

教育部—「學海築夢」102 年度獎助大專
校院選送學生赴國外專業實習計畫

多媒體視覺辨識系統技術實習計畫心得報告

國立台北科技大學電子工程系研究所—焦名楷

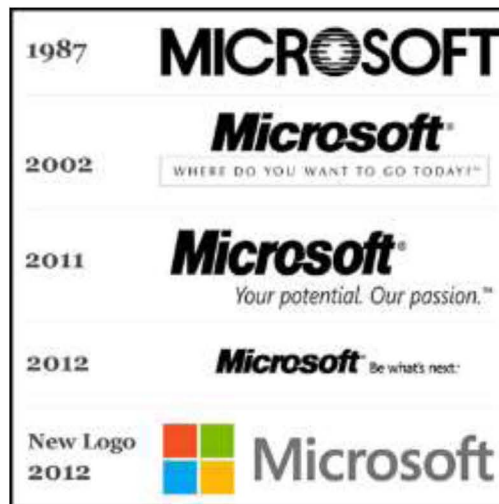
一、緣起

臺北科技大學是一所在教育方針上強調理論與實作並重並結合的學校，從學校課程開課即可發現，包括了理論、實務技術、方法工具應用等，尤其又以實作為本校學生的重要課程目標，透過積極邀請業界來學校給予講課，並結合技術性的實務練習來加強學生技術能力，如此教學特色將使得臺北科技大學的學生畢業後，能輕鬆因應業界工作的各種挑戰與世界趨勢潮流，以向業界展現臺北科技大學學生的智慧與才能兼備之能力。本次透過教育部的「學海築夢」計畫－獎助大專校院選送學生赴國外專業實習計畫，讓本校學生有機會將視野拓展至海外，以最直接的方式接觸到國際的公司與大企業，我們也很有興地聯絡到華盛頓大學的黃正能教授，由於黃教授的安排與協助下，讓我們有機會到美國華盛頓州的西雅圖，前去微軟公司於西雅圖之總部進行相關最新的多媒體的研發和技術之實習，以及多家知名公司參觀，以提升在技術上與市場上的國際視野。



二、 國外機構簡介

微軟股份有限公司 (Microsoft Corporation) 為美國一家跨國電腦科技公司，從 1975 年創辦至今，全球各地遍布了有超過 112 個國家有設立 Microsoft 分公司，總員工人數更高達 9 萬人以上，在台灣同樣也設有分部—台灣微軟，位於信義區附近主要以推廣為主，而微軟於西雅圖的總部含蓋所有的層面，從市場推廣分析、未來技術之研究到對外產品的研發皆有。研發的產品涵蓋越來越廣泛，從初期主打軟體的電腦作業系統到現在購併了 Nokia 公司而延伸出硬體產品之相關業務等，可說是造就現今的科技之重要推手。



Microsoft 微軟 1987 年至今註冊商標之演進

此公司是由兩位創辦人—比爾蓋茲 (Bill Gates)、保羅·艾倫 (Paul Allen) 於 1975 年創立，目前執行長一職由 Satya Nadella 擔任，比爾蓋茲則退居幕後從事各種慈善事業。



Microsoft 創辦人—比爾·蓋茲、保羅·艾倫

創辦初期以販售用與 Altair 8800 的 BASIC interpreters 起家作為開創的基礎，而演變至現今家家戶戶無可取代的電腦作業系統—Windows，同時也用於許多公司行號與學校單位，又以 Windows XP 為代表作在市場銷售與市占率曾創下驚人的數字，到現今仍有許多軟體只支援該系統版本。



Microsoft 微軟作業系統商標演進

遊戲、娛樂產業也是微軟相當重視的一塊市場，微軟相關與此的商標有 Xbox、Xbox 360、Surface、Zune 等，在微軟娛樂設備研發中，又以 Kinect 為近年來最大的創舉商業產品，Kinect 的發展主要是用於微軟的 Xbox 360 和 Xbox One 遊戲主機的周邊設備，它讓玩家有著不同操作體驗，完全不須使用及手持任何搖桿控制器，而是透過手勢動作與口說語音來操作遊戲與系統介面，這也是本次實習的重點之一。



Microsoft 微軟所研發的娛樂相關產品

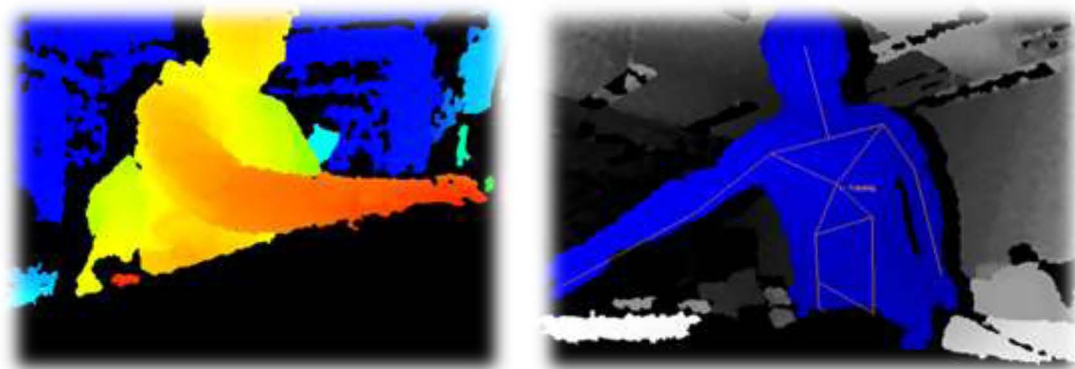
三、 國外實習企業或機構之學習心得

本次實習計畫之內容相當的充實與豐富，主要由我們的指導教授黃士嘉教授與華盛頓大學 (University of Washington, UW) 的教授 Prof. Jenq-Neng Hwang 負責幫我們進行實習安排與拜訪規劃，讓我們來美國西雅圖於此計畫的執行期間，於微軟 (Microsoft) 實作並接觸到許多視覺及影像應用之相關技術，但不僅僅只是對與微軟技術上的認識與學習，此外

也參訪了多家在西雅圖設立總部的知名科技相關大公司與機構，VoiceBox 其研發與販售許多智慧型的語音助理系統、T-Mobile 其為美國前 3 大的電信公司，以及華盛頓大學的校內，例如：圖書館，教室，實驗室，學校餐廳等，所有的一切都讓我們大開眼界。

■ 微軟公司技術實習與研究討論

本計畫主要的實習技術，主要牽涉到 Kinect 的實作與應用以及相關視覺與影像處理之演算法的思考與設計，透過應用許多強大的硬體資源，可以造就出許多非常有創意的應用，Kinect 在硬體的使用上最具重大突破的是使用深度鏡頭 (Depth)，透過打出紅外線光以及接收並偵測返回的量，進而判斷鏡頭前物體之遠近與距離，通常可透過以圖形化的介面來具體化其深度結果，以不同顏色強度來代表其對應的遠近深度，如下圖之左邊，此外 Kinect 仍具備了基本的彩色鏡頭，以利基本彩色視覺影像的採集，透過此兩種鏡頭：(1) 深度與 (2) 彩色的交互輔助，將使得物件上的辨識與偵測更加精準與容易，基於這些影像資訊，將可以做到骨架 (Skeleton) 的偵測，如下圖之右邊。



在過去 Kinect 的開發其實並不容易，必須透過第三方(非微軟釋出)的軟體開發套件 (Software Development Kit) 進行操作與研發，而現今只需要到微軟為 Kinect 網站：<http://www.microsoft.com/en-us/kinectforwindows/default.aspx>，即可去下載 Kinect for Windows SDK 來開發使用 Kinect，此官方 SDK 不只允許開發者可以使用微軟自家的語言 (C# 和 Visual Basic)，也可以使用 C++ 之程式語言，其中還內建了許多函示庫資源，像是基本的手勢辨識的支持以及語音辨識。微軟在這部分做的相當得好，在 SDK 裡面附有完整

的範例程式碼專案，其專案是用 Microsoft Visual Studio 的 Windows Presentation Foundation 為基礎架構的應用程式，這些範例程式讓我們實習的初期可以快速的看到其程式執行結果並可以快速上手進行 Kinect 的應用。整體 Kinect 的設計要點可以簡單敘述如下：

- ✓ 獲取被接入的 Kinect 實體
- ✓ 開啟彩色影像串流，從彩色鏡頭
- ✓ 開啟深度影像串流，從深度鏡頭
- ✓ 實作 VideoFrameReady 事件，處理一般影像
- ✓ 實作 DepthFrameReady 事件，處理深度影像

在整體實習上，必須先具備影像處理的知識與概念，像是影像之色彩空間 (RGB, YUV)、座標點關係、串流影像以 Frame 方式進入等，才知道其運作原理以及如何進行進階開發與應用。



技術實習之 Kinect 開發

實習期間，不只有進行實作演練，每個星期也會定期參與會議討論，特別是會固定在華盛頓大學進行一周內所作之事情，作一些進度與未來的目標

方向報告，每次的會議中都可以學到很多不同技術與研究內容，像是各類影像影片上的物件偵測與辨識技術，這是由黃正能教授在華盛頓大學帶領的實驗室之重要的一項研究主題，在會議中我見識到了許多這方面的多樣應用之可能性，例如：交通路況監控、跨越多台監控錄影機之相同人物追蹤、生產包裝工廠之捕魚產線上魚種類之辨識等，每個應用都有其背後相當可觀的經濟價值與技術深度。此外我也被邀請於其中一次會議中進行個人的研究主題報告，我的報告題目為我於台北科技大學博士期間之研究主題，此為我投稿 IEEE Transactions on Intelligent Transportation System 期刊並被接受的論文，主題名為 - A Genetic Algorithm-based Approach to Solve Carpool Service Problem in Cloud Computing - ，報告過程中與黃正能教授進行了許多的技術內容討論，因此獲得了許多非常寶貴的建議，同時也啟發了我更多的未來研究方向。



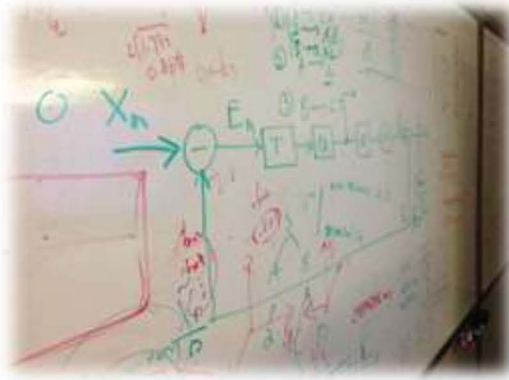
微軟的公司領地與園區相當的大，其中 Research 的區域之 Building 99，如下圖，為我們在門口的大家合照，裡面的人員進行各式各樣的研究，這些研究都是未來可能結合或成為微軟的銷售產品。



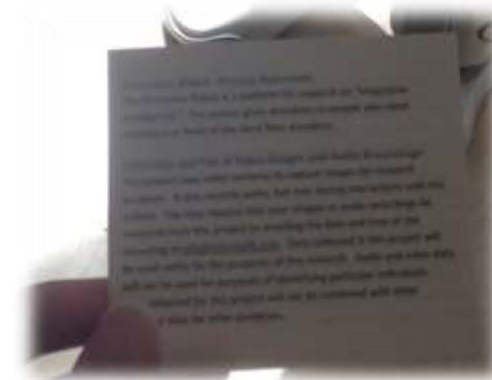
微軟的實習與進入必須有內部人員帶領才可進入，是由 Dr. Liu Zi-Cheng 帶領我們進入，在進入微軟的建築之前需在櫃檯進行登入與製作個人名牌的程序，如下圖，為我的個人名牌。



微軟的工作環境帶給人有一種不一樣的感受，它讓工作的人感覺很自由可以很隨意地作任何事情，如下圖，你可以隨意在旁邊的白板牆壁作筆記或計算等，在公司內部還會設有飲料區供大家使用，如下圖，此外隨處也設有許多討論沙發與開會室，讓員工隨時隨地有需求時就可以使用，以促進整體公司成員的工作效力。



有趣的是有時在路上走著走著，會遇到微軟研究人員開發的系統的展示，如下圖，此為一台結合日本的小機器人與 Kinect 影像偵測的應用，只要有人走到機器人的面前，它就會舉起手和你大招呼並詢問你需要什麼幫助，可以對他說「我想找某某人」、「廁所在哪裡？」等，它都會回答你以協助你解決問題。



另外我們在實習中還到了一小測試房間，裡面架設了許多攝影鏡頭，此房間是在進行 3D 視訊的開發，讓你在視訊時有著更深入其境的體驗，我們也在房間嘗試了這個開發階段的产品，如下圖，左邊為我們的同行學弟在體驗此 3D 視訊系統的情況，以及右邊的研發階段的 3D 視訊眼鏡。



最後我們抵達一間會議室，由微軟的研究人員介紹了一些他們最近的研究主題以及所使用的技術，深度學習 (Deep Learning, DL) 為他們進來喜愛使用的一種演算法技術，其為一種類神經網路的變形架構，透過將神經網路之內部隱藏層以許多層的方式設計，使得此神經網路能有深度的學習來達到更好的執行效果，此外也克服了傳統倒傳遞類神經網路的學習之缺陷，傳統學習方式是亂數設定類神經網路的初始值，再透過迭代的方式來訓練網路，而 DL 因為有多層隱含所以訓練時，是每次以一層來進行訓練稱之為 Layer-wise 的訓練，以避免傳統神經網路造成在訓練時，將誤差的結果傳遞到下一層，這將造成 gradient diffusion 的問題。



四、 國外實習之生活體驗

■ VoiceBox 公司參訪

VoiceBox ，如下圖，於前台的大家合照，參訪是由 Ms. Mei Jul Lin (Vice President) 帶領我們先參觀整間公司的工作環境。



與 Ms. Mei Jul Lin (Vice President)的大家合照

隨後帶我們到公司內的會議室，並開始由 Michael Kennewicka (CEO 兼創辦人) 作一些他們公司的技術與理念介紹，如下圖，介紹過程中包含一些尚未對外公開商業產品介紹。



Kennewicka (CEO) 的介紹以及展示第一代 VoiceBox 產品

他們主要產品是關於語音辨識 (Speech Recognition) 的應用與結合，像是結合個人助理的功能置你周邊的電子裝置中，讓使用者只需要透過直覺地說話與機器對談，就可以進行許多的操作，像是建議餐廳的查詢、今日天氣狀況的查詢、即時交通路況的查詢等，顯然地他們主要技術部分是在實作使用者與機器之間能以語音對話的方式操作，至於背後的有關各種資訊之龐大資料庫將是取至其他家公司，所以於介紹中他們也提到今年對於中國市場部分，已與多家大陸廠商簽約進而存取到多樣的資料庫，來供 VoiceBox 的使用者搜尋使用，此外也將這部分服務產品實作至 Android 系統手機上，並現場展示給我們看，類似於 Apple 的 Siri 功能，直接對手機說話，機器即可快速的回應相應的結果給他，這部分 VoiceBox 比 Apple 厲害的是使用者所說的話會被記錄下來，作為搜尋參考之用，而 Apple 是句子獨立的方式進行辨識，因此對 VoiceBox 的產品使用以下的對話方式也沒任何問題：

User 說：

今天台北天氣如何？

VoiceBox 說：

今天台北天氣溫和，最高 24 度，最低 20 度，適合出遊

User 說：

那高雄呢？（註：無須再提及“天氣”關鍵字）

VoiceBox 說：

高雄與台北天氣類似，最高 26 度，最低 22，適合出遊

這家公司讓我印相非常深刻，因為在會議期間，Ms. Mei Jul Lin (Vice President)、Michael Kennewicka (創辦人兼執行長)、Bob Kennewick (CTO) 等皆出席與我們會面。

■ T-Mobile 公司參訪

T-Mobile，如下圖，為我進入此公司的參訪名牌，這間公司可能對於台灣人很陌生，但對於我們來到美國後之後，就會發現它無所不在，

T-Mobile 是美國前三大的電信公司，主要提供服務就是各類電信服務，像是我們抵達美國時就立即去辦了 T-Mobile 的行動電話號碼，T-Mobile 的特色就是價格便宜，因此廣受許多美國人愛用。



<http://www.t-mobile.com>



在這邊的參訪從外面擺設來看，就感覺得到 T-Mobile 對顏色的執著與貫徹，從公司周圍的路燈裝飾到公司內部的牆上擺設，都是以粉紅色系進行設計。在進入公司參訪前，需先於大廳進行一些登記動作，櫃台將會給予一張名片貼紙，如上圖為有我的名字的貼紙，進入 T-Mobile 同樣地快速的瀏覽了他們的工作環境與辦公司後，就到會議室與他們的工程師以及一些公司人一起討論給予我們一些對 T-Mobile 的基本介紹與未來規劃的演講。下圖為我們在 T-Mobile 總部大廳與接待我們的 Kevin 的合影。





T-Mobile – Validation Lab

■ **華盛頓大學 (University of Washington)**

西雅圖華盛頓大學 (The University of Washington at Seattle) 建校於 1861 年，為美國西岸州立大學系統中最早設立的高等教育學校。建校時為私立學校，從 1889 年被收歸華盛頓州所有。華盛頓大學位於華盛頓湖 (Washington Lake) 與聯合灣 (Union Bay) 之間，校園佔地超過 700 英畝。交通非常便利，從學校到西雅圖市區僅需約 15 分鐘的車程。華盛頓大學主要由三個分校組成。主校區位於西雅圖市區東北角提供博士學位的學程以外，華盛頓大學還有 Bothell 與 Tacoma 兩所分校提供學生到碩士學程的學位。西雅圖主校區總共包括 17 個學院，為一年級新生到博士在讀生提供課程與研究機會。



華盛頓大學的校園美景



設於華盛頓大學校園內－微軟贊助華盛頓大學的紀念

五、 國外實習之具體效益

1. 增強外語能力

我們的指導教授在台灣時，會每周安排美國的外籍老師與我們進行一對一的英文指導，所以建立我們每位在英文上的信心，不會遇到外國人而

不敢開口說英文。長久以來的英文準備，讓對於能來美國這樣全是英文的環境感到相當興奮，來到美國一切的事物都轉變為英文，不論是溝通上或視覺上，英文聽力與對話的能力也因此有所提升，因為要辦理各項事務能使用的語言就只有英文，例如：點餐、購物、買票、問路…等，此外更讓我印象深刻的是超級市場的食物，我到超級市場最常做的一件事情就是拿出手機把所有英文牌子拍下來或是將單字打到手機裡面紀錄，方便之後查字典學習，因為如此，我學到了許多食物的英文單字，包含醬料、各種水果、肉的部位等，在這之前都只到中文如何說而已，經過了兩個月的努力與洗禮，到超級市場買食物時都可以直接確認這是否自己想要的東西，顯然地英文能力有非常具體的提升。

2. 提升國際觀

在美國微軟的實習與多家企業公司參訪中，讓我國際觀有著重大的改變，一直認為國際知名公司能這麼成功，作出這麼多有影響力的事情及產品，一定是把底下員工天天加班熬夜所帶來的結果，猶如台灣新聞上常常看到的事情－工程師熬夜過度過勞等，但實際上完全不同於台灣的工作環境，像是 Microsoft 微軟的工作環境，其目標就是希望給所有員工最方便、最舒適、最規律的生活，時間到了大家就會準時下班，工作累了可以去周邊的冰箱拿取免費的飲料，找個舒服咖啡桌休息，如此反而更加速每個員工的時間利用度，達到更好的產出結果，這些結果都很直接底反映在國際市場上。另外一件值得注意的事情是可以發現美國的科技公司內，是由來自世界各地的人種所構成，最常看見的種族不是亞裔人士就是印度裔人士，世界各地在各國的精英匯聚於此，也是美國企業能成功的關鍵因素之一。

3. 做事態度與方法的學習

在此次實習計畫期間，我們每周都要參與英文會議討論，聽取當地其他人的研究進度，此外也會與我們的指導教授每天進行群體的會議討論，透過會議來討論所面臨的問題，各自的進度報告，讓大家都能以積極與謹慎的態度準備每一次會議內容，從中也學習到一場好的開會會議，是不能花太長的時間進行，而是要盡可能在最短的時間內表達自己的所有想法與要討論的事項，避免無意義的討論，已使得會議的目的能達到並且僅需花

費少許大家的寶貴時間。

六、感想與建議

這次的學海築夢計畫是一次非常難得的體驗，許多許學生來說這可能都是他們人生第一次的出國體驗，在這非常感謝教育部能提供如此一計畫，讓我們學生能有更多的出國體驗機會。整個計畫中，非常感謝我們的指導教授—Prof. Shih-Chia Huang 努力為我們爭取而來出國的機會，以及積極地幫我們安排在美國各方面的生活事項，另外也非常感謝華盛頓大學的 Prof. Jenq-Neng Hwang 與我們進行各種技術討論，並安排了完整的微軟多媒體技術的實習，以及許多美國在西雅圖公司的參訪機會，提升我們的國際觀與世界觀，並以不同角度來看到公司運作方式。這兩個月在美國生活，讓我們大家體悟到英文其實沒那麼可怕，出國留學其實你也辦得到，只要你願意放手一搏其實機會都在你的身邊，如此也讓我對未來的方向有更多元的想法和方向，從原本的台灣國內放遠到國際。在美國最常遇到的亞裔人種為港澳地區、韓國人以及大陸地區的人，台灣人相對比較少數，希望教育部未來能一直提倡類似的相關計畫，讓更多台灣學生接觸到國外世界，增加國際視野創造出更多出國創業人才，打開台灣人在國際的知名度。