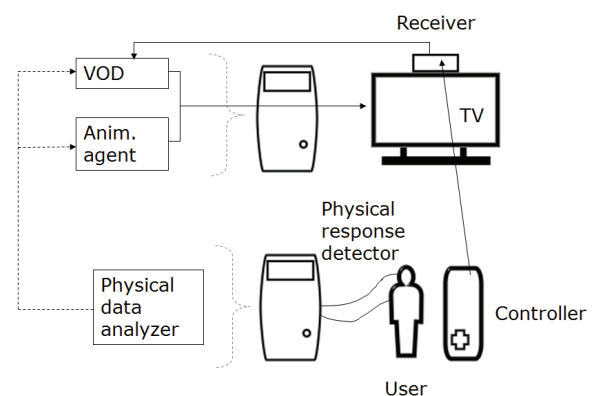
規劃量測工具以及設計互動介面：

- **視覺(Visual)** — 即是以電腦視覺的技術來輔助找出使用者目前的外顯行為狀態。傳統上，主要以臉部表情辨識(Facial Recognition)或身體姿勢辨識為主要的處理方式。臉部辨識常用於輔助辨認使用者的情緒，不過在複雜背景下如何提高臉部表情辨識的正確率，以及在一個空間中如何正確定位出人臉位置，這些相關問題目前仍是一個挑戰中的議題。身體姿勢辨識則通常用於注意力偵測或者傳達訊息的輔助技術，像是我們可以利用運動偵測(Motion Detection)來辨認使用者的注意力目前是否專注在電視上，或者人已經離開客廳了。而肢體辨識(Gesture Detection)則是可以利用一些特定手勢動作，傳達指令。例如，手勢往上或者往下對應到頻道的上下切換。
- **聽覺(Audio)** — 主要的應用多以使用者利用自然語言對電腦下達指令，以語音辨識(Speech Recognition)技術為主。例如，使用者直接利用特定的字彙語言表達目前的情緒狀態或者操控特定功能。目前語音辨識的精確度在有限字彙的限制下已越來越高，不過在觀看電視的應用中，因為電視聲音和人聲會混在一起，因此如何過濾隔離出人的聲音，這也是一個挑戰中的重要問題。
- **觸覺(Tactile)** — 傳統上，有些偵測使用者的生理狀態，如手部皮膚的導電度、肌肉繃緊程度或者心跳速度等，則是依靠使用者接觸或者穿戴特殊裝置來偵測。另外，由任天堂公司所研發的 Wii 主機，其中的 Wii Remote 包含了指向定位和動作感應，提供給使用者另一種新型態的人機互動介面。

目前上述多模方式的人機介面研究大都以人類和電腦為基礎[6]，而以人類和電視的相關研究則比較少見，因此我們會針對以電視為基礎，整合現有的技術以及工具來設計適合我們的多模介面，並

以遙控器輸入、臉部表情辨識、腦波儀、眼動儀等作為量測的來源，進行綜合分析。

在實驗環境的需求上，我們希望完善觀測使用者在不同情境下瀏覽電視的狀態，並且在實驗過程中可以製造特定事件以刺激使用者的情緒。因此，第二階段實驗觀測環境將分成在下列三個平台進行：**(a)多媒體隨選系統(MOD)**：由於這是目前互動電視中最为普及的一項應用，並且如前所述，可以觀察特定主題的題材影片，以方便我們量測短期行為的情緒特徵。而系統將延續使用第一階段實驗所開發的 MOD 系統；**(b)模擬第四台選台系統**：因為傳統第四台仍是目前主流的收視方式，但傳統第四台對於我們來說，並無法掌握媒體內容以及頻道順序，想要製造特定事件來觀測也因此較為不易，所以我們將自行開發能模擬第四台系統的電視播放系統，讓使用者可以選擇頻道，而影片有平行時間播放的觀念，在實驗過程中便可以產生特定主題事件來觀測，並掌握更多的媒體內容主導權；**(c)傳統 Live 第四台**；即是一般第四台的實驗環境，頻道媒體來源由第四台有線公司提供。因為目前媒體來源主要仍是以傳統有線電視為主，因此我們希望能測試在 EPG 配合現有有線電視的框架限制下，使用者模型可以應用到何種程度上。此外，我們將以可控制下的第四台模擬環境設計實驗組及對照組，觀察使用者模組對於受測者觀看電視的體驗經歷是否有改善，以及改善的程度為何。



圖二、第二階段實驗環境平台概觀圖

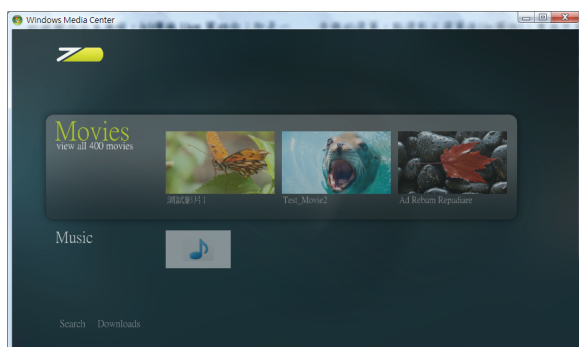
最後，第二階段整體的實驗平台如圖二所示，使用者觀看電視時，我們將會藉由觀察工具記錄使用者行為之外，也會搭配生理指標和行為的辨識系統來偵測特殊行為事件，透過此分析結果及使用者模型的資訊，適度地回饋使用者，而回饋的方式將包含以文字框訊息、動畫方式或者其他可能的介面，來呈現互動電視代理人(Agent)，輔助使用者操作上的困難或者緩和情緒壓力等效果。

5. 系統設計與實作

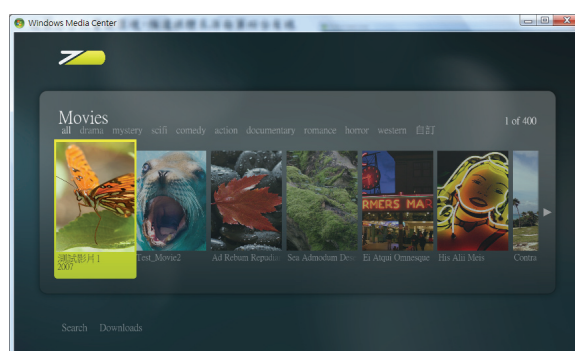
5.1 MOD 實驗平台

由於在電視觀看實驗上的需求，我們需要一個能提供影片內容及設計特殊事件的實驗平台來觀察短期行為的反應。在本研究中，我們以挑選 MCE SDK 5.0 來實作我們的 MOD 系統。MCE SDK 中附有類似 MOD 的樣本程式，提供了一些基礎的介面來輔助開發者設計出瀏覽電影電視及聽音樂的平台，因此我們以此專案程式為基礎，將系統改寫成能符合我們實驗需求的系統。

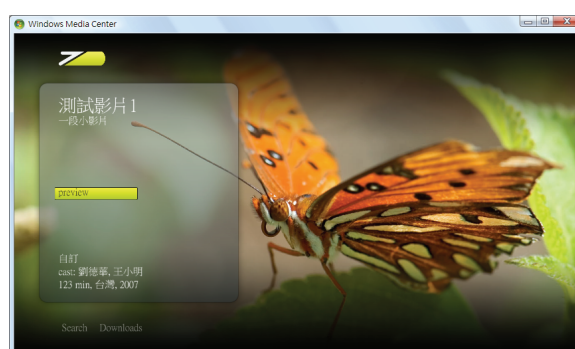
目前這個系統的介面如圖三所示，包含了瀏覽影片和音樂的選單項目，點選影片項目後(如圖四所示)，畫面中第一列項目可以依照影片後設資料(metadata)的類型(genre)進行篩選，使用者在點選有興趣的影片封面後，會進入到另一個介紹影片內容的畫面選單(如圖五所示)，顯示影片的大綱介紹、影片類型、演員陣容等相關資料，透過這個選單，使用者即可下載影片來觀看。



圖三、MOD 主選單畫面



圖四、MOD 影片瀏覽畫面



圖五、MOD 影片選單

目前我們已完成這個系統的雛形設計，未來我們將以這樣平台做為我們觀看電視的實驗環境，提供給受測者觀看特定主題影片時使用，並且也將會依照兩方面的需求繼續擴充平台；第一方面是根據實驗上的需求製造特定事件，例如電話鈴聲來了，此時電視螢幕上便出現動畫角色，詢問是否要降低音量，藉此觀察受測者的反應。第二方面為根據受測者的需求所做的輔助功能。當實驗經過一段時間後，整理出受測者的功能需求後，我們將依此設計適時的情緒調節等功能，並在未來的實驗中持續觀察受測者的體驗是否得到改善。在未來的 MOD 系統中，我們也將會根據使用者模型分析的狀態和使用者的進行其他更深層的互動。

5.2 實驗記錄與檢視工具

為了更進一步瞭解使用者在觀看電視時的行為以及引發這些行為的背後原因，我們設計了一套

實驗紀錄工具，能夠以側錄的方式收集有關使用者整個收視過程中的重要訊息，提供我們進行深入分析時所需的資料。收集的訊息包含了節目內容、受試者的肢體動作及遙控器事件。



圖六、實驗記錄所需的硬體裝置

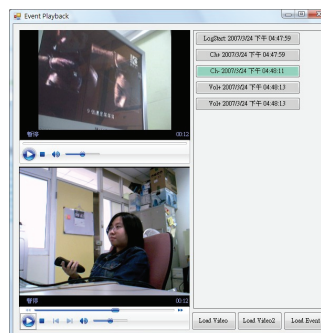
如圖六所示，實驗記錄工具的硬體部分，是由一台電腦主機，搭配兩支視訊攝影機以及一組紅外線接收器所構成。電腦主機用於運行記錄工具的軟體與做為訊息記錄的儲存空間。在兩支視訊攝影機中，一支是用來拍攝節目畫面，另一支則是用來拍攝使用者收看過程中的肢體活動行為。這兩支攝影機除了拍攝畫面之外，也會一併記錄聲音。而紅外線接收器則是用來側錄使用者操作電視遙控器時所發出的訊號，像是切換頻道或調整音量等事件。我們也實作了 MCE 的系統上的遙控器事件側錄機制，但由於 MCE 的系統尚未普及，因此目前在沒有使用 MCE 系統的家庭裡，我們仍可以紅外線接收器側錄傳統電視的遙控器訊號。透過這些機制，我們將遙控器事件及其開始與結束發生的時間一併記錄下來，以供未來透過時間的標記將這些由不同頻道所收錄的訊息同步化，以供進一步的分析。



圖七、電視觀看實驗之記錄工具

如圖七所示，電視觀看實驗之記錄工具會隨者使用者打開電視而開始運作。每當使用者按下電視遙控器上的電源鈕來開啟電視時，記錄工具就自行啟動。實驗過程中，兩支攝影機開始分別捕捉使用者收看的節目畫面與使用者本身的活動行為，而紅外線接收器也同時開始側錄電視遙控器訊息。當使用者再度按下電源鈕來關閉電視，記錄工具也會自動停止記錄，但捕捉使用者活動行為的攝影機會在電視關閉後額外多錄製一段時間後才停止，其原因是我們認為關機後的一小段時間，使用者的行為也可能隱含了一些訊息，例如，受試者關閉電視的原因。

另外我們也設計了一個檢視工具，協助我們在初期以人工的方式來分析這些紀錄工具所錄下來的訊息(如圖八所示)。在檢視工具的左邊呈現的是兩支攝影機所拍攝的畫面，上面為節目影像，下方為受試者活動影像。右邊則是遙控器事件發生列表。檢視工具在播放影片時，會自動同步我們所記錄的兩段影片與事件。例如，當你透過控制面板調整左上電視節目影片的時間時，左下的使用者活動影像的影片也會自動跳到對應的位置。而當影片播放至有遙控器事件發生的時間點時，列表中該事件的標籤也會呈現出不同顏色以資區別。我們也可以點擊某個特定事件，此時兩段影片會自動跳到對應的時間點來播放。透過這個工具，分析人員可以清楚的瞭解整個受試過程發生的事情。



圖八、已開發完成的實驗事件檢視工具

我們的實驗觀察與檢視工具最大的好處就是

通用性與擴充性非常高。無論受試者進行實驗的地點是在實驗室，或者是在自己家中，收看的節目是傳統第四台或是使用 MOD 系統，都可以進行記錄的工作。在未來，我們希望強化檢視工具，加入一些加速人工分析，甚至是自動分析的功能。例如，有些片段使用者正在專心的觀賞節目，長時間沒有產生太大肢體活動與遙控器訊號，我們可以使用影像辨識的技術來將這些片段自動過跳過，以節省分析時間。還有，我們也可以使用資料探勘的技術來自動分析使用者的按鈕行為是否有特殊的樣式，而這些樣式又是否對應到特別的節目內容，藉此瞭解使用者的習性。至於擴充性，除了目前的影像與遙控器訊號，未來也可以將其他的生理儀器的訊號整合進來，像是腦波、心跳、血壓以及體溫等。

6. 結論與未來展望

目前我們第一階段的實驗環境（包括實驗觀察錄影及檢視工具）已經建置完成。第一階段的實驗也正在進行中，希望藉由此次的實驗能夠幫助我們整理出受試者的行為特徵，往後可以透過一些工具與技術來取代人工分析。未來我們也會持續的進行實驗，擴充現有工具的功能性，以及繼續開發尚未完成的實驗觀測平台，包括模擬第四台環境以及整合生理訊號等。

最後，目前情緒偵測的議題仍是一個相當困難的問題。雖然已經有越來越多的相關研究持續的在解決這樣的問題，並且也有一些不錯的成果，但是離精確的判斷仍是有一段要走。因此，我們未來的 STVM 主要目的，並不在於能夠準確的判斷使用者的情緒狀態，而是希望藉由這樣的設計，輔助使用者在操作上遇到困難時或者情緒狀態有較大轉變時，使得操作上更便利貼心，以及理解情緒變化背後的原因，進而幫助他來改善情緒，帶給使用者在觀看互動電視上擁有更佳的使用體驗。

7. 致謝

此研究在國科會 NSC95-2627-E-004-002 計畫的支助下完成，特此致謝。

8. 參考文獻

- [1] Y. Hara, Y. Tomomune and M. Shigemori, "Categorization of Japanese TV Viewers Based on Program Genres They Watch," *User Modeling and User-Adapted Interaction*, 14(1):87-117, 2004.
- [2] A. Jaimes and N. Sebe, "Multimodal Human Computer Interaction: A Survey," in *Proc. of IEEE Int. Workshop on Human-computer Interaction*, 2005
- [3] C.-H. J. Lee, C. Chang, H. Chung, C. Dickie and T. Selker, "Emotionally Reactive Television," in *Proc. of Intelligent User Interface '07*, pp. 28-31, 2007.
- [4] J. Masthoff, "Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers," in *User Modeling and User-Adapted Interaction*, 14(1):37-85, 2004.
- [5] D. O'Sullivan, B. Smyth, D. C. Wilson, K. McDonald and A. Smeaton, "Improving the Quality of the Personalized Electronic Program Guide," *User Modeling and User-Adapted Interaction*, 14(1): 5-36, 2004.
- [6] M. Pantic and L.J.M. Rothkrantz, "Toward an Affect-Sensitive Multimodal Human-Computer Interaction," in *Proc. of the IEEE*, 91(9):1370-1390, 2003.
- [7] Z. Yu, X. Zhou, Y. Hao and J. Gu, "TV Program Recommendation for Multiple Viewers Based on user Profile Merging," in *User Modeling and User-Adapted Interaction*, 16(1):63-82, 2006.
- [8] TiVo, <http://www.tgc-taiwan.com.tw/>
- [9] 中華電信 MOD, <http://mod.cht.com.tw/MOD/Web/index.php>
- [10] 微軟 MCE, http://www.microsoft.com/windows/products/windows_vista/features/details/mediacenter.mspx