

出國報告（出國類別：參加國際研討會）

**Color CENTRIST: 應用於影像場景分  
類的顏色描述子  
以及  
發掘視覺樣版應用於建築物影像分類  
與產品影像搜尋**

服務機關：國立中正大學資訊工程學系

姓名職稱：朱威達助理教授

派赴國家：香港

出國期間：2012年6月4日至6月8日

報告日期：2012年6月15日

## 目次

一、	摘要 .....	2
二、	目的 .....	2
三、	過程 .....	2
四、	與會心得 .....	6
五、	建議 .....	7
六、	攜回資料 .....	7

## 一、摘要

此次於 2012 年 6 月 4 日到 6 月 8 日至香港參加 2012 年 ACM International Conference on Multimedia Retrieval(ACM 多媒體檢索研討會)，簡稱 ICMR，主要任務是發表兩篇論文，並且帶領學生進行國際學術交流。ICMR 是多媒體檢索領域的頂尖研討會，會中聚集來自數十國的專家學者，安排三位知名學者與業界領袖的演講，並且提供一個平台讓學者們發表與溝通學術成果。會中亦安排至香港邵逸夫創意媒體中心，以及維多利亞港旁的星光大道參觀，讓與會者充份體驗香港的文化。此次與會經驗收獲豐富，也藉此鼓勵學生多走向國際，努力研究以獲得國際肯定。

## 二、目的

此次到香港參加 2012 年 ACM International Conference on Multimedia Retrieval 發表兩篇論文，分別是「Color CENTRIST: 應用於影像場景分類的顏色描述子」(Color CENTRIST: A Color Descriptor for Scene Categorization) 以及「發掘視覺樣版應用於建築物影像分類與產品影像搜尋」(Visual Pattern Discovery for Architecture Image Classification and Product Image Search)。這兩項研究結果都是在 2010 到 2011 年經歷兩年的時間研究，在 2011 年年底投稿，2012 年三月被接受的成果。此次到香港參加此研討會的主要目的是與其他各國的研究者交流研究想法，並且希望能從參與國際會議的過程中增加中正大學與台灣學術界的知名度。另外，這次本人的一位碩士班研究生(開發 Color CENTRIST)也申請到國科會研究生出國參與國際會議的補助，因此本人此次出訪的另一個目的是帶領學生參加國際場合，讓他了解國際會議的運作與學術交流實際進行的狀況，讓學生除了平日單方面閱讀他人的論文之外，也可以直接面對面與其他研究者溝通。

## 三、過程

2012 年的 ICMR 於 6 月 5 日至 6 月 8 日舉行。根據大會的統計，今年有 145 篇論文投稿，最後由來自 18 個國家共 50 篇論文入選，接受率約 34%。此外，這次共有 2 個特別議程，包括 Socio-Video Semantics 與 Social Events in Web Multimedia，共收錄 11 篇論文，來自 10 個不同國家。這兩個特別議程特別能顯示目前多媒體檢索的最新趨勢。值得一提的是，ICMR 是由 ACM CIVR (Conference on Image and Video Retrieval) 以及 MIR (Multimedia Information Retrieval Workshop) 兩個以前已分開進形的兩個小研討會所合組而成，目的是讓該領域的研究者能更加專注，且進一步提昇多媒體檢索

領域在整個多媒體研究群中的地位。從第一屆的 CIVR 算起，今年的 ICMR 已經算是多媒體檢索研討會的 10 週年，因此大會也在 6 月 7 日的晚宴中特別慶祝。

本人帶領研究生於 6 月 4 日搭機前往香港，在會議前一天出發的原因除了為了趕上第一天會議的 tutorial 課程之外，另一個原因是此次的機票是由個人哩程累積，在 2011 年就先行更換而得，相關手續費並未申請任何補助。

6 月 5 日的會議內容是三場 tutorial，大會請到 Google Research 的 Edward Chang 與台灣大學林智仁教授講解 Foundations of Large-Scale Multimedia Information Management & Retrieval。從 Edward Chang 任職的 Google 來看，大規模的多媒體內容檢索與管理已經是必然的趨勢。台大林智仁教授則是機器學習領域中國際聲譽卓著的學者，他給予大量使用到機器學習的多媒體研究者許多建議。第二場 tutorial 的主題是 Music Information Retrieval，由來自奧地利的 Markus Schedl 與日本的 Masataka Goto 教授主講。Goto 教授是音樂檢索領域的明日之星，他所開發的 PodCastle 與 Songle 系統相當令人驚豔。本人與 Goto 教授在會中有許多時間溝通意見，並交流彼此的研究心得。本人在 2012 年 10 月底於日本奈良舉辦的一個 workshop 之中也請到他來進行專題演講。第三場 tutorial 的主題是 3D Video Segmentation, Recognition, and Retrieval，由來自 University of Texas at Dallas 的 Prabhakaran 教授主講，他同時也是 2013 年 ICMR 的會議主席。三場 tutorial 都在香港城市大學的邵逸夫創意媒體中心進行，這棟建築物由名建築師設計，是香港城市大學資訊工程與多媒體研究的大本營。從此建築我們也可看出香港的大學有來自政府與商界的經費挹注，在硬體設備與招募人才上都很有規模。

6 月 6 日早上 8:45 先進行簡單的會議開幕典禮，主辦單位請來醒獅團祝賀，並期許會議圓滿成功。會議議程在主席團簡單的“點睛”儀式後展開。九點鐘由第一個 keynote speaker 上場，他是來自美國哥倫比亞大學的 Prof. Paul Sajda，演講主題為 Cortically-Coupled Computing for Multimedia Retrieval。Sajda 教授從人觀看不同影像在磁共振造影(MRI)影像上的差別來進行多媒體檢索研究。這是一個包含醫學工程與影像處理的跨領域研究，也把多媒體研究跟人的感官結合在一起。這是一個重要的方向，但是如與會者在演講後所提到，類似的資料不容易取得，使用上還不方便，因此本研究目前還處於剛啟動的階段。演講後經過短暫休息，10:20 開始進行第一個 oral

session，主題是 **Annotation and Classification**。Image/video annotation 一直是近年來重要的研究議題，這個議程中的四篇論文有的是整理過往提出的方法與特徵進行比較，有的則是提出新的觀點，可見論文可從不同的角度切入，皆能對該研究有所貢獻。中午在會場“九龍香格里拉飯店”用午餐，緊接著下午 1:40 進行第二個 oral session: **Near-Duplicate and Copy Detection**。這同樣也是近年來的重要議題，從舊題目中開發出新意是我在這個議程中最大的收穫。3:40 開始的議程是特別由法國 Eurecom 的 Benoit Huet 教授所組織的特別議程：**Social Events in Web Multimedia**。此新興的議題在這兩年獲得高度重視，我們可藉由網路上使用者留下的大量社群資料，如分享相片或影片、標籤、註解等，建立模型來描述媒體的資訊，並進一步偵測或預測網路上的社交事件。這次有三篇都來自歐洲的論文針對此議題進行討論。5:00 進行本日的最後一個議程，主題相當有趣，名為 **Fresh Views on Multimedia Retrieval**。顧名思義，這個議程中的文章屬於重新檢視或從另一個角度來看多媒體檢索，並引發新的研究方向。其中讓人印象深刻的有來自於西班牙的 Prof. Simone Santini 提到，以列表(list)的方式來呈現搜尋效果不見得是一個直覺的呈現方式。還有來自 Yahoo Research 的 David Shamma 提到他們所開發的多人互動與視訊瀏覽系統。6:40 結束第一天緊湊的會議，與會者在 Welcome Reception 時段在會場聚會聊天，會議主席除了跟會議籌備委員表達感謝之意，也帶領與會者到飯店旁的海灣邊步道欣賞每晚八點的維多利亞港燈光秀。本人在這段時間相當榮幸與大會主席 Yong Rui (Microsoft China)一起步行聊天。他在多媒體領域的影響相當大，能將自己介紹給他認識算是此行相當重要的收穫之一。藉由這樣的機緣，他在隔天的晚宴上也將我介紹給第三場 keynote 的演講者，也是來自於台灣，目前是亞洲微軟研究院院長的 Hsiao-Wuen Hon 認識。

6月7日早上 9:00 進行第二場 keynote speech，講者是來自法國 INRIA Research Center 的 Cordelia Schmid 博士。Schmid 博士在電腦視覺所開發的特徵方法被廣泛應用於多媒體系統中，對於此領域有極大的貢獻，她的論文已經被引用數千次，可說是近幾年扮演先驅者的重要研究者之一。她提出一個簡單但相當有效的描述法，適用於廣泛的影像搜尋問題中。本人過去兩年也採用了 Schmid 博士團隊提出的方法，因此對於這場演講特別有感覺。緊接著 10:20 進行本日的第一場 oral session，主題為 **Machine Learning and Hashing for Multimedia Retrieval**，此主題著重在利用機器學習的技巧以促進多媒體檢索。尤其是雜湊函數(hashing)最近兩年帶來新的突破，也因此格外獲得大家的重視，與會者紛紛提出問題與論文發表者討論。這次會議其中一個特別的地方

是議程在兩個會場舉行，部份議程在九龍香格里拉飯店，部份議程在香港城市大學邵逸夫創意媒體中心。6月7日中午用完午餐之後，主辦單位派巴士將我們從飯店送往邵逸夫創意媒體中心以進行下午的議程。這輛巴士充分顯示香港地小人稠的特性，每排五個座位，一輛巴士居然坐了六十幾個人。在創意媒體中心，下午 1:40 進行重頭戲 Best Paper Session，由來自 Princeton、CMU、Oxford、Univ. of Amsterdam、以及 German Research Center for Artificial Intelligence 的學者競逐。其中讓人印象最深的是 Oxford University 的研究 “Name the Sculpture”，利用視覺的資訊，此研究能辨識著名的雕像與其作者的相關資訊，相當有趣且引發新的研究方向。不過後來獲得最佳論文獎的是來自德國的團隊，他們提到利用觀看者的人口統計特性來幫忙做視覺概念分析。這個新的概念獲得委員的青睞，在當天晚宴時公布得獎。3:40 進行第二個特別議程 “Socio-Video Semantics”，與第一個特別議程相似，本議程也是以分析使用者行為或特性出發，研究其對多媒體檢索相關議題的幫助。

我們的兩篇論文在 6 月 7 日下午 5:00 至 6:20 的海報論文議程中進行發表，與會者在聽完兩天的口頭報告之後有機會與論文發表者面對面一對一溝通研究成果。本次會議是 single track，也就是同一時間只進行一項議程，因此所有與會者都有充分的時間互動，也因此海報論文議程格外熱鬧。這次我們發表的第一篇論文是 “Color CENTRIST: 應用於影像場景分類的顏色描述子”，我們提出一個新的描述子來描述影像場景，並證實加入顏色資訊之後能讓場景分類的效果提昇。發表過程中，法國 INRIA 的研究團隊對本論文特別感興趣，也給予我們相當多的建議跟指導，讓我們獲益良多。第二篇論文是 “發掘視覺樣版應用於建築物影像分類與產品影像搜尋”，我們將影像中特徵點的空間關係描述為圖(graph)，並利用找尋頻繁子圖(frequent subgraph)的演算法將視覺樣板(visual pattern)找出來。利用視覺樣板，我們進行影像中的建築風格分類與產品搜尋。來自 Princeton 的研究者對這項研究特別感興趣，在了解過後也分享他過去相關的研究經驗。在與多位學者一對一講解之後，海報議程在六點半結束。接著，主辦單位帶領我們到城市大學旁邊的商場，當天的晚宴地點就在商場內的一家餐廳進行。在用餐的間隔時段，主辦單位頒發最佳論文獎，最佳評審獎，以及公布明年度的舉辦地點。此外，為了慶祝 ICMR 十週年，主席也細數過去的歷史，並安排一些慶祝活動。席中本人與學生跟來自中國、荷蘭、法國的學者同桌。看到學生從一開始與會不敢與外國人溝通，到晚宴間已能跟旁邊的人以英文溝通，我深感覺到讓學生多接觸國際場合的重要性，短時間的訓練可以讓他們成長很多。

6月8日進行本會議最後一天的行程。早上9:00 第三場 keynote speech 由亞洲微軟研究院院長 Hsiao-Wuen Hon 博士主講，講題是 **The Road to Pervasive Multimedia Search and Multimodal Interaction**。洪博士是台灣人，到美國 CMU 拿到博士學位，是國際知名的語音辨識的專家。他在演講中講了一系列亞洲微軟研究院的研究成果。亞洲微軟研究院在過去十年間聲譽卓著，在許多領域都佔有領先的地位。這天早上最特別是從早上 10:00 開始的 Demo session。每位發表者有一分鐘的時間上台用簡短扼要的方式介紹他的研究，並鼓勵大家到他的攤位參觀。之後與會者可自由與發表者面對面溝通，並由發表者展示他的系統。在這個議程中，我印象最深的是平板電腦上的應用大行其道，例如來自奧地利、希臘、愛爾蘭、日本等地的學者都以它當為展示平台。來自德國的新聞視訊分析平台也讓人印象深刻。這個議程另一個特別的地方是讓與會者票選最佳展示系統，而非由資深學者組成委員會評選，因此更能代表大家的喜好。此榮銜最後由希臘的團隊，以 “Cluster-based Photo Browsing and Tagging on the Go.” 為題獲得。本人此次在本會議的行程就到這為止，因為班機時間的關係，我中午就離開會場。其實大會下午還安排了一個 industrial session，邀請學界跟產業界的人給幾場演講，並且跟與會者溝通討論。另外還有一個 panel session，幾位資深的研究者分享他們對於 “Where is the user in Multimedia Retrieval” 的看法。

#### 四、與會心得

如同之前所述，本人覺得此行最大的收穫之一是跟大會主席 Yong Rui 有面對面的接觸與認識。會有這樣的機緣主要是本會議是多媒體檢索領先許多重量級研究者會出席的會議。因此，我再次肯定參與的會議品質一定要高，這樣才容易有明確的收穫。另外，本會議的參與人數不算多，約 100 人左右，這也讓與會者有充份的時間討論。主辦單位同時間只安排一個議程，所以可以集中與會者的注意力。

另外一個感想是學生出國參與學術活動的重要性。才不過幾天的時間，我帶領的學生就從緊張無法與別人以英語應對，到可以在宴席間與隔壁的學者聊天溝通。這對於一個碩士班學生來講相當難能可貴。所以，除了讓學生瞭解學術發表的型式跟場合之外，他們還能體會不同國家的文化，與他人溝通，拓展他的視野。當然，因為經費有限，政府或學校無法補助所有的學生出國，但如果學生的成果已確定被國際頂尖的會議接受，從教育的觀點來看應該大力補助。

## 五、建議

現在大部份的會議要求每篇論文都要有一個完整(非學生身份)的註冊，此次會議本人發表兩篇論文，所以光註冊費就花掉四萬多元。事實上，將多個研究成果集中投稿於一個頂尖會議是較為經濟的作法，我們只需花一趟機票錢跟生活費就發表了多個成果。然而，目前國科會所補助的基礎是以機票費生活費註冊費總額補助，一旦註冊費提高，相對的就排擠到機票與生活費的補助。我們很高興成果豐碩能一次發表多篇論文，但也苦於成果越多負擔越大。近年來機票與註冊費都漲，我們知道政府財政困難，但也感覺到資源分配不夠合理。若能加強審查機制，或者給不是五年五百億學校中努力且獲得不錯成績的研究者更多資源，我想更能促進良性的學術競爭。

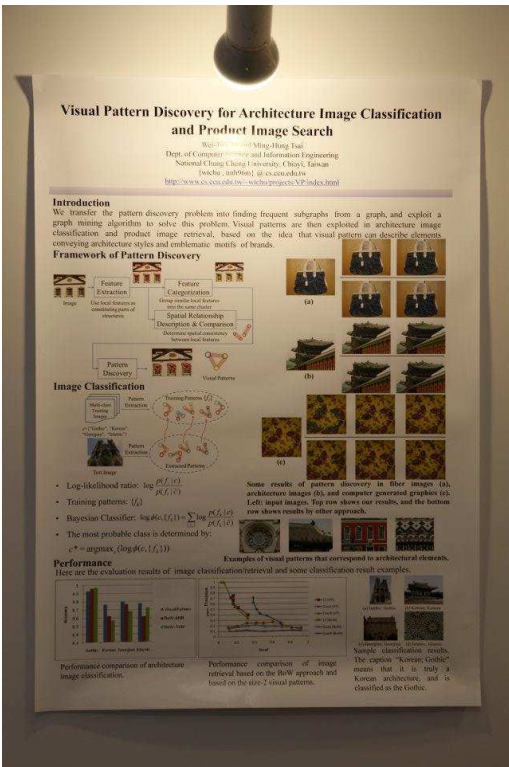
## 六、攜回資料

- Proceedings of ACM International Conference on Multimedia Retrieval, 2012.
- Call for paper, ACM Workshop on Crowdsourcing for Multimedia, 2012.
- Call for paper, Special issue on Multimedia Event Detection, Machine Vision and Applications, 2012.
- Call for paper, Special issue on Socio-Video Semantics, IEEE Transactions on Multimedia, 2012.
- Call for paper, Special issue on Learning from Multiple Evidences for Large Scale Multimedia Analysis, Computer Vision and Image Understanding, 2012.





本人與學生在會場合照



在此會議發表的兩篇論文

