

I Know You: 智慧型家庭網路影音傳送系統

黃宏宗¹ 林正倫¹ 阮明英¹ 胡誌麟¹ 蘇奕宇² 劉炳傳²

¹國立中央大學通訊工程系

²台灣工業技術研究院

{101523047, 101523036, 101523601, clhu}@cc.ncu.edu.tw; {pioneersu, flashliu}@itri.org.tw

摘要

本論文在家庭網路的環境下，提出一套新穎的智慧型家庭網路影音傳送系統，提供使用者在家庭網路環境中不受居家空間限制的多媒體影音服務。此系統之設計整合 UPnP 網路內容分享平台、人臉辨識及人機互動介面設計等技術，並搭配家庭成員資料庫之建置，使得家庭成員在不同居家空間之間移動時，系統可以即時感知家庭成員所在位置的變換，提供個人化的影音傳送及播放服務。在實作與展示中，我們經由實機展示來具體呈現出這一套智慧型家庭網路影音傳送系統的功能，達成家庭網路環境下具隨意、隨地及隨選裝置特性之即時性與個人化的多媒體影音傳送服務。

關鍵詞：UPnP, 人臉辨識, 多媒體影音, 家庭網路, 數位家庭

1. 前言

隨著家庭寬頻網路的滲透率提升，各式新興網路多媒體裝置的推成出新，數位家庭的理念因應而生，也帶動網路多媒體裝置產品的迭代出新。在廣泛的數位家庭應用範疇中使用者最感興趣的主要是在家庭影音娛樂、家電自動化控制及居家安全監視。就最貼近消費市場的家庭影音娛樂而言，家庭成員希望可以在家中不同的空間，利用各式網路家電來和家人及朋友分享各種影音內容，也希望網路家電裝置能一改舊往只能被動的接收指令，可以主動地提供家庭成員有關家庭網路中最新分享的或是經常播放的影音媒體資訊。

本文提出的智慧型家庭網路影音傳送系統，採用符合國際數位家庭產業組織 Digital Living Network Alliance (DLNA)所規範的 Universal Plug and Play (UPnP)相容性技術[1]，在 UPnP 平台架構上發展 UPnP AV 多媒體影音內容分享服務，讓家庭網路環境下的網路裝置，可以發現彼此的存在、互相溝通及分享各自擁有的多媒體檔案。另一方面，對於家庭成員在居家環境中的移動和裝置使用習慣，此一智慧型系統支援多媒體內容同步播放及顯示服務，這服務的設計是在多媒體播放器裝置端結合人臉辨識技術，偵測靠近裝置的人臉影像畫面，辨識每一位家庭成員的身份。當使用者在居家空間

內移動時，鄰近的多媒體播放器裝置可以發現某一使用者進入它的服務範圍，並立即在家庭網路中發佈這項消息，若家庭網路中有其他正在服務這位使用者的多媒體控制器，它將從網路中得知使用者位置的改變，隨即啟動同步轉置和播放功能，讓多媒體伺服器把正在播放的多媒體內容，改傳送到使用者當下所在位置附近的多媒體播放器端。

因此，本文所提出的智慧型家庭網路影音傳送系統的設計，可相容於 DLNA 和 UPnP 數位家庭產業標準，並在多媒體播放器裝置端結合人臉辨識技術，使得系統可辨識每位家庭成員，提供個人化的多媒體影音服務之目的。系統雛形的開發與實作部分，將是以消費市場既有的網路多媒體裝置為應用載具，實機展示出該系統的功能，讓使用者在家庭網路環境下體驗不同於以往的家庭網路多媒體影音服務，帶給使用者全新的使用者情境和體驗。

本論文的章節組織如下所述，第 2 章提出相關文獻與技術，第 3 章說明系統架構與功能，第 4 章描述系統離型設計與實作，最後第 5 章總結本文的研究成果和展望。

2. 文獻探討

近年來許多高階嵌入式網路通訊裝置除了配備高級的資料處理器之外，更具備有相機模組及多媒體處理器的升級，因而有越來越多的影像辨識與 UPnP 的應用服務陸續提出，以下我們將分別對 UPnP、影像辨識和這兩項技術結合之應用作探討。

2.1 UPnP 之應用

[2]首先在家庭網路的環境下引薦入行動多媒體伺服器一角色，成功的透過 UPnP AV 模式，以簡易的網路遠端操控取代傳統的手動繁瑣操作，把行動裝置上的影音媒體轉置到網路電視上播放，展現出行動影音媒體分享的使用情境。[3]提出一套結合 UPnP 和 P2P 的家庭多媒體開道器設計，讓支援 UPnP 的多媒體播放裝置可以直接執行來自外部網路的 P2P 檔案下載及串流服務，使得媒體來源不再侷限於家庭網路內部的媒體內容，延伸至網際網路上龐大的 P2P 多媒體內容來源。[4]則是在 UPnP 裝制架構上整合近場通訊技術(NFC)，提出一個新穎的手勢輔助的遠端操控方法。上述為本文作者先前

發表的部分研究成果，讀者可參閱[2]-[4]所屬的參考文獻，將可清楚 UPnP 技術與應用發展。

2.2 影像辨識之應用

影像辨識知識及技術應用相當廣泛，如車牌辨識、手勢辨識、人臉辨識等。如[5]所述，智慧型互動應用可分為兩類：人臉辨識在「智慧型環境」下的應用，以及人臉辨識在「智慧型裝置」上的應用。以上兩類是基於人的局部外貌之特徵，比對過去或事先已註冊的人臉圖資，辨識使用者的身份。然而由於影像辨識的運用相當仰賴裝置載具的資料計算能力，因此過去的應用範疇較少套用在消費性電子產品上，近年來許多高階消費型電子終端裝置，如行動通訊裝置、平板電腦及電視，其硬體規格大幅提升亦包含視訊模組，基於這類裝置的影像辨識技術與應用將很有發展性，因於是這類裝置的主要操作者，人臉辨識技術的應用特別值得留意。

2.3 UPnP 結合影像辨識之應用

由於在網路家庭的環境下可能同時存在許多的媒體播放裝置，透過 UPnP 所發現的網路裝置名稱經常是製造商所給定的名稱，對消費者來說這類的命名方式並非直覺或主觀上可以輕易辨認出所屬的特定裝置，因此[6]提出一個較直覺的裝置操作方式，透過 UPnP 結合電腦視覺，讓使用者在行動裝置鏡頭所拍攝的畫面中做裝置的選擇，然後再對該裝置提供影音媒體分享。

由以上的討論中可以看出，儘管 UPnP 和人臉辨識兩者雖非全新提出的技術，但是至今尚未有具體整合兩技術並應用在數位家庭環境下的網路媒體服務，是以本文所擬之系統旨在整合兩大技術並設計新穎的多媒體影音轉送服務。

3. 系統架構與功能

以下將會詳細說明此系統的功能設計和運作方式，3.1 節描述系統架構及各子系統的功能設計，3.2 節將從系統面說明整體的運作流程。

3.1 系統架構

如圖 1 所繪，本系統的應用範圍為一個包含有網路多媒體媒體家電、個人電腦、網路儲存設備(通稱為多媒體網路裝置)的家庭網路環境。通常這種網路是由多媒體網路裝置透過固接網路或無線網路介面連接至路由器所形成的小型區域網路，多媒體網路裝置將可以從路由器取得 IP 位址，並利用這些裝置之間共同的 UPnP 通訊協定，以 IP 網路封包傳送與接收取得不同裝置的資訊、亦經由 UPnP 訊息來操控裝置的功能。另外，多媒體網路裝置可經由視訊模組的功能，擷取前方的影像交予後方軟件進

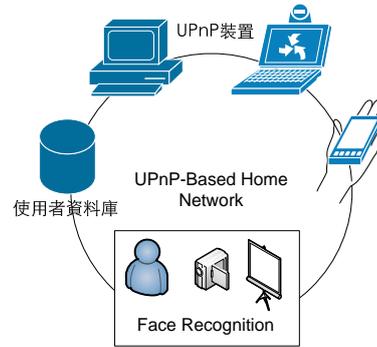


圖 1 系統環境示意圖

行影像辨識，辨識所得的資訊將可提供給區域網路中各個裝置使用。

如圖 2 所示，系統架構的組成包含四個子系統功能模組：影像辨識、控制端、媒體播放器、使用者資料庫。



圖 2 系統組成之模組

3.1.1 影像辨識

影像辨識除了具備對人的感知能力外，也具備對環境明亮度的感知能力。在此模組採用 P.Viola 和 M.Jones 所提出的人臉偵測快速演算法[6]，這個演算法的運行包括三個步驟：第一為積分影像 (Integral Image)，使用一個中間值代表一個影像，可以很迅速地計算矩形特徵，在 Viola 的系統中，每個矩形特徵值的計算，最多只需要從積分影像中取 9 個元素做加減運算，這就是其快速的原因。第二為弱分類器與 Adaboost 疊代演算法，透過從大量的弱分類器中選取最具有分類意義的來組合成強分類器，[7]使用矩形特徵作為分類的依據，共使用四種矩形特徵。第三步驟則是層級分類器，一個影像會有很多個子框架，層級分類器則由許多分類器組合而成，每一層都是由上述演算法訓練所得到的一個強分類器，配合閾值，使其排除了眾多子框架中非人臉的樣本。

[8][9]提出幾套著名的人臉辨識的方法，但考慮到嵌入式裝置有限的運算能力，以及家庭成員的組成人數不高，此處尚不需要使用過於複雜的方法，而是採用統計學中常被拿來做常態分佈比對的 Bhattacharyya 距離，即可符合辨識之所需，接著再搭配使用者資料庫後，此影像辨識模組能夠判斷其是否為家庭成員，若為已註冊的家庭成員，則針對該成員提供個人偏好的影音服務。

3.1.2 媒體播放器

媒體播放器是建構在本機端內部多媒體播放引擎之上的應用軟體和使用者界面，因此，媒體播放器可以蒐集本機裝置的資訊與媒體播放內容，系統的其他模組元件也將可已透過媒體播放器來間接操作多媒體播放裝置。

3.1.3 控制端

控制端可以藉由接收並解析家庭網路中由不同網路多媒體裝置所群播(Multicast)送出的封包，從中發現網路上存在兩種類型的 UPnP 裝置：第一為多媒體伺服器裝置(Media Server)以及多媒體播放裝置(Media Renderer)，當這些裝置離開時，控制端也可透過該裝置再次發出的群播封包得知其離開，並更新目前存在於網路中的裝置清單。

控制端能夠藉由 UPnP AV 服務，瀏覽各個多媒體伺服器本身儲存且願意分享出來的多媒體檔案目錄。對於 UPnP 多媒體播放裝置，控制端可以操控其播放狀態如：開始、暫停、停止以及播放進度。對於正在播放的多媒體內容，控制端可取得其檔案的 URI (Uniform Resource Identifier)，得知該檔案為哪一個多媒體伺服器所擁有以及所屬的媒體格式，除外，控制器也可以取得各個多媒體播放裝置可以支援的多媒體檔案類型及格式，經過比對檔案來源與接收端兩者之間的檔案處理與裝置播放能力是否相符，如此將可確保多媒體伺服器與多媒體播放器之間的相容性。

3.1.4 使用者資料庫

使用者資料庫紀錄家庭網路環境中使用者的註冊資訊。使用者第一次使用本系統時，須先拍攝照片並輸入使用者名稱以建立個人基本資料。完成後，可選擇在區域網路中常用的多媒體網路裝置和服務，以便本系統日後辨識出使用者時，可以立即提供使用者個人化的服務內容。

3.2 系統運作流程

關於本系統的運作方式，經歷一次完整的使用流程將包含：使用者註冊、進入偵測模式、後端資料相似度比對及提供個人化服務等階段；系統內部的運作則由影像辨識、控制端與使用者資料庫三個模組元件的互動來完成，圖 3 顯示其整體的運作。

影像辨識在此系統啟動後即隨之運作，藉由偵測裝置的鏡頭所拍攝到的即時影像，立刻透過快速演算法[6]篩選出人臉，與使用者資料庫內的人臉特徵去做相似度比對，之後會抓取該成員之姓名並顯示於偵測畫面中，表示系統核對成功之結果。使用者註冊時，亦會使用到此子系統，如同在偵測模式下一樣，當使用者按下擷取 ID 照後，鏡頭及人臉

偵測相繼被啟動，當成功偵測出欲註冊成員之人臉特徵時，系統會切割出只有人臉特徵的資訊，此時該使用者可做確認來儲存其頭像資訊，以利與其他個人化設定做綁定。

控制端於本系統初始化時即開啟硬體的網路傳輸介面，接收區域網路封包，取得已加入的多媒體播放裝置。在使用者建立個人名稱和照片之後，控制端將區域網路中的裝置清單提供給使用者來選取慣用的裝置。之後，當影像辨識模組辨識出已註冊的使用者後，會將使用者名稱傳送至控制端，控制端向使用者資料庫以此名稱查詢相對應的裝置名稱，並且調用該裝置的影音服務，取得播放狀態和目前播放的內容。當控制端取得這些屬於特定使用者的資訊之後，將可以進一步控制對應的多媒體播放器，讓使用者可以決定是否接續播放方才在上一個裝置上尚未播放完的多媒體內容。

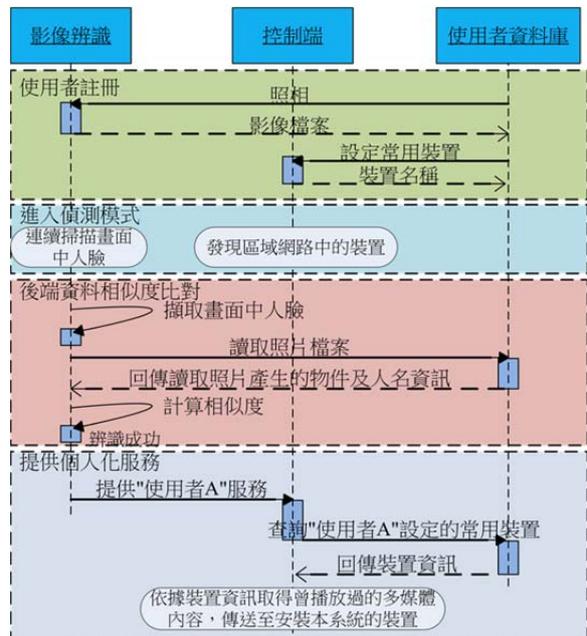


圖 3 元件互動

4. 系統雛型開發和展示

4.1 節說明系統開發平台與應用程式執行環境，4.2 節說明系統軟件的使用操作方式，4.3 節根據使用者情境的規劃，提出幾項實例以佐證之。

4.1 環境設備

此系統雛型的軟件研發採用 Java 及 C++ 程式語言，前者主要用於 UPnP 相關應用程式的開發，後者則是牽涉到 OpenCV 電腦視覺處理所需的軟體工作等。系統直接使用 Android 嵌入式行動終端裝置為，執行系統雛型所需的軟體程式和套件。以下兩小節敘述 UPnP 和影像辨識功能模組的實作方法。

4.1.1 UPnP on Android

此雛形的研發則利用 *Cling Java/Android UPnP* 程式函式庫及工具集，實作多項 UPnP 多媒體應用服務，例如：調用 AV Transport Service (AVT) 進行多媒體內容播放、瀏覽 Content Directory Service (CDS)的內容，以利於多媒體內容在區域網路中多媒體網路家電之間的傳送、轉置及播放。

4.1.2 OpenCV on Android

OpenCV (Open Source Computer Vision Library) 是一個跨平台的電腦視覺函式庫，原由一系列 C 函數和少量 C++ 函數構成，現今亦可用 Java 語言撰寫，並支援 Android 平台上軟體開發。OpenCV 可用於開發即時的圖像處理及計算機視覺等應用，它常被用於解決人機互動、物體辨別、圖像分割、人臉辨識、動作識別、運動跟蹤等圖像處理的問題。這部分的程式開發項目主要是在影像擷取、人臉偵測和辨識，以及圖像分割等，透過這些功能的串聯來實現系統感知使用者的出現和身份辨認等目標。

4.1.3 載具設備

圖 4 為系統雛型實作所使用的設備，標號 1 為一架網路電視，編號 2 是放置在網路電視上方，一架搭載 Android2.2 以上智慧型手機的影像偵測模組，編號 3 為家庭網路上各式的 UPnP 多媒體裝置，此處以筆記型電腦搭載 UPnP 軟體仿效之。



圖 4 系統雛型之環境設置

4.2 使用者操作流程

以下描述使用者在操作系統雛型的過程當中所執行的人機互動，如圖 5 所繪，這部分包含前端所產生的使用行為以及後端系統相對應的運作。

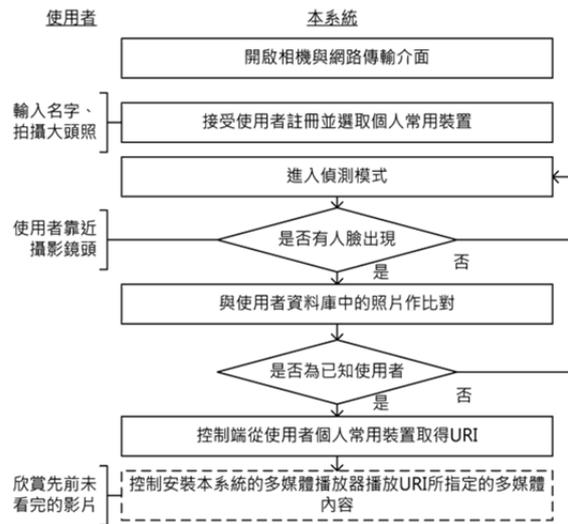


圖 5 操作流程

4.2.1 使用者註冊及常用裝置註冊

系統開啟影像攝影機，連續掃描在攝影機鏡頭前方的畫面，分析影像的特徵判斷是否有人臉出現，如果有則以一個綠色方框標示出被偵測到的人臉範圍。使用者僅需點選此方框即可將範圍內的人臉影像儲存為圖片於檔案系統中，並賦予此圖片檔案一個使用者自行定義的名稱。初步完成使用者資料的建立後，此系統尚須進一步分析該人臉資料在使用上的合適性，此時將會載入先前已經建立的人臉資料，比對人臉資料的差異性，如差異性大於系統預設的百分比值則會要求使用者更正人臉圖片。當成功建立人臉資料之後，可進一步附加上常用的多媒體裝置，系統將透過服務類型(UDA Service Type)過濾出具備多媒體播放能力的裝置，以列表方式顯示供使用者選取多媒體播放裝置。

4.2.2 人臉辨識驅動影音播放標的之轉換

在系統啟動時除了開啟攝影機鏡頭進行影像擷取之外，網路傳輸介面也會同時接受來自區域網路中的封包，以偵測其他裝置的存在，一旦辨識出已註冊使用者出現在攝影鏡頭前且從資料庫中取得該使用者名稱所對應的常用裝置，以常用裝置名稱比對目前存在區域網路與否，如果出現在區域網路中，則取得該裝置目前正在播放的多媒體檔案鏈結(URI)同時取得播放進度，設定系統指定的鄰近裝置上的多媒體播放器並準備播放的多媒體鏈結，完成後發送開始播放的指令。多媒體內容便透過人臉辨識的驅動，自動轉換播放標的從使用者常用的裝置到指定的鄰近裝置。

如圖 6 的實施例，以下說明人臉辨識和影音播放標的變換的操作細節。編號 1 為偵測模式的畫面，當系統啟動後，會直接來到此畫面，偵測畫面下方有三個選項，分別為 Members、UPnP 和 Switch：

Switch 選項是用來開關多媒體影音傳送功能，

為避免任一第三方裝置也同時偵測到該名成員時，相同的多媒體影音檔會重新被傳送，形成不斷重新播放的錯誤狀況。UPnP 選項讓使用者可以手動去選擇網路上多媒體播放裝置以及希望播放的影音媒體檔，這部分是依循 UPnP AV 模式，讀者可參閱 [1][2]。Members 選項是切換到使用者註冊介面，如編號 2 畫面，使用者可點選下方圖鈕來新增成員，選擇後會進入編號 3 的個人化設定，輸入使用者名字及建立人臉圖片，最後進行個人裝置的選擇，如圖所示，使用者可以看到目前家庭網路環境下存在有一個媒體播放裝置 BRAVIA KDL-xxx，最後在使用者完成以上的個人化設定後，點選 OK 按鈕會回到編號 1 的偵測模式畫面，此時系統的使用者資料庫已經紀錄了上述的個人化設定。



圖 7 家庭空間組成示意圖



圖 6 使用者實機操作之實施例

4.3 實例展示

本系統研發成果的展示是以居家生活空間為應用實施的場景，如圖 7 所示，基本上一個居家生活空間可由許多不同的小型空間所組成，一般的居家設計上會有客廳(編號 1)、臥房(編號 2)，也許會有書房和飯廳(編號 3 和 4)。

展示的規劃上，假定偵測裝置是安裝在客廳的電視上方，其他空間也許會有家庭成員常用或喜愛的多媒體網路裝置，例如個人電腦、筆記型電腦和智慧型行動電話等。特別解釋的一點是，客廳是家庭成員感情交流的主要場地，因此在系統離型的佈建及測試時，我們選擇把偵測裝置放在客廳裡。

4.3.1 情境 1: 單人

一位家庭成員原先在臥房內使用電腦正在觀賞影片，不過他想換個更舒適更大的螢幕來欣賞，於是他走出房間來到客廳，當他坐在沙發上面對著電視，本系統會辨識出他的身份，並發現他常用的裝置即臥房內的電腦上有一部正播放的影片，系統將影片的播放轉送到客廳的電視以供觀賞。

4.3.2 情境 2: 多人且有一位家庭成員

某日家中來了許多親戚在客廳聚會，因為這些訪客先前尚未有過註冊，人臉辨識模組的判定之後，並不會執行任何動作。此時，有位家庭成員願意和大家分享方才在房間內觀看的影片，他只需走到客廳的電視前，系統成功完成人臉辨識，啟動多媒體影音傳送的動作，實際測試結果如圖 8。



圖 8 多人之中有一位家庭成員

4.3.3 情境 3: 多人且多位家庭成員

圖 9 展示同一個空間中有多個家庭成員的測試結果，畫面中兩人都已進行註冊，當其中一方較靠近偵測裝置時，此系統就會認定其為主操控者，此時螢幕上就會顯示該成員的名字。

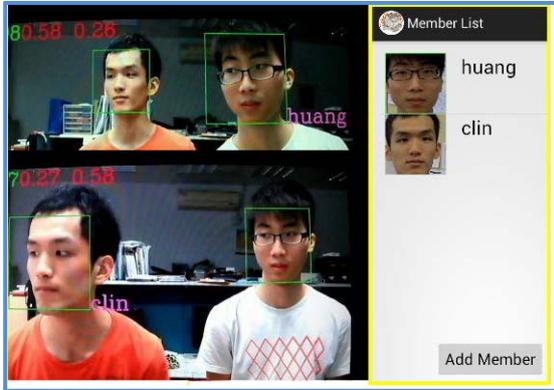


圖 9 主操作者選擇測試

4.4 系統封包分析

本系統實際運作時，利用 Wireshark 網路監測軟體取得裝置在區域網路中所發出網路封包資訊。圖 10 及 11 展示偵測裝置進入圖 3 的偵測模式時，以 4.2.2 章節所描述的運行方式，透過 HTTP 傳送訊息向使用者個人裝置取得多媒體內容資訊(圖 11a 處)，並轉送內容至指定的多媒體播放裝置，協助播放裝置向伺服器取得影片串流(圖 12b)。

表 1 各個裝置的 IP 位址

裝置	IP 位址
偵測裝置	192.168.1.116
使用者個人裝置	192.168.1.117
指定的多媒體播放裝置	192.168.1.106
多媒體伺服器	192.168.1.117

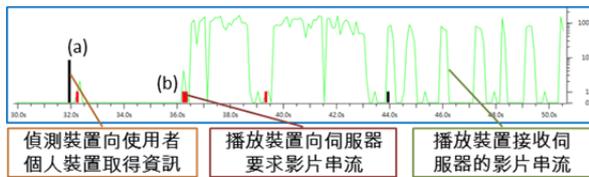


圖 10 系統運作過程之封包流量分析
(橫軸為時間(秒)，縱軸為封包數量)

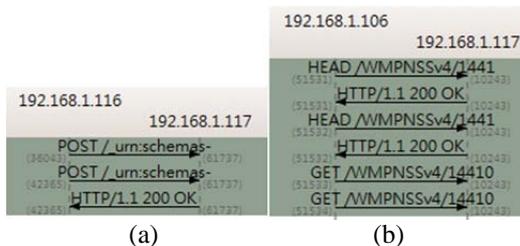


圖 11 (a)控制端與個人裝置間的 HTTP 訊息交換; (b) 指定的播放器向多媒體伺服器取得影片串流

5. 結論與展望

本論文提出在家庭網路環境下的智慧多媒體影音傳送系統之架構規劃、功能設計及雛型實作。本系統整合 UPnP 網路內容分享平台與人臉辨識兩大技術，搭配使用者資料庫，帶給使用者全新的人機互動方式，使得家庭影音分享不被居家空間的隔閡所限制。系統功能的特色在於一改傳統一個口令一個動作的手動操作模式，系統能夠自動根據家庭成員的身份，主動提動個人化的多媒體影音服務。至於系統雛型的實作，我們利用 Cling UPnP 與 OpenCV 等開放式軟體套件，完成一套可以建構在 Android 執行平台上的軟體模組並封裝為一組 Android 應用程式。最後，我們透過實機運行來真實呈現具即時性且個人化的智慧型家庭網路影音傳送服務。至於接下來的研發工作，我們將提升系統服務的個人化特色，讓使用者所屬的相關設定更為完備，並將加入客製化的多媒體內容服務，帶給使用者更多樣化的數位多媒體內容和應用。

致謝

本論文部分研究之進行，承台灣工業技術研究院之委研計畫 B5-10223-2H-1 經費支持。

參考文獻

- [1] UPnP Forum, "UPnP AV Architecture:2 Version: 1.0", Dec 2010.
- [2] C.-L. Hu, W.-S. Liao, and Y.-J. Huang, "Mobile Media Content Sharing in UPnP-Based Home Network Environment", *Information Science and Engineering Journal*, vol. 24, no. 6, pp.1753-1769, Jan 2008.
- [3] C.-L. Hu, C.-L. Lin, P.-C. Lin, and Y.-Y. Su, "Gesture-Assisted Remote Control for Media Content Sharing and Distribution in Home Networks", *Proc. of IEEE ISCE*, pp.87-88, Jun 2013.
- [4] C.-L. Hu, B.-J. Hsieh, H.-C. Lin, and Y.-F. Hsu, "Integrated home gateway design for enabling P2P content sharing in UPnP-Based home networks", *Proc. of IEEE ICCE*, pp. 229-230, Jan 2011.
- [5] H. K. Ekenel, H. Gao, J. Stallkamp, M. Fischer, and R. Stiefelhagen, "Face recognition for smart interactions", *Proc. of IEEE ICME*, pp. 1007-1010, July 2007.
- [6] Petros Belimpasakis, and Rod Walsh, "A combined mixed reality and networked home approach to improving user interaction with consumer electronics", *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, pp. 139-144, February 2011.
- [7] P. Viola and M. Jones, "Robust Real-Time Face Detection," *International Journal on Computer Vision*, vol. 57, no. 2, pp. 137-154, May 2004.
- [8] P.N Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Transactions, Pattern Analysis and Machine Intelligence*, 1997.
- [9] M.A. Turk and A.P. Pentland, "Face recognition using eigenfaces," *Proc. of IEEE CVPR*, Jun 1991.