

# 行政院國家科學委員會專題研究計畫 成果報告

## 以繪畫內容為主的影像搜尋系統

計畫類別：個別型計畫

計畫編號：NSC92-2213-E-032-022-

執行期間：92年08月01日至93年07月31日

執行單位：淡江大學資訊工程學系

計畫主持人：顏淑惠

共同主持人：林慧珍

計畫參與人員：王佳仁

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 11 月 2 日

行政院國家科學委員會補助專題研究計畫  成果報告  
 期中進度報告

以繪畫內容為主的影像搜尋系統

計畫類別： 個別型計畫  整合型計畫

計畫編號：NSC 92-2213-E-032-022

執行期間：92年8月1日至93年7月31日

計畫主持人：顏淑惠

成果報告類型(依經費核定清單規定繳交)： 精簡報告  完整報告

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、  
列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：淡江大學 資訊工程系

中 華 民 國 93 年 10 月 15 日

# 行政院國家科學委員會專題研究計畫成果報告

## 以繪畫內容為主的影像搜尋系統

計畫編號：NSC 92-2213-E-032-022

執行期間：92年8月1日至93年7月31日

計畫主持人：顏淑惠 淡江資訊工程系

### 一、中英文摘要及關鍵詞：

#### 中文摘要

早期日據時代每年一度的台灣美術展覽是當時台灣的繪畫史上一大盛事，也是台灣當地畫家開始接觸西洋畫並進入現代畫的一個關鍵時期。因為時間的久遠、戰火的摧殘以及保存技術，使得這些寶貴的畫作已經幾乎都遺失殆盡，所能取得僅有當時出版黑白印製的展覽圖錄。目前中央研究院也已經著手建立數位影像資料查詢系統（數位典藏系統 [http://dialog.iis.sinica.edu.tw/Content/MDC\\_3/Obj\\_8/query.asp](http://dialog.iis.sinica.edu.tw/Content/MDC_3/Obj_8/query.asp)）。目前此系統由使用者輸入台灣美術展覽會的屆數，系統則輸出當屆所有畫作名稱，以供使用者選取感到興趣的畫作。未來的計畫是加上以畫家姓名作為搜尋的工具。而我們的計畫則是希望直接由畫作內容為搜尋的方式，讓這些寶貴的歷史財產可以更為一般民眾所接近。

本計畫提出一個以低階的影像特徵來區分肖像畫與其他的畫（如風景、靜物、等等）。肖像畫的定義是一張繪畫以一個人上半身（有時會有到兩、三個人），或全身以站或坐為主題，臉或姿勢的表現常常是比背景來得重要。因此若能確定在畫作內有人臉，再來分辨是否為肖像畫便是一個簡單的問題。因此計畫中提出一個根據人臉特徵，透過 SVM 分類器對繪畫內容做搜尋之人臉影像偵測系統，利用此系統能查詢出影像中人臉所在的區域，並縮小在一可能範圍，如此不但能減少錯誤偵測，亦能快速的檢查出此張影像是否可能為肖像畫。

**關鍵詞：**繪畫搜尋系統, SVM (support vector machine), 邊緣偵測, 模版比對, 人臉偵測

#### Abstract

The Annual Taiwan Painting Exhibition was a very important event within the painting society dating back to seventy years ago. It was also during that time when Taiwanese was first exposed to western painting techniques, making it an important era in the art history of Taiwan. Due to various events, including World War II, much of the paintings from that era have been lost. What remains are gray-toned printed catalogs from painting exhibitions. Restoration work is being

conducted beginning with the digitization of the paintings. However, the system currently in place can only categorize paintings by names, thus it is not friendly to users who are not familiar with the paintings. In our project, we try to develop a content-based image categorization for a user-friendly painting query system.

A portrait is a painting mainly focuses on the face. Thus, if a face is detected on a painting, then it is much easier to determine whether it is a portrait. We propose a method to detect whether there is a face on the painting and thus to distinguish paintings into portraits and others (landscapes, still life, etc.). An SVM-based classifier is first applied to locate possible face candidates. A face candidate then is matched to a face template for further confirmation.

In this research, we focus our attention on edge and shape features of painting images as described in the following: (1). Edge features: A variety of edge features have been exploited. All of these features have proved to have advantages and disadvantages. In this project, we will make an attempt to collect and overview on SVM-based approach. (2). Shape features: Templates are used as models in a variety of faces.

**Keywords :** painting query system , SVM (support vector machine), edge detection, template matching, face detection

## 二、報告內容：

### (一)、前提介紹

因為近代電腦科技與網際網路的快速發展，越來越多的使用者透過網際網路來搜尋他們有興趣的資料。但是網際網路就像是一個超大型的資料庫，含有各式各樣的資料型態，如：聲音、影像、視訊…等。如何能從一個龐大的資料庫系統中快速且正確地找出使用者想要的資料，與增加搜尋速度與正確性將是重大課題。

近代有許多人臉偵測的方法被提出來。Zhao *et al.* [1] 利用類神經網路的方式來偵測臉部。Baek *et al.* [3] 提出了 partial least squares (PLS) 方法，將特徵向量的維度有效地縮減，將問題縮小化的新技術，以進行臉部辨識的工作。Czyz *et al.* [2] 則是結合了數種已知的分類器，將它們加以組合並使他們的輸出達到最佳化。但這樣的最佳化工作是有條件性的，例如各個分類器之間是相互獨立的。Li *et al.* [4] 提出了利用 SVM 來進行各種臉部姿勢的偵測與識別。

而其他重要的研究課題則是有關於 Support Vector Machines (SVMs)。Burges [5] 提出了一篇有關於 SVM 整體說明的文章“A Tutorial on Support Vector Machines for Pattern Recognition”，在文章中詳細地介紹了有關於 SVM 的基本概念、理論的介紹，是一篇想初學 SVM 的入門文章。Platt [6] 則是提出了一個快速訓練 SVM 的方法“Sequential Minimal Optimization” 簡稱 SMO。SMO 將困難且運算複雜的二次規劃的問題簡化成單純的小問題。經由 SMO 演算法能夠快速地求得近似的最佳解，進而減少運算時間。

### (二)、Proposed Method

#### ● 前處理

邊緣偵測是電腦視覺與影像處理的一項重要前處理工作。能夠正確與有效地將影像的邊緣萃取出來，對於後續處理的有效性將有大幅度的提升。而前處理的工作通常包含了許多步驟，不是單一步驟可以完成，因此，在本小節中將介紹前處理的方法與步驟說明。

##### i. 模糊化

為去除雜訊並將不必要的細微資訊去除，第一步就是將影像模糊化。一般來說模糊化有下列方法(1)中間值濾波：中間值濾波能夠濾除雜訊，同時其最大的優點是能同時保持邊緣的資訊不被模糊，這對於我們的方法而言非常重要。(2)高斯濾波：這個方法的像素值計算能用圖 1 的遮罩來表示，高斯濾波的效果很類似於低通濾波器卻又可以大致保持邊緣資訊。圖 2~4 表現出原圖及濾波之後的效果。

	1	2	1
1/16	2	4	2
	1	2	1

圖 1:高斯模糊遮罩( $\sigma^2 = 0.5$ )



圖 2:原始圖形



圖 3:中間值濾波結果(3x3)



圖 4:高斯濾波( $\sigma^2 = 0.5$ )

### ii. 邊緣偵測：(Sobel Operation)

一般來說，邊緣偵測有多種方法，而最常用的方法莫過於 Sobel method。Sobel 遮罩計算像素所在位置的導數，也就是梯度。對於像素位置  $(x,y)$  梯度值的運算可以由下列運算式定義：

$$\nabla f = \left[ \frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right]^T \quad (1)$$

梯度的反應強度可以由下列運算式來逼近：

$$\begin{aligned} \|\nabla f\| \approx & |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| \\ & + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)| \\ & + |(z_6 + 2z_8 + z_9) - (z_1 + 2z_2 + z_4)| \\ & + |(z_4 + 2z_7 + z_8) - (z_2 + 2z_3 + z_6)| \end{aligned} \quad (2)$$

利用公式 (2) 將梯度的反應強度計算出來，接下來就是要對梯度的反應強度進行過濾的動作，至於門檻值該要如何設定呢？在這裡我們提出一個能夠動態適應所有圖像的方法，說明如下：大於門檻值的點數總和需與圖像的面積呈一固定比例，在本方法中，所設定的比例是 10%，也就是大於門檻值的點數總合約為圖像面積的 10%。將小於門檻值的位置清除為 0，其餘保持其梯度值。經過門檻值過濾後的梯度影像保留了影像的主體結構，能提供後續的判別工作，如圖 5、6 所示。

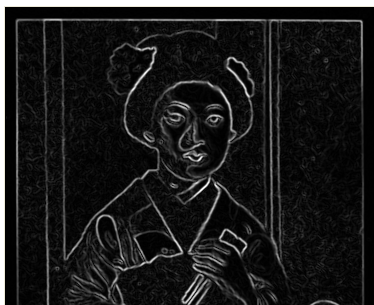


圖 5:邊緣偵測所得影像



圖 6:經過門檻值過濾後影像

### iii. 正規化

在我們實際將邊緣偵測後的影像做後續的判別時，一個很重要的步驟那就是需要做正規化。因為不管圖像的大小，拍攝影像的攝影師或是繪畫的作者，基於創作的意念所建構出來的主體大小完全不相同，因此必須將欲判別的區域正規化成固定大小，在本方法中所使用的是 bilinear 的方法，透過此種方法將任何的影像調整成固定大小，在

我們的方法中，經過正規化後的影像為 25x25。

● 分類器的說明與使用

在各種分類器中最常為大家所使用的就是類神經網路(NN)，不過類神經網路的訓練過程使屬於黑箱作業，無法預期訓練的結果會如何，在加上訓練完成後的分類器的類推效果沒有經過最佳化，因此在判別新樣本時不能保證其類推的結果是最好的，所以在本方法中並不使用類神經網路。在本計畫中所使用的分類器是 Support Vector Machines (SVMs)。因為 SVM 在訓練的過程中，會將空間的維度透過 Kernel Function 提高，並在較高的維度底下找尋一超平面，將訓練的類別分割開來，而且此一超平面會公平地分割兩塊區域，此公平地分割將有助於將來類推新樣本時的正確性。有關 SVMs 的說明如下：

i. Linear Support Vector Machines (SVMs)

線性可分割的 SVMs 能輕易的將訓練資料線性分割成兩類，假設我們將訓練用樣本標示成下列表示方式：

$$\{x_i, y_i\}, i = 1, \dots, l, y_i \in \{-1, 1\}, x_i \in \mathbf{R}^d$$

假設存在一些超平面能將上述資料分割開來，且若有資料點落在這些超平面上則需滿足下列條件：

$$w \cdot x + b = 0$$

w 是這些超平面的法線，並假設所有的訓練樣本均滿足下列條件：

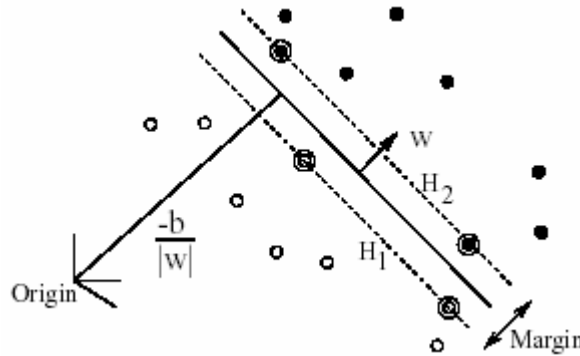


圖 7:線性可分割

$$x_i \cdot w + b \geq +1 \text{ for } y_i = +1 \tag{3}$$

$$x_i \cdot w + b \leq -1 \text{ for } y_i = -1 \tag{4}$$

將上述兩式合併可以得到下列方程式：

$$y_i(x_i \cdot w + b) - 1 \geq 0 \quad \forall i \tag{5}$$

接下來針對(5)式利用一些動態規劃的演算法就能夠求得最佳解，所求得之最佳解如圖 7 說明，也就是希望將 Margin 最大化。不過利用動態規劃來求得最佳解的方式會花費太多運算時間，在本計畫中我們使用 Platt [6] 所提出的“Fast Training of Support Vector Machines using Sequential Minimal Optimization”也就是“SMO”的方法，此方法

能在一可容許的時間範圍內求出近似的最佳解（如圖 8 說明）。

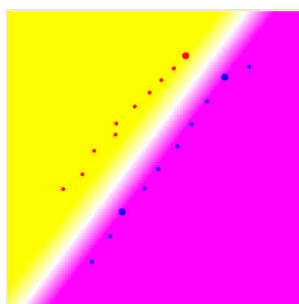


圖 8: 線性可分割的範例說明

### ii. Non-Linear Support Vector Machines

如經過上述方法仍無法將訓練資料成功地分割，那就是屬於非線性可分割的例子，因此就需要 Kernel Function 介入，Kernel Function 最主要的目的是要將原空間的維度提高，將空間的維度提高之後就能在高維度的空間中尋找一超平面，此超平面能順利將訓練樣本在該空間中做線性分割，如圖 9、10 所示。

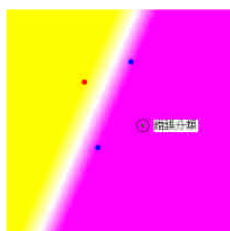


圖 9: 線性不可分割

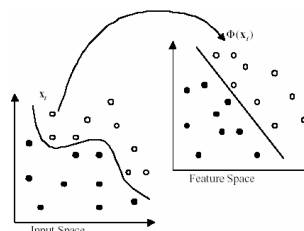


圖 10: 空間轉換

而目前所常用的 Kernel Function 有下列 3 種：

$$K(x, y) = (x \cdot y + 1)^P \quad (6)$$

$$K(x, y) = \exp\left(\frac{-1}{2\sigma^2} \|x - y\|^2\right) \quad (7)$$

$$K(x, y) = \tanh(\beta_0 \langle x_i \cdot x \rangle + \beta_1) \quad (8)$$

透過 Kernel Function 的空間轉換就能夠將線性不可分割的問題加以解決。如圖 11 所示：

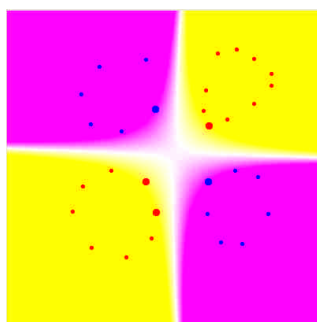


圖 11 利用 Kernel Function 解決線性不可分割的問題

### iii. 利用邊緣資訊訓練 SVM

將選定好的人臉及非人臉影像，依據上述邊緣偵測及門檻值選取方式進行前處理，再



將處理過後影像進行正規化，使其達到 25x25 像素，之後將像素當作特徵值以進行 SVM 的訓練，因此每一個用訓練樣本的維度將達到 625 維，而所使用得 Kernel Function 是方程式(7)。下列是用來訓練 SVM 的人臉及非人臉的部分樣本：

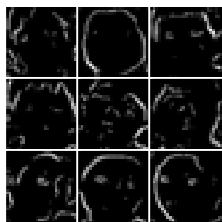


圖 12 人臉樣本

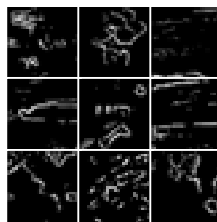


圖 13 非人臉樣本

在本方法中，人臉樣本共 71 個，非人臉樣本共 100 個，訓練過後的 SVM 可用來判別影像中人臉的位置。

### (三)、實驗與結果

本實驗結果包含幾個部分：(1)邊緣偵測及門檻值過濾後結果。(2)SVM 判別臉部位置。

#### i. 邊緣偵測

在本實驗中利用中間值濾波、高斯濾波與 Sobel 遮罩來進行前處理的工作，同時門檻值的選取為影像面積的 10%，圖 14、15 為濾波前後的結果。



圖 14 原始影像



圖 15 高斯濾波後影像

圖 16、17 為經過 Sobel 邊緣偵測與門檻值濾除之後的影像說明。



圖 16:邊緣偵測所得影像



圖 17 經過門檻值濾除後影像

#### ii. 人臉偵測

利用訓練過後的 SVM 對前處理後的影像進行搜尋，搜尋的範圍由下列條件所限制：

- (1)人臉最小為影像的 1/10，最大為影像的 1/2。
- (2)每次橫向移動為影像寬度的 1/100。
- (3)每次縱向移動為影像高度的 1/100。
- (4)搜尋範圍每次增加影像寬度的 1/10。

將所選取的範圍經過正規化後再經由 SVM 判別，若判別為臉部則將選取範圍圈選出來，如圖 18、19 所示：



圖 18: 臉部偵測結果

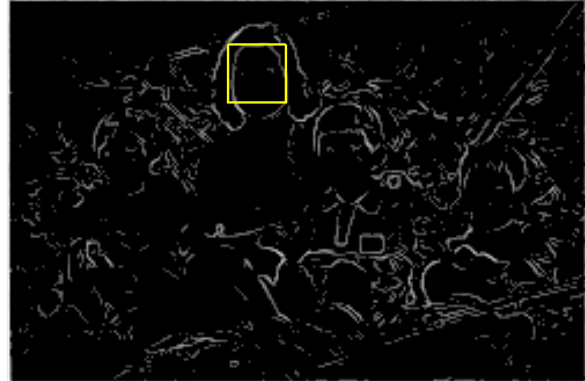


圖 19 臉部: 偵測結果

#### (四)、討論與未來方向

在本方法中提出一個新的人臉偵測方法，此方法係利用 SVM 做為臉部判別的依據，會選用 SVM 的原因是因為它經過訓練之後的結果是最佳化(圖 20)，而一般常用的類神經網路所訓練之後的結果只是針對訓練樣本的正确性，無法保證將來在進行類推時的能力(圖 21)。

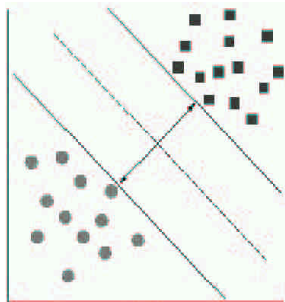


圖 20: SVM 經過最佳化之後結果

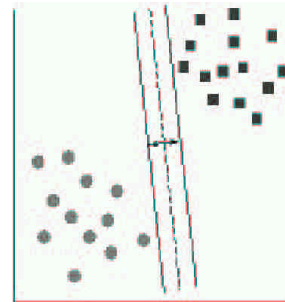


圖 21: 類神經網路未經過最佳化

利用最佳化的分類結果來進行新樣本的類推工作，較能得到正確的判別，也就是提高了正確性。雖然藉由 Kernel Function 會將空間的維度提高，但因為在提高維度的同時一併將內積運算完成，並不會增加運算的複雜度。因此我們所提出的方法不但滿足正確性，同時增加有效性。

在本方法中所提出利用 SVM 分類器來作為判別人臉的工具，雖然在訓練的過程中能保證其結果的最佳化，也能提升測試新樣本時的類推能力而增加判別的正确性，但是面對各式各樣條件不同的影像，難免還是會發生判別錯誤的情況，例如：形狀類似人類頭部的圓形物件、甚至是中間部位鏤空的部位都會被分類器判別為人臉而發生判別錯誤(圖 22)。未來的工作將針對判別錯誤的地方再加以分析，並且再對 SVM 加

強回饋的功能，能讓使用者將判別錯誤的位置回饋給 SVM 當作錯誤樣本，一步一步加強其判別的功能，期能達到最佳的判別正確性。

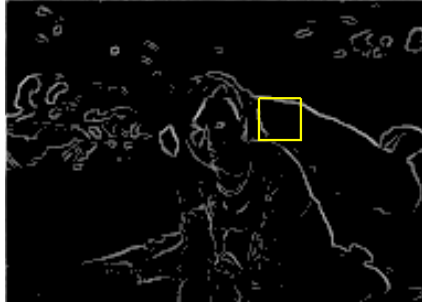


圖 22: 判別錯誤的例子

### 三、參考文獻

1. Z.-Q. Zhao, D.-S. Huang, B.-Y. Sun, "Human face recognition based on multi-features using neural networks committee" Pattern Recognition Letters 25 (2004) 1351 - 1358.
2. J. Czyz., J. Kittler, L. Vandendorpe, "Multiple classifier combination for face-based identity verification" Pattern Recognition 37 (2004) 1459 - 1469.
3. J. Baek., M. Kim, "Face recognition using partial least squares components" Pattern Recognition 37 (2004) 1303 - 1306.
4. Y. Li, S. Gong, J. Sherrah, H. Liddell, "Support vector machine based multi-view face detection and recognition" Image and Vision Computing 22 (2004) 413-427
5. C. J.C. BURGESS, "A Tutorial on Support Vector Machines for Pattern Recognition" Data Mining and Knowledge Discovery, 2, (1998) 121-167.
6. John C. Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization" Microsoft Research.
7. S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, K. R. K. Murthy, "Improvements to Platt's SMO Algorithm for SVM Classifier design" Dept. of Mechanical and Production Engineering National University of Singapore

### 四、計畫成果自評

本計畫執行成果所提之利用 SVM 來進行人臉偵測方法，能有效且正確的找尋影像中人臉部的的位置所在，並解決最佳類推能力問題，執行成果顯示與當初所提之原計畫內容大致相符，也達到預期目標。