出國報告（出國類別：國際會議）

參加 2016 IEEE IVMSP 國際會議發表論文

服務機關：國立虎尾科技大學 資訊工程系

姓名職稱：林武杰 副教授

派赴國家：法國

出國期間：105 年 7 月 5 日 – 105 年 7 月 20 日

報告日期：105 年 9 月 12 日

摘要

本報告說明計畫主持人林武杰參加 IEEE IVMSP 國際會議發表論文的過程、心得、與建議。會議期間為 105 年 7 月 11-12 日，為期 2 天，地點在法國波爾多。這次研討會的主要議題為 low dimensional models for image and video processing and analysis。本研討會論文的評審競爭激烈，僅有 37 篇被接受，其中只有 3 篇為單一作者(single author)，計畫主持人為其中一篇的單一作者。藉由研討會會議與其他國家的學者交流，計畫主持人不僅可以了解目前相關領域的技術研究現況，也激發了不少新的研究想法可以應用在解決實務問題。有關影像處理與視訊分析的相關國際研討會，目前舉辦地點仍是以美國居多，IVMSP 是少數今年在歐洲舉辦的研討會之一。今年參與會議的研究學者皆是以來自歐洲地區的為主，因此是一個接觸歐洲相關領域專家學者不錯的機會，建議將來有機會仍可以參加類似特性的研討會，以拓展國際學術人脈。

目次

目的

計畫目標

本次參加 2016 IEEE IVMSP (Image, Video, and Multidimensional Signal Processing) Workshop 國際會議的目的是希望發表計畫主持人的論文，two-class clustering of nonlinearly separable data by using shape-specific points, 了解影像與視訊處理相關領域最新的研究成果，同時也藉此會議與相關領域的歐洲學者交流，拓展國際的學術人脈。

主題

本次研討會論文的研究主題包括(但不限於)

- Image and video modeling on manifolds
- Low dimensional models and low rank methods
- Sparse and low-rank models in learning and pattern Recognition
- Graph-based methods for image and video analysis
- Image & Video Processing Techniques
- Image & Video Sensing, Modeling, and Representation
- Image & Video Analysis, Synthesis, and Retrieval
- Computational Imaging

而這次研討會的主要議題為 low dimensional models for image and video processing and analysis。除此之外，本次也邀請了一些學者進行 plenty talk，其中包括了

- Baba C. Vemuri, Efficient recursive algorithms for statistical analysis of manifold-valued data
- Pierre Vandergheynst, Sampling, inference and clustering for data on graphs

另外也有一些 invited talks，包含了：

- Tulay Adali, Data fusion using source separation: role of diversity
- Michael Bronstein, Geometric Deep Learning

- Xavier Pennec, Barycentric subspace analysis: an extension of PCA to manifolds
- Anuj Srivastava, Elastic Riemannian Frameworks for High-Dimensional Signal Processing
- Michael B. Wakin, subspace Approximations on the Continuum
- Mathews Jacob, Superresolution Imaging using Piecewise Smooth Image Models
- Oliver Cossairt, X-RAY Fluorescence Image Super-resolution using Dictionary Learning
- Saiprasad Ravishankar, LASSI: A Low-Rank and Adaptive Sparse Signal Model for Highly Accelerated Dynamic Imaging
- Yon Visell, Learning Parts of Touch in the whole Hand with Sparse Dictionary Learning

緣起

利用形狀特徵點 RWM(Radius Weighted Mean)[1]與系統中心點來進行分群的想法最早是由 Yang and Lin[1]所提出，不過當時的研究僅能處理線性可分資料。因此計畫主持人進一步將形狀特徵點擴展到處理非線性可分的資料，並且利用人工產生的資料點來驗證研究中所發展的公式解其效果。在研究的過程中，我們發現當擴展 Yang and Lin [2]的方法在處理非線性資料時，決策平面(decision hyperplane)通過 RWM 這個形狀特徵點有時並不是一個好的選擇，因此我們在研究中也開始找尋更佳的通過點，讓分群的品質可以更進一步改善。除此之外，核心函式的選擇對於分群結果影響也很大，在本研究中，我們實驗了二種常見的核心函式在相同資料的分群表現，藉以提供研究學者在選擇核心函式時的參考。同時在我們的研究中，我們也探討本研究的公式解是否對於資料的 translation, rotation, scaling 具備不變性(invariant)，可惜限於篇幅在本次論文中這部分的研究成果並未能列入。

考量本研究的成果，我們認為具有以下的學術貢獻：

(1) 可以直接利用公式(closed-form formulas)，將線性不可分割的資料點快速分為二群

(2) 本研究的公式不需初始解，因此分群結果不像 k-means 法受初始解影響很大

(3) 不需反覆改善目前解，因此不需 iterative procedure，直接使用公式即可求出解

(4) 本研究所提出的公式解可以適用於任何維度的資料

(5) 由於分群速度快，因此可以適合應用在處理大量的資料，像是影像或是視訊。

<u>預期效益</u>

藉由參加這次的國際會議，我們可以多了解不同研究學者對於本研究的想法以及建議，同時也可以與歐洲相關領域的學者交流拓展學術人脈，了解是否有較新的相關領域研究議題或是計畫正在進行。

過程

本次研討會有二個 plentary talks, 在 Vemuri 的演講中有談到了 Frechet Mean (FM), 其定義為

$$M = \arg\min_{Y \in \mathcal{M}} \sum_{i=1}^{n} d^2(Y, X_i).$$

其中 n 個資料點所形成的集合 $\{X_1, X_2, \cdots, X_n\} \subset \mathcal{M}$，而 $\mathcal{M}$ 為 Riemannian manifold。下圖可以明顯地表示 Intrinsic mean 與一般的 Arithmetic mean 有何不同

(Left): Arithmetic mean and **(Right):** Intrinsic mean of data points lying on 1−D manifold.

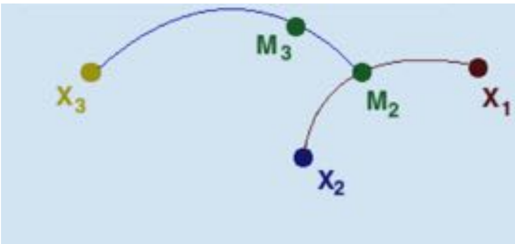
根據 Afsari et al. [3]以及 Pennec [4] 的研究， 在一個凹式球(geodesic ball)若是曲率(curvature)小於 $\pi/(2\sqrt{k})$，則 FM 存在並且唯一。當資料點的數量 n>2 時，則 FM 的計算無公式解，必須使用迭代的方式(iterative schemes)以梯度法(gradient descent)來計算。由於梯度法的缺點包括了：(1)資料點數多時，計算量龐大，(2)新增資料點時，FM 必須重新計算，(3)必須完整儲存所有資料點，方能計算 FM。有鑒於此，Vermuri 在演講

中提出了利用漸進式的方式(incremental technique)來預估 FM 的位置，並將其稱為 iFME(incremental FM estimator)，其精神是利用下列方式來預估 FM

$$
\begin{aligned}
M_1 &= X_1, \\
M_{i+1} &= \Gamma_{M_i}^{X_{i+1}}(w_{i+1}),
\end{aligned}
$$

其中 $w_{i+1} = \frac{1}{i+1}$, and $\Gamma_X^Y : [0,1] \to \mathcal{M}$

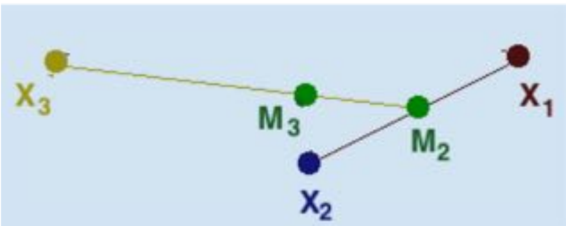
參考下圖



他的想法是源自於在尤氏空間(Euclidean space)，資料點 $X_1, X_2, \ldots, X_n$的算術中心(arithmetic mean)可以用下列漸進式的方式計算而得

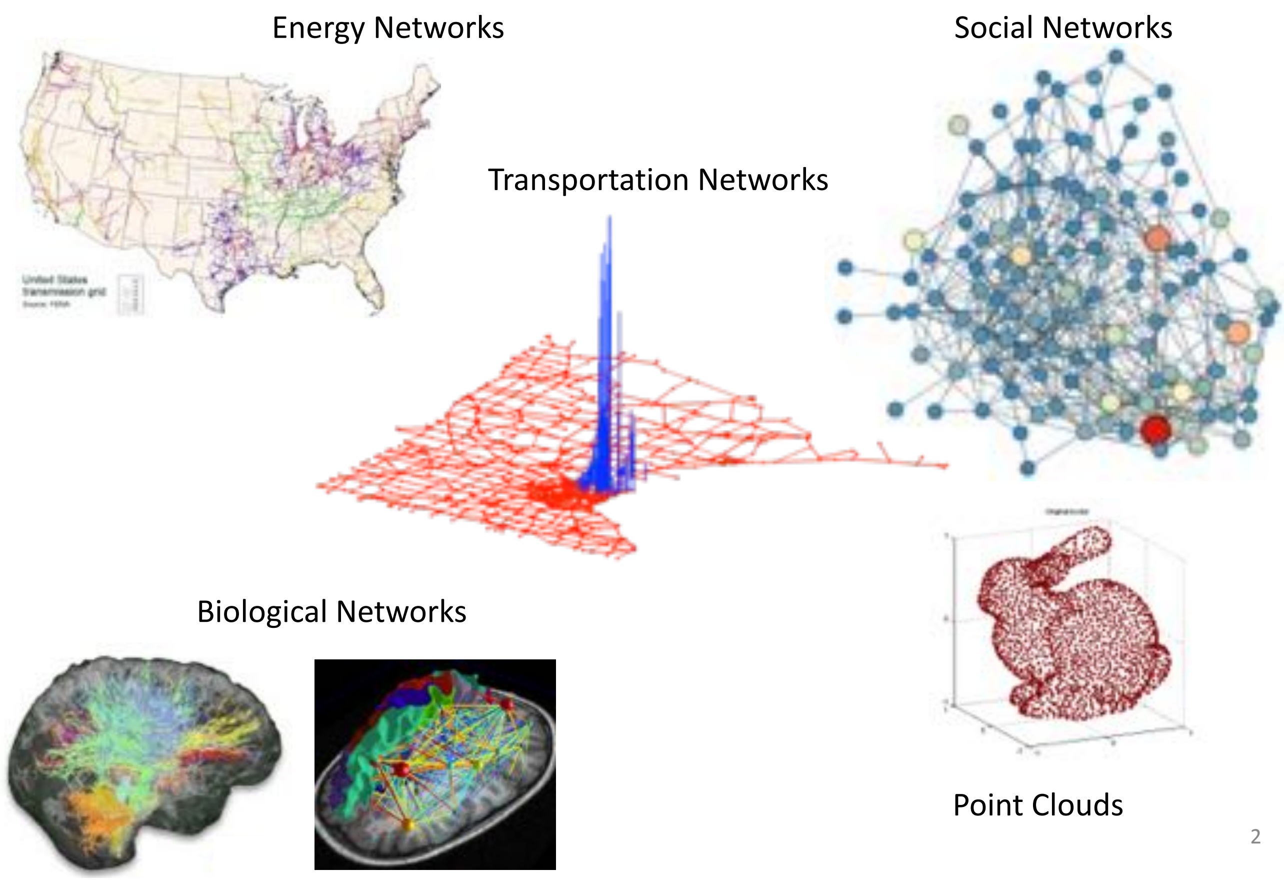
$$
\begin{aligned}
M_1 &= X_1, \\
M_{i+1} &= \frac{1}{i+1} X_{i+1} + \frac{i}{i+1} M_i.
\end{aligned}
$$

參考下圖



Vemuri 所提出的方法適合用於位在球體(hypersphere), Grasmannian, SPD(n), 以及 Stiefel manifold 上的資料點計算其 FM 位置。基於 iFME，Vemuri 進一步提出了 CCM-EPGA, 與傳統的方法 PGA 或是 exact PGA 相比，CCM-EPGA 的 projection error 非常低

Vandergheynst 的演講是提出針對圖形(graph)上的節點(vertex)，當只有少數節點被觀察到時，如何有效的推導出未觀察到的資訊(或是節點)、採用何種訊號模式、以及利用少數觀測到的節點推導出圖的完整結構等問題。這個問題在現今環境非常實用，因為在我們的日常生活中有許多這樣的例子，像是交通網路、社群網路、能源網路、生物神經網路等等，參考下圖，若能藉由少數觀察到的樣本節點及其連結，能夠推導出其



完整結構的樣態，對於實際問題的解決將會非常有效而且快速。一個圖形 G 可以用其節點 V, 邊線連接ε，以及邊線權重 W 來表示，G={V, ε, W}，其中$W \in R^{n\times n}$。令對角分支矩陣(diagonal degree matrix) D 的對角線元素為 $d_i := \sum_{i\neq j} W_{ij}$ ，則組合圖拉普拉斯為(Combinatorial graph Laplacian) L := D-W，正規化後(normalized) $\mathsf{L} := \mathsf{I} - \mathsf{D}^{-1/2}\mathsf{W}\mathsf{D}^{-1/2}$ 。由於 L 是實數值對稱 PSD, 因此 $\mathsf{L} = \mathsf{U}\Lambda\mathsf{U}^{\intercal}$，其中 U 為特徵向量形成的矩陣 $\mathsf{U} \in \mathbb{R}^{n\times n}$，稱為 Graph Fourier Matrix，而 $\wedge$ 為非負值特徵值所形成的矩陣。令 $x \in R^{n\times n}$

是 k-bandlimited signals，傅立葉係數為(Fourier coefficients) $\hat{x} = U^T x$，則 x 在圖型上為一平滑的 k-bandlimited signals,

$$x^T L x = \sum_{i \neq j} w_{ij} (x[i] - x[j])^2 = \sum_k \lambda_k |\hat{x}^k|^2$$

若 $x = U_k \hat{x}^k, \quad \hat{x}^k \in R^k$, $U_k := (u_1, \ldots, u_k) \in R^{n \times k}$, 其中 $u_1, \ldots, u_k$ 為前 k 個特徵向量。

當取樣的模型(Sampling Model)為隨機取樣，

$$\boldsymbol{p} \in \mathbb{R}^n \quad \boldsymbol{p}_i > 0 \qquad \|\boldsymbol{p}\|_1 = \sum_{i=1}^{n} \boldsymbol{p}_i = 1$$

$$\mathsf{P} := \mathrm{diag}(\boldsymbol{p}) \in \mathbb{R}^{n \times n}$$

若獨立選取 m 個隨機樣本

$$\mathbb{P}(\omega_j = i) = \boldsymbol{p}_i, \quad \forall j \in \{1, \ldots, m\} \text{ and } \forall i \in \{1, \ldots, n\}$$

則

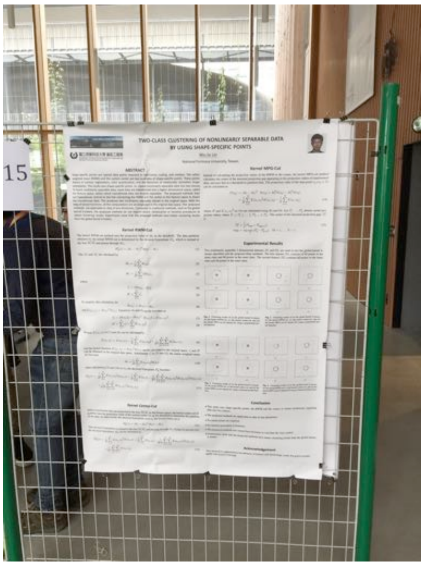
$$\boldsymbol{y}_j := \boldsymbol{x}_{\omega_j}, \quad \forall j \in \{1, \ldots, m\}$$

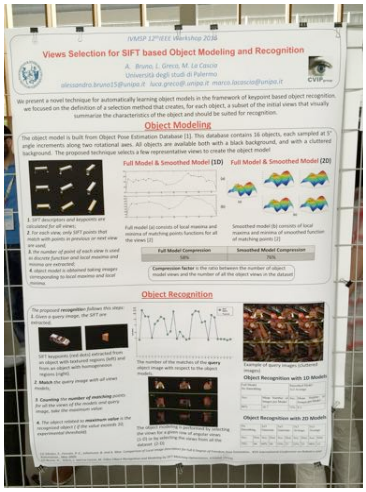
$$\boldsymbol{y} = \mathsf{M}\boldsymbol{x}$$

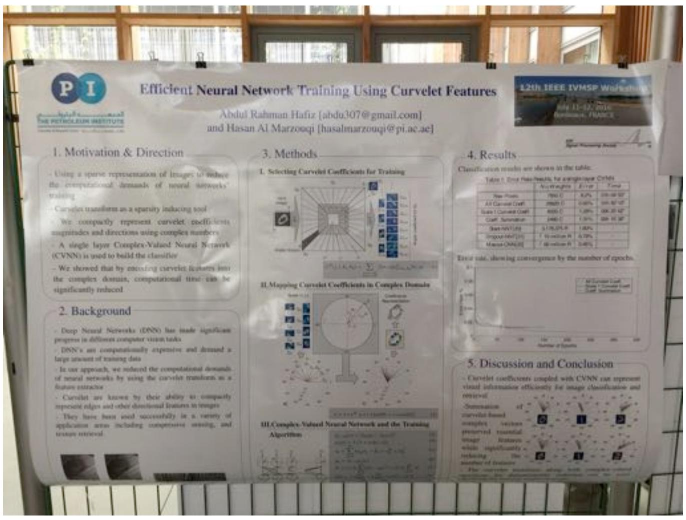
這樣的方式解釋了在實務應用上，若讓使用者對電影評分，則同一社群的使用者其評分結果傾向類似，並且相類似的電影其評分也會類似。因此，藉此理論可以建立一個電影推薦系統，向同一社群的人推薦社群其他人評分較高的電影，這些電影應該也會受到這些人的喜愛。在 Vandergheynst 的演講當中，也提到對於未被評分的電影如何推導其評價，其方法也是透過前 k 個特徵向量來進行，因為這些向量帶有圖型最多的資訊，由於這些是屬於比較細節的部分，演講時許多公式所用到的符號限於投影片的篇幅，並未一一定義，因此需要在會後再找尋其相關的論文研讀，才較能掌握其理論立基，因此在此報告中我們就不詳述。

主持人在本次會議中的論文報告中有一些人提出問題，以 plentary talks 演講者 Vemuri 為例，他對於我利用形狀特殊點來進行分群的想法感到有興趣，也進一步瞭解了本論文在學術上的貢獻，雖然他的問題並不著重在公式的細節，不過他的提問大致能掌握

到重點，讓我覺得本篇論文應該先在期刊發表才對，不然一些不錯的想法很快會被別人直接學走。其他一些學者提的問題較為一般，在此便不一一贅述。下圖為主持人在演講海報前的留影、一些我有興趣的論文海報、以及在會中交流新認識的國際學者。







心得與建議事項

IEEE IVMSP 這個會議舉辦至今 2016 年為第 12 屆，基本上會議的主題比較聚焦，所以參與的人數並不多，這是和其他影像、視訊相關研討會比較不同之處。本研討會透過投稿方式接受的論文都是以 poster 方式發表，因此與會學者可以在期間充分與其他學者交流，以及研討相關技術。今年本研討會來自台灣的論文並不多，除了主持人發表的論文之外，還有一篇是清大教授、一篇是工研院的研究，由此可見本研討會在台灣的能見度還不高，如果想要精緻經營學術人脈，或許是一個不錯的選擇，但是由於接受的論文並不多，所以若希望大量多看一些不同但相關的學術研究的話，本研討會便比較不適合。另外有趣的一點是，有別於美國舉辦的相關研討會，本次研討會來自亞洲的學者並不多，也較少看到大陸留學生的身影。由於本研討會的論文審核嚴格、參與人數不多，因此建議相關學者還是直接將論文發表在期刊會比較划算，研討會網址與評價可參考[5]與[6]。

參考文獻

[1] A. Mitichc, and J. K. Aggarwal, “Contour registration by shape-specific point for shape matching,” Computer Vision, Graphics and Image Processing, Vol. 22, pp. 396-408, 1983

[2] C.Y. Yang and J.C. Lin, “Use of radius weighted mean to cluster two-class data,” Electronic Letters, Vol. 30, No. 10, pp. 757-759, 1994

[3] B. Afsari, R. Tron, et al., On the convergence of gradient descent for finding the Riemannian center of mass, SIAM Journal on Control and Optimization, Vol. 51, No. 3, pp. 2230-2260, 2013.

[4] X. Pennec, Probabilities and satistics on Riemannian manifolds: asic tools for geometric measurements, NSIP, pp. 194-198, 1999.

[5] 研討會官方網址，http://ivmsp2016.org

[6] 研討會評價網址，

https://www.eventadvisor.com/event/ivmsp-2016-ieee-12th-image-video-and-multidimensio/