

出國報告（出國類別：實習）

多媒體視覺辨識系統技術實習
計畫成果報告

服務機關：國立台北科技大學電子工程學系

姓名職稱：黃士嘉

派赴國家：美國

出國期間：民國 103 年 07 月 16 日 至 民國 103 年 09 月 17 日

報告日期：民國 103 年 09 月 24 日

目錄

一、	緣起-----	3
二、	國外實習機構簡介-----	5
三、	國外實習機構聯繫-----	7
四、	國外專業實習活動-----	9
五、	國外企業機構參訪與交流-----	13
六、	計畫執行感想與建議-----	15

一、緣起

本多媒體視覺辨識系統技術實習計畫(以下簡稱為海外專業實習計畫)之目的係為培育具國際觀之最新互動多媒體實作技術人才。透過選送 13 位學生親身體驗到美國 Microsoft 公司-西雅圖總部之互動多媒體部門的實習過程,使選送生可學習最新發展的 Kinect 肢體感測裝置之視覺辨識系統。

近年來,Microsoft 公司發表一套全新的 Kinect 肢體感測裝置, Kinect 主要利用 RGB 彩色攝影機、紅外線發射裝置、紅外線 CMOS 攝影機和陣列式麥克風來得到彩色影像、3 維深度影像以及立體聲音訊號三種資訊來達到精準的肢體感測與辨識。其中 3 維深度影像對於要完成高精確度的影像辨識系統是非常重要的關鍵,透過利用深度資訊可以輕易地掌握物體的遠近及物體的形狀等等,並對於後續的物體辨識上,可以有效的提高其精確度。因此, Kinect 肢體感測裝置目前已經被廣泛的使用在視訊監測、醫療復健、家庭娛樂和互動生活等等。同時,因 Kinect 具備極佳的偵測與辨識能力,使得 Kinect 肢體感測裝置在視覺辨識系統中扮演了相當重要的角色。

因此,本次海外專業實習計畫係以實作方式讓選送生進行為期兩個月在美國 Microsoft 公司-西雅圖總部之互動多媒體部門的實習訓練,並透過 Kinect 肢體感測裝置與軟體平台之開發工具,使每位選送生皆可清楚掌握開發視覺辨識系統之實務能力,同時,藉由親自到知名的國際企業

Microsoft 實習，並透過該公司之主管親自帶領和交流下，每位選送生皆體會與學習到知名國際企業的工作態度以及處事方法，進而有效的培養他們的國際價值觀，並且也將此技術引入台灣扎根，為我國培養優秀之視覺辨識系統研發人才。



圖(一) 遴選生與國外生活導師於美國微軟 Microsoft 之合照

二、 國外實習機構簡介

本次海外專業實習計畫選送 13 位電子資訊專業領域之學生至美國 Microsoft 公司-西雅圖總部之互動多媒體部門執行專業實習訓練。

美國 Microsoft 微軟股份有限公司是目前全球最大的一家跨國電腦科技公司。公司於 1975 年由比爾蓋茲(Bill Gates)和保羅艾倫(Paul Allen)創立，兩位創辦人相信個人電腦將會成為辦公室或家庭中重要的工具，因此創立 Microsoft 微軟公司並研發個人電腦的軟體程式。



圖(二) 微軟商標歷史

創辦至今，在全球已超過 112 個國家設有 Microsoft 分公司，全球員工總人數高達 9 萬人左右。Microsoft 微軟公司主要以研發、製造、授權和提供廣泛的電腦軟體服務業務為主。目前 Microsoft 微軟總部位於美國華盛頓州的雷德蒙德，最為著名和暢銷的產品為 Microsoft Windows 作業系統和 Microsoft Office 軟體更為世界著名且暢銷的產品。

微軟的成功和軟體業界的興起都應歸功於個人電腦的發展，而蓋茲先生早已預見這樣的事實，同時在比爾蓋茲的領導下，微軟公司致力於不斷研發、改進軟體科技，希望能夠讓它變得更簡單好用、更平易近人、更有

趣且更具經濟效益。而 Microsoft 微軟公司除了專於提供 Windows 電腦作業平台外，也積極開發伺服器平台和信息工具，並深入了解客戶來解決客戶的商業方案，進而幫助企業提高其整體競爭力。在消費電子市場上，也陸續研發有關多媒體裝置及家庭娛樂等產品，像是 Windows Phone、Microsoft Surface、Xbox 360、Kinect、Skype 以及 Lync 等暢銷產品。Microsoft 微軟公司以其功能強大、獨特創新及易於操作的產品與技術，不斷改善人們的生活、工作和交流方式，讓人們在任何時間、任何地點、使用任何裝置，都能輕鬆取得資訊，以技術變革的領導者角色帶領全球走向人為導向的生活目標。



圖(三) 微軟相關產品

三、 國外實習機構聯繫

本次海外專業實習計畫之美國 Microsoft 公司實習負責人為 Yi-Min Wang 博士，Yi-Min Wang 博士為微軟雷德蒙德研究院的主任，透過 Yi-Min Wang 博士的聯絡下與美國 Microsoft 西雅圖總部之互動多媒體部門建立起本次海外專業實習之機會。並於實習過程中，透過研發部門主管 Ming-Chien Lee 博士的帶領下，讓 13 位選送生皆順利的習得開發基於 Kinect 肢體感測裝置之高精確度的影像辨識系統的關鍵技術。以下說明為本次選送生實習過程之相關協助及輔導介紹。

● 提供實習學生之待遇

1. 實習學生住所
2. 企業實習導師-Ming-Chien Lee 博士
3. 實習學術導師-Jenq-Neng Hwang 教授
4. 國外生活導師：黃士嘉教授與 Jenq-Neng Hwang 教授

● 實習環境

1. 提供辦公桌與網路設備。
2. 提供相關實習程式開發工具。
3. 企業導師針對實習期間傳授專業技術以及成績考核事項。
4. 國外生活導師協助處理實習期間之生活問題。

● 實習學分採計方式

依據『國立台北科技大學研究所校外實務研究課程開設準則』，海外實習時數需滿 320 小時，待選送生實習結束回國後，由計畫主持人黃士嘉副教授與企業實習導師-Ming-Chien Lee 博士依照實習成效評定成績。修習『校外實務研究』之課程，其相關規定如下：

1. 課程名稱：校外實務研究
2. 學分數：3 學分
3. 校外實習類型：海外實習
4. 學生於海外從事實習課程，需在同一機構實務研究，並累積不得低於 320 小時為原則
5. 實習成績為輔導老師佔 50%、實習研究單位佔 50%之比例為原則。
6. 研究所『校外實務研究』課程之實施，採用『國立台北科技大學研究所校外實務研究課程開設準則』之相關規定。
7. 研究生至海外從事實習者，仍視為『全時間修讀』研究生。

四、 國外專業實習活動

本次海外專業實習計畫中，13 位遴選生於美國 Microsoft 公司-美國西雅圖總部之互動多媒體部門的專業實習活動分為五個階段，於每個階段皆有國外生活導師黃士嘉教授和企業實習導師 Ming-Chien Lee 博士的陪同與指導，其五個實習階段分為如下說明：

- 實習規範與環境介紹：於實習第一週，我們安排互動多媒體部門說明實習規範與環境介紹，讓遴選生熟悉實習環境。同時，在每週晚間時段，安排國外會話導師，分別一對一教導學生實用科技英文應用與口語對話之訓練，並制定實習目標，讓遴選生秉持嚴謹的心態面對實習單位所賦予的任務，並自我期許完成預期成效。
- 第一階段實習：讓遴選生實習『3 維深度計算』實作，透過紅外線發射與接收裝置進行連續近紅外線的發射，來測量空間深度進行編碼計算後，對於編碼後資訊進行影像深度演算法之設計，並實作於 Kinect 肢體感測裝置上，培養 3 維深度計算的實務能力。
- 第二階段實習：讓遴選生實習『人體偵測』實作，基於背景與移動物體兩者間不同特性與動靜態的關係進行判斷分析，並根據特性關係設計其演算法，並實作於 Kinect 肢體感測裝置上，進而驗證相關特性，藉此培養人體偵測的實務能力。
- 第三階段實習：讓遴選生實習『人體關節辨識與骨架追蹤』實作，透

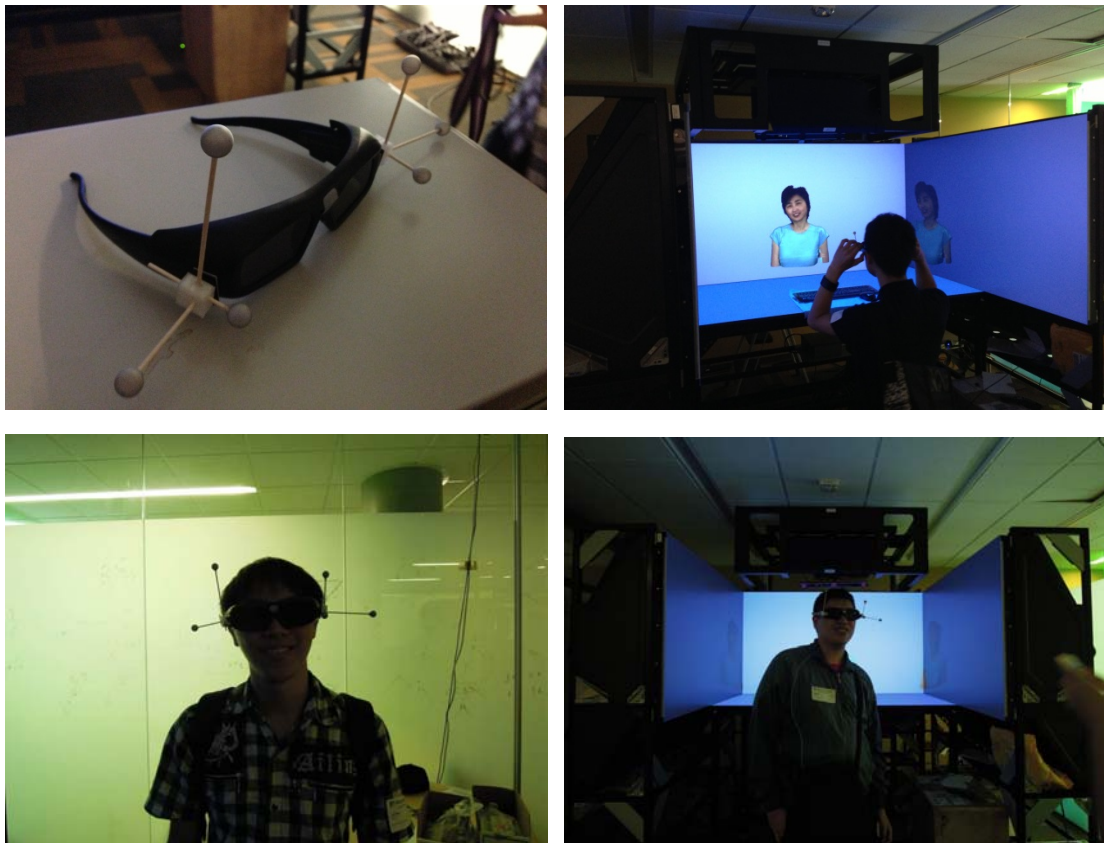
過人體特有的關節特徵與四肢特性進行判斷分析後，根據關節位置分別進行人體骨架的追蹤並進行演算法的設計，並實作於 Kinect 肢體感測裝置上，藉此培養人體關節辨識與骨架追蹤的程式設計能力。

- 第四階段實習：讓遴選生實習『角度變化計算與偵測』實作，透過人體四肢關節以及人臉轉向的位移角度特性進行演算法的設計，並實作於 Kinect 肢體感測裝置上，進而驗證相關特性，藉此培養人角度變化計算與偵測的程式設計能力。
- 第五階段實習：讓遴選生實習『手掌特徵擷取與手勢辨識』實作，透過進行人體關節辨識與骨架追蹤的實作後，根據手掌的輪廓與手指關節的特性利用程式方式進行擷取，並根據不同手勢的特徵進行演算法的撰寫，並實作於 Kinect 肢體感測裝置上，藉此培養手掌特徵擷取與手勢辨識的程式設計能力。
- 實習成效評定與測驗：結合上述實習：『3 維深度計算實作』、『人體偵測』、『人體關節辨識與骨架追蹤』、『角度變化計算與偵測』和『手掌特徵擷取與手勢辨識實作』五項實作做綜合性的應用並以專題方式驗收成果，企業實習導師 Ming-Chien Lee 博士以此成果作為本次海外實習成績的評定標準。

- 其他相關實習過程介紹

基於 Kinect 之 3D 顯像通訊應用

此應用係利用 Kinect 肢體感測裝置移植到電視的 3D 顯像通訊之應用。此應用尚需配合特殊的 3D 立體眼鏡來達到真實的 3D 效果，讓使用者可以於通話的過程中，可以不同角度觀看 3D 立體成像的各個角度位置，感受到真實的臨場感。下圖中，為我們參與實際測試之過程。



圖(四) 基於 Kinect 之 3D 顯像通訊技術

基於 Deep Learning 之自然語音處理與應用

由於 Kinect 肢體感測裝置裝設有陣列式麥克風，因此，此應用系利用此 Kinect 裝置與 Deep Learning 之演算法來達到準確地將文字歸類及預測每個文字間的關聯。本次的應用先透過研究員 Xiao-dong He 向遴選生們說明「Deep Learning for Natural Language Processing and Related Applications」的演算法。預期透過 Deep Learning 之演算法，打破人與人之間的語言隔閡，讓各國人民之間的溝通皆可以用自己最熟悉的母語來確切的表達出語意。過程中，係透過 Deep Learning 之演算法來即時顯示以及翻譯出使用者所說的每一句話。



圖(五) Deep Learning 演算法教學之工作會議



圖(六) 遴選生執行實習過程及相關工作會議

五、 國外企業機構參訪與交流

透過本次海外專業實習的機會，我們也特別安排遴選生們透過休假期間到美國其他大型科技公司參訪與交流，藉此引導遴選生快速適應國外的實習環境與當地生活，同時也可以建立與該公司的溝通管道，便於日後台灣各校之優秀學生海外實習之機會與聯繫。以下為各大參訪科技公司介紹。

● VoiceBox Technologies

VoiceBox 公司係為自然語言辨識和對話語音科技的先鋒。該公司之主要產品為自然語音辨識系統，其系統可裝載於各大車載系統、智慧型手機、智慧型電視和其他家用系統上運行。該公司曾獲得 32 項專利於自然語音應用技術。且該公司所擁有的專利曾被知名專利價值評估網站(Patent Power Scorecard)評估為全球最具創新力和影響力的智慧財產權。該公司之產品已被豐田、先鋒、克萊斯勒、道奇、飛雅特、TomTom、雷諾和馬自達等公司裝載於各式產品上。



圖(七) VoiceBox Technologies 企業參訪

● T-Mobile

T-Mobile 是一家跨國的電信公司，「T」在公司名字裡面代表「電信」所有德國電信的子公司都用「T」開頭，屬於 Freemove 聯盟。T-Mobile 在西歐和美國營運 GSM 網路並通過金融手段參與東歐和東南亞的網路營運。該公司擁有 1.09 億用戶，是世界上知名的行動電話公司之一。該公司之總部就設在華盛頓 Bellevue。T-Mobile USA 是美國市場上第三大營運商，也是繼 Verizon Wireless 之後的第二大成長最快的公司。T-mobile 也是唯一一家在歐洲和美國使用統一品牌的行動電話公司。通過國際漫遊協定連接到德國電信並與其他 GSM 網路兼容，T-Mobile USA 比任何一家美國無線營運商提供更多的全球範圍內更多的覆蓋面積。2004 年，T-Mobile USA 被授予了 J.D.Power 和 Associates 的幾個獎項，包括「2004 無線用戶服務最高評價獎」和「2004 無線零售服務海外用戶最滿意獎」。



圖(八) T-Mobile 企業參訪

六、 計畫執行感想與建議

本計畫之目的係提供 13 位遴選生於美國 Microsoft 公司執行專業實習活動。首先，實習過程主要實作一套基於 Kinect 之高精確性的視覺辨識系統，依照本次的實習內容，我們可以將實作方式分為五個主要的實習階段，其包含 3 維深度計算實作、人體偵測實作、人體關節辨識與骨架追蹤實作、角度變化計算與偵測實作、手掌特徵擷取與手勢辨識實作，透過上述五個階段的專業實習訓練，成功的讓每位遴選生習得最新且尚未公開之視覺辨識系統開發之技術。

同時，每位遴選生透過實習交流機會，讓他們可以實際的使用英文溝通，體驗該如何將想要表達的意思確切的以英文表達出來，並讓聽者可以確實了解與掌握。明顯地，在實習交流和工作會議之經驗，每位遴選生皆有大幅度的成長，英語溝通能力明顯增進。

每位遴選生經過專業實習經驗，皆親自體驗到知名國際企業的運作模式與工作環境，並從過程中學習來自各國的菁英之研究態度與經驗，相信遴選生們已具備多元思維的能力，並具備有廣闊的國際視野與國際價值觀。特別值得一提的是，台灣學生一直缺乏主動發問的習慣，在這次的實習過程中，每位遴選生於 Microsoft 公司積極開放的環境中的薰陶，皆變的非常的主動發問，一有不懂的細節，皆會於工作會議中以英文來提出並詢問，並勇於與來自全球各地的專業菁英切磋砥礪，這便是每位遴選生一項

非常明顯的收穫。

最後，在此由衷感謝教育部-學海築夢之經費補助，讓來自台北科技大學的 13 位遴選生有機會能夠參與海外實習經驗。這些學生於此趟的實習之旅不僅在專業研究方面獲益良多，更是學習到不同的國家文化，也感受到來自於不同國家專家的思維，更能夠從中獲得不同的看法與見解。這趟實習不但可以提升台灣和國立臺北科技大學在多媒體專業發展上的知名度，亦可吸收最新的工業科技相關之技術，未來期望能有更多的相關計畫機會能讓更多的學生參與及體驗國外工作實習的經驗。

