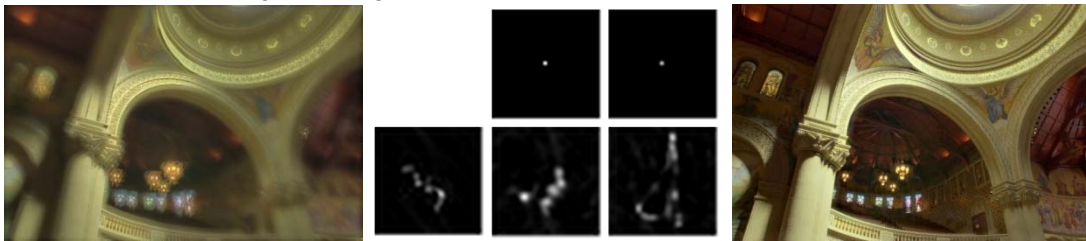


## 莊永裕老師研究簡介

我的研究領域為電腦圖學、電腦視覺及多媒體，近年來的研究主軸為計算攝影學、三維影像處理、全域照明及視訊索引。

### 1. 計算攝影學(computational photography)

計算攝影學是近年來的一個新興研究領域，討論如何利用計算的方式改善攝影的結果，其研究範圍包括提昇解析度(super-resolution)、去除模糊(deblurring)、提升動態範圍(high dynamic range imaging)等。我們的 CVPR 2009 論文” High Dynamic Range Image Reconstruction from Hand-held Cameras”提出第一個可用於手持式相機的高動態範圍影像重建方法，克服在長曝光時間下所得的相片會因手震而產生模糊的問題，從而由模糊的不同曝光相片中取得一個清晰的高動態範圍影像。在理論上，我們提出一個成像模型，描述清晰影像、模糊核心(blur kernel)及相機反應曲線(response curve)之間的關係，並以貝式模型(Bayesian model)對其求解，是第一個將去模糊及提升動態範圍的問題置於同一模型下求解的方法。在應用上，使用者在拍攝高動態範圍影像時可以不必使用笨重的腳架，使得拍攝變得更容易。此外，我們於 ICCV 2009 提出一個新的視訊穩定(video stabilization)的方法，跳脫傳統運動估計(motion estimation)、運動補償(motion compensation)、重新合成(composition)的三步架構，由一個新的角度來處理視訊穩定問題，避免運動估計誤差造成的不穩定，從而得到一個更強健的(robust)視訊穩定方法，提升家用相機所取得的視訊品質。另外，在 CVPR 2012 我們提出一個考慮維持線性結構的影像大小改變(image resizing)演算法。



未去除模糊之 HDR 影像

估計之模糊核心

去除模糊之 HDR 影像

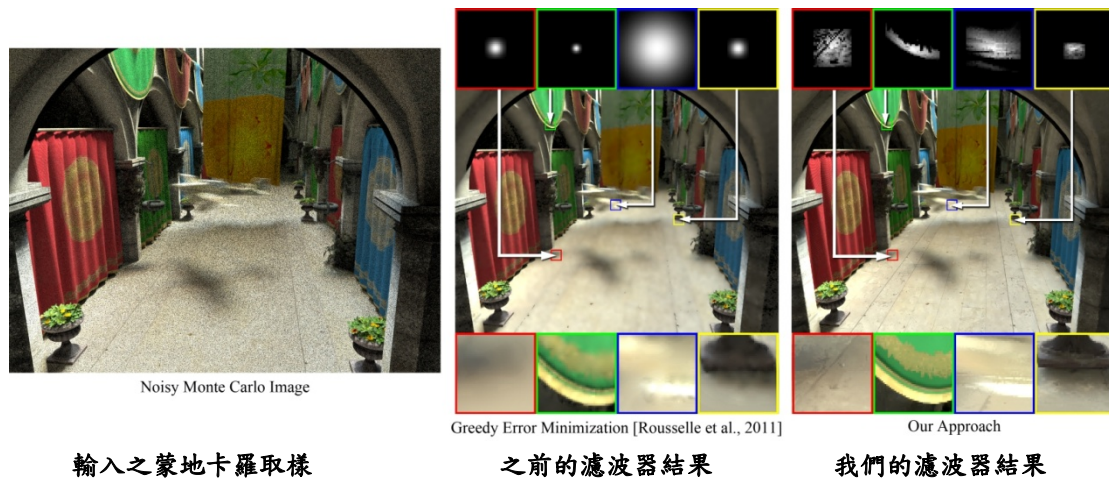
### 2. 立體影像處理(stereoscopic media processing)

三維立體影像為近年電子消費市場之一重要趨勢，惟其進展過去多著重於硬體設備，體認到三維立體影像編輯將於硬體普及後成為一個重要的研究趨勢，我們自 2009 年開始即投入三維立體影像編輯之研究，我們的 IEEE Transactions on Multimedia 論文”Content-Aware Display Adaptation and Interactive Editing for Stereoscopic Images”提出一個立體影像縮放及深度編輯的方法，讓使用者可以自由調整影像之大小及物體之深度，而不會帶來扭曲，以保持影像之自然，為三維立體影像編輯的研究先驅之一。CVPR 2012 的論文則提出一個基於圖層(layer)的立體影像大小改變演算法，進一步提升影像的品質。此外，我們整理三維電影製作原理之立體影像編輯原則(3D cinematography principles)，並將其應用於多項立體影像編輯應用中，如三維視訊穩定(stereoscopic video stabilization)、三維相簿瀏覽(stereoscopic photo slideshow)等，這些原則可以廣泛應用於許多不同的立體影像編輯應用中，該文發表於 ACM Multimedia 2011。在 ACM SIGGRAPH Asia 2012 提出一個立體影像複製(cloning)演算法，讓使用者可以輕易地由一張影像剪貼物體至另一張影像，並維持合成影像的真實性。



### 3. 全域照明(global illumination)

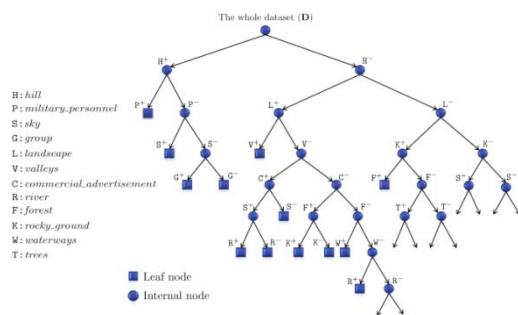
我們的 SIGGRAPH Asia 2012 論文改善蒙地卡羅追光法(Monte Carlo Ray Tracing)中由於取樣不足之變異(variance)所造成之雜訊影像(noisy image)問題。利用多個不同尺度之濾波器(filter)來減少雜訊，大尺度的濾波器雖可大幅減少雜訊，但是可能使得影像過於平滑而失去細節；反之，小尺度的濾波器雖可保持影像細節，但是其雜訊抑制的效果較差，我們利用統計學公式 Stein's Unbiased Risk Estimator (SURE)來估計不同尺度之濾波器之可能誤差，從而得以選取最佳之濾波器以重建出具細節且抑制雜訊之影像，並可由此一誤差估計找出需要增加取樣之位置，而達到調適取樣(adaptive sampling)的目的，得到一個簡潔優雅的取樣與重建架構，對於快速產生高品質影像極具實用價值。



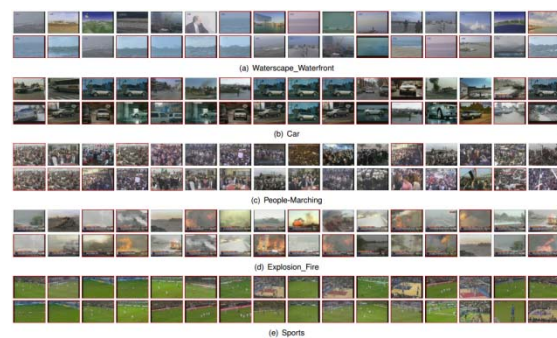
### 4. 視訊索引(video indexing)

由於數位視訊擷取設備的普及、儲存裝置的容量迅速地成長以及資訊分享平台的蓬勃發展，加快了視訊的複製、傳播和交換，個人擁有的視訊內容不斷地增加，為能有效地保存和使用這些數位內容，視訊索引遂成為近年來多媒體及資訊擷取研究的一個重要課題。視訊索引分析視訊的內容，以對其加註索引，如汽車、郊區、大樓、天空等語意標記(semantic index)，以便利視訊之蒐尋。傳統上，要判斷某一影像是否具有某一語意概念時，一般都是由語意偵測器針對單一語意概念以及使用從單一影像上所擷取的特徵進行預測，這種方法完全依賴從影像上所擷取的特徵是否具有高度代表性以及事先訓練好之語意偵測器的可靠度，很不幸地，對部分語意概念來說，偵測的效果並不理想。我們利用視訊的語意關聯性及時間連續性來提升索引的準確率。我們首先於 ACM Multimedia 2008 提出一個監督學習演算法(supervised learning)由使用者標記中學習視訊的前後文關係(inter-concept association)及時間關係(inter-shot relationship)提升索引的準確率。所謂的

前後文關係指的是，同一個影像上若出現語意概念 A 的話，通常會伴隨著語意概念 B 或者語意概念 C 就一定不會同時出現。另外，時間關係指的是當語意概念 A 出現在某一影像時，一般來說，接下來的數個影像以至於數十個影像也會同樣出現語意概念 A，這兩種關係可以用來補強語意偵測器不足的地方，提升語意偵測的正確性達 30%。我們亦提出一個跨域學習(cross-domain learning)的演算法可以由一個類別的視訊學習前後文關係及時間關係，再將其用於另一類別的視訊中，該文發表於 IEEE PAMI。最後，我們提出了一個全自動、不需標記之非監督式學習(unsupervised learning)方法，可以直接由不精確的語意偵測器中利用協同過濾(collaborative filtering)的概念來提升語意偵測的正確性，由於其不需人工標記訓練資料之特性使得其更易於使用。此一系列研究聚焦於利用語意關聯性及時間連續性來提升語意偵測的準確率，先後提出了監督式學習、非監督式學習及跨域學習方法，為語意偵測結果融合領域中最全面的成果之一。



學得的前後文關係



語意偵測的結果

更多研究說明請參考 <http://www.csie.ntu.edu.tw/~cyy/projects/index.html>