

修平學報

第二十二期

修平技術學院 編印

中華民國一〇〇年三月出版

第二十二期學報中文目錄

1. 移動機器人融合主動式 RFID 和雷射掃描器之定位
..... 林鴻興、蔡清池 1
 2. 二維圓形薄疊層曲樑之有限變形研究 林秋文 19
 3. 獨立伯努利變數和之變異數與香儂熵 繆紹昌、陳志賢、劉家頤 35
 4. 氣動式手臂之設計及氣壓缸動態模型推導 許銘全、陳建銘、劉芳志 45
 5. 直立圓柱表面微極流體薄膜在磁場效應影響下之非線性破裂理論
..... 宋鴻明、徐仲亭 63
 6. 女性消費者對銀行授信業務刻板印象之研究：以苗栗縣為例 陳泓予 81
 7. 以分散式雜湊機制設計無線隨意網路之需求遠徑協定 張兆村 105
 8. 燒結溫度對於二氧化鈦染料敏化太陽能電池之影響 高銘政 125
 9. 智慧型行車防撞安全系統
..... 麥毅廷、陳振庸、陳怡伸、楊大威、高資淵、紀佳瑋 135
 10. 以化學氣相沉積法合成一維氧化銻鋅奈米線結構及其特性分析
..... 陳志榮 153
 11. 創新動態學：以習慣領域為基礎探討可變空間之決策制定 陳彥曲 163
-
-

Contents

1. Localization of an Autonomous Mobile Robot by Fusing Active RFID and Ranging Laser Scanner	
..... Hung-Hsing Lin, Ching-Chih Tsai	1
2. Finite Deformation of 2-D Thin Circular Curved Laminated Beams	
..... Chiu-Wen Lin	19
3. On the Variance of Sum of Independent Bernoulli Random Variables and Shannon Entropy	
..... Shao-Chang Miao, Chih-Sheng Chen, Chia-Yee Liu	35
4. Design of Pneumatic Robotic Arm and Dynamic Model of Pneumatic Cylinder Actuators	
..... Ming-Chiuan Shiu, Chien-Ming Chen, Fang-Jr Liou	45
5. Nonlinear rupture theory of a thin micropolar liquid film on a cylinder under a magnetic field	
..... Hung Ming Sung, Chung-Ting Hsu	63
6. The Study on Female Consumer Stereotypes towards Bank's Loan Business : Take Miaoli County as Research Sample	
..... Hung-Yu Chen	81
7. Distributed Hashing Mechanisms for On-demand Routing Protocol in Wireless Ad Hoc Networks	
..... Chao-Tsun Chang	105
8. Effects of annealing temperature of TiO₂ thin films for application in dye-sensitized solar cells	
..... Ming-Cheng Kao	125
9. Intelligent Safety Vehicle Anti-collision System	
..... Yi-Ting Mai, Jeng-Yueng Chen, I-Shen Chen, Da-Wei Yang, Zi-Yuan Gao, Jia-Wei Chi	135

10.Synthesis and Characterization of One-Dimensional Indium Zinc Oxide	
Structure by Chemical Vapor Deposition	Chih-Jung Chen 153
11.Innovation dynamics : decision making in changeable spaces based on	
habitual domains	Yen-Chu Chen 163

移動機器人融合主動式 RFID 和 雷射掃描器之定位

林鴻興、蔡清池

摘要

本文目的是發展出整合主動式無線射頻辨識系統 (RFID)、里程計和雷射掃描器等感測系統的室內移動機器人 (AMR) 全域姿態定位方法。

首先探討主動式定位系統是從選擇的電子標籤到讀取器之間所接收到發射訊號強度數據轉換的距離，利用即時最小平方演算法，可估測出機器人在起動時的未知位置和方位角姿態初值。最後在任何環境中，主動式無線射頻辨識系統利用即時最小平方方法，並融合雷射掃描器使用模糊增廣訊息濾波策略，可以估測移動機器人全域姿態。並透過實驗結果，證實本文所提之定位系統是有效性可行的。

關鍵詞：增廣訊息濾波器、雷射掃描器、定位、無線射頻辨識系統、感測器融合。

Localization of an Autonomous Mobile Robot by Fusing Active RFID and Ranging Laser Scanner

Hung-Hsing Lin, Ching-Chih Tsai

Abstract

The purpose of this paper aims at developing a global pose localization method for an autonomous mobile robot (AMR) in indoor environments by fusing measurements including an active radio frequency identification device (RFID), one odometric device, and a ranging laser scanner.

First, to explore the distance from transpose of signal intensity between receive and transmit from selected tags of a reader. Use a real-time least-square method to estimate the unknown position and orientation when the mobile robot start-up. Last, use fuzzy extended information filtering (FEIF) scheme of RFID with real-time least-sequence method fuse laser scanner to estimate global pose of a mobile robot moving in any circumstance. Experimental results are provided to show the effectiveness and merits of the proposed localization systems.

Keywords: Extended Information Filter, Laser Scanner, Localization, RFID, Sensor Fusion.

一、緒論

移動機器人的一個基本功能是可以自我確定在環境中的定位。移動機器人已經被廣用來指導參觀者在擁擠的環境中服務，特別是博物館或商業的展覽會場。在規畫預定的環境從事導覽定位的任務[1]中，移動機器人的定位能力有兩個主要問題是關於初值靜態定位和移動中的姿態定位技術，以及在未知的環境(地圖)中，如何尋找最初的定位。使用內部和外部感測器來考慮初值化姿態和保持機器人的連續即時姿態對導航機器人是很重要的，因為不僅在起動中的初值定位問題，而且在移動期間中，倘若當機器人定位一旦發生分歧時，如何給予機器人作適時的定位補償[2]。

定位是利用先驗環境地圖資訊、機器人初值姿態估計以及感測器的觀測值等輸入資訊，經過一定的處理和變換，產生更加準確地對機器人姿態的估計。目前，室內移動機器人大多採用機載感測器定位。它主要由移動姿態、地圖獲取組成。應用感測器感知的資訊，實現可靠的定位是當前移動機器人研究中倍受關注、富有挑戰性的一個重要研究主題。使用外部感測器(例如：里程計、攝影機、紅外線、微波雷達、陀螺儀、雷射掃描器、超音波感測器、和RFID等等)和配合所設計的

定位演算法，可以量測到有關的定位問題。雖然定位的研究領域很寬廣，在定位方面的文獻中，相關研究曾在[2-6]中被提出。儘管移動機器人定位問題已經研究多年，但現有的各種方法仍存在種種不足，在解決定位問題中，都會遇到不確定性的問題，各種方法都有其優缺點，對於主動式訊號的研究中，作者[3]曾提出較低成本的超音波定位系統(使用兩個超音波發射器和3個接收器系統有關機器人的自我定位系統)。在文獻[2]所提出的定位方法中，雖然被應用於移動服務型機器人，但頗為複雜且需要耗費許多時間方可獲得機器人的最初姿態定位；而超音波定位系統在[3]中，乃探討有關初值的姿態問題，準確度方面卻是有需要再改進的缺點。

無線射頻辨識系統是非接觸式利用無線電的識別系統，透過附著在人身上或物品上面的ID身分識別電子標籤(Tag：由IC晶片、天線與包材所構成)和讀取器(Reader)進行通訊，以無線方式進行必要資訊的交換。換言之，也就是利用無線電波來傳送識別資料，以達到身份識別的目的。主動式電子標籤裡因為裝置了電池，能將標籤的訊號發射到100m、甚至更遠，一般應用主要為10~30m的感應距離，圖1為一般主動式的RFID系統。

近年來使用雷射掃描器已有研究出

機器人定位的相關問題。Dubrawski *et al.* [7]提出在環境裡，利用一台雷射掃描器尋求機器人的移動，探討位置和方位角的變化。Jensfelt 和Christensen [8]曾提出 2-D雷射掃描器使用在minimalistic環境模型中，以卡門濾波器為基礎達到姿態的研究方法。但是，前述所使用卡門濾波器的定位通常會受到融合多面牆時，測量技術會遭遇到困難。

其次，透過增廣卡門濾波器（EKF）[9]方法廣泛應用在機器人姿態過程中，但由於在過程中會產生噪音變化以及模型化時所產生的訊息錯誤，已經被證實會產生濾波器發生分歧的問題。為了考慮這些不利的因素，最近訊息濾波器（IF）已被用在強而有力的濾波技術，雖然訊息濾波器代數模式相當於卡門濾波器（KF），但是前者擁有較多的優點，例如有低計算的複雜性，更容易作到分散式和初值化的處理，也常被應用於多重感測器的融合技術。但是，如同增廣卡門濾波器演算法一樣，由於噪音變化或者模型化時所產生的錯誤，增廣訊息濾波方法還是多少會受到濾波器分歧問題的影響。為了要防止增廣訊息濾波器（EIF）演算法免受到分歧，利用模糊邏輯（Fuzzy logics）[10]來處理改善噪音和測量協方差兩個矩陣的相關問題。作者[3]曾提出模糊增廣訊息演算法中，由模糊調諧器（Fuzzy tuner）組

成可以自動調整指數權重的濾波器來處理噪音和測量協方差兩個矩陣所產生分歧的問題。因此本研究目的是先利用無線射頻辨識系統（RFID）的標籤校正與使用最小平方方法估測出移動機器人初值定位技術，並融合雷射掃描器使用模糊增廣訊息濾波器演算法為基礎的量測技術，發展機器人即時移動姿態技術，透過所提議的演算法，不僅很容易為機器人作最初姿態定位，而且可使用在移動中的機器人定位估測，為導航提供一個更簡單和有效率定位系統。

二、融合雷射掃描器與里程計之姿態

本系統首先利用雷射掃描器所掃描到環境的資料與所建環境模型來做比對，並透過檢驗門檻（Validation Gate）和最小平方線段擬合（Least Square Line Fitting）此兩種方法，來擷取雜亂的模型參數。圖 2 為所提議的以模糊增廣訊息濾波為基礎的姿態追蹤演算法方塊圖，其中 M 代表為環境模型。

2.1 環境模型

如圖 3 所示，設 m_i 為一矩行的環境模型， M 是由若干個牆所組成的，而每個牆是以 $(\rho_i^m, \alpha_i^m, l_i^m)$ 三個參數所表示。其中 ρ_i^m 表示為原點到牆的垂直距

離， α_i^m 為原點到牆之間的夾角， l_i^m 為牆的長度。

2.2 牆（線）的擷取

其次有關擷取牆（線）的方法，包括二部份為檢驗門檻（Validation Gate）和最小平方線段擬合（Least Square Line Fitting）方法，如圖 2 所示。（1）檢驗門檻（Validation Gate）如圖 4 所示， G 以 4 個參數 $G = (\hat{\rho}, \hat{\alpha}, \delta, \gamma)$ 表示，其中 $\hat{\rho}$ 為導航位置到預估牆的垂直距離， Y^R 是機器人前進的方向，且 $\hat{\alpha}$ 定義為機器人前進的夾角， δ 是檢驗門檻的寬度， γ 是檢驗門檻的開角。檢驗門檻最主要的目的是將雷射所掃到的資料中所屬於雜訊的部份先行濾掉。（2）使用最小平方線段擬合方法，最主要的目的是要擷取在環境模型中，每個牆相對於機器人垂直距離 ρ 和相關的角度 α 。將以上所提的方法，在已知環境模型內的牆，可以經由線的擷取演算法，容易地被引出 [11]。

2.3 模糊增廣訊息濾波演算法（FEIF）

這個部分要闡明模糊增廣訊息濾波（FEIF）的方法技術。以下內容為非線性離散時間的系統模型和測量模型：

$$X(k+1) = f(X(k), U(k), k) + W(k) \quad (1)$$

$$Z(k) = h(X(k), k) + V(k) \quad (2)$$

其中 $f(\cdot)$ 和 $h(\cdot)$ 是狀態 $X(k)$ 和輸入 $U(k)$ 的非線性函數及二次微分。而 $W(k) \sim N(0, Q(k))$ 和 $V(k) \sim N(0, R(k))$ 分別是有關雜訊過程和測量雜訊矩陣。為了獲得最佳的姿態估測和移動機器人的追蹤，本研究所提的模糊增廣訊息濾波演算法中，有關雜訊過程和測量雜訊協方差矩陣說明如下：

$$R(k) = R \cdot \alpha^{-\beta(k+1)}, Q(k) = Q \cdot \alpha^{-\beta(k+1)} \quad (3)$$

式中 α 是很接近於 1 的權重因子， β 是指數函數變化的程度。以下為 FEIF 演算法預估和估測方程式說明：

(i) 一步的預估

$$\hat{y}_\alpha(k|k-1) = Y_\alpha(k|k-1) \times f(\hat{X}(k-1|k-1), U(k-1), (k-1)) \quad (4)$$

$$Y_\alpha(k|k-1) = [\alpha^\beta \nabla f_X(k-1) \times Y_\alpha^{-1}(k-1|k-1) \times \nabla f_X^T(k-1) + Q]^{-1} \quad (5)$$

(ii) 估測（測量更新）

$$\hat{y}_\alpha(k|k) = \hat{y}_\alpha(k|k-1) + \nabla h_X^T(k) \cdot \left(\frac{R}{\alpha^\beta}\right)^{-1} \cdot [r(k) + \nabla h_X(k) \cdot \hat{X}(k|k-1)] \quad (6)$$

$$Y_\alpha(k|k) = Y_\alpha(k|k-1) + \nabla h_X^T(k) \cdot \left(\frac{R}{\alpha^\beta}\right)^{-1} \cdot \nabla h_X(k) \quad (7)$$

$$\text{其中 } r(k) = Z(k) - h(\hat{X}(k|k-1)) \quad (8)$$

$$\nabla f_X(k-1) = \left. \frac{\partial f}{\partial X} \right|_{X=\hat{X}(k-1|k-1)}$$

$$\nabla h_x(k) = \frac{\partial h}{\partial X} \Big|_{X=\hat{X}(k|k-1)}$$

注意如果當 $\alpha=1$ 時，則模糊增廣訊息濾波器就變為一般的增廣訊息濾波器。

利用模糊增廣濾波器訊息是有必要的，而濾波器的處理是由模糊調諧設備組成的自動調整指數權重和雜訊協方差兩矩陣測量。所提議利用模糊調諧設備來監控更新，利用 FEIF 演算法在 (6-8) 式中，可以調 α 值來避免濾波器分歧所產生的問題。模糊的調諧器設計是由三個輸入和一個輸出所組成；平均值 (mean)、二級動差值 (Second-order Moment) 和二級動差值的變化率 (Second-order Mmoment Slew Rate of The Innovation) 用來監控濾波器的分歧。透過選擇一定取樣 l ，模糊調諧器的三個輸入則描述如下：

統計平均值 (Mean)：

$$\bar{r} \cong \frac{1}{l} \sum_{k=1}^l r(k) \quad (9)$$

統計更新二級動差值 (Second-order Moment)：

$$\sigma_r^2 \cong \frac{1}{l} \sum_{k=1}^l r^2(k) \quad (10)$$

統計更新二級動差值的變化率 (Second-order Moment Slew Rate of The Innovation)：

$$\text{變化率 (Slew rate)} = \sigma_r^2(k) - \sigma_r^2(k-k_1) \quad (11)$$

其中參數 k_1 是正值並且可由設計者選擇，其中參數的選擇設計是 $k_1 = 50$ 和 $l = 500$ 。隸屬函數則透過遺傳演算法 (Genetic Algorithms)，輸入變數由三角形的函數定義分別如圖 5-7 所示，模糊調諧器的輸出權重因子 α 值如圖 8 所示。對於調諧器的設計想法，詳細說明在 FEIF 演算法裡，作者曾發表在 [11] 中。在表 1-3 中總共使用 45 條模糊的規則，其中 NEGATIVE LARGE (NL)、NEGATIVE MEDIUM (NM)、SMALL (S)、MEDIUM (M)、LARGE (L)、POSITIVE、ZERO 和 NEGATIVE 分別代表所對應的模糊語言集合。提議的模糊調諧器是採取 Mamdani-style 的推斷方法和 Max-min 方法來解模糊化。以人類的直覺和經驗判斷，調諧器的設計共使用 45 條模糊的規則，以下簡述其中舉例，說明如下：

如果平均值 \bar{r} 是小的，動差 σ_r^2 是大的，且變化率是正的，則 α 是適中。

三、應用無線射頻辨識系統 (RFID) 之姿態追蹤

3.1 電子標籤 (Tag) 的校正方法

RFID 定位是利用 RSSI 值來分析 Tag 和 Reader 之間的距離，雖然 RSSI 值會因為某些因素干擾，如多路徑影響、環境溼

度、溫度、穿透衰減...等產生 RSSI 浮動，故需要進行電子標籤校正，但可以確定的是 RSSI 值在某個特定環境下，和距離會產生某種關係。利用波傳導理論和實驗經驗，所求得的 RSSI-Distance 特性曲線後，利用此校正曲線，轉換成相對應距離後，再使用最小平方法求得機器人姿態座標。對於 m 個電子標籤而言，RSSI 值和相對應距離的數學關係式可表示為

$$\text{RSSI}_i = \frac{K}{d_i^n}, \quad i = 1, \dots, m \quad (12)$$

其中 d_i 為距離，參數 K 和 n 為校正參數。由每個電子標籤實驗 n 公尺的 RSSI-Distance 關係，可以得到下列方程式：

$$\begin{bmatrix} 1 & -\log_{10} d_1 \\ 1 & -\log_{10} d_2 \\ \vdots & \\ 1 & -\log_{10} d_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \log_{10} \text{RSSI}_1 \\ \log_{10} \text{RSSI}_2 \\ \vdots \\ \log_{10} \text{RSSI}_n \end{bmatrix}$$

$$\Rightarrow \mathbf{JX} = \mathbf{K} \quad (13)$$

其中參數 a 、 b 可經由最小平方法求得

$$\mathbf{X} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{K} = \begin{bmatrix} a \\ b \end{bmatrix} \quad (14)$$

其中 $K=10^a$, $n=b$ 。

3.2 最小平方定位法

由 3.1 所得到的校正曲線，可以求出一個別 Tag 和 Reader 的相對應距離，設定 m 個電子標籤， $\mathbf{T}_1 = (x_1, y_1, z_1)^T$ ， $\mathbf{T}_2 = (x_2, y_2, z_2)^T, \dots, \mathbf{T}_m = (x_m, y_m, z_m)^T$ ， $d_i = (K_i / \overline{\text{RSSI}_i})^{1/n_i}$, $i = 1, \dots, m$ 轉換每個電子標籤的 RSSI 值成為相對應距離。由圖 9 所示，可以得到下列聯立方程式：

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 = d_1^2 \\ (x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 = d_2^2 \\ \vdots \\ (x_m - x)^2 + (y_m - y)^2 + (z_m - z)^2 = d_m^2 \end{cases} \quad (15)$$

經由化簡，可得下列方程式：

$$\mathbf{AX} = \mathbf{B}$$

其中

$$\mathbf{A} = \begin{bmatrix} 2(x_1 - x_2) & 2(y_1 - y_2) & 2(z_1 - z_2) \\ \vdots & & \\ 2(x_1 - x_i) & 2(y_1 - y_i) & 2(z_1 - z_i) \\ \vdots & & \\ 2(x_1 - x_m) & 2(y_1 - y_m) & 2(z_1 - z_m) \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} d_2^2 - d_1^2 + (x_1^2 - x_2^2) + (y_1^2 - y_2^2) + (z_1^2 - z_2^2) \\ \vdots \\ d_i^2 - d_1^2 + (x_1^2 - x_i^2) + (y_1^2 - y_i^2) + (z_1^2 - z_i^2) \\ \vdots \\ d_m^2 - d_1^2 + (x_1^2 - x_m^2) + (y_1^2 - y_m^2) + (z_1^2 - z_m^2) \end{bmatrix}$$

由最小平方法解矩陣方程式：

$$\mathbf{X} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B} \quad (16)$$

\mathbf{X} 即為所求機器人位置座標 $(x, y, z)^T$ 。

3.3 機器人方位的計算

在室內定位中角度問題或許不用在意，但對放置於展覽場的機器人來說，方位是很重要的。譬如當機器人要走到下一個會場時，如果只知道自己的位置卻不知道方位，有可能立刻就喪失定位的功能。經 3.2 所求的機器人位置座標 $\mathbf{X} = (x, y, z)^T$ 外，當機器人前進一小段距離時，可由兩點位置決定一直線方程式如 (17) 式。

$$y_i = mX_i + c \quad (17)$$

其中 m 為直線斜率、 c 為截距。

假設從 (16) 式求得多點機器人座標，(17) 式可寫成下列矩陣形式。

$$\begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} c \\ m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \Rightarrow \mathbf{CY} = \mathbf{D},$$

$$\mathbf{Y} = \begin{bmatrix} c \\ m \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}, \mathbf{D} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (18)$$

再利用最小平方方法求解可得

$$\begin{bmatrix} c \\ m \end{bmatrix} = \mathbf{Y} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{D} \quad (19)$$

斜率及截距即可從 (19) 式求得，而 $m = \tan \theta$ ，其中 θ 為機器人方向和 X 軸的夾角，即可計算 $\theta = \arctan(m)$ 。故利用 (16)、(19) 式至少使用四個電子標籤方可求得機器人的姿態座標 $\mathbf{X} = (x, y, z, \theta)^T$ 。

3.4 機器人初值定位

當機器人在 RFID 有效訊號幅射的範圍內，在全域架構下作自我定位或校正當時的姿態時，利用即時初值定位演算法，機器人可以正確找到起動時的初值定位，演算法步驟說明如下：

步驟 1: 測量到全部電子標籤 (T_1, \dots, T_l)

和讀取器 (R_1) 之間的 $RSSI_1(j), \dots, RSSI_l(j), j=1, 2, 3, \dots, n$ ，全部 n 組的 $RSSI$ 數據。

步驟 2: 測量到全部電子標籤 (T_1, \dots, T_l)

和讀取器 (R_1) 之間的 n 組 $RSSI$ 數據後，利用下列公式求平均值：

$$\overline{RSSI}_i = \frac{1}{n} \sum_{j=1}^n RSSI_i(j), i=1, \dots, l.$$

步驟 3: 使用 (12) 式校驗的參數 K_i 和

m_i ，可以把步驟 2 求得到的 $RSSI$ 測量平均值轉變成所相對的距離：

$$d_i = (K_i / \overline{RSSI}_i)^{1/m_i}, i=1, \dots, n.$$

(20)

步驟 4：利用最小平方法 (16) 式，計算出機器人起動時的估測位置 $(\hat{x}_0, \hat{y}_0, \hat{z}_0)$ 。

步驟 5：讓機器人沿著直線移動 Δd 距離時，利用 (16) 式求得機器人下一步的位置 $(\hat{x}_1, \hat{y}_1, \hat{z}_1)$ 時，然後透過下式計算出機器人起動時的方位：

$$\hat{\theta}_0 = A \tan 2(\hat{y}_1 - \hat{y}_0, \hat{x}_1 - \hat{x}_0) \quad (21)$$

式中 $\hat{x}_1 - \hat{x}_0 = v \cos \hat{\theta}_0$ 、
 $\hat{y}_1 - \hat{y}_0 = v \sin \hat{\theta}_0$ 以及 v 為機器人的線性速度。

四、融合 RFID、里程計和雷射掃描器即時姿態跟蹤定位演算法

主動式的無線射頻辨識系統利用最小平方法原理，可求得機器人的姿態估測座標。同時在增廣訊息濾波技術的框架下，針對定位而言，從機器人當時的姿態和環境模型，把牆壁的參數計算引出是必要的。假設從雷射掃描一個整體週期的數據資料，可獲得 s 面所擷取到牆的資料，而整體系統的測量方程式如 (22) 式所示：

$$z(k) = \bar{h}(X(k), M) + V(k) \quad (22)$$

式中

$$z(k) = [d_1 \ d_2 \ \dots \ d_m \ \rho_1 \ \alpha_1 \ \rho_2 \ \alpha_2 \ \dots \ \rho_s \ \alpha_s]^T$$

$$X(k) = [x(k) \ y(k) \ \theta(k)]^T$$

$$V(k) = [v_1 \ v_2 \ \dots \ v_m \ w_1(k) \ w_2(k) \ \dots \ w_{2s}(k)]^T$$

$V(k)$ 是 Zero-mean 量測雜訊的變異差矩陣 $R(k) = \text{diag}\{\delta_1^2, \delta_2^2, \dots, \delta_s^2\} \cdot \alpha^{-\beta(k+1)}$ 。經線性化後測量函數 $\bar{h}(X, M)$ 表示如下：

$$\bar{h}(X, M) = [D_1 \ D_2 \ \dots \ D_m \ P_1 \ Q_1 \ P_2 \ Q_2 \ \dots \ P_s \ Q_s]^T \quad (23)$$

式中

$$D_j = \sqrt{(x_j - x)^2 + (y_j - y)^2 + (z_j - z)^2},$$

$$j = 1, \dots, m.$$

$$P_i = \rho_i^m - \sqrt{x^2 + y^2} \cdot \cos(\alpha_i^m - \psi),$$

$$Q_i = \alpha_i^m - \theta(k), \quad i = 1, 2, 3, \dots, s,$$

ρ_i^m 是表示到第 i 面牆壁的距離以及 α_i^m 是表示所對應的角度。

在已知 AMR 最初姿態，車上使用編碼器、RFID 和雷射掃描器尋求連續姿態。如 RFID 利用最小平方法融合雷射掃描器並結合模糊增廣訊息濾波方法來校正車體位置的估測誤差時，除了 (24) 式外，另外需要描述下列車體的運動狀態方程式：

$$X(k+1) = f(X(k)) + W(k) \quad (24)$$

式中

$$f(X(k)) = \begin{bmatrix} x(k) + \Delta d \cos \theta(k) \\ y(k) + \Delta d \sin \theta(k) \\ \theta(k) + \Delta \theta \end{bmatrix},$$

$$W(k) = \begin{bmatrix} w_x(k) \\ w_y(k) \\ w_\theta(k) \end{bmatrix},$$

$$\Delta \theta = \Delta T \cdot w_c, \quad \Delta d = \Delta T \cdot v_c.$$

式中 $w_x(k)$ 和 $w_y(k)$ 分別表示位置推算器的位置誤差；和 $w_\theta(k)$ 表示地面不平的隨機干擾量。使用離散時間的模糊增廣訊息濾波器來估算移動機器人的姿態，融合演算法其預估、估測方程式和步驟說明如下：

步驟 1：初始化。

當時間 $k=0$ 時，選 $\beta = -2$ ，位置之初始訊息狀態估測值為 $\hat{y}_\alpha(0|0)$ ，以及訊息矩陣 $Y_\alpha(0|0)$ ，訊息狀態的初值 $\hat{y}_\alpha(0|0) = Y_\alpha(0|0) \cdot \hat{X}(0|0)$ 其中 $\hat{X}(0|0) = [\hat{x}(0|0), \hat{y}(0|0), \hat{\theta}(0|0)]^T$ 。

步驟 2：一步的預估。

時間為 k 時之最佳的估測 $Y_\alpha(k)$ 和訊息矩陣是 $Y_\alpha(k+1|k)$ 。由 (25-26) 式之計算，可得到最佳預測值為 $\hat{y}_\alpha(k+1|k)$ ，以及訊息狀態矩陣 $Y_\alpha(k+1|k)$ 。

$$\hat{y}_\alpha(k|k-1) = Y_\alpha(k|k-1) \times f(\hat{X}(k-1|k-1), U(k-1), (k-1)) \quad (25)$$

$$Y_\alpha(k|k-1) = [\alpha^\beta \nabla f_x(k-1) Y_\alpha^{-1}(k-1|k-1) \times \nabla f_x^T(k-1) + Q]^{-1} \quad (26)$$

式中 Jacobian 矩陣

$$\begin{aligned} \nabla f_x(k) &= \left. \frac{\partial f}{\partial X} \right|_{X=\hat{X}(k|k)} \\ &= \begin{bmatrix} 1 & 0 & -\Delta d \sin \theta \\ 0 & 1 & \Delta d \cos \theta \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

步驟 3：當定位控制器量測到一個新的數值 $Z_i(k+1)$ 後，計算更新值 $r_i(k+1)$ 利用 (27) 式，並經由模糊調諧器的演算法，可以產生一個適當權重因子 α 值。

步驟 4：估測（測量更新）。

在時間為 k 時，系統定位控制器讀到新的測量資料 $Z_i(k+1)$ 和使用 (6-7) 式計算得到更新的訊息狀態估測值 $\hat{y}_\alpha(k+1|k+1)$ 以及訊息矩陣 $Y_\alpha(k+1|k+1)$ 。其中更新值 $r_i(k+1)$ 如下式所示：

$$r_i(k+1) = Z_i(k+1) - \bar{h}_i(\hat{X}(k+1|k), M) \quad (27)$$

而且

$$\hat{X}(k+1|k) = Y_\alpha(k+1|k)^{-1} \cdot \hat{y}_\alpha(k+1|k)$$

其中 Jacobian 矩陣 $\nabla \bar{h}_X(k+1)$ 為

$$\nabla \bar{h}_x(k+1) = \begin{bmatrix} \frac{x-x_1}{D_1} & \frac{y-y_1}{D_1} & \frac{z-z_1}{D_1} \\ \dots & \dots & \dots \\ \frac{x-x_m}{D_m} & \frac{y-y_m}{D_m} & \frac{z-z_m}{D_m} \\ H_1^1 & H_1^2 & 0 \\ 0 & 0 & -1 \\ H_1^2 & H_1^2 & 0 \\ 0 & 0 & -1 \\ & \vdots & \\ H_1^s & H_1^s & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (28)$$

式中

$$H_1^i = -\frac{x}{r} \cdot \cos(\alpha_i^m - \psi) + \frac{y}{r} \cdot \sin(\alpha_i^m - \psi)$$

$$H_1^i = -\frac{y}{r} \cdot \cos(\alpha_i^m - \psi) - \frac{x}{r} \cdot \sin(\alpha_i^m - \psi) \quad (29)$$

$$i = 1, 2, 3, \dots, s.$$

$$r = \sqrt{x^2 + y^2} \quad \text{和} \quad \psi = \arctan\left(\frac{y}{x}\right).$$

步驟 5：重複步驟 2 到步驟 4。

當車體到達所設定的終點時，就停止執行。

五、實驗結果與討論

5.1 移動機器人

定位是機器人完成導航之相當關鍵技術，用以隨時知悉該車體相對於展示空

間世界座標系之姿態（含位置與車頭方向），進而完成定位任務。移動機器人定位研究早已行之多年，尤其是使用 RFID、雷射掃描器與全域（Global）定位技術研究論文，更是不勝枚舉。本研究擬將數個（三個以上）電子標籤置於展示空間的天花板（牆）上，將 RFID 的讀取器（Reader）架在車體上，藉由校正，透過接收訊號強度（RSSI）轉成距離，再使用最小平方方法將分別的距離資料，換算成概略的機器人全域位置值，並融合雷射掃描器，運用模糊增廣訊息濾波器（FEIF）完成精確的機器人定位，機器人平台機構如圖 10 所示。

5.2 融合 RFID，里程計與雷射掃描器之即時姿態跟蹤定位實驗

本研究應用主動式的無線射頻辨識系統，融合以矩形模型 2D 雷射掃描器的外部感測器，研發機器人定位方法與技術，同時也可以解決行動平台姿態起始值、姿態估測以及同時定位等相關問題。

實驗流程分為電子標籤校正和定位程序兩方面如圖 11 所示，電子標籤校正程序是為了解決每個 Tag 因製程與使用環境不同（溫度、溼度、訊號干擾...等）所產生的個別差異，在不同環境中都必須對每個電子標籤進行校正，建立精確的

RSSI-Distance 特性曲線。定位程序必須先設定 4 個電子標籤的座標，架在機器人上的 Reader 即時接收到 RSSI 值，依電子標籤校正程序中所建立的 RSSI-Distance 特性曲線轉換成相對應距離，再使用最小平方法可求得機器人即時姿態座標。

實驗 1：將四個電子標籤分別放在牆上位置：

$$(x_1, y_1, z_1) = (47\text{cm}, 544\text{cm}, 300\text{cm}),$$

$$(x_2, y_2, z_2) = (400\text{ cm}, 680\text{ cm}, 299\text{ cm}),$$

$$(x_3, y_3, z_3) = (473\text{cm}, 313\text{cm}, 299\text{ cm}),$$

$$\text{和 } (x_4, y_4, z_4) = (595\text{cm}, 2\text{cm}, 300\text{cm}),$$

機器人的真實位置於 (117.5 cm, 160 cm) 及方位角是 90° 。實驗前四個電子標籤都經過校準程序如圖 12 所示；其中紅色曲線代表從 0 到 10 公尺中，每增加 1 公尺所對應的 RSSI 量測值。藍色曲線是利用校正公式計算所得到的特性曲線。如圖 13 所示，機器人靜態實驗的估測值為 (128.43 cm, 171.06 cm, 74.36°)，圖中 O 符號是代表機器人的實際位置，X 符號是代表機器人利用最小平方法的估測值位置，□ 符號是代表電子標籤所放置的位置。本實驗所提出的 RFID 全域定位方法中，位置誤差少於 20 cm 和方位角少於 25° ，實驗結果證明有效地可被運用在機器人的定位。

實驗 2：當機器人在平坦環境進行直線

移動時，將所提議 2D 雷射掃描器估測演算法做性能評估。在實驗過程中，移動距離是 508 cm，車體速度是 $v_c = 5\text{ cm/sec}$ 。實驗環境是由 10 面平的牆所圍成如圖 14 (a) 所示，大約是 $760\text{ cm} \times 330\text{ cm}$ 。環境模型中參數分別是：

$$(\rho_1^m, \alpha_1^m) = (0\text{ cm}, 0^\circ), (\rho_2^m, \alpha_2^m) = (0\text{ cm},$$

$$90^\circ), (\rho_3^m, \alpha_3^m) = (145\text{ cm}, 0^\circ), (\rho_4^m, \alpha_4^m) =$$

$$(34\text{cm}, 90^\circ), (\rho_5^m, \alpha_5^m) = (234\text{cm}, 0^\circ), (\rho_6^m,$$

$$\alpha_6^m) = (364\text{cm}, 90^\circ), (\rho_7^m, \alpha_7^m) = (334\text{cm},$$

$$0^\circ), (\rho_8^m, \alpha_8^m) = (441\text{ cm}, 90^\circ), (\rho_9^m, \alpha_9^m) =$$

$$(234\text{ cm}, 0^\circ) \text{ 和 } (\rho_{10}^m, \alpha_{10}^m) = (760\text{ cm}, 90^\circ)$$

。考慮利用 FEIF 初始化時，其中有關實驗參數和誤差協方差距陣分別為 $Y(0/0) = \text{diag}\{1/20, 1/20, 1/5\}$, $R = \text{diag}\{0.0000025, 0.0000025, 0.0000025, 0.0000025, 0.0000025, 0.0000025, 0.0000025, 0.0000025, 0.0000025, 0.0000025\}$

和 $Q = \text{diag}\{0.001, 0.001, 0.001\}$ ；以及機器人從起點在 (117.5cm, 150cm) 位置出發，並沿 90° 度的方位角移動行走走到終點 (117.5cm, 658cm) 位置上停止。實驗過程的取樣期間 ΔT 是 0.4 秒。機器人分別以 EIF 和 FEIF 為不同演算法基礎量測的 x-y 軸姿態估測軌跡圖，其中的初值是取用實驗 1 靜態定位的結果，並選擇檢驗門檻參數： $\delta = 2\text{ cm}$ 和 $\gamma = 2^\circ$ 。圖 14 (b) 是車頭角度估測軌跡圖。從圖 14 觀察了解，儘管在初值尚存有誤差的情形，但機

器人估測性能迅速收斂於真實值。由實驗圖中比較，以 FEIF 為基礎的姿態方法，有明顯降低誤差並獲得很好的姿態估計值 $(\hat{x}, \hat{y}, \hat{\theta}) = (117.6643 \text{ cm}, 657.7315 \text{ cm}, 89.8712^\circ)$ 和穩定誤差是 $(0.1643 \text{ cm}, 0.2685 \text{ cm}, 0.1288^\circ)$ 。

從實驗結果足以證明以 FEIF 為基礎的連續移動姿態定位演算法，有令人滿意的定位性能。

六、結論

本研究提出移動機器人融合主動式的無線射頻辨識系統和雷射掃描器的定位，先把電子標籤安裝在已知不同牆上位置，利用 RSSI-Distance 特性曲線轉換量測到的 RSSI 值到電子標籤相互所對應距離，使用最小平方法估測出機器人最初姿態，再融合里程計和雷射掃描器等感測，利用模糊增廣訊息濾波器的演算法，量測機器人移動時的即時定位姿態估測。值得一提的是當機器人在 RF 波傳播的有效範圍內，能使用所提議方法即時作校正定位。由實驗結果證實機器人的初值化定位和以低速在平坦地形上連續移動姿態的方法和技術是有效的。

七、參考文獻：

- [1] B. Graf, M. Hans, and R. D. Schraft, "Mobile robot assistants," *IEEE Robotics and Automation Magazine*, vol. 11, no.2, pp.67-77, 2004.
- [2] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, vol.1, New York, Academic Press, 1979.
- [3] H. H. Lin, C. C. Tsai, J. C. Hsu, "Ultrasonic localization and pose tracking of an autonomous mobile robot via fuzzy adaptive extended information filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no.9, pp.2024-2034,2008.
- [4] H. H. Lin, C. C. Tsai, and J. C. Hsu, "Ultrasonic localization and pose tracking of an autonomous mobile robot via fuzzy adaptive extended information filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 9, pp. 2024-2034, 2008.
- [5] S. Se, D. G. Lowe, and J. J. Little, "Vision-based global localization and mapping for mobile robots," *IEEE Transactions Robotics and Automation*, vol. 21, no. 3, pp. 364-375, 2005.
- [6] David C. K. Yuen, Bruce A. MacDonald, "Vision-based localization algorithm based on landmark matching, triangulation, reconstruction, and comparison," *IEEE Transactions on robotics*, vol. 21, no. 2-3, pp. 217-221, 2005.
- [7] A. Dubrawski, and I. Siemiatkowska, "A method for tracking the pose of a

- mobile robot equipped with a scanning laser range finder,” *IEEE Transactions on Robotics and Automation*, vol.3, no.3, pp. 2518-2523, 1998.
- [8] P. Jensfelt, and H. I. Christensen, “Pose tracking using laser scanning and minimalistic environmental models,” *IEEE Transactions on Robotics and Automation*, vol.17, no.2, pp. 138-147, 2001.
- [9] Q. H. Meng, Y. C. Sun and Z. L. Cao, “Adaptive extended Kalman filter (AEKF) -based mobile robot localization using sonar,” *Robotica*, vol.18, no.5, pp.459-473, 2000.
- [10] L. A. Zadeh, “Fuzzy logic,” *Computer*, vol. 21, no.4, pp.83-93, 1988.
- [11] C. C Tsai, Y. S. Huang, H. H. Lin, Y. C., “Map-based posture tracking of a nursing-care walking assistant using laser scanner,” *Chinese Automatic Control Conference*, Nov.18-19, 2005. pp.2518-2523, 1998.
-

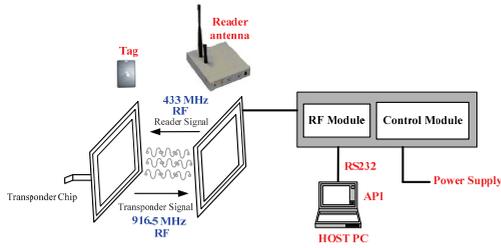


圖 1 基本的主動式 RFID 系統

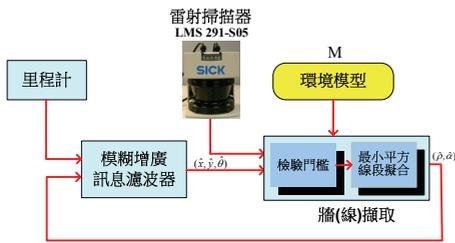


圖 2 以模糊增廣訊息濾波為基礎的姿態追蹤演算法方塊圖

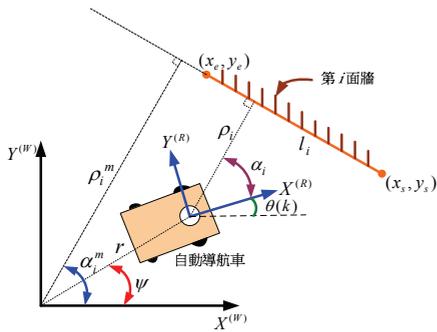


圖 3 模型參量定義圖

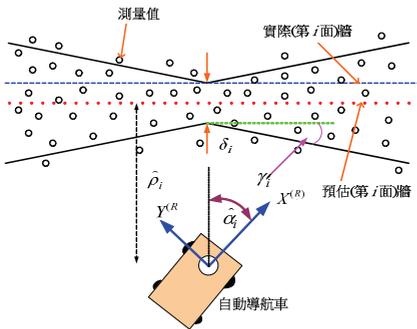


圖 4 檢驗門檻 (Validation Gate)

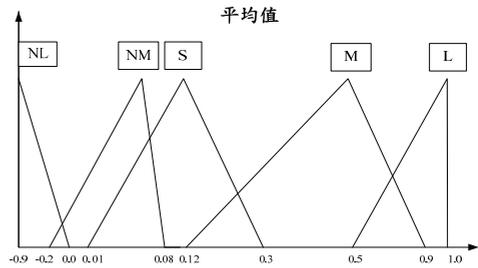


圖 5 模糊輸入平均值的歸屬函數

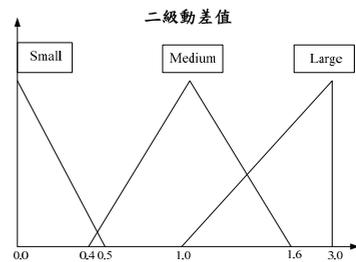


圖 6 模糊輸入更新二級動差值的歸屬函數

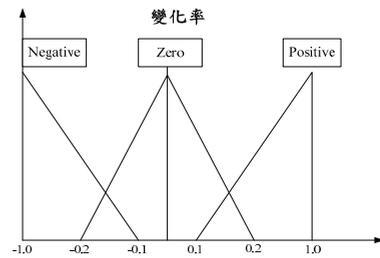


圖 7 模糊輸入更新二級動差值變化率的歸屬函數

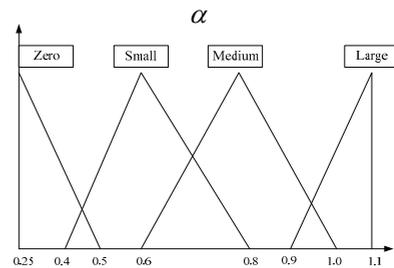


圖 8 模糊輸出變數權重因子 α 值歸屬函數

表 1 正變化率的模糊規則庫

$\sigma_r \backslash \bar{r}$	NL	NM	S	M	L
S	Large	Small	Small	Small	Large
M	Large	Large	Small	Large	Large
L	Large	Large	Medium	Large	Large

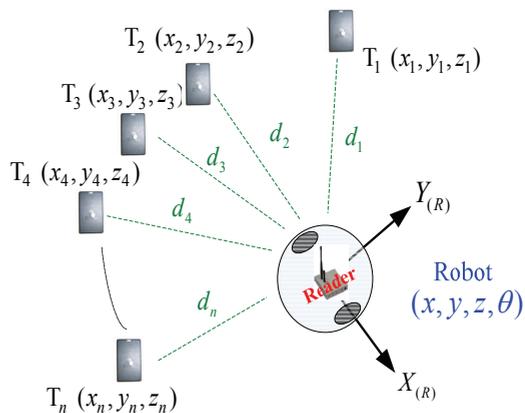


圖 9 機器人 RFID 定位示意圖

表 2 零變化率的模糊規則庫

$\sigma_r \backslash \bar{r}$	NL	NM	S	M	L
S	Large	Small	Zero	Small	Large
M	Large	Medium	Small	Medium	Large
L	Large	Medium	Small	Medium	Large

表 3 負變化率的模糊規則庫

$\sigma_r \backslash \bar{r}$	NL	NM	S	M	L
S	Medium	Small	Small	Small	Medium
M	Large	Large	Small	Large	Large
L	Large	Large	Medium	Large	Large

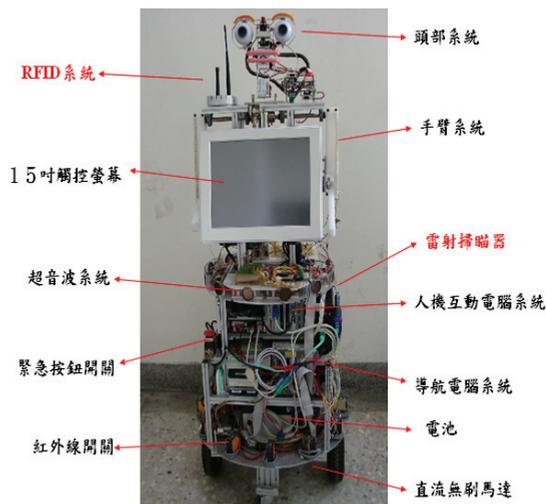


圖 10 移動機器人平台機構圖

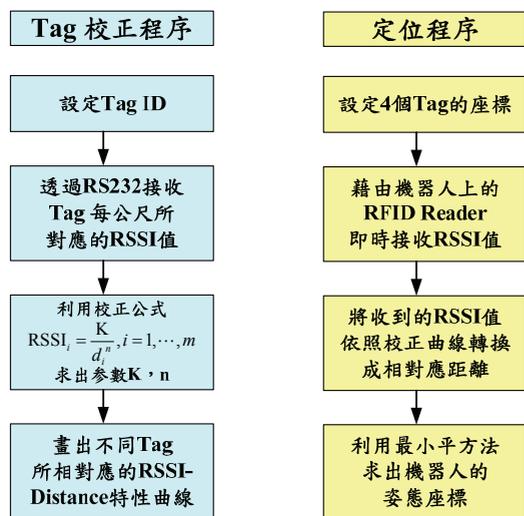
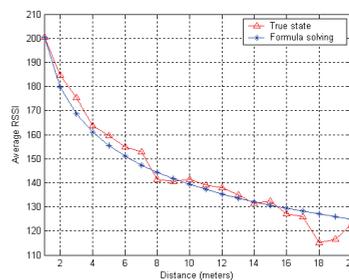
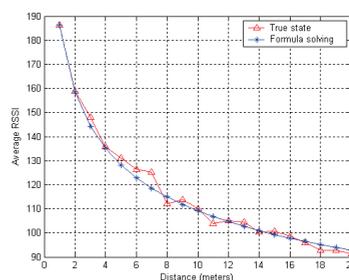


圖 11 電子標籤校正和定位程序流程圖

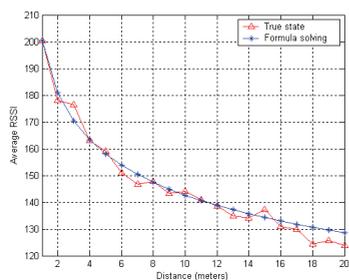


(c)

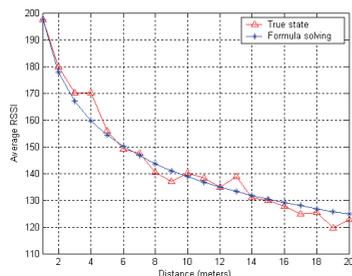


(d)

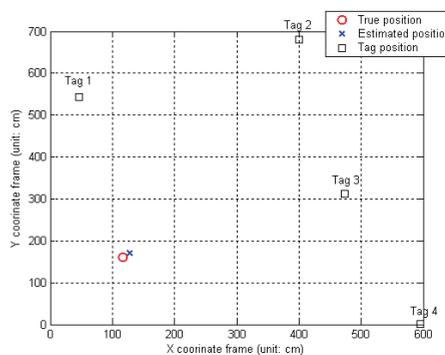
圖 12 (a) Tag 1 (b) Tag 2 (c) Tag 3 (d) Tag 4 RSSI-Distance 特性曲線



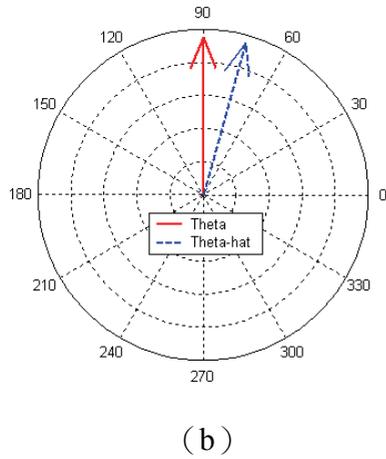
(a)



(b)



(a)



(b)

圖 13 靜態機器人的 (a) 位置和 (b) 方位角

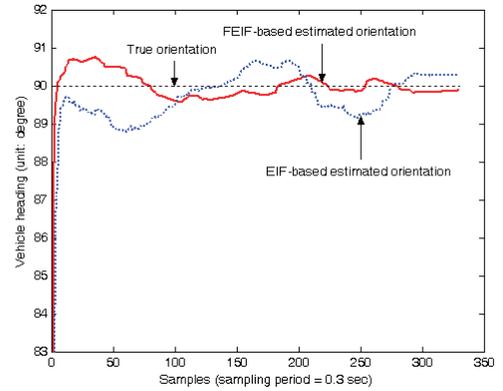


圖 14 (b) 移動車頭角度估測的軌跡圖

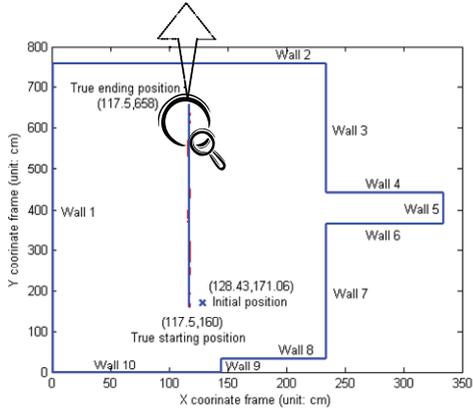
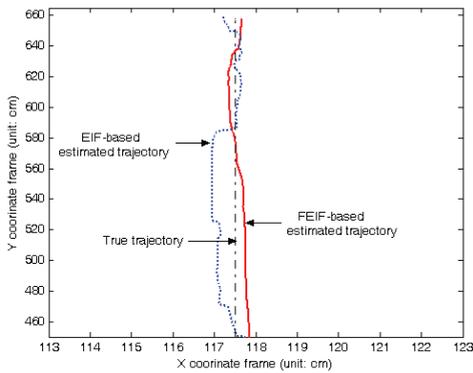


圖 14 (a) 機器人 x-y 軸姿態估測軌跡圖

Finite Deformation of 2-D Thin Circular Curved Laminated Beams

Chiu-Wen Lin

Abstract

An analytical method is derived for obtaining the finite deformation of 2-D thin curved laminated beams. The general solutions are expressed by fundamental geometric quantities. As the radius of curvature is given, the fundamental geometric quantities can be calculated to obtain the closed form solutions of the axial force, shear force, bending moment, rotation angle, and deformed or un-deformed displacement fields. The closed-form solutions of the circular curved laminated beams under pure bending moment case are presented. It shows the consistency of the results of present study with those by ANSYS.

Keywords : Finite deformation, Curved laminated beams, Variable curvatures, Analytical solutions, Non-linear behavior.

二維圓形薄疊層曲樑之有限變形研究

林秋文

摘要

本研究應用解析方法，分析研究二維圓形薄疊層曲樑之有限變形；其一般通解以曲樑之基本幾何特性值表示之。當曲率半徑確定時，則曲樑之基本幾何特性值可以計算出；並藉以求得曲樑之剪力、軸向力、彎矩、旋轉角、變形位移場與未變形位移場等物理量之閉合型式解。本文發表了懸臂圓形薄曲樑承受純彎矩作用之閉合型式解，並與有限元素法套裝分析軟體 ANSYS 分析結果比較；結果非常一致。

關鍵詞：有限變形理論、疊層曲樑、解析解、非線性行為。

1. Introduction

The rod theory is one of the most developed parts of the elasticity theory. The finite deformation of rods in space, are always related nonlinear geometric behavior. There are two approaches which are very common. One is three dimensional rods theory. The rod is treated as a three dimensional elastic body to which the methods of three dimensional elasticity theory are applied (Green and Naghdi [1]). The other is one dimensional director theory. The rod is treated as a curve (Green and Naghdi [2]). Naghdi [3], Green [4] showed the nonlinear behavior of rods in both ways. Green [5] showed some relationship between two approaches.

Due to the complexity of mathematical models, most studies have to adopt some kind of simplification, such as small displacement, small shearing strain, small rotation or small shearing effect. By using finite element method, Li. [6,7] derived a finite deformation theory based on total Lagrangian description for 2-D and 3-D beams of zero Poisson's ratio without all the simplifications. Some studied the finite deformation under dynamic loading. Oguibe [8] investigated the numerical study of the elastic plastic response of multilayer aluminum cantilever beams subjected to an impulse loading. The

numerical results were compared with the experimental results. Attard [9] studied finite strain of an isotropic hyper-elastic Hookean beam. He used an appropriate strain energy density. The shear effect was included. The solution was also applied to stability behavior design of a helical spring. Mauget [10] applied Lagrangian coordinates to derive an isotropic constitutive law for a large displacement formulation of woods. Toi [11] used total Lagrangian approach for the super-elastic large deformation analysis of a shape memory alloy helical springs.

Most studies focus on straight rods. Only few investigate curve rods. Atanackovic [12] analyzed the finite deformation of a circular ring under uniform pressure. Brush [13] derived a finite deformation stability equation for circular ring under various pressures. He also investigated the stability of nonlinear equilibrium equations for fluid pressure loading. Due to the complexity of mathematical models, analytical solutions are very limited. Timoshenko [14] showed the large deformation of an *Elastica*. It also showed the stability of a straight beam of large deformation. In this paper, we apply the tangent slope coordinate theory by Lin [15, 16, 17, 18] and lamination theory by Herakovich [19] to study finite deformation 2-D of curved laminated beams with

circular curvature.

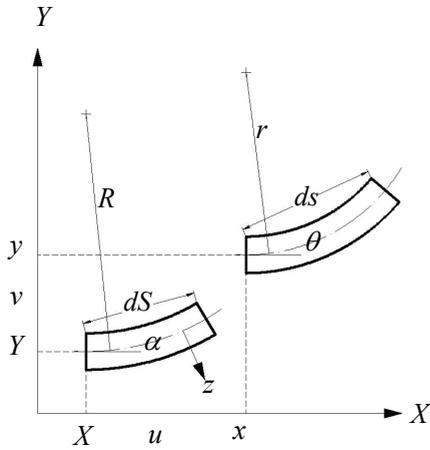


Fig. 1. Deformation of length element

2. Fundamental Equations

2.1 Displacement Field and Kinematics Relationship

Consider an elastic curved laminated beam with variable curvature whose axis lies on a 2-D plane. Assume the rod is made of elastic material such that the stress is linear to the strain even for finite deformation of the beam. Since the strain is finite, the displacement of a point, the extension of the axis and the rotation angle of any cross section are not necessarily small. To simplify the analysis, assume cross sections do not change the shape and size and the cross section is always orthogonal to the axis in the deformed state.

To describe the laminated curve beam on a 2-D reference configuration, the

un-deformed length element dS after deformation become the deformed length element ds . The coordinate of end point (X, Y) in the un-deformed state deforms to (x, y) shown as Fig. 1. At the un-deformed state, the tangent slope angle at (X, Y) is denoted by α . At the deformed state, the tangent slope at (x, y) is denoted by θ . The deformation at (X, Y) is denoted by (u, v) where u is the horizontal displacement, and v is the vertical displacement. Hence

$$x = X + u, \quad y = Y + v. \quad (1)$$

The rotation angle φ can be found by

$$\varphi = \theta - \alpha. \quad (2)$$

Since the strain at the centroid axis is defined by $\varepsilon = (ds - dS) / dS$, or

$$ds = (1 + \varepsilon)dS. \quad (3)$$

As in the case of in-extensional curved beam, $\varepsilon = 0$. For any length element dS , there is a corresponding radius of curvature R , such that

$$dS = R d\alpha. \quad (4)$$

Here the radius of curvature R does not have to be a constant. Most well known curves can be determined by specifying the radius of curvature, such as circle, ellipse, parabola, cycloid, hyperbola, centenary, spiral curves, etc.

For the deformed length element ds , the corresponding radius of curvature is denoted by r , i.e.

$$ds = rd\theta. \quad (5)$$

At a distance z from centroid axis, the un-deformed length element is denoted by

$$dS_z = (R+z)d\alpha. \quad (6)$$

And the deformed length element is

$$ds_z = (r+z)d\theta. \quad (7)$$

The strain at a distance of z defined by

$$\varepsilon_z = \frac{ds_z - dS_z}{dS_z} = \frac{\varepsilon + z \frac{d\varphi}{dS}}{1 + \frac{z}{R}}. \quad (8)$$

Equation (8) can be simplified to

$$\varepsilon_z = \varepsilon + z \frac{d\varphi}{dS}. \quad (9)$$

Here assume $z \ll R$ so that $z/R \ll 1$ can be neglected. In other words, the curved beam is slender in the sense that dimension of cross section is much less than the dimension of radius of curvature.

2.2 Constitutive Relationship

Consider a thin curved laminated beam composed of layers. The layers are arranged on the centroid plane even the fiber can be in different direction. Here the effect of shear deformation is neglected. The stress-strain relationship for a layer along axis direction is

$$\sigma = Q_{11}\varepsilon_z. \quad (10)$$

Where σ is the normal stress component along tangent direction and Q_{11} is the elastic stiffness coefficient of the laminated material. The coefficient Q_{11} is expressed

by

$$Q_{11} = \frac{E_1}{1 - \nu_{12}\nu_{21}}. \quad (11)$$

Where E_1 is the longitudinal Young's modulus along fiber direction, ν_{12} is the longitudinal Poisson's ratio, and ν_{22} is the transverse Poisson's ratio. For a n -layers of laminated beam, the stress-strain relationship of K_{th} layer due to rotation transformation

$$[\sigma]_k = [\bar{Q}_{11}]_k [\varepsilon]_k. \quad (12)$$

Where

$$\bar{Q}_{11} = Q_{11} \cos^4 \gamma + 2(Q_{12} + 2Q_{66}) \cos^2 \gamma \sin^2 \gamma + Q_{22} \sin^4 \gamma \quad (13)$$

With

$$Q_{12} = \frac{\nu_{12}E_2}{1 - \nu_{12}\nu_{21}}, \quad (14)$$

$$Q_{66} = G_{12}, \quad Q_{22} = \frac{E_2}{1 - \nu_{12}\nu_{21}}.$$

In the expression, E_2 is the transverse Young's modulus perpendicular to fiber direction, G_{12} is longitudinal shear modulus, and γ is the angle between tangential direction and fiber direction, respectively. The coefficients Q_{11} , Q_{12} , Q_{66} , Q_{22} are also the elastic stiffness coefficients.

Assume that the width of the cross section is b and the total thickness is $h=nt$ (Fig. 2.). At each cross section there are n -layers from bottom to the top. At the bottom it is the first layer and at the top it is n -th layer.

At any cross section, the resultant force and moment are obtained by integrating the stress in each layer through the thickness. Here the bending moment is positive if it intends to decrease the radius of curvature of the centroid axis. The notation and sign convention of axial force N , moment M together with shear force Q , external

distributed tangential force q_α and radial force q_R are shown in Fig. 3.

The axial force N and moment of the laminate are obtained by integration of stresses in each layer through the thickness,

$$\begin{bmatrix} N \\ M \end{bmatrix} = \begin{bmatrix} A_{11} & B_{11} \\ B_{11} & D_{11} \end{bmatrix} \begin{bmatrix} \varepsilon \\ \phi' \end{bmatrix}. \quad (15)$$

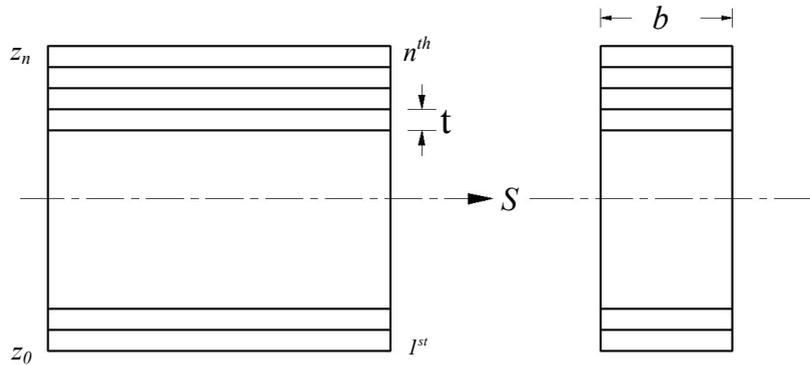


Fig. 2. Layers of laminated curved beam

Where

$$\begin{aligned} A_{11} &= \sum_{k=1}^n b \int_{z_{k-1}}^{z_k} [\bar{Q}_{11}]_k dz, \\ D_{11} &= \sum_{k=1}^n b \int_{z_{k-1}}^{z_k} [\bar{Q}_{11}]_k z^2 dz, \\ B_{11} &= \sum_{k=1}^n b \int_{z_{k-1}}^{z_k} [\bar{Q}_{11}]_k z dz. \end{aligned} \quad (16)$$

The variable ϕ' denotes $d\phi/dS$, the A_{11} represents the in-plane stiffness, the B_{11} defines the bending-stretching coupling and the D_{11} is the bending stiffness. Since in engineering application, most used materials are carbon fiber and glass fiber. The thickness of carbon fiber is almost the same. Therefore in application it is

convenient to use same thickness. However even if different thicknesses are used as the curved laminated beam, the equations of Eq. (16) are still valid. In the case of symmetric laminated, the integral $B_{11}=0$. Then the strain and curvature change will be decoupled, i.e.

$$N = A_{11}\varepsilon, \quad M = D_{11}\phi'. \quad (17)$$

Here it is noted that if the longitudinal Young's modulus and the transverse Young's modulus are equal in each layer, that implies $Q_{11} = Q_{22} = E'/(1-\nu^2)$, then in the laminate $A_{11} = E'A/(1-\nu^2)$, $B_{11} = 0$, D_{11}

$= E'/(1-\nu^2) = EI$. Hence the anisotropic material will reduce to be an isotropic material.

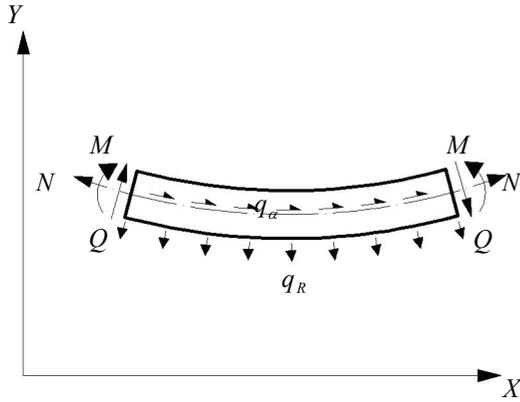


Fig. 3. Sign convention of forces and moment

2.3 Kinetics

To demonstrate the force balance, there are several ways, such as in X - Y plane and expressing all quantities in terms of X - Y components. The other choice is to express all quantities in terms of reference configuration, expressing all quantities in axial and shear components or tangent and normal components. Another choice is in terms of spatial configuration or deformed configuration.

The force balance in the reference configuration can be expressed by

$$\begin{aligned} \frac{dN}{dS} + \frac{Q}{R} &= -q_\alpha, & -\frac{N}{R} + \frac{dQ}{dS} &= -q_R, \\ \frac{dM}{dS} &= Q. \end{aligned} \quad (18)$$

The three equations show the balance of forces along tangential direction, radial direction and moment. The equilibrium equations can be obtained by taking free body of a curved element. These equations are the same as the force balance equation in small deformation [15], since the effect of finite deformation does not effect equilibrium equation in the reference configuration.

The equilibrium equations can also be expressed by deformed configuration,

$$\begin{aligned} \frac{dN}{ds} + \frac{Q}{r} &= -\frac{q_\alpha}{1+\varepsilon}, \\ -\frac{N}{r} + \frac{dQ}{ds} &= -\frac{q_R}{1+\varepsilon}, & \frac{dM}{ds} &= \frac{Q}{1+\varepsilon}; \\ dx &= (1+\varepsilon)\cos\theta dS, & dy &= (1+\varepsilon)\sin\theta dS; \end{aligned} \quad (19)$$

$$N = A_{11}\varepsilon + B_{11}(1+\varepsilon)\frac{d\varphi}{ds},$$

$$M = B_{11}\varepsilon + D_{11}(1+\varepsilon)\frac{d\varphi}{ds}.$$

There are three configurations to describe the balance equations. Later it will show that to solve the system equations, some configuration are needed to combined and solved. The complicated term arises from $1+\varepsilon$ term. This term is retained due to the conservative field of external loads. However this term also induces a nonlinear effect. Note that, from Eq. (19c), there is no stability term. Axial induced moment is not included. Hence the system equation should promote no buckling mode. It shows the finite deformation.

Here assume the external distributed loads q_α , q_R change direction but not in magnitude. There are several ways to describe equilibrium equations. So far there are five equations : three forces Eqs. (19a, b, c) and two constitutive equations Eqs. (19f, g) for N , Q , M , and ε , φ . To complete the analysis of finite deformation, the displacements and Constitutive Relationships are included. Once taking the integrals of Eqs. (19d, e), from Eq. (1), the displacements u , v can be found.

3. Applications

3.1 Pure Bending Solutions

Here to demonstrate the analytical solution of a finite deformation of angle-ply curved laminated beam, consider a curved beam symmetric with respect to Y axis, and extending from $\alpha=-\beta$ to $\alpha=\beta$. The un-deformed radius of curvature may be different. Hence this curved beam can be many kinds of curved beams. A couple of concentrated moments are applied at both ends. Assume that the origin (un-deformed state) is located at $\alpha=0$ (Fig. 3a).

Here the deformed configuration is used. In the absence of distributed loads q_α , q_R in Eqs. (19a, b)

$$\frac{dN}{ds} + \frac{Q}{r} = 0, \quad -\frac{N}{r} + \frac{dQ}{ds} = 0, \quad (20)$$

$$\frac{dM}{ds} = Q;$$

$$\frac{dx}{ds} = \cos \theta, \quad \frac{dy}{ds} = \sin \theta;$$

$$N = A_{11}\varepsilon + B_{11}(1+\varepsilon)\frac{d\varphi}{ds},$$

$$M = B_{11}\varepsilon + D_{11}(1+\varepsilon)\frac{d\varphi}{ds}$$

By using changing of variable Eqs. (20a, b) can be induced to

$$\frac{dN}{d\theta} + Q = 0, \quad -N + \frac{dQ}{d\theta} = 0. \quad (21)$$

Taking derivative of Eq. (21a), combining with Eq. (21b) and eliminating the variable Q , the equation yields,

$$\frac{d^2N}{d\theta^2} + N = 0. \quad (22)$$

The solution of N is

$$N = A_1 \cos \theta + A_2 \sin \theta. \quad (23)$$

Where A_1 , A_2 are two constants to be determine by suitable boundary conditions. Taking derivative of Eq.(23), with the help of Eq. (21a), the variable Q is

$$Q = A_1 \sin \theta - A_2 \cos \theta. \quad (24)$$

At the un-deformed ends $\alpha=\beta$, $N(\pm\beta) = Q(\pm\beta) = 0$. At the deformed ends, $\theta = \theta_\beta = \beta + \varphi(\beta)$, $N(\beta + \varphi(\beta)) = Q(\beta + \varphi(\beta)) = 0$, or

$$\begin{aligned} N &= A_1 \cos(\beta + \varphi(\beta)) + A_2 \sin(\beta + \varphi(\beta)) = 0 \\ Q &= A_1 \cos(\beta + \varphi(\beta)) - A_2 \sin(\beta + \varphi(\beta)) = 0 \end{aligned} \quad (25)$$

Where $\varphi(\beta)$ is the deformed rotation angle at $\alpha=\beta$. It is still an unknown and

needed to find. The solution of Eqs.(25) is

$$A_1 = A_2 = 0.$$

Hence

$$N = Q = 0. \quad (26)$$

With the help of Eq. (26) , integrate Eqs.

(20c) , once to obtain

$$M = A_3. \quad (27)$$

Where A_3 is a constant.

At the free end, the boundary condition is

$$M(\beta + \varphi(\beta)) = M_0. \quad (28)$$

Hence the solution of M is

$$M = -M_0. \quad (29)$$

Therefore Eqs. (20f,g) yield

$$A_{11}\varepsilon + B_{11}(1 + \varepsilon)\frac{d\varphi}{ds} = 0, \quad (30)$$

$$B_{11}\varepsilon + D_{11}(1 + \varepsilon)\frac{d\varphi}{ds} = M_0.$$

The solution of ε and $(1 + \varepsilon) d\varphi/ds$ is

$$\varepsilon = \frac{-M_0 B_{11}}{A_{11} D_{11} - B_{11}^2}, \quad (31)$$

$$(1 + \varepsilon)\frac{d\varphi}{ds} = \frac{M_0 A_{11}}{A_{11} D_{11} - B_{11}^2}.$$

The solution Eq. (31) show that even under pure bending moment, the strain at the centroid line is still deformed. This is caused by material properties. For a symmetric curved laminated beam, $B_{11} = 0$ and $\varepsilon = 0$.

Then there is no deformation along centroid line. It becomes inextensible. Since the coefficients of A_{11} , B_{11} and D_{11} are all positive, and denominator $A_{11}D_{11} - B_{11}^2$ is

normally positive, ε is then negative. In other words, the curved laminated beam becomes longer under the positive pure bending case. Eq. (31b) is expression in terms of un-deformed state for the sake of integration. Integrating Eq. (31b) once to obtain

$$\varphi = \int_0^\alpha \frac{A_{11} M_0}{A_{11} D_{11} - B_{11}^2} R d\alpha + \varphi_0. \quad (32)$$

Where φ_0 is a constants, determined by

boundary conditions. Here the deformation

angle is expressed in term of reference

configuration. Due to symmetric, $\alpha=0$, φ

$(0) = 0$. Hence $\varphi_0 = 0$.

The Eq. (32) implies

$$\varphi = \frac{A_{11} M_0}{A_{11} D_{11} - B_{11}^2} S. \quad (33)$$

The deformed angle can be expressed in the reference configuration. Furthermore the deformed shape can be evaluated by the deformed coordinates x , y . The integral of Eqs. (20d, e) ,

$$x = \int_0^\alpha \left(1 - \frac{M_0 B_{11}}{A_{11} D_{11} - B_{11}^2} \right) \cos \left(\alpha + \frac{A_{11} M_0}{A_{11} D_{11} - B_{11}^2} S \right) R(\alpha) d\alpha, \quad (34)$$

$$y = \int_0^\alpha \left(1 - \frac{M_0 B_{11}}{A_{11} D_{11} - B_{11}^2} \right) \sin \left(\alpha + \frac{A_{11} M_0}{A_{11} D_{11} - B_{11}^2} S \right) R(\alpha) d\alpha.$$

The integration constants are vanished due to the symmetric conditions of $\alpha=0$, $x=0$, $y=0$. There is no deformation at the origin.

The integrals Eq. (4) depend on the form in the integral. Since the arc length is calculated based on the given curvature.

Once the radius of curvature is assigned, the arc can be evaluated. Then the integrals of Eq. (34) can be found. However, only some given radius of curvature can be carried out to explicit form (instead of integral form), due to the difficulty of integral. In this analysis, the results show at least some curves can be calculate explicitly, some are still not, even all curves can be evaluated by numerical techniques. For an angle-ply symmetric curved laminated beam, $B_{11}=0$, $\varepsilon = 0$.

The deformed coordinated Eq. (34) can be simplified to

$$\begin{aligned} x &= \int_0^\alpha R(\alpha) \cos\left(\alpha + \frac{M_0}{D_{11}} S\right) d\alpha, \\ y &= \int_0^\alpha R(\alpha) \sin\left(\alpha + \frac{M_0}{D_{11}} S\right) d\alpha \end{aligned} \quad (35)$$

Here the procedures show that no need to calculate the deformations u , v but directly evaluate the deformed coordinates x , y .

Consider a cantilever angle-ply symmetric curved laminated beam with variable curvature. The curve starts from fixed end $\alpha=0$ to free end $\alpha=\beta$ shown as Fig. 3a. The origin is located at $\alpha=0$. A concentrated moment M_0 is applied at the end of $\alpha=\beta$. It is equivalent to a free curved beam which is symmetric with respect to y -axis. The curved beam starts from free end $\alpha=-\beta$, to $\alpha=\beta$. A pairs of concentrated moment M_0 are applied at both ends shown as Fig. 3b. Due to symmetry, only the

portion from $\alpha=0$ to $\alpha=\beta$ is needed to be considered. Due to fixed end or symmetry, at $\alpha=0$ the boundary conditions are

$$\begin{aligned} \varphi &= 0, \quad u = 0, \quad v = 0, \\ \theta &= 0, \quad x = 0, \quad y = 0. \end{aligned} \quad (36)$$

At the free end ($\alpha=\beta$), the axial force, shear force and moment are

$$N(\beta)=0, \quad Q(\beta)=0, \quad M(\beta)=M_0. \quad (37)$$

Substituting Eqs. (37) into Eqs. (23, 24), the constants can be found $A_1=A_2=0$, $A_3=M_0$. Substituting Eq. (37a) into Eqs. (34a, b), the deformed angle of slope becomes

$$\theta = \alpha + \lambda S, \quad \lambda = \frac{M_0}{D_{11}}. \quad (38)$$

The Eq. (35) can be expression by

$$\begin{aligned} x &= \int_0^\alpha R \cos(\alpha + \lambda S) d\alpha, \\ y &= \int_0^\alpha R \sin(\alpha + \lambda S) d\alpha. \end{aligned} \quad (39)$$

In Eqs. (39), the constants from integrating vanish because at $\alpha=0$, $x=y=0$. Once the radius of curvature of the curve beam is specified, the deformed coordinates can be found. The displacements u , v with the help of Eq. (1) is then

$$\begin{aligned} u &= \int_0^\alpha 2R \sin\left(\alpha + \frac{1}{2}\lambda S\right) \sin\left(\frac{1}{2}\lambda S\right) d\alpha, \\ v &= \int_0^\alpha 2R \sin\left(\frac{1}{2}\lambda S\right) \cos\left(\alpha + \frac{1}{2}\lambda S\right) d\alpha. \end{aligned} \quad (40)$$

Note that the radius of curvature r after deformation can be calculated by taking derivatives of Eqs. (39). It can be

simplified to

$$r = \frac{R}{1 + \lambda R}. \quad (41)$$

It is seen that if the curve is a circle, or R is a constant, then r is a constant as well

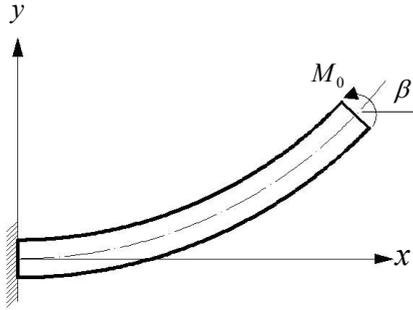


Fig. 3a. A cantilever curved beam under a concentrated moment M_0 .

However for a curve of variable curvature, R is not a constant. The deformed radius of curvature changes to another function by $1/(1 + \lambda R)$. Therefore the curve changes type as it is deformed.

3.2 Finite deformation of circular laminate curve beam under pure bending moments

For an angle-ply circular laminated beam, choose the radius as the characteristic radius, i.e.

$$R = 1. \quad (42)$$

It is equivalent to $X = \sin \alpha$, $Y = 1 - \cos \alpha$ in parametric form. The solution of displacements u , v in Eq. (40) can be integrated to yield

and only changes magnitude. The magnitude factor is $1/(1 + \lambda R)$.

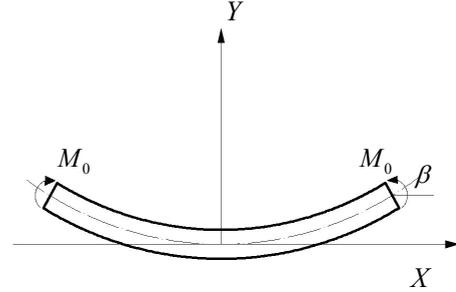


Fig. 3b. Simple curved beams under a pair of moment M_0 .

$$\begin{aligned} u &= \frac{\sin(1 + \lambda)\alpha}{1 + \lambda} - \sin \alpha, \\ v &= -\frac{\lambda}{1 + \lambda} \cos \alpha - \frac{\cos \alpha}{1 + \lambda}. \end{aligned} \quad (43)$$

The deformed coordinates from Eqs. (39) are

$$\begin{aligned} x &= \frac{1}{1 + \lambda} \sin(1 + \lambda)\alpha, \\ y &= \frac{1}{1 + \lambda} [1 - \cos(1 + \lambda)\alpha]. \end{aligned} \quad (44)$$

The coordinates of Eqs. (44) are a circle in parametric form. The radius of curvature is $1/(1 + \lambda)$. In other words as the moment M_0 increases, the radius of curvature decreases by the scale of $1/(1 + \lambda)$. Since the strain at the centroid axis is zero, the circumference can be calculated to close the circular curve. The required moment to close the circle is

$$M_{\alpha_{close}} = \frac{D_{11}}{R_0} \left(\frac{\pi}{\beta} - 1 \right).$$

Or, the required λ_{close} to close the circle is

$$\lambda_{close} = \frac{\pi}{\beta} - 1. \tag{45}$$

For instance to close a quadrant symmetric curve laminated beam (from $-\pi/4$ to $\pi/4$), the required moment is $\lambda_{close}=3$. The deformed curves of a quarter circular at various closed moments are shown in Fig. 4.

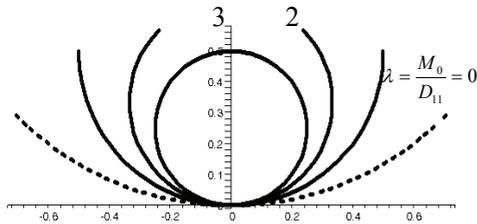


Fig. 4. Symmetric angle-ply circular curved laminated beams from $-\pi/4$ to $\pi/4$ under a couple of moment M_0 .

The final angle versus close moment of angle-ply laminates are shown in Fig. 5.

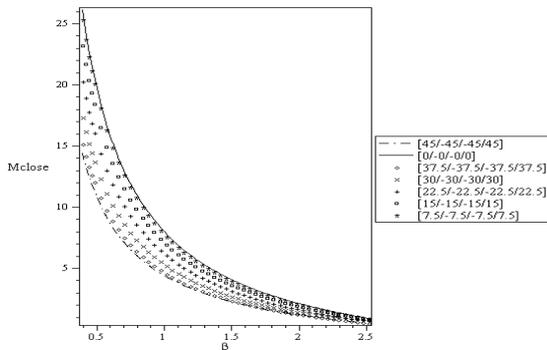


Fig. 5. The final angle versus close moment of various 4-layers stacking sequences.

Consider a cantilever angle-ply symmetric circular curved laminated beam

subjected to a pure bending moment $M_0=D_{11}/R_0$, the curve starts from fixed end $\alpha=0$ to free end $\alpha=\pi/2$. Beams cross-section consist with 2 layers of carbon fiber T300/5208 $[-\theta, \theta]$ Material properties are $E_1=132\text{Gpa}$, $E_2=10.8\text{Gpa}$ and $\nu_{12}=0.24$, $G12=5.65\text{Gpa}$, $M=1.867572156\text{N} \cdot \text{m}$ each layer thickness $h=1\text{mm}$, the width is $b=20\text{mm}$.

The fiber orientation angles versus equivalent λ_{eq} of different stacking numbers are shown in Fig. 6. And the fiber orientation angles versus bending stiffness of different stacking numbers are shown in Fig. 7.

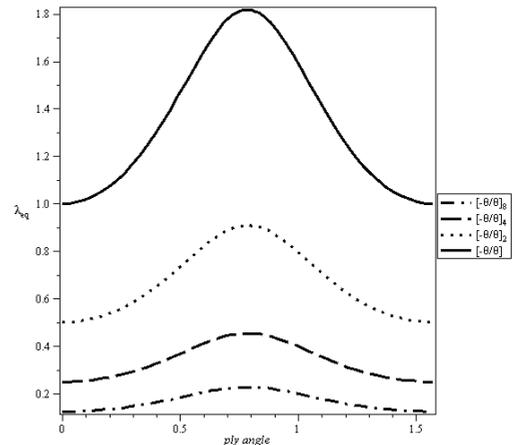


Fig. 6. The fiber orientation angles versus equivalent λ_{eq} of different stacking numbers ($\gamma=0$ to $\gamma=\pi/2$)

Figure 6, shows the cantilever angle-ply symmetric circular curved laminated beam under same apply moment. The different fiber orientation angles stacking sequences

from 0 to $\pi/2$, it can be seen the equivalent λ of curved laminated beam will increase to the tip at $\pi/4$. Fig. 7, shows the cantilever angle-ply symmetric circular curved laminated beam.

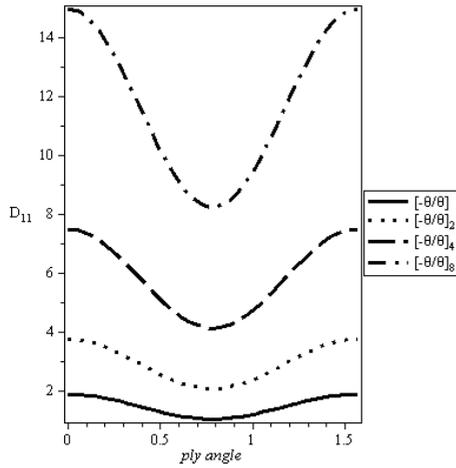


Fig. 7. The fiber orientation angle γ versus bending stiffness of different layer numbers ($\gamma=0$ to $\gamma=\pi/2$)

The different fiber orientation angles and layer numbers from 0 to $\pi/2$, it can be seen the bending stiffness of curved laminated beam, D_{11} , exhibit minima at $\gamma = \pi/4$. With the helps of Eq. (13), the bending stiffness D_{11} can be expressed in terms of the fiber orientation angles and layer numbers n using Eq. (46),

$$D_{11} = \left(\frac{b \cdot h^3}{12} \bar{Q}_{11} \right). \quad (46)$$

Noted that, laminate consisting of an equal number of equal-thickness layers at

$+\theta$ and $-\theta$ fiber orientations are called angle-ply laminates, such laminates are specially orthotropic. Because such laminates do not exhibit coupling between in-plane extensional and shear response. For a symmetric curved laminated beam, $B_{11} = 0$. The deformation is determined by D_{11} .

The deformations of different stacking sequences are shown in Fig. 8.

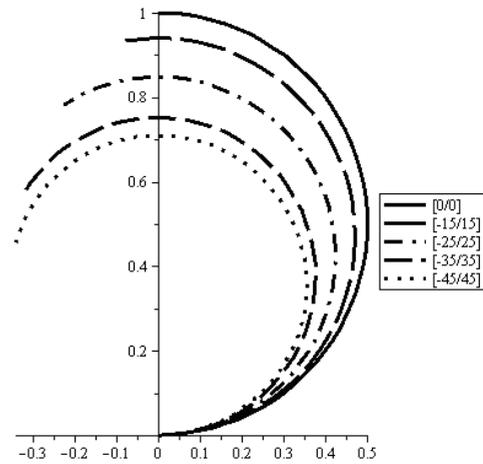


Fig. 8. The deformation shapes of different stacking sequences

Figure 9, shows the free end deformed x -coordinates and y -coordinates with various λ from 0 to 10 . As λ increases to $7, 15, \dots$. The cantilever angle-ply symmetric circular curved laminated beam will always spin to fixed end $(x, y) = (0, 0)$.

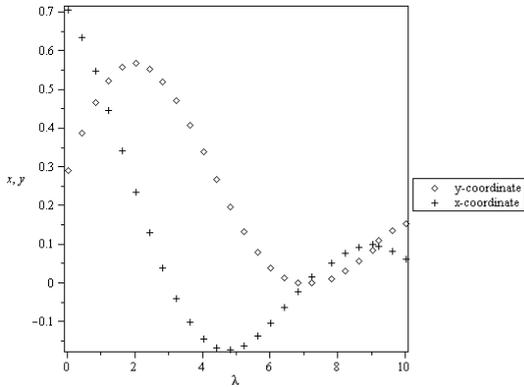


Fig. 9. The free end deformed coordinate of cantilever circular curved beam (from 0 to $\pi/4$)

The deformation shapes shown as Fig. 10.

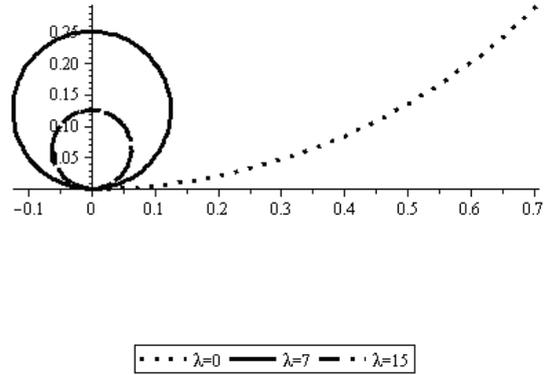


Fig. 10. The deformation of circular curved beam (from 0 to $\pi/4$) under various λ .

3.3 ANSYS results

Consider a cantilever angle-ply symmetric circular curved laminated beam subjected to a pure bending moment $M_0=D_{11}/R_0$, the curve starts from fixed end $\alpha=0$ to free end $\alpha=\pi/2$. Beams cross-section consist with $2n$ layers of carbon fiber T300/5208 [0_{2n}], Material are

$E_1=132\text{Gpa}$, $E_2=10.8\text{Gpa}$ and $\nu_{12}=0.24$, $G_{12}=5.65\text{Gpa}$, each layer thickness $t=1\text{mm}$, the total thickness is $h=2nt$ and the width is $b=20\text{mm}$. Using ANSYS large deformation static analysis, the various R/h versus deformed displacements u is shown in Table 1.

Table 1. The various R/h versus deformed displacements u .

R/h	2500	1250	1000	500	200	100	10
u	0.999737	0.999987	0.999987	0.999751	0.999751	0.999751	0.999751
Error (%)	0.0263%	0.0013%	0.0013%	0.0249%	0.0249%	0.0249%	0.0249%

Where $h/b = 1/10$.

It shows the consistency of the results of present study with those by ANSYS.

4. Conclusions

This paper has presented an analytical method for obtaining the finite deformation of 2-D circular curved laminated beam. The rod axis is either inextensible or extensible. The curved beam is slender in the sense that dimension of cross section is much less than the dimension of radius of curvature. To derive the analytical method for the general solutions, one can introduce the coordinate system defined by the radius of centroidal axis and the angle of tangent slope.

The general solutions expressed by fundamental geometric quantities form a set of equations having seven unknown constants. The seven constants can be directly determined by suitable boundary conditions. As the radius in terms of the tangent slope angle is given, the fundamental geometric quantities can be calculated to obtain the closed form solutions of the axial force, shear force, bending moment, rotation angle, and displacement fields at any cross-section of circular curved laminated beams. These results of the applications indicate that the closed-form general solutions derived by the analytical method would be valid for in-plane thin angle-ply circular curved laminated beams. Thus the analytical method would be useful to engineers

attempting to obtain the exact expressions for thin angle-ply circular curved laminated beams in engineering applications. Especially, helical spring and clockwork spring production and process.

References

- [1] Green A. E., Naghdi P.M., Wenner M. L., *One the theory of rods I. Derivations from the three-dimensional equations*, Proc. Roy. Soc. Lond. A. 337, 1974, 451-483.
- [2] Green A. E., Naghdi P.M., Wenner M. L., *One the theory of rods II. Developments by direct approach*, Proc. Roy. Soc. Lond. A. 337, 1974, 485-507.
- [3] Naghdi P.M., *Finite deformation of elastic rods and shells*, Proc. Iutam Symp. On Finite Elasticity, Lehigh Univ., Aug. 1980, 47-103.
- [4] Green A. E., Naghdi P. M., *Non-isothermal theory of rods, plates and shells*. Int. J. Solids Strucs. 1970, 6, 209-244.
- [5] Green A. E., Laws N., *Remarks on the theory of rods*, J. Elasticity, 1973, 3,

- 179-184.
- [6] Li M., *The finite deformation theory for beam, plate and shell Part I. The two-dimensional beam theory*, Computer Methods Appl. Mech. Engrg. 146 (1997) 53-63.
- [7] Li M., *The finite deformation theory for beam, plate and shell Part III. The three dimensional beam theory and FE formulation*.
- [8] Oguibe C.N., Webb D. C., Large deflection analysis of multilayer cantilever beams subjected to impulse loading, Computers and Structures, 78, 2000, 537-547.
- [9] Attard M. M., *Finite strain – beam theory*, International J. Solids and Structures, 40, 2003, 4563-4584.
- [10] Mauget B., Perré P., *A large displacement formulation for anisotropic constitutive laws*, Eur. J. Mech. A/ Solids 18, 1999, 859-877.
- [11] Toi Y., Lee J., Taya M., *Finite element analysis of super-elastic large deformation behavior of shape memory alloy helical springs*, Computers and Structures 82, 2004, 1685-1693.
- [12] Atanackovic T. M., *Stability theory of elastic rods*, World Scientific, 1997.
- [13] Brush D. O., Almroth B. O., *Buckling of bars, plates and shells*, McGraw-Hill New York, 1975, 121
- [14] Timoshenko S. P., *Theory of elastic stability*, McGraw-Hill, 1961, 76.
- [15] Lin K. C., Huang S. H., *Static Closed-form Solutions for In-plane Thick Curved Beams with Variable Curvatures*, J. Solid Mechanics and Material Engineering, 2007, 1, 1026-34.
- [16] Lin K. C., Huang S. H., *Static Closed-form Solutions for In-plane Shear Deformable Curved Beams with Variable Curvatures*, J. Solid Mechanics and Material Engineering, 2007, 1, 1362-73.
- [17] Lin, K.C. and Hsieh, C.M., *The closed form general solutions of 2-D curved laminated beams of variable curvatures*, Composite Structures Vol.79 (2007) , pp.606-618.
- [18] Lin C. W., *Finite Deformation of 2-D Thin circular curved beams*, Hsiuping Journal, Vol. 19 (2009) , pp.203-216.
- [19] Herakovich C. T., *Mechanics of Fibrous Composites*, John Wiley & Sons, 1998.

獨立伯努利變數和之變異數與香農熵

繆紹昌、陳志賢、劉家頤

摘要

假設某國中某班共有 n 名學生，令 p_i ($0 \leq p_i \leq 1$) 為第 i 位學生能順利進入理想高中之機率。令 X_i 為參數 p_i 之伯努利隨機變數，則 $S = X_1 + X_2 + \cdots + X_n$ 為能進入理想高中之總人數。在 $p_1 + p_2 + \cdots + p_n$ 為一固定常數的限制下，以兩種方法找出 $Var[S]$ 之極大值與極小值之條件，也建立起 $Var[S]$ 之極值與香農熵之關係。

關鍵詞：蓋理論、蕭爾-凸性質、香農熵、凸多邊形區域。

On the Variance of Sum of Independent Bernoulli Random Variables and Shannon Entropy

Shao-Chang Miao, Chih-Sheng Chen, Chia-Yee Liu

Abstract

There are n students in a class. The i th student is evaluated and assigned a constant p_i ($0 \leq p_i \leq 1$) reflecting the student's probability of being admitted to an ideal high school. Let X_i ($1 \leq i \leq n$) be independent Bernoulli random variables with parameters p_i . Then $S = X_1 + X_2 + \cdots + X_n$ is the number of the students in the class who will be admitted to an ideal high school. Assuming that $p_1 + p_2 + \cdots + p_n$ is a fixed constant, the maximum and minimum values of $\text{Var}[S]$ are obtained using two different methods. The notions of majorization and Shannon entropy relevant the problem are defined and discussed. The relationships between the extremal values of $\text{Var}[S]$ and Shannon entropy are also established.

Keywords : Majorization, Schur-Convexity, Shannon Entropy, Convex polygonal Region.

1. 前言

在統計學領域裡一個重要的課題是在合理的限制下，建立出一個變異數最小之點估計；如數理統計中之最小變異數不偏點估計或迴歸分析中之最佳線性不偏點估計。令 X_i ($i=1,2,\dots,n$) 為參數 p_i 之獨立伯努利 (Bernoulli) 隨機變數， $S = X_1 + X_2 + \dots + X_n$ 。本文假設 p_i 的總和為一定值的限制下，研究 S 之變異數性質，並建立出 $\mathbf{p}=(p_1, p_2, \dots, p_n)$ 應滿足之條件，使得 S 之變異數為極大或極小。為瞭解本題之隨機性質，也定義與研究兩個相關之香農熵 (Shannon entropy)，並利用香農熵解釋 S 之變異數之極值。本文使用之研究方法為凸分析 (convex analysis) 和蓋理論 (majorization) 與相關之蕭爾-凸性質 (Schur convexity) 理論。

李天岩[1]有關香農熵部分避開艱深的數學語言，對於香農熵和相關的議題，作了詳細的介紹；香農熵之定義與精神，在 Ross [9]有清楚之說明；Bose[3]香農熵相關之計算與在資訊科學上的應用為一可讀性很高之資料來源。另外 Rao 等人[8]也針對傳統香農熵定義上之缺點，提出一個替代測量值並建立出此測量值的一些良好性質。蓋理論和蕭爾-凸性質理論部分，最完整的參考資料為 Marshall 與

Olkin [7]和經典之 Hardy 等人[6]。楊重駿與楊照崑[2]也介紹蓋理論之背景及其應用。在凸分析和凸規劃 (convex programming) 部分，Valentine [11]詳細的研究凸集合的結構；最近之 Boyd 與 Vandenberghe [4]有系統的討論以 n 維歐氏空間為背景之凸規劃問題；另外 Gass [5]在第二章和第十二章分別針對凸多面體 (convex polyhedron) 的結構和非線性規劃作了清楚的介紹。

本文結構如下：第 2 節描述問題之結構；第 3 節利用凸分析方法解出相關之最佳化問題；第 4 節利用蓋理論的方法提出另一個解法；第 5 節定義二個相關之香農熵，說明智力資產分配平均程度和 $Var[S]$ 大小之關係；第 6 節為結論。

2. 問題的描述

假設某國中某班共有 n 名學生，每位學生經評估後均給予一個常數 p 值 ($0 \leq p \leq 1$) 代表此學生能夠順利進入理想高中之機率。假設第 i 位學生之 p 值為 p_i ，令 X_i 為參數 p_i 之伯努利隨機變數 ($P[X_i = 1] = p_i$ ， $P[X_i = 0] = 1 - p_i$)。因為 X_i 代表第 i 位學生是否能進入理想高中，可以合理的假設 X_1, X_2, \dots, X_n 為獨立隨機變數，則 $S = X_1 + X_2 + \dots + X_n$ 為全班 n 名學生

中，能順利進入理想高中的總人數。就該班導師而言，一個自然發生的問題是如果機會允許的話，要如何選擇一組最佳的 p_i 值。在理想狀況之下，選擇每個 p_i 值均等於 1，則 $S \equiv n$ ；這樣的選擇會滿足任何“好”的準則（criteria of goodness）。

因此，一個比較合理的問題是假設全班期望順利進入理想高中人數 $E[S] = p_1 + p_2 + \dots + p_n$ 為一定值 L 時，考慮要如何選取一組 p_i 使得 $Var[S]$ 為最小。一般人的直覺反應是推測選取 $p_1 = p_2 = \dots = p_n = \frac{1}{n}L = \bar{p}$ 時， $Var[S]$ 會有最小值。但簡單的計算可得

$$\begin{aligned} Var[S] &= n\bar{p} - \sum_{i=1}^n p_i^2 = \\ & n\bar{p}(1 - \bar{p}) - \sum_{i=1}^n (p_i - \bar{p})^2 \end{aligned} \quad (2-1)$$

從上式公式可知，在 $p_1 = p_2 = \dots = p_n = \frac{1}{n}L = \bar{p}$ 時， $Var[S]$ 會有極大值而不是極小值。換言之，若執行嚴格的能力分班，使得班上 n 位學生之 p 值均為 \bar{p} ，將使得 S 的離散程度為極大。

根據上述之說明，一個自然延伸的問題是在假設 $E[S] = \sum_{i=1}^n p_i$ 為一固定常數 L 時，如何選取一組 p_i 值，使得 $Var[S]$ 有極小值。

3. 最佳化（一）—凸分析

根據 (2-1)，在 $\sum_{i=1}^n p_i = L$ 的限制條件

下， $Var[S]$ 的極大值與極小值是由 $\sum_{i=1}^n p_i^2$ 而決定。因此必須解出下列非線性規劃問題

$$\max/\min f(\mathbf{p}) = p_1^2 + p_2^2 + \dots + p_n^2 \quad (3-1)$$

$$\text{s.t. } p_1 + p_2 + \dots + p_n = L \quad (\equiv k + d) \quad (3-2)$$

$$p_i \leq 1 \quad (1 \leq i \leq n) \quad (3-3)$$

$$p_i \geq 0 \quad (1 \leq i \leq n) \quad (3-4)$$

(3-2) 中之常數 L 滿足 $L \in [0, n]$ 。為方便討論，令 $L = k + d$ ； k 為 L 之整數部分（integral part，小於或等於 L 之最大整數）與 $d = L - k$ 為 L 之分數部分（fractional part）。

當 $L = k$ 時，利用 $0 \leq p_i^2 \leq p_i \leq 1$ 即可得

$$p_1^2 + p_2^2 + \dots + p_n^2 \leq p_1 + p_2 + \dots + p_n = k$$

因此目標函數 $f(\mathbf{p})$ 的極大值為 k 和 $Var[S]$ 的極小值為 0， $Var[S]$ 的極小值發生於 $\mathbf{p} = (1, 1, \dots, 1, 0, \dots, 0)$ (k 個 1 和 $(n-k)$ 個 0) 和此點分量之所有排列，共

$\binom{n}{k}$ 個點。由 (2-1) 可以直接看出，當 $L=k$ 時， $Var[S]$ 的極大值為 $n\bar{p}(1-\bar{p})$ ， $\bar{p} = k/n$ 。

當 $L=k+d$ ($d \neq 0$) 時，找 $f(\mathbf{p})$ 的極大值的困難度較高；但可以利用目標函數和可行區的凸性質找出 $f(\mathbf{p})$ 的極大值。此非線性規劃問題的可行區是由 (3-2)，(3-3) 與 (3-4) 所定義出一個座落於超平面 (hyperplane) (3-2) 上的一個有界之閉凸多邊形區域 (closed and bounded convex polygonal region)。此可行區之頂點 (vertices)，是由 (3-2)，(3-3) 與 (3-4) 共 $(2n+1)$ 個超平面，任取 n 個的交點 (如果存在和可行) 而決定。考慮這 $(2n+1)$ 個超平面的方程式，可立刻結論此可行區之頂點座標為 $(1, 1, \dots, 1, d, 0, \dots, 0)$ (k 個 1，1 個 d 與 $(n-k-1)$ 個 0) 和此點之分量之所有排序，共 $(n-k)\binom{n}{k}$ 個點。根據克林-米爾曼 (Krein-Milman) 定理，見 page 207，(Royden, 1968) 或 page 36，Gass [5]，此可行區內的每一個點都可以寫成這 $(n-k)\binom{n}{k}$ 個頂點之凸組合 (convex combination)。換言之，可行區內的任一點 \mathbf{p} 均可寫成

$$\mathbf{p} = \sum_{i=1}^{(n-k)\binom{n}{k}} a_i \mathbf{u}_i \quad (3-5)$$

$$\text{其中 } \mathbf{u}_i = \text{頂點}, a_i \geq 0, \sum_{i=1}^{(n-k)\binom{n}{k}} a_i = 1 \quad (3-6)$$

目標函數 (3-1) 是個定義於凸集合可行區上之凸函數 (convex function) (見 Marshall 與 Olkin [7])，即

$$\begin{aligned} & f(b_1 \mathbf{u}_1 + b_2 \mathbf{u}_2 + \dots + b_m \mathbf{u}_m) \\ & \leq b_1 f(\mathbf{u}_1) + b_2 f(\mathbf{u}_2) + \dots + b_m f(\mathbf{u}_m) \end{aligned} \quad (3-7)$$

其中 $\mathbf{u}_i \in$ 可行區， $b_i \geq 0$ ， $\sum_{i=1}^m b_i = 1$ ， $m \geq 1$ 。

目標函數 (3-1) 在每一個頂點的函數值均為 $k+d^2$ ，由 (3-6) 與 (3-7) 可知目標函數在可行區內任一個點的函數值均小於或等於 $k+d^2$ 。因此 $f(\mathbf{p})$ 的極大值為 $k+d^2$ 和 $Var[S]$ 之極小值為 $d-d^2$ 。整合上述結果，可得下列定理：
定理一：

在 (3-2)，(3-3) 與 (3-4) 的限制下， $Var[S]$ (2-1) 的極小值與極大值分別為：

(1) 若 $p_1 + p_2 + \dots + p_n = k$ ， $Var[S]$ 的極小值為 0，發生於 $\mathbf{p} = (1, 1, \dots, 1, 0, 0, \dots, 0)$ 和此點之分量之所有排列，共 $\binom{n}{k}$ 個點； $Var[S]$ 的極大值為 $n\bar{p}(1-\bar{p})$ ，發生於 $\mathbf{p} = (\bar{p}, \bar{p}, \dots, \bar{p})$ 。

(2) 若 $p_1 + p_2 + \dots + p_n = k+d$

($d \neq 0$)， $Var[S]$ 的極小值為 $d \cdot d^2$ 發生於 $\mathbf{p} = (1, 1, \dots, 1, d, 0, \dots, 0)$ 和此點之分量之所有排列，共 $(n-k) \binom{n}{k}$ 個點； $Var[S]$ 的極大值為 $n\bar{p} (1-\bar{p})$ ，發生於 $\mathbf{p} = (\bar{p}, \bar{p}, \dots, \bar{p})$ 。

4. 最佳化 (二) — 蓋理論

從 $Var[S]$ 極小值與極大值發生的點的型式，可以對此問題的結構作更進一步的瞭解。可以將 p_i 解釋為第 i 位學生的智力資產，因此本題可以解釋為在全班總智力資產為 $k+d$ (即 (3-2)) 和每位學生的智力資產必需滿足 (3-3) 與 (3-4) 的智力資產門檻限制下，要如何分配總智力資產 $k+d$ ，使得 $Var[S]$ 極小或極大。根據定理一，在智力資產分配最不平均時 (見定義一後之說明)， $Var[S]$ 有極小值；在智力資產分配最平均時， $Var[S]$ 有極大值。這個觀察，加上目標函數 (3-1) 是個定義在凸集合 (可行區) 上之蕭爾-凸函數 (Schur convex function, 見定義二)，清楚的指出，可以用蓋理論的觀點來研究此問題。為節省篇幅，將只考慮 $p_1 + p_2 + \dots + p_n = k + d$, $d \neq 0$ 的狀況。

首先介紹和蓋理論相關的數個定義與觀念。令

$\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n$ ，將 \mathbf{x} 之分量由大而小排序得 $\bar{x}_1 \geq \bar{x}_2 \geq \dots \geq \bar{x}_n$ ，令

$$\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) ;$$

稱 $\bar{\mathbf{x}}$ 為 \mathbf{x} 之遞減排序。 \mathbf{y} 與 $\bar{\mathbf{y}}$ 之定義亦同。

定義一：(Marshall 與 Olkin [7])
假設 (1)

$$\sum_{i=1}^k \bar{x}_i \leq \sum_{i=1}^k \bar{y}_i, \quad i=1, 2, \dots, n-1,$$

假設 (2)

$$\sum_{i=1}^n \bar{x}_i = \sum_{i=1}^n \bar{y}_i$$

則稱 \mathbf{y} 蓋過 (majorizes) \mathbf{x} 或 \mathbf{x} 被 \mathbf{y} 蓋過 (\mathbf{x} is majorized by \mathbf{y})，以符號 $\mathbf{x} \prec \mathbf{y}$ 或 $\mathbf{y} \succ \mathbf{x}$ 表示。

從定義一中可知 \mathbf{x} 與 \mathbf{y} 分量之順序，在 $\mathbf{x} \prec \mathbf{y}$ 關係中並沒有扮演任何角色。若將 \mathbf{x} 與 \mathbf{y} 視為兩種資產或所得分配，則 \mathbf{x} 對應的勞倫司曲線 (Lorenz curve) 比 \mathbf{y} 對應的勞倫司曲線更接近於 45 度線；即 \mathbf{x} 所代表的分配比 \mathbf{y} 代表的分配更平均 (Marshall 與 Olkin [7])。

利用定義一，可以很容易的證明，可行區 (3-2), (3-3) 與 (3-4) 內的任一點 \mathbf{p} 均滿足下列關係式：

$$(\bar{p}, \bar{p}, \dots, \bar{p}) \prec \mathbf{p} \prec (1, 1, \dots, 1, d, 0, \dots, 0) \quad (4-1)$$

定義二：(Marshall 與 Olkin [7])

假設 $A \subseteq \mathbf{R}^n$ 。令 \mathbf{g} 為定義於 A 之實

數值函數，若 g 滿足下列條件，則稱 g 為定義於 A 之蕭爾-凸函數：

$$\mathbf{x}, \mathbf{y} \in A, \mathbf{x} \prec \mathbf{y} \Rightarrow g(\mathbf{x}) \leq g(\mathbf{y}) \quad (4-2)$$

有關蕭爾-凸函數成立之條件與相關之運算 (Marshall 與 Olkin [7])，根據目標函數 (3-1) 是一個定義於可行區之蕭爾-凸函數。由 (4-1) 與 (4-2) 可得，對可行區內的任何一點 \mathbf{p} ，下列不等式成立：

$$f(\bar{p}, \bar{p}, \dots, \bar{p}) \leq f(\mathbf{p}) \leq f(1, 1, \dots, 1, d, 0, \dots, 0)。$$

因此 $Var[S]$ 的極小值與極大分別成立於 $(1, 1, \dots, 1, d, 0, \dots, 0)$ 與 $(\bar{p}, \bar{p}, \dots, \bar{p})$ 。這提供另一種找 $Var[S]$ 之極值的方法。

5. 香農熵與 $Var[S]$

香農熵的觀念是由美國電腦工程師和數學家 Claude E. Shannon [10] 於 1948 年發表之著名論文中所建立 (Shannon, 1948)，此論文對資訊科學建立出一個完整的數學基礎。

令 $\mathbf{u} = (u_1, u_2, \dots, u_m)$ 代表一個離散機率測度 (discrete probability measure)；即 $u_i \geq 0$ 和 $u_1 + u_2 + \dots + u_m = 1$ 。則 \mathbf{u} 之香農熵為 (Ross [9])：

$$H(\mathbf{u}) = -\sum_{i=1}^m u_i \log u_i \quad (5-1)$$

($u_i = 0$ 時， $0 \cdot \log 0 \equiv \lim_{x \rightarrow 0} (x \log x) = 0$)。

可以很容易的證明，當 $u_i = \frac{1}{m}$ 時， $H(\mathbf{u})$ 有極大值。在機率學的領域中， $H(\mathbf{u})$ 代表測度 \mathbf{u} 之不確定性程度 (degree of uncertainty) 或亂度。可以很容易的證明出 $H(\mathbf{u})$ 是一個蕭爾-凹函數 (即 $-H(\mathbf{u})$ 是蕭爾-凸函數，見 Marshall 與 Olkin [7])。因此，假設 $\mathbf{u} = (u_1, u_2, \dots, u_m)$ 和 $\mathbf{v} = (v_1, v_2, \dots, v_m)$ 為兩個離散機率測度，則會有

$$\begin{aligned} \mathbf{u} \prec \mathbf{v} &\Rightarrow H(\mathbf{u}) \geq H(\mathbf{v}), \\ \text{和 } H(1, 0, \dots, 0) &\leq H(\mathbf{u}) \leq H\left(\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m}\right)。 \end{aligned}$$

從資產或所得分配的觀點而言，隨著測度 \mathbf{u} 分配的平均程度 (degree of evenness) 增大， $H(\mathbf{u})$ 也隨著增大。

在經濟學領域中，在不考慮 $u_1 + u_2 + \dots + u_m = 1$ 的離散機率測度限制下， \mathbf{p} 之香農熵 $H(\mathbf{p}) = -\sum_{i=1}^n p_i \log p_i$ 也被提議作為所得平均程度的指標值 (measure of equality 見 Marshall 與 Olkin [7])。因此，全班 n 名學生智力資產最平均時， $Var[S]$ 有極大值；智力資產最不平均時， $Var[S]$ 有極小值。

當 p_i 值均相等時, S 的機率分佈為二項式分佈。但 p_i 不相等時, S 的機率分佈則是相當的複雜。令

$$\begin{aligned} u(k) &= u(k; p_1, p_2, \dots, p_n) \\ &= P[S = k], \quad k=0, 1, \dots, n \end{aligned} \quad (5-2)$$

S 分佈之香農熵為

$$\begin{aligned} H_s(p_1, p_2, \dots, p_n) &= \\ &= -\sum_{k=1}^n u(k) \log(u(k)) \end{aligned} \quad (5-3)$$

根據定理 E.1. (Marshall 與 Olkin [7])， $H_s(p_1, p_2, \dots, p_n)$ 為 (p_1, p_2, \dots, p_n) 之蕭爾-凹函數。這說明隨著學生智力資產越趨於平均, S 之香農熵也跟著變大。

令 D 為可行區 (3-2), (3-3) 與 (3-4) 和集合 $\{(p_1, p_2, \dots, p_n) : p_1 \geq p_2 \geq \dots \geq p_n\}$ 之交集。定義一之二元關係 \prec 為集合 D 上之一個偏排序 (partial order); $H(p_1, p_2, \dots, p_n)$, $H_s(p_1, p_2, \dots, p_n)$ 和 $Var[S]$ (將定義域限制到 D) 均是 \prec 之逆序函數 (order reversing function, 即

$$\mathbf{x} \prec \mathbf{y} \Rightarrow \varphi(\mathbf{x}) \geq \varphi(\mathbf{y}))$$

從統計的觀點而言, 只要知道集合 D 中任二個點之偏排序 \prec 關係即可推論出 $H(p_1, p_2, \dots, p_n)$, $H_s(p_1, p_2, \dots, p_n)$ 和 $Var[S]$ 在這二個點的函數值的大小關係。在集合 D 中根據偏排序 \prec 所決定之

最小元素 (minimal element) 和最大元素 (maximal element) 分別是 $(\bar{p}, \bar{p}, \dots, \bar{p})$ 和 $(1, 1, \dots, 1, d, 0, \dots, 0)$; 這二個點分別是集合 D 中智力資產分配平均程度最高和最低之二個點; $H(p_1, p_2, \dots, p_n)$, $H_s(p_1, p_2, \dots, p_n)$ 和 $Var[S]$ 在這二個點分別有極大值和極小值。

從編班的觀點而言, 在 $E[S]=L$ 的限制下, 隨著班上 n 名同學的同質性增強 (即其智力資產分配更趨於平均), 將使 S 之香農熵與 $Var[S]$ 變大。一般人誤以為會分攤風險之最平均之智力資產分配 $(\bar{p}, \bar{p}, \dots, \bar{p})$, 事實上會使 S 分佈之香農熵 $H_s(p_1, p_2, \dots, p_n)$ 和 $Var[S]$ 為極大。若政策允許, 選取 $(1, 1, \dots, 1, d, 0, \dots, 0)$ 將使 $H_s(p_1, p_2, \dots, p_n)$ 和 $Var[S]$ 為極小。

6. 結論

本文從教育的觀點說明在 $E[S]=p_1 + p_2 + \dots + p_n$ 為一個常數時, $Var[S]$ 的大小和向量 $\mathbf{p}=(p_1, p_2, \dots, p_n)$ 依蓋理論所定義之分配平均程度呈相同方向之變動。從香農熵的觀點而言, $\mathbf{p}=(p_1, p_2, \dots, p_n)$ 之分配平均程度變大造成 \mathbf{p} 之香農熵與 S 分佈之香農熵 (5-3) 變大, 這直接的解釋 \mathbf{p} 之亂度變大, 造成 S 之分佈亂度變大, 也造成 $Var[S]$ 變大。

分攤風險（即 p_i 值均同）是一般人認為避免重大損失（但也避免重大收益）的保守中庸之道。本文證明，當某些風險因素可以完全掌控時（即 $p_i=1$ 或 0 ），應充分善用此契機，在 $E[S]$ 為一常數限制時， $Var[S]$ 會有最小值。

參考文獻

中文部份

[1] 李天岩（1989）熵（Entropy），數學傳播第十三卷第三期，中央研究院數學研究所發行。

[<http://www.math.sinica.edu.tw/math-media/d133/13301.pdf>]

[2] 楊重駿，楊照崑（2002），蓋理論（Theory of Majorization）及其在不等式上的應用，數學傳播第六卷第四期，中央研究院發行。

[<http://www.math.edu.tw/math-mdia/d64/d6402.pdf>]

英文部分

[3] Bose, R. (2003), Information Theory, Coding and Cryptography, McGraw Hill, Singapore.

[4] Boyd, S. and Vandenberghe, L. (2004), Convex Optimization, Cambridge University Press, New York.

[5] Gass, S. (1994), Linear Programming, 5th ed., McGraw Hill, Singapore.

[6] Hardy, G., Littlewood, J. E., and Polya, G. (1952), Inequalities, Cambridge University Press, New York.

[7] Marshall, A. W. and Olkin, I. (1979), Inequalities: Theory of Majorization and Its Applications, Academic Press, New York.

[8] Rao, M., Chen, Y., and Vemuri, B. C. (2004), Cumulative Residual Entropy: A New Measure of Information, IEEE Trans. On Information Theory, Vol. 50, No. 6, June, 2004, pp 1220-1228.

[9] Ross, S. (2006), A First Course in Probability, 7th ed., Pearson Education International, Singapore.

[10] Shannon C. E. (1948), A Mathematical Theory of Communication, Vol 27, pp 379-423, 623-656, The Bell System Technical Journal. [<http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>]

[11] Valentine, F. A. (1968), Convex Sets, McGraw Hill, New York.

氣動式手臂之設計及氣壓缸動態模型推導

許銘全、陳建銘、劉芳志

摘要

本文介紹一個以氣壓缸動態模型為基礎之氣動式手臂 (Pneumatic Robotic Arm ; PRA)，本 PRA 為一種讓上肢動作不便之患者所使用的輔助器具。由於手對於身體來說是相當重要的部位，假使手失去了原有的功能，對於日常生活是有相當嚴重的影響。因此，設計此款 PRA，使手臂可以達到有效的復健功能，即為本文之目的。本文首先建立一個由雙動式單桿氣壓缸和電磁閥所組成的氣動式系統模型，其目的是為了解氣動式驅動器的設計與控制要求。利用該模型我們取得氣壓缸動態模型的方程式表示式。證實氣壓缸用於 PRA 是可行的，可以取代傳統馬達，減輕其重量。同時，在 PRA 上安裝壓力感測器來感知患者受到機構帶動時的力量，以提供系統做出最適合患者之動作。將整套系統運用在復健醫學中，透過此 PRA 引導患者動作，以達到復健之目的。

關鍵詞：氣壓缸、動態模型、氣動式、手臂、復健。

Design of Pneumatic Robotic Arm and Dynamic Model of Pneumatic Cylinder Actuators

Ming-Chiuan Shiu, Chien-Ming Chen, Fang-Jr Liou

Abstract

A robotic arm is developed to act as a part of physical treatment of rehabilitation to restore and maintain functional ability for injured upper extremities. By combination of relay circuit, DSP, the pneumatic cylinders and the electromagnetic valves, the pneumatic controlled robotic arm is proposed in this paper. Via a proper decision of rehabilitation profile of robotic arm, computer program can be compiled first offline. After downloading the profile into DSP, DSP can drive the robotic arm acting as the desired rehabilitation profile by the relay circuit, the pneumatic cylinders and the electromagnetic valves. This paper first develops an exact model of a pneumatic system consisting of a double-acting single-rod pneumatic cylinder and an electromagnetic valve, with the goal of providing an insight into the design and control requirements for pneumatically actuated systems. Use the model to derive the group equations representing the dynamic behavior of the pneumatic cylinder. In addition, the pressure sensors are used to detect the force of patients when the pneumatic robotic arm is action. These pressure sensors can provide the appropriate action for the system. Some basic experiments of rehabilitation poses are demonstrated the performance of the pneumatic robotic arm.

Keywords : rehabilitation, robotic arm, pneumatic control.

壹、緒論

衛生署調查發現，台灣每年約有 68000 人罹患中風、13000 人死於中風，存活的中風患者約有 85% 會有不同程度的肢體障礙，中風五年內，二次中風的發生率也高達 30%，是國內「成人殘障」首要病灶，也是次於惡性腫瘤第二大死因。根據統計高達 75% 的患者有輕度至重度的活動障礙，吃飯、穿衣、行動等需要依賴別人照顧。病情嚴重患者約有 15%，有危及生命的可能性。另外有 10% 的患者會恢復至中風前的活動能力[1]。因為絕大多數的中風患者都有行動障礙的問題，所以復健工作是需要相當龐大的人力資源。中風後的黃金復健治療期為出院後六個月內到一年之間，錯過黃金時期未復健，身體機能就會急速的惡化。由於許多家庭礙於經濟和交通因素的考量，因而延遲就診復健，錯過了黃金治療時期。基於上述的理由，本文主要針對上肢的部分做研究，我們設計出一個專為上肢不便之患者所使用的復健輔具，且具有攜帶方便、低成本等優點。由於手對於身體來說是相當重要的部位，如果手失去了功能，對於日常生活的影響是相當嚴重的，更別說是謀生的能力。因此，我們設計此款氣動式手臂（Pneumatic Robotic Arm；PRA）如圖 1 所示。以幫助其肩、肘受過傷或手術過後，需要復健的患者，藉由此種復健手

臂的幫助，使其手臂可以達到有效的復健功能。

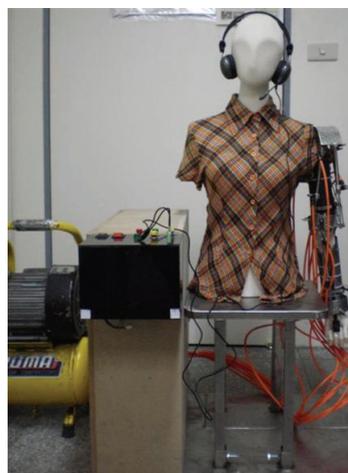


圖 1 氣動式手臂（PRA）實體圖

早期氣動式驅動器使用僅限於簡單和非精確性定位的應用。因為空氣的壓縮性和非線性在轉換的過程中，會產生不穩定狀態。在過去幾年當中，有了突破性的發展，使得氣動系統在設計方法和先進控制上有更良好的進展[2]。1988 年 Liu 等 [3] 制定有效的位置控制器基於線性驅動器位置的中間行程。1991 年 Bobrow 等 [4] 說明使用適應性控制用於力量驅動和軌跡跟蹤。1995 年 Ben-Dov 等 [5] 發展有效的力量控制用於氣動式驅動器，使驅動控制力量高達 16 赫茲。

在復健機器人的物理治療方面，應用已經超過十年。許多文獻提出了他們對於復健機器人在上肢治療上的一些看法。在最近的研究中，1991 年 MIT [6] 設計完成第一台上肢復健訓練機器人，此系統稱為

MIT MANUS，由五連桿機構所組成，末端阻抗小，利用阻抗控制實現訓練的穩定性和安全性，用於患者的肩、肘動作。2005年 Sanchez 等[7]設計出一套命名為 Pneu-WREX 的氣動式機器人，其機器人使用了氣動式驅動器（氣壓缸與活塞）和非線性控制，利用彈性帶纏繞在兩個四桿機構以平衡手臂，WREX 有 5 個自由度，能夠完成簡單的手臂動作。2007 年 Nef 等[8]在瑞士蘇黎世大學研發出復健手臂機器人 ARMin，其機器人主要功能是重複且密集的訓練手臂，增強受傷過後手臂的復健和治療上肢癱瘓的中樞神經系統，此結構是由四個主動和兩個被動的自由度組成，加上轉矩與力矩感測器，能夠進行肘部屈伸和肩膀的空間動作。2007 年由台灣大學資工系 等[9]研發出乳癌機器人，其機器人擁有九個關節手臂，設計了手腕、上臂運動、抬舉等復健動作。並利用無線生理訊號感測器，將患者的體溫、心跳和復健效果，傳回主機，即時判別復健的成效和心理狀態。由上述所見，相關文獻都已證實機器人用於復健治療上是可行的，經由電腦的運算法則可以達到準確的位置和力量控制，並在電腦中整合各種不同的量測數據，由不同量化的物理量來評估復健治療的成效[10]。

在過去的文獻當中，關於復健手臂大多為電控系統，也就是利用馬達來作為主

要的驅動元件，在此我們使用氣壓缸來取代馬達，減輕其重量。本創作的目的是希望能幫助在手臂活動有困難的患者，讓患者能夠藉由復健器材的幫助，加強復健的治療效果。除此之外，復健工作是需要相當多的人力資源協助。因此，這方面非常耗費時間及人力，透過 PRA，可以減少時間及人力的耗費，而達成快速且有效的復健工作。

本篇報告共分為七章。第二章介紹系統特性及其設計概念。第三章則是探討系統控制流程。第四章針對我們設計之硬體架構做一完整說明。第五章為氣動式驅動器的動態模型推導。第六章呈現系統之完整動作。第七章是本文之結論，主要為本 PRA 的設計心得以及討論用於醫療方面的可行性。

貳、設計概念

目前在醫療復健機器人的研究上，已經被廣泛的討論。而且它們的實用性與應用性已在醫療方面得到良好的評價。在這些研究中主要的問題是在於“機械的柔軟度”，這是發展這類型機構的重點，醫療復健機器人的出現在人群之中是自然且可以被接受的。機械的柔軟度被廣泛的要求，必須具有下列兩特點：

- (1)可像人類般具有活動性。
- (2)可像人類般具有敏感性。

本文是根據上述特性，在機械柔軟度的想法上提出我們的設計，特別是在人類手臂上的動作。這次研究結合的範圍相當廣泛，其志在於將醫療電子技術、醫學復健、氣壓系統與機電整合的專業知識結合，製造出專門對於手臂復健的氣動式手臂。目的是希望幫助上肢活動有障礙的患者，讓患者能藉由復健器材的幫助，加強復健治療的效果。

在我們所設計的構想上，結構方面不採用以往的線控或是馬達，而是採用氣壓控制，用氣壓缸取代馬達，以減輕重量，減少患者的負擔。首先，將程式寫入 DSP 晶片中，由 DSP 晶片控制固態繼電器的 ON 或 OFF，使電磁閥將氣體導入氣壓缸，進而控制氣壓缸動作。在肩關節部分利用四組電磁閥與氣壓缸(氣壓缸尾端再與換向接頭連接)相互搭配，利用最基本的線性直線運動達成肩關節位移的動能，肘關節則用一組電磁閥與氣壓缸來完成擺幅動作，並藉由壓力感測器透過 DSP 晶片偵測患者施力的大小與位置，以便執行不同需求的調整，以達成良好的治療效果。

如果再配合專業的復健醫師來指導，更能詳盡的規劃出清楚的治療流程，相信對於患者的復健能夠有相當大的幫助。

接下來，介紹 PRA 的系統架構，架

構是由 PC、RS-232 傳輸線、DSP 晶片、固態繼電器、電磁閥與氣壓缸，以及壓力感測器和座椅機構所組合而成，來完成一些簡單基本的手臂動作。

參、系統架構

系統架構分為控制與驅動兩方面。在控制方面是由 PC 和 DSP 晶片所組成，以便進程式撰寫以及感測控制。而驅動方面則是由固態繼電器透過匯流排連接電磁閥以便驅動氣壓缸。在此藉由空氣壓縮機給予足夠的氣體，使 PRA 整體架構得到充分的氣壓來推動五組氣壓缸，即可完成所設定的基本動作。實體圖如圖 1 所示，架構中有控制與驅動電路、模特兒本體、座椅機構以及 PRA。圖 2 為整體系統控制流程圖，利用 PC 將程式透過 RS-232 傳輸線寫入 DSP 晶片，由 DSP 晶片控制固態繼電器的 ON 或 OFF(繼電器前端通入 DC 23V)，當繼電器位於 ON 時，DC 23V 通入電磁閥打開讓氣體導入氣壓缸，進而控制氣壓缸，使其動作。由繼電器端子連接於各別的電磁閥與氣壓缸，將氣壓缸設置在氣動式手臂上，利用程式編輯使其相互動作，並藉由壓力感測器透過 DSP 晶片偵測患者施力的大小與位置，以便執行不同需求的調整，即可做出與真正手臂相同的基本動作。

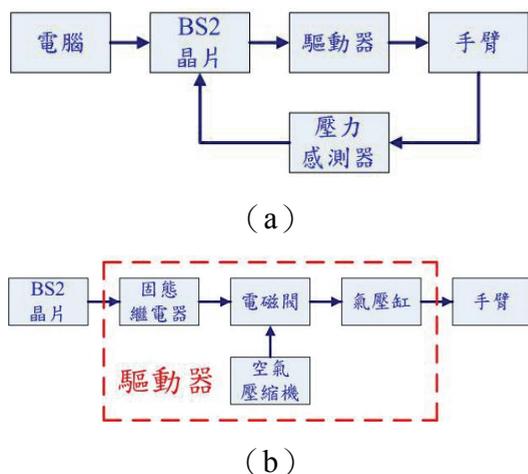


圖 2 (a) 系統控制流程圖，
(b) 驅動器內部方塊圖

接下來，我們將介紹此 PRA 的硬體架構。

肆、硬體架構

首先，我們針對機構設計之硬體架構來做探討。圖 3 為 PRA 的外部整體架構圖，架構中有氣壓缸、空氣壓縮機以及座椅機構等，接著將對各機構加以詳細說明。



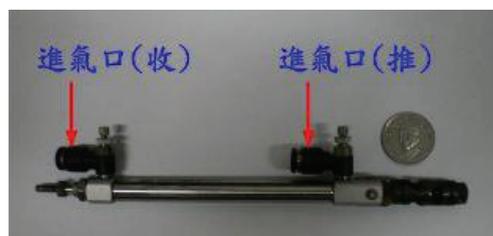
圖 3 PRA 的外部整體架構圖

4.1 氣壓缸

PRA 最重要的致動元件是氣壓缸，本文是用氣壓缸取代馬達的笨重感。由於

PRA 將氣壓缸固定在人體的手臂上。因此，氣壓缸與手臂的重量是要考慮的因素。而使用氣壓缸的優點為可以使用較小的氣壓來推動較大的重量，同時其兼具了體積小、重量輕的優勢，所以選擇氣壓缸來設計。氣壓缸可分為單動式、雙動式、無接觸感應式及無桿氣壓缸，我們實驗所採用的是雙動單桿氣壓缸，壓縮空氣可分別供給活塞兩側，並藉助電磁閥如圖 9

(a) 所示，改變氣壓缸的前進與返回，此種最常被採用。本研究所使用的氣壓缸為 MCMJ，如圖 4 (a) 所示。此款氣壓缸特點：(1) 免給油 (2) 高品質；耐久佳 (3) 多樣化支架 (4) 氣密性佳。圖 4 (b) 為氣壓缸的伸展長度。



(a)



(b)

圖 4 (a) 氣壓缸實體圖，
(b) 氣壓缸的伸展長度

表 1 為我們所使用之 MCMJ 氣壓缸的產品規格。

表 1 MCMJ 氣壓缸規格表

氣壓缸內徑	6、10、16
配管口徑尺寸	M5*0.8
使用流體	氣壓
最大使用壓力	7 kgf/cm ²
最小使用壓力	單動：引入型 -1.5~2.5 kgf/cm ² ，押出型 -1.5~2.0 kgf/cm ² 複動：0.6~1.2kgf/cm ²
耐壓力	10 kgf/cm ²
周圍溫度	-5°C~+60°C (不凍結)
感應開關	RCM、RCS
感應開關束帶	BJ6~BJ16、BM10~BM16

4.2 空氣壓縮機

我們所設計之 PRA 為氣壓控制系統，因此將利用空氣方式來傳達能量，其特性是運用氣體來轉換能量及傳遞動力。透過空壓機將機械能轉換成氣體能，在對此氣體能做壓力、方向以及流量控制，再藉由氣壓缸將氣體能轉換成已控制好的出力大小、動作方向與運動速度之機械能。

整體而言，因為經過多次的能量轉換（機械能→氣體能→機械能），故氣壓本身並不是一種高效率之系統。但由於氣壓能夠提供一種高密度之能量，又可以不經過齒條、螺桿或機構轉換而直接提供快

速、順暢、有力之旋轉、擺動或直線往復運動來作功，對過負載有相當的安全性，而且出力及速度容易控制等等之特性，使氣壓成為目前工業界所廣泛使用的動力來源。

本實驗所採用的空壓機規格為 AC 110V 1 hp，如圖 5 所示。



圖 5 空壓機實體圖

由於空壓機在運轉過程中會產生噪音，在此我們只需要加裝消音器或隔音箱，就能有效降低噪音問題。另外，透過空壓機的儲存輸送方便之特性，將空壓機設置於遠處，只需將動力配管導至需要之處，就能排除噪音問題。因此，在我們所設計的 PRA 中，空壓機的噪音問題是能夠被有效解決的。

4.3 座椅機構

復健治療是種相當費時且耗費人力的工作。因此，我們考量到患者的體能狀

況，通常不適合長時間久站，所以特別為患者設計一套可調式座椅來搭配 PRA，讓患者能夠在體能無負擔的狀況下進行復健，達到良好的治療效果。座椅機構設計圖，如圖 6 所示。

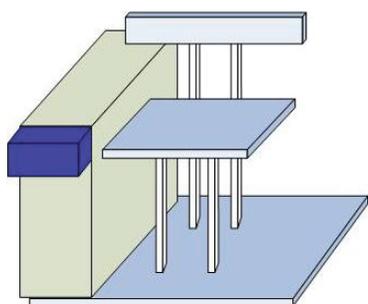


圖 6 座椅機構設計圖

我們所設計的 PRA 架設於座椅的左邊，控制箱架設於座椅的右邊，以方便大多數患者使用右手進行調整控制。由於每位患者的體型不盡相同，所以藉由椅背來調整高度，配合不同體型患者的手臂位置，達到可調式座椅的效用。座椅機構實體圖，如圖 7 所示。



圖 7 座椅機構實體圖

介紹完 PRA 的機構設計之後，接著來談控制箱內部的電子設計之硬體架構。

4.4 DSP 晶片

本研究採用 BASIC Stamp 2 晶片。如圖 8 (a) 所示。它是 Parallax 公司 BASIC Stamp DSP 系列中的一種微處理器。此系列共有的特點如下：

- (1) 功能強，操作簡單，應用整合領域廣的微處理器控制晶片。可支援涵蓋所有機器人的應用範圍，包括各式感測器，馬達與控制器，機器人套件，通訊與無限運用領域。
- (2) BASIC Stamp 微處理器整合 PLC，多工控制器及單晶片的處理功能，具備 Visual Basic，Java，Assembly 等語言編譯功能，免除使用單晶片系列工作時所需的周邊裝置，如模擬器、燒錄器等。
- (3) 具備 EEPROM（程式儲存），最多可支援 4000 行指令（8 x 2K Bytes）；處理速度可達到 12000 行/秒。

由於控制命令必須透過電腦來與 DSP 晶片溝通。因此，使用 RS-232 傳輸線來連接電腦與 DSP 晶片。圖 8 (b) 便顯示出 DSP 晶片如何透過 RS-232 傳輸線與電腦溝通的接腳說明。



(a)

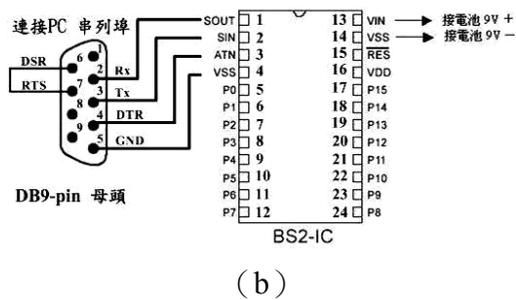


圖 8 (a) DSP 晶片圖, (b) DSP 晶片與 RS-232 傳輸線串接之接線圖

經由了解 RS-232 傳輸線以及 DSP 晶片的各項接腳以及功能,所規劃的各項控制命令便可以輕易的透過 PC 傳送給控制器來控制電磁閥,進而達成控制 PRA 的目的。

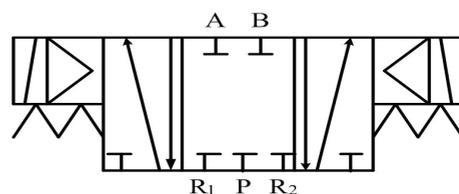
4.5 電磁閥

電磁閥是氣壓缸重要的輸出控制單元,利用氣壓缸的兩個出氣口來變換手臂的伸展位置,調整電磁閥的輸出大小,決定氣壓缸的伸展長度。電磁閥可分為單電磁線圈電磁閥 (single solenoid valve) 及雙電磁線圈電磁閥 (double solenoid valve) 兩類,它是以電磁線圈的磁力,操作換向之方向控制閥。最常用於啟動、感測、停止與變換壓縮空氣的方向。電磁閥所具有的優點為：(1) 反應速度快 (2) 可遠距離控制。氣壓系統中使用單線圈電磁閥或雙線圈電磁閥,需視實際工作需求與使用場所而定。原則上,使用單線圈電磁閥時,為了使用電氣訊號保持,在電氣迴路

上就須增加自保持的迴路,這使得電氣迴路的設計複雜許多。我們所使用的電磁閥為雙線圈電磁閥。圖 9 (a) 為五口二位電磁閥實體圖,圖 9 (b) 為五口二位電磁閥示意圖。五口二位中第一個數字五代表有五個進出氣口 (各有英文代號),第二個數字二代表有兩個工作位置。圖 9 (b) 中 A、B 為二個出氣口,而 R1、P、R2 三口中,P 為進氣口,R1 為 A 的排氣,R2 為 B 的排氣。當電源未通電時,A 為常開,當 P 進氣時,會由 A 出氣,B 常閉。當通電時,磁簧會做出吸附的動作,出氣口會切換成 B,A 原先的氣體,會從 R1 排出。因此,透過控制電磁閥之開關,即可達到控制氣壓缸之伸縮。



(a)



(b)

圖 9 (a) 五口二位電磁閥實體圖, (b) 五口二位電磁閥示意圖

表 2 為我們所使用之 MVSD-180 電磁閥的產品規格。

表 2 MVSD-180 電磁閥規格表

本體寬幅	18mm
配管口徑代號	6A
閥種類	五口二位
本體型式	6 直配管型
配管口徑尺寸	PT1/8
使用流體	空氣
使用壓力範圍	1.5~8kgf/cm ² , 2~7kgf/cm ² , (5/3)
耐壓力	10kgf/cm ²
流量	533 NI/min , 400 NI/min (5/3)
有效斷面積	9mm ² ~12mm ²
應答時間	30ms
周圍溫度	-5°C~+50°C (不凍結)
定格電壓	AC110,220V,DC24V
電力消耗	AC=3.7/3.1VA,DC=2.9W
電壓容許範圍	±10%
絕緣等級	F 級
重量	120g~190g

4.6 壓力感測器

壓力感測器廣泛的運用在機械設計結構中，作為基本的感測元件。通常市面常見的壓力感測元件多為剛性結構。我們實驗所採用的是 Tekscan 公司生產之撓性壓力感測器 (Flexiforce Sensor) [11]。Flexiforce Sensor 屬於壓阻效應型的壓力感測器，其優點包括負重與電阻變化具有

高度線性關係，感測器本體非常薄 (0.127mm) 且具有可撓性，屬於薄片式感測器，如圖 10 所示。



圖 10 Flexiforce Sensor 實體圖

Flexiforce Sensor 依據負載大小分成 Low (25 lbf)、Medium (150 lbf)、High (1000 lbf) 三種等級，而 Low 等級中又可細分為 1lbs、25lbs、以及 100lbs 三種規格。本實驗測試所採用的規格為 100lbs 規格。詳細規格如表 2 所示。

表 3 Flexiforce Sensor 規格表

規格	
額定電壓	DC 5V
長度	203 mm
寬度	14 mm
厚度	0.127 mm
感測區域	9.53mm
線性誤差	< +/-5%
重覆性	< 滿刻度的 ±2.5%
遲滯現象	< 滿刻度的 4.5 %
漂移	< 3% /相對時間
上升時間	< 20 μsec
操作溫度	15°F - 140°F (-9°C - 60°C) *

在此，我們將 Flexiforce Sensor 分別安裝於 PRA 的上臂與手肘，如圖 11 所

示。當 Flexiforce Sensor 感知患者受到機構帶動時所反應的力量，透過 DSP 晶片做處理，進而控制固態繼電器的 ON 或 OFF，使 PRA 的動作依照程式指令進行修正，如圖 2 所示。

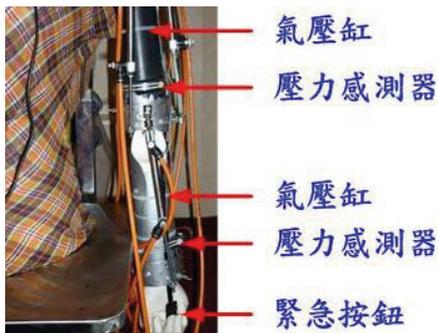


圖 11 Flexiforce Sensor 安裝位置圖

Flexiforce Sensor 屬於壓阻效應式壓力感測器，在無負載的情況下，電阻值非常大（約 $20M\Omega$ ）。隨著負載的增加，電阻值開始下降。藉由不同負載時產生的電阻變化，配合使用電路，轉換成電壓變化，以便得到電壓—負載關係圖，如圖 12 所示。

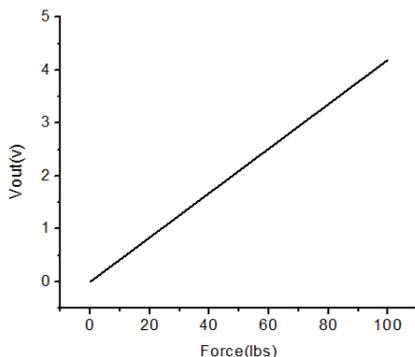


圖 12 輸出電壓—負載關係圖

我們使用圖 12 之電壓與負載關係圖來建

立查表，以提供患者在使用設備時所反應在設備上之力量所對應出的電壓關係。這個關係可以讓設備提供一個可使患者更加舒適的復健動作。

4.7 緊急按鈕

由於我們考量到患者在使用 PRA 時，所產生的安全性問題。因此，在機構上設置緊急按鈕，如圖 11 所示。利用薄膜開關安裝於 PRA 的手腕上，當患者感覺到疼痛時，只需按下緊急按鈕，PRA 的動作就會立刻停止；當患者疼痛減輕後，只需按下重置鍵，PRA 的動作就會回到初始狀態。

透過壓力感測器及緊急按鈕，這兩種安全防護的機制，因此可以提升我們所設計的 PRA 之安全性。

簡單介紹完硬體架構之後，在下一章節我們所要探討的是氣動式驅動器的動態數學模型推導，經由此方程式表示式，讓我們進一步了解氣壓缸動態與方程式表示式之間的關係。

伍、氣動式驅動器的動態數學模型

由於我們所設計的氣動式手臂最主要的致動元件為氣壓缸。因此，在本章中我們的目的是要取得氣壓缸動態模型的方程式表示式。由於我們的系統具有五組完全相同之雙動式單桿氣壓缸，它們是

經由五口二位電磁閥導入氣體，所以我們僅需針對其中一組氣壓缸與電磁閥來討論其動態模型。最後以求得五組完整之氣壓缸數學模型。

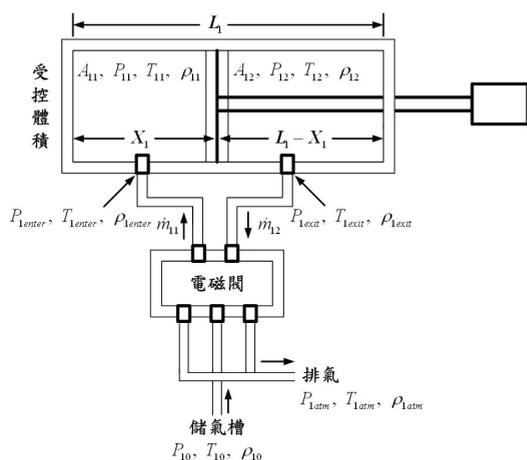


圖 13 雙動式單桿氣壓缸和電磁閥示意圖

因為氣壓缸的伸展長度變化即是改變缸體內部之體積大小。根據氣壓缸之示意圖（如圖 13），第一個驅動器的缸體內部可被分為氣室 1 和氣室 2。其中氣室 1 為進氣氣室，氣室 2 為出氣氣室，根據熱力學第一定律可以表示如下[2,12,13]：

$$\dot{Q}_1 + \dot{m}_{11} \left(h_{1enter} + \frac{v_{1enter}^2}{2} \right) = \dot{m}_{12} \left(h_{1exit} + \frac{v_{1exit}^2}{2} \right) + \frac{\partial E_1}{\partial t} + \dot{W}_1 \quad (1)$$

\dot{Q}_1 ：第一個驅動器的受控體積之熱效率；

\dot{W}_1 ：第一個驅動器的受控體積到活塞組件之輸出功率；

$(\partial E_1 / \partial t)$ ：第一個驅動器的受控體積之總能量變化率；

\dot{m}_{11} ：第一個驅動器的質量流率流入受控

體積；

\dot{m}_{12} ：第一個驅動器的質量流率流出受控體積；

h_{1enter} ：第一個驅動器的氣體流入氣室 1 之焓；

v_{1enter} ：第一個驅動器的氣體流入氣室 1 之速率；

h_{1exit} ：第一個驅動器的氣體流出氣室 2 之焓；

v_{1exit} ：第一個驅動器的氣體流出氣室 2 之速率。

接下來，我們將討論方程式（1）中之各項元素是如何決定的。

5.1 $(\partial E_1 / \partial t)$ 的推導

由於氣壓缸內受控體積的動能與位能變化率比起其體積內的能量變化率是微小的。因此，可以將其忽略不計。得到第一個驅動器的受控體積之總能量變化率為：

$$\frac{\partial E_1}{\partial t} = \frac{\partial (U_{11})}{\partial t} + \frac{\partial (U_{12})}{\partial t} \quad (2)$$

假設系統中的氣體為理想氣體， U_{11} 和 U_{12} 是氣室 1 和氣室 2 的內部能量，可被分別定義為方程式（3）和（4），

$$U_{11} = C_v \rho_{11} V_{11} T_{11} \quad (3)$$

$$U_{12} = C_v \rho_{12} V_{12} T_{12} \quad (4)$$

C_v 是定容比熱， ρ_{11} 、 V_{11} 、 T_{11} 、 ρ_{12} 、 V_{12} 、

T_{12} 分別為氣室 1 和氣室 2 的氣體密度、氣體體積，以及氣體溫度。

由於理想氣體定律可寫成如下：

$$\rho T = \frac{P}{R} \Rightarrow P = \rho RT \quad (5)$$

在此， P 是理想氣體壓力， ρ 是理想氣體密度， R 是理想氣體常數， T 是理想氣體溫度。根據方程式 (5)，方程式 (3) 和 (4)，可分別被改寫如下：

$$U_{11} = \left(\frac{C_V}{R} \right) P_{11} V_{11} \quad (6)$$

$$U_{12} = \left(\frac{C_V}{R} \right) P_{12} V_{12} \quad (7)$$

此處的 P_{11} 和 P_{12} 是氣室 1 和氣室 2 的內部氣壓。接下來，把方程式 (6) 和 (7) 代入到方程式 (2) 中，取代 U_{11} 和 U_{12} 。則第一個驅動器的受控體積之總能量變化率可如下所示。

$$\begin{aligned} \frac{\partial E_1}{\partial t} &= \frac{\partial \left(\frac{C_V}{R} P_{11} V_{11} \right)}{\partial t} + \frac{\partial \left(\frac{C_V}{R} P_{12} V_{12} \right)}{\partial t} \\ &= \left(\frac{C_V}{R} \right) (\dot{P}_{11} V_{11} + P_{11} \dot{V}_{11}) + \left(\frac{C_V}{R} \right) (\dot{P}_{12} V_{12} + P_{12} \dot{V}_{12}). \end{aligned} \quad (8)$$

5.2 \dot{W}_1 的推導

氣體經由活塞組件影響，所造成的功率變化為：

$$\dot{W}_1 = P_{11} \dot{V}_{11} + P_{12} \dot{V}_{12} \quad (9)$$

\dot{V}_{11} 和 \dot{V}_{12} 分別為氣室 1 和氣室 2 的體積變化率。

5.3 輸入輸出焓的推導

氣體從儲氣槽進入氣壓缸(通常儲氣槽連接至空氣壓縮機)由於氣體在儲氣槽的速率為零，其焓表示為滯焓 h_{10} 。方程式 (10) 描述 h_{10} 和氣體流入氣室 1 之間的關係為：

$$h_{1enter} + \frac{v_{1enter}^2}{2} = h_{10} = C_p T_{10} \quad (10)$$

T_{10} 是氣體在儲氣槽的溫度。 C_p 是定壓比熱，經由 $C_p = C_V + R$ ，可知 C_p 與上述的理想氣體常數是有關聯的。同樣地，氣體在氣室 2 的速率較小時，比較氣體流出通過電磁閥的速度 (v_{1exit})。假設，方程式 (10) 符合氣室 2 的焓 h_{12} ， h_{12} 到氣體流出電磁閥之間的關係可表示為：

$$h_{1exit} + \frac{v_{1exit}^2}{2} = h_{12} = C_p T_{12} \quad (11)$$

5.4 方程式 (1) 的重新推導

我們可將方程式 (8) 取代方程式 (1) 中之 $(\partial E_1 / \partial t)$ ，再將方程式 (9)、(10)、(11) 代入到方程式 (1) 中，重新推導可得：

$$\begin{aligned} \dot{m}_{11}(C_p T_{10}) = \dot{m}_{12}(C_p T_{12}) + \left(\frac{C_v}{R}\right)(\dot{P}_{11}V_{11} + \dot{P}_{12}V_{12}) \\ + \left(1 + \frac{C_v}{R}\right)(P_{11}\dot{V}_{11} + P_{12}\dot{V}_{12}). \end{aligned} \quad (12)$$

由於 $C_p = C_v + R$ ，所以可簡化方程式 (12) 為：

$$\begin{aligned} \dot{m}_{11}T_{10} = \dot{m}_{12}T_{12} + \left(\frac{C_v}{RC_p}\right)(\dot{P}_{11}V_{11} + \dot{P}_{12}V_{12}) \\ + \left(\frac{1}{R}\right)(P_{11}\dot{V}_{11} + P_{12}\dot{V}_{12}). \end{aligned} \quad (13)$$

假設，極少的熱交換發生在驅動器與周邊環境間，則方程式 (14) 可代表第一個驅動器的熱力學第一定律，如下所示：

$$\begin{aligned} \dot{m}_{11}(RT_{10}) = \dot{m}_{12}(RT_{12}) + \left(\frac{1}{k}\right) \\ (\dot{P}_{11}V_{11} + \dot{P}_{12}V_{12}) + (P_{11}\dot{V}_{11} + P_{12}\dot{V}_{12}). \end{aligned} \quad (14)$$

其中 $k = C_p / C_v$ 。

如圖 13 所示，最左側為 ($X_1 = 0$) 和 L_1 為氣壓缸的行程總長度，所以氣室體積及氣室體積變化率為：

$$\begin{cases} V_{11} = A_{11}X_1 \\ V_{12} = A_{12}(L_1 - X_1) \end{cases} \Rightarrow \begin{cases} \dot{V}_{11} = A_{11}\dot{X}_1 \\ \dot{V}_{12} = -A_{12}\dot{X}_1 \end{cases} \quad (15)$$

將方程式 (15) 代入到方程式 (14) 可得：

$$\begin{aligned} \dot{m}_{11}(RT_{10}) - \dot{m}_{12}(RT_{12}) = \left(\frac{1}{k}\right)(\dot{P}_{11}A_{11} - \dot{P}_{12}A_{12})X_1 \\ + \left(\frac{1}{k}\right)\dot{P}_{12}A_{12}L_1 + (P_{11}A_{12} - P_{12}A_{12})\dot{X}_1. \end{aligned} \quad (16)$$

根據方程式 (16)，我們便可調整不同質量流率 \dot{m}_{11} 和 \dot{m}_{12} 的電磁閥開口來獲得

X_1 。換句話說，改變 \dot{m}_{11} 和 \dot{m}_{12} 便可以直接控制第一個驅動器的活塞桿位置，進而得到不同的氣壓缸伸縮長度。而整體五組氣壓缸與電磁閥之動態模型可表示成方程式 (17)。

$$\begin{aligned} \mathbf{RT}_0\dot{\mathbf{M}}_1 - \mathbf{RT}_2\dot{\mathbf{M}}_2 = \left(\frac{1}{\mathbf{k}}\right)(\dot{\mathbf{P}}_1\mathbf{A}_1 - \dot{\mathbf{P}}_2\mathbf{A}_2)\mathbf{X} \\ + \left(\frac{1}{\mathbf{k}}\right)\dot{\mathbf{P}}_2\mathbf{A}_2\mathbf{L} + (\mathbf{P}_1\mathbf{A}_1 - \mathbf{P}_2\mathbf{A}_2)\dot{\mathbf{X}} \end{aligned} \quad (17)$$

在方程式 (17) 當中，我們將繼電器與電磁閥視為理想元件，以簡化推導。在此，假設我們所設計的氣動式驅動器是處於理想環境下，所以不加以考慮驅動器的延遲時間以及摩擦係數。本章節主要目的是為了解氣壓缸伸縮與氣體流量之關係，以便進行氣壓缸伸展長度的調整，而達到有效的手臂位置之控制。

陸、實驗動作

我們將動作流程設計好並使用電腦撰寫程式，透過 RS-232 傳輸線寫入 DSP 晶片，由 DSP 控制固態繼電器的 ON 或 OFF，當繼電器位於 ON 時，DC 23V 通入電磁閥，此時電磁閥打開讓氣體導入氣壓缸，進而控制氣壓缸，使其動作。由繼電器端子連接五組電磁閥與氣壓缸，將氣壓缸裝置在復健手臂上，達到需求的動作。圖 14 (a) 到 (e) 為手臂做平舉以及側彎之實驗圖。



(a)



(b)



(c)



(d)



(e)

圖 13 (a) 初始狀態 (放下), (b) 手肘上舉, (c) 手肘上舉後放下, (d) 將手臂平舉, (e) 回到初始狀態

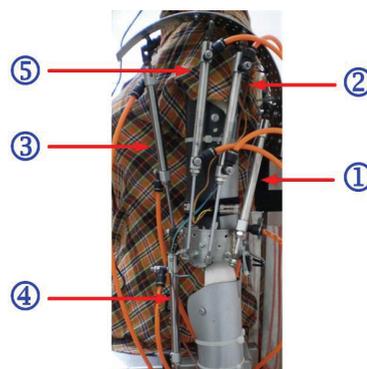


圖 15 氣壓缸位置編號圖

由於我們設計的 PRA 還在測試階段，基於安全考量，尚未對人體進行相關實驗。在此我們使用了人體模特兒進行動作實驗，模特兒本體有相當程度的靈活性，故能模擬真正手臂動作。

圖 14 (a) 到 (e) 為復健治療動作所模擬出的手臂執行動作。接著我們將介紹氣壓缸變換的動作過程，由圖 15 來對照氣壓缸編號。圖 14 (a) 為 PRA 的初

始狀態，在未供電的情況下，氣壓缸 ① 呈現收縮狀態，其餘四支氣壓缸為伸展狀態。圖 14 (b) 為電源供應器提供 DC 23V 通入電磁閥，此時電磁閥打開讓氣體導入氣壓缸，氣壓缸 ④ 由伸展轉為收縮，其餘氣壓缸不變動，手肘慢慢往上抬。圖 14 (c) 由 DSP 控制固態繼電器並藉助電磁閥改變氣壓缸 ④，從收縮轉為伸展，其餘氣壓缸不變動，做出手肘上舉後放下，回歸初始狀態的動作。圖 14 (d) 為氣壓缸 ① 由收縮轉為伸展，氣壓缸 ②③⑤ 由伸展轉為收縮，其餘那支缸不變動，手臂慢慢往上抬。圖 14 (e) 為氣壓缸 ① 縮回，氣壓缸 ②③⑤ 伸展，返回初始狀態。經由上述之動作規劃，以及程式反覆執行，即可達到重覆使用之目的。同時，亦可規劃出不同需求之復健動作來透過本 PRA 實現。

柒、結論

本文以設計 PRA 為目標，希望能將復健手臂整合氣壓控制系統以取代傳統馬達，再將整套系統運用在復健醫學中，透過此機器人引導患者動作，以達到復健之目的。

本研究採用 BASIC Stamp 2 晶片，該晶片整合了 PLC (可程式控制器)，多工控制器及單晶片的處理功能，具備 EEPROM (程式儲存) 之能力，功能強大

且操作簡單。整體架構由五組氣壓缸來推動 PRA，利用程式編輯使其相互動作，即可做出與真正手臂相同的基本動作。同時，在 PRA 上安裝壓力感測器來感知患者受到機構帶動時的力量，以提供系統做出最適合患者之動作。

我們亦建立了一組完整的氣動式驅動器的動態模型，經由模型推導我們得到此 PRA 驅動器的方程式表示式。透過此方程式，我們可以適當改變氣動式驅動器的氣體質量來改變氣壓缸的伸展長度進而得到 PRA 的動作。

本文主要目的為設計一款氣動式手臂，而我們所設計的 PRA 目前只是設計雛型，經由雛型的實驗來證實 PRA 用於復健治療是可行的，可使患者手臂達到有效的復健功能。至於 PRA 的機構動態模型推導，以及量化的實驗數據，將列為未來的研究重點。

參考文獻

- [1] 許銘全、陳建銘，氣動式復健手臂，第四屆智慧生活科技研討會，2009。
- [2] H. Kazerooni, "Design and Analysis of Pneumatic Force Generators for Mobile Robotic Systems," *IEEE/ASME Trans. on Mechatronics*, vol. 10, pp. 411-418, 2005.
- [3] S. Liu and J. E. Burrows, "An analysis of a pneumatic servo system and its

- application to a computer-controlled robot,” *J. Dyn. Syst., Meas., Contr.*, vol. 110, pp. 228–235, 1988.
- [4] J. E. Bobrow and F. Jabbari, “Adaptive pneumatic force actuation and position control,” *J. Dyn. Syst., Meas., Contr.*, vol. 113, pp. 267–276, Jun. 1991.
- [5] D. Ben-Dov and S. E. Salcudean, “A force-controlled pneumatic actuator,” *IEEE Trans. Robot. Autom.*, vol. 11, no. 6, pp. 906–911, Dec. 1995.
- [6] J. C. Rosier, J. A. van Woerden, L.W. van der Kolk, B. J. F. Driessen, H. H. Kwee, J. J. Duimel, J. J. Smits, A. A. Tuinhof de Moed, G. Honderd, and P. M. Bruyn, “Rehabilitation robotics : The MANUS concept,” in *Proc. 5th Int. Conf. Advanced Robotics*, Pisa, Italy, 1991, pp. 893–898.
- [7] R. J. Sanchez, Jr., E. Wolbrecht, R. Smith, J. Liu, S. Rao, S. Cramer, T. Rahman, J. E. Bobrow, D. J. Reinkensmeyer, “A Pneumatic Robot for Re-Training Arm Movement after Stroke : Rationale and Mechanical Design,” *9th International Conference on Rehabilitation Robotics*, pp. 500-504, June 28 - July 1. 2005.
- [8] T. Nef, M. Mihelj, R. Riener, “ARMin – A Robot for Patient-Cooperative Arm Therapy,” *Medical and Biological Engineering & Computing*, vol. 45, no. 9, pp. 887-900, 2007.
- [9] 臺大醫院醫學工程部，醫療儀器發展新知，[http : //www.ntuh.gov.tw/BMED/equipment/DocLib/%E9%A6%96%E9%A0%81.aspx](http://www.ntuh.gov.tw/BMED/equipment/DocLib/%E9%A6%96%E9%A0%81.aspx)。
- [10] 林棟煌，“肘關節神經復健用機器人之改良和臨床測試”，國立成功大學機械工程系碩士論文，2001。
- [11] 田朝元，“以 PC 為基礎之頸椎牽引機力量感測器設計”，逢甲大學電機工程學系碩士論文，2007。
- [12] 王豐、吳黎民、吳順治、林振森、林進益，熱力學概論，大揚出版社，1997。
- [13] 翁文爐、夏一民編譯，化工熱力學，高立圖書有限公司，2001。

Nonlinear rupture theory of a thin micropolar liquid film on a cylinder under a magnetic field

Hung Ming Sung, Chung-Ting Hsu

Abstract

This paper studies the nonlinear rupture process of the thin micropolar liquid film on a cylinder under the influence of the uniform magnetic field. The long-wave perturbation method is employed to derive the generalized kinematic equations with free film surface conditions. The generalized nonlinear evolution equation is solved numerically. The effects of magnetic field, the van der Waals potential, the surface tension, the micropolar parameter, K , and the radius of the cylinder on the rupture behavior are extensively investigated. The results show that the film rupture time will be delayed by increasing the associated parameters including the Hartmann number, the micropolar parameter, and the radius of the cylinder, or to decrease the initial disturbance.

Keywords : micropolar, magnetic field, rupture, thin film, surface tension, van der Waals potential.

直立圓柱表面微極流體薄膜 在磁場效應影響下之非線性破裂理論

宋鴻明、徐仲亭

摘要

本研究探討在磁場效應的影響下，直立圓柱表面微極流體薄膜的非線性破裂現象。首先使用長波微擾法推導薄膜的自由面方程式，利用電腦數值計算方法解薄膜的自由面方程式，研究微極流體薄膜受定值且均勻的磁場影響下，在直立圓柱表面的非線性破裂機制，分析磁場效應的大小、表面張力、凡得瓦爾勢能及圓柱半徑大小對破裂機制的影響，另外針對微極流體之微極參數 K 對薄膜破裂機制的影響，做深入的探討與分析。結果顯示，隨著磁場無因次參數、微極參數及圓柱半徑之增加，或減少初始擾動，將會延緩薄膜之破裂。

關鍵詞：微極流體、磁場、破裂、薄膜、表面張力、凡得瓦爾勢能。

1. Introduction

The hydrodynamic stability of a thin film flow is a subject of great importance commonly needed in biomedical, mechanical, and chemical engineering industries. The liquid film will become unstable when the film is thinner than 1000 \AA [1]. The instability is associated with a new term, called van der Waals potential, which results in the rupture of the film. The rupture of liquid films occurs in many industrial applications involving disperse and colloid systems and in biological phenomena [2]. It is known that the macroscopic rupture process will influence the film flow significantly. It is thus highly desirable to understand the underlying film flow characteristics and the associated time-dependent properties so that suitable conditions for homogeneous film growth can be developed for different industrial applications.

Ruckenstein and Jain [3] studied the spontaneous rupture of a liquid film on a planar solid wall. They modeled the liquid film as a Navier-Stokes continuum with an extra body force due to the van der Waals attractions and used quasi-steady lubrication approximation to obtain the results of linear dynamic instability. From their results, one can obtain rough estimates for the rupture time. Williams

and Davis [4] examined the nonlinear evolution equation and numerically treated it as an initial value problem with periodic boundary conditions. The results indicated that the nonlinearities of the system would accelerate the rupture process. Hwang and Chang [5] analyzed the film rupture problem of a power-law non-Newtonian fluid. The results revealed that the rupture time of the film decreases with the decreasing magnitude of the power-law exponent, n . However, the mode of the most unstable disturbance would not be changed by the value of the index n . Prokopiou and Cheng [6], Alekseenko et al. [7] and Hwang et al. [8] applied the integral technique with a specified velocity profile to obtain a system of nonlinear evolution equations. These equations were used to simulate the nonlinear rupture processes. The numerical analysis results obtained from the model was systematically compared with those obtained by using long-wave model. Chen and Hwang [9] studied the nonlinear rupture theory of a thin liquid film on a cylinder. They showed that the decrease of a cylindrical radius will accelerate the rupture process.

The vast majority of studies on thin-film flow problems were devoted to Newtonian fluids. The film flow of non-Newtonian fluids has attracted only

modest attention. Apparently, most of the above studies were addressed on the Newtonian fluid model. However, the rheological behavior of the fluid during the manufacturing process or the glue in biological chemistry cannot be completely described by the Newtonian fluid. In recent years, the general research interest on stability problems has been directed toward the microstructure of fluid flows. A subclass of these fluids was named micropolar fluids by Eringen [10] who first proposed the theory of micropolar fluids. Micropolar fluids exhibit certain microscopic effects arising from the local structure and micromotion of the fluid element. In application, the micropolar fluids may be used to model some man-made fluids, such as the polymeric fluids, animal blood, fluids with additives, and liquid crystals, etc. The extension of the theory of micropolar fluids to cover the thermal effect was developed by Eringen [11]. Liu [12] studied the instability of micropolar fluids. He found that the initiation of instability is delayed due to the presence of particles in the fluid. Hung et al. [13] employed the methods of nonlinear analysis to study the stability of the thin micropolar liquid film flowing down along a vertical plate. Their studies indicated that the micropolar parameter K plays an important role in stabilizing a film flow.

They also showed that the micropolar coefficients, $\Delta(=h_0^{*2}/J)$ and $\Lambda(=\gamma/\mu J)$, have very little effect on the stability of the micropolar film. Hung and Tsai [14] studied the rupture of a thin micropolar liquid film. The results displayed that the rupture time of the film increases as the vortex viscosity of micropolar fluid increases. Cheng et al. [15] studied the nonlinear stability analysis of the thin micropolar liquid film flowing down on a vertical cylinder. They found that the film flow can become relatively more stable by increasing either the micropolar parameter K or the radius of the cylinder.

Hsieh [16] studied the linear stability problem of fluid flowing down an inclined plane with a transversely applied magnetic field. It is found that the presence of the magnetic field tends to stabilize the system. Renardy and Sun [17] investigated the linear stability of a viscous fluid film flowing down an inclined plane under the effects of gravity and the tangential magnetic field. The results indicated that the presence of the magnetic field has stabilizing effects on both the film surface and the shear modes. That results in a relatively more stable film flow condition. Tsai et al. [18] studied the nonlinear hydromagnetic stability of a condensation

film flowing down a vertical plate. They showed that the magnetic field can stabilize the film flow. Hung and Tsai [19] analyzed the film rupture problem under the magnetic field. They showed that the enhancement of the magnetic effect can delay the rupture time. Coons et al. [20] made the predictions from the bounding scaling equations and it is shown that the retarded Hamaker constants derived from approximate Lifshitz theory underestimate the critical thickness of foam and emulsion films.

Detailed film flow analysis is, indeed, of great importance for the development of rupture theory in characterizing various film flow conditions. The rupture analysis of film flow, however, is very essential for both theoretical development and practical applications in disperse, colloid, and biological systems. To our best knowledge, the rupture of the thin micropolar liquid film on a cylinder under the influence of the magnetic field has not been comprehensively investigated so far. Since the types of film rupture problems are so important in many practical applications that the behavior of the rupture of thin micropolar liquid films on a cylinder under the magnetic field should be more carefully studied. This study employs the newly developed advanced numerical schema to investigate the onset of the film rupture and

also the influence factors of the applied magnetic field, the micropolar parameter, the initial disturbance, and the radius of the cylinder. In an attempt to verify computational results and to illustrate the effectiveness of the proposed modeling approach, several numerical examples are presented.

2. Analysis

2.1 Basic electromagnetic concepts

It is notable that an electrical conductor moving in a magnetic field generates an electromotive force that is proportional to the speed of motion and the strength of the magnetic field. The field of magnetohydrodynamics is complicate, since it involves the solutions of both the Navier-Stokes equations for fluid flows and the Maxwell's equations for a magnetic field. In magneto fluid mechanics, Maxwell's equations are presented as [21]

$$\nabla \cdot B = 0 \quad (1)$$

$$\nabla \cdot D = 0 \quad (2)$$

$$\nabla \times H = J_e \quad (3)$$

$$\nabla \times E = -\frac{\partial B}{\partial t} \quad (4)$$

The magnetic flux B and dielectric displacement field D are expressed by

$$B = \mu_e H , \quad (5)$$

$$D = \varepsilon E, \quad (6)$$

where J_e is the current density, μ_e is the magnetic permeability and H is the magnetic fields. By Ohm's law, the total current flow can be defined as

$$J_e = \sigma_e (E + U \times B), \quad (7)$$

where U is the velocity vector and σ_e is the electrical conductivity. By combining the above equations, the magnetic induction equation can be expressed as

$$\frac{\partial B}{\partial t} = \nabla \times (U \times B) + \frac{1}{\sigma_e \mu_e} \nabla^2 B. \quad (8)$$

In momentum equation, the electromagnetic force must be included and can be expressed as

$$F_m = J_e \times B = \sigma_e (U \times B) \times B. \quad (9)$$

2.2 Generalized kinematic equation

Figure 1 shows the configuration of the thin micropolar liquid film on a vertical cylinder. A uniform magnetic field with constant strength B_0 is applied normal to the liquid film. Both the electrical double-layer and the Marangoni effect are neglected. The film thickness is assumed to be thin enough to neglect the gravity effect, while it is still under the influence of van der Waals attraction. The perturbation of film flow is assumed to be axisymmetrical so that the azimuthal component of the flow is neglected. All physical properties

are assumed to be constant. The magnetic Reynolds number is assumed to be small so that the induced magnetic field can be neglected as compared to the applied magnetic field. Based on the above-mentioned assumptions, the principles of mass, momentum and angular momentum conservation for an axisymmetric isothermal incompressible micropolar fluid leads one to a set of system governing equations. Let u^* and w^* be the velocity components in r^* and z^* directions, respectively, and N^* is the angular microrotation momentum. The governing equations for the film flow generally can be expressed in terms of cylindrical coordinates (r^*, z^*) as

$$\frac{1}{r^*} \frac{\partial(r^* u^*)}{\partial r^*} + \frac{\partial w^*}{\partial z^*} = 0, \quad (10)$$

$$\begin{aligned} & \rho \left(\frac{\partial u^*}{\partial t} + u^* \frac{\partial u^*}{\partial r^*} + w^* \frac{\partial u^*}{\partial z^*} \right) \\ &= -\frac{\partial}{\partial r^*} (p^* + \phi^*) \\ &+ (\mu + k) \left(\frac{\partial^2 u^*}{\partial r^{*2}} + \frac{1}{r^*} \frac{\partial u^*}{\partial r^*} \right. \\ &+ \left. \frac{\partial^2 u^*}{\partial z^{*2}} - \frac{u^*}{r^{*2}} \right) - k \frac{\partial N^*}{\partial z^*}, \end{aligned} \quad (11)$$

$$\begin{aligned} & \rho \left(\frac{\partial w^*}{\partial t^*} + u^* \frac{\partial w^*}{\partial r^*} + w^* \frac{\partial w^*}{\partial z^*} \right) \\ &= - \frac{\partial}{\partial z^*} (p^* + \phi^*) \\ &+ (\mu + k) \left(\frac{\partial^2 w^*}{\partial r^{*2}} + \frac{1}{r^*} \frac{\partial w^*}{\partial r^*} \right. \\ &\left. + \frac{\partial^2 w^*}{\partial z^{*2}} \right) + k \left(\frac{\partial N^*}{\partial r^*} + \frac{N^*}{r^*} \right) - w^* B_0^2, \end{aligned} \quad (12)$$

$$\begin{aligned} & \rho J \left(\frac{\partial N^*}{\partial t^*} + u^* \frac{\partial N^*}{\partial r^*} + w^* \frac{\partial N^*}{\partial z^*} \right) \\ &= \gamma \left(\frac{\partial^2 N^*}{\partial r^{*2}} + \frac{1}{r^*} \frac{\partial N^*}{\partial r^*} + \frac{\partial^2 N^*}{\partial z^{*2}} \right. \\ &\left. - \frac{N^*}{r^{*2}} \right) + \kappa \left(\frac{\partial u^*}{\partial z^*} - \frac{\partial w^*}{\partial r^*} - 2N^* \right) \end{aligned} \quad (13)$$

where ρ is the density of a fixed-value, μ is the molecule fluid viscosity, κ is the vortex viscosity, γ is the spin-gradient viscosity, p^* is the flow pressure, ϕ^* is the van der Waals potential, B_0 is the strength of magnetic field, and J is the micro-inertial density.

The fluid motion is presumably affected by (i) viscous action, which is measured by μ , (ii) the effect of couple stresses, measured by γ , and (iii) the direct coupling of the microstructure to the velocity, measured by κ . Stokes [22] showed that each of the constants μ , γ and κ , can have value greater than, or equal to, zero, so that the ratios γ/μ and κ/μ , measured respectively by the relative strengths of the couple stress to the

viscous effect and the microstructure coupling to the viscous effect, can have any value greater than, or equal to, zero. Kolpashchikov et al. [23] determined the viscosity coefficients of the micropolar fluid on the basic of viscometric and thermal measurements. The order of magnitude of the micro-inertial density, J , is comparable to l^2 , where l is a material length depending on the size of microstructure. For most suspensions, the order of magnitude of l is about 10^{-3} to 10^{-5} cm (Eringen [24]).

The boundary conditions for the film flow system at the cylinder surface of $r^* = R^*$ can be expressed as

$$u^* = 0 \quad (14)$$

$$w^* = 0, \quad (15)$$

$$N^* = \frac{1}{2} \left(\frac{\partial u^*}{\partial z^*} - \frac{\partial w^*}{\partial r^*} \right) \quad (16)$$

The boundary conditions for the film flow at free surface of $r^* = R^* + h^*$ can be given as

$$\begin{aligned} & 2 \left(\frac{\partial u^*}{\partial r^*} - \frac{\partial w^*}{\partial z^*} \right) \frac{\partial h^*}{\partial z^*} + \\ & \left(\frac{\partial u^*}{\partial z^*} + \frac{\partial w^*}{\partial r^*} \right) \left[1 - \left(\frac{\partial h^*}{\partial z^*} \right)^2 \right] = 0 \end{aligned} \quad (17)$$

$$\begin{aligned}
& p^* + 2\mu[1 + (\frac{\partial h^*}{\partial z^*})^2]^{-1} [(\frac{\partial u^*}{\partial z^*} + \frac{\partial w^*}{\partial r^*}) \frac{\partial h^*}{\partial z^*} \\
& - \frac{\partial w^*}{\partial z^*} (\frac{\partial h^*}{\partial z^*})^2 - \frac{\partial u^*}{\partial r^*}] + S^* \{ \frac{\partial^2 h^*}{\partial z^{*2}} [1 \\
& + (\frac{\partial h^*}{\partial z^*})^2]^{-3/2} - \frac{1}{r^*} [1 + (\frac{\partial h^*}{\partial z^*})^2]^{-1/2} \} = p_a^*, \quad (18)
\end{aligned}$$

$$\frac{\partial h^*}{\partial t^*} + \frac{\partial h^*}{\partial z^*} w^* - u^* = 0, \quad (19)$$

$$N^* = 0, \quad (20)$$

where p_a^* is the ambient pressure, S^* is the surface tension, and h^* is the local film thickness. The variables associated with the superscript “*” stand for dimensional quantities. The boundary conditions at the interface, Eqs. (17) and (18), are the balance of tangential and normal stresses (Edwards et al. [25]). Equation (19) is the kinematic equation at free surface. The boundary conditions for the angular microrotation momentum N^* at solid wall and at free surface, as shown in Eqs. (16) and (20), were also discussed extensively in Datta and Sastry [26] and Ahmadi[27].

In order to simplify the analysis procedure, it is customary to define dimensionless variables as

$$t = \frac{\mu}{\rho h_0^{*2}} t^* ; \quad r = \frac{r^*}{h_0^*} ; \quad z = \frac{z^*}{h_0^*} ;$$

$$u = \frac{\rho h_0^*}{\mu} u^* ; \quad w = \frac{\rho h_0^*}{\mu} w^* ;$$

$$P = \frac{\rho h_0^{*2}}{\mu^2} (P^* - P_a^*) ; \quad K = \frac{\kappa}{\mu} ;$$

$$\Delta = \frac{h_0^{*2}}{J} ; \quad S = \frac{S^* \rho h_0^*}{3\mu^2} ;$$

$$\phi = \frac{\rho h_0^{*2}}{\mu^2} \phi^* ; \quad N = \frac{\rho h_0^{*2}}{\mu} N^* ;$$

$$\Lambda = \frac{\gamma}{\mu J} ; \quad h = \frac{h^*}{h_0^*} ; \quad (21)$$

where Λ , K , and Δ are the dimensionless micropolar parameters, and h_0^* is the film thickness of local base flow. By so doing, the non-dimensional governing equations and associated boundary conditions can now be presented as

$$\frac{1}{a + \zeta} [(a + \zeta)u]_{\zeta} + w_z = 0, \quad (22)$$

$$\begin{aligned}
& u_t + uu_{\zeta} + ww_z = -p_{\zeta} \\
& + (1 + K) \left\{ \frac{1}{a + \zeta} [(a + \zeta)u_{\zeta}]_{\zeta} \right. \\
& \left. + u_{zz} - \frac{u}{(a + \zeta)^2} \right\} - KN_z, \quad (23)
\end{aligned}$$

$$\begin{aligned}
& w_t + uw_{\zeta} + ww_z = -(p_z + \phi_z) \\
& + (1 + K) \left\{ \frac{1}{a + \zeta} [(a + \zeta)w_{\zeta}]_{\zeta} \right. \\
& \left. + w_{zz} \right\} + K \left(N_{\zeta} + \frac{N}{a + \zeta} \right) - m^2 w, \quad (24)
\end{aligned}$$

$$N_t + uN_\zeta + wN_z = \Lambda \left\{ \frac{1}{a + \zeta} [(a + \zeta)N_\zeta]_\zeta + N_{zz} - \frac{N}{(a + \zeta)^2} \right\} + K\Delta(u_z - w_\zeta - 2N). \quad (25)$$

The associated boundary conditions on the cylinder surface ($\zeta = 0$) can be given as

$$u = 0 \quad ; \quad w = 0 \quad (26)$$

$$N = \frac{1}{2}(u_z - w_\zeta). \quad (27)$$

The boundary conditions at free surface of the film flow ($\zeta = h$) can also be given as

$$2(u_\zeta - w_z)h_z + (u_z + w_\zeta)(1 - h_z^2) = 0, \quad (28)$$

$$p + 2(1 + h_z^2)^{-1}[(u_z + w_\zeta)h_z - w_z h_z^2 - u_\zeta] + 3S(1 + h_z^2)^{-\frac{3}{2}}[h_{zz} - \frac{1}{a + h}(1 + h_z^2)] = 0, \quad (29)$$

$$N = 0, \quad (30)$$

$$h_t + wh_z = u, \quad (31)$$

where u and w are the dimensionless velocity components in r and z directions, respectively, p is the dimensionless pressure, N is the dimensionless angular microrotation momentum,

$m = (\sigma_e B_0^2 h_0^{*2} / \mu)^{1/2}$ is the Hartmann number. $\phi = Ah^{-3}$ [4] is the

non-dimensional van der Waals potential in which $A = A^* / (6\pi h_0^* \rho \nu^2)$, and A is the dimensionless Hamaker constant. It is noted that this potential is independent of r , and greater than zero. The r -coordinate is scaled as $r = a + \zeta$, where a is given by R^* / h_0^* . In case of $K = 0$, the vortex viscosity effect of particles in the micropolar fluid is neglected. The fluid flow thus becomes a typical classical Newtonian film flow.

Since the wavelengths of unstable disturbances are much longer than the film thickness [3], one can simplify the above equations by using the long wave perturbation method. By choosing α as the perturbation parameter for variable expansion, the new scaled variables can be defined as

$$z = \frac{\xi}{\alpha} \quad ; \quad t = \frac{\tau}{\alpha}. \quad (32)$$

The order of the parameters are also presumed as

$$A = O\left(\frac{1}{\alpha}\right) \quad ; \quad S = O\left(\frac{1}{\alpha^3}\right) \quad ; \quad u = O(\alpha) \\ w = O(1) \quad ; \quad p = O\left(\frac{1}{\alpha}\right); \\ \phi = O\left(\frac{1}{\alpha}\right) \quad ; \quad a = O\left(\frac{1}{\alpha}\right), \quad (33)$$

The variables, u , w , p , N , and ϕ can now be expanded as follows

$$u = \alpha[u_0 + \alpha u_1 + O(\alpha^2)], \quad (34)$$

$$w = w_0 + \alpha w_1 + O(\alpha^2), \quad (35)$$

$$p = \frac{1}{\alpha} [p_0 + \alpha p_1 + O(\alpha^2)], \quad (36)$$

$$N = N_0 + \alpha N_1 + O(\alpha^2), \quad (37)$$

$$\phi = \frac{1}{\alpha} \phi_0. \quad (38)$$

By plugging the above equations into Eqs. (22) - (31), the system governing equations can now be collected and solved order by order. Hung et al. [13,14] showed that the micropolar parameter $\Lambda (= \gamma / \mu J)$ has little effect on the stability of the micropolar film. By neglecting the Λ term in Eq. (25), one can obtain the equations of the zeroth-order as

$$u_{0\zeta} + w_{0\zeta} = 0, \quad (39)$$

$$p_{0\zeta} = 0, \quad (40)$$

$$p_{0\zeta} + \phi_{0\zeta} = (1 + K)w_{0\zeta\zeta} + KN_{0\zeta} - m^2 w_0, \quad (41)$$

$$K\Delta(w_{0\zeta} + 2N_0) = 0; \quad (42)$$

and the boundary conditions at $\zeta = 0$, are given as

$$u_0 = w_0 = 0, \quad (43)$$

$$N_0 = -\frac{1}{2} w_{0\zeta} \quad (44)$$

also, at $\zeta = h$,

$$w_{0\zeta} = 0, \quad (45)$$

$$p_0 = -3Sh_{\zeta\zeta} + 3Sa^{-2}(a - h), \quad (46)$$

$$N_0 = 0, \quad (47)$$

$$h_\tau + w_0 h_\zeta = u_0. \quad (48)$$

Since the higher-order solutions are of little influence on the system [4], they can be neglected. By solving Eqs. (39) - (42) simultaneously with the associated boundary conditions, the zeroth-order solutions of the film flow velocity can now be given as

$$w_0 = -\frac{3}{m^2} (Sh_{\zeta\zeta\zeta} + Sa^{-2} h_\zeta + Ah^{-4} h_\zeta) [\cosh \beta m \zeta - (\tanh \beta m h) (\sinh \beta m \zeta) - 1], \quad (49)$$

$$u_0 = \frac{3}{m^2} (Sh_{\zeta\zeta\zeta} + Sa^{-2} h_{\zeta\zeta} + Ah^{-4} h_{\zeta\zeta} - 4Ah^{-5} h_\zeta^2) \left(\frac{\sinh \beta m \zeta}{\beta m} - \frac{\tanh \beta m h}{\beta m} \cosh \beta m \zeta - \zeta + \frac{\tanh \beta m h}{\beta m} \right) + \frac{3}{m^2} (Sh_{\zeta\zeta\zeta} + Sa^{-2} h_\zeta + Ah^{-4} h_\zeta) [(1 - \cosh \beta m \zeta) h_\zeta \sec^2 \beta m h] \quad (50)$$

where $\beta = \sqrt{\frac{2}{K+2}}$. Substituting the above equations into Eq. (48) and rewriting the equation in terms of the original variables, the evolution equation can be expressed as

$$h_t + \left[\frac{3}{m^2} \left(h^{-2} - \frac{h^{-3} \tanh \beta m h}{\beta m} \right) (Sh^3 h_{zzz} + Sa^{-2} h^3 h_z + Ah^{-1} h_z) \right]_z = 0. \quad (51)$$

The parameters S and A in Eq.(51) can also be removed by rescaling. By redefining the following new variables

$$Z = \left(\frac{A}{S}\right)^{\frac{1}{2}} z \quad ; R = \left(\frac{A}{S}\right)^{\frac{1}{2}} a \quad ; T = \left(\frac{A^2}{S}\right)t, \quad (52)$$

the nonlinear evolution equation can be obtained as

$$h_T + \left[\frac{3}{m^2} \left(h^{-2} - \frac{h^{-3} \tanh \beta m h}{\beta m} \right) (h^3 h_{ZZZ} + R^{-2} h^3 h_z + h^{-1} h_z) \right]_Z = 0. \quad (53)$$

In the cases of $K = 0$ and $a \rightarrow \infty$, the above equation can be reduced to the same result as given in Hung and Tsai [19]. In the case of $m = 0$, the electromagnetic force can be neglected in momentum equation. The nonlinear evolution equation can be derived as

$$h_t + \frac{2}{K+2} (Sh^3 h_{ZZZ} + Sa^{-2} h^3 h_z + Ah^{-1} h_z)_z = 0 \quad (54)$$

The variation of film thickness in the base flow is found to be very small, so it is reasonable to assume that the local dimensionless film thickness is equal to one. The dimensionless film thickness as expressed in perturbed state can be presented as

$$h(Z, T) = 1 + \eta(Z, T), \quad (55)$$

where η is a perturbed quantity of stationary film thickness. By inserting the above equation into Eq. (53), the linearized equation can be obtained and

expressed as

$$\eta_T + \frac{3}{m^2} (\eta_{ZZZZ} + R^{-2} \eta_{ZZ} + \eta_{ZZ}) = 0 \quad (56)$$

The normal mode analysis is performed by assuming that

$$\eta = H_0 \exp(\omega T + iqZ) + c.c., \quad (57)$$

where ‘‘c.c.’’ is used to denote the complex conjugate, the characteristic equation can thus be obtained as

$$\omega = \frac{3}{m^2} (q^2 + R^{-2} q^2 - q^4), \quad (58)$$

where ω is the growth rate of the disturbance and q is the wave number. The wave number of the maximum growth rate of the disturbance q_m is given as $d\omega/dq = 0$ and can be expressed as

$$q_m = \left(\frac{1}{2} + \frac{1}{2R^2} \right)^{\frac{1}{2}}. \quad (59)$$

The nonlinear partial differential Eq. (53) is solved numerically by using the finite-difference method. Central differences are employed in space, and the midpoint rule is used for time. The Newton-Raphson method is introduced to solve the resulting system of difference equations. The convergent tolerance is set to be 10^{-5} . It is treated as an initial-value problem with spatially periodic boundary conditions within the fixed interval $0 \leq Z \leq 2\pi/q_m$. In this paper, we use the spatial grid points of $n=40$ and a time step

of $\Delta T = 0.001$ in each case, and the initial condition [9] is given as

$$h(Z, 0) = 1.0 + H_0 \cos(q_m Z) \quad (60)$$

$$0 \leq Z \leq 2\pi / q_m$$

where H_0 is the initial disturbance.

3. Results and discussion

A numerical example is presented here to illustrate the effectiveness of the proposed modeling approach. To deal with the nonlinear rupture problem, a thin micropolar liquid film, attached on a cylinder and subjected to a magnetic field normal to the film, is carefully studied. The effects of the Hartmann number, m , the micropolar parameter, K , the initial disturbance, H_0 , and the radius of the cylinder, a , on the rupture behavior are extensively investigated. Physical parameters that are selected for study include (i) the values of the Hartmann number for $m = 0, 0.2$, and 0.5 , (ii) the values of micropolar parameter for $K = 0, 0.3$, and 0.6 , (iii) the values of the initial disturbance for $H_0 = 0.1, 0.2$, and 0.3 , (iv) the values of the dimensionless radius of the cylinder for $a = 50, 100$, and 200 , (v) the values of parameters S and A are 0.1 and 10^{-4} [4].

Figure 2 shows the minimum film thickness h_{\min} versus time for $a = 100$,

$H_0 = 0.2$, and $K = 0.3$ with various values of m . The results indicate that the rupture time for the micropolar film flow in the absence of the magnetic field ($m = 0$) is shorter than that of the film under the influence of the magnetic field, and the rupture time is delayed as the value of m increases. This is because the large value of m can result in a large induced magnetic force and stabilize the film flow. Figure 3 shows the film profiles for $a = 100$, $H_0 = 0.2$, and $K = 0.3$ at the rupture time of $m = 0$. The results indicate that the micropolar film flow is ruptured in the absence of the magnetic field ($m = 0$) while the film are not ruptured at $m = 0.2$ and $m = 0.5$. The minimum film thickness h_{\min} versus time for $a = 100$, $H_0 = 0.2$, and $m = 0.2$ with various values of K is presented in Fig. 4. The figure indicates that when the film flow is under the influence of the magnetic field, the rupture time for the Newtonian fluid ($K = 0$) is shorter than those of the micropolar fluids ($K = 0.3$ and 0.6), and the rupture time is delayed as the value of K increases. This is because the large value of K can result in a large vortex viscosity so that the convective motion of the fluid is reduced. This causes the smaller unstable factor and makes the liquid film rupture slowly. The

film profiles for $a = 100$, $H_0 = 0.2$, and $m = 0.2$ at the rupture time of $K = 0$ is shown in Fig. 5. The results indicate that when the film flow is affected by the magnetic field, the Newtonian fluid film ($K = 0$) can be ruptured while the micropolar fluid films are still all right.

The minimum film thickness h_{\min} versus time for $m = 0.2$, $K = 0.3$, and $H_0 = 0.2$ with various values of a is shown in Fig. 6. The figure indicates that the rupture time is delayed as the radius of the cylinder increases while the micropolar film flow is under the influence of the magnetic field. It is noted that the cylinder with a smaller radius makes the flow relatively more unstable [15]. This is due to the surface tension introduced by lateral curvature. The effect of lateral capillary force will increase as the radius of cylinder decreases. Therefore, it introduces a destabilizing effect into the system. This destabilizing effect occurs since the radius of the trough of waves have a smaller value than that at the crest of the waves, and the surface tension will produce large capillary pressure at a smaller radius of curvature. This induces the capillary pressure and forces the fluid from trough to move upward to the crest. Thus, the amplitude of the wave is increased. The film profiles for $m = 0.2$, $K = 0.3$, and $H_0 = 0.2$ at the

rupture time of $a = 50$ is shown in Fig. 7. The results indicate that as the micropolar film flow is under the influence of the magnetic field, the liquid film with $a = 50$ is ruptured, however, the rupture does not occur at $a = 100$ and $a = 200$. Figure 8 shows the minimum film thickness h_{\min} versus time for $a = 100$, $m = 0.2$, and $K = 0.3$ with various values of H_0 . The results reveal that as the micropolar film flow is under the influence of the magnetic field, the rupture time is delayed as the initial disturbance, H_0 , decreases. The smaller value of H_0 makes the liquid film system more stable. Figure 9 shows the film profiles for $a = 100$, $m = 0.2$, and $K = 0.3$ at the rupture time of $H_0 = 0.3$. The results reveal that as the micropolar film flow is under the influence of the magnetic field, the liquid film with $H_0 = 0.3$ is ruptured, however, the rupture does not occur at $H_0 = 0.1$ and $H_0 = 0.2$.

4. Conclusion

The rupture theory of the thin micropolar liquid film on a cylinder under the influence of the magnetic field is studied in this paper by using the method of long wave perturbation. Eringen's micropolar fluids model is employed to

formulate the governing equations with microstructures in the fluid. The generalized nonlinear evolution equations is derived and numerically estimated to investigate the fields of film flow rupture associated with different value of the Hartmann number, m , the dimensionless micropolar parameter, K , the radius of cylinder, a , and initial disturbance, H_0 , respectively. The results show that the film rupture time will be delayed by increasing the associated parameters including the Hartmann number, m , the micropolar parameter, K , and the radius of the cylinder, a , or to decrease the initial disturbance, H_0 .

References

1. Sheludko, A. : Thin liquid film. Adv. Colloid Interface Sci. **1** (1967) , pp.391-494.
2. Dimitrov, D. S. : Dynamic interactions between approaching surfaces of biological interest. Prog. Surf. Sci. **14** (1983) , pp.295-300.
3. Ruckenstein, E., Jain, R. K. : Spontaneous rupture of thin liquid films. J. Chem. Soc. Faraday Trans. II **70** (1974) , pp.132-146.
4. Williams, M. B., Davis, S. H. : Nonlinear theory of film rupture. J. Colloid Interface Sci. **90** (1982) , pp.220-228
5. Hwang, C. C., Chang, S. H. : Rupture theory of thin power-law liquid film. J. Appl. Phys. **74** (1993) , pp.2965-2967
6. Prokopiou, T. M., Cheng, H.C. : Long waves on inclined films at high Reynolds number. J. Fluid Mech. **222** (1991) , pp.665-691.
7. Alekseenko, S. V., Narkoryakov, V. Y., and Pokusaev, B. G. : Wave formation on a vertical falling liquid film. A.I.C.H.E. J. **31** (1991) , pp.1446-1460.
8. Hwang, C. C., Chen, J. L., and Shen, L. F. : Strong nonlinear dynamic rupture theory of thin liquid films. Phys. Rev. E. **54** (1996) , pp.3013-3016.
9. Chen, J. L., Hwang, C. C. : Nonlinear rupture theory of a thin liquid film on a cylinder. J. Colloid Interface Sci. **182** (1996) , pp.564-569.
10. Eringen, A. C. : Simple microfluids. Int. J. Eng. Sci. **2** (1964) , pp.205-217
11. Eringen, A. C. : Theory of micropolar fluids. J. Math. Mech. **16** (1967) , pp.1-18.
12. Liu, C. Y. : Initiation of instability in micropolar fluids. Phys. Fluids **14** (1971) , pp.1808-1809

13. Hung, C. I., Tsai, J. S., and Chen, C. K. : Nonlinear stability of the thin micropolar liquid film flowing down on a vertical plate. *J. of Fluids Eng.* **118** (1996) , pp.498-505.
 14. Hung, C. I., Tsai, J. S. : Rupture of thin micropolar liquid film. *Acta Mech.* **122** (1997) , pp.217-223.
 15. Cheng, P. J., Chen, C. K., and Lai, H. Y. : Nonlinear stability analysis of the thin micropolar liquid film flowing down on a vertical cylinder. *ASME J.Fluids Eng.* Vol. **123** (2001) , pp.411-421.
 16. Hsieh, D. Y. : Stability of conducting fluid flowing down an inclined plane in a magnetic field . *Phys. Fluid* **8** (1965) , pp.1785-1791.
 17. Renardy, Y., Sun, S. M., : Stability of a layer of viscous magnetic fluid flow down an inclined plane. *Phy. Fluid* **6** (1994) , pp.3235-3246
 18. Tsai, J. S., Hung, C. I., and Chen, C. K. : Nonlinear hydromagnetic stability analysis of condensation film flow down a vertical plate. *Acta Mech.* **118** (1996) , pp.197-212
 19. Hung, C. I., Tsai, J. S. : Rupture of thin film under the magnetic field. *J. appl. Phys.* **80** (1996) , pp.4220-4222.
 20. Coons, J.E., Halley, P.J., McGlashan, S.a. and Tran-Cong, T. : Scaling laws for the critical rupture thickness of common thin films. *Colloids and Surfaces A* **263** (2005) , pp.258-266.
 21. Kakac, S., Shah, R. K., and Aung, W., *Handbook of single-phase heat transfer*, New York, J. Wiley, 1987.
 22. Stokes, V. K., ” *Theories of fluids with microstructures- An introduction*, Berlin : Springer-Verlag, 1984.
 23. Kolpashchikov, V. L., Migun, N. P., and Prokhorenko, P. P. : Experimental determination of material micropolar fluid constants. *Int. J. of Eng. Sci.*, **21** (1983) , pp.405-411.
 24. Eringen, A. C. : Theory of anisotropic micropolar fluids. *Int. J. of Eng. Sci.*, **18** (1980) , pp.5-17.
 25. Edwards, D. A., Brenner, H., and Wasan, D.T. : *Interfacial transport processes and rheology*. Boston : Butterworth-Heinemann, 1991.
 26. Datta, A. B., Sastry, V. U. K. : Thermal instability of a horizontal layer of micropolar fluid heated from below. *Int. J. of Eng. Sci.*, **14** (1976) , pp.631-637.
 27. Ahmadi, G. : Stability of a micropolar fluid layer heated from below. *Int. J. Eng. Sci.*, **14** (1976) , pp.81-89
-

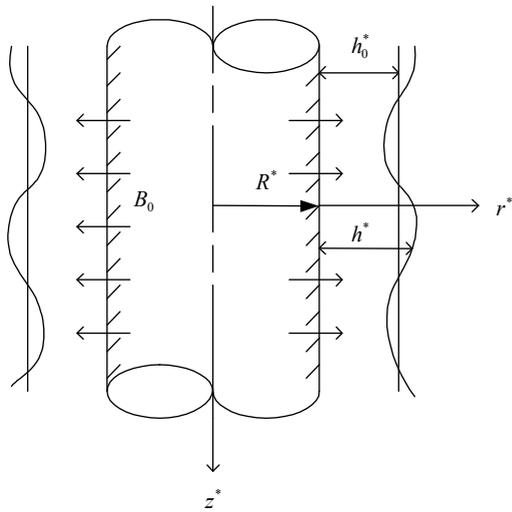


Fig. 1 The physical model of the system

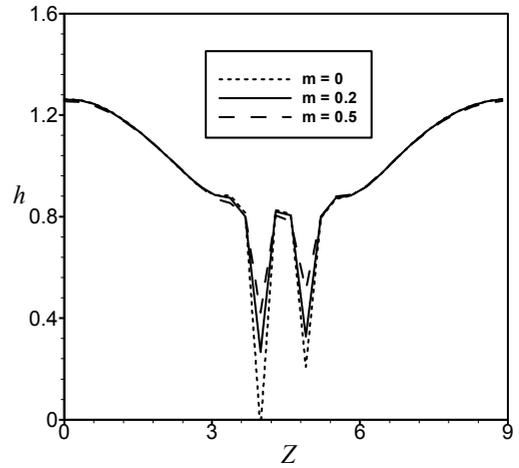


Fig. 3 Film profiles for $a = 100$, $K = 0.3$, and $H_0 = 0.2$ at the rupture time of $m = 0$.

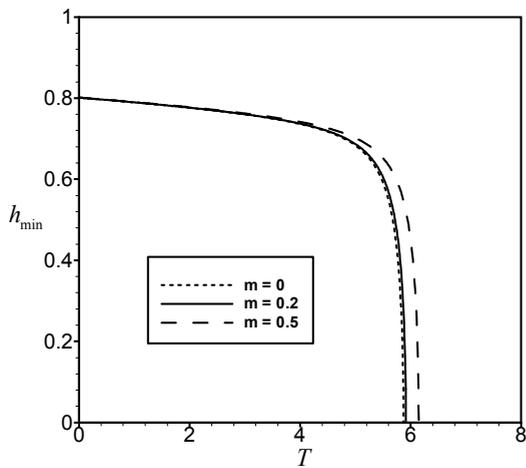


Fig. 2 The minimum film thickness h_{\min} versus time for $a = 100$, $K = 0.3$, and $H_0 = 0.2$ with various values of m .

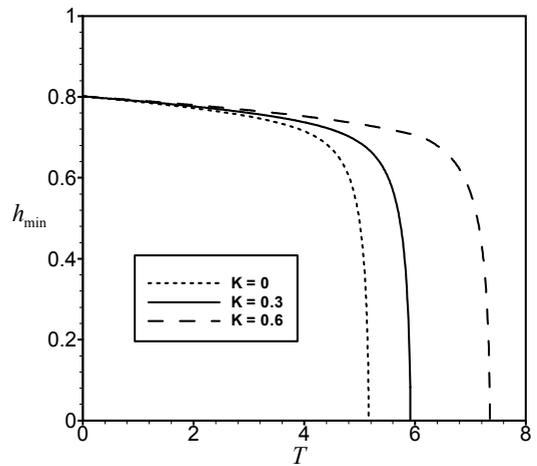


Fig. 4 The minimum film thickness h_{\min} versus time for $a = 100$, $m = 0.2$, and $H_0 = 0.2$ with various values of K .

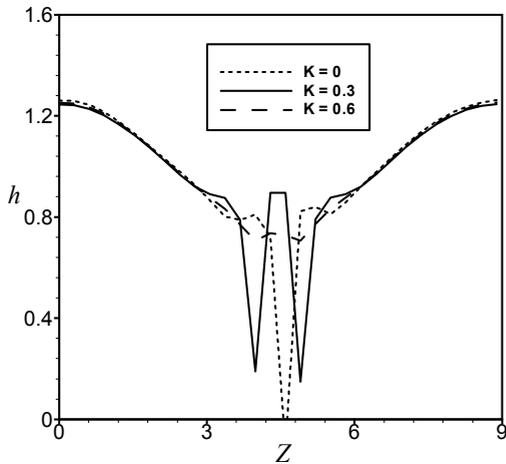


Fig. 5 Film profiles for $a = 100$, $m = 0.2$, and $H_0 = 0.2$ at the rupture time of $K = 0$.

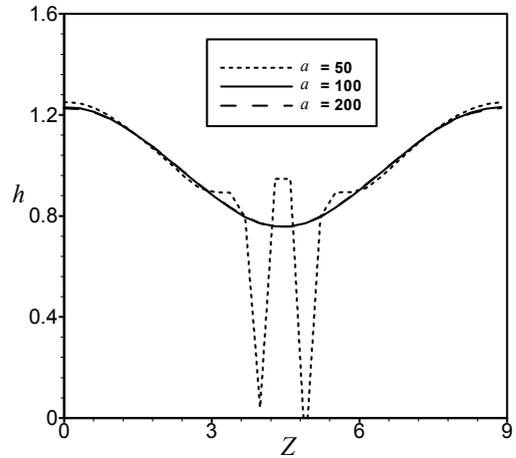


Fig. 7 Film profiles for $K = 0.3$, $m = 0.2$, and $H_0 = 0.2$ at the rupture time of $a = 50$.

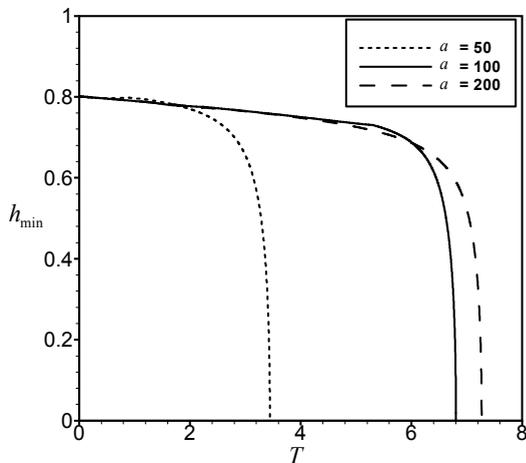


Fig. 6 The minimum film thickness h_{\min} versus time for $K = 0.3$, $m = 0.2$, and $H_0 = 0.2$ with various values of a .

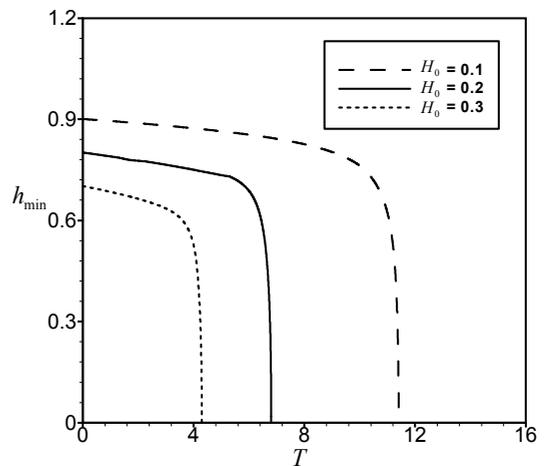


Fig. 8 The minimum film thickness h_{\min} versus time for $K = 0.3$, $m = 0.2$, and $a = 100$ with various values of H_0 .

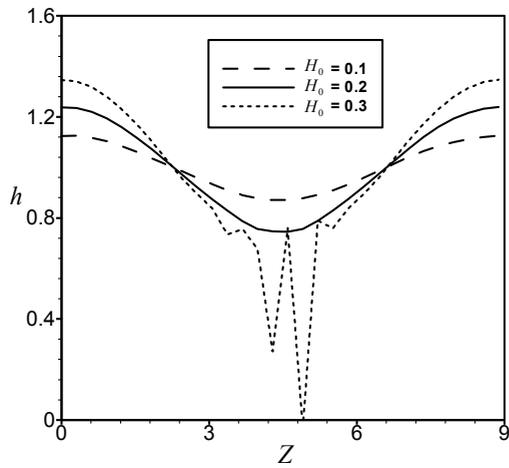


Fig. 9 Film profiles for $K = 0.3$, $m = 0.2$, and $a = 100$ at the rupture time of $H_0 = 0.3$.

女性消費者對銀行授信業務 刻板印象之研究：以苗栗縣為例

陳泓予

摘要

本文以女性消費者為研究對象，藉由「次級資料分析法」和「深度訪談法」來進行女性消費者對銀行授信業務刻板印象之研究，研究目的如下：

- 一、探知女性消費者對於銀行授信業務的刻板印象。
- 二、了解女性消費者對於銀行授信業務刻板印象的原因為何？
- 三、歸納上述兩點，提供銀行業者在經營與行銷領域上的建議。

經研究後發現：女性消費者對銀行的聯想是「錢」，並無其他刻板印象；對銀行服務的第一印象是：公營銀行態度比較不好，感覺很官僚；民營銀行比較有禮貌，但太做作，很假；對銀行授信業務的印象仍停留在「利用關係」、「人情」、「特權」的思考模式，對銀行授信業務刻板印象形成的原因為：受到「他人口碑影響」、「報章雜誌」、「主觀意識」等三個方向的資訊來源影響。

關鍵詞：女性消費者、刻板印象、授信業務。

The Study on Female Consumer Stereotypes towards Bank's Loan Business : Take Miaoli County as Research Sample

Hung-Yu Chen

Abstract

In our daily life, there are all kinds of stereotypes. In regard to all walks of life, occupations, Stereotypes may imply a preconceived impression.

Similarly, we may also have stereotyped ideas about the banking industry, system, or the attitude of the staff in banks because of universal cognition, previous impressions, experiences from others, etc.

This study aims to explore whether female consumers have stereotyped ideas about marketing staff in bank's loan business and what lead them to do this?

Keywords : Female Consumer, Stereotypes, Bank's Loan Business.

壹、前言

一、研究背景

揭開臺灣銀行產業發展史，要從馬關條約簽訂後開始談起，1895年9月，日本大阪中立銀行首先在基隆設立「出張所」，代辦國庫業務並經營商業銀行存、放款業務，自此臺灣開始有了銀行。歷經100年的演進，在1991年為因應自由市場競爭，開放新銀行設立後，臺灣銀行業正式邁入競爭的新紀元。在這段時間裡，過去基於特殊背景需要而成立的民營金融機構，紛紛改制為商業銀行；屬於地方金融機構的信用合作社也在此時加入轉型的行列，以致於臺灣銀行業的競爭在2000年達到白熱化。再則，為維護金融秩序，促使金融機構達到合理的經濟規模，發揮綜合經營效益與提升國際競爭力，在2000年通過「金融機構合併法」，隔年又通過「金融控股公司法」。在這些法源架構下，臺灣銀行產業自此跨入整合的新世紀。

日常生活之中，消費者對於銀行的意象可能是和「錢」有關的地方、行業，甚至認為銀行是有錢人和有權人，從事權、益交易的地方，或是銀行服務人員態度不

佳、對尋常小老百姓十分刻薄等等，造成消費者認為銀行的負面刻板印象。不諱言的，在產業開放經營之後，國內銀行的發展與從業人員素質參差不齊，雖然大多數銀行均面臨經營困境，然而其中亦不乏有消費者始終對某些銀行有極具正面的評價，忠誠度不減。消費者對於銀行的負面刻板印象，可能來自於服務人員提供服務時的態度不佳，或是傳媒報導利益輸送的不法案件，或是資訊不對稱所引起的誤解，但都不足以代表全體銀行的共同特性。

二、研究動機與目的

在我們的日常生活之中，充斥著各式各樣的刻板印象，無論是對於人、事、物或是地點、行業、職業等等，都可能隱含了先入為主的主觀刻板印象。同樣地，我們對於銀行業的經營、制度或從業人員的服務態度或是銀行的業務也可能因為既有的認知、過去片段印象、他人口碑的傳達或是傳播媒體報導的關係，形成了先入為主主觀的刻板印象。

柯淑敏（2001）提出：性別刻板印象讓人刻板的認為女性會傾向選擇陰柔的工作，在就學科系的選擇上，女性會選擇文科或商科。而在傳統的觀念中，從事會

計工作以及與錢財有關的工作，大都數是女性的人數比率較多，一般家庭中掌管錢財和經濟問題者也多以女性為主。另，沈美杏（2007）提出：兩性社會行為上的差異受限於社會角色理論之影響，因社會角色的期待，造就了可被觀察的刻板印象。根據研究者觀察每天前往銀行辦事的消費者，亦以女性人數所佔比率較高（會計小姐、老闆娘）。這個現象不禁讓研究者好奇，不知女性消費者對銀行授信業務是否有著既定的刻板印象？而女性消費者為何會有這樣的刻板印象？以及這些刻板印象是否會因為不同的銀行而有所差異？因此以女性消費者為研究對象，寄望能夠由本研究提出一些實務上與學術上的意見，即為本研究之動機。

由於研究者係銀行從業人員，深知在銀行所經營的業務中，「授信業務」所創造的利潤是銀行收入的主要來源之一，女性既是和銀行接觸頻率較高的消費族群，則女性消費者對銀行的聯想，除了錢之外，不知是否仍有其他刻板印象？再者，對銀行授信業務的認知，是否仍然停留在「利用關係」、或「特權」的思考模式，而對於這些刻板印象形成的原因，是否亦受「口碑」的影響？觀諸國內有關於銀行產業的研究，大多是聚焦在政策面、

滿意度、忠誠度和消費者行為等面向，針對銀行授信業務的刻板印象進行研究者付之闕如，因此，本研究嘗試進行探索性研究。而本研究欲探討的目的有三：

- （一）探知女性消費者對於銀行授信業務的刻板印象。
- （二）了解女性消費者對於銀行授信業務刻板印象的原因為何？
- （三）歸納上述兩點，提供銀行業者在經營與行銷領域上的建議。

三、研究範圍與限制

在有限的時間、人力、成本以及研究者本身學術背景的影響下，有其適用範圍與限制存在。因本研究的研究地區與樣本有限，無法作全面性推論，且囿於質性研究的特性，有部份結果受研究者主觀判斷影響，再加上本研究為探索性的研究，對於訪談結果的驗證，尚未加以反覆淬鍊，較難更具體與完整的呈現研究的結果，本研究僅以一次的訪談結果即寫為研究結論，此亦為本研究限制之一。

貳、文獻探討

一、銀行主要業務概述

依據銀行法第3條的規定，銀行業能夠經營的業務共計有22項，但歸納而言，

銀行業為從事收受存款、辦理放款、外匯等業務的行業；其中存款業務又稱為受信業務，而放款業務又名授信業務。

二、授信之定義與原則

依據銀行法第 5-2 條：本法稱授信，謂銀行辦理放款、透支、貼現、保證、承兌及其他經中央主管機關指定之業務項目。邱潤容（1998）提出：所謂「授信」，係指銀行對於客戶授與信用，並擔負風險，以賺取利息的業務。銀行授予信用之型態主要有二種，其一為銀行以其自有資金放貸給資金需求者，以賺取利息（即為直接融資）；其二為應社會經濟發展及交易型態改變的需要，除了直接授信之外，另有如票據承兌、保證、簽發信用狀等非直接資金授予的間接授信業務。這二種型態在形式上雖然不同，但承擔風險的本質並無差異。

銀行從事資金之貸予或信用之提供，風險之發生均在所難免。資金貸予的風險為資金之損失，而信用提供之風險則為代履行其債務，其結果亦為資金之損失，故辦理直接授信與間接授信，均須審慎處理，以期債權之安全。銀行授信業務為銀行營利的主要來源，授信業務之主要功能為一方面配合國內經濟發展的要

求，滿足工商企業活動及社會大眾的需要；另一方面則在維持銀行的正當收益及優質的放款資產，確保存戶權益。依據邱潤容提出銀行在受理承做授信業務時，應遵循下列原則：

（一）安全性：

銀行辦理授信，其主要之資金來源及信用基礎，乃是建立在存款戶之存款及股東之投資上。所以為了確保社會廣大存款戶之存款資金的安全，以及銀行股東之權益，授信業務安全性，也就是授信之品質乃成為授信的基本原則之一。

（二）流動性：

由於銀行授信的資金多數來自於短期負債，所以辦理授信業務應盡量避免資金之固定呆滯，因此維持適度之流動性以配合營運需要，亦為授信的基本原則之一。

（三）公益性：

銀行放款為將社會儲蓄導向投資之主要管道，為工商企業及個人資金融通之主要來源，所以經濟發展有賴銀行之協助，辦理授信業務是否符合公益性原則，遂成為從事授信業務所應把握的原則。

（四）收益性：

銀行的經營，雖然負有調節社會金融，穩定國民經濟之使命，但銀行也是營利單位，也必須負擔行員薪金、利息支出、稅捐等龐大開支，所以必須有收益方足以維持。

（五）成長性：

銀行為追求永續的經營與發展，各項業務的成長亦相當重要，有交易對象企業的成長才有銀行的成長，因此，要求助於交易對象企業成長的融資，就是「成長性原則」。

三、授信業務的原則

對於借戶信用評估不當，不但影響銀行債權之確保，亦將損及銀行之收益性與流動性。邱潤容（1998）：為了確保銀行債權以及滿足借戶之資金需求的雙重任務，處理授信業務時應把握授信之五項原則（即通稱之授信5P原則）：

（一）借戶（People）：

對借戶之評估，可由其責任感、企業經營成效、企業經營者及其銀行往來情形加以了解分析。

（二）資金用途（Purpose）：

一個健全的授信業務，非僅重視借

戶提供的擔保品，更重要的是要看借款資金的運用計劃是否合情、合理與合法，缺乏具體與積極資金用途之放款，先天上就已註定可能會形成收回困難的結果。此外，貸放之後，更應追查資金動向是否依照原核定的計劃運用，以免資金流用。

（三）還款來源（Payment）：

還款來源是確保授信債權本利回收的前提要件。因此，分析借戶償還授信的資金來源，毋寧是銀行評估信用的核心，也是考核授信主管人員能力強弱的指標。

（四）債權保障（Protection）：

就債權保障而言，可分為內部保障與外部保障兩方面。內部保障，專指銀行與借戶之間的直接關係；而外部保障，係指由第三者對銀行承擔借戶的信用責任而言，通常銀行債權的外部保障，有保證、背書等方式。外部保障的關鍵乃在於保證人、背書人信用、資力等條件。惟銀行債權的確保應以內部保障列為優先考慮的主要因素。

（五）授信展望（Perspective）：

銀行對於授信案件的判斷，除應照上述四個原則加以評估之外，尚應注意借款

戶事業的展望，包括借款戶行業別的前途及借款戶本身將來的發展性，再據以在風險與利益的衡量下作成貸款與否的決定。

四、刻板印象的定義

陸谷孫（1992）指出：「刻板印象」的英文為 *stereotypes*，其源自於早期的印刷排版工人是按照字母的順序查找字模來排版的，在這樣的過程中，為了方便起見，他們把經常聯合使用字詞的字模網綁起來，每次遇到連用就直接使用網綁的字模而不必分別查找幾次，以此加快排版速度。這些被捆綁起來的字模就叫做 *stereotypes*。但時至今日，關於「刻板印象」的詞意已經延伸出許多不同的解釋。

學者們對於「刻板印象」各自以不同的觀點加以詮釋，李美枝（1979）認為：刻板印象指一種對人或事抱持著固定不變的看法。也就是說，人們常須以部分的資訊類推到整體，用來類化或理解他們所接觸的事物。Devine（1989）提出：刻板印象是自動的。（一）刻板印象的激發可以被暗中地、自動地激發，影響後續的想法、感覺與行為，即使接受者的偏見很低。（二）動機性因素會引起刻板印象的認知處理過程，亦是指當接受者在特定情境中有特定目標時。李美枝（1990）認為

刻板印象是指社會對某些特定族群的人，有一種簡化、僵化及過度類化的看法，是一種先入為主的判斷。Basow（1992）提出：刻板印象是指對社會某一特定群體中的人強烈地過度類化，它是一種固定、特定且不易改變的形式，常常忽略個別差異、過度簡化，具有相當一致性的偏誤態度或觀念，且普遍存在於社會中，強烈的支持了個人的觀念，其特點為（一）刻板印象是粗淺、偏差的，也可能是真實的，不一定有經驗或事實根據，當然也不一定是壞的；（二）刻板印象是個人對社會團體或人群做歸因。溫世頌（1992）將刻板印象定義為：（一）共同的觀念。（二）以團體屬性描述其該團體成員特質。（三）個人對他人持有不正確的固板想法。張春興（1996）認為刻板印象是指不以直接經驗為根據，又不以事實資料為基礎，單憑一己之見或道聽塗說，以粗淺、過度概括化的方式，形成對某人、某事或某團體武斷的評定。劉淑雯（1996）認為刻板印象是一種過度簡化、忽略個別差異、具有相當一致性的偏誤態度或觀念，且刻板印象持久、僵固、不易改變並強烈支持個人的觀念。Silverblatt & Eliceri（1997）認為：刻板印象為對於個人、團體以及事物持有過於簡化的概念。李卓夫（2000）亦認為

刻板印象代表了我們是如何看待自己與他人所處之團體，影響我們與他人互動的方式。黃德祥（2000）提出：若是刻板印象用在對人對事的判斷上，卻也常形成僵硬、主觀或武斷的看法。林麗珊（2002）認為個體在成長的過程中，經歷他人的期許、社會的鼓勵，造成我們表現出符合期待的男女特質，同時排斥不符合期待的行為，形成性別刻板化的結果。林文瑛（2002）認為大眾眼中因為性別不同而產生對男、女性的不同印象與假設，就像是男、女性有著不同的特質，適合從事不同的工作。黃葳威（2002）將刻板印象視為一種過度簡化、負面、片面、不完全、甚至誇大的印象。其認為電視媒體由於傳遞人事物相關議題的偏頗，一般大眾在其長期的薰陶下，對特定人事物形成的刻板印象，可能在不自覺中影響我們判斷思維。存有刻板印象人通常不認為刻板印象只是抽象概念或認知的歸類，而以為刻板印象便是真實的。羅燦慶（2003）則認為：刻板印象是無所不在，如：國家及種族間，各種職業、地域、意識型態、性別、及經濟領域（對某特定企業或產品的印象）等，而刻板印象在跨文化溝通中，具有導引的功能。因此，人們可以在這複雜的世界中化簡為繁，對外面的人事地物有

初步的認識。晏涵文（2004）認為一般人在社會化的過程中，如果學習到社會文化所賦予性別的規範，就會對性別角色及其行為的信念與態度形成一種固定、刻板和概念化的標記，且產生相對應的行為傾向。

綜合上列學者的見解之後，不難窺見刻板印象是指我們對某樣人、事、物，乃至於某一個行業、事件、職務、制度等所形成的一種先入為主的觀念，是一種過度僵化、不易改變且根深蒂固的觀念，面對所想認知的人、事、物尚不了解時，刻板印象提供了最簡單的認知與分類的方法，但它卻未必都是正確的看法，因而刻板印象有時雖然有助於我們快速將共同特徵的東西加以分類或了解，但卻也容易造成僵化的認知，而忽略個別差異。

雖然在日常生活當中，「刻板印象」一詞通常給人負面的感覺，但是刻板印象對於任何事物來說卻不見得是件壞事，以本研究而言，銀行業者即可藉由瞭解消費者對於銀行授信業務的刻板印象，進而調整授信人員接待顧客的態度、說明的口氣和用語，或是規劃一套更適合的流程來滿足來到銀行借款消費者的期望。因此，本研究較認同中立的定義，認為刻板印象的形成是受到消費者本身對於人事物既有

的認知、過去片段的印象、他人口碑的傳達，或是傳播媒體報導的影響，是種籠統而概括的印象或觀念。

叁、研究方法與設計

在企業管理實務上的研究，可以將研究設計的類型分為探索性研究（research design）與結論性研究（exploratory research）二大類。Zikmund（1994）提出：探索性研究是指對相關情境所知甚少，沒有足夠相關資訊可供參考時，為了釐清與定義問題本質所做的初始化研究。當研究者對於所欲研究的特定主題或現象沒有足夠的認識時，必須採取探索性研究以獲得未來研究所需的資訊。

一、方法選取與資料蒐集

古永嘉（2003）指出：探索性研究可以使用定性方法，或定量方法，但通常較著重使用定性方法。黃俊英（2006）提出：探索性研究的主要目的在發掘初步的見解，並提供進一步研究的空間，並不力求證明或推翻其他相關理論，通常在探索性研究方面，會配合進行資料整理分析、深度的訪談、參與觀察等。由於國內針對銀行授信業務的刻板印象進行研究者付之闕如，因此，本研究嘗試以探索性研究中

的「次級資料分析法」和「深度訪談法」來進行研究。

Earl Babbie（1998，李美華等譯）指出：次級資料分析法（Secondary Data Analysis）可以適用於任何形式的傳播媒介，這些可以用來學習的人為事實可以是書本、雜誌、詩、報紙、歌曲、繪畫、演講、信件、法律以及制度，和有關於任何成份或集合。次級資料分析是一種最基礎而實用的探索性研究方法，依據楊國樞（1993）的說法，即蒐集有關前人所做研究，分析其研究結果與建議，指出需要驗證的假說，並說明這些建議的假設是否具備實用價值，而作為自身研究的基礎。本研究初期先以「次級資料分析法」進行蒐集與整理相關資料，研究者藉由了解銀行業的主要業務、授信的定義，再擴展到理解授信業務的流程及原則，並深入了解學者對於刻板印象的見解及定義，蒐集的資料包括期刊、論文、專書、銀行業業務作業手冊、網頁等資料，蒐集後進行分析、整理與綜合，以期對本研究進行初步瞭解。

由於本研究主題中的「刻板印象」問題是極為主觀的意象，無法透過直接觀察而得到結果，因此本研究另外採用質化研究的「深度訪談法」（in-depth interview）

進行資料蒐集。Earl Babbie指出：深度訪談又稱為「質化訪談」，採用較少但具代表性的樣本，在訪問者和受訪者之間針對研究主題進行互動，而不是一組特定的問題、必須使用一定的字眼和順序來詢問的方式。質化訪談在本質上由訪問者建立對話的方向，再針對由受訪者所提出的若干特殊主題加以追問。深度訪談法可以讓研究者以「中性化的探索技巧」執行非操作化的訪談，透過蒐集、組織及分析資料，提出綜合性及系統性的有價值及較深入的問題，針對複雜性的研究對象深入了解。

Miller與Crabtree（1992）將訪談法分為「結構式」、「非結構式」及「半結構式」訪談，半結構式訪談乃是研究者預先擬定訪談大綱，對於訪談結構有初步的規劃與控制，接著再依照與受訪者的互動中，調整提問的順序。Wimmer & Dominick（1995，李天任等譯）指出：深度訪談法最大的優點是受訪者可以提供豐富詳盡的資料，並回答敏感性的問題，而訪問者與受訪者之間的關係，使訪問者易於接近其它研究方法中受到限制的話題。本研究採用深入訪談的原因除了希望可以有更多的發現與獲得豐富的資料外，另一方面由於研究者本身的職業背景即為銀行從

業人員，職場生涯中深入參與銀行徵信、授信與催收訴訟等面向的工作，因此長期吸收相關產業訊息，由研究者關注銀行業的現狀與未來發展趨勢，所建立的背景知識及專業領域的熟悉，以十年以上工作時間的歷練，透過學術論述的思維進行訪問，期能以實務界的角度切入，分析女性消費者對於銀行授信業務刻板印象的關鍵因素。

二、訪談流程與研究對象

本研究的目的，是為了瞭解女性消費者對於銀行授信業務的刻板印象，並以苗栗縣地區為研究的區域，在訪談的流程部份略述如下：

（一）準備工作：

根據書籍、期刊資料、相關研究結果等文獻整理後，擬定訪談大綱，並於訪談前預先提供給受訪者。

（二）進行預訪：

雖然事先設計訪談問題，但不一定按照問題順序來訪問，保留受訪者有餘地決定如何作答，以此法導引出更多細節，達到研究的深度與廣度、確保訪談的品質，以得到深入詳盡的資料。

（三）進行訪談：

每位受訪者的訪談時間均為15到20分鐘，以面對面訪談方式進行，並隨時觀察受訪者的反應，掌控現場氣氛、並調整訪問方式，以獲得研究所需要的資料。除了針對預設的問題大綱進行深度訪談，訪問過程中之後將把受訪者當場回答的內容詮釋之後，再提出進一步的訪問。

(四) 整理資料：

在訪談結束之後，撰寫訪問稿，並就訪談對象的訪問稿，進行初步歸納，再與次級資料分析法所整理的資料交叉閱

讀，整合後撰寫報告。

高敬文(1988)指出：在質化研究中，對於研究樣本的選擇主要是根據研究問題的特性、理論的需要及研究資料的豐富性來選擇受訪者。由於本研究旨在探討女性消費者對於銀行授信業務的刻板印象，因此在研究對象的選擇方面，係以苗栗縣境內銀行的顧客為主，並以訪問當日銀行結束營業前到行的女性顧客為選取對象，訪談對象詳見表1：

表1 訪談對象

編號	受訪日期	受訪地點	稱謂	年齡	職業	有無向銀行借款經驗
1	2010/03/04	台銀苗栗分行	李小姐	25	會計	無
2	2010/03/12	一銀苗栗分行	王小姐	32	工程師	有
3	2010/03/18	合庫苗栗分行	陳小姐	29	公務員	無
4	2010/03/26	中國信託頭份	AMY	22	大學生	無
5	2010/03/30	兆豐頭份分行	李小姐	35	廠務助理	有
6	2010/04/01	竹南信社頭份	洪小姐	38	家管	有
7	2010/04/07	台灣企銀竹南	梅子	23	待業中	無
8	2010/04/16	玉山竹南分行	殷小姐	40	作業員	有
9	2010/04/20	華銀竹南分行	張小姐	36	教師	有
10	2010/04/28	彰銀苑裡分行	李小姐	26	技術員	無
11	2010/05/03	台中商銀苑裡	劉太太	52	企業主	有
12	2010/05/11	土銀通霄分行	邱小姐	30	服務業	無
13	2010/05/19	渣打後龍分行	許媽媽	56	家管	無
14	2010/05/27	渣打三義分行	JUDY	24	服飾業	無
15	2010/06/02	渣打公館分行	小可	22	店員	無

三、訪談大綱

本研究訪談結構採用「半結構性訪談」，根據預先擬定好的訪談大綱，對於受訪者進行訪談，配合受訪者的回答調整提問順序，再加以追問，以達到資料蒐集的目的。訪談大綱的設計如下：

(一) 基本資料：包含如何稱呼、年齡、職業、是否有向銀行借錢的經驗等。

(二) 對銀行第一印象發想及形成原因：

1. 當提到銀行您會想到什麼？為何會想到這些人、事、物？
2. 您認為銀行業者或服務人員的服務態度如何？為何會有這樣的感受？

(三) 銀行業授信業務刻板印象發想及形成原因：

1. 您覺得銀行辦理放款業務的訊息蒐集容易或困難？為何會有如此的想法？
2. 您認為政府、銀行業者、傳播媒體對消費者提供銀行借款的訊息是否充分，您有何看法？
3. 您認為由週遭朋友所提供銀行借款的訊息是否充分，您有何看法？
4. 您認為向銀行申辦借款的利息和收費價格如何？為何會有這樣的感受？
5. 您認為銀行辦理放款業務的准貸或不准貸是如何決定的，您有何看法？

6. 您認為向銀行申請辦理放款是不是一定要有抵押物，您有何看法？

四、信度與效度

王文科(2000)提出質性研究是歸納邏輯，藉著分析以及提出漸呈抽象和綜合形式的資料中，「建立」起得自資料的發現，藉著歸納推理，從分析觀察結果，而建立敘述性描述(narrative descriptions)，並歸納邏輯從敘述性描述中，產生概念，因此，質性研究比較偏重效度、強調主觀性，並注重全貌分析。探討質性研究的信度與效度有其困難之處，尤其對於效度的探討，雖有部分質性研究的學者認為質性研究與量化研究的認識論、本體論不同，不必探討信度與效度問題，但多數的研究者則認為質性研究優於其他方法，仍必須針對此問題加以探討，以顯示其重要性。胡幼慧、姚美華(1996)引述 Lincoln 與 Guba 的見解，在質性研究的信度與效度上，提出確實性(credibility)、可轉換性(transferability)、可靠性(dependability)等質性研究信度、效度的控制方法。而為提昇本研究的品質，本研究即採用其所定之方法，以提高研究的信度與效度，詳如表 2：

表 2 提高研究的信度與效度的方法

方法	表彰	意涵	本研究具體作法
確實性	內在效度	質性研究資料真實的程度。	使用多種證據來源：如書籍、期刊資料、相關研究結果等文獻。
可轉換性	外在效度	研究者能詳實的記錄現象的脈絡、情境、對話。	針對所蒐集相關資料進行重點分類系統化管理，以便於研究分析。再請受訪者協助檢視資料整理初稿，並尊重其看法對有瑕疵處進行修正。
可靠性	內在信度	資料來源的可靠性，被研究者個人經驗的重要性與唯一性。	受訪者均非銀行從業人員，並涵括多個年齡層，由受訪者主觀的刻板印象，進而探究其造成產生刻板印象的原因。

五、資料分析的技術

由於在本研究資料的蒐集上，蒐集了大量的語言資料，故在資料的分析技術上，本研究採用了當前質性研究分析方法上快速發展的方法－「敘述分析法」(narrative analysis)。Arsenault & Anderson (1998) 認為，分析質的資料是一系統化的過程，將資料組織成可以管理的單位，結合並綜合觀念、發展概念、主題或理論，並闡明研究的重要發現。本研究首先將訪問內容整理成文字稿，根據文字的內容及訪談時的筆記，進行資料的歸納、比較和概念化。然後在研究發現的主題呈現上，採用林明地 (2000) 建議的原則：(一) 出現次數比較多的主題內容。

(二) 與文獻發現異同之意見。(三) 指出事件發生的原因者。(四) 內容值得討論者。(五) 意見較獨特者。將受訪者的陳述直接引用於文中，以清楚表達研究者的分析概念，最後加以意涵上的詮釋，呈現出最終結果。

肆、資料整理與分析

本研究試圖探索出女性消費者對於銀行授信業務具有哪些刻板印象？以及瞭解造成女性消費者對於銀行授信業務產生刻板印象的原因為何？期望能夠協助銀行業改善或保持授信業務的經營。因此，在整理訪談結果之後，將女性消費者對於銀行具有的第一印象、銀行授信業務具有的刻板印象以及造成女性消費者對

於銀行授信業務產生刻板印象的原因加以分析，共分成二部份。第一部分說明女性消費者對於銀行的第一印象有哪些？第二部分說明女性消費者對於銀行授信業務具有哪些刻板印象？並且說明造成女性消費者對於銀行授信業務產生刻板印象的形成原因為何？

一、對銀行的第一印象

女性消費者對於銀行的第一印象多為有很多錢，是可以存錢、領錢、匯款的地方，也是可以用新台幣換外國貨幣的地方，缺錢的時候可以到銀行借錢，在銀行可以申請信用卡，也可以申請支票。

- 1：銀行就是存錢、領錢、匯款的地方阿，我的工作每天都要跑銀行，銀行就是和錢有關的地方。
- 2：創造信用的機構吧。
- 3：我想到的是可以換外匯的地方耶，因為我最近很注意外幣匯率的變動。
- 4：銀行裡錢很多，就是可以讓我們從事財務操作的地方阿，例如可以在銀行買基金。
- 5：提款機。
- 6：就是存錢、領錢的地方阿。
- 7：搶匪的最愛（笑）。
- 8：每個月都要固定繳房貸的地方，想存錢都存不到。
- 9：銀行是創造信用的地方，像我有很多張信用卡，買東西可以延遲付款，真好。
- 10：有錢的地方阿；冷氣很冷。
- 11：我所想到的就是可以存錢、借錢，還有跑三點半。
- 12：都是和錢有關的事吧！就是去存錢啦、借錢啦、匯款啦，換美金啦。
- 13：銀行是我藏私房錢最好的地方。
- 14：我想到的是刷卡耶。我的客戶大部分都刷卡，有的人信用卡很多張呢。
- 15：我換零錢的地方。

女性消費者，對銀行的服務態度第一印象是：公營銀行職員態度比較不好，感覺很官僚，職員大多有晚娘面孔，但是也

有少部份例外。民營銀行的行員感覺上比較有禮貌，但是也時候感覺太做作，很假；但是對熟識和不熟識的人，在態度上

有差別。銀行的職員都是白領階級，待遇對銀行業者及服務人員態度問題，仍屬正面。不錯，要考進銀行很不容易，是令人稱羨面。的工作。歸納而言，大多數的女性消費者

- 1：可能因為臺銀是國家的，櫃員的臉都很臭，要不是老闆在台銀往來，我才不要來呢。
- 2：態度還可以啦，我其實也不常來。
- 3：服務態度還可以啦，現在銀行是服務業，大致上來說都還不錯。不過也時候到不熟的銀行，就不一定了。
- 4：中國信託的制服很漂亮，感覺服務態度不賴，也很有禮貌，但是有時候假了一點，講了一大堆廢話還聽不到重點。
- 5：對認識的老顧客態度都不錯丫，但是對不認識的人就不一定了。很多人都說民營銀行的服務態度比較好，尤其是玉山銀行。
- 6：我是常客，和行員都很熟，服務態度很好阿。
- 7：有人說公營行庫的老行員態度比較差，但我覺得應該不會，服務態度應該是因人而異，和那家銀行應該比較沒有關係，不過我對台銀的印象很不好，之前去台銀，他們的櫃員做的很慢，很官僚。
- 8：銀行員是金飯碗耶，我也很想考銀行，但是不好考。不過有人說現在銀行員不好做，也勸我別考。
- 9：態度比以前好了，不知道是不是因為我是老師的關係？
- 10：彰銀的服務態度不好，我是從彰銀移到這裡的。老銀行的員工都以為他們還是金飯碗摔不壞的。
- 11：服務很親切，行員和我感情不錯，像我的子女一樣。
- 12：等待時間太久了，可能我是陌生人，行員的態度就很一般阿。
- 13：服務還不錯啦，一進門就奉茶了。

- 14：金飯碗，誰叫我以前不讀書，不然在銀行工作肯定不錯。聽說公家銀行的人都很冷，不過應該也會有例外的人吧。
- 15：夏天銀行裡的冷氣很冷，但是人也很冷，可能是銀行有要求吧，很多銀行的員工感覺很有禮貌，但都太做作了，很制式化，我覺得不是很舒服。但是在服務態度上還不錯啦。

二、對銀行授信業務的刻板印象與原因

說明女性消費者對於銀行授信業務具有哪些刻板印象？並且說明造成女性消費者對於銀行授信業務產生刻板印象的形成原因為何？

(一) 銀行授信業務資訊蒐集的刻板印象

女性消費者普遍認為銀行授信業務資訊的蒐集並不容易的，銀行業者對於授信業務資訊的公開管道並不多，甚至大多數的人都說政府沒做過這些事。另外，有部分女性消費者則都是聽取來自親朋好友的親身經歷，她們認為聽別人的口碑比較實際，但依台灣人好面子的特性，通常不太喜歡告訴別人自己有向銀行借錢，除非是自己的親人或特別好的朋友。

- 1：我覺得很困難，因為我沒有看過政府的宣傳耶，廣告也沒有阿，到銀行去的時後，他們會大約講一下啦。
- 2：應該是很困難。政府宣傳，有嗎？我沒印象。廣告，好像只記得以前萬泰銀行的現金卡，但現在應該沒了吧。別人有借錢應該不會告訴我吧。
- 3：銀行會有DM，網站資訊也有阿。
- 4：有阿，中國信託的廣告不是挺多的。
- 5：不是很容易，我有聽我同學說過他辦貸款的那家銀行速度很快，廣告就好像沒看過耶。政府對銀行授信業務的宣傳，我沒印象有看過或聽過耶。
- 6：太困難了。銀行好像都只有廣告理財業務吧，授信業務我沒有印象。政府部門有做過這些事嗎？
- 7：我認為不簡單，因為朋友在銀行借錢應該不會告訴我吧。而且如果我要借錢也不會去問朋友。

- 8：有阿，玉山的形象廣告不錯阿，不過好像只是形象廣告，沒有提到銀行業務。
- 9：並不是太容易。我比較會相信有辦過貸款的人說的耶，廣告應該都不準吧，政府部門，我沒印象我們的政府有做過這些事，然後媒體都是選舉的消息，所以我也沒印象，倒是常常聽到銀行員叫我借錢，不過細節就沒講了。
- 10：很多民營銀行的行員會傳真DM到公司啦。其他的好像就沒印象了。
- 11：行員會主動推薦，廣告就沒甚麼印象。
- 12：算是可以啦，如果有心要蒐集的話應該不難，但是政府部門主動的告訴消費者，好像沒有耶。銀行借錢的業務可以廣告嗎？我不知道耶。我是有聽過幾個同事去銀行借錢的經驗，聽他們說應該比較準吧。
- 13：沒有印象。
- 14：我覺得不容易。我有看過銀行主動寄給我創業貸款的DM，也記得很久以前有聽過銀行在廣播節目中廣告貸款的商品，但現在好像都沒了。政府部門的宣導我沒有印象有耶；然後我姐有辦過房貸啦，如果我要向銀行借錢，應該會去問她。
- 15：我的電子信箱很多廣告啦，都是借錢的廣告。

（二）銀行授信業務的利息和收費價格的刻板印象

女性消費者普遍認為銀行授信業務的利息和收費情形，是很不透明的。大部分的人都認為：和銀行業者有關係，或者認識銀行裡面的人（辦理授信業務者），利息和費用會收的比較低，僅有大學生AMY認為利率的收取應該有公式計算，不會大家都一樣，和有向銀行借款的王小

姐、張老師知道銀行會根據職業、收入等等因素考量利率高或低，而且現在辦貸款都會收手續費。顯見女性消費者對於銀行授信業務的訂價，存在著資訊不對稱的現象。而形成刻板印象的原因主要是傳媒報導及聽他人的口語傳播，次要的原因是主觀的判斷。

- 1：應該是大公司或者是和銀行有關係的人利息會比較低吧。收費應該也是這樣阿。報紙不是常登嗎？
- 2：會根據不同的職業有不同的標準吧。

- 3：我認為信用借款的利息比較高，而且聽說還會收手續費，但是收多少我就不知道了。
- 4：利率的收取應該有公式計算吧，應該不會大家都一樣。報紙上有看過辦理貸款要收手續費。
- 5：我是團體房貸耶，公司員工都統一，細節我並不是很清楚。
- 6：我的貸款是我侄子辦的，我也沒有問過。
- 7：報紙不是說現在的房貸利率都很低嗎，至於怎麼算我就知道了。收費？要收什麼費用？
- 8：我是辦勞貸，政府不是有補貼利息，別人的我不知道耶。
- 9：以我的經驗銀行會根據職業、收入等等因素考量利率高或低，而且現在辦貸款都會收手續費。
- 10：我想和銀行有認識的話，應該利率會比較低。
- 11：現在利率應該都很低吧，聽別人說很競爭喔。
- 12：聽我同事說，利息一定要向銀行爭取。既然是信用借款應該就不用抵押吧，如果是房貸當然就要。
- 13：我覺得要貨比三家，要會向銀行爭取利率才會低。收費應該也是要爭取才會算比較少。
- 14：聽說銀行很競爭，利息也可以殺價吧。
- 15：聽說應該是有關係的人利息會比較低吧。有關係的人還要付什麼費用嗎？

（三）銀行授信業務准駁與有無抵押品的刻板印象

女性消費者普遍認為銀行授信業務的准駁，必須透過和銀行有關係的人來談，會比較容易核准；並且普遍的認同要向銀行借款必須提銀行擔保品，因為銀行是很勢利的。僅有少部分的人認為銀行會

有一套準則來決定授信的准駁。大部分的人都認為：和銀行業者有關係，或者認識銀行裡面的人（辦理授信業務者），應該會很快就核准，而且必須要有不動產提供給銀行抵押，顯見女性消費者對於銀行授信業務的准駁，也存在著資訊不對稱的現象。而形成刻板印象的原因主要是主觀的判斷及聽他人的口語傳播。

- 1：聽說和銀行經理多交流，利息應該會比較好談吧。
- 2：房貸就一定要有抵押阿，信用借款就不用，但收入不高的人應該不好核准，利率會有計算公式吧。
- 3：准或不准銀行應該會有一套準則有所依循，不過我想如果有不動產給銀行抵押，應該會比較容易核准。
- 4：老師說：銀行有標準作業流程。
- 5：我認為要向借錢一定要有東西給銀行啦，不然銀行怎麼會放心借錢給你。
- 6：當然要抵押，利息就和銀行經理談阿。
- 7：我想如果我有正常和穩定的收入，銀行應該會比較容易核准吧。不然如果我沒有工作，銀行怎麼會敢把錢借給我。
- 8：只要和銀行夠熟，甚麼條件都不是重點。
- 9：我覺得有房子抵押給銀行，銀行應該就會核准，所以要和銀行借錢一定要有不動產。
- 10：和銀行有認識有關係的話，應該就會准，有沒有抵押就不一定。報紙不是登很多大公司不是借了很多的信用借款嗎？
- 11：不一定啦，信用放款也是有阿，只不過利息會高一點。
- 12：聽說都要抵押吧。有抵押就會准阿。
- 13：我認為准不准應該是要看借的人是不是還得起，但是聽說沒有房子給銀行抵押比較難借。
- 14：這我不知道耶。
- 15：大家知道當然要有不動產抵押阿，銀行最現實了，沒東西他們是不會准的。

伍、結論與建議

本研究經由針對女性消費者進行深度訪談後，探索出女性消費者對於銀行的第一印象，及女性消費者對於銀行授信業務刻板印象形成的原因受到「口語傳播」、「報章雜誌」、「主觀意識」等三面向資訊來源的影響。

一、對銀行的第一印象

經彙總訪談的大量語言資料後可知，當提到「銀行」這二個字時，女性消費者普遍想到的就是「錢」，以及和「錢」有關的工作，諸如：存錢、領錢、匯款，也是可以用新台幣兌換外國貨幣的地方，缺錢的時候可以到銀行借錢，在銀行可以申請信用卡，也可以申請支票。此部分與銀行法第3條規定銀行業能夠經營的業務是相同的，顯見女性消費者對於「銀行」的認知正確，並沒有其他特殊的刻板印象。

但是女性消費者對銀行服務態度的第一印象，普遍認為公營銀行職員態度比較不好，感覺很官僚，但是也有少部份人員例外。民營銀行的行員感覺上則比較有禮貌，但是也時候感覺太做作，很假。歸納而言，大多數的女性消費者對銀行業者及服務人員態度的看法，仍屬正面。

二、對於銀行授信業務刻板印象

女性消費者普遍認為對於銀行授信

業務資訊的蒐集是不容易的，甚至大多數的人都說政府沒做過這些事，而且銀行業者本身對於授信業務資訊的公開管道也不多。在授信業務的利率計算、收費情形，以及授信准駁的核定方面，女性消費者普遍認為是很不透明的，大部分的人都認為：大公司、和銀行業者有關係的，或者是認識銀行裡面的人（辦理授信業務者），會比較容易核定借款的申請，而且利率和費用也會比較低，甚至有可能不需要費用。並且普遍認為要向銀行申請借款必須提供不動產給銀行設定抵押。

女性消費者在銀行授信業務利率和收費情形的思考上，並不若邱潤容所提出，銀行在受理承做授信業務時，應遵循的五項原則。女性消費者對這部份的刻板印象仍停留在「利用關係」、「人情」或「特權」的思考模式。此外女性消費者對銀行辦理授信業務時，是否一定要有抵押物的刻板印象，大多數都認為銀行是很勢利的，所以必須要有抵押物，此部份與銀行業者處理授信業務時所思考的授信原則（即通稱之授信5P原則），亦有值得玩味之處。蓋因女性消費者普遍認為銀行僅看有無擔保品，即得以決定准駁，殊不知銀行業與「當舖」不同，除了「債權保障」這個因素之外，上有其他四項重要因素亦應一併考量。

三、對於銀行授信業務刻板印象形成的原因

刻板印象的形成，必然有其原因。經彙總訪談的大量語言資料後可知，女性消費者普遍認為從親朋好友獲得的資訊，相較於報紙或媒體所報導的較為真實與實際，也有人認為來自親朋好友的經驗比較客觀，因此造成女性消費者對於銀行授信業務刻板印象的主因為「他人口碑宣傳」。

在日常生活之中，大眾傳播媒體是接觸頻率最高的資訊來源媒介，無論是報紙、電視、雜誌，都可能造成女性消費者對於銀行授信業務刻板印象。在本研究中，不少女性消費者對於銀行授信業務的資訊來源是報紙，因此傳媒的報導是造成本項刻板印象的原因之一。

人類是具備思考能力的，在本研究中得知，無論是否具有向銀行申請借款的經驗，多數的女性消費者對於銀行授信業務，是用「想當然爾」的思考邏輯來判斷。因此「主觀意識」亦是造成女性消費者對於銀行授信業務刻板印象的原因之一。

四、對經營與行銷建議

自從國內開放新銀行的設立之後，銀行業即走進了經營的困境，開拓新的顧客勢必將成為解決經營困境的主要生存之道。對於銀行業者而言，授信業務所產生

的利差，仍是銀行業獲利的主要來源，在傳統的觀念中，從事會計工作以及與錢財有關的工作，大都數是女性的人數比率較多，一般家庭中掌管錢財和經濟問題者也多以女性為主，但是銀行的授信業務對於掌管金錢大權的女性消費者而言仍是個陌生的市場環境，在開拓女性消費者市場之前，有必要先瞭解女性消費者對於銀行的授信業務具有哪些刻板印象？及這些刻板印象形成的原因為何？因此，根據本研究的發現，提出如下的建議，給予銀行業者在經營與行銷領域上做為參考：

- (一) 銀行業者應與政府有關部門，共同加強傳媒報導，使資訊公開透明，將銀行辦理授信的各項原則教育消費者於潛移默化之中。
- (二) 銀行業者可善用「口碑行銷」力量，加強內部服務人員的訓練，提高專業度與改善服務品質，藉由既有客戶協助行銷，但亦應適度減少不必要的或過度的服務，以防適得其反。
- (三) 銀行業者可藉傳媒之力，將承辦的成功案例廣為宣傳（如輔導創業成功案例），除社會責任之外，亦可提升企業專業形象。

五、對後續研究建議

本研究為探索女性消費者對於銀行的授信業務板印象以及刻板印象形成原

因的質性研究，在有限的時間、人力、成本以及研究者本身學術背景的影響下，有其適用範圍與限制存在。因本研究的樣本有限，無法作全面性推論，且囿於質性研究的特性，有部份結果受研究者主觀判斷影響，因此建議後續研究者：

- (一) 增加更多地區的研究樣本，若能以全球各地區的國家為樣本地區，則研究將會更加完整。
- (二) 擴大研究層面，因本研究僅從女性消費者的角度探討對於銀行授信業務的刻板印象，因此後續研究者若能針對全部消費者，或是由銀行經營者的角度進行深度訪談，再比較服務提供者與接收者兩者之間對於刻板印象的異同，如此亦能夠有更全面性的瞭解。
- (三) 本研究以深度訪談法整理出女性消費者對於銀行授信業務刻板印象，以及口語傳播」、「報章雜誌」、「主觀意識」三面向資訊來源的影響。後續研究者可以量化研究的方式發展量表，以利全面性與進階性的推論。

參考文獻

- [1] 王文科(2000)，質的教育研究法。台北：師大書苑。
- [2] 古永嘉(2003)，企業研究方法。台北：美商麥格羅希爾。
- [3] 李美枝(1979)，社會心理學。台北：大洋。
- [4] 李美枝(1990)，女性心理學。台北：大洋。
- [5] 李天任、藍莘譯(Wimmer & Dominick原著，1995)。大眾媒體研究。台北：亞太圖書。
- [6] 李美華等譯(Earl Babbie原著，1998)，社會科學研究。台北：時英出版社。
- [7] 李卓夫(2000)，大學生性別角色刻板印象與兩性平權態度關係之研究。國立成功大學教育研究所未出版碩士論文。
- [8] 邱潤容(1998)，銀行實務。台北：三民。
- [9] 林明地(2002)，質的研究實例舉隅：校長領導的參與觀察。國立中正大學教育研究所主編，質的研究方法，P51-87。高雄：麗文。
- [10] 林文瑛(2002)，性別角色與性別角色發展。載於晏涵文、黃富源，性別平等教育，35-55。台北：展智。
- [11] 林麗珊(2002)，性別差異與性別平等的教育。載於晏涵文、黃富源，性別平等教育，16-34。台北：展智。
- [12] 胡幼慧、姚美華(1996)，一些質性方法上的思考。在胡幼慧編，質性研

- 究。臺北：巨流。
- [13] 柯淑敏 (2001) 兩性關係學。台北：揚智文化。
- [14] 高敬文 (1988)，質的研究派典之理論分析與實際應用。屏東：東益。
- [15] 晏涵文 (2004)，性、兩性關係與性教育。台北：心理。
- [16] 張春興 (1996)，教育心理學：三化取向的理論與實踐。台北：東華。
- [17] 陸谷孫主編 (1992)，英漢大辭典。台北：東華。
- [18] 黃德祥 (2000)，青少年發展與輔導。台北：五南。
- [19] 黃葳葳 (2002)，從刻板印象看網路色情，臺北市終身學習網通訊(18)。
- [20] 黃俊英 (2006)，行銷研究概論。台北：華泰文化。
- [21] 溫世頌 (2006)，心理學辭典。台北：三民。
- [22] 楊國樞 (1993)，社會及行為科學研究法。台北：東華。
- [23] 劉淑雯 (1996)，平面與立體的歷程：兩性平等教育之卷宗評量。教育資料與研究，13，50-54。
- [24] 羅燦慶 (2003)，突破雜誌(二月號)，P42-43。
- [25] Basow, S. L. (1992). Gender stereotypes and roles, California : Brook/Cole Publishing Company.
- [26] Devine, P. G. (1989). Stereotypes and prejudice : Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, P5-18
- [27] Miller, W. L., & Crabtree, B. F. (1992). Primary Care Research : A Multi-Method Typology and Qualitative Road Map, in Miller & Crabtree (eds.), *Doing qualitative research*. Newbury Park, California : Sage Publications.
- [28] Silverblatt, A. & Eliceri E. M. E. (1997). *Dictionary of Media Literacy*. Westport : Green wood press.
- [29] Zikmund, W. G. (1994). *Business Research Methods*, 4th ed., Texas : Dryden.
- [30] 銀行法，取自全國法規資料庫網站
- [31] 沈美杏 (2007)，論性別角色刻板印象之內涵。網路社會學通訊，第 61 期，2007 年 3 月 15 日。取自：http://www.nhu.edu.tw/~society/e-j/61/61_88.htm

以分散式雜湊機制設計無線隨意網路之需求 遶徑協定

張兆村

摘要

在無線的 Ad Hoc Network 中，需求遶徑協定（On-demand Routing Protocols）是一種利用封包達到即時尋找及建立通訊路徑的協定。在先前以 DSR 為基礎的相關研究 [1][2][5][8][9] 中，遶徑找尋（Route Discovery）通常會為探知網路的現況而產生大量的需求封包，但也因此完成有用資訊的收集，以做為未來遶徑決策（如最短路徑計算）的依據。然而，通信裝置上的大量遶徑快取（Routing Cache）需求，將因此大幅增加空間需求及遶徑計算的複雜度。本篇論文將提出一個新的分散式雜湊遶徑協定（Disributed Hashing Routing Protocol; DHRP），此協定藉由雜湊式快取（Hashing Cache）方式以降低主機快取需求量及遶徑計算的複雜度。經實驗效能之結果顯示，所提出的協定相較於相關研究，具有明顯地降低快取儲存體之需求量、遶徑計算之複雜度、重新遶徑所產生的控制封包之負擔、及平均遶徑長度等的優點。

關鍵詞：雜湊、雜湊式遶徑、需求遶徑、無線隨意網路。

Distributed Hashing Mechanisms for On-demand Routing Protocol in Wireless Ad Hoc Networks

Chao-Tsun Chang

Abstract

In wireless ad hoc network, on-demand routing protocols are proposed for discovering and establishing a route only when a source host originates a data packet addressed to the destination host. In the DSR-based study [1][2][5][8][9], route discovery usually raises a large number of request packets for exploring the current state of the network, but it also performs the collection of useful information for the future routing decisions, such as the computations for the shortest path. However, large routing caches in devices will increase the spatial and routing complexities. This paper presents a novel distributed hashing routing protocol (DHRP) by applying the hashing cache scheme to reduce the cache size of the hosts and the routing complexity. Performance results reveal that the proposed protocol significantly reduces the required size of cache storage, increases routing complexity and reduces the control packet overhead for rerouting and the average route length.

Keywords : hash, hash-based routing, on-demand routing, ad hoc network.

壹、簡介

無線隨意網路 (Wireless Ad Hoc Network) 是由一群無線行動主機所動態構組而成的網路，其網路並無明顯的建構形狀或集中式的管理機制。在此種行動式的網路環境中，由於每個行動主機的無線通訊距離有限，致使在此網路內之行動主機若要傳送訊息給其他主，可能必須獲得沿途上其它行動主機的幫忙以代傳訊息。無線隨意網路相當適合應用於：分散式協同計算 (Collaborative Computing)、多媒體教室 (Multimedia Classroom)、行動教室 (Mobile Classroom)、災害搶救 (Disaster-relief)、搜尋與救難 (Search-and-rescue)、法律的執行 (Law Enforcement)、距離的學習與推算 (Distance Learning)、軍事及戰場運用、及其它特殊用途上的運用[3]。

遠徑協定在網路通訊上，負責尋找及建立最佳傳輸路徑的任務。尋找及建立路徑的方法，大致可分為表格驅動 (table driven) 及依需求 (on-demand) 等兩種。表格驅動的遠徑協定是藉由集中式蒐集所有的遠徑資訊，以利於建立最佳的通訊路徑。而依需求遠徑協定 (On-demand Routing Protocol) 是一種依需求即時尋找及建立通訊路徑的協定。然而，在無線通訊的行動性應用上，

對於尋找及建立路徑的代價、遠徑的最佳化與路徑的維護效益等因素上更顯得遠徑協定設計的重要。

在以往的文獻中，已有許多依需求機制的遠徑協定應用於隨意網路 [1][2][3][4][5] 等。其中，DSR (Dynamic Source Routing Protocol) 是利用歷史記錄來簡化建立通訊路徑時所需的負擔。當某個主機 s 開始要傳送資料封包給目的主機 d ，且主機 s 並不知道有已知的歷史路徑可到達目的地時，主機 s 將啟始一個新的遠徑找尋 (Route Discovery) 動作，並於網路內以氾流 (Flooding) 的方式來發送 RREQ (Route Request) 封包。一旦 RREQ 封包傳到主機 d 或任一個已知道如何遠徑到主機 d 的主機時，該主機即將其所儲存遠徑到主機 d 的資訊傳回主機 s ，以 RREP (Route Reply) 封包回應遠徑找尋的需求。DSR 依靠兩種主要機制來運作：遠徑找尋及遠徑維護。遠徑找尋程序被應用於當主機 s 要傳送資料封包到主機 d ，而尚不知道有已知的歷史路徑可到達主機 d 時。

依需求遠徑協定是一種與目的主機建立通訊路徑的協定，然而此協定之不同處在：主機於平時並不需要建立通訊路徑或維護路徑，當來源主機要送出資料封包時，才動態的即時建立通訊路徑。為了避免每次要

傳送資料封包前，皆須完成費時的繞逕找尋程序，因此，以 DSR 為基礎的依需求繞逕協定必須儲存先前已被使用繞逕資訊。由於主機快取的空間有限，無法儲存大量的歷史路徑，故其快取的管理策略亦有待改善。

隨著繞逕找尋次數的成長，各主機儲存繞逕的空間需求量亦增加，然而要求每個主機備有無限量儲存空間的假設，有其實際運作的困難。再者，在文獻[2]所提出的方法中，面對大量的繞逕計算（Routing Computation），其複雜度約為 $O(n^2)$ ，計算能力較弱的主機亦將在繞逕延遲上，被視為一項重要的挑戰。若對快取空間做限制[2]，則快取管理的置換（Replacement）將會面臨在效能與路徑學習能力上的其它挑戰。

目前個人無線通信設備廣泛被使用，且逐漸於該些設備內建稍弱的數據運算能力及少量記憶空間，如行動手機、個人數位助理（PDA）、配具藍芽技術之智慧家電（IA）等，以提昇無線通訊之應用範圍及降低設備價格。若由該些無線通訊設備構組成隨意網路，在繞逕協定上採用簡單的 DSR 為基礎的繞逕協定是直接可行的方法。

在 DSR 的實現方面，基本上將有兩個主要的關鍵因素尚待改善：計算能力（Computing Power）及空間限制（Space Limitation）。近年來 Yi-Chun Hu 及 David B.

Johnson[2]針對 DSR 的記憶體管理問題提出快取管理策略（Caching Strategies）以節省記憶體空間，及 M. Spohn 及 J.J. Garcia-Luna-Aceves[5]亦提出強化繞逕修護（Route Repair）的作法，以提昇該類協定之實用性。然而[1][2][5]於時間複雜度（Time Complexity）及快取置換（Cache Replacement）方面仍有可供精進之處。另一種快取繞逕選擇的方法 DSR-Cache[8]被提出，動態選擇高剩餘電量的主機做為虛擬主幹（Virtual Backbone），以提高繞逕封包對目的主機的高快取命中率及達到網路電量的最佳化利用，此方法雖有效降低繞逕封包廣播所造成的電量消耗，然而卻無法找到較短的繞逕路徑。此外，研究[9]提出一 Ant-DSR 的繞逕協定，以分散式拓樸找尋方式來維護 DSR 的繞逕資訊，以降低主機移動所造成的繞逕失效機會及網路負擔，然而此協定仍須依賴大量快取空間及耗時的繞逕找尋運算。

雜湊技術已有效地被廣泛利用於資料的快速搜尋應用上，且網際網路的 WWW 上也有相關的研究[6][7]被討論。本論文將以雜湊的機制為基礎，並同時利用關聯式硬體（Associative Hardware）的記憶體存取與平行比對功能設計，使繞逕找尋計算的時間複雜度由 DSR 的 $O(n^2)$ 大幅降低為 $O(1)$ ；

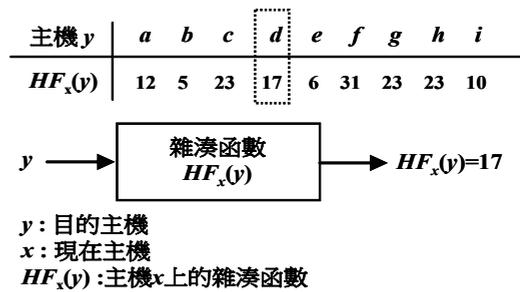
若以軟體實現，亦可控制於 $O(C(k_i, 2))$ ，其中 $host_i$ 及 k_i 分別代表主機 i 及其相鄰主機的數目。而主機採用雜湊技術亦可確保快取的儲存數量是有限的 (Bounded)，若快取的列大小為 r ，則快取大小的限制將為 $\sum C(k_i, 2) * r$ ，其各別 k_i 值可透過隨意走動 (Random Walk) 方式及主機分佈實驗獲得可行之值。因此，本論文所提出的雜湊機制將可有效減少快取所需的儲存空間、降低遠徑找尋的複雜度及提昇 DSR 為基礎遠徑協定的整體效率。

本論文的其餘內容安排如下，在第二節中，我們將探討分散式雜湊技術及雜湊快取的設計，並於第三節舉例說明分散式雜湊遠徑協定 (DHRP) 的遠徑運作細節，第四節將討論與相關研究的定性分析，而第五節探討 DHRP 的實驗與結果討論，最後一節做簡單的結論。

貳、雜湊函數及雜湊快取設計

DSR 為基礎的相關遠徑協定主要是利用每個行動主機的快取來記錄以往曾經建立過的歷史路徑，以增進未來建立通訊路徑的效率。然而其龐大的遠徑資訊儲存空間、可能造成多個來源主機建立局部共同的遠徑及過長的遠徑結果等問題。Yi-Chun Hu 及

David B. Johnson[2]對於 DSR 的快取管理策略採取不同的設計，藉以分析及觀察在封包傳送率、遠徑負擔、延遲、及路徑最佳化等各方面的效能表現。他們藉由快取結構、快取容量、及快取計時更新等三項策略的設計來改善 DSR 在快取方面儲存與管理的能力。



(a) 雜湊函數 $HF_x(y)$ 的用法

		主機 x 的雜湊表 (第 i 位元值)															
pre	suc	0	1	...	5	6	...	10	11	12	...	17	18	...	23	...	31
a	c	0	0	...	1	0	...	0	0	0	...	1	0	...	0	...	0
f	c	0	0	...	0	0	...	0	0	0	...	1	0	...	1	...	0
f	a	0	0	...	0	0	...	1	0	1	...	1	0	...	0	...	0
c	a	0	0	...	0	0	...	0		1	...	0	0	...	0	...	0
a	f	0	0	...	0	0	...	0	1	0	...	0	0	...	0	...	1
c	f	0	0	...	0	1	...	0	0	0	...	0	0	...	0	...	1
...

(b) 主機 x 的雜湊快取結構範例

圖 1、雜湊函數與雜湊快取的範例

在快取結構方面，雖然連結快取 (Link Cache) [2] 的設計較路徑快取 (Path Cache) [1] 更具連結分享及學習 (Link Sharing and Learning) 的優點，而採用較小的快取空間

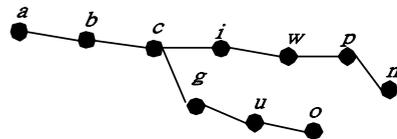
亦可間接提昇效能，然而，採用連結快取且使用複雜的圖形搜尋演算法（如 Dijkstra's 最短路徑演算法），將使計算的複雜度增加。另外，[2]在置換策略方面，採用固定計時（Static Timeout）與調整式計時（Adaptive Timeout）兩種模式，發現各有其優缺點，但若未加入計時，則明顯在效能上會表現較差，此現象說明，主機的行動性對 DSR 設計的影響相當重要。

本論文所提出的 DHRP 協定中，假設無線通訊設備已內建雜湊快取（雜湊快取）與關聯式硬體的記憶體存取與平行比對功能設計。以下，我們分別討論雜湊函數及雜湊快取的設計。

一、以雜湊函數表示邊徑資訊

在無線隨意網路中，假設主機 x 有個 RREQ 封包正在尋找往目的主機 y 的通訊路徑，當主機 x 收到此封包時，將檢查自己是否為目的主機，若否，則應有能力經雜湊快取的查詢而得知該將此封包往其相鄰的那一個主機傳送，才能正確地將 RREQ 封包傳送至目的主機 y 。因此，雜湊函數的設計將採用兩參數 x 及 y ，其中 x 表示主機自己的 ID，而 y 表示目的主機 ID，而雜湊函數 $HF_x(y)$ 所運算出來的值，應能用來決定要將 RREQ 封包傳送給那個相鄰的主機。舉例而

言，如圖 1 (a) 所示，若主機 x 收到一個 RREQ 封包，而在此封包中所描述的目的為 d ，則經 $HF_x(d)$ 的運算可知結果為 17。因此，主機 x 應能透過雜湊快取，如圖 1 (b) 所示，查詢各快取列內第 17 位元，若該位元值=1，則表示決定將此 RREQ 封包送往該列中 *suc* 欄內所指定的鄰近主機 a 及 c 傳送，若該位元值=0，則表示往 *suc* 欄內所指相鄰主機的走法並無經驗可循。而此種雜湊快取內對某位元的平行比對與記憶體存取功能，可由內建的關聯式硬體達成。透過雜湊快取對歷史經驗的儲存及雜湊函數對路徑的計算，將可達到有效率的邊徑找尋目的。



(a) 兩條路徑 an 及 co

		主機 g 的雜湊表						
pre	suc	...	HF_g (u)	...	HF_g (c)	...	HF_g (o)	...
c	u	0	1	0	0	0	1	0
u	c	0	0	0	1	0	0	0

(b) 圖 (a) 中路徑 co 在主機 g 內的快取記錄內容

圖 2：路徑於雜湊快取的儲存範例

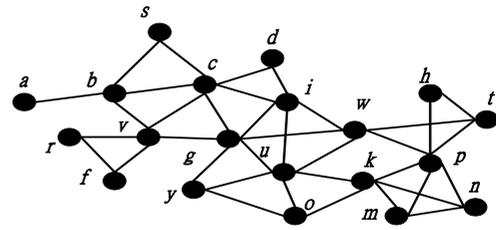
二、雜湊快取的設計

定義一：雜湊快取 (HC_i)

以 HC_i 表示主機 i 實現 DHRP 協定的歷史遠徑資訊集合，亦稱為主機 i 的雜湊快取。雜湊快取是由許多快取列所組成，且每一快取列是由三個欄位資訊所構成：前主機 (Previous) 欄、後主機 (Successive) 欄及雜湊表欄。對於某歷史路徑而言，在快取的結構中，前兩欄為主機 x 的前主機 (以 pre 代表) 及後主機 (以 suc 來代表)，而於雜湊表內則以 $HF_x(y)$ 運算方式表示其雜湊位置 (Slot) 的 index 值。

在 DHRP 協定中，歷史路徑被儲存於雜湊快取內。以下我們將描述如何搭配雜湊快取的設計，使主機 x 有能力決定將 RREQ 封包傳到適當的相鄰主機。如圖 2 (a) 所示，若在主機 c 的快取內已記錄兩條歷史路徑 $an : a-b-c-i-w-p-n$ 及路徑 $co : c-g-u-o$ 。而此時，若主機 c 發出 RREQ 封包欲建立通訊路徑至目的主機 w ，對主機 c 而言，必須能決定要將此 RREQ 封包送往相鄰主機 i 、 g 、 b 中的那些主機。主機 c 必須區分走往下一個主機可到達那些目的主機，以供遠徑決策。例如：走往主機 b 可到達主機 a ，而走往主機 g 可到主機 u 及主機 o ，走往主機 i 可到主機 w 、主機 p 及主機 n 。

以下將針對一歷史路徑說明其路徑上的各主機如何以節省快取空間的方式，將路徑資訊記錄於雜湊快取中。



(a) 一個隨意網路結構之範例

path an : $a-b-c-i-w-p-n$
 path st : $s-c-g-u-k-p-t$
 path fd : $f-v-c-d$
 path co : $c-g-u-o$
 path yt : $y-g-w-t$

(b) 圖 (a) 網路結構之路徑範例

圖 3、多條路徑於隨意網路結構之範例

如圖 2 (a) 的路徑 co 在雜湊快取的設計中，此歷史路徑將被記錄成如圖 2 (b) 內容：我們進一步以 $HF_g(y)$ 表示在主機 g 的雜湊快取內可到達主機 y 所對應的雜湊位元位置，其中當 pre 欄為 c 且 suc 欄為 u 列的 $HF_g(u)$ 位置及 $HF_g(o)$ 位置內的值皆為 1，其表示走往主機 u 將可到達主機 u 及主機 o 。在雜湊快取內同時亦須記錄反向資訊，例如從主機 u 進入主機 g 並往主機 c 走，將可到達主機 c ，即 pre 欄為 u 且 suc 欄為 c 列中 $HF_g(c)$ 值應為 1，其餘各欄內位元值將為 0。

三、雜湊快取的表示方式及相關定義

為方便表達及陳述雜湊函數與雜湊快取在遶徑找尋的運作，以下將給予定義本論文所使用的符號。

定義二：雜湊表的位置索引 (N_{xy})

以 N_{xy} 表示雜湊快取中雜湊表內位置為 $HF_x(y)$ = 第 N 位元的位置索引。

定義三：雜湊表內的位置值 ($SV_{ps}[j]$)

以 $SV_{ps}[j]$ 表示雜湊快取中， pre 欄位為 p 且 suc 欄位為 s 的快取列中，位置為第 j 位元的位置值。

定義四：雜湊表的快取列 ($psVi$)

以 $psVi$ 表示主機 i 內雜湊表的一個快取列內所有值為 1 的位置索引集合，即 $psVi \in HC_i$ ，且 $psVi = \{HF_i(v) \mid SV_{ps}[HF_i(v)] = 1, \forall \text{主機 } v\}$ 。

定義五： $psVi$ 的碰撞 (Collision)

若主機 i 的雜湊快取內，將可到達不同主機 x 及 y 的歷史遶徑資訊，儲存於相同的快取列及相同的位置索引內時，稱主機 i 到主機 x 及 y 的歷史遶徑資訊發生碰撞。亦即 $psVi \in HC_i$ ， N_{ix} 及 $M_{iy} \in psVi$ ，當 $x \neq y$ 且 $N_{ix} = M_{iy}$ ，則稱 $psVi$ 內 N_{ix} 及 M_{iy} 發生碰撞。

如上之定義，假設圖 1 (a) 中的 $HF_x(y)$ 是代表 pre 欄為 x 且 suc 欄為 y 列的位置索引，則 $HF_x(d) = 17$ ，將以 17_{xy} 表示該快取列中雜湊表的第 17 個位置索引。同時圖 1 (b) 中 faV_x 的資訊集合為：

$$faV_x = \{HF_x(v) \mid SV_{fa}[HF_x(v)] = 1, v \in \{i, a, d\}\}$$

，現將改以 $faV_x = \{10_{xi}, 12_{xa}, 17_{xd}\}$ 表示。在 psV_x 中發現 23_{xc} 、 23_{xg} 及 23_{xh} 皆表示相同該雜湊表內第 23 位元的雜湊值為 1，這表示該快取列的雜湊表內發生碰撞，稍後章節中將會再介紹碰撞的相關討論。同時，假設圖 2 (b) 中 $HF_g(u) = 12$ 、 $HF_g(o) = 25$ 及 $HF_g(c) = 6$ ，則歷史路徑 co 在主機 g 的兩個快取列的雜湊表將表示成 cuV_g 、 $ucV_g \in HC_g$ 、且 $cuV_g = \{12_{cu}, 25_{co}\}$ 及 $ucV_g = \{6_{uc}\}$ 。

各主機的雜湊快取內容主要依據為歷史路徑。以下說明歷史路徑為建構內容之範例，依圖 3 (a)、(b) 來介紹各主機的雜湊快取內記錄的資訊。如圖 4 (a)、(b) 所示，其中為便於說明，內容 N_{xy} 皆為假設值。在圖 4 (a) 內，我們會發現快取內並不需要儲存 baV_a 、 csV_s 、 gcV_c 、 uoV_o 、 gyV_y 、 wiV_i 、 piV_i 、 pnV_n 等資訊，因為如 baV_a 表示走向主機 a 後將可到達主機 a 的資訊，明顯是多餘的。基本上，DHRP 協定對快取空間的需求與所儲存的遶徑路徑數量相關，但是，因為相同快取列內的雜湊值可共用空間相同的位元。若

k_i 為主機 i 的相鄰主機數量且快取列大小為 r ，從 pre 及 suc 欄內的所有可能情況組合，則可估算出主機 i 的最大快取列數量為 $C(k_i, 2)$ ，亦可估算得快取大小的限制為 $\sum_i C(k_i, 2) * r = \sum_i [k_i * (k_i - 1) / 2] * r$ ，其各別 k_i 值可透過隨意走動及主機分佈實驗獲得可行之值。

四、雜湊函數的設計實例

本文採用式 (1) 為實驗中之雜湊函數。假設 x 為 RREQ 封包目前所在的主機 ID，而 y 為目的主機 ID：

$$HF_x(y) = y \bmod A[x \bmod t] \quad (1)$$

其中 $A[0..t-1]$ 為一正質數陣列且

$$A[i] < A[j] \quad \forall i, j \quad 0 \leq i < j \leq t-1,$$

式中 $A[0]$ 及 $A[t-1]$ 分別表示正質數陣列內的最小與最大之質數，而各主機內所建立的雜湊表其平均位置數目為

$$S_{ht} = (A[0] + A[t-1]) / 2.$$

S_{ht} 值的大小，顯然與雜湊碰撞的機率相關。

Set	Elements of Set	Set	Elements of Set
HC_a	abV_a	HC_w	ipV_w, piV_w
HC_b	acV_b, caV_b	HC_u	$goV_u, ogV_u, gkV_u, kgV_u$
HC_c	$sgV_c, gsV_c, biV_c, ibV_c, cgV_c$	HC_g	$ywV_g, wyV_g, cuV_g, ucV_g$
HC_s	scV_s	HC_k	upV_k, puV_k
HC_o	ouV_o	HC_p	$kaV_p, ikV_p, wnV_p, mvV_p$
HC_y	ygV_y	HC_t	nvV_t, ipV_t
HC_i	cvV_i, wcV_i	HC_n	npV_n

(a) 圖 3 中五條 path 對應的雜湊快取

Set	Elements of Set
abV_a	$12_{ab}, 23_{ac}, 5_{ai}, 18_{av}, 27_{ap}, 3_{an}$
acV_b	$4_{bc}, 9_{bi}, 16_{bv}, 15_{bp}, 25_{bn}$
caV_b	13_{ba}
sgV_c	$11_{cg}, 18_{cu}, 2_{ck}, 5_{cp}, 17_{ct}$
gsV_c	6_{cs}
biV_c	$17_{ci}, 12_{cw}, 7_{cp}, 19_{cn}$
ibV_c	$4_{cb}, 18_{ca}$
cgV_c	$11_{cg}, 18_{cu}, 7_{co}$
scV_s	$3_{sc}, 15_{sg}, 25_{su}, 31_{sk}, 10_{sp}, 21_{st}$
ouV_o	$12_{ou}, 17_{og}, 32_{oc}$
ygV_y	$23_{yg}, 29_{yw}, 8_{yt}$
cvV_i	$8_{iv}, 13_{ip}, 19_{in}$
wcV_i	$18_{ic}, 21_{ib}, 30_{ia}$
ipV_w	$1_{wp}, 22_{wn}$
piV_w	$14_{wi}, 28_{wc}, 29_{wb}, 20_{wu}$
goV_u	21_{uo}
ogV_u	$14_{ug}, 25_{uc}$
gkV_u	$12_{uk}, 23_{up}, 31_{un}$
kgV_u	$14_{ug}, 25_{uc}, 25_{us}$
ywV_g	$0_{gw}, 26_{gt}$
wyV_g	20_{gv}
cuV_g	$15_{gu}, 19_{go}, 1_{gk}, 5_{gp}, 26_{gt}, 17_{ip}$
ucV_g	$6_{gc}, 30_{gs}$
upV_k	$4_{kp}, 18_{kt}$
puV_k	$12_{ku}, 23_{kg}, 31_{kc}, 7_{ks}$
kaV_p	29_{pt}
ikV_p	$2_{pk}, 10_{pu}, 21_{pg}, 4_{pc}, 16_{ps}$
wnV_p	23_{pn}
mvV_p	$21_{pw}, 14_{pi}, 27_{pc}, 26_{pb}, 0_{pa}$
nvV_t	$17_{tv}, 9_{tg}, 18_{ty}$
ipV_t	$28_{tp}, 30_{tk}, 14_{tu}, 9_{tg}, 19_{tc}, 23_{ts}$
npV_n	$5_{np}, 2_{nw}, 27_{ni}, 31_{nc}, 15_{nb}, 7_{na}$

(b) 圖 (a) 每一 cache row 組成

圖 4：圖 3 (b) 內各路徑的雜湊快取表示範例

採用雜湊技術做為快取策略，一般情況下都必須面對碰撞引起的相關問題。若我們採用的雜湊函數為 $HF(y) = y \bmod S_{ht}$ ，亦即，僅以 y 當作雜湊函數的參數，而不將 x 納入參數的考量。如此，將更易造成在主機 x 對不同目的主機 t 及 t' 代入雜湊函數計算時，產生 $HF(t) = HF(t')$ 的機率增加，而發生誤將要到目的主機 t 的 RREQ 封包引導到主機 t' 時才發現錯誤。因為該誤導情形發生於相同快取列下的碰撞狀況，故我們估計發生此種誤導機率 P_{error} 為 $(C(k, 2) * S_{ht})^{-1}$ ，其中 k 為平均各主機的相鄰主機個數。為了有效解決此問題，我們採用公式(1)已討論之 $HF_x(y)$ ，即所謂之分散式雜湊機制，將主機 x 本身的 ID 納入雜湊函數的參數中，以造成各主機內記錄可到達主機 y 的 $HF_x(y)$ 值相同的機率降低，而將要到目的主機 t 的 RREQ 封包誤導到主機 t' 的機率幾乎降至 0。假設因碰撞造成連續 i 個主機皆無法辨識誤傳，則其誤導機率為

$$P_{error} = (C(k, 2) * S_{ht}^i)^{-1} \quad (2)$$

。連續 i 個主機皆發生誤導的機率如圖 5 所示，明顯發現 P_{error} 有效控制程度為 i 優先於 S_{ht} 且 S_{ht} 又優先於 k 。

當然，假使希望 $i \leq 3$ 以控制發生錯誤的機率，則必須有效控制 S_{ht} 值，而質數陣列

中質數值越大將使得錯誤的機率越小，但是卻使得快取大小變得越大。我們在經驗上的建議是，當主機上的快取大小是受限制時，依 DHRP 快取大小限制的估計值，反推以計算出快取列的大小及質數陣列中的 $A[0]$ 值，並選擇出適當的 t 值。

i	k	S_{ht}	P_{error}
2	5	200	$2.5 * 10^{-6}$
2	5	300	$1.1 * 10^{-6}$
	5	200	$1.25 * 10^{-8}$
3	5	300	$3.7 * 10^{-9}$
3	6	300	$2.46 * 10^{-9}$

圖 5：連續 i 個主機皆發生錯誤的機率

參、DHRP 協定的運作

DHRP 基本上是採取連結狀態 (link-state) 方式來表達路徑的邊徑資訊。本節將運用前面已說明的路徑的儲存與雜湊快取的表示方式，說明 DHRP 的邊徑找尋演算法，及以圖 3 及圖 4 範例資訊為例，說明它的運作情形與相關之特徵探討。

一、DHRP 的封包型態及邊徑資訊交換

DHRP 會使用三種類型的控制封包以進行邊徑資訊交換，包括 RREQ (Route Request)、RREP (Route Reply)、及 RERR (Route Error)。其中 RREQ、RREP 用於邊

徑找尋，RERR 用於產生遠徑失誤。在傳輸封包時，RREQ 是不可信賴的廣播方式傳遞，RREP 及 RERR 則都是可信賴的單點傳遞方式（unicast）傳輸封包。

RREQ 封包格式為（廣播 ID, 來源主機 ID, 目的主機 ID, 最大跳躍數）。廣播 ID 由來源主機在每次送出 RREQ 封包時逐次累增，它可與來源主機 ID 構成 RREQ 的唯一辨識資訊。來源主機 ID 及目的主機 ID 用於記錄 RREQ 封包所要找尋遠徑的來源主機及目的主機位址。而最大跳躍數則用以控制 RREQ 廣播封包的汨流範圍。

當 RREQ 封包傳遞到目的主機後，目的主機會以 RREP 封包回應及蒐集 RREQ 封包所產生之遠徑結果。RREP 封包格式為（廣播 ID, 來源主機 ID, 遠徑路徑）。因廣播 ID 與來源主機 ID 構成 RREQ 的唯一辨識資訊，故回覆遠徑結果時，可用以反向蒐集遠徑沿途之各主機 ID 資訊於遠徑路徑內，直到來源主機為止。如此，來源主機便可知道完整的遠徑結果了。

二、DHRP 的遠徑找尋範例

本節將以圖 3、圖 4 為例，說明 DHRP 協定的遠徑找尋如何進行。若以圖 3 (a) 中的主機 b 為來源主機欲找尋遠往主機 t 的路徑，如以 DHRP 遠徑找尋演算法運作，其

RREQ 封包遠徑過程將如圖 6 (a) 所示，而其 RREP 封包的回覆過程將如圖 6(b) 所示。

[DHRP 遠徑找尋 (來源主機) 演算法]

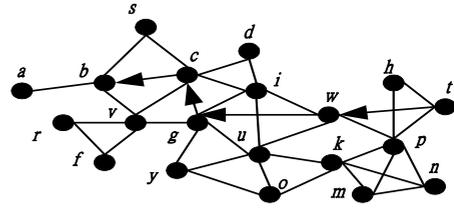
```
//Algorithm for the source host
//visited ( $B_{ID}, x$ ) is a Boolean function to test
//whether host  $x$  had received the RREQ packet
//  $B_{ID}$ 
//assume the destination host is the host  $d$ 
1. compute  $N = d \bmod A[s \bmod t]$ ;
2. compute  $T = \{p \mid N_{sd} \in_{kp} V_s, \forall k\}$ ;
3. if  $T = \emptyset$  then set  $T = \{p \mid p \in neighbor(s)\}$ ;
   (assume  $T = \{p_1, p_2, \dots, p_{|T|}\}$ )
4. for  $i = 1$  to  $|T|$ 
   send a RREQ packet ( $B_{ID}, s, d, h_{count}$ ) to host  $p_i$  ;
```

[DHRP 遠徑找尋 (回覆) 演算法]

```
//Algorithm for relay hosts
//visited ( $B_{ID}, x$ ) is a Boolean function to test whether
//host  $x$  had received the RREQ packet  $B_{ID}$ 
//assume the destination host is the host  $d$ 
1. host  $y$  receives a RREQ packet ( $B_{ID}, s, d, h_{count}$ ) from
   host  $r$ ;
2. if  $y = d$  then {reply a RREP packet ( $B_{ID}, s, d$ ) to track
   the routed path from host  $s$  to host  $d$  by  $B_{ID}$  and
    $s$ ; return;}
3. if  $h_{count} = 1$  then drop packet and exit;
4. if visited ( $B_{ID}, y$ ) then drop packet and exit;
5. compute  $N = d \bmod A[y \bmod t]$ ;
6. compute  $T = \{p \mid N_{yd} \in_{kp} V_y, \forall k\}$ ;
7. if  $T = \emptyset$  then set  $T = \{p \mid p \in neighbor(y) - \{r\}\}$ ;
   (assume  $T = \{p_1, p_2, \dots, p_{|T|}\}$ )
8. for  $i = 1$  to  $|T|$ 
   send a RREQ packet ( $B_{ID}, s, d, h_{count}-1$ ) to host  $p_i$  ;
```

遠徑找尋過程中，不是來源的主機皆會執行遠徑找尋（回覆）演算法，依圖 3、圖

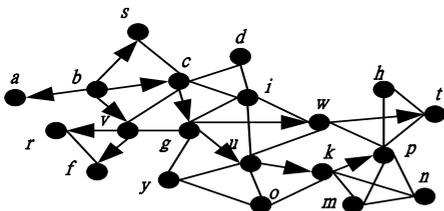
4 四的範例數據，如主機 c 於步驟 6 之 T 集合為 $\{g\}$ ，故主機 c 會傳送 RREQ 封包給主機 g 。而接著主機 g 於步驟 6 之 T 集合為 $\{u, w\}$ ，故主機 g 會傳送 RREQ 封包給主機 u 及主機 w 。同理，主機 s 於步驟 6 之 T 為空集合，故執行步驟 7 及步驟 8 以廣播方式傳送 RREQ 封包給鄰近的主機 c 。對主機 c 而言，該 RREQ 封包可由 $visited(B_{ID}, x)$ 函數驗證出前已處理過及予以丟棄，故於圖 6 (a) 中並不標示重覆拜訪的 RREQ 封包傳遞過程，以免造成混淆。當主機 t 收到 RREQ 封包時，於步驟 2 發現已找到目的主機，故將由主機 t 依據廣播 ID 與來源主機 ID 所構成 RREQ 的唯一辨識資訊，傳送 RREP 封包給主機 w 以反向蒐集邊徑沿途之各主機 ID 資訊於邊徑路徑內，直到來源主機為止。其 RREP 封包的回覆過程將如圖 6 (b) 所示。圖 6 (a) 中因邊徑找尋 (回覆) 演算法的步驟 7，由於發現雜湊快取內有歷史邊徑資訊，故造成主機 d, i, h, n, m, o 及 y 等不會收到 RREQ 封包。



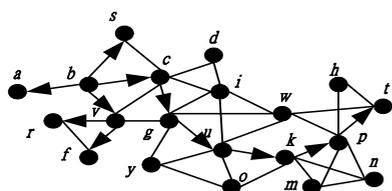
(b) DHRP 的 RREP 封包的回覆過程

圖 6、DHRP 的邊徑找尋範例

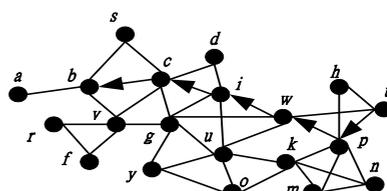
DHRP 在處理 RREQ 封包的傳遞方式上，不同於其它 DSR 基礎邊徑協定。RREQ 封包必須一直傳遞到目的主機，才由目的主機以 RREP 封包做回覆動作。亦即，當 RREQ 封包傳入已記錄可到達目的地之歷史邊徑的主機時，DHRP 會依雜湊快取的過去邊徑經驗繼續傳向目的主機，但是此時的傳遞方式不再是廣播給相鄰的主機。此點除可以省去原有 DSR 協定再次確認路徑的代價，同時也因為 DHRP 中碰撞現象無法確認轉傳的次個主機真的可到目的主機。不過在分散式雜湊機制下可以迅速控制 RREQ 被誤傳的發生機率。一但發現 RREQ 被誤傳，則步驟 3 會從下個主機起又回覆廣播傳遞 RREQ 運作。當然對每個主機而言，已傳遞過的 RREQ 封包將不再重複處理，以免造成網路的負擔。



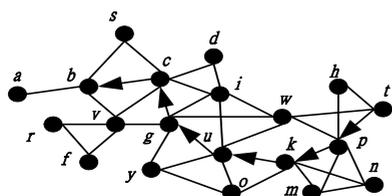
(a) DHRP 的 RREQ 封包邊徑過程



(a) DSR 之 RREQ 封包遠徑過程



(b) G_DSR 之 RREP 封包的 reply 過程



(b) DSR 之 RREP 封包的回覆過程

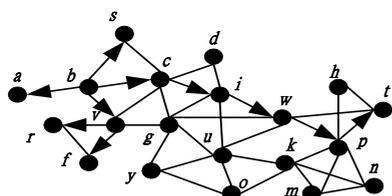
圖 7：DSR 的遠徑找尋範例

圖 8、G_DSR 的 route discovery 範例

各種協定在範例中的結果，可發現一些有趣的差異，而這些差異正可說明其間不同的策略觀點。由主機 b 遠徑到主機 t ，在 G_DSR 協定內可利用路徑 an 及 st 遠徑資訊計算得主機 c 到主機 t 的最佳路徑為 $c-i-w-p-t$ 。而利用 DSR 協定則會得到建議路徑為 $c-g-u-k-p-t$ 。若利用 DHRP 協定則會得到路徑為 $c-g-w-t$ 。不過，此例尚不足以說明 DHRP 協定一定會得到比 G_DSR 協定更短的路徑。畢竟比 G_DSR 協定找到更短路徑的機率是較小的，因為 G_DSR 協定在主機 c 找到可通往主機 t 的歷史路徑 $c-g-u-k-p-t$ ，且同時有另一歷史路徑 $c-i-w-p-n$ 與其相交於主機 p 上，且 $c-i-w-p$ 距離較 $c-g-u-k-p$ 短，則 G_DSR 協定就會有路徑學習的較佳結果出現。反觀 DHRP 協定則須在歷史路徑 $c-g-u-k-p-t$ ，發現路上主機 g 亦有其它歷史路徑 $y-g-w-t$ 可通往主機 t 。如此適當的配合下，才有產生路徑學習的效果且比 G_DSR 協定找到的距離更短的機會，稍後會進一步介紹此路徑學習的特性。

三、DSR 基礎遠徑協定比較及說明

為了便於比較及說明，本文將簡稱[2]中所採用以 Dijkstra's 最短路徑演算法為基礎之連結快取方式 DSR 協定為 G_DSR。如同樣以圖 3 (a) 中的來源主機為 b ，欲找尋遠往主機 t 的路徑，則 DSR 協定之 RREQ 封包遠徑過程將如圖 7(a)所示，而其 RREP 封包的回覆過程將如圖 7 (b) 所示。同例若以 G_DSR 執行之，則其 RREQ 封包遠徑過程將如圖 8 (a) 所示，而其 RREP 封包的回覆過程將如圖 8 (b) 所示。



(a) G_DSR 之 RREQ 封包遠徑過程

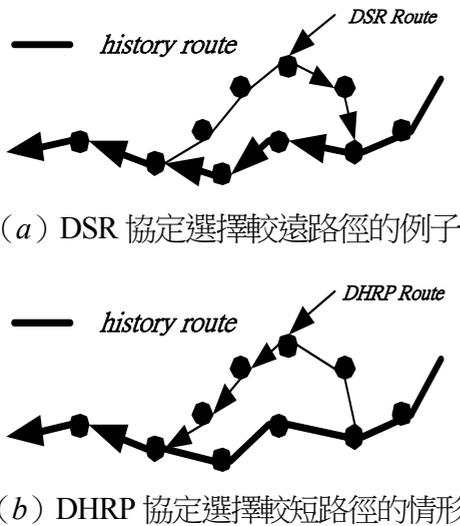


圖 9：DHRP 協定選擇較短路徑範例

在 DHRP 協定中，RREQ 封包一定要到達目的主機後才做回覆 RREP 封包動作來確認路徑。而 DHRP 遶徑找尋（回覆）演算法中的步驟 4 顯示出另一該協定的特徵，其意義是：當 RREQ 封包延著歷史遶徑資訊前進時，若該封包的廣播 ID 與來源主機 ID 已出現過，則表示已有更短的遶徑方式可取代此封包之遶徑結果，故將封包丟棄。此種協定的特徵，使得 DHRP 協定有機會選擇脫離因延著歷史遶徑而找到的較遠的路徑結果。如圖 9 (a) 所示 DSR 協定選擇較遠路徑的例子，而圖 9 (b) 則顯示 DHRP 協定選擇較短路徑的情形。

四、遶徑分享及學習

在[2]中採用連結快取方式，即各主機皆

以圖形的資料結構來儲存歷史遶徑資訊。此種快取管理策略至少表現兩種改善 DSR 協定的優點：

- (1) 分享：減少連結資訊的重覆儲存量。
- (2) 學習：整合連結資訊，可得原始個別路徑資訊無法得知之新遶徑方式。

DHRP 在遶徑分享及學習方面的設計，可由下列說明其特徵：

(1) 分享：在雜湊快取內相同的快取列，可以儲存不同遶徑路徑的資訊，藉以降低重覆儲存量。如圖 3 及圖 4 中的路徑 st 及 co ，分別於快取列 caV_g 內共用 15_{gu} 、及快取列內 ucV_g 內共用 6_{gc} 等儲存空間。

(2) 學習：[2]中以圖形資料結構來連結多條遶徑路徑的連結資訊，以圖形的表示方式將可同時考量兩個以上的歷史路徑，而經計算產生新遶徑，具有路徑學習的功能。如圖 8 (a) 所示，G_DSR 協定的 RREQ 封包當傳入主機 c 後，便可依歷史遶徑資訊及 Dijkstra's 演算法計算得傳輸到目的主機 t 的最佳路徑 $ct : c-i-w-p-t$ 。而此路徑 ct 乃是綜結原歷史遶徑資訊的路徑 an 的 $c-i-w-p$ 片段資訊及路徑 st 的 $p-t$ 片段資訊而成的新路徑。而如圖 6 (a) 所示，DHRP 亦學習到另一條新路徑 $bt : b-c-g-w-t$ 。此路徑 bt 乃是綜結原歷史遶徑資訊的路徑 an 的 $b-c$ 片段資

訊、路徑 st 的 $c-g$ 片段資訊及路徑 yt 的 $g-w-t$ 片段資訊而成的新遠徑。不過，在學習的特徵上相較於 DSR 及 G_DSR，DHRP 的快取大小明顯節省很多。在學習的特徵上，DSR 並無新遠徑的學習能力，而 G_DSR 則比 DHRP 花費較長的計算複雜的 Dijkstra's 演算法，以學習到較短的新遠徑做為傳輸路徑。

肆、相關研究分析

在這節中，我們以定性方式分析 DSR 相關文獻與本論文所提出的 hashing 機制。在遠徑快取所採取的方法上，大致可歸納成路徑快取、連結快取及雜湊快取等三類，其資料結構分別代表二維串列 (List)、圖形 (Graph) 及雜湊 (hashing) 的基本特性的不同。

如表 1 所示，在快取容量特徵列中，顯示出連結快取及雜湊快取可以實現於預先規範的儲存環境，其中的關鍵因素就是快取分享。分享是將相同的遠徑資訊利用相同儲存空間表示與管理。DSR 相關的協定必須有效解決面對遠徑資訊的增加與儲存空間的需求問題，而快取分享正是可以實現於控制儲存空間的主要因素。快取密度 (Cache Density) 的特徵列中，表示雜湊快取高於連結快取，主要有兩項因素所造成：(1) 雜湊

是以位元為儲存單位，而圖形是以位元組為儲存單位。(2) 雜湊可以允許資訊的重疊 (Overlap) 表示，此是因雜湊碰撞所造成之現象。

若觀察比較遠徑找尋的時間複雜度時，不論 DHRP 採取硬體或軟體方式執行快取列的比對均有較好的表現，此預期中的結果乃源於雜湊機制的明顯好處。此外，路徑快取乃採取線性方式比對可能的遠徑結果。而連結快取[2]則於每個 RREQ 封包到達的主機，皆以 Dijkstra's 最短路徑演算法計算出較短的遠徑結果。後兩者在時間複雜度雖皆為 $O(n^2)$ ，但連結快取[2]所找尋到的路徑長度卻是最短的，不過它僅表示歷史遠徑資訊中可到達目的地的最短路徑。

表 1：DSR 方法及其快取結構分類特徵比較表

	Path Cache	Link Cache	Hashing Cache
Data Structure	2-D List	Graph	Hashing Table
Cache Capacity	Unlimited	$O(n^2)$ or Controlled	$O(pk^2)$ or Controlled
Cache Density	Low	Medium	High
Time Complexity	$O(n^2)$	$O(n^2)$	$O(1)$
Cache Learning	No	Yes	Yes
Path Length	Long	Short	Medium

註： p 表示快取列大小， k 表示與主機相鄰且有邊徑經過的平均鄰居數目

從訊號的競爭、主機的處理與雜訊的干擾延遲來觀察，在不考慮頻寬的限制因素下，則汨流方式所找到的邊徑路徑會有較佳的整體通信效能。連結快取[2]所執行的最短路徑演算法，實際上是一種快取學習的具體表現，也因此能較路徑快取[1]具有整體邊徑學習的優點。雜湊快取亦具有快取學習的能力，然而，此種學習的能力卻內建於 DHRP 協定的邊徑方式，特徵及範例已於第三節描述。

因此，DHRP 協定同時考慮到歷史邊徑資訊的學習、雜湊的誤邊機率及訊號的競爭與延遲等特徵，並明確於 DSR 與汨流方式間取得效能在邊徑結果上的平衡。

伍、DHRP 的實驗及效益

我們以模擬實驗的方式進行定量分析與觀察。模擬實驗是以 1000 部主機隨機移動於 1.5 公里 x 1.5 公里的正方形空間中，並由隨機分佈的方式設定成隨意網路。假設每個主機的無線通訊範圍為 100 公尺，且於實驗區域內無障礙物阻隔通訊。主機的行動性是以隨機走動方式進行，移動的個數分別假設為 10%、20%、30%、40%、50%，移動的最大速率分別假設為 10、30、或 50 公尺/

秒。模擬對象是以 DSR 為基礎的相關協定 [1][2]、DHRP 及汨流方式等，藉由相對性的數據比較，以評估 DHRP 協定的效益。

實驗中 DHRP 協定所採用的雜湊函數如式 (1)，其中質數陣列由 $A[0]$ 至 $A[49]$ 組成， $A[0]=101$ 且 $A[i+1]$ 為 $A[i]$ 的相鄰質數。主機 ID 皆設成小於 $A[0]^2$ ，即 101^2 ，故假設 $1234 \leq \text{主機 ID} \leq 2233$ 。歷史邊徑經驗是相關實驗協定的主要控制因素。歷史邊徑經驗又分為有效資訊量、無效資訊量及擴展資訊量等。所謂無效資訊是指重覆出現的歷史邊徑經驗資訊，而有效資訊則是指對於邊徑找尋有直接幫助的歷史邊徑經驗資訊。擴展資訊是指整合不同歷史邊徑經驗，所學習擴展得知的經驗資訊。歷史邊徑經驗又可分別由歷史邊徑的數目及距離做為觀察控制。

模擬實驗所要觀察的結果有快取表現率、快取資訊量、複雜程度、效能及快取大小。令表現連結數、不同連結數、比較次數分別代表邊徑資訊的連結數量、邊徑資訊的不同連結數量、及邊徑時所做的邊徑資訊比較次數其。其中各觀察結果的定義如下：

(1) R_Ratio : (表現率)

= 表現連結數/不同連結數，用以表現連結資訊被重覆儲存的比率。

(2) IOC : (快取資訊量)

= 不同連結數/快取大小 (位元組)，用

以表現快取分享的程度。

(3) Complexity (複雜程度):

=比較次數/路徑，用以評估各種遠徑找尋協定上所需的計算能力。

(4) Performance (效能):

= (遠徑長度 - 汨流遠徑長度) / 遠徑數量，用以評估協定的平均遠徑結果，結果是以相對於汨流方式找尋遠徑長度的平均差值來表現。

(5) 快取大小: (位元組數/主機)，用以評估遠徑協定對於遠徑資訊在儲存空間上的需求程度。

我們以隨機選取主機 ID 方式，決定遠徑來源主機及目的主機的座標，並觀察各實驗結果的變化。圖 11 至圖 15 是以隨機建立 100 條遠徑的平均數據結果，其中圖 14 是以 30% 的移動個數為實驗基礎。

圖 11 顯示出 DHRP 於相同 *cur* 欄的快取列間存有相同的連結遠徑資訊，但仍小於 DSR 約 12%，而 G_DSR 則因為以圖形方式儲存遠徑資訊，保持比率等於 1。DHRP 為了降低誤遠機率及支援雜湊的可逆計算能力所負出的花費，而 G_DSR 亦需於每個主機上實現較複雜的 graph 資料結構及演算法，計算能力的負荷會相對沉重。圖 12 中看出無論以軟體或硬體方式實現快取列間的比對計算，DHRP 的比較次數皆明顯較

G_DSR 少。

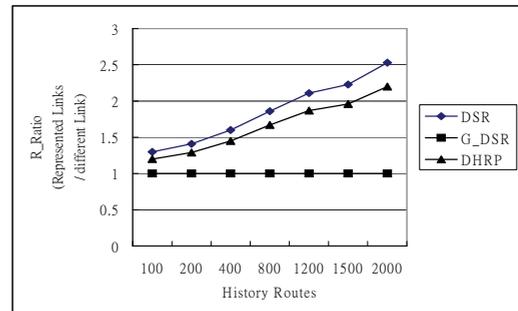


圖 11、歷史遠徑與 R_Ratio 的實驗結果

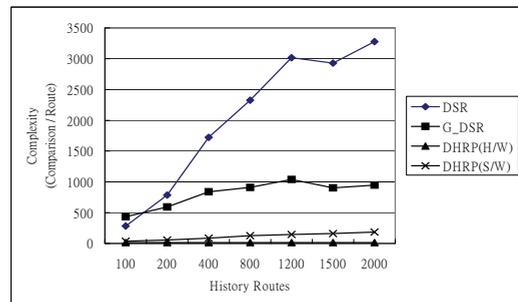


圖 12、歷史遠徑與時間複雜度的實驗結果

圖 13 明顯表示了 DHRP 可平均於每 1 位元組內儲存最多的連結遠徑資訊。雖然 G_DSR 及 DHRP 皆有分享能力，但 DHRP 的儲存單位為位元，而 G_DSR 則為位元組。除此之外，DHRP 的雜湊碰撞亦是另一種分享的潛在因素。相對於圖 14 可觀察出各協定對於儲存空間的成長及需求程度，其中 G_DSR 及 DHRP 對儲存空間的需求成長是較平緩的，此表示可運用限制大小的儲存空間來設計遠徑快取，[2]亦已探討 G_DSR 的此類研究及其可能性。

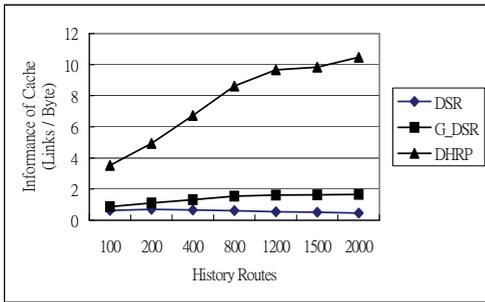


圖 13、歷史遶徑與 IOC 的實驗結果

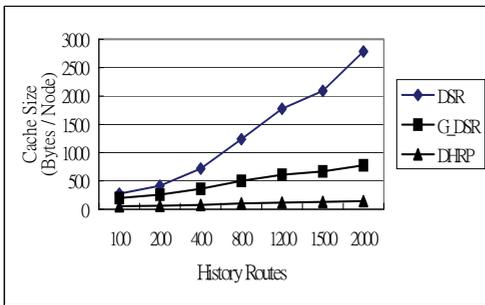


圖 14、歷史遶徑與快取大小的實驗結果

實驗中將以遶徑的長度做為遶徑效率的觀察重點，圖 15 表示各協定相對於氾流方式的遶徑長度差值比較。氾流方式的遶徑通常可以找到網路中最短的遶徑，然而卻會造成大量的路徑找尋控制封包及網路負擔。圖 15 中，G_DSR 以複雜的 Dijkstra's 最短路徑演算法找出接近氾流遶徑的結果，但會造成主機的較重計算負荷及較多的快取需求，結果如圖 12 及 14 所示。雖然，DHRP 在遶徑長度方面稍高於 G_DSR，介於 DSR 及 G_DSR 之間，但當歷史遶徑多於 1200 條時，則可明顯拉近與 G_DSR 遶徑效

能之差距，而 DSR 遶徑效能則遠比 G_DSR 及 DHRP 較差。此結果對於採用小量快取及較少的計算負荷設備而言，確顯示出 DHRP 具有不錯的實用性。

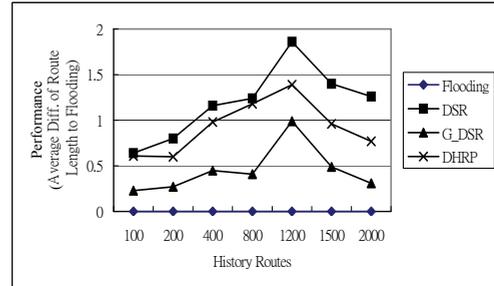


圖 15、歷史遶徑與效能的實驗結果

陸、結論

目前個人無線通信設備內建稍弱之數據運算能力及少量記憶空間，當此些設備因需求而構組成隨意網路時，在遶徑協定上採用簡單的 DSR 遶徑協定是直接可行的實現方法。本論文提出了 DHRP 遶徑協定，以改善[1][2]於儲存空間及計算能力之可能應用限制，我們亦實驗結果具體說明 DHRP 相較於其它相關研究[1][2]在快取表現率、快取資訊量、複雜程度、實用性及快取大小方面，皆有較好的運作效果。另外，從定性及定量的分析中，DHRP 也皆有不錯的表現。但在移動性的維護上，顯示 DHRP 較適合慢速的移動性網路應用環境。

參考文獻

- [1] D. Johnson and D. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks", *Computer Communications Review Proceedings of SIGCOMM 96*, Aug. 1996.
- [2] Yi-Chun Hu and David B. Johnson, "Caching Strategies in On-Demand Routing Protocols For Wireless Ad Hoc Networks", *Proceedings of the 5th International Conference on Mobile Computing and Networking (MobiCom)*, pp.231-242, 2000.
- [3] Koji Nakano, Stephan Olariu, "Randomized Initialization Protocols for Ad Hoc Networks", *IEEE Transactions on Parallel and Distributed Systems*, Vol.11, No.7, pp. 749-759, 2000
- [4] Vincent D. Park and M. Scott Corson, "A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks", *Proceedings of INFCOM*, pp.1405-1413, April 1997.
- [5] Marcelo Spohn and J. J. Garcia-Luna-Aceves, "Neighborhood Aware Source Routing", *Proceedings of the 6th International Conference on Mobile Computing and Networking (MobiCom)*, pp.11-21, 2001.
- [6] David Karger, Eric Lehman, Tom Leighton, Rina Panigrahy, Matthew Levine, and Daniel Lewin, "Consistent hashing and random trees : distributed caching protocols for relieving hot spots on the World WIDE Web", *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pp.654-663, 1997.
- [7] Blaise Gassend, G Edward Suh, Dwaine Clarke, Marten van Dijk, and Srinivas Devadas, "Caches and Hash Trees for Efficient Memory Integrity Verification", *Proceedings of the IEEE Ninth International Symposium on High-Performance Computer Architecture*, pp.295-306, 2003.
- [8] K Murugan, S Balaji, P Sivasankar, S, and S. Sbanmugavel, "Cache based energy efficient strategies in mobile ad hoc networks", *Proceedings of the IEEE International Conference on Personal Wireless Communications (ICPWC)*, pp.90-94, 2005.
- [9] Aissani, M.; Fenouche, M.; Sadour, H.; Mellouk, A.; "Ant-DSR : Cache

Maintenance Based Routing Protocol for Mobile Ad-Hoc Networks”, *Proceedings of the 3rd Advanced International Conference on Telecommunications (AICT)* , pp.35-40, 2007.

Effects of annealing temperature of TiO₂ thin films for application in dye-sensitized solar cells

Ming-Cheng Kao

Abstract

TiO₂ thin films have been deposited on FTO-coated glass substrates using sol-gel technology for application as the work electrode for the dye-sensitized solar cells (DSSC). The effects of annealing temperature (400~700 °C) on the microstructure, morphology and optical properties of TiO₂ thin films were studied. The electrode of DSSC fabricated with TiO₂ thin films were characterized by X-ray diffraction (XRD), scanning electron microscopic (SEM) and Brunauer-Emmett-Teller (BET) analysis. Based on the results, the TiO₂ films annealed at 700 °C shows better crystallization, higher porosity and surface area than those of TiO₂ films annealed at lower temperature. The higher efficiency (η) of 1.57 % with J_{sc} and V_{oc} of 5.6 mA/cm² and 0.65 V, respectively, was obtained by the TiO₂ film annealed at 700 °C.

Keywords : TiO₂, sol-gel, dye-sensitized solar cells, short-circuit current, open-circuit voltage.

燒結溫度對於二氧化鈦染料敏化 太陽能電池之影響

高銘政

摘要

本論文以溶膠-凝膠法製作二氧化鈦薄膜，並作為染料敏化太陽能電池之工作電極，研究不同燒結溫度（400~700°C）對於二氧化鈦薄膜之微結構、表面型態及光學特性等影響，藉由 XRD 繞射儀、電子式掃描顯微鏡（SEM）及比表面積分析儀（BET）等儀器量測染料敏化太陽能電池之二氧化鈦工作電極相關物理特性。實驗結果顯示，隨著燒結溫度的增加，二氧化鈦薄膜的孔隙度及比表面積也會增加，並且改善二氧化鈦表面之染料 N3 吸附量及光電轉換效率，同時在太陽能電池之特性方面，當燒結溫度 700°C 時，可以得到較高之短路電流、開路電壓及轉換效率，分別為 5.6 mA/cm²、0.65 V 及 1.57 %。

關鍵詞：二氧化鈦、溶膠凝膠法、染料敏化太陽能電池、短路電流、開路電壓。

1. Introduction

Recently, nanocrystalline titanium oxide (TiO₂) is well known as a metal oxide semiconductor, has been extensively studied in many applications. For example, it is used for photo-electrodes, photocatalysts and dye-sensitized solar cells (DSSC) applications with high performance due to its fine physical, chemical and optical properties [1-3]. The DSSC consist of sensitizing dye, nanoporous metal oxide film, electrolyte and opposite electrode. The metal oxide film plays a key role in the enhancement of photoelectric conversion efficiency of DSSC, many studies focus on the relation between film structure and photocatalytic activity as well as the power conversion efficiency of DSSC [4-10]. A high incident photon to current conversion efficiency can be expected for TiO₂ films with better crystalline and higher specific surface area. Therefore, the control of the film structure is very important for applications of TiO₂ films.

In this paper, we report the preparation of high-quality TiO₂ nanocrystalline thin films with high quality on ITO-coated glass substrates, using the sol-gel method. The

properties of TiO₂ thin films are much influenced by not only the growth processes but also the heat treatment parameters. Therefore, it is necessary to investigate the effect of annealing process on the microstructure, transparent properties, and morphology of TiO₂ films. The synthesis conditions and characteristics of the TiO₂ thin films are examined through the thermal analysis (DTA/TGA), X-ray diffraction (XRD), scanning electron microscopy (SEM) and Brunauer-emmett-Teller (BET) analysis to clarify and report on the effect of the deposition parameters. The photoelectric performance of DSSC were studied by I-V curve, the influence of annealing temperature of TiO₂ thin films on the performance of DSSC was also discussed.

2. Experiment

Fig. 1 shows the sol-gel process for preparing the precursor solutions of TiO₂ with a general chemical formula. Titanium diisoprop-oxide bis (2,4-pentanedionate) (TIAA), Ti(OC₃H₇)₂(CH₃COCHCOCH₃)₂ (Alfa, 99.9%+ purity), was used as precursors and 2-methoxyethanol, C₃H₈O₂

(Fluka, 99.9%+ purity) was used as solvent. The gravimetrically assayed $\text{Ti}(\text{OC}_3\text{H}_7)_2$ ($\text{CH}_3\text{COCHCOCH}_3$)₂ reagents were dissolved in a mixture of 2-methoxyethanol solution at room temperature. The mixture was refluxed at 120 °C for 1 h under ambient atmosphere and then cooled to 80 °C for 2 h to promote its homogeneity, and then a stock solution of ~1M concentration with a golden color was obtained.

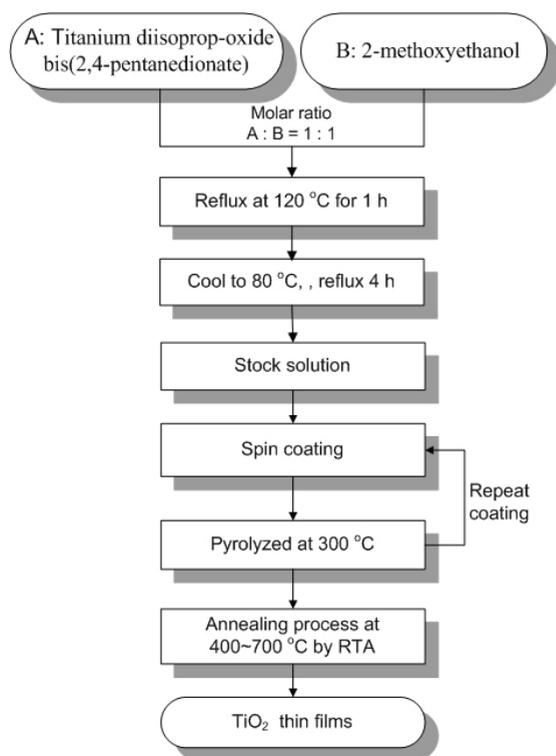


Fig. 1. Flow diagram of the preparation of TiO_2 films.

The stock solutions were spin-coated on

ITO-coated glass substrates ($2\text{ cm} \times 2.5\text{ cm}$) at a spin rate of 3000 rpm for 30 s using a commercial photoresist spinner. The precursor solutions were deposited onto the substrates via a 0.2 μm syringe filter, thus avoiding particulate contamination. After each coating step, the gel films were pyrolyzed on the hot plate at 300 °C for 2 min before final annealing. The average thickness of a single-coated as-fired layer, measured by an α -step surface profiler, was found to be about 0.1 μm . The desired film thickness of approximately 2 μm was achieved by repeating the spin-coating and heating cycles. After multi-coating, TiO_2 thin films were annealed at 400 °C, 500 °C, 600 °C and 700 °C for 2 min by the RTA in the oxygen atmosphere. The film-deposition process is also shown in Fig. 1.

Thermal analysis of precursor gels using differential thermal analysis (DTA) and thermo gravimetric analysis (TGA) was carried out to identify the major evaporation, decomposition and phase crystallization steps during pyrolysis. The gels were pre-dried at 100 °C for 10 h, and then analysed under a heating rate of 5 °C /min. The thickness of TiO_2 film obtained was measured by an α -step surface profiler. The

crystallization and microstructures of thin films were analysed by XRD, SEM and BET analysis, respectively. In order to sensitize TiO₂ film, the TiO₂ electrodes were immersed in a 3×10^{-4} M solution of RuL₂(NCS)₂ (Solaronix, N3 dye) in ethanol for 24 h at room temperature. The electrolyte was composed of 0.5M lithium iodide (LiI) /0.05M iodine (I₂) in acetonitrile. The short-circuit photocurrent (J_{sc}) and open-circuit voltage (V_{oc}) were measured using a Keithley Model 2400 source measure unit under an illumination of 100 mW/cm² (AM 1.5 simulated light radiation) from a 1000 W xenon lamp.

3. Results and Discussion

The thermolysis behavior of TiO₂ gel predried at 100 °C is shown in Fig. 2. In TGA analysis, the gel exhibited approximately 50% weight loss at temperatures ranging from 100 °C to 300 °C. This result is due to the elimination of adsorbed water, solvent and the decomposition of organic by-products. The DTA data indicate that two of exothermic peaks are associated with the former weight loss at temperatures between 100 °C and 300 °C. The final peak which appeared at about 500 °C is

associated with a weight loss that is probably due to crystallization of TiO₂. This indicates that the TiO₂ gel decomposes and crystallizes at temperature as low as 500 °C.

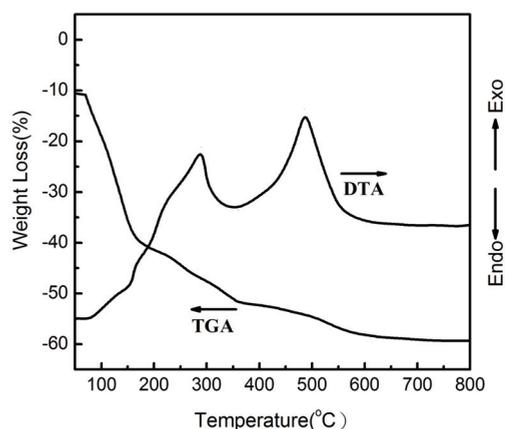


Fig. 2. Thermo gravimetric analysis(TGA)and the corresponding DTA for gel of TiO₂.

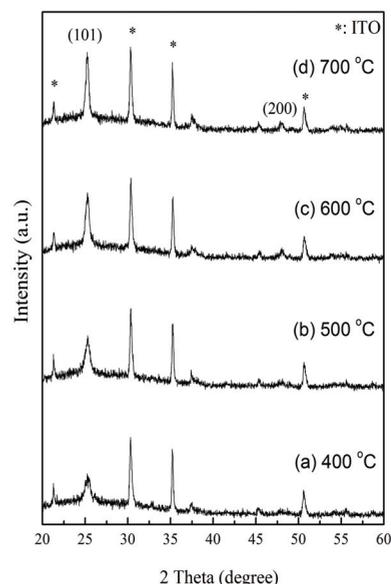


Fig. 3. XRD results as a function of annealing temperature for TiO₂ films.

Figure 3 shows the XRD patterns of TiO_2 thin films annealed grown on ITO-coated glass substrates at different heating temperatures, 400, 500, 600 and 700 °C for 2 min, respectively. TiO_2 films show a anatase structure with preferred (101) orientation. As the annealing temperature was increased, the peaks of the (101) -plane for samples became sharper and the full-width at half-maximum (FWHM) decreased indicating better crystallinity and an increase in grain size with increasing annealing temperature. The (101) peaks obtained for TiO_2 films indicated a strong (101) -axis orientation to the substrate surface. These results demonstrate that the structure of TiO_2 films is improved when annealed at higher temperatures.

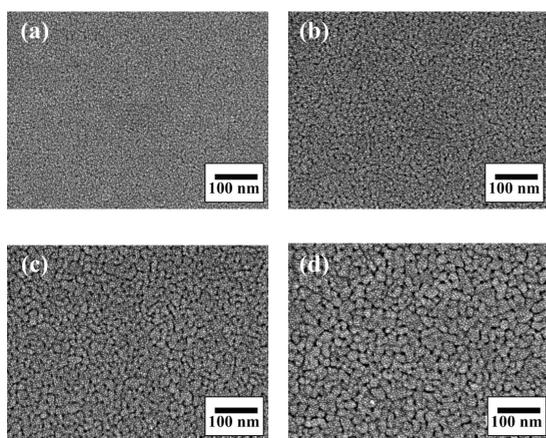


Fig. 4. Surface microstructures of TiO_2 films annealed at various temperatures; (a) 400 °C, (b) 500 °C, (c) 600 °C and (d) 700 °C. (bar = 100 nm)

The evolution of surface microstructure of the TiO_2 thin films at annealing temperatures from 400 to 700 °C for 2 min are shown in Fig. 4. According to the SEM results, the mean sizes of the TiO_2 crystallites were estimated to be 8, 12, 15, and 21 nm for Figs. 4 (a) – (d), respectively. These results demonstrate that the grain size of TiO_2 thin films increases as the increased heating temperature. In addition, the surface micrograph of TiO_2 thin film exhibits nanocrystalline and nanoporous structure which is composed of interconnected nanoparticles.

Table 1. The specific surface area and porosity of TiO_2 films obtained from BET measurements.

Annealing temperature (°C)	BET Surface area (m^2/g)	Porosity (%)
400	25	12
500	34	18
600	44	24
700	56	35

The Brunauer-emmett- Teller (BET) analysis of thin film is an important factor because it will affect the absorption of dye molecules onto the TiO_2 film. The surface area and porosity of the TiO_2 thin films annealed at various annealing temperature are shown in

Table 1. It is obvious that both surface area and porosity of films increase with the increase of annealing temperature. This phenomenon appears to be the results of the increased surface mobility of the atoms and molecules absorbed on the substrate at higher substrate temperature. In addition, with an increase of the annealing temperature, a crystalline phase starts to appear and leads to an appearance of the grains. Therefore, the surface area thus increases and the film becomes relatively coarse. The maximum surface area and the porosity of 56 m²/g and 35 %, respectively, was obtained by the TiO₂ film annealed at 700 °C.

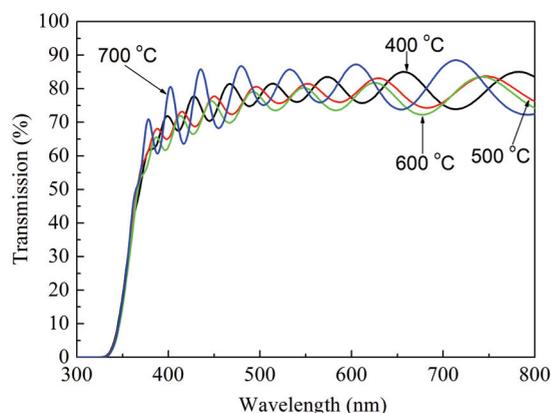


Fig. 5. Optical transmittance spectra of TiO₂ films annealed at 400 °C, 500 °C, 600 °C and 700 °C

Figs. 5 shows the UV–visible optical

transmittance spectra of the TiO₂ films between 300 and 800 nm in wavelength. It can be seen that all films have high transmittance and the absorption edge is at about 300 nm. The average transmittance of the films annealed at 400 °C, 500 °C, 600 °C and 700 °C in the visible range was about 85%, 80%, 75% and 60%, respectively. From the results, the TiO₂ film annealed at 700 °C has highest transmittance spectra, and it can enhance the incident light intensity to obtain larger photocurrent of DSSC.

Table 2. The performances of DSSC fabricated with TiO₂ thin films annealed at various temperatures.

Annealing temperature (°C)	J_{sc} (mA/cm ²)	V_{oc} (V)	Fill factor (%)	Efficiency (η %)
400	2.6	0.59	32	0.51
500	4.5	0.61	33	0.92
600	5.1	0.62	41	1.23
700	5.6	0.65	43	1.57

Table 2 shows the various parameters of DSSC with TiO₂ thin films annealed at various temperatures. From the results, the values of J_{sc} and V_{oc} of DSSC increased with the increase of annealing temperature. It indicates that the higher surface area and porosity of the TiO₂ thin films can absorb easily the dye. Therefore,

the light incident on the ITO/TiO₂/RuL₂(NCS)₂ electrodes generates electrons by exciting dye molecules absorbed onto the TiO₂ film. A high incident photon to current conversion efficiency can be expected under illumination through TiO₂ film in the application of DSSC. In addition, the *I-V* curve for DSSC with TiO₂ film annealed at 700 °C is shown in Fig. 6. The higher efficiency (η) of 1.57% with J_{sc} and V_{oc} of 5.6 mA/cm² and 0.65 V, respectively, was obtained by the TiO₂ film annealed at 700 °C.

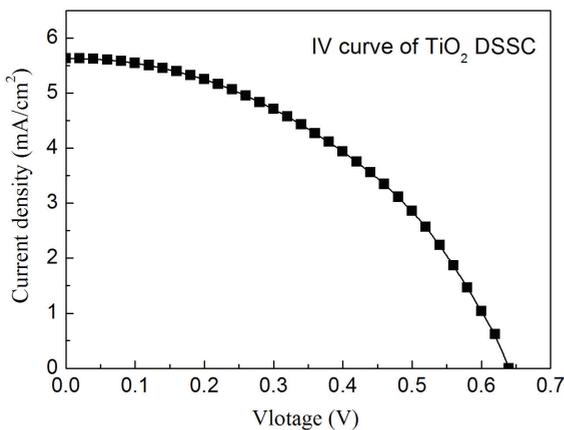


Fig. 6. I-V curve for DSSC with TiO₂ thin film annealed at 700 °C.

4. Conclusions

The influence of annealing temperature (400~700 °C) on the performance of DSSC with TiO₂ films was studied. With increasing

the annealing temperature, the intensities of (101) peaks of TiO₂ thin films gradually increased. From the results of BET analysis, the increase in surface area and porosity of TiO₂ films crystallized at the higher temperature of 700 °C contributed to the improvement on the short-circuit photocurrent (J_{sc}), open-circuit voltage (V_{oc}) and efficiency (η) of DSSC.

5. References

- [1] M. Gratzel, "Applied physics : solar cells to dye for," Nature, 421 (2003), pp. 586-587.
- [2] B. O'Regan, M. Graetzel, "A low-cost, high efficiency solar cell based on dye-sensitized colloidal TiO₂ films," Nature, 353 (1991), pp. 737-740.
- [3] N. G. Park, J. V. Lagemaat, A. J. Frank, "Comparison of dye sensitized rutile and anatase based TiO₂ solar cells," J. Phys. Chem. B 104 (2000), pp. 8989-8994.
- [4] P. Basudev, S. K. Batabyal, A. J. Pal, "Vertically aligned ZnO nanowire arrays in Rose Bengal-based dye-sensitized solar cells," Solar Energy Mater. Solar Cells, 91 (2007), pp. 769-773.

- [5] J. Lagemaat, K. D. Benkstein, A. J. Frank, "Relation between particle coordination number and porosity in nanoparticle films : implications to dye-sensitized solar cells," *J. Phys. Chem. B*, 105(2001), pp. 12433-12436.
- [6] M. Durr, A. Bamedi, A. Yasuda, G. Nelles, "Tandem dye-sensitized solar cell for improved power conversion efficiencies," *Appl. Phys. Lett.*, 84 (2004) , pp. 3397-3399.
- [7] B. A. Gregg, S. G. Chen, S. Ferrere, "Enhanced dye-sensitized photoconversion efficiency via reversible production of UV-induced surface states in nanoporous TiO₂," *J. Phys. Chem. B*, 107 (2003) , pp. 2019-3029.
- [8] M. Okuya, K. Nakade, S. Kaneko, "Porous TiO₂ thin films synthesized by a spray pyrolysis deposition (SPD) technique and their application to dye-sensitized solar cells," *Solar Energy Mater. Solar Cells*, 70 (2002) , pp. 425-435.
- [9] Y. Saito, S. Kambe, T. Kitamura, Y. Wada, S. Yanagida, "Morphology control of mesoporous TiO₂ nanocrystalline films for performance of dye-sensitized solar cells," *Solar Energy Mater. Solar Cells*, 83 (2004) , pp. 1-13.
- [10] P. Wang, L. Wang, B. Ma, B. Li, Y. Qiu, "TiO₂ Surface Modification and Characterization with Nanosized PbS in Dye-Sensitized Solar Cells," *J. Phys. Chem. B*, 110 (2006) , pp. 14406-14409.
-

智慧型行車防撞安全系統

麥毅廷、陳振庸、陳怡伸、楊大威、高資淵、紀佳瑋

摘要

隨著科技的進步，車輛也可以結合資訊科技成為智慧型車輛，而行車駕駛安全(Driving Safety)是駕駛者最重視且期待的，行進中車輛如果能夠在交通事故發生前即有預警告給駕駛者，使駕駛者可以提早反應；或是直接讓車子自行作緊急處理如自動減速或是煞車之動作，藉此可以有效避免行車事故的發生。為了能有效發出碰撞預警及車輛自動處理機制，本研究希望設計並實做一系統，稱為智慧型行車防撞安全系統(Intelligent Safety Vehicle Anti-collision System, ISVAS)，能有效的在碰撞發生時即發出警示訊息通知附近的車輛，可以不必等到駕駛者來處理，這樣可以有效提昇安全性。ISVAS 系統的目標將利用車載資通訊(Telematics)中的車間通訊方式(Vehicle To Vehicle, V2V)的概念，讓車輛週期性的廣播本身的相關資訊，例如位置資訊(Location Information)、速率(Velocity)等等，使行進中之車輛能交換彼此的相關資訊，而當事故發生時，其他車輛一旦收到事故的訊息就可以加以閃避或是作進一步的處理，將可以有效防範未來。目前本篇論文將以簡單的智慧型機器車來模擬車輛，並在各車上安裝無線 ZigBee 介面，而目前初步的成果是透過簡單的 V2V 溝通，當發生事故時，將緊急訊息通知其它車輛，透過 ZigBee 來相互通訊及溝通以實現我們初步提出的 ISVAS 智慧型行車防撞安全系統。

關鍵詞：車載資通訊、車載網路、行車安全、車間通訊。

麥毅廷：修平技術學院資訊網路技術系助理教授
陳振庸：修平技術學院資訊網路技術系助理教授
陳怡伸：修平技術學院資訊管理系講師
楊大威：修平技術學院資訊網路技術系學生
高資淵：修平技術學院資訊網路技術系學生
紀佳瑋：修平技術學院資訊網路技術系學生
投稿日期：990907 接受刊登日期：991026

Intelligent Safety Vehicle Anti-collision System

Yi-Ting Mai, Jeng-Yueng Chen, I-Shen Chen,
Da-Wei Yang, Zi-Yuan Gao, Jia-Wei Chi

Abstract

Vehicle-to-vehicle (V2V) communications are enabling technologies for vehicles to cooperative driving for enhanced traffic safety and efficiency. V2V communications refer to communications between road vehicles (cars, trucks, motorcycles, etc.) without involving a central control unit. Early warning is crucial for safety-related V2V applications. To achieve safety driving, the Intelligent Safety Vehicle Anti-collision System (ISVAS) is proposed. Our proposed intelligent system will send the warning message immediately after there are car accidents occurring. According to V2V communications, the other cars could have enough time to avoid car accident and make an appropriate decision such as slow down, stop, and detour after receiving the warning messages. To achieve our proposed ISVAS system, the robot vehicles have been designed to simulate vehicles on the road. Besides, vehicles also apply ZigBee wireless interface to communicate with each other. The experiment has demonstrated our proposed intelligent system can provide safety driving for vehicles when cars accident occurred.

Keywords : Telematics, Vehicle, VANET, Driving Safety, V2V.

Yi-Ting Mai, Assistant Professor, Department of Information and Networking Technology, HIT.

Jeng-Yueng Chen, Assistant Professor, Department of Information and Networking Technology, HIT.

I-Shen Chen, Instructor, Department of Information Management, HIT.

Da-Wei Yang, Student, Department of Information and Networking Technology, HIT.

Zi-Yuan Gao, Student, Department of Information and Networking Technology, HIT.

Jia-Wei Chi, Student, Department of Information and Networking Technology, HIT.

壹、緒論

現今社會的科技進步，交通建設也因而更先進、便利，因此交通運輸的便利性也是一年比一年好，台北到高雄開車上高速公路也只需要四個小時。如此良好的交通運輸可以增進人們生活的便利生，因此如何在便利、多元化的交通運輸上提供行車上最重要的基本要求「安全行車」將是許多用路人、車輛駕駛共同的目標。就如同現行高速公路的車輛駕駛，由於路況的改善，可以順暢的行車，但在這種高速且長時間的汽車上，駕駛人的精神往往因為長時間駕駛也都消耗得很快，常常都會有精神不濟的狀況，也因而導致臨時狀況發生時反應變慢的情形，而產生等事故。由此可見，行車之中結合現今智慧型資訊系統[1]，可以提供行車駕駛多一分安全功能，我們希望此系統用於車與車之間的通訊，以即時的速度告知其他車輛自己的狀況，其他車輛要是收到不良的訊息就可以加以閃避，防範未來有可以發生連環車禍的可能性。這樣不只可以減少傷害到別人也可以更加保護自己的行車安全。

在行車用路上，目前車載資通訊（Telematics）系統也是因為車輛之間需要通訊因此而產生一種新的通訊模

式，而其原來的技術就是現有通訊、資訊與車輛工業的整合，因此是一個多元化的技術，而要能夠在這個領域提供相關的服務，就提供者而言，必定是跨領域的，因此由內容提供者、內容整合者、電信業者、軟體業者、硬體業者、車載資通訊系統服務供應商（Telematics Service Providers, TSP）及系統整合業者（汽車製造商）等所組成，分別藉由協力合作的方式來提供相關服務及技術。在駕駛安全的相關研究領域中，可再細分為主動式安全（Active Safety）與被動式安全（Passive Safety）[2]。被動式安全系統是我們目前十分常見的行車配置，這是一種當事故發生時用以減少或降低其對人駕駛和乘客之傷害程度，常見之被動式安全系統例如安全氣囊（Airbag）、安全帶（Safety Belt）等等。另一方面的主動式安全系統為較新的議題，其中包括三大部份，分別為車輛穩定性系統（Vehicle Stability Systems）、駕駛警告與訊息資訊系統（Driver Warning and Information Systems）及防碰撞系統（Collision Avoidance Systems），這是為利用各種裝置、設備、技術等，目的是能夠在事故發生前即產生相關動作，以期保持車輛的操控狀態，避免事故發生的機會或降低其損害程度。

而本論文也是以設計主動式安全系統為基本概念，希望設計一套智慧型行車防撞安全系統 (ISVAS)，希望透過此系統可以用來加強駕駛人員在行車方面的安全性，基本概念及作法就是讓每台車定時的發送資訊給附近周遭的汽車或是各種車輛，當一台車發生意外時，如車禍碰撞時，附近的車子系統會立即收到訊號並做緊急剎車，而遠方的車子系統會立即做出辨識並改換行進車道或是繞道等。由於智慧型判斷系統於車輛上，事故車輛可即時的傳送訊號給其他車輛，使其他車輛可以迅速的加以閃避或者停靠路肩等進一步安全措施。同時也可以在藉由車子上的無線通訊系統將訊號進一步傳送給管制中心：如高公局，將車禍情報迅速傳到中心後，可以更快的派出警力支援車禍現場，並將車禍消息發送到所有的車子上，讓系統可以做出辨識，結合區域定位系統可以重新選擇路徑以避免碰撞和塞車。本文後續在第二節中說明車載資通訊之文獻探討；第三節陳述本論文設計之 ISVAS 系統的架構及演算法流程；第四節說明系統實作之相關功能；最後在第五節提出簡單結論及未來發展。

貳、文獻探討

一、車載資通訊

車載資通訊 (Telematics) 一詞是由

電信學 (Telecommunication) 與資訊學 (Informatics) 所組成。汽車製造聯盟 (Alliance of Automobile Manufacturers) 將其定義如下：結合汽車電子與通訊技術，提供駕駛者行車導引與資訊交換服務，改善個人的行車安全與保全。車載資通訊系統是由車內的資訊平台 (On Board Unit, OBU)、通訊業者、服務內容提供者三端所串連，其目的是提供車內通訊與資訊服務。其服務的提供方式是藉由全球衛星定位系統確定車子所在的位置，利用車載資訊平台收集行駛資訊與接受使用者的需求，透過無線通訊網路與服務內容連結，滿足使用者資訊傳遞與交換的需求，如圖 1。而台灣的國道高速公路行車旅程之應用如圖 2 [3]。運用如而有許多學者也在此新興的跨領域中進行研究與探討，如車際無線通訊架構[4]、智慧型運輸[5]、行車車間移動通訊協定[6]等方面的探討。而在[7]探討在 V2V 的無線通訊上如何在實體層透過 multiple-input-multiple-output (MIMO) 的方式將無線傳輸的頻寬作提升及有效利用。而在[8]中提出較佳通訊方式在 V2V 的車間通訊上，面對緊急的需求時可以仍然可以保證訊息的傳達，但是此論文上仍是透過類似 WiMax [9][10]的技術，在韓國專有的 WiBro [11]網路來傳送，也就是有固定點之基地台

(Infrastructure) 的架構來進行通訊，而非 V2V 僅車輛間之通訊方式。

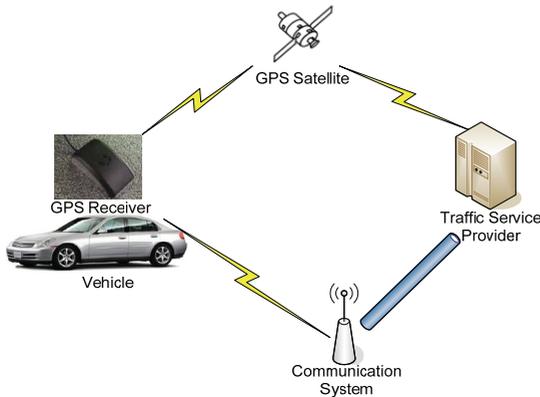


圖 1 車載資通訊系統示意圖

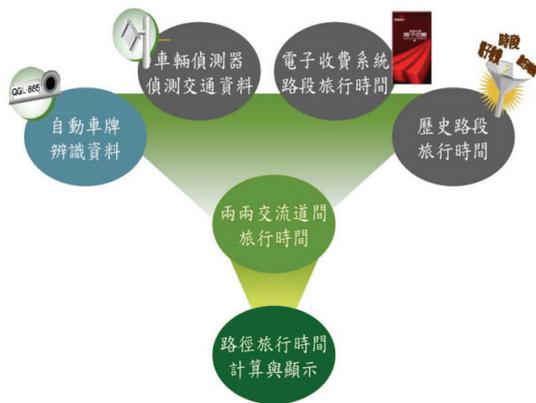


圖 2 國道高速公路行車時間預測系統

二、ZigBee

ZigBee 是由是 IEEE 802.15.4 國際標準組織與 ZigBee Alliance 兩個組織，分別制定硬體與軟體標準[12][13]，ZigBee 是一種短距離、架構簡單、低消耗功率與低傳輸速率之無線通訊技術，其傳輸距離可達 100 公尺，其工作頻率分為 2.4 GHz、915 MHz 和 868 MHz 等三種頻段，

分別提供 250 Kbps、40 Kbps 和 20 Kbps 之傳輸速率，網路架構具備 Master/Slave 屬性，並可達到雙向通信功用[14][15]。根據 ZigBee 之技術，ZigBee 無線網路本身具有下列之特性：省電、高可靠度、高擴充性、多樣性網路拓撲、體積小等等。ZigBee 無線網路中有三種身份，分別是協調者 (Coordinator)、路由器 (Router) 及終端裝置 (End Device) 所組成，而在此無線網路中也支援三種拓撲：Mesh、Star 和 Tree [12]如圖 3 所示。而 ZigBee 的命名，源自於蜜蜂在發現花粉時，展現如同 Z 字形狀的舞蹈[16][17]。由於 ZigBee 一開始也是當成無線網路傳輸架構的一種模式，因此有一些論文也在探討在 ZigBee 無線網路上的路由協定[18]。而後有一些家庭自動化設備，希望將傳統採用線控制之方式，進而研發無線 (Wi-Fi) [19]、藍芽 (Bluetooth) [20][4] 等無線技術，直到後來 ZigBee 技術發展成熟，而有一些論文也將此技術用於居家通訊之中[21]-[24]。而在行車安全系統的議題上，在[25]中提出基本想法，並將 V2V 的行車通訊實作於 ARM9 的設備上；[26][27]文中，也針對 V2V 的行車安全架構作了初步的規畫與設計，並將 V2V 的行車通訊透過 IEEE 802.11p [28] 進行溝通，並透過 GPS 進行車輛間的定位作為車間 V2V 的行車距離的參考依

據。前面論文中提到一些 V2V 的行車安全架構，但是在車輛位置的計算上，必須透過 GPS 來取得資訊，而本篇論文希望提出在車間輛上有簡單的無線偵測設備如超音波 Sensor，即可以透過簡單及低成本的方式了解車輛間的行車距離，此外透過 ZigBee 來進行 V2V 的通訊傳輸來作行車安全系統相較於 IEEE 802.11p 的設備，可以有較低的建置成本，而在此架構上現有的論文中，並沒有完整提出實作的論文。

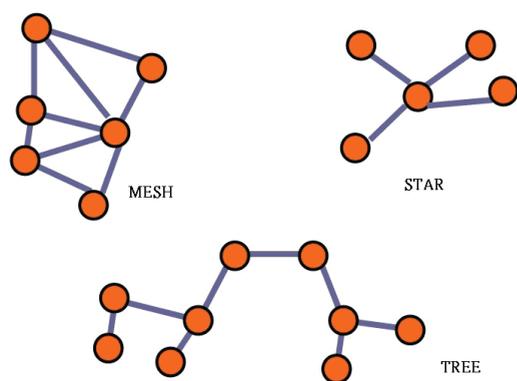


圖 3 ZigBee 無線網路中三種拓樸

參、智慧型行車防撞安全系統

一、智慧型行車防撞安全系統架構

本論文提出一個 ISVAS 的架構，是希望讓車輛彼此可以相互溝通，而一般車載通訊最主要的無線通訊協定為 IEEE

802.11p [28]，但是本論文實際模擬之環境僅為數十公尺，因此將以 ZigBee 無線網路來當通訊介面。圖 4 為一般車輛行進間透過無線訊號廣播（如藍色圈）相關行車資訊。

圖 5 為交通道路上發生事故時，模擬車輛相撞的時候，當發生車輛相撞，這時肇事與受害者兩方的車輛會在第一時間發送訊息給其他車輛，其他車輛接受到訊息會馬上解析資料做判斷。如果前面車輛有發生小擦撞時，將會通知其它車子變換車道來閃避行駛，但前面要是發生重大車禍，左右和後面來車會立即停駛。此外在每台車輛上得左右邊都安裝上超音波感測裝置，如果是前側小擦撞進行閃避時在加以偵測左右來車已保障用路者的安全。圖 6 為模擬偵測左右來車示意圖。當然前側也有裝上超音波感測以防與前車過近而發生意外。圖 7 為車輛過近時車輛減速的示意圖。

由於一般車輛中主要是透過人力駕駛來作行車的操控，我們在此系統中，希望讓車輛具有如人之操控智慧可以讓在行車中發生特別狀況時，可以自動處理，因此我們將各種車輛上增加了無線通訊模組、碰撞偵測及無線超音波感測，並將這些感測器加入的流程演算法說明如下。圖 8 是模擬一般車輛，而增加感應設備為觸碰偵測及 ZigBee 無線通訊模組，

智慧車的動作流程為車輛啟動後就會先偵測觸碰感測器有無狀況信號及此車專在於接收訊號，如果有訊號就停駛無訊號就行駛，再來決定行駛或停駛。圖 9 為超音波車系統流程圖，此車也有裝設無線訊號接收裝置，此外還會透過車前的超音波感測器去偵測與前車的距離，來決定是否停駛。我們假設超音波測得與前車之距離為 D ，而當 D 變小時，我們必須確認是否小於 D_{th1} ，如果小於 D_{th1} 時，我們就必須進行減速，而當 D 仍持續減少，那表示可能會讓兩車距離過近，而 D 的值未達 D_{th2} 時車子將自動停止行駛，判斷方式詳如圖 9。圖 10 為偵測兩側來車系統流程圖，會因為接收到訊息或者是前端觸碰裝置受到碰撞而開啟兩側超音波感測器，偵測兩旁是否有來車來助於車子判斷是否可以改道行駛。

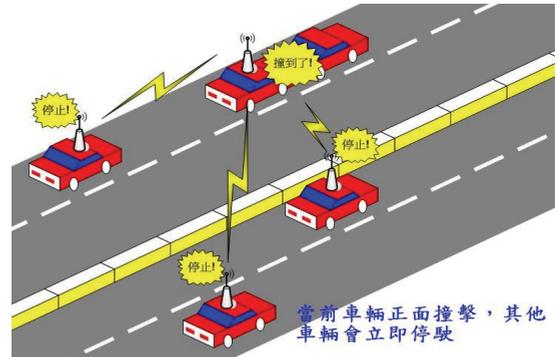


圖 5 車輛行車事故通知示意圖

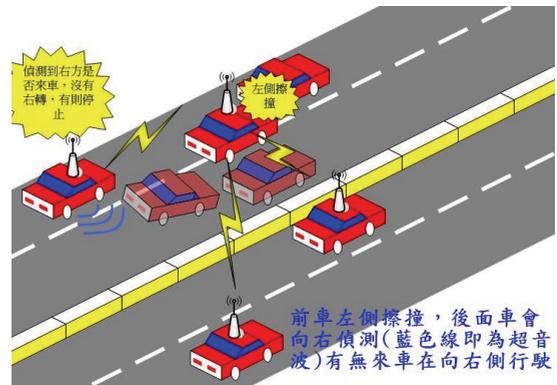


圖 6 模擬偵測左右來車示意圖

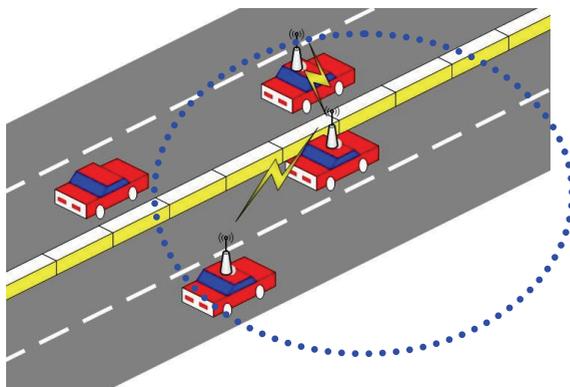


圖 4 車輛間通訊示意圖

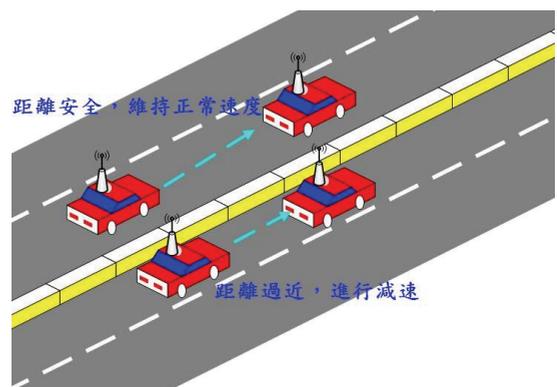


圖 7 車輛過近減速示意圖

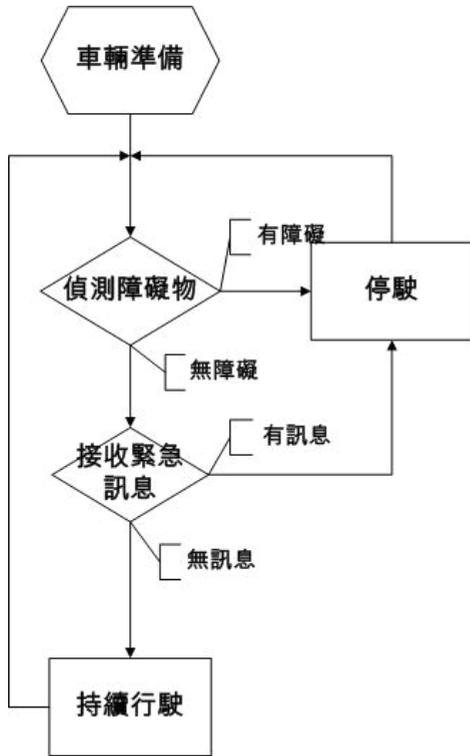
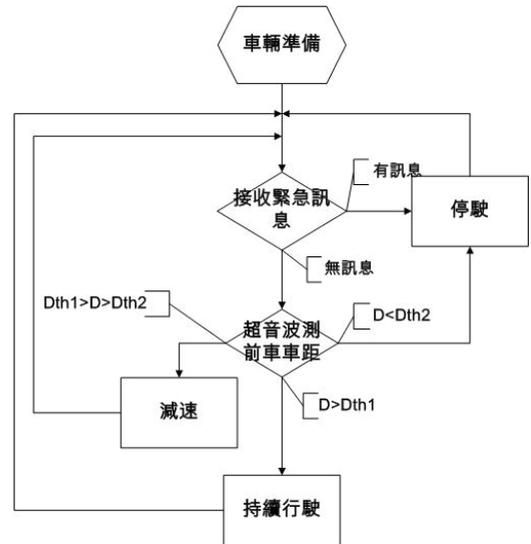


圖 8 觸碰及訊號接收車演算法流程



D→超音波測得與前車車距
 Dth1→車據接近應減速臨界值
 Dth2→車距過近應停止臨界值

圖 9 超音波及訊號接收車演算法流程

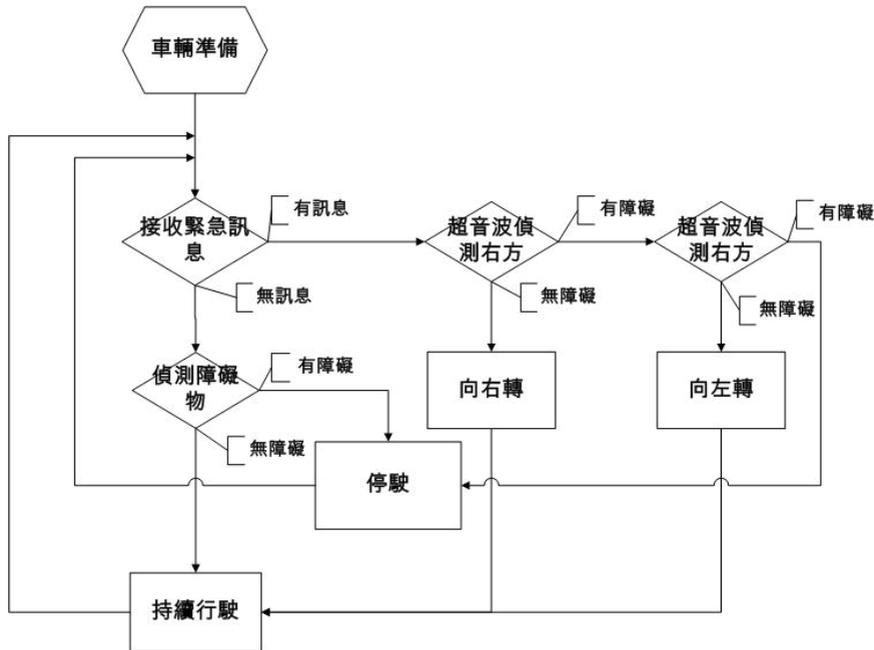


圖 10 雙側超音波車演算法流程

肆、模擬實驗

一、智慧型行車防撞安全系統實作設備簡介

這個系統實作採用颯公司的寶博智慧型機器車 04 台來模擬實際車輛行動，而無線網路通訊介面由於實驗模擬範圍不大，因此將透過 ZigBee 來進行車輛間溝通之無線協定，而寶博機器車可以外加 XBee 介面和超音波感測器，XBee 介面這是一個非常普遍的 2.4GHz XBee 模組 (Digi formally Maxstream)，如圖 11。XBee 模組是一個平價、低功率的無線感測網路，超音波感測器是感應距離的模組，簡單且低功率的感測裝置，如圖 12。這個模組利用簡單的 Serial command 來執行在 802.15.4 上的 ZigBee 通訊協定。因此可以支援車輛與車輛間或是固定服務基地台的通訊，不論是 point-to-point 或 multi-point networks，XBee 模組都可以支援，因此可以讓本論文提出之智慧型行車防撞安全系統 ISVAS 的架構可以達成。另外其它相關零件如下，如圖 13：

1. Whisker 鐵絲
2. 7/8 吋的 4-40 Phillips 平頭螺絲
3. 1/2 吋圓形墊片
4. 四號尼龍墊圈
5. 3PIN 排針
6. 電阻, 220 Ω (紅-紅-棕) *2

7. 電阻, 10 k Ω (棕-黑-橙)

8. TTL 介面 (讓 TX、RX 容易接到不同介面上)

9. XBEE 版 *1

10. LED (紅) *1

透過上面相關機器車零件及 XBee 模組組合後，我們初步完成模擬具有無線通訊及智慧型判斷功能之機器車將如圖 14。



圖 11 XBee 模組

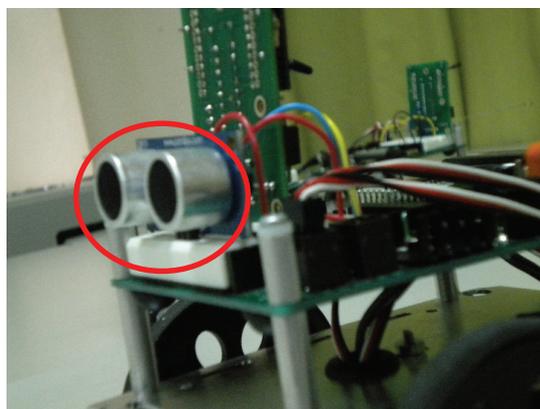


圖 12 超音波感測器

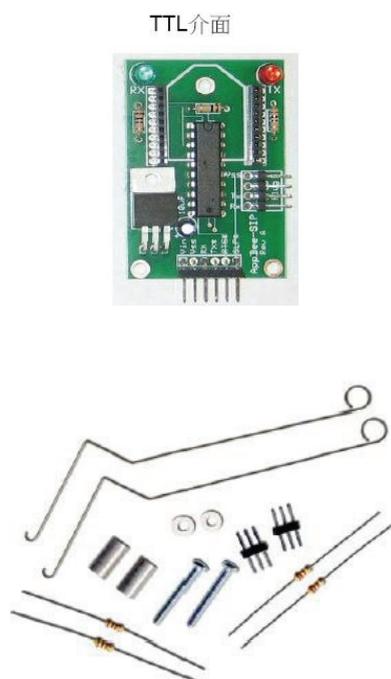


圖 13 機器車部分零件

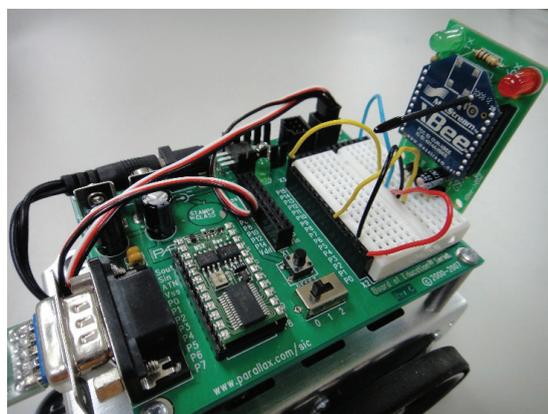


圖 14 模擬車輛之機器車及 Wireless 通訊介面

二、智慧型行車防撞安全系統程式簡介

由於是模擬車輛間互動，因此我們設計

機器車 1 號為前面之行車，並且準備碰觸裝置，一旦碰到障礙物時，就會自動停止，而在此時通知其它車輛，以便提供訊息來讓其它車輛可以進行安全行車自動預警之判斷。機器車 2 號接受到訊息會立即停駛，並判斷閃避路線繼續行駛，但如果左邊或右邊用超音波有感測到障礙物便會再向另一邊行駛，以避免車輛間之碰撞。機器車 3 號則只是簡單的做訊號收發，來確定自己是否繼續向前。而機器車 4 號，車頭裝載超音波感測器，感測是否與前側車輛保持適當距離，如果發現距離過短會馬上減速或立即停駛。如圖 15 左上角為 1 號車、左下為 2 號車，右上角為 3 號車、右下角為 4 號車。四台車都是正常行走。圖 16、圖 17、圖 18 分別為 4 號車判斷距離遠近並開始加減車速，LED 燈亮黃燈表示距離正常，亮黃綠燈表示距離太近，亮黃紅燈表示距離過近立即停止行駛，圖 19 為 1 號車正面撞及全體車輛立即停止圖 20 為 1 號車輛擦撞右邊，2 號車輛則偵測左方無來車自動向左方改道閃避，4 號車輛則因距離過近而停止行駛。

機器車 1 號：

此程式是採用觸鬚來判斷是否撞到障礙物，正常行走時 IN5 和 IN8 值為 1，XBEE 傳送給二號和三號車 light 為 0，當觸鬚撞到障礙物時（不管 IN5 或 IN8）則回傳值 1 改為 0，此時代表一號車撞到障礙物，一號車的 LED 燈亮起，同時間原本傳給二號和三

號車的 light 值從 0 改為 2，告知二號和三號車，一號車撞到障礙物。

部分程式碼如下：

```
' {$STAMP BS2}
' {$PBASIC 2.5}

FREQOUT 4, 2000, 3000
RX PIN 0
TX PIN 2
light VAR Word

HIGH TX
DO{

  IF (IN5 = 1) AND (IN8 = 1)
  THEN
  {
    LOW 14
    PAUSE 10
    light = 0
    GOSUB Forward_Pulse
  }ELSE{
    GOSUB Stop_Stop
    light = 2
    HIGH 14
    PAUSE 10
  }
  ENDIF

  SEROUT TX,84, [DEC
light,CR,CR]
  PAUSE 10

}LOOP

Forward_Pulse :
  PULSOUT 12,650
  PULSOUT 13,850
  PAUSE 20
  RETURN
```

```
Stop_Stop :
  PULSOUT 12,750
  PULSOUT 13,750
  PAUSE 20
  RETURN
```

機器車 2 號：

此程式是要接收一號車給的資訊，知道一號車是否正常行走，藉由 XBEE 的接收，傳回的值是 0 或 2，接收到 0 代表正常行走，接受到 2 則代表 1 號車撞到障礙物此時 2 號車會改變行駛方向變偵測改變之方向有無來車，再確定行駛方向，同時 LED 燈亮表示擦撞的方向。部分程式碼如下：

```
' {$STAMP BS2} '
' {$PBASIC 2.5} '
timeleft VAR Word
timeright VAR Word
pulseCount VAR Byte
Re :
DO
  LOW 4
  LOW 2

  IF (IN9 = 0) OR (IN11 = 0)
  THEN
    DO
      PULSOUT 3,5
      PULSIN 3, 1, timeleft
      PULSOUT 2,5
      PULSIN 2, 1, timeright
      IF (timeleft < 500) AND
(timeright < 500) T
    HEN
      HIGH 4
      HIGH 2
```

```

        GOSUB Back_Up
        GOSUB Turn_Left
        GOSUB Re
    ELSEIF (timeleft < 500)
THEN
    HIGH 4
    GOSUB Back_Up
    GOSUB Turn_Right
    GOSUB Re
    ELSEIF (timeright < 500)
THEN
    HIGH 2
    GOSUB Back_Up
    GOSUB Turn_Left
    GOSUB Re
    ELSE
    LOW 4
    LOW 2
    GOSUB Forward_Pulse
    ENDIF
    LOOP
ELSE
    LOW 4
    LOW 2
    GOSUB Forward_Pulse
    ENDIF
LOOP
'
-----subprogram-----
-----
Forward_Pulse : ' Send a single
forward pulse.
PULSOUT 13,850
PULSOUT 12,650
PAUSE 20
RETURN
Turn_Left : ' Left turn, about
90-degrees.
FOR pulseCount = 0 TO 20
PULSOUT 13, 650
PULSOUT 12, 650

```

```

PAUSE 20
NEXT
RETURN
Turn_Right :
FOR pulseCount = 0 TO 20 ' Right
turn, about 90-degrees.
PULSOUT 13, 850
PULSOUT 12, 850
PAUSE 20
NEXT
RETURN
Back_Up : ' Back up.
FOR pulseCount = 0 TO 40
PULSOUT 13, 650
PULSOUT 12, 850
PAUSE 20
NEXT
RETURN

```

機器車 3 號：

此程式是要接收一號車給的資訊，知道一號車是否正常行走，藉由 XBEE 的接收，傳回的值是 0 或 2，接收到 0 代表正常行走，接受到 2 則代表 1 號車撞到障礙物此時 3 號車馬上停止行走，同時 LED 燈亮。部分程式碼如下：

```

' {$STAMP BS2}
' {$PBASIC 2.5}
RX PIN 0
TX PIN 2
a VAR Byte
light VAR Byte
counter VAR Byte
a = 0
HIGH TX
DO
SERIN RX, 84, [DEC light]
DEBUG DEC light, CR

```

```

IF (light = 3) THEN
  PULSOUT 13, 750
  PULSOUT 12, 750
ELSEIF (light = 2) AND (a = 0)
THEN
  FOR counter= 1 TO 20
  PULSOUT 12,850
  PULSOUT 13,850
  PAUSE 20
  NEXT

  FOR counter= 1 TO 45
  PULSOUT 13,850
  PULSOUT 12,650
  PAUSE 20
  NEXT

  FOR counter= 1 TO 20
  PULSOUT 12,650
  PULSOUT 13,650
  PAUSE 20
  NEXT

  a = 1
  ELSEIF (light = 1) AND (a = 0)
  THEN
    FOR counter= 1 TO 20
    PULSOUT 12,650
    PULSOUT 13,650
    PAUSE 20
    NEXT

    FOR counter= 1 TO 45
    PULSOUT 13,850
    PULSOUT 12,650
    PAUSE 20
    NEXT

```

機器車 4 號：

此程式是要接收超音波感測器給的資訊，感測前方車輛是否與本車輛保持是當距離，藉由感測器測量的數值，來判斷車輛該不該減速行駛或者立即停止。部分程式碼如下：

```

' {$STAMP BS2}
' {$PBASIC 2.5}

Time VAR Word
DO
  PULSOUT 15,5
  PULSIN 15, 1, time

  'DEBUG HOME, " time = ",DEC5
  time

  IF time < 300 THEN
    PULSOUT 13,750
    PULSOUT 12,750
    PAUSE 10
  ELSEIF time < 650 THEN
    PULSOUT 13,770
    PULSOUT 12,730
    PAUSE 10
  ELSEIF time < 1000 THEN
    PULSOUT 13,780
    PULSOUT 12,720
    PAUSE 10
  ELSEIF time < 1200 THEN
    PULSOUT 13,790
    PULSOUT 12,710
    PAUSE 10
  ELSE
    PULSOUT 13,850
    PULSOUT 12,650
    PAUSE 10
  ENDIF
LOOP

```

除了前面 ISVAS 系統基本實作外。我們也同時作了初步不同方式在安全停車上的比較，比較方式如圖 16 中 3、4 號車，如果兩台車一台有 ZigBee 無線通訊模組而另一台僅有無線超音波模組，當前面發生事故時開始，兩車實際停下車所需要耗費之時間量測，相關參數（表 1）及實驗結果如

表 2。我們可以發現透過無線通訊發送警訊有不錯的效果

表 1 無線通訊與超音波停車時間比較

項目	數值
ZigBee 頻率	2.4GHz
ZigBee 訊號範圍	半徑 5 公尺
超音波感應距離	最遠 2 公尺
D_{th1}	30 公分
D_{th2}	10 公分
機器車車速（原始）	40cm/s
機器車車速（減速時）	13cm/s

表 2 無線通訊與超音波停車時間比較

	無線通訊車	超音波車
停車時間	0.5 秒	4.6 秒

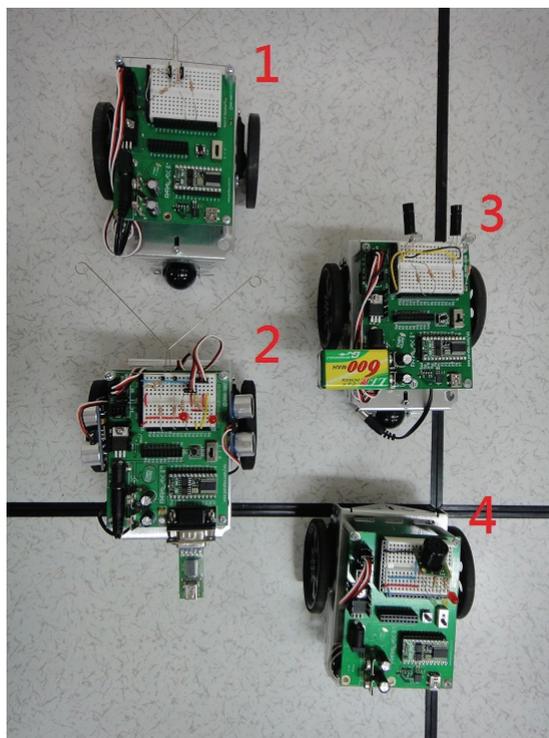


圖 15 模擬正常行車之車輛

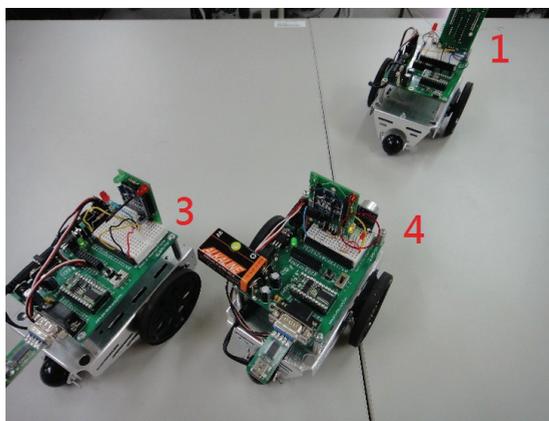


圖 16 模擬 4 號車輛距離維持正常，亮黃燈

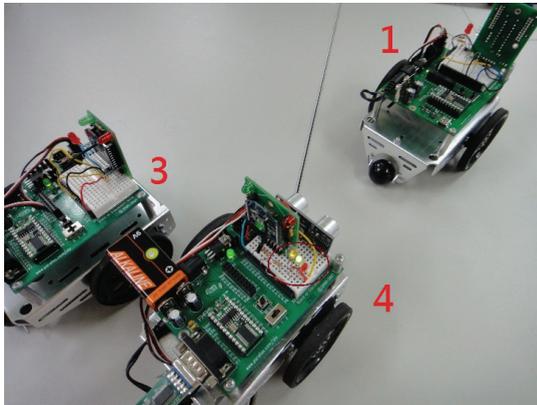


圖 17 模擬 4 號車輛距離變短並減速，亮黃綠燈

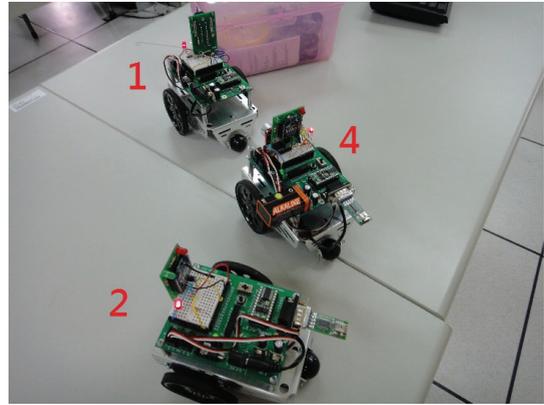


圖 20 模擬行車事故及車輛通訊動作

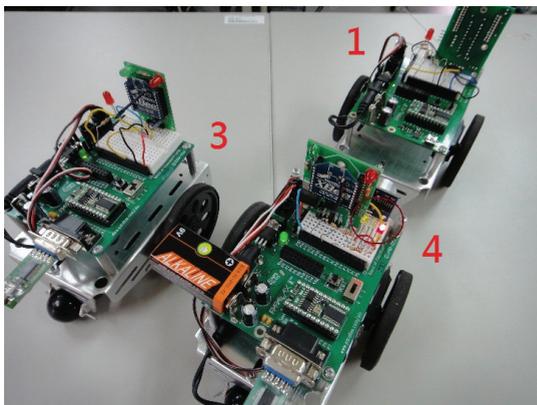


圖 18 模擬 4 號車輛距離過近，立即停止，亮黃紅燈

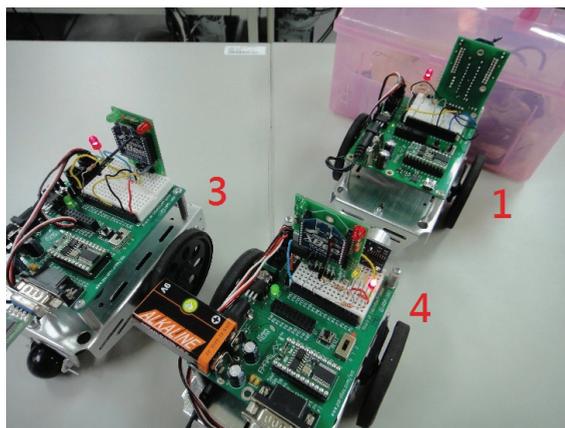


圖 19 模擬行車事故及車輛通訊動作

伍、結論

行車駕駛安全是一直長久存在的交通安全議題，相關學者、專家及研究人員，也一直不斷地以更先進的技術提供更佳的安全系統來保護駕駛者與乘客。在本論文中，為了在車輛發生碰撞時，可以即時的發送警示給其它車輛和駕駛者，因此提出了一個簡單的智慧型行車防撞安全系統 (Intelligent Safety Vehicle Anti-collision System, ISVAS)，希望讓行車車輛間可以除了夠駕駛者肉眼觀察行車狀況外而作行車車輛控制外，可以讓車輛本身具有智慧功能，可以作一些事故的即時處理或是通知，這樣將可以有效增進行車安全性，並且將本系統簡單實作呈現，而在實作結果中可以發現，在非行車發生事故之車輛，仍然可以透過無線訊號通知後，讓其它車輛可以即時的煞車，以避免車輛碰撞。而在未來的研究中，將持續以系統開發、並希望延伸 ISVAS 的效能和增加

其它狀況之判斷等為主要研究目標，以解決 ISVAS 目前只能適用特定的情況。也希望可以讓 ISVAS 有機會實際運作於交通運輸環境之中。

參考文獻

- [1] [http : //safety.iot.gov.tw/techknowledge/files/先進大客車安全配備簡介.pdf](http://safety.iot.gov.tw/techknowledge/files/先進大客車安全配備簡介.pdf)
- [2] [http : //www.autooo.net/utf8-classid36-id3188.html](http://www.autooo.net/utf8-classid36-id3188.html)
- [3] 交通部台灣區國道高速公路局, [http : //www.freeway.gov.tw/](http://www.freeway.gov.tw/)
- [4] Y. Chen, and L. Chen, "Using Bluetooth Wireless Technology in Vehicles," Proceedings of the IEEE International Conference on Vehicular Electronics and Safety, Oct. 2005, pp.344-347.
- [5] K. Selvarajah, A. Tully, and P.T. Blythe, "ZigBee for Intelligent Transport System Applications," Proceedings of the Road Transport Information and Control (RTIC 2008) , May 2008, pp. 1-7.
- [6] Y.S. Chen, C.H. Cheng, C.S. Hsu, and G.M. Chiu, "Network Mobility Protocol for Vehicular Ad Hoc Networks," Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC 2009) , April 2009, pp. 2162-2167.
- [7] A. Alonso, A. Paier, T. Zemen, N. Czink, and F. Tufvesson, "Capacity Evaluation of Measured Vehicle-to-Vehicle Radio Channels at 5.2 GHz," Proceedings of the IEEE International Conference on Communications Workshops (ICC) , May 2010, pp. 1-5.
- [8] S. Shin, M. Kang, J.-S. Jin, S.-U. Kim, H.-L. Kim, and S. Moh, "Vehicle-to-Vehicle Emergency Message Dissemination through the WiBro Network," Proceedings of the 6th International Conference on Networked Computing (INC) , May 2010, pp. 1-6.
- [9] IEEE Std. 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks—Part 16 : Air Interface for Fixed Broadband Wireless Access Systems," Oct. 2004.
- [10] IEEE Std. 802.16e-2006, "IEEE Standard for Local and Metropolitan Area Networks—Part 16 : Air Interface for Fixed Broadband Wireless Access Systems—Amendment : Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands," Feb. 2006.
- [11] TTAS.KO-06.0082/R1 WiBro Standard Phase-II, "Specifications for 2.3GHz band Portable Internet Service (PHY & MAC Layer) ," Dec. 2005.
- [12] IEEE Std 802.15.4-2006, Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications

- for Low-Rate Wireless Personal Area Networks (LR-WPANs). New York, NY: IEEE, Sep. 2006.
- [13] J. S. Lee, "Performance Evaluation of IEEE 802.15.4 for Low-rate Wireless Personal Area Networks," IEEE Transactions on Consumer Electronics, vol. 52, no. 3, Aug. 2006, pp. 742-749.
- [14] W. Zhu, "Wireless Technology and Web," Proceedings of the IEEE 6th Circuits and Systems Symposium, vol. 1, June 2004, pp. 1-26.
- [15] N. Baker, "ZigBee and Bluetooth Strengths and Weaknesses for Industrial Applications," Computing & Control Engineering, vol. 16, no. 2, April 2005, pp. 20-25.
- [16] 李俊賢, "ZigBee 技術規格與測試方案之發展", 電腦與通訊, vol. 119, 2007年3月, pp. 19-28。
- [17] 施松村、項家茜, "應用 Zigbee 技術的無線感測網路", 第七屆離島資訊技術與應用研討會論文集, 2008年5月, pp.310-318。
- [18] R. Peng, S. Mao-heng, and Z. You-min, "Zigbee Routing Selection Strategy Based on Data Services and Energy-balanced Zigbee Routing", Proceedings of the IEEE Asia-Pacific Conference on Services Computing (APSCC'06), Dec. 2006, pp.400-404.
- [19] IEEE Std 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," June 2003.
- [20] IEEE Std 802.15.1-2002, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs), April 2002.
- [21] 江叔盈、陳弘斌、林重州、羅雅畔、楊靜安、李郁溥, "無線感測網路於環境及居家生活之應用", 研討會論文, 2004年。
- [22] 林國煌、黃偉特、詹博宇, "ZigBee 無線感測網路技術於家庭自動化控制之實現", 全國優質家庭學術研討會論文集 (UHC2008), 2008年11月。
- [23] 張祐城、林欣緯、陳朝陽, "ZIGBEE 無線感測網路於廚房瓦斯洩漏偵測之應用" Proceedings of the Information Education and Technological Applications Conference 2009 (IETAC 2009), Taichung, Taiwan, Nov. 2009, pp. 449-452.
- [24] 張祐城、陳朝陽、林欣緯, "ZIGBEE 無線感測網路技術結合紅外線應用於數位家電控制" Proceedings of the Information Education and Technological Applications Conference 2009 (IETAC 2009), Taichung, Taiwan, Nov. 2009, pp. 453-458.

-
- [25] M. Zhou, H. Zhu, and Shan Zhu, “A Design of Automotive Anti-Collision Control System Based on ARM9,” Proceedings of the Second International Conference on Intelligent Computation Technology and Automation, 2009 (ICICTA '09) , Oct. 2009, pp. 109-112.
- [26] W. Cho, K.-S. Han, H. K. Choi, and H. S. Oh, “Realization of Anti-Collision Warning Application Using V2V Communication,” Proceedings of the IEEE Vehicular Networking Conference (VNC 2009) , Oct. 2009, pp. 1-5.
- [27] A.-R. Huang, Z.-T. Xiang, W.-R. Jiang, and Y.-F. Chen, “Vehicle Auxiliary Anti-Collision Warning System Based on Positioning in Electronic Map,” Proceedings of the International Conference on Information Engineering and Computer Science, 2009 (ICIECS 2009) , Dec. 2009, pp. 1-4.
- [28] IEEE Std 802.11-2007 : IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,” June 2007.
- 飊 機 器 人 - 寶 博 機 器 車 ， [http :
//playrobot.com/](http://playrobot.com/)
-

以化學氣相沉積法合成一維氧化銦鋅 奈米線結構及其特性分析

陳志榮

摘要

本論文主要利用化學氣相沉積法經由氣液固 (VLS) 機制在低製程溫度下成長一維氧化銦鋅 (IZO) 奈米線於 ITO 基板上，並使用掃描電子顯微鏡觀察一維氧化銦鋅奈米線樣本的表面形態、結構、長度、線徑與分佈情形，並利用 X 光單晶繞射儀鑑定其結構的結晶性，再利用 X 光能量散譜儀 (EDS) 做一維氧化銦鋅奈米線的成份分析。本實驗在成長一維氧化銦鋅奈米線的最佳實驗結果 (參數為銦 (In) 顆粒的製程溫度為 700 °C、鋅 (Zn) 粉末的製程溫度為 550 °C、成長壓力為 1 Torr、氫氣流量為 50 sccm、氧氣流量為 10 sccm)，由 FE-SEM 的截面分析發現，本實驗氧化銦鋅奈米線為中空結構，奈米線線徑約為 100 nm 到 150 nm，長約為 5 μ m，電阻值約 20 - 150 Ω 。

關鍵詞：化學氣相沉積法、氧化銦鋅、奈米線。

Synthesis and Characterization of One-Dimensional Indium Zinc Oxide Structure by Chemical Vapor Deposition

Chih-Jung Chen

Abstract

In this study, we report the fabrication of one-dimensional Indium Zinc Oxide (IZO) on an ITO glass substrate by chemical vapor deposition at low temperature via vapor-liquid-solid (VLS) mechanism. The morphologies and structural properties of the IZO nanowire were characterized by field emission scanning electron microscopy (FE-SEM), X-ray diffraction (XRD) and energy dispersive spectrometer (EDS). The IZO nanowires optimal experimental parameters are growth pressure is 1 Torr, argon gas flow of 50 sccm and oxygen flow of 10 sccm. In this condition, the idea temperatures are 700 °C for indium (In) particles and 550 °C for zinc (Zn) particles. The SEM images reveal that IZO nanowires have uniform diameters of approximately 100 nm to 150 nm, and their lengths are up to 5 μm

The XRD results demonstrate that the nanowires are crystalline with highly preferential orientation. And our EDS analysis also proves that the composition ratio and this material are IZO nanowires.

Keywords : chemical vapor deposition, IZO, nanowires.

壹、前言

透明導電膜的演進是在1902年時從第一個被發現具有透明導電性質材料的氧化鎘(CdO)[1]開始，而第一個將氧化鎘(CdO)薄膜沉積在玻璃上是在1907年由Badeker所發表，直到1942年二氧化錫(SnO₂)[2]作為透明導電膜專利被發表提出，而後來隨著許多科技人才以及有關方面的研究學者持續不斷的在這幾十年不斷的以透明導電膜作為研究方向，已經開始將對透明導電膜的注意力著重於氧化銦錫(Tin-doped Indium Oxide, ITO)及氧化鋅(ZnO)系列的材料，並且開始非常廣泛以及普遍的地應用在各種光電以及其他領域上。

透明導電氧化物(Transparent Conduction Oxide, TCO)[3]是一種同時具備透光性以及導電性兩種特性的材料，例如：氧化銻錫(Antimony Tin Oxide, ATO)、氧化銦錫(Indium Tin Oxide, ITO)的優點為高透光率、低電阻、高載子遷移率，而其中在科技業以及工業界所用之透明導電玻璃皆以氧化銦錫為主要材料，在可見光區範圍其透光度較高，並且其電阻係數值低於 $10^{-3} \Omega\text{-cm}$ ，由於此種材料具有高透光率以及較低的電阻率，而另一項優點是此材料在化學使用上的穩定性也相對的提高許多，所以透明導電

材料至今一直在工業界以及相關研究領域中被廣泛使用。

大部分的透明導電氧化物的能階都是由於自由載子的改變而導致能階也有所變化，而其多數載子為電子，所以一開始先激發的電子會先佔據導帶之最低能階處，剩下的電子則需要更多的能量(例如：光、熱)才能使得電子可以從價電帶被激發至傳導帶，而在一般常見的單一種元素所形成的氧化物其能隙範圍大都在3 eV以上，但是此一範圍在可見光波範圍中的光波並沒有辦法將電子從價電帶激發到傳導帶，導致此一類型的氧化物半導體無法對可見光波範圍的光產生吸收，所以單一種元素氧化物中的透明導電氧化物(TCO)材料在可見光範圍是屬於透明的材料，但其電子的傳導性卻不甚良好，所以此類的氧化物都需要透過異質離子的摻雜增加半導體材料的導電性或製造氧化不完全的半導體化合物亦或者是半導體內有離子空缺才能有較高的導電性，也因此本實驗可以透過異質離子的摻雜以增加半導體的導電性，並且選擇摻入比原化合物陽離子價數多一價的金屬離子，或是添加其陰離子少一價的非金屬離子[4]。

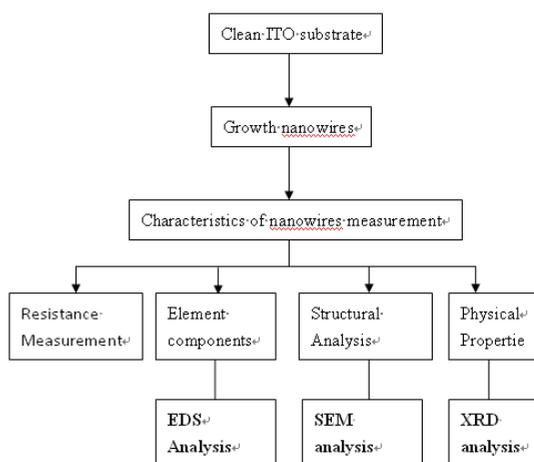
透明導電氧化物(TCO)是屬於一種可透光性及可導電性兩種特性混合的材料。目

前透明導電氧化物已被廣泛應用在太陽能電池 (Solar Cell) [5]、燃料電池 (Fuel cell) [6]、發光二極體 (Light-emitting diode, LED) [7] 以及光感測器 (Photonic sensor) [8]... 等方面, 而氧化銦鋅 (Indium-Zinc Oxide, IZO), 其透光率在可見光範圍與氧化銦錫 (ITO) 類似, 而其導電性也較優異, 再者鋅 (Zn) 是屬於一種無毒性的材料, 因此提供了一種可以取代傳統氧化銦錫 (ITO) 的方法。

而本實驗研究中, 利用一種不需要提高太多成並且在實驗上比較簡單能夠可以提供較低溫的製程方式-化學氣相沉積法 (Chemical Vapor Deposition, CVD), 來成長氧化銦鋅 (IZO) 奈米線。

貳、實驗方法

本實驗主要研究方向是製作出氧化銦鋅 (IZO) 奈米線, 並對其結構型態、成份及分佈情形做分析。



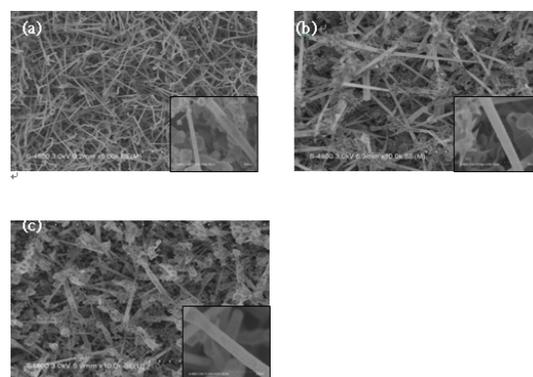
圖一、一維氧化銦鋅 (IZO) 奈米線製備流程圖

圖一為本實驗製備氧化銦鋅 (IZO) 奈米線的實驗流程圖, 一開始我們將 ITO 玻璃基板切割適當大小的尺寸, 並且將清洗完畢之後的氧化銦錫 (ITO) 玻璃黏在石英載玻片上, 然後再將實驗所需要的銦 (In) 和鋅 (Zn) 粉末調配為實驗所需的重量比分別置於坩鍋中, 將銦 (In) 粉末放於儀器的上風處並設定位置所需的實驗溫度, 鋅 (Zn) 粉末放於上風處 550 °C, 另外將基板放於下風處, 溫區範圍為 500 °C - 400 °C, 之後將整個實驗機台內部的腔體抽到背景壓力為 5×10^{-2} Torr 以下開始通氬氣 (Ar) 並且開始加溫, 當儀器的溫度到達製程的溫度時, 開始通入氬氣和氧氣, 並調整成長壓力, 反應 1 小時, 等待高溫爐冷卻後將實驗樣品取出以利後續量測。

在利用化學氣相沉積法 (CVD) 製備完實驗樣品之後，本實驗在觀察樣本的形貌以及巨觀結構時是利用逢甲大學共同貴重儀器中心的 HITACHI S4800 場發射型掃描式電子顯微鏡 (Field Emission SEM, FESEM) 來輔助觀察一維氧化銦鋅 (IZO) 奈米線的表面型態並且從樣本的橫截面得知一維氧化銦鋅 (IZO) 奈米線的長度以及寬度。另外本實驗在評定一維氧化銦鋅 (IZO) 奈米線晶體結構的變化情形時，是利用 X 光射線繞儀來分析，使用的機台型號為 D8 SSS，其操作如下：X-ray 的光源為銅靶 ($\text{Cu K}\alpha$ ， $\lambda = 1.54051 \text{ \AA}$)，分析的條件為 40 KV、20 mA，掃描範圍 (2θ) 為 $20^\circ - 70^\circ$ ，掃描速度為 $6^\circ/\text{min}$ 。將所得的實驗結果透過 JCPDS (Joint of Committee on Powder Diffraction Standards) 卡資料做結晶類型比對。而場發射型掃描式電子顯微鏡 (FE-SEM)，其型號為 HITACHI S4800，操作時加速電壓為 3KV，藉由訊號接收器擇取二次電子 (Secondary electron images, SEI) 或背向散射電子 (Backscattered Electron) 方式就可以觀察到一維氧化銦鋅 (IZO) 奈米線在 ITO 玻璃基板上的成線情形以及表面分佈情況並且得知奈米線的長度及寬度的尺寸，另外本實驗為了進一步探討氧化銦鋅 (IZO) 奈米線其銦 (In) 和鋅 (Zn) 所含

有成份的比例，所以藉由 X 光能量分散光譜 (Energy Dispersive Spectroscopy, EDS) 觀測氧化銦鋅 (IZO) 奈米線所包含的成份，其 X 光能量分散光譜分析原理為當電子束通過試片時，會將材料內層軌域的電子打掉，此時較高能階的電子會跳入低能階，並放出以 X 光形式出現的能量，不同的元素能階差不同，放出的 X 光能量也不同，故此稱為特徵 X 光。EDS 就是利用偵測特徵 X 光來分析試片中的成份，EDS 採用矽偵測器，矽偵測器在 3 - 15 keV 的範圍內接收效率接近 100%，所以本實驗藉由 X 光能量分散光譜來作為奈米線的成份分析。

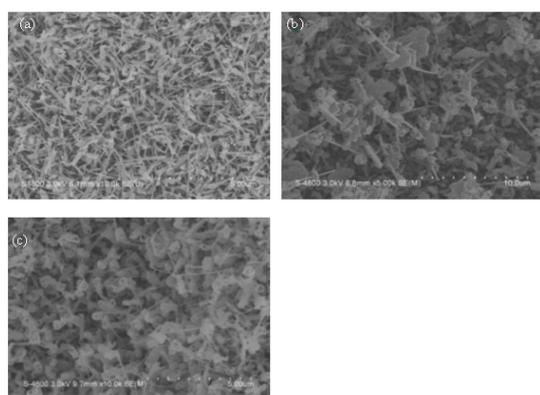
參、結果與討論



圖二、一維氧化銦鋅 (IZO) 奈米線在相同條件下以不同的銦 Source 溫度成長之 SEM 圖 (a) 700°C ，(b) 800°C ，(c) 900°C

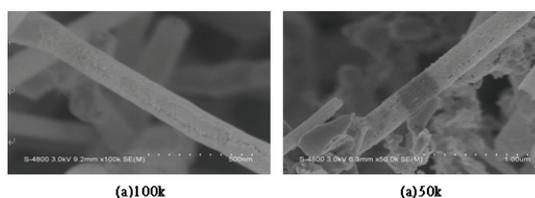
圖二為氧化銦鋅 (IZO) 奈米線以不同的銦 Source 材料溫度成長的 SEM

圖，由SEM圖得知銻(In)顆粒溫度的變化，會造成氧化銻鋅(IZO)奈米線表面沉積的變化，這結果是從本系列的實驗結果推論當基板溫度變化時會影響銻和鋅這兩種材料蒸氣量的凝聚會形成金屬顆粒的大小，而從實驗結果得知當銻(In)顆粒溫度達到900 °C時，氧化銻鋅(IZO)奈米線表面所吸附的附著物會比銻(In)顆粒溫度為700 °C及800 °C來的多，而會造成這種情形是由於銻的蒸氣量會隨著溫度的變化而改變，所以隨著銻(In)顆粒溫度降低，則銻的揮發量也會減少以致於氧化銻鋅(IZO)奈米線表面的附著物會逐漸減少，如圖二所示，這是由於銻(In)粉末材料的蒸氣量會隨著溫度提升而增加，進而造成一維氧化銻鋅(IZO)奈米線形貌的變化。



圖三、氧化銻鋅(IZO)奈米線在相同條件下不同的氧氣(O₂)流量之SEM圖(a) 10 sccm, (b) 5 sccm, (c) 0 sccm

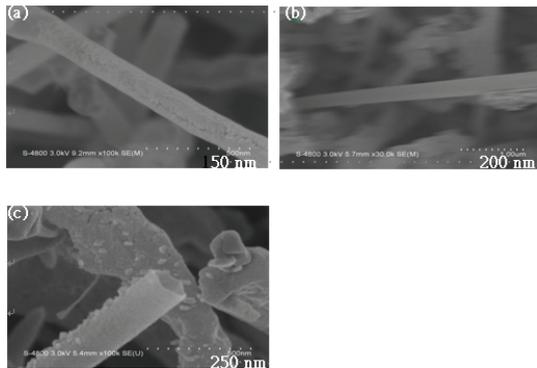
圖三為氧化銻鋅(IZO)奈米線在相同條件下利用不同的氧氣(O₂)流量所做的實驗之SEM圖，由圖可知在基板溫度相同的情況下，如果改變通氧(O₂)量，會改變奈米線表面上的形貌和線徑的大小，這是由於氧氣流量大小會造成氧氣在爐管內的含量有所改變，進而直接影響銻材料粉末以及鋅材料粉末的蒸汽量，導致如果氧氣量的增加也會提高氧化銻鋅(IZO)奈米線的成长速度，進而造成奈米線較好的成長位置會往較低溫位置移動。因此，由實驗結果得知其氧氣(O₂)流量大小會影響氧氣在爐管內的含量，當氧氣(O₂)流量增加，爐管內的氧氣(O₂)會增加，使得銻(In)和鋅(Zn)蒸氣量變多，因而造成奈米線上會有不同的沉積導致形貌上的改變。



圖四、氧化銻鋅(IZO)奈米線在相同條件下不同的氬氣(Ar)流量之SEM圖(a) 50 sccm, (b) 75 sccm

圖四為氧化銻鋅(IZO)奈米線在相同條件下通入不同的氬氣(Ar)流量之SEM圖，從實驗中可以觀察出當氬氣

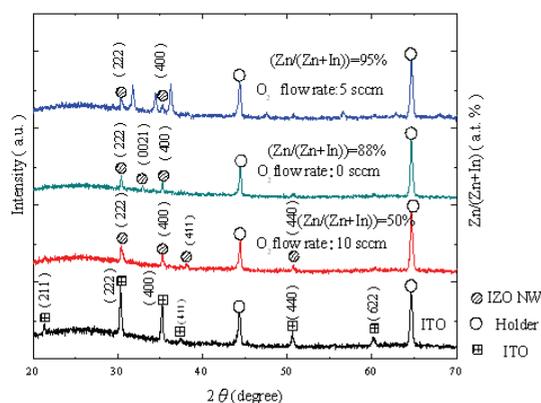
(Ar) 流量為25 sccm時，氧化銦鋅 (IZO) 奈米線無法形成線狀，反而形成一點一點的顆粒狀；而當氬氣 (Ar) 流量達到50 sccm 以上才能形成線狀，這是由於氬氣 (Ar) 在成長奈米線時，會帶送銦 (In) 和鋅 (Zn) 的蒸氣量往基板上沉積並且往低溫處移動，而氬氣流量越大會使得銦和鋅蒸氣往低溫處移動以使得最佳奈米線的成長位置也會往低溫處方向移動，而當基板在相同的位置時，從實驗結果可以得知奈米線的線徑會隨著氬氣流量的改變造成線徑大小變化，如圖四所示。



圖五、氧化銦鋅 (IZO) 奈米線在相同條件下不同成長壓力之 SEM 圖 (a) 1 Torr, (b) 10 Torr (c) 100 Torr

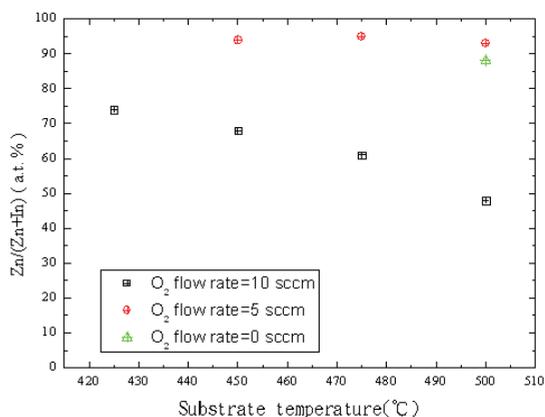
圖五為氧化銦鋅 (IZO) 奈米線在相同條件下以不同的成長壓力下實驗的SEM圖，從製程的成長壓力之下觀察，本實驗結果發現奈米線的線徑變化以及

密度會被成長壓力所影響，如圖五(a)~(c)所示。當成長壓力為1 Torr時，奈米線的線徑約為150 nm。當成長壓力為10 Torr時，奈米線的線徑約為200 nm，而成長壓力為100 Torr時，奈米線的線徑約為250 nm。因此，從實驗結果得知，成長壓力會直接影響奈米線的線徑大小，當成長壓力越大，奈米線的線徑會越粗，而成長壓力越小時，奈米線的線徑會越細，本文推論成長壓力會直接影響奈米線的線徑大小，當成長壓力越大，奈米線的線徑會越粗，而成長壓力越小時，奈米線的線徑會越細，其原因就在於當成長奈米線時在低壓下過飽和濃度較高，其銦 (In) 和鋅 (Zn) 粉末不易蒸發，而在高壓下過飽和濃度就會減少，所以在低壓下更多的銦和鋅的粉末會較容易產生蒸氣量而揮發，而在本實驗過程當中在低的成長壓力下約1 Torr，成長奈米線的基板溫度較低 (約在500 °C- 400 °C)，而在高的成長壓力下約100 Torr，成長奈米線的基板溫度較高 (約在500 °C- 475 °C)，而在相關文獻探討[9]中也提出了奈米線在較高的壓力底下，其奈米線的冷凝範圍會轉化到較高的溫區，出現這種情況是因為在較高的壓力底下，必須有更多的蒸氣，以保持正確的濃度，所以在高溫較易成長奈米線，所以成長壓力會有可能改變奈米線的直徑及密度。



圖六、不同氧氣 (O_2) 流量之氧化銦鋅 (IZO) 奈米線之 XRD 圖

圖六為不同氧氣 (O_2) 流量之氧化銦鋅 (IZO) 奈米線之 XRD 圖，從 X-ray 繞射分析可以得知隨著通氧量的變化，奈米線的結構會有明顯的改變，如圖六所示，通 10 sccm 時，鋅所佔的原子百分比 ($Zn/(Zn+In)$) 為 50 % 時整個晶相為氧化銦結構，三個主訊號分別為 (222)、(400)、(411)，JCPDS 編號為 65-3170，隨著通氧氣 (O_2) 量變化，晶相會轉變，當鋅粉末材料含量為 88 % 時，出現了 $Zn_5In_2O_8$ 的訊號，晶相為 (0021)，鋅粉末材料含量為 95 % 時，訊號整個轉變為氧化鋅訊號，三個主訊號分別為 (100)、(002)、(101) JCPDS 編號為 65-3170。因此，結合 EDS 和 XRD 確認奈米線為氧化銦鋅 (IZO) 奈米線。



圖七、通氧氣 (O_2) 量的大小與鋅 (Zn) 和銦 (In) 原子百分比 ($Zn/(Zn+In)$)

圖七為 EDS 分析，從圖七可以知道通氧 (O_2) 量的大小與鋅 (Zn) 和銦 (In) 的原子百分比，並且由本實驗的數據可以得知，氧氣 (O_2) 流量的增加，會使的銦 (In) 蒸氣量明顯的增加，當氧氣 (O_2) 流量為 5 sccm 及 0 sccm 時，鋅 (Zn) 含量均在 90 % 以上，但是當氧氣 (O_2) 流量為 10 sccm 時，鋅 (Zn) 的比率明顯下降，因此，推論氧氣 (O_2) 流量會導致銦 (In) 和鋅 (Zn) 材料粉末的蒸氣有所揮發，而氧氣 (O_2) 流量越大時，銦 (In) 和鋅 (Zn) 材料粉末的蒸氣量揮發就越多，而從文獻[9]當中亦可以得知氧氣越多會造成生長加速，使得不同型態的奈米線形成。

肆、結論

本實驗利用化學氣相沉積法經由VLS機制 (Vapor-liquid-solid, VLS)，成功在低溫下成長氧化銦鋅 (IZO) 奈米線在ITO玻璃上。

經由FE-SEM觀察發現，奈米線會隨著溫度變化而影響奈米線沉積的密度、形貌以及線徑，其中基板溫度越低則奈米線的線徑越細，反之，基板溫度越高則奈米線的線徑越粗，且藉由調控成長壓力及氬氣 (Ar) 流量的大小也會影響奈米線的線徑大小。而經由本實驗的結果可以了解成長壓力為1 Torr時，線徑約為100 nm - 150 nm。而在100 Torr時，線徑約為250 nm。本實驗最佳的實驗結果 (參數為銦 (In) 顆粒為700 °C、鋅 (Zn) 粉末為550 °C、成長壓力為1 Torr、氬氣流量50 sccm、氧氣流量10 sccm)，本實驗所備製之一維氧化銦鋅錫 (IZO) 奈米線，由於具有較低的製程溫度、低真空以及較低成本的製作等優點，所以由此評估氧化銦鋅 (IZO) 適合作為透明導電氧化物 (TCO) 之應用。

伍、參考文獻

- [1] F. Streintz, *Annals of Physics*, (Leipzig), 9 854, (1902) .
- [2] J.T. Littleton, U.S. Patent, 2,118,795, (1938) .
- [3] Tadatsugu Minami, "Substitution of transparent conducting oxide thin films for indium tin oxide transparent electrode applications" *Thin Solid Films*,516, (2008) , pp.1314-1321.
- [4] 黃奕盛 (2006)。「含奈米金為例之AZO透明導電膜的研究」。材料工程研究所碩士論文，私立大同大學。
- [5] 顧鴻壽 (2008)。太陽能電池元件導論。全威圖書。
- [6] N. Saito, H. Haneda, T. Sekiguchi, N. Ohashi, I. Sakaguchi and K. Koumoto, "Low-Temperature Fabrication of Light-Emitting Zinc Oxide Micropatterns Using Self-Assembled Monolayers," *Adv Mater.***14** (2002) , pp. 418-421.
- [7] W. Y. Weng, S. J. Chang, C. L. Hsu, T. J. Hsueh, and S. P. Chang, "A Lateral ZnO Nanowire Photodetector Prepared on Glass Substrate," *Journal of The Electrochemical Society*, **157**(2)(2010),

pp. K30-K33.

- [8] 劉吉平、赫向陽編著，林鴻明審校，
（2003）。奈米科學與技術。世貿出版社。
- [9] S H Dalal, D L Baptista, K BKTeo, R G Lacerda, D A Jefferson and W I Milne, “Controllable growth of vertically aligned zinc oxide nanowires using vapour deposition,” *Nanotechnology*, 17 (2006), pp. 4811–4818
-

Innovation dynamics : decision making in changeable spaces based on habitual domains

Yen-Chu Chen

Abstract

Human behaviors involve dynamic, evolving, interactive, adaptive processes. Important decision making, as a part of human behaviors, is usually dynamic and involves changeable parameters. These parameters can interact with each other and vary with time, the situation, and changes in the psychological states of the decision makers involved. According to the habitual domains theory, decision making can reach a steady state and exhibit habitual patterns as time passes. As a consequence, people may unwittingly assume that the decision parameters have fixed known dimensions and ranges. However, in real life, the parameters might or might not be noticed. Even when they are noticed, their dimensions and ranges cannot be predetermined. Decision making with this kind of feature, is called “decision making in changeable spaces”. Corporate innovation problems are of this type.

Innovation itself is a dynamic process, which includes transforming competence sets for innovation, producing products or services to release the pains and frustrations of target groups, and creating and distributing value. In the field of innovation studies, no framework has systematically described these processes in the past. This research is the first attempt to integrate these components into a single system.

Business innovation itself involves decision making problems in changeable space, including transforming competence set for innovation, producing products or services to release the pain and frustration of target groups, creating values, and distributing values. Based on habitual domains theory and its related competence set analysis, this article introduces the concepts of habitual domains and decision makings in changeable spaces as to describe the dynamics of human behavior and the changing nature of decision making problems. It explores the expansion of competence sets and creation of value, and proposes an integrated framework, *Innovation Dynamics*. Two corporate cases are discussed to verify the framework. By examining the operations of each link in innovation dynamics, corporations can understand if each and all links are properly developed, so that they can continually upgrade their products/services and create maximal value by releasing the pains and frustrations in potential domains of customers.

Keywords : Habitual Domain, Decision Making in Changeable Space, Competence Set Analysis, Innovation Dynamics.

創新動態學：以習慣領域為基礎探討 可變空間之決策制定

陳彥曲

摘要

人類行為乃由動態且複雜的流程所組成。決策行為係人類行為的一部份，因此亦含有許多動態變化的參數。這些參數互相影響，且會隨著時間、情境，乃至於決策者的心理狀態變化而有所改變。根據習慣領域理論，決策行為雖因不同參數的變化而改變，卻會隨著時間的經過而達到穩定狀態，並且具有習慣性；這使得人們常不自覺地假設決策參數的維度或範圍是固定的。然而在現實生活中，決策參數的變化並非固定在某特定範圍，甚至可能不為人所注意。具備這種動態特性的決策，稱為可變空間下的決策。企業創新問題即屬於可變空間下的決策問題。

企業創新本身是一個動態過程，包含了能力集合的擴展轉化，提供產品及服務以解除特定族群的痛苦與煩惱，創造價值並進一步分配價值；其中每個環節都涉及可變空間之決策制定。過去的創新研究中，缺乏整體性的架構以描述此一動態流程。本研究將以習慣領域理論為基礎，提出「創新動態學」此一動態循環架構，探討企業如何深入潛在領域，有效獲取、轉化其能力集合，滿足顧客需求，並且創造價值。文中並以知名企業為例進行驗證，期望透過個案探討方式幫助企業了解如何落實創新與創造價值，以提升企業競爭力。

關鍵詞：習慣領域、能力集合分析、可變空間下的決策、創新動態學。

I. INTRODUCTION

Human behaviors involve dynamic, evolving, interactive and adaptive processes. Important decision makings, as part of human behaviors, usually are dynamic and involving changeable parameters, too. Each decision problem may be characterized by various dimensions of parameters. It involves a number of elements such as decision alternatives, decision criteria, decision outcomes, decision preference, and decision information inputs. It also involves with the following four environmental facets : decisions as a part of the behavior mechanism, stages of the decision process, players involved, and unknowns in decision making. These parameters can interact with each other and vary with time, the situation, and changes in the psychological states of the decision makers involved. Although dynamic in nature, decision making, as a part of human behaviors, may reach a steady state and exhibit habitual patterns as time goes by. As a consequence, in mathematical programming or ordinary decision-making problems, we may unwittingly assume that the decision

parameters (or variables) have fixed known dimensions and ranges. However, in real life, the parameters might or might not be noticed. Even when they are noticed, their dimensions and ranges cannot be predetermined. Decision making with these features is called “decision making in changeable spaces” [14-16].

Corporation innovation itself, which involves setting corporate goals, evaluating states, understanding customers’ needs, producing and providing products and services, and creating value for targeted customers and themselves, is of the type of challenging decision-making problems in changeable spaces. If corporate decision makers are not aware of the existence and changing nature of the relevant parameters in decision making, they may fall into decision blinds and traps [14, 17] and make serious mistakes.

In the last two decades, the researches about innovation have covered a wide range of concepts, methodologies and application areas which lead to an abundant literature [2-8]. However, those related studies on innovation seldom explored the key factor of successful innovation from the perspective of “satisfying or releasing potential needs, pains, and

frustration". Innovation itself is a dynamic process which includes transforming a competence set for innovation, producing products or services to release the pains and frustrations of target groups, and creating and distributing values. In the field of innovation studies, there previously was no framework that systematically described these processes. This prompts us to study business innovation from the viewpoints of habitual domain and decision making in changeable spaces.

The rest of this article is organized as follows. In section II, the idea of decision making in changeable spaces is discussed. The concepts of habitual domain and competence set are introduced in section III and IV respectively. Decision blinds and traps are also discussed. In section V, the framework of Innovation Dynamics is proposed, and its verification is provided in section VI. Conclusion and further researches are summarized in section VII.

II. DECISION MAKING IN CHANGEABLE SPACES

To facilitate the presentation, let us use the following example to illustrate how the

challenge problems are solved by looking into the possible changes of the relevant parameters.

A. Example 1 : The 1984 Olympics in Los Angeles

The 1984 Summer Olympics, officially known as the Games of the XXIII Olympiad, were held in 1984 in Los Angeles, California, United States. Following the news of the massive financial losses of the 1976 Summer Olympics in Montreal, Canada, and that of 1980's Game in Moscow, USSR, few cities wished to host the Olympics. Los Angeles was selected as the host city without voting because it was the only city to bid to host the 1984 Summer Olympics.

Due to the huge financial losses of the Montreal and that of the Moscow, the Los Angeles government refused to offer any financial support to the 1984 Games. It was then the first Olympic Games that were fully financed by the private sector in the history. The organizers of the Los Angeles Olympics, Chief Executive Officer Peter Ueberroth and Chief Operating Officer Harry Usher, decided to operate the Games like a commercial product. They raised fund from corporations

and a great diversity of activities (such as the torch relay) and products (for example, “Sam the Eagle”, the symbol and mascot of the Game), and cut operating cost by utilizing volunteers. In the end, the 1984 Olympic Games produced a profit of over \$220 million.

In the above example, new players, all the potential customers to the Olympic Games besides the athletes, were brought into the decision problems. When the organizers of the Los Angeles Olympics decided to focus on introducing more potential players into the decision problems and making good use of their competence, the old solution and alternatives, such as asking financial support from the government, were dropped, and the decision problems began to move in more effective and stable direction.

Example 1 shows us that in reality, the players, criteria, alternatives, perception of rules of games, and outcomes (part of decision parameters) are not fixed. The parameters, including their dimensions, are dynamically changed depending on how deep, how far and how broad we look into the potentiality (or *potential domains* as to be introduced shortly). The fact that the parameters can themselves be

the control or decision variables is a main feature of decision making in changeable spaces (of parameters).

B. Dynamic Changes of Decision Parameters

Mathematically, decision making in changeable spaces can be described as follows:

Assume that changeable decision parameters involve the following decision elements (extension to include other parameters can be done similarly):

- (i) the alternative set at time t , denoted by X_t ;
- (ii) the criteria at time t , denoted by F_t ;
- (iii) the outcome measured in terms of the criteria at time t , denoted by \mathcal{F}_t ;
- (iv) the preference of decision maker at time t , denoted by D_t ; and
- (v) the information inputs at time t , denoted by I_t .

Each decision element is a set which can vary with time, situation, and the decision maker's perception to the decision problems. The alternative set at time $t+\Delta t$ can be denoted by

$$X_{t+\Delta t} = G(X_t, F_t, \mathcal{F}_t, D_t, I_t, HD_t) \quad (1)$$

where HD_t , consisting of actual domains (AD_t), reachable domains (RD_t), potential

domains (PD_t) and activation probability (AP_t), is the habitual domain at time t as to be described in section III. As in (1) , $X_{t+\Delta t}$ not only depends on X_t , but also on the other decision elements, $F_t, \mathcal{F}_t, D_t, I_t$ as well as HD_t .

Note that X_t and $X_{t+\Delta t}$ can be set functions, and the difference between X_t and $X_{t+\Delta t}$ would describe the changes due to time and situation. Also note that X_t and $X_{t+\Delta t}$ can have different dimensionality.

Similarly, we can write the dynamic change of other parameters as follows :

$$F_{t+\Delta t} = H (X_t, F_t, \mathcal{F}_t, D_t, I_t, HD_t) \quad (2)$$

$$\mathcal{F}_{t+\Delta t} = J (X_t, F_t, \mathcal{F}_t, D_t, I_t, HD_t) \quad (3)$$

$$D_{t+\Delta t} = K (X_t, F_t, \mathcal{F}_t, D_t, I_t, HD_t) \quad (4)$$

$$I_{t+\Delta t} = L (X_t, F_t, \mathcal{F}_t, D_t, I_t, HD_t) \quad (5)$$

Note, (1) – (5) describe the fact that the decision elements (or parameters) not only vary with time, but also mutually interact with each other through time. For further discussion see Ch 7-8 of [11, 12].

III. HABITUAL DOMAINS

The collection of ideas and operators (including ways of perceiving, thinking, responding, acting, and memory) in our brain, together with their formation and dynamics, is

called our habitual domains (HDs). Over time, unless extraordinary events occur or purposeful effort is exerted, our HDs will become stabilized within a certain domain. This phenomenon can be mathematically proved [1, 9].

Being aware of the habitual ways of our decision making is important for us to make better decisions and avoid costly mistakes. To better understand the concept of HD, let us briefly introduce the elements of HD, which are important parameters in the human behavioral systems.

HD at time t , HD_t , include the following four sub-concepts :

- (i) *Potential domain*, designated by PD_t , is the collection of all ideas and operators which can be potentially activated with respect to specific events or problems by one person or by one organization at time t . In general, the larger the PD_t , the more likely that a larger set of ideas and operators will be activated, holding all other things equal.
- (ii) *Actual domain*, designated by AD_t , is the collection of ideas and operators which are actually activated in our minds at time

t. Note that not all the ideas and operators in the potential domain can be actually activated. Also note that the AD_t is a subset of the PD_t , that is $AD_t \subset PD_t$.

(iii) *Activation probability*, designated by AP_t , is defined for each subset of PD_t and is the probability that a subset of PD_t is actually activated or is in AD_t . For example, people who emphasize profit may be more likely to activate the idea of money, while people who study mathematics may be more likely to generate equations.

(iv) *Reachable domain*, designated by RD_t , is the collection of ideas and operators which can be generated from a given set in an AD_t . In general, the larger the idea set and/or operator set in AD_t , the larger the RD_t .

At any point in time, without specification, HD_t is the collection of the above four subsets. That is,

$$HD_t = \{PD_t, AD_t, AP_t, RD_t\}$$

When there is no confusion, the subscript “*t*” may be dropped as to simplify the presentation. In general, the AD is only a small portion of the RD, while the RD is only a small

portion of the PD, and only a small portion of the AD is observable. This makes it very difficult for us to observe other people's HDs and/or even our own HDs. For more details, see [11-13].

IV. COMPETENCE SETS AND DECISION BLINDS

The study on competence set analysis began with Yu [10], as a derivative of HD theory. The competence set (CS) for a given decision problem is defined as a collection of ideas, knowledge, skills and resources for its effective solution. Such a set, like HD, implicitly contains PD, AD, RD, and AP as discussed in section III. When the decision maker thinks he/she has already acquired and mastered the CS as perceived, he/she would feel comfortable making the decision and/or undertaking the challenge. Anything or anyone, including a product or service that can release the pain, frustration and charge, has competence. Everyone, and every corporation, has its competence sets.

To analyze the competence sets of individuals or corporations, we can decompose the CS as follow :

$$CS_t = (CS_t^1, CS_t^2, CS_t^3, \dots, CS_t^n) \quad (6)$$

where CS_t^k , denotes the k th item of the CS at time t . Note that CS will be dynamically changed as time (t) goes by.

Because of HDs and being unaware of the decision parameters and their changing nature, people would easily have decision blinds or even get into decision traps. Let us denote the *truly needed* CS for solving problem E successfully by $CS_t(E)$ at time t , and its perception by decision makers, by $CS_t^*(E)$. Then $CS_t(E) \setminus CS_t^*(E)$ would be the *decision blinds*, the set of all the competences required but not seen by the decision makers at time t . See the illustration of Figure 1. Note that the larger the decision blind is, the more likely decision makers might make dangerous mistakes.

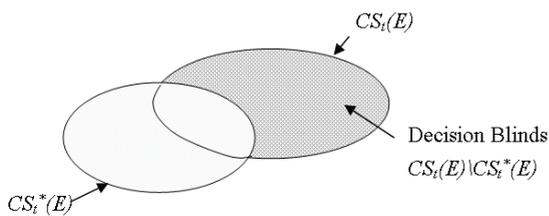


Figure 1. Decision blinds

Usually, $CS_t(E)$ and $CS_t^*(E)$ can be changed with time. Suppose that $CS_t^*(E)$ is fixed or trapped in a certain domain and CS_t

$(E) \setminus CS_t^*(E)$ is large, then we tend to make mistake in decision and we are in a decision trap. Decision trap (i.e. $CS_t^*(E)$ is fixed, independent of t) can lead to dangerous mistake, especially when $CS_t(E)$ changes rapidly with time and $CS_t(E) \setminus CS_t^*(E)$ becomes very large.

By changing our actual domains (ADs), we can change and expand our reachable domains (RDs). We can reduce decision blinds and/or avoid decision traps by systematically changing the ADs. For illustration, assume that $CS(E)$ and RDs are given, as depicted in Figure 2. Then as we move the AD from A to B , then to C , our decision blind reduces progressively from $CS(E) \setminus RD(A)$ to $CS(E) \setminus (RD(A) \cup RD(B))$ then $CS(E) \setminus (RD(A) \cup RD(B) \cup RD(C))$.

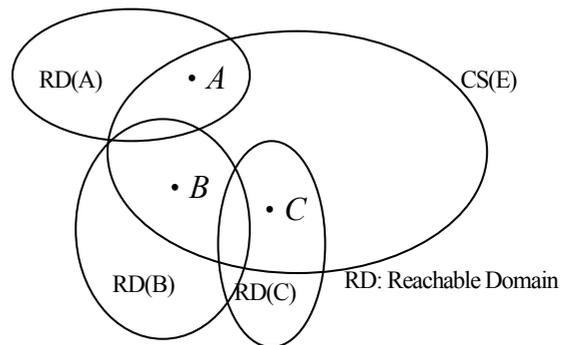


Figure 2. Decision blind reduces as we move our AD from A to B then to C

For challenging decision problem, we can treat the decision parameters as different points for ADs. Systematically moving over the parameters and pondering their possible RDs can expand our RDs for dealing with the challenging problems. As a consequence, CS^* (E) is expanded and our decision blinds, $CS(E) \setminus CS^*(E)$, reduced. In additions, the HD tools [11-13] can work with the individual decision parameters as to reduce the decision blinds and avoid decision traps. They can expand and enrich our ADs and RDs and look into the depth of the PDs, they can also expand and enrich our perception on the decision problem and its related parameters.

V. INNOVATION DYNAMICS

According to the HD theory and CS analysis, all humans and things can release pains and frustrations for certain group of people at certain situations and time. Thus, all humans and things carry the competence (in broad sense, including skills, resources, functionalities, even attitudes). If we regard all humans and things as a set of different CSs, then producing new products or services can be

regarded as a transformation of the existent CS to a new form of CS. Based on this, we could depict a comprehensive and integrated framework, called the *Innovation Dynamics* (see Figure 3), to help people understand corporate innovation and creation of maximal values for the targeted customers and themselves.

The dynamics can be interpreted clock-wise, according to the indices of Figure 3, as follows :

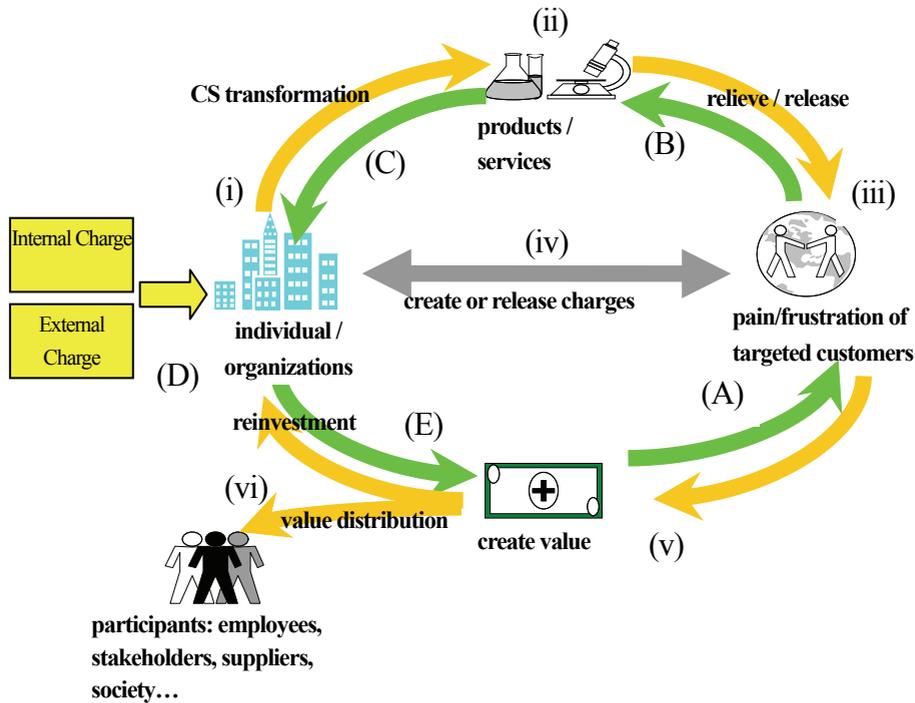


Figure 3. Innovation Dynamics

(i) According to HD Theory, when there exists unfavorable discrepancies between the current states and the ideal goals of individuals or organizations (for instance, the corporations are losing money instead of making money, or they are technologically behind, instead of ahead of the competitors), it will create charges which can prompt the individuals or corporations to work harder to reach their ideal goals.

(ii) The transformation of CSs will be

presented in visible or invisible ways, which results in a new set of the products or services produced by the corporations.

(iii) The products or services produced by corporations must carry the capability to relieve / release the pain and frustration of targeted customers. Note that there are actual domains, reachable domains, and potential domains for the targeted customers, and for their pains, frustrations, and problems.

(iv) Besides discharge, corporations or

organizations can create charges to the targeted customers by means of marketing, advertisement or promotion, and vice versa.

- (v) The targeted customers will experience the change of charges. When their pains and frustrations are relieved, the customers become happy. By their buying the products or services, the products and services create their value.
- (vi) The value will be distributed to the participants such as employees, stakeholders, suppliers, society, etc. In addition, to gain the competitive edge, products and services have to be continuously upgraded and improved. The reinvestment therefore is needed in order to develop and produce new products and services.

In contrast to the clockwise cycle, the Innovation Dynamics can be interpreted counter-clockwise, according to the indexing of Figure 3, as follows :

- (A) To create values, the corporations must consider who will be the targeted customers, and what kind of pain and frustration they have, both in actual and

potential domains.

- (B) In order to ease the pains and frustrations for the targeted customers, what products or services, in actual and potential domains, are needed? Competitiveness becomes an important issue in the final selection of the products and services to produce.
- (C) How do the corporations transform their internal and external competence and resource to develop or provide the selected products and services effectively and efficiently?
- (D) When the transformation of CSs succeeds, the corporation's internal and external charge will be released, at least partially.
- (E) New goals as to create new values can be reestablished. The innovation cycle : (A) → (B) → (C) → (D) → (E) → (A) will go round and round.

The concept of Innovation Dynamics describes the dynamics of how to solve a set of problems with our existent or acquired competence (to relieve the pains or frustrations of targeted customers or decision makers at certain situations) as to create value, and how to distribute this created value so that we can continuously expand and enrich the CS to solve

more challenging problems and create more value. Observe that each links, clockwise or counterclockwise, in Figure 3 involves decision makings in changeable spaces.

VI. VERIFICATION

Let us consider the following two cases to verify the framework of Innovation Dynamics.

A. *Case I : YouTube*

At the end of 2006, a 1.5 year-old video sharing company, YouTube, was merged by Google, the well-known search engine, at the price of \$1.65 billion. This high-profile event was the biggest merger case for Google at that time. It drew a great deal of discussion and attention.

The three YouTube founders originally intended to provide their product to eBay as a way to introduce auction products, in addition to pictures and text descriptions. The idea appeared to meet a demand, but never succeeded. They then extended their product to the PDs of potential users, and found there are a large number of people with a strong desire to express themselves to online friends. These desires in PDs were not discovered until

YouTube dropped their original idea and explored the real need in the PDs of the potential customers.

From the initial garage venture to a website worth over US\$1.6 billion, the growth and development of YouTube is a process of continuous integrating and transforming CSs. In the process, they released people's potential pain and frustration by providing effective product and service that others could not, or would not. By doing so, they enhanced their corporate competitiveness and used it to create value.

B. *Case II : Wii*

Nintendo started working in the game console industry in the 1970s, when there were not many design alternations for game consoles. Players must operate the gamepad with two hands, and they could only use their thumbs to control movement. Until 2006, Satoru Iwata, who had been the president of Nintendo for less than five years, led Nintendo to break the three-decade old design. Wii, with the simple creativity of "operate with one hand", was born. The appearance of Wii has created a new generation of games. It brings a new

entertainment experience, and the innovative interface of game control has successfully reduced the time needed for new players to learn how to play a game. The remote control is equipped with sound effects, vibrates and has orientation functionality. It allows players to simulate the behavior of real games and brings users an unprecedented gaming experience.

Nintendo had lost its leading position in the gaming industry before Wii entered into the market, which caused the company market share to fall behind. The Wii is Nintendo's innovation breakthrough. Its innovation is not only in "subverting traditional design", but more importantly, it satisfied the desire of people "wanting to experience realistic gaming" in PDs. In the past, the game console industry was always committed to pursuing exquisite graphics, sound and light effects, attempting to satisfy the desire of player. However, luxurious graphics, sound and light effects are the needs in the ADs for gamers (desires of audio and visual aspects) . Allowing body movements and feeling the speed, direction or even power with the game are the strong needs hidden in the players' PDs. By satisfying the needs in PDs, Wii recreated

the interaction between gamers and games, which not only creates value for Nintendo, but also allows the company to regain its competitive advantage.

VII. CONCLUDING REMARKS

In this article, the concepts of decision making in changeable spaces, HDs, CS analysis and Innovation Dynamics and its verification are introduced. The activities over each link of the Innovation Dynamics involve decision making in changeable spaces. By examining the operations of each link in Innovation Dynamics, corporations can understand if each and all links are properly developed, so that they can continuously upgrade their products or services and maximally create value by releasing pains and frustrations for the customers in the PDs. The Innovation Dynamics can help them to be as successful and competitive as YouTube and Nintendo.

Many research problems remain open. For instances, how to systematically analyze the invisible potential domains as to find effective method to acquire, adjust and allocate resources

in potential domains; and how to effectively transform the hidden resources in potential domains into products/services to effectively release the pains and frustrations of the people concerned. Mathematical analysis for specific cases would be of great interest to study. These would bring value to practical business innovation and to academic research as well.

REFERENCES

- [1] S. J. Chan and P. L. Yu, "Stable habitual domains : existence and implications", *Journal of Mathematical Analysis and Applications*, vol. 110, pp. 469-482, 1985.
- [2] W. M. Cohen and D. A. Levinthal, "Absorptive capacity : a new perspective on learning and innovation", *Administrative Science Quarterly*, vol. 35, no. 1, pp. 128-152, 1990.
- [3] D. Edquist, *Systems of innovation. Technologies, institutions and organizations*, London : Pinter, 1997.
- [4] J. Hauser, G. J. Tellis, and A. Griffin, "Research on innovation : a review and agenda for marketing science", *Marketing Science*, vol. 25, no. 6, pp. 687-717, 2006.
- [5] D. Leonard-Barton, *Wellsprings of Knowledge : Building and Sustaining the Sources of Innovation*, Harvard Business School Press, 1995.
- [6] R. Mcadam, "Knowledge management as a catalyst for innovation within organizations : a qualitative study", *Knowledge and Process Management*, vol. 7, no. 4, pp. 233-241, 2000.
- [7] S. Shapiro, *24/7 innovation : a blueprint for surviving and thriving in an age of change*, M. McGraw-Hill, 2001.
- [8] J. Tidd, J. Bessant, and K. Pavitt, *Managing Innovation*, John Wiley & Sons Ltd, Chichester, 2005.
- [9] P. L. Yu, *Multiple Criteria Decision Making : Concepts, Techniques and Extensions*, Plenum Press, New York, 1985.
- [10] P. L. Yu, *Understanding Behaviors and Forming Winning Strategies*, Monograph, School of Business, University of Kansas, 1989.
- [11] P. L. Yu, *Forming Winning Strategies, an Integrated Theory of Habitual Domains*. Berlin, Heidelberg, New York, London, Paris, Tokyo : Springer-Verlag, 1990.
- [12] P. L. Yu, *Habitual Domains and Forming*

- Winning Strategies*. Taiwan : NCTU Press, 2002.
- [13] P. L. Yu, *HD : Habitual Domains, the Human Softwares That Determine Your Life*, Taipei, Taiwan : Reading Times, 2009
- [14] P. L. Yu and C. Y. Chiang-Lin, “Decision traps and competence dynamics in changeable spaces”, *International Journal of Information Technology and Decision Making*, vol. 5, no. 1, pp. 5–18, 2006.
- [15] P. L. Yu and Y. C. Chen, “Dynamic MCDM, Habitual Domains and Competence Set Analysis for Effective Decision Making in Changeable Spaces,” in : *Trends in Multiple Criteria Decision Analysis, International Series in Operations Research & Management Science, Volume 142*, M. Ehrgott *et al.* Eds, Springer, 2010, ch 1, pp. 1–35.
- [16] P. L. Yu and Y. C. Chen, “Dynamic multiple criteria decision making in changeable spaces : from habitual domains to innovation dynamics”, *Annals of Operations Research*, in press, doi : 10.1007/s10479-010-0750-x.
- [17] P. L. Yu and Y. C. Chen, “Blinds, fuzziness and habitual domain tools in decision making with changeable spaces”, *Human Systems Management*, in press.

修平技術學院《修平學報》徵稿要點

- 一、本刊為純學術性之刊物，專供本校同仁及校外人士發表研究成果及論著之用。
 - 二、論述及研究報告文字（含圖表），以 20 頁為度，來稿須以電腦排版，以 A4 紙列印，請交文稿電腦磁片及紙本各一式三份（論文格式及版面規格，請依照出版組提供之說明資料或上網下載）。
 - 三、無論中文或英文稿件皆須附上中英文題目、摘要，並註明作者姓名及科系名稱。他國文字稿件須附中文題目摘要，其字數以 500 字為度；並應列舉中、英文或他國文字之關鍵詞（keywords）。
 - 四、文稿之審查依據「修平技術學院學報評審辦法」。
 - 五、稿件格式按各專業學門標準格式或參考「修平學報論文格式」。
 - 六、作者投稿後，若在作業程序中因故取消投稿者，則其後兩期不再接受其投稿。
 - 七、所投稿件經編審委員審查同意刊登，稿件經刊印後，不得在他處刊印發表。如果已在其他刊物正式公開發表後，轉投本刊物，本委員會不負查核之責，相關著作權問題，由當事人自行負責。稿件若涉及一稿兩投或抄襲者，本學報得拒絕作者稿件五年。
 - 八、經審查採用之文章，排版後送請作者校稿，作者僅能修正排版印刷之錯誤，且不得擅自於校稿過程中增減內容。
 - 九、本刊每期以刊登二十篇論文為原則，經審查後決議可刊登者，如超過篇數，則按最後定稿時間先後排序，安排至下一期刊登。
 - 十、本刊文之作者應對論文之內容及同意發表權之取得，負全部之責任。並請於投稿時即將「修平學報投稿授權聲明書」（附件）填妥後一併寄交。
-
-

-
-
- 十一、本刊文之作者應對論文之內容及同意發表權之取得，負全部之責任。若著作人投稿於本刊經收錄後，同意授權本刊得再授權國家圖書館或其他資料庫業者，進行重製、透過網路提供服務、授權用戶下載、列印、瀏覽等行為。並請於接受刊登時即將「著作授權同意書」(附件)填妥後一併寄交。
 - 十二、來稿經採用者，送當期學報一份及其著作抽印本二十份，作者如需增印，增印之成本由作者自付。
 - 十三、賜稿請送交圖書館出版組。
 - 十四、本要點經學報編審委員會議通過，陳請校長核定後公布實施，修正時亦同。
-
-

修平技術學院學報評審辦法

第一條 來稿之評審係由學報編審委員會遴聘校外相關領域之專家學者擔任。

第二條 由執行編輯（出版組）收稿、登錄及分類後，交由學術副校長挑選兩名校外專業人士或學者進行評審。

第三條 每篇稿件原則上由兩位評審，每位評審除於評審意見表上陳述意見外，並需對稿件作出下述三項之一建議：

- 一、接受刊登。
- 二、修正後再審。
- 三、不予刊登。

第四條 依據前述審查意見，處理方式如下表：

處理方式		第 二 位 評 審 意 見		
		接受刊登	修正後再審	不予刊登
第 一 位 評 審 意 見	接受刊登	刊登	寄回修改	* 第三位評審
	修正後再審	寄回修改	寄回修改	* 第三位評審
	不予刊登	* 第三位評審	* 第三位評審	退稿

*1.若第三位評審意見為「接受刊登」或「修正後再審」時，則請作者對不予接受之審查意見進行答覆外，將採兩正方評審意見予以刊登。

*2.若第三位評審意見為「不予刊登」時，將採兩負方評審意見予以退稿。

第五條 本刊將針對審查意見及結果函送投稿人，並說明處理方式。

第六條 評審作業相關人員，對評審委員身份應予以保密，以避免紛爭。

第七條 投稿人不得有打聽及干涉評審委員之言行。

第八條 本辦法經學報編審委員會議通過，陳請校長核定後公布實施，修正時亦同。

修平學報

中華民國一〇〇年三月出版

發行人 鍾瑞國

出版者 修平技術學院圖書館出版組

地 址 41280 臺中市大里區工業路十一號

電 話 04-24961100

傳 真 04-24961187

編輯者 修平學報編審委員會

召 集 人—陳培中

編審委員—方世榮 江可達 林婉芳 陳清祺

張志凌 鄧作樑 盧志偉

(依姓氏筆劃排序)

執行編輯—陳清祺

印刷者 天空數位圖書有限公司

地 址 40255 台中市南區忠明南路 787 號 30 樓

電 話 04-22623893

傳 真 04-22623863

版權所有 請勿翻印