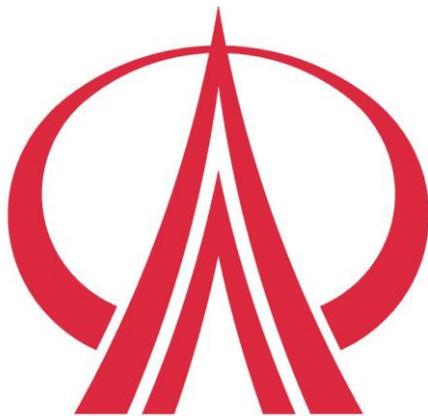


修平科技大學 電機工程系

DEPARTMENT OF ELECTRICAL ENGINEERING
HSIUPING UNIVERSITY OF SCIENCE AND TECHNOLOGY

實務專題報告書

語音辨識應用紅外線於家電控制



指導老師：魏嘉延

專題製作學生：四技電機四甲 王清輝 BD97020

四技電機四甲 許家齊 BD97901

中華民國 100 年 12 月 7 日

摘要

蓋茲說，未來人們使用電腦時將透過「自然介面」，直接以手指動作或說話的「自然」方式操縱電腦。如微軟推出的「Surface」觸控桌面電腦，另外也介紹了與福特汽車公司合作的「Sync」技術，可在車上聲控各種電子裝置。所以「語音聲控」的時代即將來臨。本專題主要是將語音辨識系統應用紅外線傳輸控制在與生活息息相關的家電中，使我們的日常生活可以更加便利輕鬆。

本專題共分為六大章節，第一章主要是在研究動機方法，以及語音辨識的歷史演進和帶給人們的便利，第二章節主要是描述語音辨識的原理，接下來第三章則是說明我們的語音辨識模組的硬體電路，介紹電路中主要晶片，及相關模組的功能，第四章主要在說明無線及紅外線傳輸控制模組，第五章是系統設計說明，第六章則是結論。最後則是參考文獻。

目 錄

摘要.....	I
目錄.....	II
圖目錄.....	IV
第一章 緒論	
1-1 語音辨識簡介.....	1
1-2 研究方法.....	1

1-3 預期結果	2
第二章 語音辨識的原理	
2-1 語音辨識流程及分類	3
2-2 特徵參數求取	
2-2.1 數位取樣	4
2-2.2 語音分段	4
2-2.3 消除靜音	5
2-2.4 語音前處理-特徵值的求取	5
第三章 聲控系統模組	
3-1 關於RSC-300/364	7
3-2 VCMM語音系統模組	8
3-2.1 VCMM模組電路	9
3-2.2 8051 控制電路	9
3-2.3 29C020晶片	12
3-2.4 LM386 聲頻放大器	13
3-2.5 5V穩壓IC	14
3-2.6 RS232位準轉換晶片	14
第四章 紅外線收發和編解碼模組	
4-1 紅外線發射器	16
4-2 紅外線LED驅動電路	16
4-3 紅外線接收器	17
4-4 紅外線應用方式	18
4-5 電器使用說明	19
第五章 系統設計說明	
5-1功能說明	20
5-2系統設計方塊圖	22
5-2.1電路方塊圖	22
5-2.2聲控技術	23
5-2.38051單晶片	25
5-3實作電路設計	28

5-4 軟體流程圖	32
5-5 系統實作	35
第六章 結論	36
參考文獻	36
作者簡介	
附錄	

圖目錄

圖2-1 語音訓練流程和處理步驟圖	3
圖 2-2 語音數位取樣示意圖	4

圖 2-3 音框範圍示意圖	5
圖 2-4 線性預測倒頻譜係數之特徵值參數求取流程	6
圖3-1 RSC-364硬體架構圖	8
圖3-2 VCMM聲控實驗應用模組	8
圖 3-3 VCMM 聲控模組方塊圖	9
圖 3-4 RSC-364 外部接線圖	9
圖3-5 AT89C2051外部接腳圖	11
圖 3-6 W29C020 電路圖	13
圖 3-7 LM386 接腳圖	13
圖 3-8 LM386 電路圖	14
圖 3-9 5V 電源產生電路圖	14
圖 3-10 MAX232 電路圖	14
圖 4-1 以直流電壓推動電路圖	16
圖4-2以數位電路推動電路圖	16
圖4-3以脈波方式驅動電路圖	17
圖4-4紅外線發射和接收示意圖	18
圖4-5紅外線發射和反射接收示意圖	18
圖4-6 繼電器電路圖	19
圖 5-1 聲控模組方塊圖	22
圖 5-2 發射端的工作方塊圖	22
圖 5-3 接收端的工作方塊圖	23
圖 5-4 8051 接腳說明	28
圖 5-5 8051 基本控制電路	29
圖 5-6 串列介面	30
圖 5-7 繼電器介面	31
圖5-8發射端軟體流程圖	32
圖5-9接收端軟體流程圖	33
圖 5-10 紅外線聲控發射模組	34
圖 5-11 紅外線接收模組	35
圖 5-12 語音辨識應用紅外線家電控制模組正視圖	35

第一章 緒論

1-1 語音辨識簡介

未來人們使用電腦時將透過「自然介面」，直接以手指動作或說話的「自然」方式操縱電腦。如微軟推出的「Surface」觸控桌面電腦，另外也介紹了與福特汽車公司合作的「Sync」技術，可在車上聲控各種電子裝置，所以「語音聲控」的時代即將來臨。本應用主要是利用 RSC-364[1] 發展的 VCMM 系統模組，來完成特定語者的語音辨識系統，發展一完整語音控制器，並藉由 LED 燈號和風扇轉動來驗證語音控制的可行性如此將語音辨識功能活用在與生活息息相關的家電中，使我們的日常生活可以更加的便利輕鬆。

語音技術牽涉的範圍極為廣大，包括語言學、音韻學、統計學、聽覺效能、訊號處理以及硬體設備等等，都是影響語音技術的因素。就因為語音技術是如此的複雜，所以即使現今的語音知識及硬體技術發展如此迅速，但成熟的語音技術還是被認為是一項艱鉅的挑戰。所以語音技術的核心在於語音辨識，而在語音辨識的領域包含了數字辨認、語音識別、句子萃取。一套好的語音辨識系統必須達到人性化的要求，例如只要任意講一句話它馬上就可以辨識，不需要有特定的條件，也就是在程式中它對關鍵字的萃取很敏感，但是遇到不適合的關鍵字他也會自動不採用，近年來在市場上已有一些不錯的應用軟體，例如在語音辨識系統的應用上就有無人化自動轉接系統、信用卡的身分辨識和一些消費性的電子產品等，進而提供了生活領域的方便性，而中文語音輸入系統也已經發展的很成熟了，相信日後聲控應用並能在技術上更有所突破。

1-2 研究方法

本專題是採用人工智慧特定語者語音辨認技術即利用 VCMM 聲控模組，並以 89S51-C 語言控制程式和經由紅外線收發模組來完成設計一套聲控系統，用語音來控制想要受控的裝置，訓練 4 組語音辨識詞組分別操控 4 個繼電器，即達到利用聲音來控制 4 個家電開關。特定語者是指誰訓練後，該人來辨認會很準確。如是男生錄音內容，男生來辨認音調接近，可以直接控制，若不準確，則自行重新錄音便可以得到高辨認率。

1-3 預期結果

本次的專題，預計達到的目標，首先是將預設 4 組聲控辨識詞組，但可提升到 10 組以上來符合家庭內諸多電器的需要，再來是增加紅外線控制系統，藉由著紅外線傳輸，將我們的聲控指令直接傳達到家電，最後則是讓這組聲控系統能符合家中每個人的聲紋，讓家裡的每一份子都能享受到這無與倫比的方便感。系統採用聲控模組做整合，例如測試錄音內容若為"電燈"，若說出"電燈"，則系統會回應電燈的語音登錄序號，此為辨認結果，發射此序號，到遠端接收。實驗結果，可以驗證以聲控模組做家電電源開關的功能，可以得到不錯的效果。本專題已經建立基本的聲控家電實驗平台，經由修改控制程式，可以應用在需要聲控的使用場合上。

第二章 語音辨識的原理

隨著人們對語音信號的研究，今天我們已經能夠憑藉著語音和機器溝通了，也就是說，機器可以分辨出人類所說的話。因此想要操作機器除了可以經由按鈕傳達命令外，最直接的方式便是用口下命令。

2.1 語音辨識流程及分類

語音辨認就是讓電腦或電子產品聽得懂人們所講的話，其基本架構如下圖 1 所示。圖 1 是分成語音訓練和語音處理辨認的部份。圖 1 所示為訓練的部份：將已知的語音信號經由「端點偵測」(End Point Detection)及「特徵參數求取」(Feature Extraction)而產生標準的語音參考樣本。

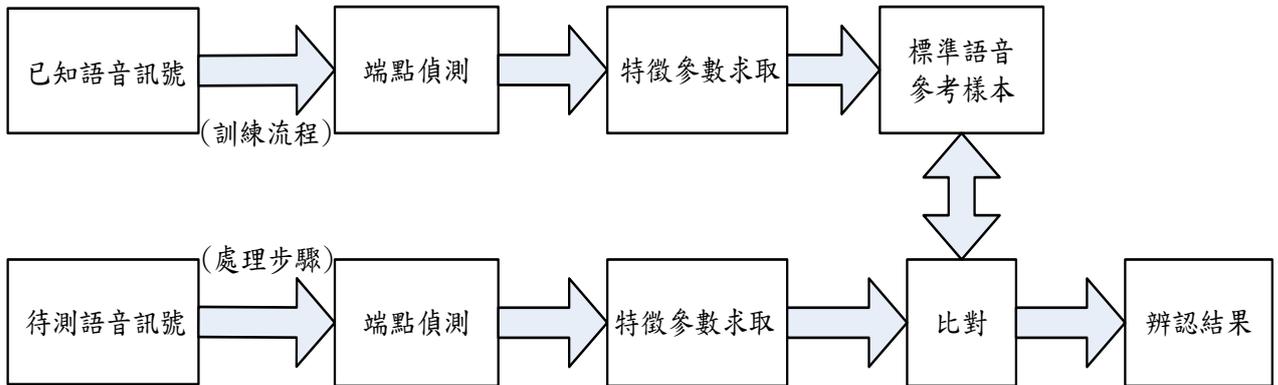


圖 2-1 語音訓練流程和處理步驟圖

在圖 1 中，將待測的語音信號，經同樣的處理步驟求得特徵參數後再與前述的標準語音參考樣本比對，找出最相似的參考樣本做為辨認的結果。

一般語音辨認依辨認的字彙數量多寡可分為：

1. 特定字彙：數個單字、詞句或者片語。
 2. 少量字彙：100 個單字左右。
 3. 大量字彙：以中文語音辨認而言則為全部中文字時，當辨認字彙數越多時候，混淆的程度也會越來越嚴重，因此辨認的困難度也會隨著提高許多。
- 若依使用者的限制而言又可分為特定語者與不特定語者：

1. 特定語者：

使用特定語者辨認系統前，須先把使用者的語音參考樣本存入當成比對的資料庫，意思是說特定語者的辨認系統在使用前就必須先進行圖 1 的訓練

學習步驟。

2.不特定語者：

使用本系統前根本不須要先學習，便能直接使用。一套最佳的語音辨認系統是不須經過學習便能進行語言辨識，但通常辨識率都較低。

另外若以說話方式的連續是否又可分為：

1.非連續語音辨認：所說的每一個字必須分開辨認。

2.連續語音辨認：以一般自然流利的說話方式來進行人性化的語音辨視，但是由於關係到相連音的問題，很難達到很好的辨認效果。

2.2 特徵參數求取

2.2.1 數位取樣

在類比信號轉為數位信號過程中，每隔一小段時間去做類比信號的取樣動作稱之數位取樣，而這一小段時間即為取樣速率(Sampling Rate)。例如8KHz 數位取樣，意思是每隔 0.125 毫秒(ms)讀取一次數位信號。如此原來連續的訊號，經過數位化的處理，變成一種不連續的訊號，如圖 2 所示。

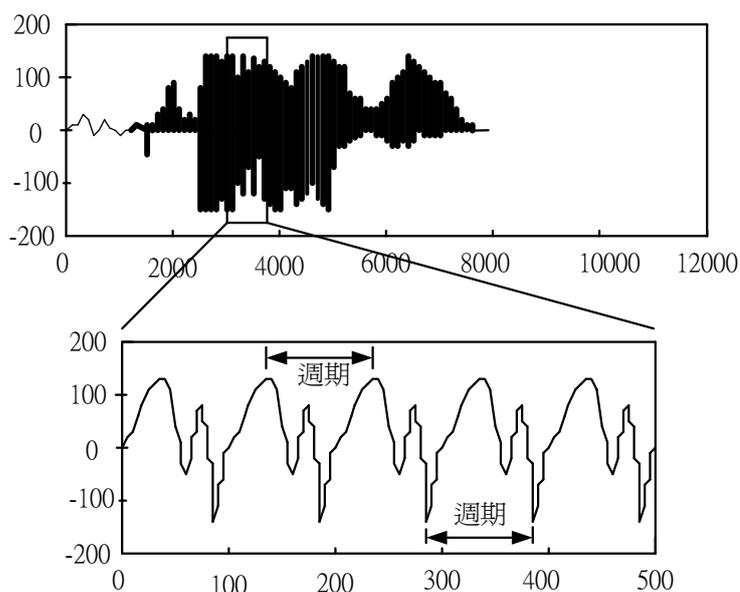


圖 2-2 語音數位取樣示意圖

2.2.2 語音分段

處理語音訊號這種不固定週期式的訊號時，最好先將其分段處理，把每一段當成一個固定的訊號，而這一小段部份就被稱為“音框”。

因為一整段語音的變化很大，而語音內每一個小區段都擁有不同的特性，將整個語音資料檔案分成為一個個的音框，並假設每個音框內的語音特性是固定的。如此一來，音框就成為語音處理的最小單位、大小沒有一定標

準視使用者自訂及一個音框可以計算出一個語音特徵值等特性。在以下說明例子中則是選擇 240 個取樣點當成一個音框作為標準，在之後的特徵值求法上，都是以音框為單位。一個音框具有一組特徵值，因此可以認定音框為辨識或者建立模型上的基本單位。

另外，我們需考慮音框與音框之間會有重疊的部份。如圖 3，在此使用 160 個重疊取樣點。例如語音第 1~240 個取樣點為第一個音框，80~320 個取樣點為第二個音框，而音框可由下面公式可得：

總音框數 = (取樣點數目 - 一個音框的取樣點數目) ÷ (一個音框的取樣點數目 - 音框和音框之前重疊的取樣點數目) + 1

以 8000 個取樣點的語音來說明，總音框數目為：

$$(8000-240)/(240-160)+1=7760/80+1=98$$

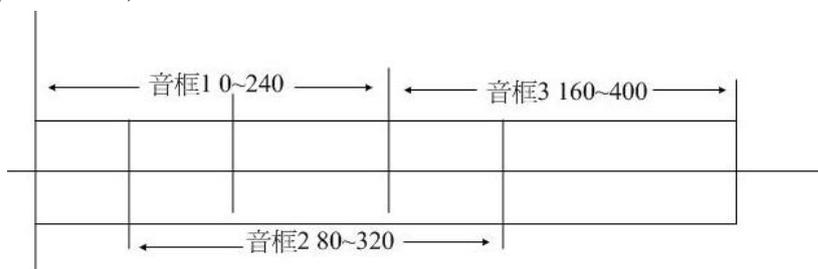


圖 2-3 音框範圍示意圖

2.2.3 消除靜音

每一個語音檔案錄製時，語音的前後一定有靜音的部份或是背景雜訊。因此靜音或是背景雜訊的部分我們並不需要，而將靜音或背景雜訊除去之後的辨識及產生的模型也會比較正確。

如何消除靜音?在此我們使用能量量測法:

- 1.找出一整個音框能量總和的平均值
- 2.設定門檻來區分有聲或無聲區域
- 3.判斷有聲無聲

2.2.4 語音前處理-特徵值的求取

由於語音資料的龐大，我們不可能將原始的語音資料全部儲存，作為模型訓練及辨識的工作，必須針對語音的特性，求取適當的特徵參數來做比對辨識。如果參數取的適當、穩定，則辨識的效果相對的會提高許多，因此訓練模型的第一步，就是語音的前處理—特徵值的求取。

在語者辨認的研究領域中，語音特徵(Speech Feature)的萃取是非常重要的

的處理步驟並且關係到系統的整體效能，在過去多種語音特徵中的研究文獻中[2-3]，如線性預測倒頻譜係數(Linear Predictive Cepstral Coefficients, LPCC)[4]、向量量化(Vector Quantization, VQ)[3,5-6]、類神經網路(Neural Network, NN)[7-13]、動態時間扭曲技術(Dynamic Time Warping, DTW)、隱藏式馬可夫模型(Hidden Markov Models, HMM)[3,14]及高斯混合式模型(Gaussian Mixture Models, GMM)等分類器應用在語者辨認上。其中倒頻譜係數(Cepstrum)是最被廣泛運用並且最具代表性的語音參數。線性預測倒頻譜係數被運用於語者辨認，主要是考慮其計算簡單且能有效表現母音(vowel)特性等優點(Matsui, & Furui, 1991)，其特徵值參數求取流程圖如圖 4 所示。

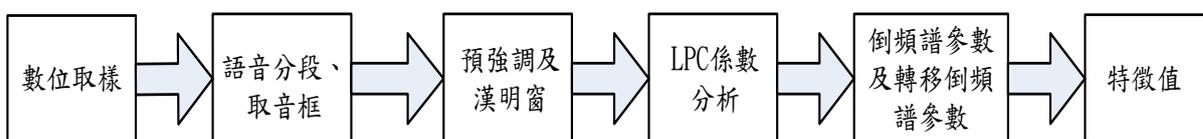


圖 2-4 線性預測倒頻譜係數之特徵值參數求取流程

隱藏式馬可夫模型(Hidden Markov Model, HMM)，HMM 的特色就是對於一段話不一定要判斷結尾，因為在訓練的時候就會自動的加入結尾狀態，使得一段話中有很多個結尾狀態，也增加辨識成功的機率。

而類神經網路所擅長的，與人類相似，具有以下幾種特性：

1. 平行處理的特性
2. 容錯(fault tolerance)特性
3. 聯想式記憶(Associative Memory)的特性
4. 解決最佳化(Optimization)問題
5. 超大型積體電路實作(VLSI Implementation)
6. 能處理一般演算法難以處理的問題

所以在上述的特性中，我們選擇使用 RSC-364 語音辨識晶片來做專題的研究，也藉此省略一些繁雜的軟硬體開發流程，直接進入語音控制應用的世界。

第三章 聲控系統模組

聲控家電開關設計，採用特定語者語音辨認技術，用語音來控制繼電器介面，特定語者是指誰訓練後，該人來辨認會很準確。如是男生錄音內容，男生來辨認音調接近，可以直接控制，若不準確，則自行重新錄音便可以得到高辨認率。

3-1 關於RSC-300/364

本專題應用RSC-364 晶片來做語音辨識，而Sensory 公司依據RSC-364亦提供一個VE 發展主機板(即Voice extreme development board) ，而硬體架構如圖5 所示。此晶片是以8位元MCU為核心的CMOS器件，內建有A/D、D/A、AGC、Pre-amplifier、16 個I/O 線、麥克風、喇叭等，使用的辨識方法為類神經網路法。根據手冊所示其辨識機率可達95%，相當實用。除此之外，RSC-364具有準確、快速的反應時間、低成本，且多功能，只要加上很少的外部元件，就可以組成一個語音識別系統。其運算能力為4MIPS(Million Instructions Per Second)為了提高運算能力，晶片上還多了一個24bit×24bit的乘法器。而RSC-364還具有5~15kb/s的語音合成功能，其語音合成是由Sensory專門設計，其音質較一般的好。它還具有改進的ADPCM(自適應差分脈衝調製)語音編解碼功能。

RSC-364 的設計，包括麥克風信號擴大，數據轉換，識別和綜合功能性，還有在ROM 儲存器(僅RSC-364 晶片具有)中，有一單晶片CPU 的核心，因此，RSC 300/364 能在14.32 MHz 提供整數性能的4 MIPS。這能使消費者以最小的費用取得最大的效能。

RSC-364 指令表非常類似於微處理器的8051 族群[15-16]。其處理器避免限制專用記憶體，透過有完全對稱來源和目的，適合全部指令。在2.5 Kbytes 的內部SRAM 中，2 Kbytes 為數據空間，0.5 Kbytes 為暫存空間。算術運算指令可以被用於任何暫存器。任何連結暫存器可能被用作16-bit 指示到任一來源或者目的一數據移動指令。Sensory 公司亦依據此發展主機板提出VE-C 的開發軟體，此軟體的應用類似C-語言，所以在開發上相當具有親和力。

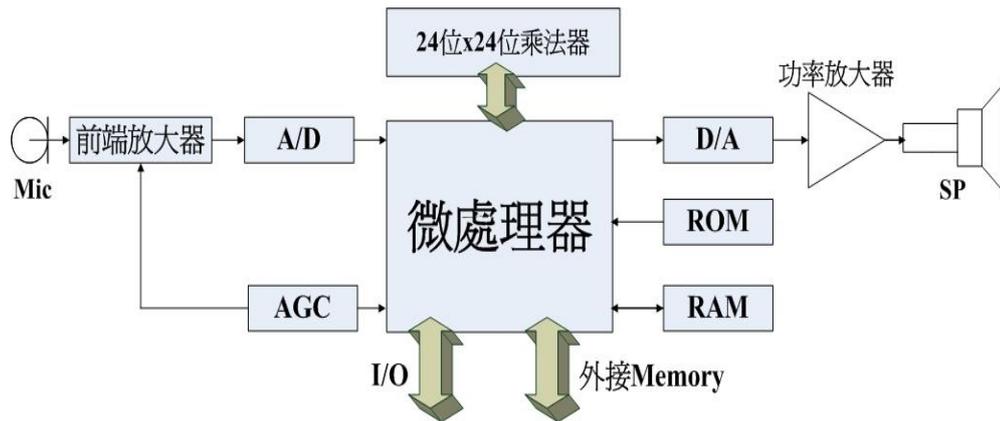


圖3-1 RSC-364硬體架構圖

建構上，RSC-364 的分離數據以及允許使用標準 EPROMs 的位址匯流排，ROMs 和 SRAMs 等，每一個外部儲存空間有另一個單一介面提供讀取和寫入。

3-2 VCMM 語音系統模組

本專題主要利用台灣偉克多微電腦工作室提供之開放式聲控實驗應用 VCMM 系統模組作為語音辨識單元如圖 6 所示，其中包含 RSC364 型辨識晶片做語音訊號之辨識處理，此 VCMM 系統模組可以獨立操作，不必倚靠 PC，也亦可在電腦中利用 Voice ExtremeIDE 撰寫程式，經由 RS-232 串列通訊埠下載到模版中，控制相關的輸出動作，如發出語音、觸發訊號產生，所以整套系統可做跨平台聲控應用，可由串列介面來下命令到 VCMM 做聲控處理。



圖3-2 VCMM聲控實驗應用模組

3-2.1 VCMM模組電路

圖 3-3 為 VCMM 聲控模組方塊圖。圖 3-4 RSC-364 外部接線圖

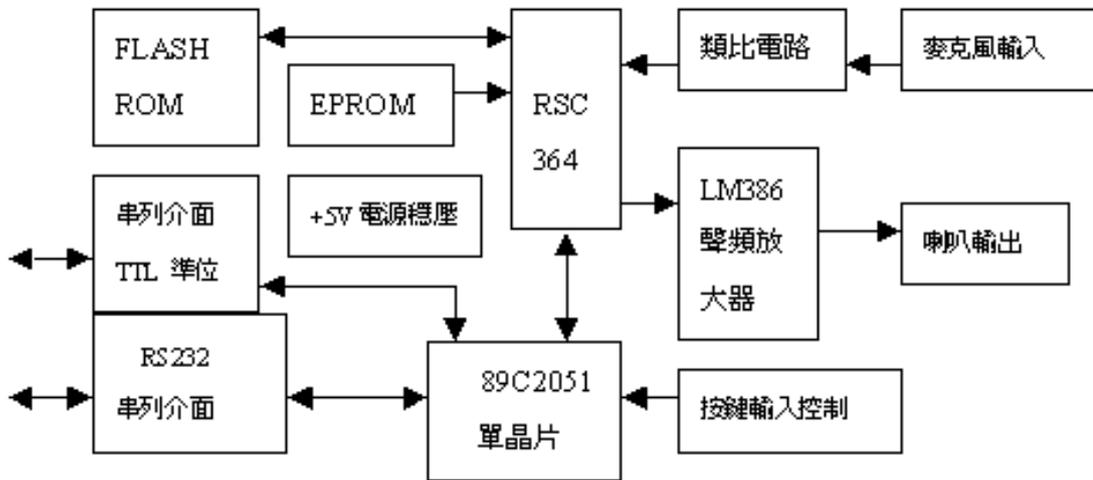


圖 3-3 VCMM 聲控模組方塊圖

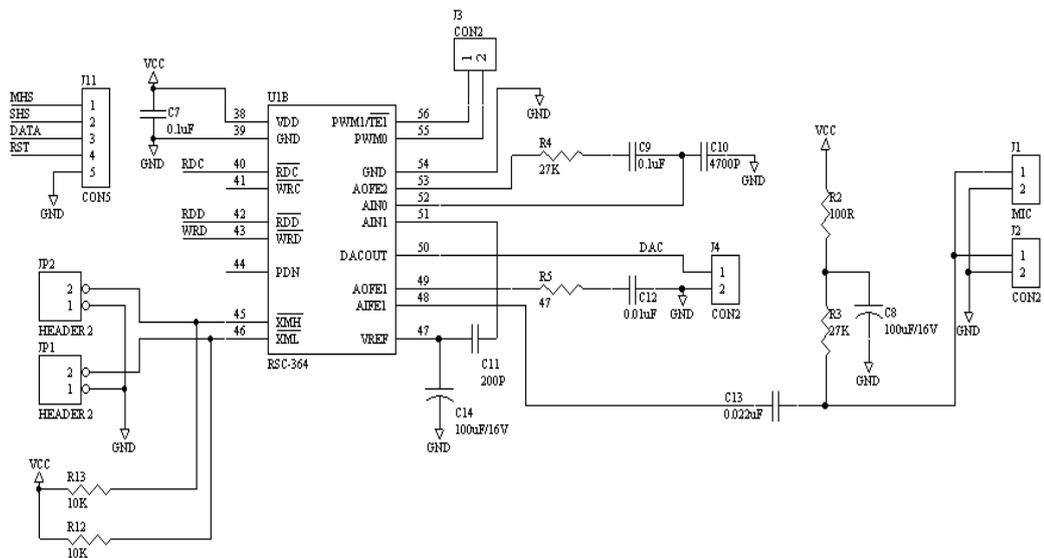


圖 3-4 RSC-364 外部接線圖

3-2.2 8051 控制電路

3-2.2.1 微電腦基本結構

微電腦系指由中央處理單元(CUP)，記憶體單元(Memory)及輸入/輸出單元(I/O)三大部份所組成。其中 CPU 可分成兩部份，即算術邏輯單元(ALU)及控制單元(CPU)。CPU 透過匯流排(BUS)執行程式碼的提取(Fetch)、解碼(Decode)、算術邏輯運算及讀寫時序信號的控制。記憶體單元提供存放程式與資料之空間，包含唯讀記憶體(Read Only Memory；ROM)。而輸入/輸出單元(I/O)提供 CPU 與外界週邊設備或元件溝通的管道，(如鍵盤、滑鼠、

顯示器等)。

3-2.2.2 單晶片微電腦

所謂單晶片微電腦(Single Chip)係指將中央處理單元(CPU)、記憶體單元(Memory)、及輸入/輸出單元(I/O)製作在同一晶片上，而形成一個能夠獨立運作的控制系統。其主要優點說明如下：

- (一).體積小：單晶片微電腦係 CPU、記憶體、I/O 埠等製作於同一晶片上，體積比一般微電腦明顯縮小。
- (二).使用簡單：單晶片內部的硬體架，與微電腦之複雜系統相比，其架構簡單，在使用與學習上更容易。
- (三).硬體接線容易：單晶片已主要功能濃縮於單一晶片上，使用時，只需接少許元件即可作廣泛的控制，除了提高硬體接線的可靠度之外，在裝配和維護上也相當容易。
- (四).擴充性佳：由於 HMOS 與 CHMOS 技術的發展，使單晶片不論在功能上、包裝密度上都增強許多，同時單晶片仍提供讀(Read)、寫(Write)等控制信號，在作外部記憶體或 I/O 埠的擴充時，相當簡便。

3-2.2.3 MCS-51 單晶片微電腦

MCS-51 族系單晶片是 INTEL 公司的產品，其中 8051 是 INTEL 公司於 1981 年生產製造的原始晶片。主要的功能及特性說明如下：

(一).8051 單晶片

1. 八位元微電腦控制晶片。
2. 4KB 內部程式記憶體，最大可外接擴充至 64KB。
3. 128 Bytes 內部資料記憶體，最大可外接擴充至 64KB。
4. 具有布林代數運算能力(位元邏輯)。
5. 四組可位元定址 I/O 埠 P0、P1、P2、P3。
6. 二組 16 位元計時 / 計數器 T0、T1。
7. 五個中斷源 INT0、INT1、T0、T1、RXD 或 TXD。
8. 一組全雙工串列埠 UART

(二).8052 單晶片

1. 八位元微電腦控制晶片。
2. 8KB 內部程式記憶體，最大可外接擴充至 64KB。
3. 256 Bytes 內部資料記憶體，最大可外接擴充至 64KB。

圖 3-5 AT89C2051 外部接腳圖

3-2.3 29C020 晶片

29C020 是 Flash ROM(快閃 ROM)，屬於單電壓晶片。Flash ROM 和 EEPROM 晶片最大的區別是，在刪除資料時，不需加特定的編程電壓。另外，Flash ROM 在刪除資料時，並非以 Byte 為基本單位，而是以 Sector(又稱 Block)為最小單位。Sector 的大小隨廠商的不同而有所不同，只有在寫入時，才以 Byte 為最小單位寫入。主板上用的還有 27 系列的晶片。27 系列的晶片屬於 EPROM，其內的資料是用 EPROM 抹除器(EPROM Eraser)發出的紫外線照射來擦除的，並且要用專用的編輯器才可以把資料重新再寫入，往晶片寫內容時也必須加一定的編輯電壓。

現在的主板 BIOS 幾乎都採用 Flash ROM(快閃 ROM)，它其實就是一種可快速讀寫的 EEPROM(Electrically Erasable Programmable ROM)，顧名思義，它是一種在一定的電壓、電流條件下，可對其 Firmware 進行更新的體積電路塊。

有很多晶片廠商都在生產 Flash ROM 晶片，我們在主板上常見的有 Winbond、SST、Intel、MXIC、ATMEL 等品牌的產品，這些廠商又提供了很多種型號的晶片，型號不同，晶片的存儲容量和讀寫電壓也不同。Flash ROM 晶片大致分為 28、29 兩大系列 28 系列的 Flash ROM 晶片是雙電壓設計的，它可以在 5V 的電壓的條件下讀取，而寫入則必須提供 12V 的電壓。

採用這種晶片的主板在升級時，會給普通的電腦用戶造成不小的麻煩---要開機箱、改跳線設置，太麻煩了。29 系列的 Flash ROM 晶片則相對簡單，由於其採用單電壓設計，讀寫都採用 5V 電壓，因此只動用軟體就可以完成讀寫 Firmware 的操作。在主板說明書中，主板廠商還列出了 Flash ROM 晶片的容量，其中有 1M 和 2M 兩種容量的型號。這裏，“M”的單位是指“Mbit”，1M 的 Flash ROM 晶片實際能存儲的容量為 1Mbit=8*128Kbyte(1Byte=8bit)，2M 的晶片為 256K。以上這些技術參數都可以通過晶片正面的編號來區分，這個編號是嚴格遵循體積電路編號規則來標注的，如 Intel 生產的 Flash ROM 晶片，它的晶片編號為“28F010”，由此可知該晶片是 5V 讀、12V 寫，容量為 1Mbit 的 Flash ROM 晶片。

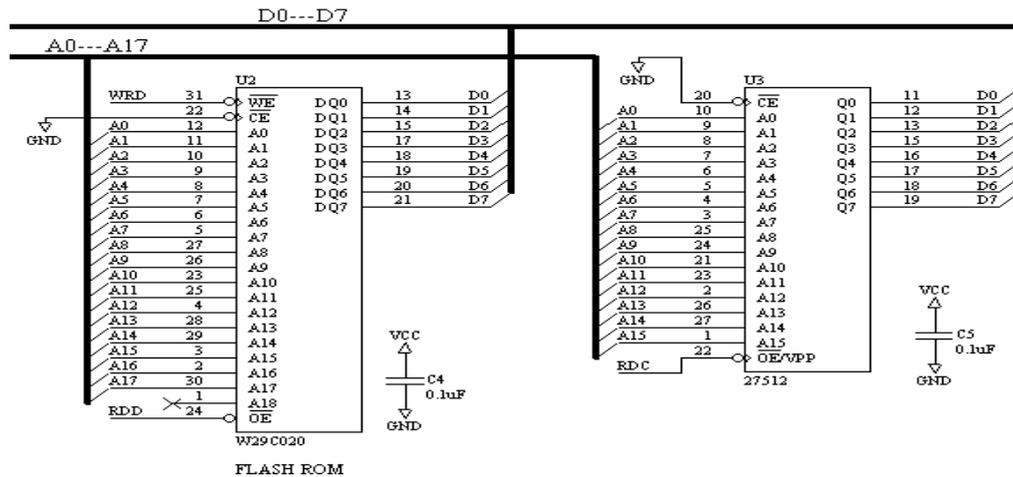


圖 3-6 W29C020 電路圖

3-2.4 LM386 聲頻放大器

3-2.4.1 LM386 概述

LM386 是美國國家半導體公司生產的音頻功率放大器，主要應用於低電壓消費類產品。為使外圍元件最少，電壓增益內置為 20。但在 1 腳和 8 腳之間增加一只外接電阻和電容，便可將電壓增益調為任意值，直至 200。輸入端以地位參考，同時輸出端被自動偏置到電源電壓的一半，在 6V 電源電壓下，它的靜態耗值為 24MW，使得 LM386 特別適用於電池供電的場合。LM386 的封裝形式有塑封 8 引線雙列直插式和貼片式。

3-2.4.2 LM386 特性

1. 靜態功耗低，約為 4mA，可用於電池供電。
2. 工作電壓範圍寬，4-12V or 5-18V。
3. 外圍元件少。
4. 電壓增益可調，20-200。
5. 低失真度。

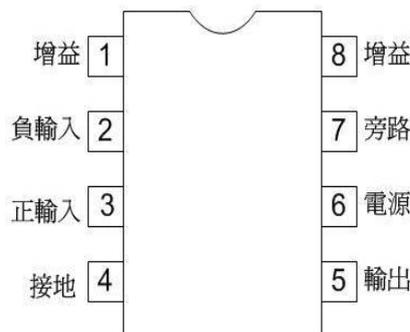


圖 3-7 LM386 接腳圖

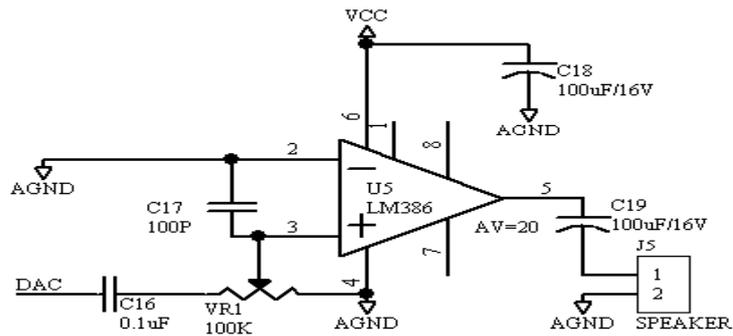


圖 3-8 LM386 電路圖

3-2.5 5V 穩壓 IC

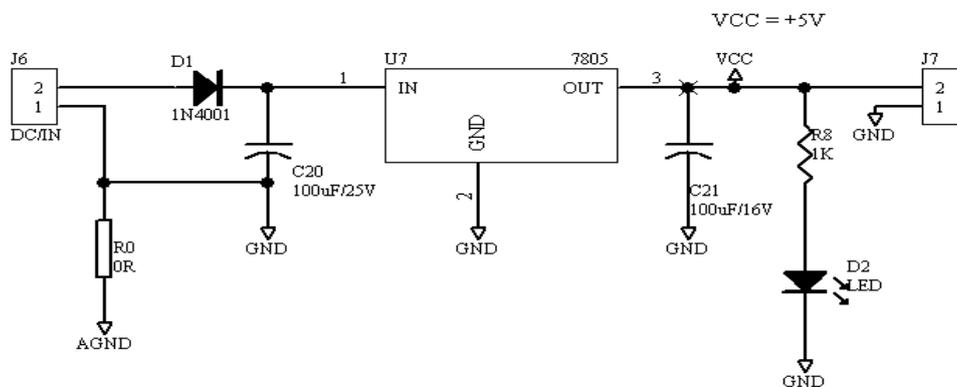


圖 3-9 5V 電源產生電路圖

3-2.6 RS232 位準轉換晶片

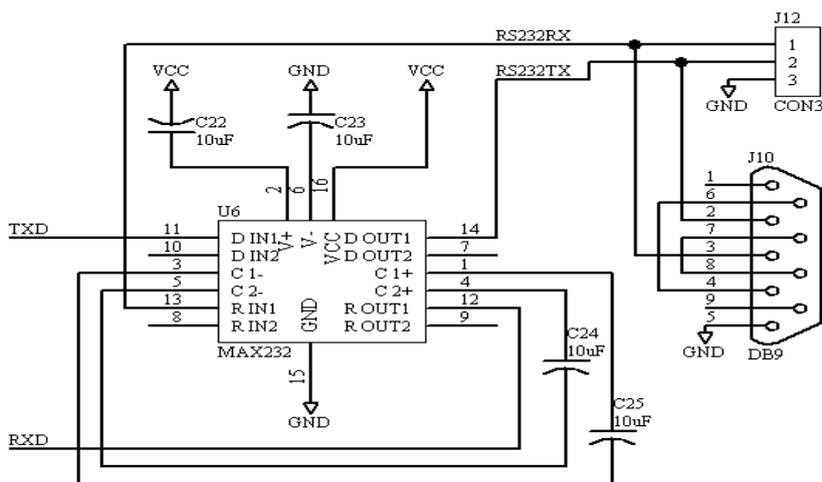


圖 3-10 MAX232 電路圖

因此VCMM的特性及功能為:

利用本套系統可以自行設計獨立操作型特定語者聲控系統，產品附加價值高。系統由 89C2051 及聲控晶片 RSC-364 (TQFP 64 PIN 包裝)。89C2051 只用 4 條 I/O 線來控制聲控晶片。適合特定語者的單音、字、詞語音辨識。不限定說話語言，國語、台語、英語皆可。有自動語音輸入偵測的功能，特定語者辨識率可達 95% 以上，反應時間小於 1 秒。系統參數及語音參考樣本一旦輸入後資料可以長久保存。系統採用模組化設計，擴充性佳，可適合不同的硬體工作平台。線上訓練輸入的語音可以壓縮成語音資料而由系統說出來當作辨認結果確認。包含有英文的語音提示語做語音動作引導。系統展示 4 組語音辨認功能，最多可以擴充控制到 60 組語音辨認。需外加+5V 電源供電。內建 DTMF 信號產生功能，可做電話撥號的實驗。內建 4 只按鍵開關及 RS232 串列通訊介面。提供完整 8051 控制介面及聲控晶片電路圖。含 8051 組合語言串列應用範例原始程式及 PC VB RS232 原始程式。可擴充軟硬體功能做進一步產品設計或聲控專題製作。

第四章 紅外線收發和編解碼模組

4-1 紅外線發射器

紅外線 LED 發光波長在 940nm 左右，是以 GaAa 組成的二極體裝置，被加上順向偏壓時，可發出紅外光，電壓約為 1.2V，順向電流最大值一般可達 50mA，工作的順向電流越大，發出的紅外線越強。

4-2 紅外線 LED 驅動電路

1. 以直流電壓推動：與 LED 的推動方式相同，紅外線 LED 也是電流驅動元件，發射強度與順向電流程比例 變化，通常會加上限流電阻以防燒毀。
2. 若以數位推動可加上電晶體。
3. 以脈波調變驅動紅外線 LED：採用脈波調變的好處是當脈波寬度越小時，其工作周期越小，可使紅外線 LED 承受較大的峰值電流，以產生較強的發射光束。通常可採用不穩態多諧振盪器來做。

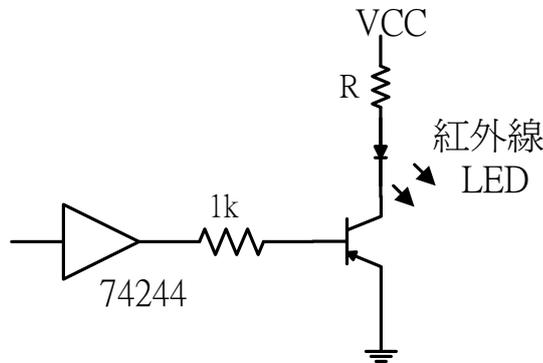


圖4-1 以直流電壓推動電路圖

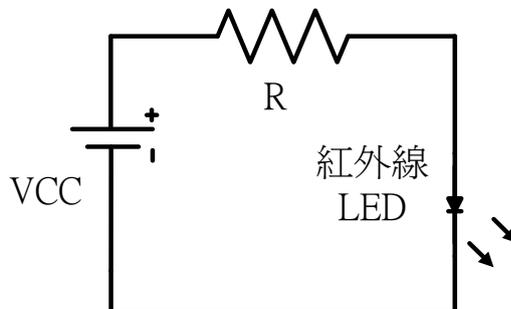


圖4-2 以數位電路推動電路圖

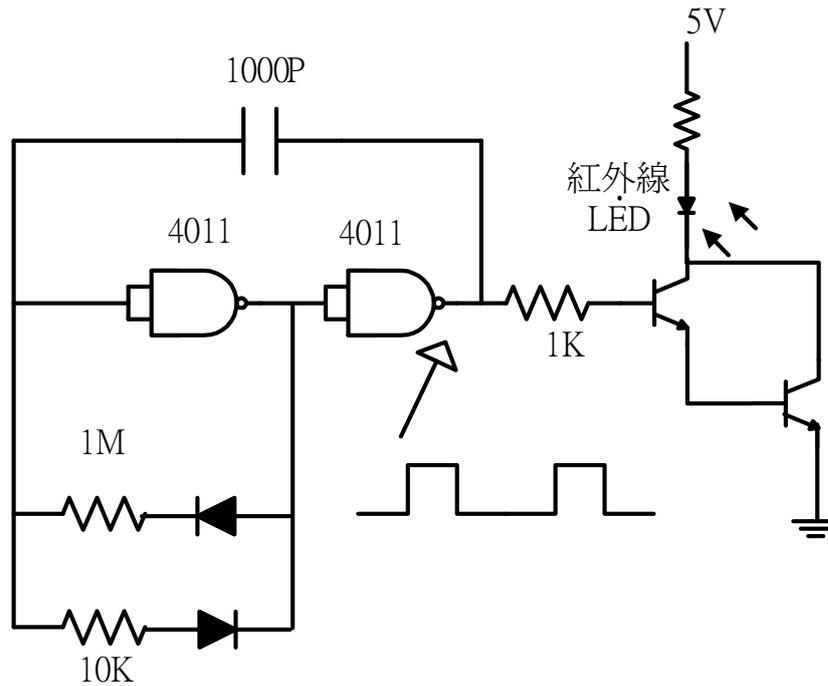


圖4-3 以脈波方式驅動電路圖

上圖是以脈波方式驅動紅外線LED，是有2組不穩態多諧振盪器，OSC1的振盪頻率較低，OSC2的振盪頻率較高，利用OSC1去控制OSC的工作，以形成脈波束去驅動紅外線LED。

4-3紅外線接收器

紅外線接收二極體工作原理紅外線的接收部分可以採紅外線接收二極體或光電晶體來使用，紅外線接收二極體是一種光電二極體元件，其對紅外線的反應靈敏，當有紅外線照射時，二端阻抗下降。光電晶體也可當作紅外線的接收器，但其靈敏度及效果較差。

4-4紅外線的應用方式

圖4-4為紅外線持續發射時，若有障礙物檢出，輸出端可得到一脈波信號，此信號可送至微電腦I/O，可作防盜器或計數器使用。圖4-5 是利用反射的方式，可作為物體偵測，或近接開關。

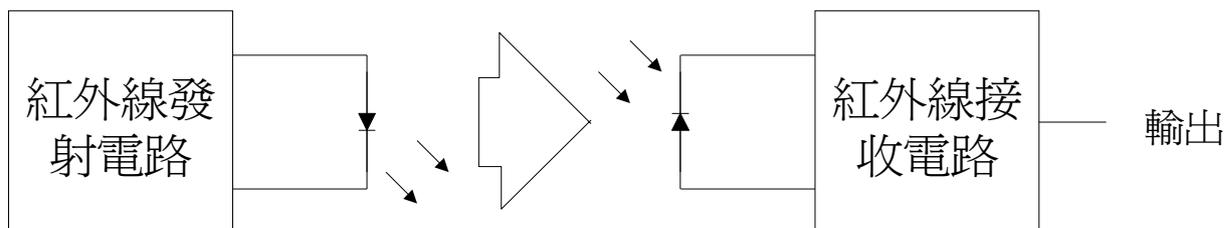


圖4-4 紅外線發射和接收示意圖

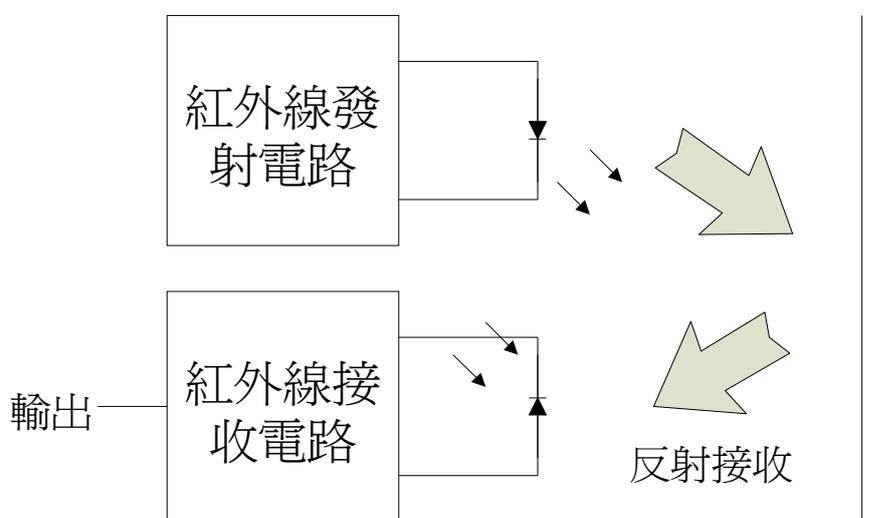


圖4-5 紅外線發射和反射接收示意圖

4-5 繼電器使用說明

在模組的擴充電路方面，外接了一組由四組繼電器所組成的開關，可以讓家電開關更多元化且多樣化，且透過對晶片的程式擴充，可以有效的控制 9 組家電。當然也必須要了解繼電器的使用方式，才能有效的連接到模組上，並完全的吻合我們的需要。

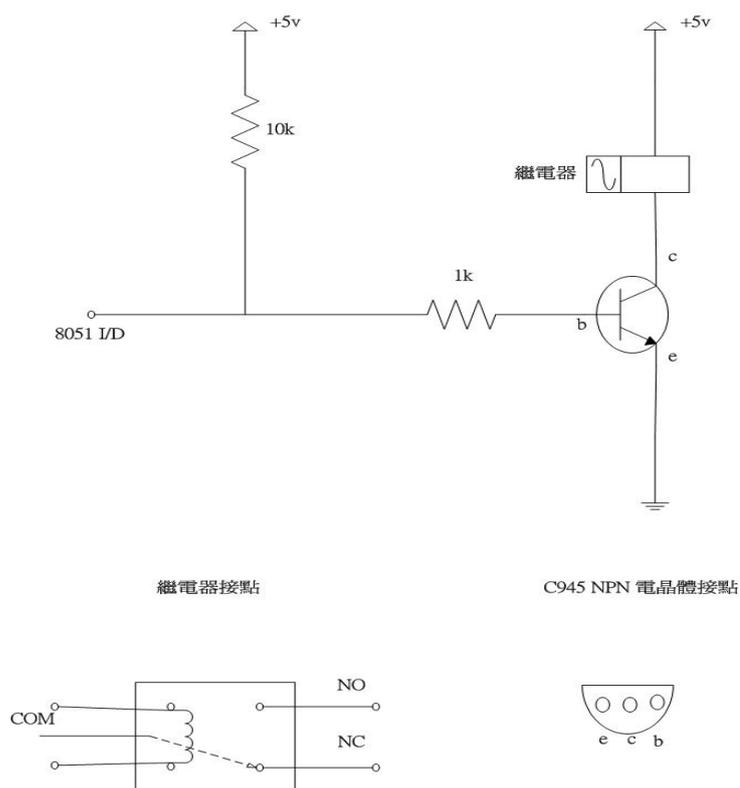


圖4-6 繼電器電路圖

透過語音辨識模組來控制繼電器ON/OFF，即可控制家電ON/OFF。繼電器接點說明如下：

NC：Normal Close常閉點。以COM為共同點，NC與COM在平時是呈導通狀態。

COM：Common，共通點。輸出控制接點的共同接點。

NO：Normal Open常開點。NO與COM在平時是呈開路狀態，當繼電器動作時，NO與COM則呈開路狀態。

第五章 系統設計說明

在本專題中，利用單晶片 8051、按鍵、結合聲控模組、紅外線發射接收介面，進而遙控繼電器介面，來控制家電電源開關的功能。以 8051 c 語言控制程式來設計一套聲控系統，採用人工智慧特定語者語音辨認技術，用語音來控制想要受控的裝置，特定語者是指誰訓練後，該人來辨認會很準確。如是男生錄音內容，男生來辨認音調接近，可以直接控制，若不準確，則自行重新錄音便可以得到高辨認率。系統採用聲控模組做整合，測試錄音內容若為"電燈"，若說出"電燈"，則系統會回應電燈的語音登錄序號，此為辨認結果，發射此序號，到遠端接收。實驗結果，可以驗證以聲控模組做家電電源開關的功能，可以得到不錯的效果。本專題已經建立基本的聲控家電實驗平台，經由修改控制程式，可以應用在需要聲控的使用場合上。

5-1 功能說明

聲控家電開關設計，採用特定語者語音辨認技術，用語音來控制繼電器介面，特定語者是指誰訓練後，該人來辨認會很準確。如是男生錄音內容，男生來辨認音調接近，可以直接控制，若不準確，則自行重新錄音便可以得到高辨認率。系統採用 VCMM 聲控模組做整合。

1. 首先在控制板上按鍵 K1--K4 動作如下:

(a) 操作鍵 K1: 語音訓練

(b) 操作鍵 K2: 聆聽系統已存在的語音內容。展示程式為編號 0--4，重複循環。

(c) 操作鍵 K3: 進行辨認 1 次。當工作 LED 亮起表示系統正在等待語音輸入，此時可以說出命令來做控制。

(d) 操作鍵 K4: 連續進行辨認。

2. 錄音有聲控內容，設定範例如下：(共有 4 組)

NO	語音內容	聲控家電動作
0	電燈	電燈開/關
1	電扇	電扇開/關
2	門燈	門燈開/關

3 樓燈 樓燈開/關

3. VCMM 按鍵操作 S1--S4 功能鍵如下:

<1>按鍵 S1: 做語音參考樣本訓練輸入，一次訓練一組，展示系統為 5 個辨認的單音。已訓練的語音會永久保存在 FLASH 的記憶晶片中，即使關機還是有效，語音訓練輸入需要輸入 2--3 次，一般正常情況輸入 2 次語音便可

按下 S1 鍵，操作過程如下:

系統說出"SAY NAME" (說一單音) -----> 第 1 次錄音

系統說出"REPEAT NAME" (重覆一遍)(非必要) -----> 第 2 次錄音

系統說出"REPEAT NAME TO RECORD"(重覆一遍錄音用)----> 第 3 次錄音

第 1 2 次錄音做為產生語音參考樣本，第 3 次錄音則是做聆聽已存在的語音內容用。若訓練成功後，系統會說出您剛剛輸入的語音，以做確認。

<2>按鍵 S2: 修改原先已存在的語音參考樣本

先按 S3 聆聽系統已存在的某組語音內容。再按 S2 則該組內容會先被刪除，再執行語音輸入訓練過程來建立新的語音參考樣本。若在語音輸入訓練過程中失敗，可以使用 S1 鍵來輸入新的語音樣本。

<3>按鍵 S3: 聆聽系統已存在的語音內容。展示程式為編號 0--4，重複循環。

<4>按鍵 S4: 進行辨認。

4.當在 VCMM 上已訓練有語音時，則可以直接進行辨認，

例如說出"電燈"，系統說出"電燈"，表示辨認正確。接著電燈打開，再說出 "電燈"，則電燈會關閉。例如編號 0，內容為"電燈"，對應到的繼電器為 RY0 說出電燈-->繼電器為 RY1 ON--> 電燈電源開，再次說出電燈-->繼電器為 RY1 OFF -->電燈電源關當然繼電器接點要配線至電源端，並連接到電燈插頭,並將開關 ON，直接由 AC110 市電來控制。

5.此控制程式是燒錄在 89C51 中，原始程式以 8051 c 語言來設計。

6.系統設計採用 2 片控制板設計，一片紅外線發射板，一片紅外線接收介面，進而遙控繼電器介面，來控制家電電源開關的功能。無線電發射板發射語音辨認結果，當紅外線接收介面收到信號後，推動相對繼電器動作。

5-2 系統設計方塊圖

5-2.1 電路方塊圖

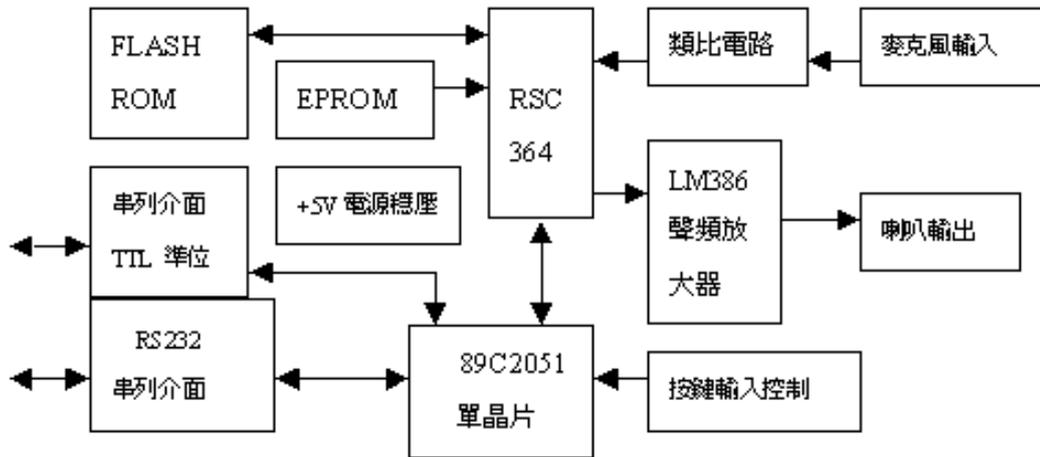


圖 5-1 聲控模組方塊圖

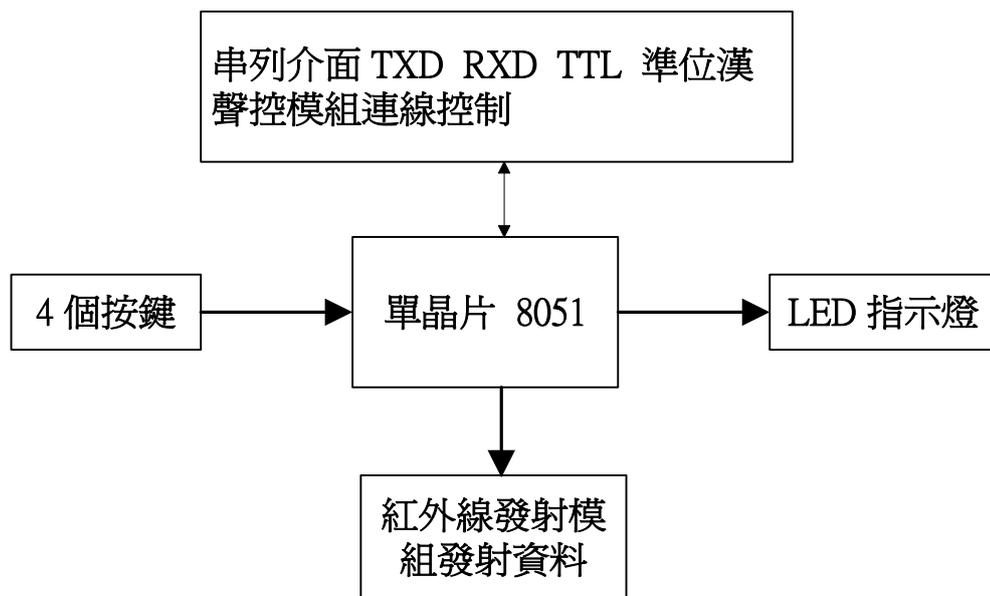


圖 5-2 發射端的工作方塊圖

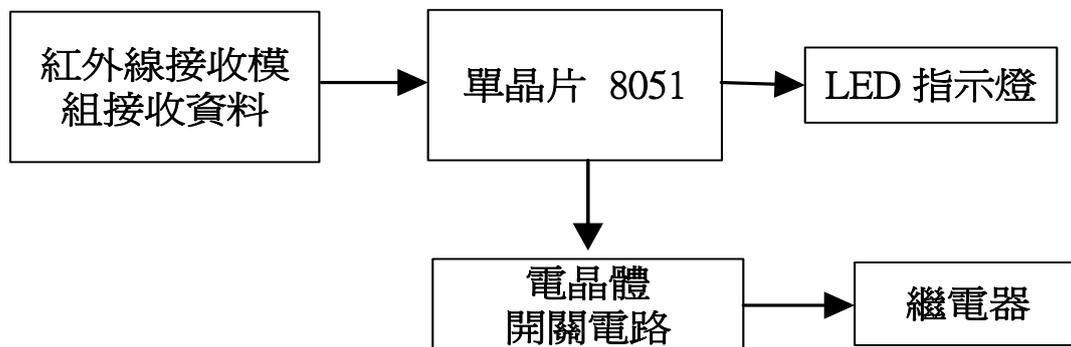


圖 5-3 接收端的工作方塊圖

5-2.2 聲控技術

聲控的技術是用來設計一台會聽話的電腦，只要對著麥克風說話，便可以指揮電腦動作，也就是要實現"芝麻開門"聲控電腦的夢想，當然現在這已不再是夢想了。這部份是本節將要介紹的重點，並且將以軟體及硬體來實現。語音辨識系統應用的範圍相當廣範，現在隨著許多關鍵技術的突破及 VLSI 技術的進步，市面上已出現許多方便使用的聲控應用產品，如中文語音輸入系統，聲控大哥大語音撥號，聲控汽車音響，您只需動口，不必動手便可以享受科技帶來的方便，相信未來還會有更多有趣的聲控電子產品上市。

聲控電腦依系統所能辨認單字多寡可以分類為以下三種：

□ 特定字彙

幾個單字、詞或是片語。

□ 少量字彙

數十個單字、詞或是片語。

□ 大量字彙

含蓋所有的單字、詞或是片語發音。以中文語音辨認而言便是所有中文字。

聲控電腦的分類，依使用者是否需要事先做訓練分為三種：

□ 特定語者

辨認系統只能辨認某一特定使用者的聲音，使用者在第一次使用此系統

時需將所有要辨認的字彙唸過一到二次，當做語音參考樣本。此過程稱為語音訓練，手機聲控撥號便是特定語者語音辨認的應用，使用手機的主人原先所輸入的人名，下回辨認時，只需說出人名，便可以辨認人名及出現對應的電話號碼並撥出電話。本章所介紹的聲控模組，便是此一技術的應用，誰來訓練說出語音，辨認時會很準確，當然如果訓練時是男生的語音，若其他的男生來辨認，只要腔調及音頻不要差異太大，仍然可以辨認出來。

□ 語者調適

使用者只要曾經對辨認系統訓練過，此系統便可以辨認出他的聲音，是一種比較有彈性的做法，使用者不需要唸完所有的音，只需要唸過一部份的單音後，系統會自動將語音參考樣本做調整。

□ 不特定語者

任何使用者不需要事先對辨認系統訓練，皆可以使用聲控系統，此時系統中已經包含不同種性別、年齡的口音，這種聲控系統是一種最完美實用的系統。

聲控電腦依語者說話的方式分類可以分為二種：

□ 單音辨認：

系統只能辨認單音，因此使用者所說出的每一個字必需分開來。

□ 連續音辨認：

系統可以接受語者連續發音。

由以上的說明，讀者可以了解，一套最理想的聲控電腦系統應該是大量字彙，不特定語者連續音語音辨認系統，一般人不需要經過學習，便可以讓電腦聽懂他發出的語音，也就是說只要對著電腦說話便可以直接來控制電腦動作了，但是要完成這樣的一套高辨認率的系統實在不是一件容易的工作。

一般在應用上，特定語者，少量字彙的單音辨認系統便可以滿足我們的特定需求，若能先完成這樣一套簡單而高辨認率的聲控系統，在不影響辨認率的情況下而後再逐漸加大字彙量，或是修改語者訓練的方式，採用語者調適的方法，也可以提升聲控系統的整體效能，增加使用的方便性。以下列舉一般設計聲控系統的基本規格供者參考，此一系統是較易完成的系統，並且要有不錯的辨識效果。

聲控電腦基本規格

1. 辨認率高
2. 特定語者
3. 少量字彙
4. 單音辨認

5-2.3 8051 單晶片

8051 主要功能列舉如下：

- 為一般控制應用的 8 位元單晶片
- 晶片內部具時脈振盪器（傳統最高工作頻率可至 12MHz）
- 內部程式記憶體（ROM）為 4K 位元組
- 內部資料記憶體（RAM）為 128 位元組
- 外部程式記憶體可擴充至 64K 位元組
- 外部資料記憶體可擴充至 64K 位元組
- 32 條雙向輸入輸出線，且每條均可以單獨做 I/O 的控制
- 5 個中斷向量源
- 2 組獨立的 16 位元計時器
- 1 個全多工串列通信埠
- 8751 及 8752 單晶片具有資料保密的功能
- 單晶片提供位元邏輯運算指令

8051 接腳說明

圖 5-4 是 8051 的接腳圖，說明如下：

□VCC

8051 電源正端輸入，接+5V。

□VSS

電源地端。

□XTAL1

單晶片系統時脈的反相放大器輸入端。

□XTAL2

系統時脈的反相放大器輸出端，一般在設計上只要在 XTAL1 和 XTAL2 上接上一只石英振盪晶體系統就可以動作了，此外可以在兩接腳與地之間

加入一 20PF 的小電容，可以使系統更穩定，避免雜訊干擾而當機。

□RESET

8051 的重置接腳，高電位動作，當要對晶片重置時，只要對此接腳電位提昇至高電位並保持兩個機器週期以上的時間，8051 便能完成系統重置的各項動作，使得內部特殊功能暫存器之內容均被設成已知狀態，並且至地址 0000H 處開始讀入程式碼而執行程式。

□EA/Vpp

"EA"為英文"External Access"的縮寫，表示存取外部程式碼之意，低電位動作，也就是說當此接腳接低電位後，系統會取用外部的程式碼（存於外部 EPROM 中）來執行程式。因此在 8031 及 8032 中，EA 接腳必須接低電位，因為其內部無程式記憶體空間。如果是使用 8751 內部程式空間時，此接腳要接成高電位。此外，在將程式碼燒錄至 8751 內部 EPROM 時，可以利用此接腳來輸入 21V 的燒錄高壓（Vpp）。

□ALE/PROG

ALE 是英文"Address Latch Enable"的縮寫，表示位址栓鎖致能信號。8051 可以利用這支接腳來觸發外部的 8 位元栓鎖（如 74LS373），將埠 0 的位址匯流排（A0~A7）栓鎖進入栓鎖器中，因為 8051 是以多工的方式送出位址及資料。平時在程式執行時 ALE 接腳的輸出頻率約是系統工作頻率的 1/6，因此可以用來驅動其他週邊晶片的時基輸入。此外在燒錄 8751 程式碼時，此接腳會被當成程式規劃的特殊功能來使用。

□PSEN

此為"Program Store Enable"的縮寫，其意為程式儲存致能，當 8051 被設成為讀取外部程式碼工作模式時（EA=0），會送出此信號以便取得程式碼，通常這支腳是接到 EPROM 的 OE 腳。8051 可以利用 PSEN 及 RD 接腳分別致能存在外部的 RAM 與 EPROM，使得資料記憶體與程式記憶體可以合併在一起而共用 64K 的定址範圍。

□PORT0（P0.0~P0.7）

埠 0 是一個 8 位元寬的開路汲極（Open Drain）雙向輸出入埠，共有 8 個位元，P0.0 表位元 0，P0.1 表位元 1，餘此類推。其他三個 I/O 埠（P1、P2、P3）則不具有此電路組態，而是內部有一提升電路，P0 在當做 I/O

用時可以推動 8 個 LS 的 TTL 負載。如果當 EA 接腳為低電位時（即取用外部程式碼或資料記憶體），P0 就以多工方式提供位址匯流排（A0~A7）及資料匯流排（D0~D7）。設計者必須外加一栓鎖器將埠 0 送出的位址栓鎖住成為 A0~A7，再配合埠 2 所送出的 A8~A15 合成一完整的 16 位元地址匯流排，而定址到 64K 的外部記憶體空間。

□PORT2（P2.0~P2.7）

埠 2 是具有內部提升電路的雙向 I/O 埠，每一支接腳可以推動 4 個 LS 的 TTL 負載，若將埠 2 的輸出設為高電位時，此埠便能當成輸入埠來使用。P2 除了當做一般 I/O 埠使用外，若是在 8051 擴充外接程式記憶體或資料記憶體時，也提供位址匯流排的高位元組 A8~A15，這個時候 P2 便不能當做 I/O 來使用了。

□PORT1（P1.0~P1.7）

埠 1 也是具有內部提升電路的雙向 I/O 埠，其輸出緩衝器可以推動 4 個 LS TTL 負載，同樣地若將埠 1 的輸出設為高電位，便是由此埠來輸入資料。如果是使用 8052 或是 8032 的話，P1.0 又當做計時器 2 的外部脈波輸入腳，而 P1.1 可以有 T2EX 功能，可以做外部中斷輸入的觸發腳位。

□PORT3（P3.0~P3.7）

埠 3 也具有內部提升電路的雙向 I/O 埠，其輸出緩衝器可以推動 4 個 TTL 負載，同時還多工具有其他的額外特殊功能，包括串列通訊、外部中斷控制、計時計數控制及外部資料記憶體內容的讀取或寫入控制等功能。其接腳分配如下：

P3.0：RXD，串列通訊輸入。

P3.1：TXD，串列通訊輸出。

P3.2：INT0，外部中斷 0 輸入。

P3.3：INT1，外部中斷 1 輸入。

P3.4：T0，計時計數器 0 輸入。

P3.5：T1，計時計數器 1 輸入。

P3.6：WR：外部資料記憶體的寫入信號。

P3.7：RD，外部資料記憶體的讀取信號。

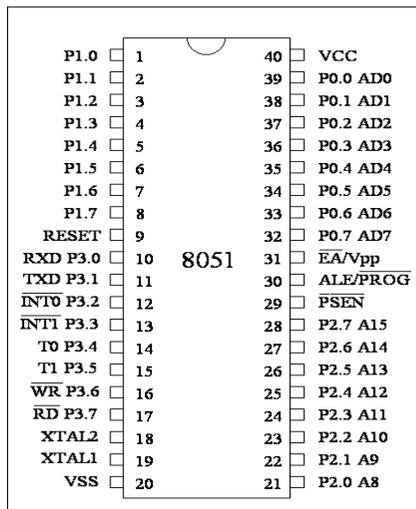


圖 5-4 8051 接腳說明

5-3 實作電路設計

電路設計分為以下幾部分：

- 單晶片 8051
- 串列介面
- 紅外線發射接收電路
- 繼電器介面

8051 控制電路如圖 5-5 所示。石英振盪晶體提供 8051 工作所需時脈。系統重置是任何微處理機系統執行的第一步，使整顆控制晶片回到預先設定的硬體狀態下。8051 單晶片的系統重置是由 RESET 接腳來做控制，當此接腳送入高電位超過 24 個振盪週期後，8051 即進入晶片內部重置的狀態下，而且一直在此狀態下等待，直到 RESET 接腳轉為低電位後，才檢查 EA 接腳是高電位或低電位，若為高電位則執行晶片內部的程式碼，若為低電位便會執行外部的程式。

而 8051 在系統重置時，將其內部的一些重要暫存器設定為某些特定的值，對 8051 而言，在晶片內部的 RESET 接腳接有一史密特觸發電路，以及一個電阻到地端，在外部接上一個 10uF 的電容加上電阻，便可達成 RESET 的重置電路。果要使程式碼重新執行時，可以按下外接的按鈕開關，

將 RESET 接腳接至+5V 電源，使系統重置而重新執行程式。LED 是當作工作指示燈用，程式一執行，工作指示燈閃動，表示程式已經開始正常執行了。

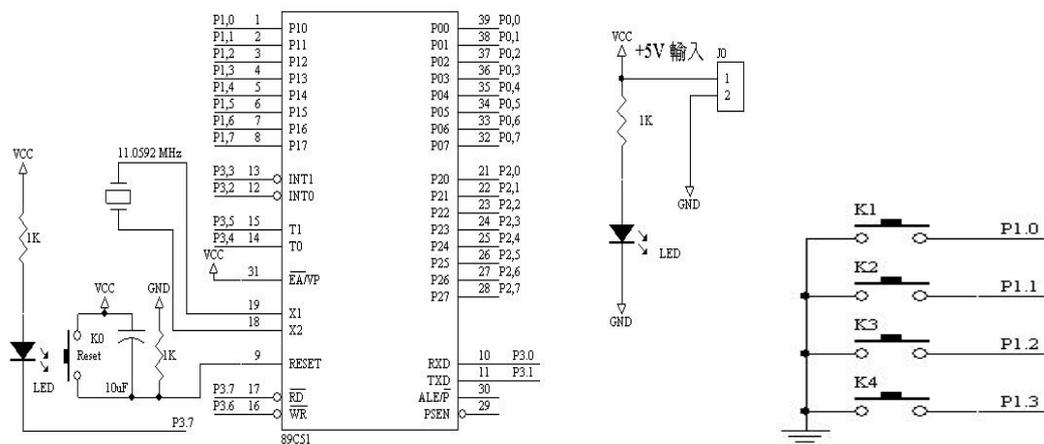


圖 5-5 8051 基本控制電路

在控制電路中，如果是按鍵數不多時可以使用一個按鍵對應一條輸入位元線控制，圖 5-5 包含一個 4 個按鍵的控制輸入電路，使用 8051 埠 14 條 I/O 線當做輸入用，由程式來控制，平時輸入端為高電位，當有按鍵按下時則相對位元會呈現低電位，經由輪流掃描判斷輸入端是否低電位，便可以知道按下了那一按鍵。

串列介面

VCMM 本身便可以做語音辨認實驗，若想進一步由單晶片 8051 設計聲控的功能也可以由外部連線來控制 VCMM 動作。所有應用主控系統都是經由串列介面來控制 VCMM 動作，進一步完成聲控的目的地。

圖 5-6 串列介面電路，由外部任何單晶片(如 8051) 經由串列介面來控制 VCMM，其電氣準位為 TTL 信號，皆可以由 J8(89C2051 串列介面 I/O 輸出接頭)連到 VCMM。以下是其控制連接方式：

外部單晶片	VCMM 89C2051
	J8
TXD	RXD
RXD	TXD
接地	接地

VCMM 串列介面聲控指令控制碼如下：

- 控制碼 't': 語音訓練，做語音參考樣本訓練輸入，一次訓練一組，展示

系統為 5 個辨認的單音。

- 控制碼 'm': 語音修改，修改原先已存在的語音參考樣本。
- 控制碼 'l': 語音聆聽，聆聽系統已存在的某組語音內容，展示程式為編號 0--4，重複循環。
- 控制碼 'r': 語音辨認。

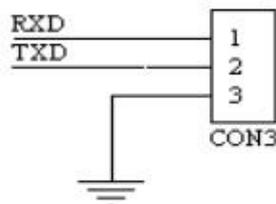
以上所有動作操作如同操作 VCMM 上的 4 個按鍵一樣。

8051 送出辨認控制碼'r'後，等待約 1 秒後，VCMM 送出控制碼 '@ab' 用以表示辨認結果，ab 為所辨認的語音樣本編號編碼，實際辨認結果編號為 no

$no=10xa+b$ no 有效值為 0--59

若辨認錯誤則會傳回'X'做回應。

例如辨認出語音樣本編號 1，則會傳回 '@01'，使用者可以再設計程式做進一步應用控制。



8051 串列介面接頭

圖 5-6 串列介面

紅外線發射及接收

一般紅外線遙控設計，為簡化製作，採用紅外線發射及高頻接收部分配對的組件設計，本實驗採用 SO315，發射器及接收器的高頻載波工作頻率約 315 MHz。加上密碼控制，可以有效完成紅外線簡易資料傳送。遙控編解碼分為發射端及接收端，內含 DIP 開關，可以提供組密碼設定，經由編碼晶片轉換為不同的數位資料，再經由高頻載波電路而發射紅外線資料出去。編碼晶片一般使用 HT12E 或是相容晶片。內含 8 組 DIP 開關，可以提供有 256 組密碼設定。

接收端由高頻接收電路、解碼 IC、DIP 開關及單晶片幾部份組成。高頻接收電路接收來自發射器送來的高頻信號，並取出其中的數位信號，而送至解碼 IC。解碼 IC 一般使用 HT12D 或是相容晶片做解碼，負責將原先編碼的資料進行解碼，取出原先發射的資料而送到單晶片上，密碼的設定

可由 DIP 開關加以調整，必須與發射機上的一致方可動作。單晶片控制程式則控制完成相對的按鍵遙控動作。

8051 控制紅外線接收模組電路，主要的元件為紅外線接收模組及數位解碼控制晶片 HT12D。為了方便實驗，紅外線發射及高頻接收部分使用配對的組件稱為 SO315，發射器及接收器的高頻載波工作頻率約 315 MHz。紅外線接收模組將接收發射器所送出的高頻載波信號，經由解調去除高頻載波信號，並取出還原數位編碼信號，再送由解碼 IC HT12D 負責解碼，取出原先的 4 位元資料，經由 8051 做進一步處理應用。

繼電器介面

圖 5-7 繼電器介面控制電路，一般在直流線圈的兩端都會加上一保護二極體，用以保護驅動輸出端的電晶體，因為在繼電器 ON OFF 之間，在線圈上會產生相當大的反電動勢，加上二極體便可迅速將此反向高壓吸收掉。當 8051 控制線 P3.4 送出低電位時，電晶體截止繼電器不導通 OFF，反之當控制線 P3.4 送出高電位時，電晶體則飽和使繼電器導通便啟動電氣連接迴路。

繼電器控制接點動作說明如下：

- L1 L2：線圈接點，加上工作電壓後可以聽見繼電器接點切換接通的聲音。
 - COM：Common，共通點。輸出控制接點的共同接點。
 - NC：Normal Close 常閉點。以 COM 為共同點，NC 與 COM 在平時是呈導通的狀態。
 - NO：Normal Open 常開點。NO 與 COM 在平時是呈開路的狀態，當繼電器動作時，NO 與 COM 導通，NC 與 COM 則呈開路（不導通）狀態。
- 繼電器一般的使用是串接在電氣迴路中，當作可程式控制的電源開關切換用，由繼電器 ON OFF 的動作，可以用來控制家電(AC 110V)開啟或關閉。在圖中當繼電器 ON 時，使電燈電源迴路接通，因此電燈會亮起。

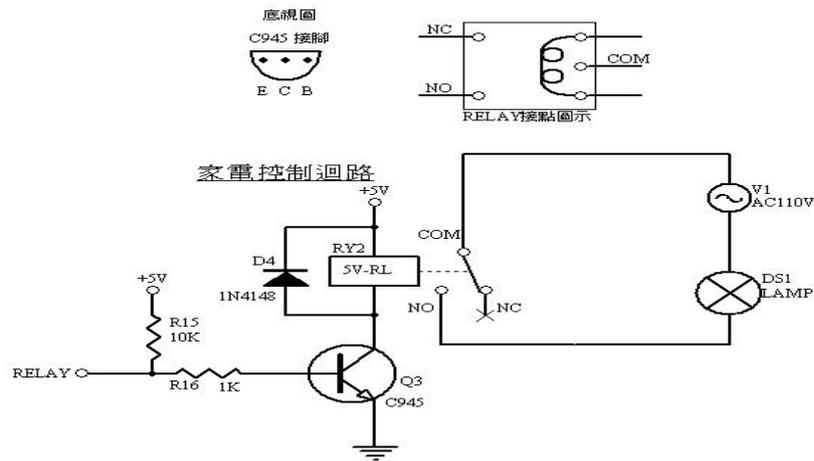


圖 5-7 繼電器介面

5-4 軟體流程圖

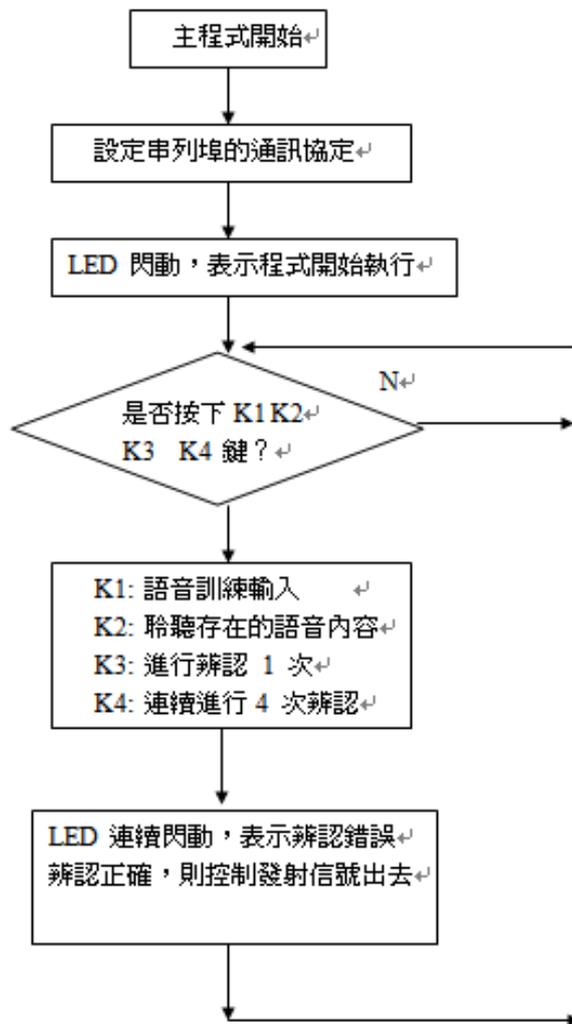


圖 5-8 發射端軟體流程圖

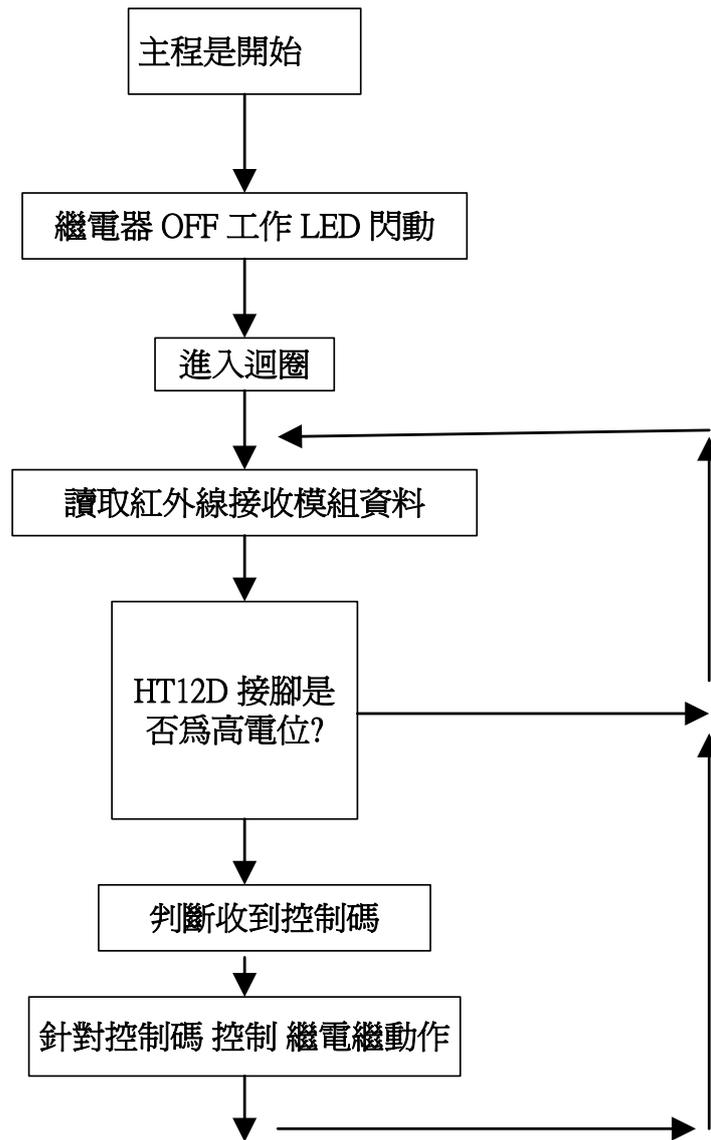


圖 5-9 接收端軟體流程圖

本專題使用 C 語言撰寫程式，請參閱附錄。

5-5 系統實作

在本專題中，利用單晶片 8051、按鍵、結合聲控模組、紅外線發射接收介面，進而遙控繼電器介面，來控制家電電源開關的功能。以 8051 c 語言控制程式來設計一套聲控系統，採用人工智慧特定語者語音辨認技術，用語音來控制想要受控的裝置，特定語者是指誰訓練後，該人來辨認會很準確。如是男生錄音內容，男生來辨認音調接近，可以直接控制，若不準確，則自行重新錄音便可以得到高辨認率。系統採用聲控模組做整合，測試錄音內容若為"電燈"，若說出"電燈"，則系統會回應電燈的語音登錄序號，此為辨認結果，發射此序號，到遠端接收。實驗結果，可以驗證以聲控模組做家電電源開關的功能，可以得到不錯的效果。本專題已經建立基本的聲控家電實驗平台，經由修改控制程式，可以應用在需要聲控的使用場合上。

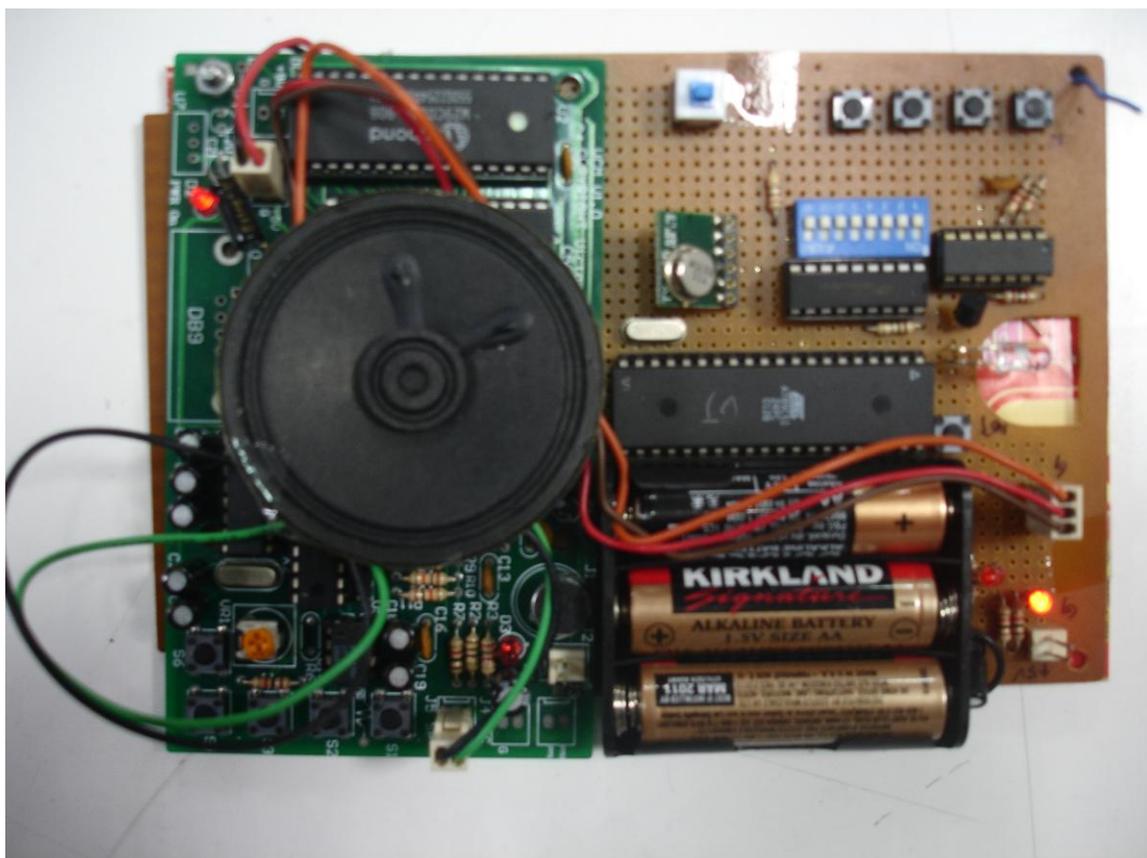


圖 5-10 紅外線聲控發射模組

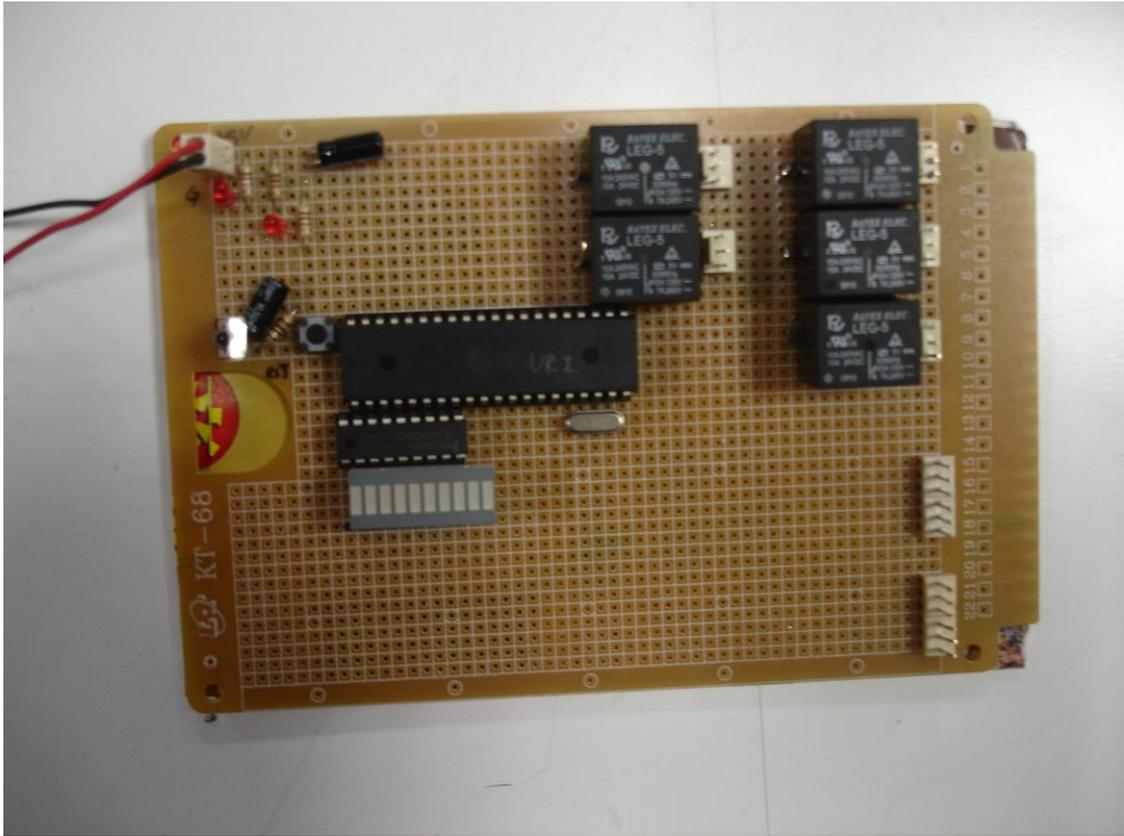


圖 5-11 紅外線接收模組



圖 5-12 聲控應用紅外線家電控制模組正視圖

第六章 結論

在本專題中，利用單晶片 8051、按鍵、結合聲控模組、紅外線發射接收介面，進而遙控繼電器介面，來控制家電電源開關的功能。以 8051 C 語言控制程式來設計一套聲控系統，採用人工智慧特定語者語音辨認技術，用語音來控制想要受控的裝置，特定語者是指誰訓練後，該人來辨認會很準確。如是男生錄音內容，男生來辨認音調接近，可以直接控制，若不準確，則自行重新錄音便可以得到高辨認率。系統採用聲控模組做整合，測試錄音內容若為"電燈"，若說出"電燈"，則系統會回應電燈的語音登錄序號，此為辨認結果，發射此序號，到遠端接收。實驗結果，可以驗證以聲控模組做家電電源開關的功能，可以得到不錯的效果。本專題已經建立基本的聲控家電實驗平台，經由修改控制程式，可以應用在需要聲控的使用場合上。

參考文獻

- [1] RSC-364 *Development Kit Manual* and RSC-364 *Data Sheet* ,Sensory Inc.
- [2] Furui,S., *Comparison of speaker recognition methods using statistical features and dynamic features. IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 342-350. 1981.
- [3] Matsui,T., & Furui,S. *Comparison of text-independent speaker recognition methods using VQ-distortion and discrete/continuous HMM's. IEEE Trans. Speech Audio Processing*, 2, pp. 456-459. 1994.
- [4] Matsui,T., & Furui,S. *A text-independent speaker recognition method robust against utterance variations. Proc. IEEE ICASSP (Toronto, Ont., Canada)*, pp. 377-380.1991, Apr..
- [5] Furui,S. *Vector-quantization-based speech recognition and speaker recognition techniques. Proc. IEEE ICASSP*, pp.954-958. 1991.
- [6] Y. Linde, A. Buzo, and R.M. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Trans. On Communication*, Com-28(1), pp. 84-95, 1980

- [7] A. Bendiksen and K. Steiglitz, "Neural Networks for Voiced/Unvoiced Speech Classification," IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Vol. 1, No. 90, pp. 521-524, 1990.
- [8] T. Ghiselli-Crippa, A. El-Jaroudi, "A Fast Neural Net Training Algorithm and Its Application to Voiced-Unvoiced-Silence Classification of Speech," IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Vol. 1, No. 91, pp. 441- 444, 1991.
- [9] IEEE TSAP, "Special Issue On Neural Networks for Speech", IEEE Transactions on Speech and Audio Processing, January 1994, Vol. 2, No. 1
- [10] Iso and T. Watanabe, "Speaker-independent word recognition using a neural prediction model", Proc. ICASSP, Albuquerque, NM, April 1990
- [11] Waibel and K-F. Lee, editors. "Readings in Speech Recognition", Vol. 1, Morgan Kaufman Publisher, Inc., San Mateo, California, 1990
- [12] 葉怡成, 類神經網路模式應用與實作, 儒林出版社, 1993.
- [13] 蘇木春, 張孝德, 機器學習 : 類神經網路、模糊系統以及基因演算法則, 全華科技圖書, 1997。
- [14] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. IEEE, Vol. 77, No.2, pp. 257-286, Feb. 1989.
- [15] 吳金成, 郭庭吉, 8051 C 語言應用與實習, 松崗電腦圖書資料股份有限公司出版, 1995 年3 月。
- [16] 陳明熒, 單晶片 8051 實作入門-專題製作篇, 初版, 文魁資訊股份有限公司, 2003 年 3 月。
- [17] 8051 單晶片 徹底研究 -- 入門篇, 林伸茂, 旗標出版
- [18] 單晶片 8051 KEIL C 實作入門, 陳明熒, 文魁出版
- [19] 8051/8751 原理與應用, 蔡朝陽, 全華書局
- [20] <http://www.chipware.com.tw>, 旗威科技有限公司
- [21] 微電腦專題製作應用電路, 鄧明發、陳茂璋, 知行出版社