Speech Driven User Interface For An Intelligent House

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Engineering
in
Information Engineering

at Massey University, [Albany],

New Zealand.

Zhenqing Liu

2005 March

Table of content

List of Figures

Abstract

Speech driven user interface for an intelligent house is one of a number of Graduate research projects at Massey University. It is part of Project 'Smart House'. This thesis details development of a control system whose inputs are speech signal rather than manual.

The control system consists of several sub-systems including speech recognition, command generation, signal transmission, signal reception and command manipulation. The completed speech driven user interface should operate in conjunction with Real-time implementation of a Microphone Array beam-former and Personal identity recognition that were developed concurrently with this project.

The speech recognition and command generation subsystems are implemented on a PC whereas the signal transmission, signal reception and command manipulation subsystems are designed at embedded board level. The remote controller can control some electrical appliances, such as TV and CD player, and switch and dim the light.

Acknowledgements

I wish to express my gratitude to the following people for their help during the course of this project. Special thanks go to my supervisor, Dr. Tom. Moir, who offered me this project, and who provided invaluable advice and motivated me to succeed. Thanks also to Joy. Wang, Tom Qi for their technical help in the process of building PCB board . Thanks to my parents for encouraging sponsoring me to get the M.S. degree. Thanks to my brother Dr. Zhenguo Liu for his technical help in process of designing the antenna. Thanks to Samir for his administrative help. This would not have been possible without the support of my friends and family.

Chapter 1 Smart house

1-1 Smart house overview

When people talk about "smart homes," "home automation," or "integrated home systems," they're talking about more or less the same thing—ways to automate and efficiently control various technologies around the house. It has been the dream of human being for a long time. Since computer engineering, electronic engineering and communication network engineering became available, we have seen a huge development. Here some projects supported by different universities are discussed in generally rather than in detailed. The purpose is to understand the wide meaning and know what they done so far.

1. Agent-based Intelligent Reactive Environments - a Research Group at the MIT Artificial Intelligence Laboratory. "AIRE" is dedicated to examining how to design pervasive computing systems and applications for people. To study this, aire designs and constructs Intelligent Environments (IEs), which are spaces augmented with basic perceptual sensing, speech recognition, and distributed agent logic."[51]

2. Aware Home Research Initiative at the Georgia Institute of Technology. The Aware Home Research Initiative (AHRI) is a focused research program and whose goal is to develop the requisite technologies to create a home environment that can both perceive and assist its occupants. The scope of the projects carried out within this program range from fundamental technical development to cognitive and ethnological studies that assess

the most appropriate and compelling technological strategies [51].

3. The MavHome Smart Home project is a multi-disciplinary research project at the University of Texas at Arlington. This focused on the creation of an intelligent home environment. The approach is to view the smart home as an intelligent agent that perceives its environment through the use of sensors, and can act upon the environment through the use of actuators [51].

4. Smart Medical Home at the University of Rochester's Center for Future Health. This project's overall goal is to develop an integrated Personal Health System, so all technologies are integrated and work seamlessly. This technology will allow consumers, in the privacy of their own homes, to maintain health, detect the onset of disease, and manage disease. The data collected inside the home will augment the data collected by physicians and hospitals. The data collection modules in the home will start with the measurement of traditional vital signs including blood pressure, pulse, respiration etc and work to include measurement of 'new vital signs', such as gait, behavior patterns, sleep patterns, general exercise, rehabilitation exercises, and more [51].

5. Stanford Interactive Workspaces Project is exploring new possibilities for people to work together in technology-rich spaces with computing and interaction devices on many different scales. This project is staffed by faculty and students from the Interactivity Lab, Software Infrastructures Group, and Graphics Lab [51].

From these projects, some general ideas about smart house can be achieved. The Intelligent house is a highly interactive environment that uses embedded computation to observe and participate in normal, everyday events. All kinds of sensors are able to detect and collect some technological parameters then transfer them to the computer. Through using different softwares, these parameters are processed and stored the current state of the intelligent environment. Speech and vision, rather than keyboards and mice, provide the main modes of interaction in smart house. Multimode integration increases the effectiveness of these perceptual technologies, for example, by using vision to aid speech understanding by recognizing facial expressions, lip movement, and gaze. As the core technologies, Speech recognition and image recognition enable to listen to people and observe what they do. People can speak with, gesture to, and interact with it in other complex ways. Smart sketching and design tools enable people to express their ideas most of them are from different conception. A robust agent-based software infrastructure supports the operation of these tools. A smart house system should master the following technical challenges, which are derived from MIT project oxygen [52]:

- ***pervasive***—it must be everywhere, with every portal reaching into the same information base;
- ***embedded***—it must live in our world, sensing and affecting it;
- ***nomadic***—it must allow users and computations to move around freely, according to their needs;
- ***adaptable***—it must provide flexibility and spontaneity, in response to changes in user requirements and operating conditions;

- ***powerful,*** *yet efficient*—it must free itself from constraints imposed by bounded hardware resources, addressing instead system constraints imposed by user demands and available power or communication bandwidth;
- ***intentional***—it must enable people to name services and software objects by intent, for example, "the nearest printer," as opposed to by address;
- ***eternal***—it must never shut down or reboot; components may come and go in response to demand, errors, and upgrades, but it as a whole must be available all the time.

In order to fulfill these challenges easily and systematically, some technologies should be applied to the smart house project. Generally, Device technologies, Network technologies, Software technologies and Perceptual technologies can be integrated together to accomplish the tough tasks in an intelligent environment.

However, one point should be noticed that the interface between human and the intelligent house should be unobtrusive and require no special interactions. Any complicated or inconvenient interface will cut down the value of smart house even that is excellent in other parts. It is deserved that inhabitants operate the house as they would an ordinary home using light switches, thermostats, and volume controls like what they are accustomed [1]. Furthermore, in contrast to existing automated homes that can be programmed to perform various functions, the real smart house should develop a home that essentially programs itself by observing the lifestyle and desires of the inhabitants and learning to anticipate their need. This means the house has the ability of adaptation, in other words, smart house can learn by itself. This house's intelligence lies in its ability to adapt its operation to accommodate the inhabitants. Thus, the adaptive ability is the core demand and technology.

1-2 Home Automation Network

Currently, it is still quite few appliances to use the services of a home communication network. A home automation network HAN allows these appliances to be integrated. [2] [3]. In this section, some theories are discussed. The home automation network HAN is focused.

Though there are many standard HANs, such as Cebus [4], Lonworks [5] and EIB [6], they are not only too expensive but also too complicated to install. Meanwhile, the protocols of these HANs are so different that they can not be interconnected in an intelligent house system. A good protocol can make the program and implementation easy and tidy. The article 'Home automation network supporting plug-and-play [7] discussed an open, low-cost and easy to use protocol to solve these problems. The protocol has five layers and supports two physical medias (twisted pair of bus TP and radio frequency RF). Actually, the wired and wireless include all medias in physical layers. A special feature is supporting plug-and-play (PnP), which means when a new appliance is plugged into HAN, the end user uses a hand-held remote controller (RC) to register it and can operate it immediately

The HAN protocol model conforms to the Open System Interconnection (OSI) model except for the session layer and the presentation layer. HAN implements the following five layers: physical layer, data link layer, network layer, transport layer, and application layer. Each layer provides services for its upper layer and calls services provided by its lower layer via service primitive. Fig.1 [7] shows two nodes communicate on a physical media, where node 1 sends a message to node 2. This message passes through the five layers in node 1 down to the physical media where node 2 receives and passes it through the five layers up to application software.

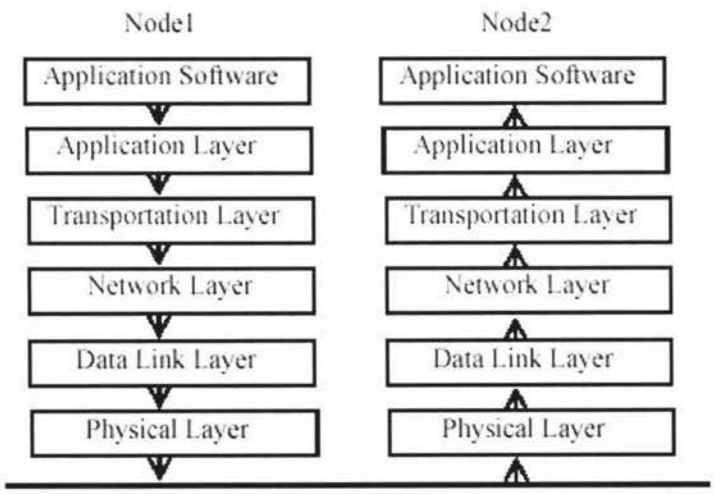


Fig.1 Communication flow in two nodes [7]

In the physical medium particular RF, the carrier frequency should be considered. Recently, several radio bands have been freed in Europe and allocated to so call ISM (Industrial, Scientific and Medical) applications needing communication ability over short distances (inside a building). It is easy and cheap to operate in these bands because no license is needed.

The 433.05 - 434.79 MHz band has been considered for the wireless Teledomotis medium, for reasons of cost, available bandwidth and propagation. This regulation restricts the power of propagation to 10mw and limits the valid distance in few tens of meters [8]. The experiment of propagation reveals that the obstacle fading is much lower at 433 MHz, and that the waves circumvent the obstacles much more easily at this frequency compared to other frequency such as 900MHz and 1.7GHz [9].

Furthermore, It ensues from the analysis that a transmitter can experience problems in reaching near receivers (fading hole), at the same time as being very disruptive for some receiver in a neighboring house. It therefore becomes almost indispensable to adjust the power of transmission according to the propagation conditions in order to limit the disturbance to other networks, right up to several hundred meters away [9] [10]. Power control, modification for data link layer and modification for network layer are applied and a whole new protocol stack has to be attached to each new wireless option in order to distinguish neighboring networks and to manage routing functions. Each of these modifications induces new constraints for the throughput and makes the throughput evaluation more complex. If these modifications for wireless solution are considered in the procedure of design HAN protocol described in section one, the reliability should be better.

In HAN, wireless communication is needed between device (DT) that has ability of telecommunication and network. All these electronic products have led to an increased demand for an effective and low cost approach for multi-user communications.

Multi-user spread spectrum (SS) communication [10] [11] system based on chaos theory for multimedia transmission in a wireless home network environment is proposed [12]. Using the ergodic property of chaos, the proposed chaotic SS system is shown to be able to modulate and transmit the multimedia signal effectively for point-to-point communication. Compared to the conventional CDMA (Code Division Multiple Access) system used recently, it is shown that the proposed system has improved and satisfactory performance in transmitting analog multimedia signals even when the channel measurement noise is strong.

DTs should be share the same radio medium in an intelligent house. The medium access method presented by Nadir Hakem and Michel Misson[13] describes a solution for wireless home networks which can share dynamically the radio channel without planning constraints as frequency or code allocation. It is very important for wireless home network in the ISM frequency band.

In conclusion, HAN with wired and wireless devices is the core of the intelligent environment. So the hierarchical design makes the implementation of intelligent environment easier.

Chapter 2 Speech signal processing

In this chapter we will generally talk about the speech signal processing which is used in speech recognition. Firstly beam-forming is discussed in Section-1 including the basic theory fundamental. Estimation methods will be talked in Section-2 and MUSIC (Multiple signal classification) will be discussed in depth. In section-3 and section-4 constant beam-width beam-forming and matched filter are the focus. In follows section, speech recognition will be discussed in generally.

2-1 Beam-forming

Conventional microphones have to be near the user at all times, forcing the user to either wear the microphone or have it move with the speaker (e.g., telephone, teleconferencing). Microphone beam-forming has eliminated the need for a movable microphone or telephone. Currently, there are ways to use many microphones to create beam patterns that will focus on one speaker in a room. The motivation for steerable microphones comes mainly from teleconferencing applications. In smart house, any inconvenience is not desired to meet especially using speech recognition. So microphone array as adaptive signal capturer will be used. However, the acoustic characteristics of a house required we deal with a lower SNR and have to change the beam direction often to catch up with the speaker, as well as requiring larger spatial coverage. Beam-forming, as a versatile approach to spatial filtering, is widely applied in radar, sonar and communication systems. The directivity pattern can be adapted to null out signals arriving from the specified direction. The selectivity of spatial direction is very valuable for propagation and reception in signal processing.

The concept of an adaptive antenna system was first thoroughly treated by Widrow et al. [14]. The motivation at that time was applications in radar and sonar. The two major problems associated with array design are direction-of-arrival estimation for sources, and subsequently creating an optimal output. Significant direction-of-arrival estimators came from Capon in 1969 [15], Schmidt in 1980 [16], and McDonough in 1983[17]. Howells developed the first step in adaptive interference canceling in 1965[18]. Since then, Griffith and Jim generalized the most prominent noise rejection structure: Sidelobe Canceler [19], which drew from the work of Levin [20] and Frost [21] and others. Similarly, the adaptive microphone array can be applied the concept from the adaptive antenna. However when the beam forming is applied in the speech signal processing by using microphone array, several remarkable considerations, with exception of two major problems touched before, should be taken into account according to the characteristics of speech signal. One is the beam width, which is the function of frequency. An approximate constant beam width is desired. The other is the multi-path. In a real reverberant environment, multi-path is very serious to distort the signal from the source. In the later two sections, these two key points are discussed and some methods to resolve it are proposed. First we discuss the theory of beam forming.

2-1-1 Theory of beam forming

John [46] gives the spatial pattern of the linear arrays of n isotropic point when excitation is equal in amplitude and spacing, shown in Fig.2-1. The total field E at far field is given by :

$$E = \sum_{i=0}^{n} \exp(jn\varphi) = \frac{\sin(n\varphi/2)}{\sin(\varphi/2)} \tag{2-1}$$

where $\varphi = \frac{2\pi d}{\lambda}\cos\phi + \delta$, which is the total phase difference of the fields from adjacent point source. The two parts contributed the total phase difference are phase difference of adjacent source δ, and phase difference according to spatial distance, $\frac{2\pi d}{\lambda}\cos\phi$. d is the adjacent distance of source and the ϕ is the angle between the array line and the desired direction. The mainlobe is $2\varphi_{01}$, where the φ_{01} is the angle which makes the E first equal to zero and the maximum of E is n when the total phase difference is zero. Duo to the duality of electromagnetic field, no matter antenna used in propagation or reception, the pattern is same. So if δ can be changed in advanced, the maximum direction can be got on the condition that $\frac{2\pi d}{\lambda}\cos\phi = -\delta$. Then the beam can be shifted and scanned.

Now we back to the microphone array. A line array of 2N+1 omni-directional microphones spaced equally by distance d is shown in Fig.2-1 [22].

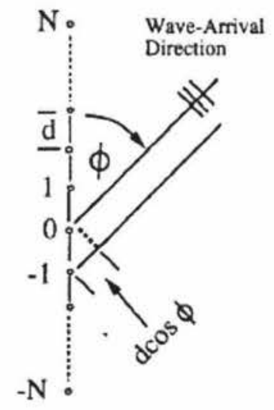


Fig.2-1 One dimensional microphone array with 2N+1 elements [22]

All microphones are assumed to have identical sensitivity. The impulse response for a plane wave from the direction of angle ϕ with the axis of line microphone is given as [23]

$$h_{(t)} = \sum_{n=-N}^{N} \delta(t + nT) \tag{2-2}$$

where t is time, δ is the Dirac delta function, T is the time difference adjacent microphone element by the spatial position. The frequency response of $h_{(t)}$ is easily got as

$$H_{(j\omega)} = \sum_{n=-N}^{N} \exp(j\omega nT) = \frac{\sin[(2N+1)\omega T/2]}{\sin(\omega T/2)} \tag{2-3}$$

The line array can be steered to a desired angle ϕ_s if the output of the n^{th} microphone is delayed by $nd\cos\phi_s / c$, where c is sound speed. Therefore the spatial response for a 2N+1 elements of microphone array steered angle ϕ_s is given as follow

$$H(\phi,\omega) = \frac{\sin[(2N+1)d\omega(\cos\phi - \cos\phi_s)/(2c)]}{\sin[d\omega(\cos\phi - \cos\phi_s)/(2c)]} \tag{2-4}$$

The beam width of main lobe is given as follows [9/7 1]

$$\theta_{BW} = 2\sin^{-1}(\frac{2\pi}{(2N+1)\omega d/c}) \approx \frac{4\pi}{(2N+1)2\pi d/\lambda} = \frac{2}{(2N+1)d/\lambda} \tag{2-5}$$

where the λ is the wavelength of sound. So the beam width is the function of frequency and linearly decreased when frequency increasing. In other words the higher adjacent distance ratio to wavelength, the narrower beam width of mainlobe gets. A pattern with 4 elements and half wavelength in adjacent distance is shown in Fig.2-3 as an example.

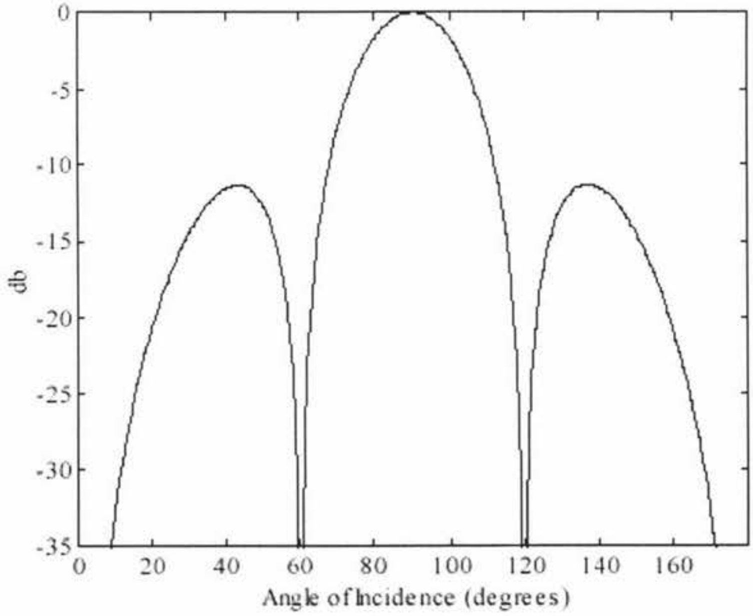


Fig.2-2 Normalized beam pattern for 4-element array with d/δ=0.5

In a narrowband beamformer, the affection of inverse proportion of beamwidth respected to frequency is generally insignificant [24]. However the beamformer applied to acoustic signal processing is broadband. So the affection is very serious in broadband situation.

If the all output of element sensor are weighted as binomial polynomial coefficient, then the sidelobe is disappear at the expense of increase of beam width of main lobe[antenna john]. This means the directivity of microphone array is inclined. However, Dolph-Chebyshev distribution gets the optimum relationship between the beam width and the sidelobe level. That means if a specified beam width id desired then the sidelobe has the minimum level and if a specified sidelobe level is desired then the beam with of mainlobe is minimum. The beam width is given by [24]

$$\theta_{BW} = 2\sin^{-1}[\frac{2}{\omega d / c}\cos^{-1}(\frac{1}{x_0}\cos\frac{\pi}{4N})] \tag{2-6}$$

where x_0 determines the sidelobe threshold and 2N+1 is the number of total elements.

2-2 Estimation of direction-of-arrival

To estimate the direction of sound source is the key point while using the microphone array. Maximum likelihood (ML) method is given in [26]. Difference among several methods is given for the estimation of direction [25]. We can find Multiple signal classification (MUSIC) has robust ability to identify the direction compared to other methods. Though CSP(cross-power spectrum phase analysis) can accurately estimate the DOA(delay of arrival) particularly when signals from different source are uncorrelated[27], and a proposed method based on CSP is described in [28] [29]. CSP only provides limited information from the received signals. So firstly, some discussion of MUSIC will be given as follows.

Firstly, we assume that there are M elements sensor with same sensitivity in a microphone array and D signal sources distributed in all space. During the propagation the noise is added to the signals. So the signal vector received from M elements are defined as follows:

$$X = AF + W \tag{2-7}$$

where A is MxD matrix and the element aij represents the function of signal arrival angles and the array element position. That is aij depends on the i^{th} element's position related to the reference of coordinate system and the response of i^{th} element from the j^{th} signal. F is D dimensional column vector. The noise appears as the vector W. The MxM covariance matrix of signal vector X is defined by

$$C = \overline{XX^*} = A\overline{FF^*}A^* + \overline{WW^*} = APA^* + \lambda C_0 \tag{2-8}$$

where $*$ means complex conjugate. When the number of incident wavefronts D is less than the number of array elements M, the APA^* is singular, say its determinant is zero. Then covariance matrix has D eigenvalue $\lambda_i + \rho$, the range of i is from 1 to D, and M-D eigenvalue ρ, which represents the power of noise. So when decomposing the covariance matrix, it is easy to get the number of signal souces according to the number of smallest and equal eigenvalue. The covariance matrix can be decomposed as follow

$$C = \sum_{i=1}^{D} \lambda_i V_i V_i^H + \rho \sum_{i=1}^{M} V_i V_i^H = \sum_{i=1}^{D} (\lambda_i + \rho) V_i V_i^H + \sum_{i=D+1}^{M} \rho V_i V_i^H \tag{2-9}$$

all vector composes M dimensional space and are orthogonal each other. The eigenvector V_i, i from D+1 to M, is orthogonal and composes a M-D dimensional noise subspace. So each column vector in matrix A containing some parameter from j^{th} signal will be orthogonal to the eigenvector in noise subspace. Then the expression shown as follow will be infinite

$$P = \frac{1}{\sum_{i=D+1}^{M} \left| A_j V_i^H \right|^2} \tag{2-10}$$

However, in practice the covariance matrix is estimated from the truncated data, the error is inevitable and the M-D eigenvalue may not be equal each other instead of being close. The expression (2-10) will not be infinite but a peak will appears. Compared to other methods applied to identify the direction of arrival of signals, Schmidt [25] gives the difference shown in Fig.2-3. From this figure we can find that MUSIC algorithm is better than other three methods. It provides asymptotically unbiased estimation of 1) number of signals; 2) directions of arrival; 3) strengths and cross correlations among the directional waveforms; 4) strength of noise and interference; 4) the frequency component of signals. All this is very important for the later processing of speech recognition. In addition, in [30], author compared MUSIC algorithm with other methods in distinguishing power spectrum and the former is better than auto-correlation based on AR model.

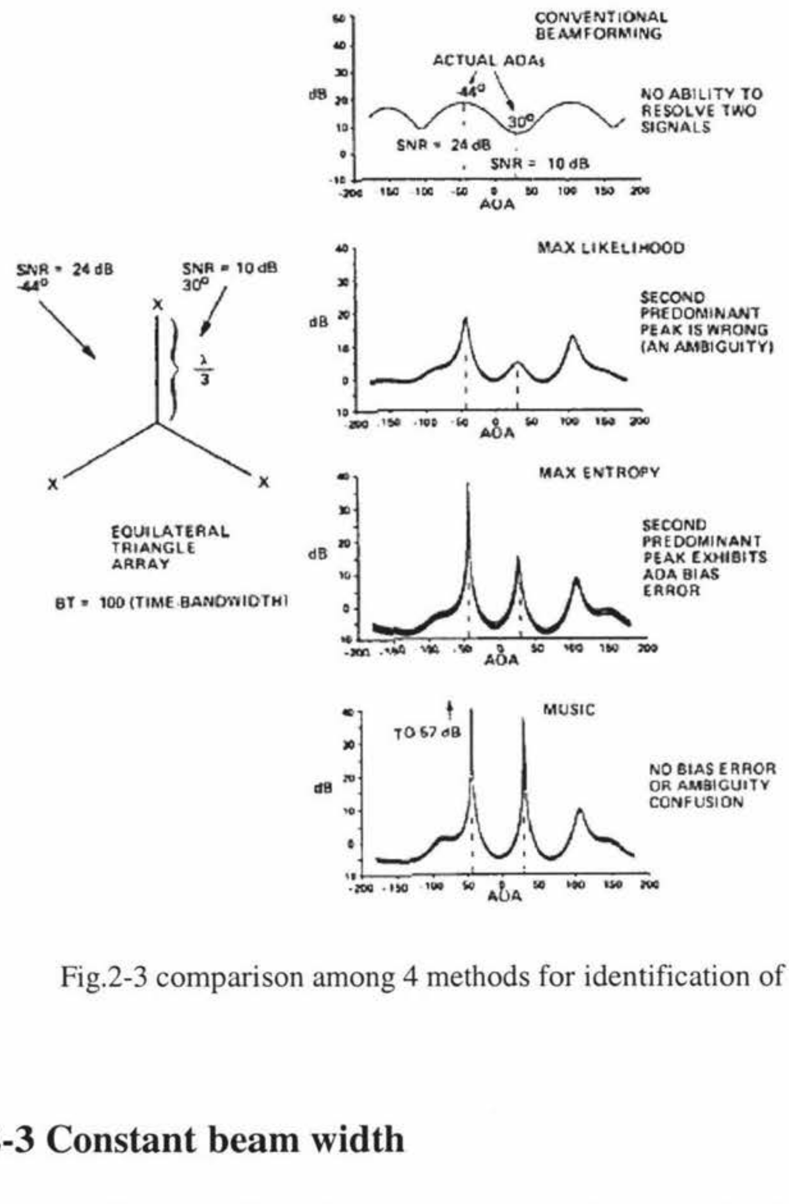


Fig.2-3 comparison among 4 methods for identification of signal's direction [25]

2-3 Constant beam width

Once the direction of speech signal is estimated by some algorithms, for example Music, the next step is speech recognition. No distortion during this procedure is desired. However, the speech signal is a broadband and the frequency range is from 60Hz to 4000Hz, spacing over four octave in frequency domain. From the expression of (2-5), the beam width is the function of frequency and linearly decreased when frequency increasing. So the magnificent change in beam width would seem to be unacceptable for the application of beam-forming in speech recognition. The reason is that the beam-forming, acts as a spatial filter, will distort the speech signal through the different attenuation depending on different frequencies. Fig.2-4 shows the difference in directivity pattern of a single beam-forming when two single frequencies are chosen.

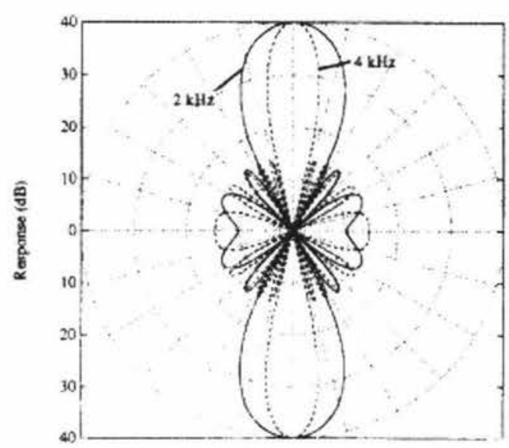


Fig.2-4 directivity pattern of two single-tone frequencies [24]

The dash represents the directivity pattern of a 4kHz single tone while the solid one is for the 2kHz single tone. It is very clean that both mainlobe and sidelobe are decreased sharply to be narrow when the frequency increasing from 2kHz to 4kHz. Furthermore, this distortion is more serious in near-field. In a real house environment, even if we change the microphone size, aperture size, the speech environment still should be considered as a near-field due to the architectural size of house. The person can not speak to a microphone array so far away, typically 3-5 meters in distance. So we must consider the distortion of beam width. A constant beam width is hoped.

Goodwin [24] and Khim [30] analyze the two different methods to accomplish a constant beam width beam-forming. They focused on the far-field compared to Pirz [31] who gave the method for near-field. A method of implementing FIR filter based the theory of ideal continuous aperture is presented by Darren B.Ward [32].

Through using these methods, an approximately constant beam width can be obtained. So far there are some different methods talking about constant beam width beamforming.

2-4 Matched filter

In a real acoustic environment, the speech signal received by the microphone array mixed up with ambient noise and multipath reverberant signals. This led the quality of speech signal too low, in other words low SNR. In order to improve the SNR, a matched filter is used because of its desirable property.

If a signal denoted as $S(t)$, which is assumed to be confined in the time interval $0 \le t \le T$, passes a LTI system and the system's response of impulse is $S(T-t)$, then this system is called Matched filter. The output from this system is given as follows

$$Y(t) = \int_0^T S(\tau)S(T-t+\tau)d\tau \qquad (2\text{-}10)$$

This equation is basically the time-autocorrelation function of the input signal $S(t)$. The output SNR obtained with the Matched filter is given as follow

$$SNR_{out} = \frac{2}{N_0}\int_0^T S^2(t)dt = 2\mathrm{E}/N_0 \qquad (2\text{-}11)$$

where E is the power of input signal, $N_0/2$ is the power spectral density of noise. So the output SNR from the matched filter just depends on the energy of the input signal waveform but not on the detailed characteristics of the input signal.

In practical speech signal processing, the matched-filter array is applied with the microphone array for the capturing sound from spatial volume in a noisy, reverberant environment. The principle of a matched filter array (MFA) is discussed in [33].

Now back to the model of a practically acoustic environment with multipaths. Assume that the speech signal is $S(t)$, the acoustic source to sensor pressure response is of the form

$$h(t) = \sum_{j=1}^{\infty} p_j \delta(t - \tau_j) \tag{2-12}$$

where Pj can thought as the attenuation of each multipath and τ_j is the delay. The purpose of MFA algorithm is to compose a number of filters each have the response of the time reverse of the corresponding source-to-sensor impulse response. For a sound source located at the focus of microphone array, the effect of the matched filter is to convolve the undistorted signal with the autocorrelation of the response of source-to-sensor. The output of each matched filter is of the form

$$y(t) = s(t) * h(t) * h(T - t) = s(t) * \sum_{j=1}^{\infty} p_j \delta(t - \tau_j) * \sum_{j=1}^{\infty} p_j \delta(T - t + \tau_j) \tag{2-13}$$

then the output of MFA is the summation of each individual matched filter. If we assume the total path of reflect is K but not infinite and one path for the direct. The attenuation of reflected path is zero. Then the output of MFA with N microphone element and N matched filter element is shown in Fig.2-5 [33].

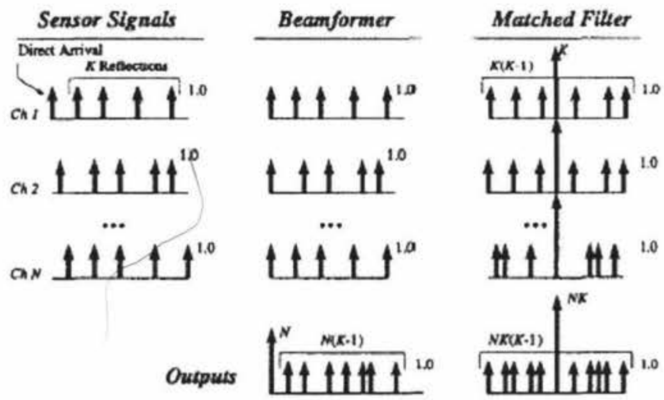


Fig.2-5 Alignment of captured signals using "delay-and –sum"

beamforming and using MFA processing [33]

The impulse response of each matched filter has a maximum value at time index 0 with amplitude K, and each has K(K-1) arrivals which are distributed in the time axis. By the summation of n matched filter elements, a main contribution with value KN at time index 0 and K(K-1)N contribution with unit value are got. The distinct advantage of MFA is the considerable improvement of SNR, which is shown as follows

$$SNR_{MFA} = \frac{P_{signal}}{P_{noise}} = \frac{(NK)^2}{NK(K-1)} = \frac{NK}{K-1} \qquad (2\text{-}14)$$

Compared to the system without using MFA, the SNR of a MFA system increases by K times.

In addition, truncation of the matched filter response is usually required because the amount of computation and intolerable system delays being equal to the room reverberation time. In [34] [35], some research is done about matched filter processing on the condition optimal truncation time and random distribution of microphone element in a real environment.

2-5 Speech recognition

As one of the main perceptual technologies, speech recognition [39] is applied to the intelligent house without any exceptions. If the head-set microphone is the sole equipment to catch the speech signal, then the inconvenience is the big problem because no matter where you are in the house you have to take it. Therefore, a microphone array is suggested to solve this problem at the expense of complexity and cost. Meanwhile vision recognition will be used as well as speech recognition in smart house in almost all projects.

However, background noise and room reverberations seriously degrade the performance of sound capture in real acoustical environment. So the method of beam-forming as forming a directive pattern sensitive to the target sound source is discussed in previous sections. Some practical methods and algorithms such as MUSIC and CSP [36] [37] are applied in the front end of speech recognition by using microphone array beamformer. In particular, accurate sound source localization becomes more important as the directive pattern is sharpened.

Once detecting the correct direction of a speaker, speech recognition automatically extracts what has been said. During this procedure, lots of methods are applied to improve the accuracy and reduce the degradation. Cross-spectral Estimators are used to enhance the quality of speech degraded by coherent and incoherent noise [38]. In some analysis, incoherence is assumed for noise, but in a real environment in which several people speak at the same time, these kinds of noise is not incoherent. So enhancement of speech degraded by coherent noise is much more important. Results shows that two-speech enhancement systems known to be suited for uncorrelated noises: the coherence function based technique and the Wiener-filter-based technique fail upon the correlation of noise component. However the cross power spectrum of correlated noise component is learnt during silent intervals (noise alone) to compute a more accurate power spectrum in the presence of speech.

As to how to localize the speaker in multiple sound sources, some excellent work has already been done by Taknobu Nishiura [40] [41].

The next task is speaker identification that automatically determines whether an utterance has been spoken by one of a population of speakers and to identify the talker. If the talker is known, in other word, is matched in voice characteristics with a person whose file is stored in a database, then the content what he or she said can be recognized. In fact the procedure is done simultaneously with the speech recognition. It seems so complicated, but this is true because so far the speech recognition has high accuracy only for a specified person.

The LPC (linear prediction coding) coefficient has been widely used in both speaker identification and speech recognition. HMM (hidden markov model), DTW (dynamic time warping and FSVQ (finite state vector quantization) are main methods used currently in speech recognition [39].

Now back to the localization of a sound source. Audio-only steering techniques often perform poorly in the presence of multiple sound sources or strong reverberation even though some enhancement techniques are applied. So vision-speech perception is the choice to achieve high quality of recognition particularly in a real intelligent environment. The experimented results shows that Video-only techniques can achieve high spatial precision but require that the audio and video subsystems be accurately calibrated to preserve this precision [42].

Audio-visual speech recognition is a novel extension of acoustic speech recognition and has received a lot of attention in the last few decades. The main motivation behind bimodal speech recognition is the bimodal characteristics of speech perception and production systems of human beings. The effect of the modeling parameters of hidden Markov models (HMM) on the recognition accuracy of the bimodal speech recognizer is analyzed. Acoustic features are merged with the visual features that carry relevant information about speech recognition and build a joint feature vector. During this procedure, different sampling rate should be taken into account. After that the task is to train the HMM model. Compared to acoustic only, the audio-visual speech recognition scheme has a much more improved recognition accuracy, especially in the presence of noise [43].

Chapter 3 ActiveX and LabVIEW

In this chapter we change our focus from speech recognition to ActiveX in depth because it acts the link between different programs and the key role in this Smart House project. With ActiveX Automation, a Windows application, such as Labview, provides a public set of objects, commands, and functions that other Windows applications can access. We can use Labview as an ActiveX client to access the objects, properties, methods, and events associated with other ActiveX-enabled applications. Labview also can act as an ActiveX server, so other ActiveX-enabled applications can access Labview objects, properties, and methods.

3-1 ActiveX

ActiveX is the general name for a set of Microsoft Technologies that allows users to re-use code and link individual programs together to suit their computing needs. Based on COM (Component Object Model) technologies, ActiveX is an extension of a previous technology called OLE (Object Linking and Embedding). The principle is that components need not be regenerated by each program, but rather, reused to fully give the user the power to combine applications together.

Therefore, to certain extent, ActiveX is a computing technology composed of several different components and each of them can perform a specialized task with common elements that are programmed similarly. All the components use methods that are, if not capable of being shared among each other, at least similar to each other. All use Object Linking and Embedding (OLE) technology to some degree to perform their tasks and are usually programmed in a language fine-tuned to Windows and OLE, such as Microsoft's Visual C++.

The fundamental element of ActiveX technology is the ActiveX control, formerly called an OLE control. An ActiveX control is a modular piece of software that performs tasks and computes information, communicates to other programs, modules. Some programming environment that can "contain" ActiveX controls, such as Web browsers like Internet Explorer or programming environments like Visual Basic, Labview, can be called as ActiveX container. Through container, an desired ActiveX control file can be embedded in it and easily entire and invoke other programs that are associated with the ActiveX control file. For instance, Excel has powerful capability to process table and charts, while Labview is robust in measurement and data sampling and analysis but not good enough to generate table and charts. A good method to combine these two merits both Excel and Labview is to use ActiveX control that is associated with Excel in Labview and then transfer the measurement data or analysis data to Excel to generate table or charts in Excel.

Currently, ActiveX technology is widely used in some fields related to the Internet and allows us innovative programs and Web programs that we can explore with Internet Explorer and other programs available from Microsoft. Multimedia play online is a good example. When we explore the internet by using Internet Explorer or other explorers, some ActiveX control are built in the web page such as the Real Player's ActiveX control. Consequently, without downloading related multimedia files, we can listen or/and watch.

So, with ActiveX technologies, Windows programmers will have a much easier time combining traditional Internet connectivity programs with powerful desktop software packages, and Internet users will find it easier to use the Internet to create business documents and use Web browsers to perform computing tasks. Some components add computing power to Web pages, others allow document-objects to call viewing programs from across the Internet, and others add new capabilities to Web servers. Furthermore information can be referred in books [44] [45].

3-2 LabVIEW

LabVIEW (Figure 3.1) is a program development application based on the graphical programming language, G, developed by National Instruments. It is designed primarily for test and measurement purposes, making it useful as an interface to the data acquisition (DAQ) hardware. LabVIEW allows developers to create programs, called virtual instruments (VIs) to recreate the appearance and functionality of real instruments such as amplifiers and filters.

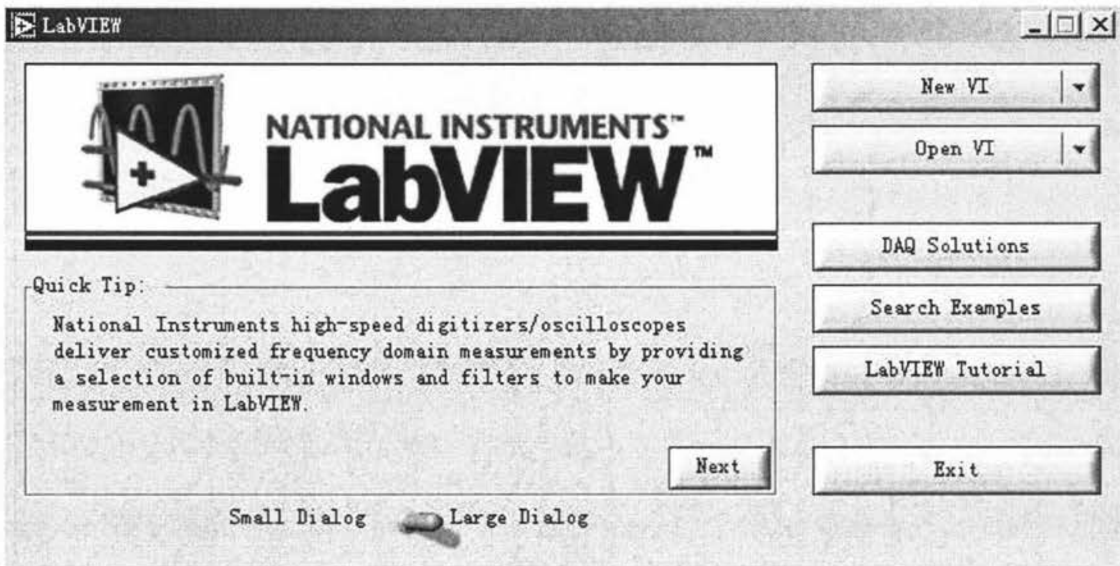
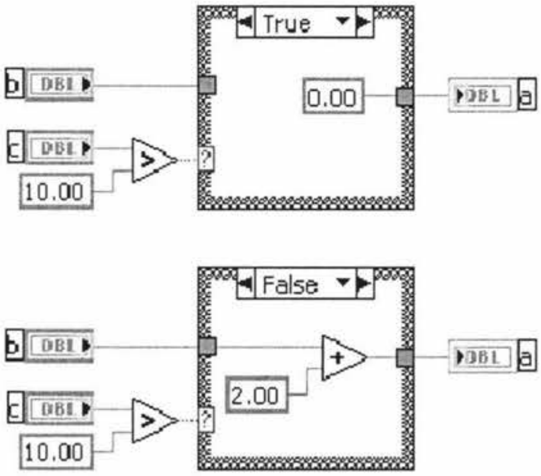


Fig.3-1 The interface of LabVIEW 6.0

Data objects are represented by blocks that are linked together with wires on a block diagram rather than lines of text. This form of programming allows developers with limited programming experience to create simple programs that perform useful tasks.

However, an experienced software developer would take longer to implement most algorithms in G than they would in a text-based language. G code also tends to be more difficult to follow at a glance than its text-based equivalent because many G structures consist of multiple sub-diagrams that cannot be observed at the same time. Figure 3.2 gives an example of this.




```
if c > 10
    a = 0;
else
    a = b + 2;
end;
```

LabVIEW structure MATLAB language

Fig.3-2 Comparison between LabVIEW and MATLAB for "IF"structure

In section 3-1 we discussed the ActiveX technology and known that ActiveX makes access, invoke and communicate with other programs possible and much easy. Although ActiveX encompasses a very broad range of technologies, the only one that is supported in LabVIEW is ActiveX Automation. ActiveX Automation in LabVIEW will be discussed now.

ActiveX Automation allows one program (the client) to control another (the server), but an Automation server is generally not embedded in the client application. The client simply calls the server like an ordinary function, and it must wait for the server to finish its task before it can continue. Version 5.0 and up of Labview offeres support for ActiveX automation as a server as well as client ActiveX Events were supported as well from the release of LabVIEW 5.1 and upwards.

In the Smart House project, ActiveX is applied to the link between Dragon naturally speaking and LabVIEW. By using it, LabVIEW allows a speech signal to trigger some VIs to accomplish some desired event. Then, we will discuss some other applications about Activex in the Smart House project. So far, we have the software of Dragon naturally speaking performing the task of speech recognition. In order to use this software to control other software which can achieve part of the intelligence in smart house, LabVIEW is chosen because ActiveX refers to the process of controlling one program from another via ActiveX. Much like networking, one program acts as the client and the other as the server. LabVIEW supports automation both as the client and the server. It means that programs (client and server) exist independent of each other, but they are able to share information between each other. This sharing of information is achieved through communication of the automation client with the ActiveX objects that the automation server exposes. The objects have properties and methods that can be accessed by the automation client. Properties are simply attributes of an object whose values can be set or retrieved from other programs. Similarly, methods are functions that are performed on objects, and they can be invoked from other programs. An example of an ActiveX property would be the program name, height or width, and likewise, an example method could be the save or print method.

The most common usage of ActiveX is via ActiveX controls, which are embeddable components that exist inside ActiveX Containers. Any program that can function as an ActiveX container will allow users to "drop" ActiveX controls into them. From these containers, the ActiveX controls have their own functionality and properties. LabVIEW is an ActiveX container and can house ActiveX controls. Again, the embedded control is manipulated via properties and methods.

ActiveX is an event-driven technology. This means that programs report when certain events occur so that other programs or the user can react accordingly. So we can create some events such as speech or browse from internet to trigger some sever programs. The choices of for sever programs in LabVIEW is very abundant. The more detailed information is given in a tutorial of LabVIEW.

Chapter 4 Speech driven Universal remote controller

Introduction

As a part of the project 'SMART HOUSE', a speech driven universal remote controller is needed. The requirement for it is that almost all electrical appliances which use remote controller, such as TV, Hi-Fi, home theatre etc should be able to be controlled by it. In the 'SMART HOUSE' project, all commands are from human speech rather than the action of a hand. DRAGON recognize what the human said, then a program written in LABVIEW generates the related binary string which is the same as what the normal remote controller creates when some buttons in it are clicked. Finally through the RS232 port, the binary strings are transmitted to a microcontroller and the corresponding pulses modulated in 36kHz or 40 kHz depending on the IR protocol are propagated. In addition, the command to control a light dimmer is modulated in 433.92 MHz AM carrier. After demodulating AM RF signal, the light can be controlled in switch mode or dimming mode.

4-1 Software Interface

In this project three software are utilized on two different hardware platforms (PC and remote controller board with microcontroller) that each require a means to communicate with each other. LabVIEW and Dragon Naturally Speaking are operated in the PC, whereas the assembly programming is embedded in the microcontroller. The discussion of two software application except of LabVIEW is provided as follow.

4-1-1 MPLAB IDE

MPLAB IDE is an Integrated development environment provided by Microchip company to support programmer developing PIC series microcontroller particularly in embedded system applications. MPLAB IDE v6.xx is a Windows-based Integrated Development Environment for the Microchip Technology Incorporated PIC microcontroller (MCU) and dsPIC digital signal controller (DSC) families. In the MPLAB IDE, functions can be done:

- Create source code using the built-in editor.
- Assemble, compile and link source code using various language tools. An assembler, linker and librarian come with MPLAB IDE. C compilers are available from Microchip and other third party vendors.
- Debug the executable logic by watching program flow with a simulator, such as MPLAB SIM, or in real time with an emulator, such as MPLAB ICE. Third party emulators that work with MPLAB IDE are also available.
- Make timing measurements.
- View variables in Watch windows.
- Program firmware into devices with programmers such as PICSTART Plus or PRO MATE II.

For PICmicro MCU's	For dsPIC Devices
MPLAB SIM Simulator	dsPIC SIM30 Simualtor
MPASM Assembler	MPLAB ASM30 Assembler
MPLINK Object Linker	MPLAB LINK30 Linker
MPLIB Object Librarian	MPLAB LIB30 Archiver/Librarian
MPLAB C17 C Compiler MPLAB C18 C Compiler	MPLAB C30 C Compiler
MPLAB ICE 2000 Emulator	MPLAB ICE 4000 Emulator
MPLAB ICD 2 Debugger	MPLAB ICD 2 Debugger
PICSTART Plus Development Programmer	
PRO MATE II Device Programmer	
Third Party Tools	Third Party Tools

Table 4-1 MPLAB IDE V6.0 provides development tools [49]

4-1-2 Dragon Naturally Speaking

Dragon NaturallySpeaking software is the premier product for large-vocabulary continuous recognition of natural speech. With Dragon Naturally Speaking, documents can be created quickly and easily with voice, without being slowed by typing on keyboard. The interface of Dragon NaturallySpeaking is shown in Fig.4-1.The goal of Dragon NaturallySpeaking is to make it easier to interact with a computer without using a keyboard or mouse. To succeed at this goal, the program must be able to analyze an incoming stream of sounds and interpret those sounds as commands and dictation. This process of interpretation is called speech recognition, and its success is measured by the percentage of correct interpretations, or recognition accuracy.

To achieve high recognition accuracy, Dragon NaturallySpeaking relies on several sources of information:

- Acoustic model—a mathematical model of the sound patterns used by the speaker's language.
- Vocabulary—a list of words that the program can recognize. Each word in the vocabulary has a text representation and a pronunciation.
- Language model—statistical information associated with a vocabulary that describes the likelihood of words and sequences of words occurring in the user's speech.

When user creates and trains Dragon NaturallySpeaking, user starts with a standard set of models and then customizes them for the way user speak (acoustic model) and the way user use words (vocabulary and associated language model). After training, during using Dragon NaturallySpeaking each time, the program loads customized user files that covers three models information mentioned above to guess the words that user said. So the training will influence to the accuracy of recognition to great degree. A good point is that as user subsequently use Dragon NaturallySpeaking, misrecognitions may still encounter occasional, though these misrecognitions will decrease over time.

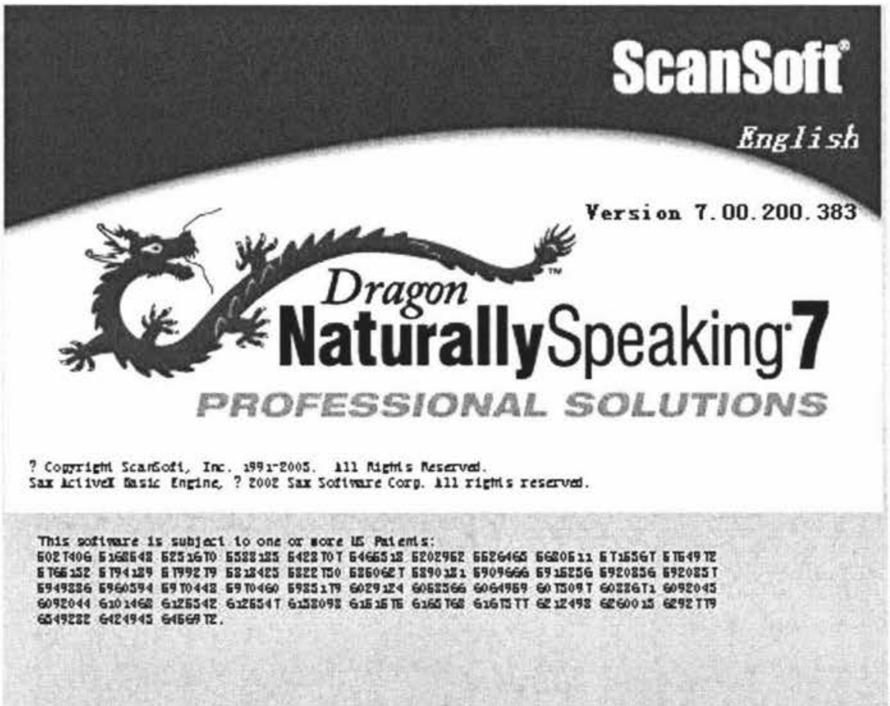


Fig.4-1 Interface of Dragon Naturally Speaking 7

Powerful commands in Dragon NaturallySpeaking can control the keyboard and mouse on the desktop, for example moving a mouse pointer anyway in the screen and clicking and dragging the mouse for moving or copying objects. Moreover, Dragon NaturallySpeaking can work in other programs such as Microsoft Word, E-mail Programs, even Internet Explorer. All this software can be performed and operated by voice rather than by hand. This perfect characteristic is what we want in the smart house. Although Dragon NaturallySpeaging can not directly generate the IR and RF signal, labview can work as a sever when ActiveX Automation of Dragon NaturallySpeaking worked as events. This means Dragon NaturallySpeaking can trigger LabVIEW and run some VIs files.

So the speech driven universal remote controller can be divided into two parts in line with the function. One is the command generation part and the other one is the command manipulation part. The former consists of Dragon NaturallySpeaking, LabVIEW, and works in PC platform while the latter consists of microcontroller embedded system that can translate commands received from PC to corresponding pulse depending on different IR protocols and propagate RF signal to around area..

4-2 IR Protocols

In order to complete universal remote control using IR LED, first the different IR protocols should be researched. So far there are lots of IR protocols to be used in different brands of electrical appliances. It is impossible to cover all kinds of IR protocols in one universal remote controller. So just some conventional protocols are considered in this project such as the Sony protocol and the Phillips RC-5 protocol [48].

The Sony protocol has 3 versions: 12-bit (described on this page), 15-bit and 20-bit versions. In this project 12-bit is considered.

- 5-bit address and 7-bit command length (12-bit protocol)
- Pulse width modulation
- Carrier frequency of 40kHz

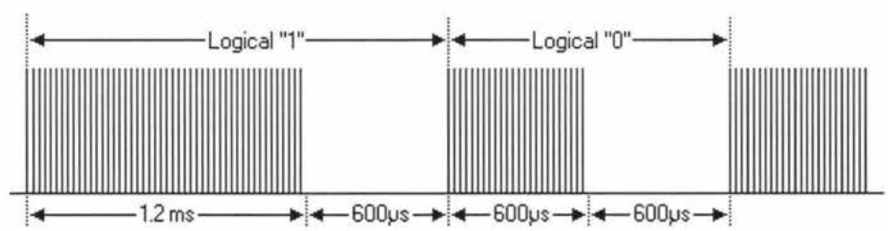


Fig.4-2 SIRC protocol bit structure [48]

The SIRC protocol uses a pulse width encoding of the bits. The pulse representing a logical "1" is a 1.2ms long burst of the 40kHz carrier, while the burst width for a logical "0" is 0.6ms long. All bursts are separated by a 0.6ms long space period. The recommended carrier duty-cycle is 1/4 or 1/3.

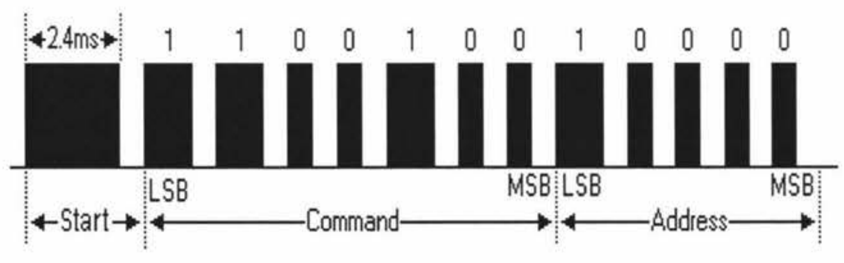


Fig.4-3 A typical SIRC protocol IR command waveform [48]

The picture above shows a typical pulse train of the SIRC protocol. With this protocol the

LSB is transmitted first. The start burst is always 2.4ms wide, followed by a standard space of 0.6ms. Then the 7-bit Command is transmitted, followed by the 5-bit Device address. In this case Address 1 and Command 19 is transmitted.

The protocol is well defined for different device types ensuring compatibility with your whole entertainment system. Lately Philips started using a new protocol called RC6 that has more features.

- 5 bit address and 6 bit command length
- Bi-phase coding (Manchester coding)
- Carrier frequency of 36kHz
- Bit time of 1.8ms
- Manufacturer Philips

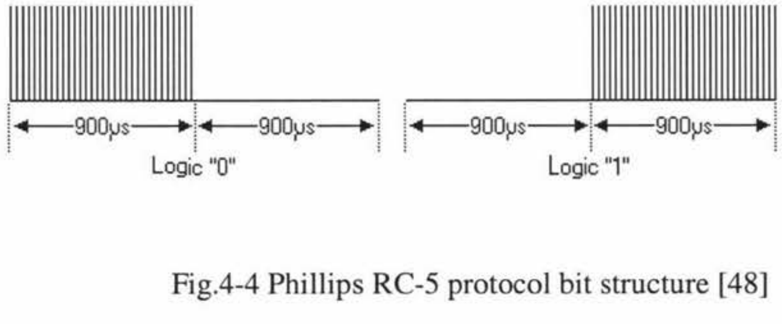


Fig.4-4 Phillips RC-5 protocol bit structure [48]

The protocol uses bi-phase modulation (or so-called Manchester coding) of a 36kHz IR carrier frequency. All bits are of equal length of 1.8ms in this protocol, with half of the bit time filled with a burst of the 36kHz carrier and the other half being idle. A logical zero is represented by a burst in the first half of the bit time. A logical one is represented by a burst in the second half of the bit time. The pulse/pause ratio of the 36kHz carrier frequency is 1/3 or 1/4, to reduce power consumption.

Fig.4-5 A typical Phillips RC-5 protocol IR command waveform [48]

The figure above shows a typical pulse train of an RC5 message. The first two pulses are the start pulses, and are both logical "1". Please note that half a bit time is elapsed before the receiver will notice the real start of the message.

Extended RC5 uses only one start bit. Bit S2 is transformed to command bit 6, providing for a total of 7 command bits.

The 3d bit is a toggle bit. This bit is inverted every time a key is released and pressed again. This way the receiver can distinguish between a key that remains down, or is pressed repeatedly. But in this project, this bit is maintain as "1" and is not necessary to change because just one command is generated by speech simultaneously.

The next 5 bits represent the IR device address, which is sent with MSB first. The address is followed by a 6-bit command, again sent with MSB first.

4-3 Microcontroller

To choose the microcontroller is not easy job according to the desired function. In this project, the main work that the microcontroller does is timing and reception of data from the PC. Although the 89c2051 series MCU has a function to reload a value in the timing mode of the Timer, which can make very accurate timing, lack of PWM function makes generation of pulse train very difficult. The reason is that a period of 36kHz is about 26 μs, considering 1/3 duty-cycle, about 8 μs is burst. This so short time is about duration of 8 instruction when the crystal is 11.0592mHz and impossible to execute some instruction particularly in the edge of a bit. Even the Attiny28L has the feature of an internal calibrated RC-5 oscillator, not only the RC-5 protocol but also Sony SIRC protocol should be covered. So finally the PIC16F627 is chosen to be the MCU. The features can be gain in the microchip website [49] and shown as follow:

		PIC16F627A	PIC16F628A	PIC16F648A	PIC16LF627A	PIC16LF628A	PIC16LF648A
Clock	Maximum Frequency of Operation (MHz)	20	20	20	4	4	4
Memory	Flash Program Memory (words)	1024	2048	4096	1024	2048	4096
	RAM Data Memory (bytes)	224	224	256	224	224	256
	EEPROM Data Memory (bytes)	128	128	256	128	128	256
Peripherals	Timer module(s)	TMR0, TMR1, TMR2	TMR0, TMR1, TMR2	TMR0, TMR1, TMR2	TMR0, TMR1, TMR2	TMR0, TMR1, TMR2	TMR0, TMR1, TMR2
	Comparator(s)	2	2	2	2	2	2
	Capture/Compare/PWM modules	1	1	1	1	1	1
	Serial Communications	USART	USART	USART	USART	USART	USART
	Internal Voltage Reference	Yes	Yes	Yes	Yes	Yes	Yes
Features	Interrupt Sources	10	10	10	10	10	10
	I/O Pins	16	16	16	16	16	16
	Voltage Range (Volts)	3.0-5.5	3.0-5.5	3.0-5.5	2.0-5.5	2.0-5.5	2.0-5.5
	Brown-out Reset	Yes	Yes	Yes	Yes	Yes	Yes
	Packages	18-pin DIP, SOIC, 20-pin SSOP, 28-pin QFN	18-pin DIP, SOIC, 20-pin SSOP, 28-pin QFN	18-pin DIP, SOIC, 20-pin SSOP, 28-pin QFN	18-pin DIP, SOIC, 20-pin SSOP, 28-pin QFN	18-pin DIP, SOIC, 20-pin SSOP, 28-pin QFN	18-pin DIP, SOIC, 20-pin SSOP, 28-pin QFN

Table 4-2 PIC16F627A/628A/648A Family Device Feature [49]

The high performance of the PIC16F627A/628A/648A family can be attributed to a number of architectural features commonly found in RISC microprocessors. To begin with, the PIC16F627A/628A/648A uses a Harvard architecture, in which program and data are accessed from separate memories using separate busses. This improves bandwidth over traditional Von Neumann architecture in where program and data are fetched from the same memory. Separating program and data memory further allows instructions to be sized differently than 8-bit wide data word. Instruction opcodes are 14-bits wide making it possible to have all single word instructions. A 14-bit wide program memory access bus fetches a 14-bit instruction in a single cycle. A two-stage pipeline overlaps fetch and execution of instructions. Consequently, all instructions (35) execute in a single-cycle (200 ns @ 20 MHz) except for program branches.

The architecture diagram of PIC16F627A is shown in Fig.4-6.

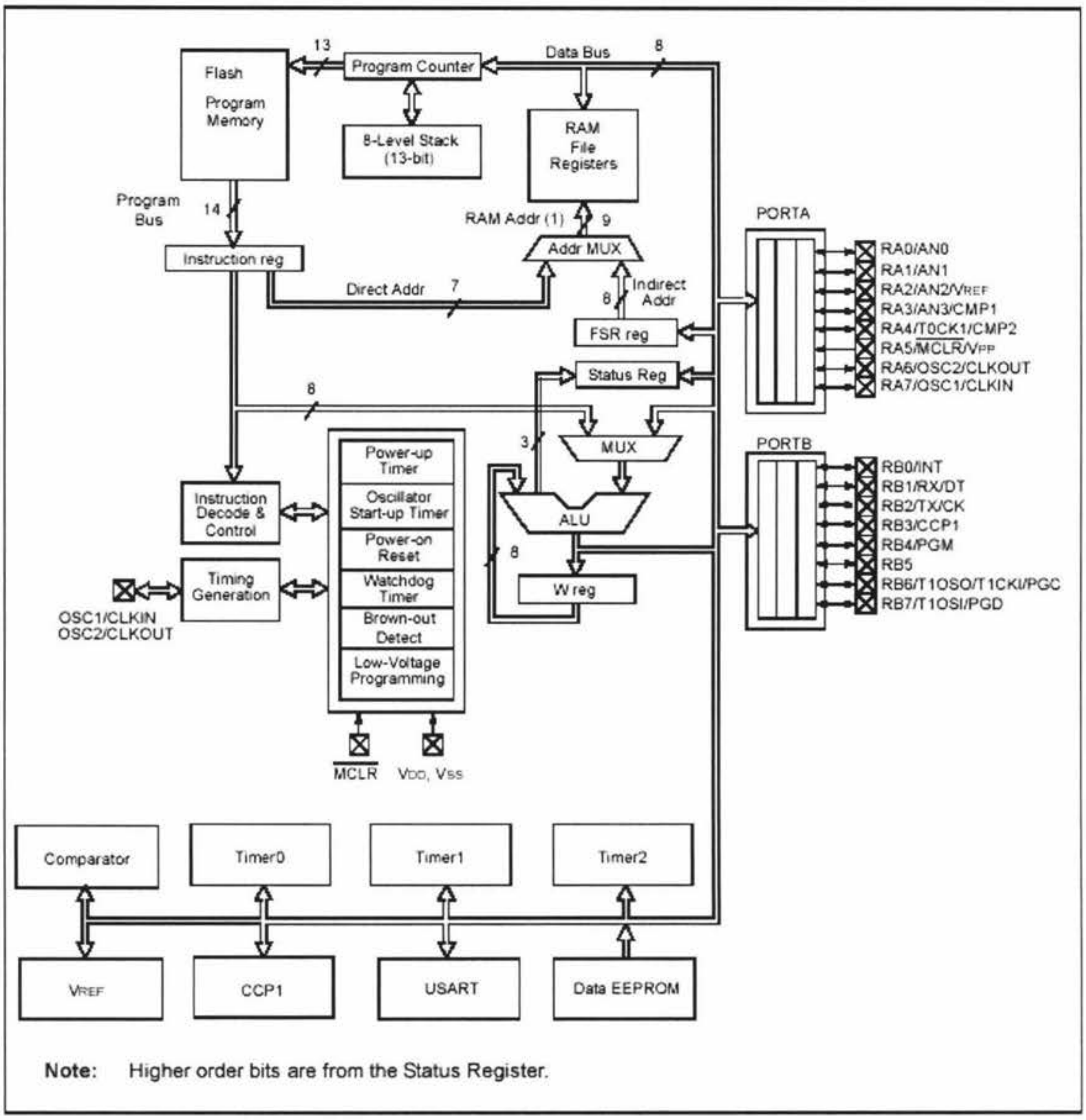


Fig.4-6 Architecture block diagram of PIC16F627A/628A/648A Family Device [49]

As shown from the block diagram, two ports (PortA and PortB) is provided by this Microcontroller family and each pin has multiple functions except for RB5. According to the preliminary design, the IR command signal is generated by CCP1 that works in the PWM mode so the pin RB3 is occupied. Then the least significant bit in PortA (RA0-RA3) is connected to the HT-12E as the address bit for appliances here especially for the dimmer. RA6, RA7 are two pins to connect the crystal supported the oscillate clock source. Although the communication between microcontroller and PC is single direction without using shaking hand protocol, RB2 is suggested not to use as other function except for transmission. RB1 is used as the receive signal from RS-232 port and RB5 is connected to the pin of transmit enable of HT-12E. Table 4-3 shows the pin function of PIC16F627A among PortA and PortB.

RA0-RA3	Address bit connecting to HT-12E A0-A3 respectively
RA4-RA5	Not using
RA6-RA7	Connects to crystal
RB0	Not using
RB1	RX receiving command from PC
RB2	Not using
RB3	PWM output pin
RB4	Not using
RB5	Control the transmit enable pin of HT12-E
RB6-RB7	Not using

Table 4-3 Each pin's function of PortA and PortB for PIC16F627

The MPLAB IDE unfortunately can not support the development for the PIC16f6XX serial family microcontroller (refer to Section 4-6-2 Programming the MCU) directly.

Finally the new paragraph of PIC18F452 was chosen in the final circuit board. The pin diagram is shown in Fig.4-7 below.

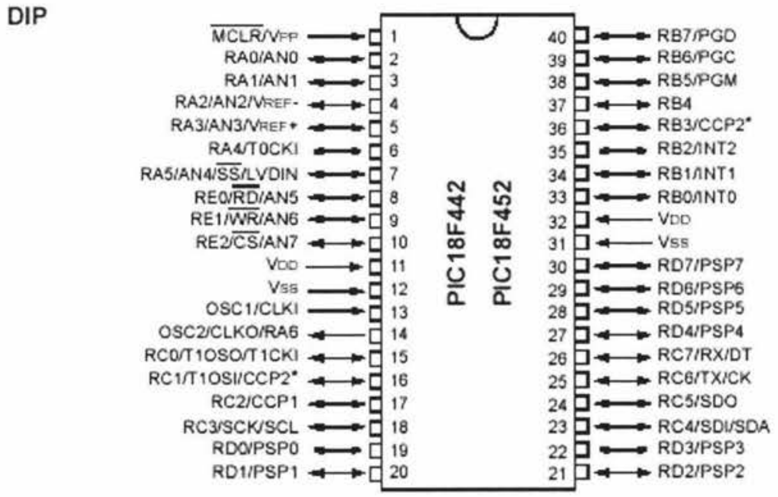


Fig.4-7 Pin diagram of PIC18F452 [49]

Compared to the PIC16FXX serial family microcontroller, the PIC18f452 is a much more powerful microcontroller with more I/O port and A/D converter channel. The architecture diagram is shown in Fig.4-8 below.

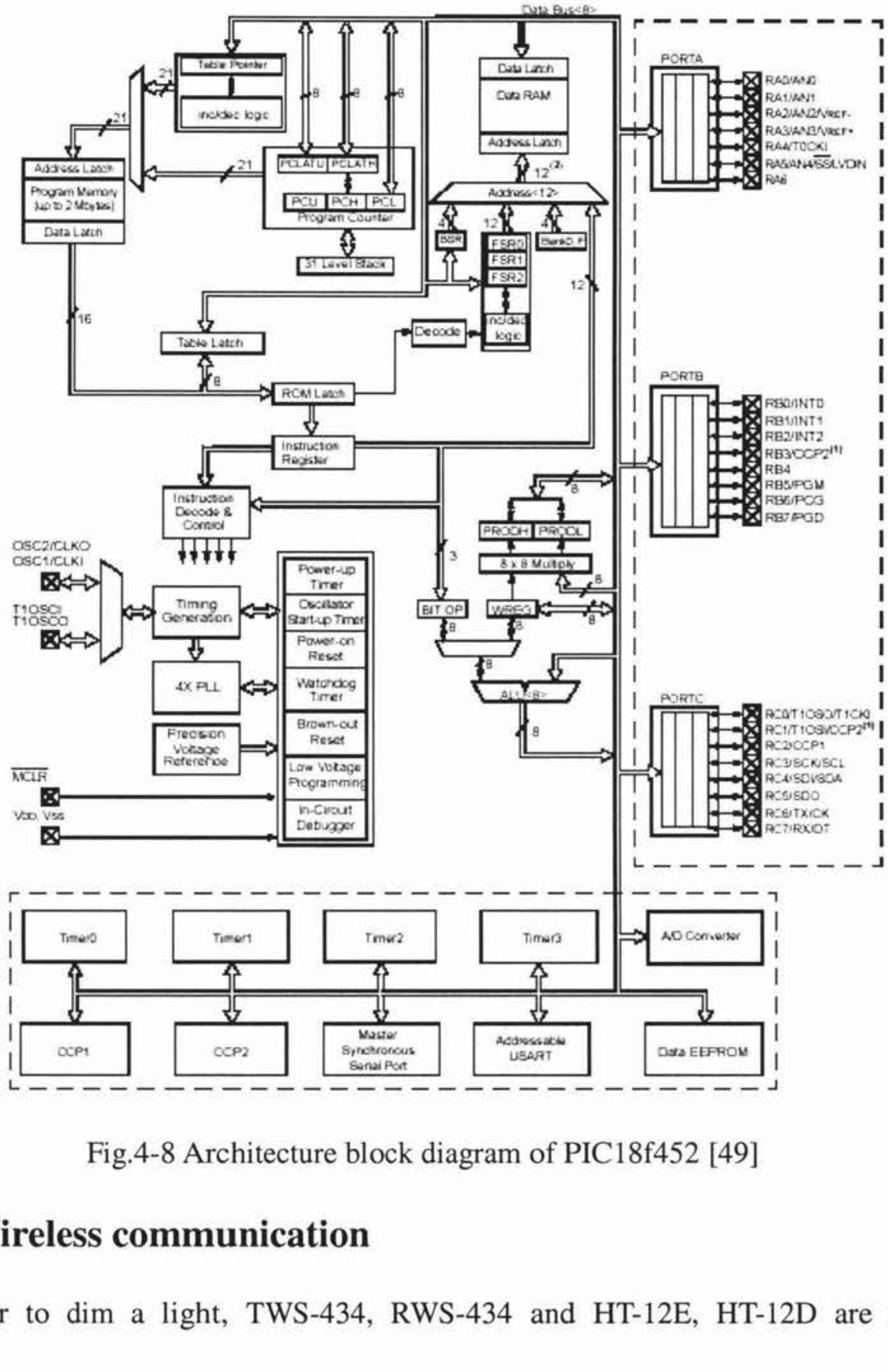


Fig.4-8 Architecture block diagram of PIC18f452 [49]

4-4 Wireless communication

In order to dim a light, TWS-434, RWS-434 and HT-12E, HT-12D are used as a transmitter, receiver encoder and decoder respectively [50]. HT-12E is an addressable encoder with 8-bit address and 4-bit data. Each address/data can be set to one of the two logic states. HT-12D is an addressable decoder with 8-bit address and 4-bit data. Both HT-12E and HT-12D are widely used in remote control toy in which not so many data is transmitted and received. A significant advantage of using HT-12E and HT-12D is to omit the microcontroller and relevant outlying components. This situation is similar to this project. TWS-434 is the 434MHz AM signal transmitter module whereas RWS-434 is the 434MHz AM signal receiver module. This pair of RF signal modules is small in size compared to corresponding microprocessor with RF transmission, reception function such as rfPIC12f and rfRXD0420 respectively. A more important point is that rfRXD0420 is too complicated and redundant to be used in the project due to the limited data communication.

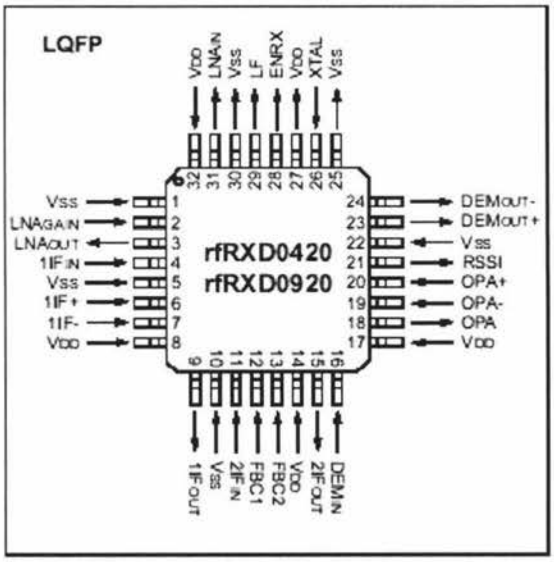


Fig.4-9 Pin diagram of rfRXD0420/rfRXD0920 [49]

Compared to the rfRXD0420, RWS-434 only has 8 pins so it can make PCB design much easy.

A signal on Din pin of HT-12D activates the oscillator that in turn is checked by a decoder continuously. The transmission is not valid until the received address is matched with the decoder's local address. Once lights in different rooms are encoded corresponding address, then valid transmission with address will switch on or off even dim specified light. The HT-12E encoder begins a 4-word transmission cycle upon receipt of a transmission enable (TE active low). This cycle will repeat itself as long as the transmission enable (TE) is held low. Once the transmission enable returns high the encoder output completes its final cycle and then stops as shown below.

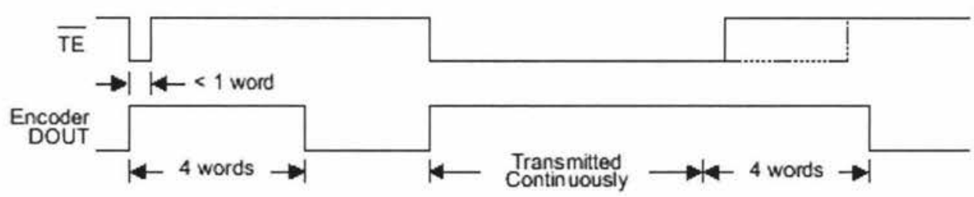


Fig.4-10 Transmission timing for HT-12E [50]

Here the duration of one word means the time to transmit 1 lead bit followed with 8-bit address and 4-bit data. This value is changed depending on the oscillator frequency of HT-12E and HT-12D. Through choosing different ohm value resistor, the oscillator frequency can be changed. The relevant relation curves are shown in Fig.4-11 and Fig.4-12.

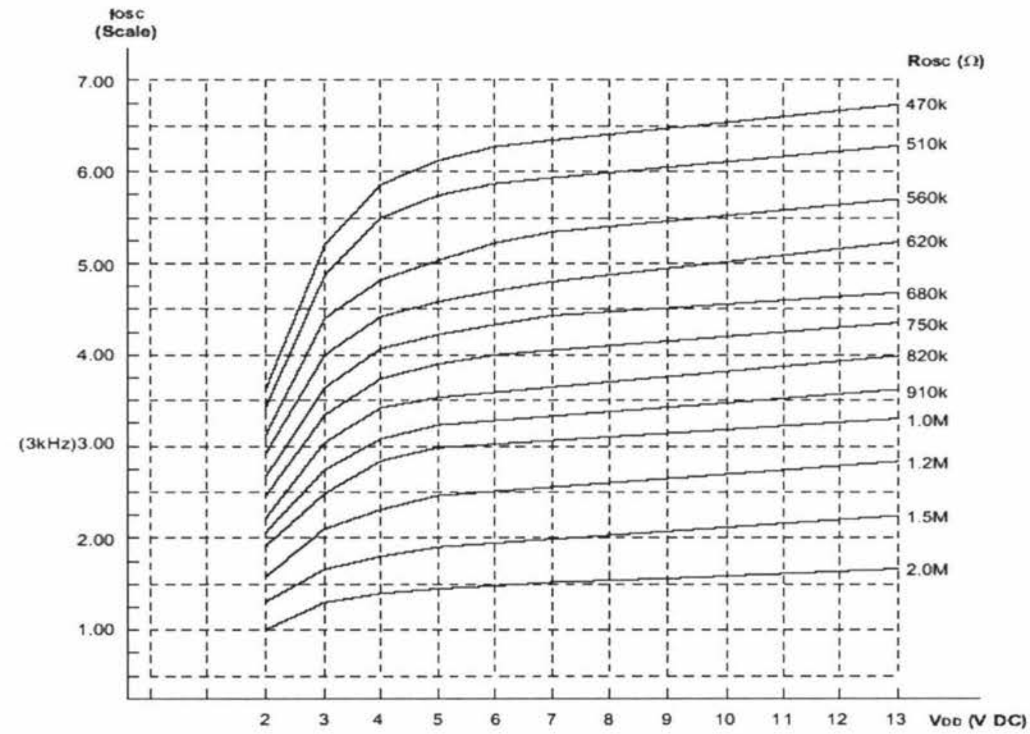


Fig.4-11 Oscillator frequency function with Rocs and

supply voltage for HT-12E [50]

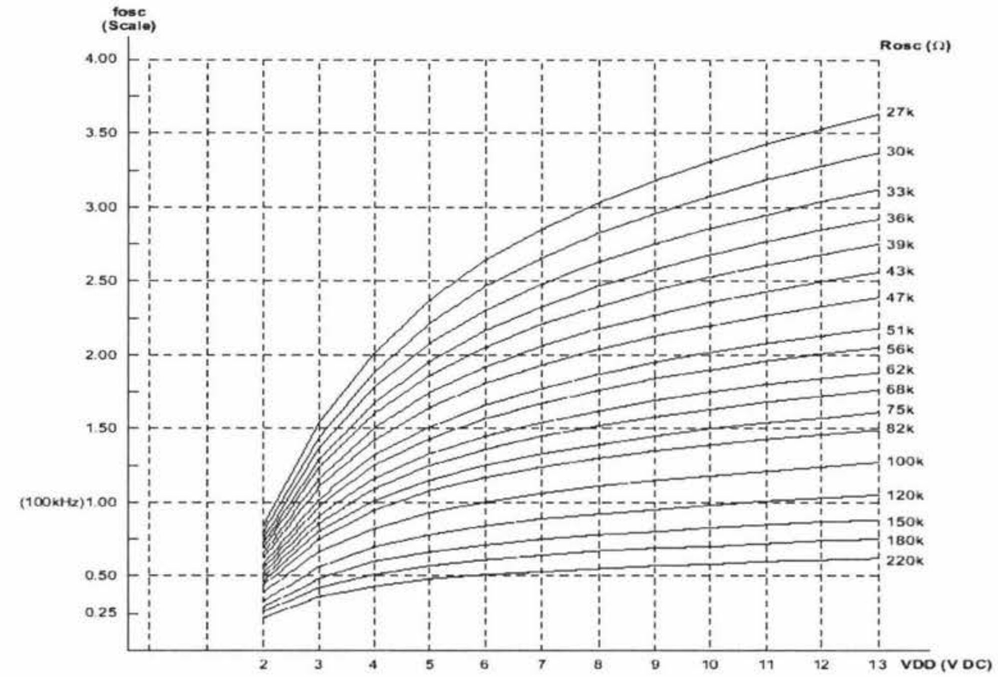


Fig.4-12 Oscillator frequency function with Rocs and

supply voltage for HT-12D [50]

A noticeable point is the oscillator frequency of HT-12D is suggested to be 50 times higher than that of HT-12E. Only both encoder and decoder match this frequency demand, they can work well.

During the experiment, two sets of resistors are used to analyze the wave form of relevant pins of HT-12E and HT-12D. The result is shown in Table 4-4.

	Width of pulse in pin VT of HT-12D (ms)	Duration of one word in pin Dout of HT-12E (ms)	Duration of one word in pin Din of HT-12D (ms)
Roscd=33kΩ Foscd=210kHz Rosce=670kΩ Fosce=4.25kHz	58	16.5	16.0
Roscd=150kΩ Foscd=70kHz Rosce=2150kΩ Fosce=1.35kHz	60	52.4	51.8

Table 4-4 Timing of wave form of three pins in two sets of resistor value

In both conditions shown in the table above, HT-12E and HT-12D match the frequency demand, the address and data are received correctly from pin Din of HT-12D. In these two situations, the trigger negative pulse excited to the pin TE of HT-12E is 6.5ms in width. This value is less than one word duration. However, if the TE holds low more than one word in time, the width of positive pulse in pin VT is far more than 50ms and is not fixed depending on the transmission data length. So the mode in which trigger negative pulse in pin TE is less than one word duration in time is chosen.

4-5 Touch control switch and dimmer

In order to switch and dim the light, the conventional wall switch cannot be used. The reason for this is that the mechanical switch device is used in conventional mode and it has memory. This means unless you change the switch state by hand, the switch will maintain in the last state(ON or OFF), but in a smart house, we just want to control the light by voice. So if we design a circuit to control the conventional switch by voice, some confusing in state of switch is happened. For example if light is switched on by light switch, then the toggle button of switch is held in the ON state, at this time, if we turn off the light by voice, as a result, the toggle button is still in the ON state, while the light is in the OFF state actually. So, the solution is to use the touch control switch mode instead of the conventional light switch.

LS7631/7632 is a 8-pin DIP integrated chip made in LSI/CSI company. This chip can be used to control a light, motor through thyristor or other power electronic semiconductor. Compared to the previous product such as LS7231, the supply voltage is much lower (5v), and the valid voltage in some control pins is much low. So it is much easy to interface with other components or circuits.

Two control modes are provided by the LS7631/LS7632. One is short touch mode, the other is long touch mode. Only long touch mode can dim the light. The short touch mode switches the light. However there are some functional differences between LS7631 and LS7632. The main difference is for the LS7631, when a long touch is applied, the dimming direction automatically reverses whenever maximum or minimum conduction angle are reached. When pin SENS or EXT are held low or high level respectively for no less than 410 ms, then dimming mode is applied. Holding low or high level in pin SENS or EXT range from 50ms to 400ms triggers the switch mode.

The dimmer board should be embedded in the wall switch box, so the whip antenna, attached to the TWS434 and RWS434, with a length of quarter wavelength, were redesigned and replaced by other antenna such as a loop antenna in a PCB.

Another problem is the metal touching plate which will reflect most parts of the electromagnetic wave propagated from the MCU board with information. The solution is to change the surface feature of the plate or (and) coupling mode of the dielectric.

4-6 Solution

4-6-1 PCB design

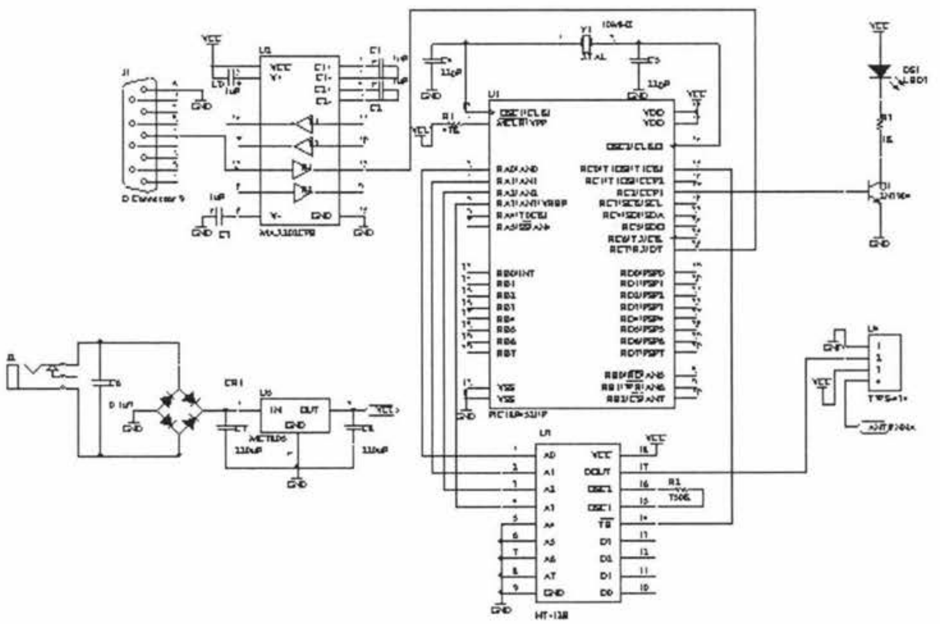


Fig.4-13 Schematic of universal remote controller

The Fig.4-13 shown above is speech driven universal remote controller using the PIC16F62A as the microcontroller. D9 connector is a female RS-232 interface. Max202cpe converts the high range of voltage, which is used during two or more devices communication via the RS232 port, to match the CMOS voltage level. The serial data is received via pin 7, denoted as RX, by MCU. After processing of MCU, corresponding signal feeds to the pin 9 of MCU and controls those electrical appliances that are controlled by IR signal. The dimming and switch light signal is transferred to antenna from pin 11 of MCU. Port A (RA0-RA3) is connected to the HT12-E from A0 to A3 respectively to provide address for different lights.

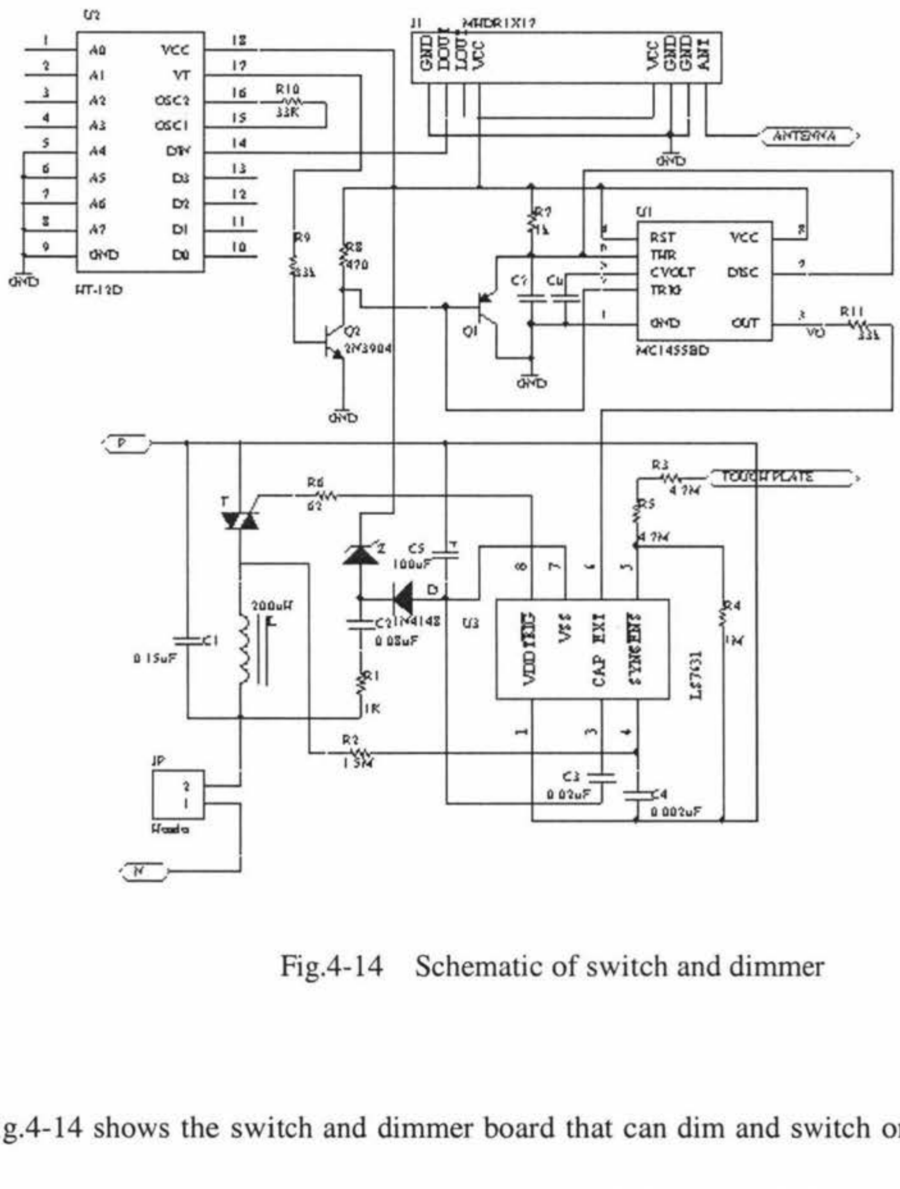


Fig.4-14 Schematic of switch and dimmer

The Fig.4-14 shows the switch and dimmer board that can dim and switch on/off lights. In order to save space and not use transformer which is part of DC power supply in conventional circuit, 5.6V zener diode, denoted as Z, capacitor C5, C2, and resistor R1 comprise the 5V DC supply. The corresponding waveform is simulated by Protel Dxp and shown as Fig.4-15 Fig.4-16. Once HT12-d detects a valid transmission, the pin VT of HT12-D is set and last till no signal is received. According to the result of experiment, the positive pulse in pin VT is about 58ms. This value is valid to trigger LS7631 as short touch mode. However, because of the fact that in PIC16f serious microcontroller, the all timers only can reach the maximum 209.712ms timing, it is impossible to directly use internal timer of microcontroller to generate more than 410ms pulse wave form. As a result, the 555 circuit is applied to act as a pulse generator fed to the LS7631 and the timer's error can reduce less than 1%. The simulation waveform is shown in Fig.4-17 and Fig.4-18.

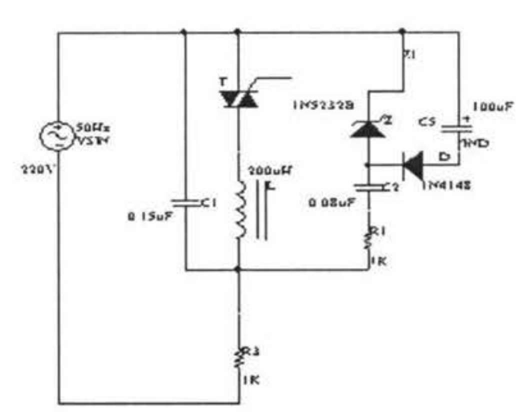


Fig.4-15 Schematic of DC supply

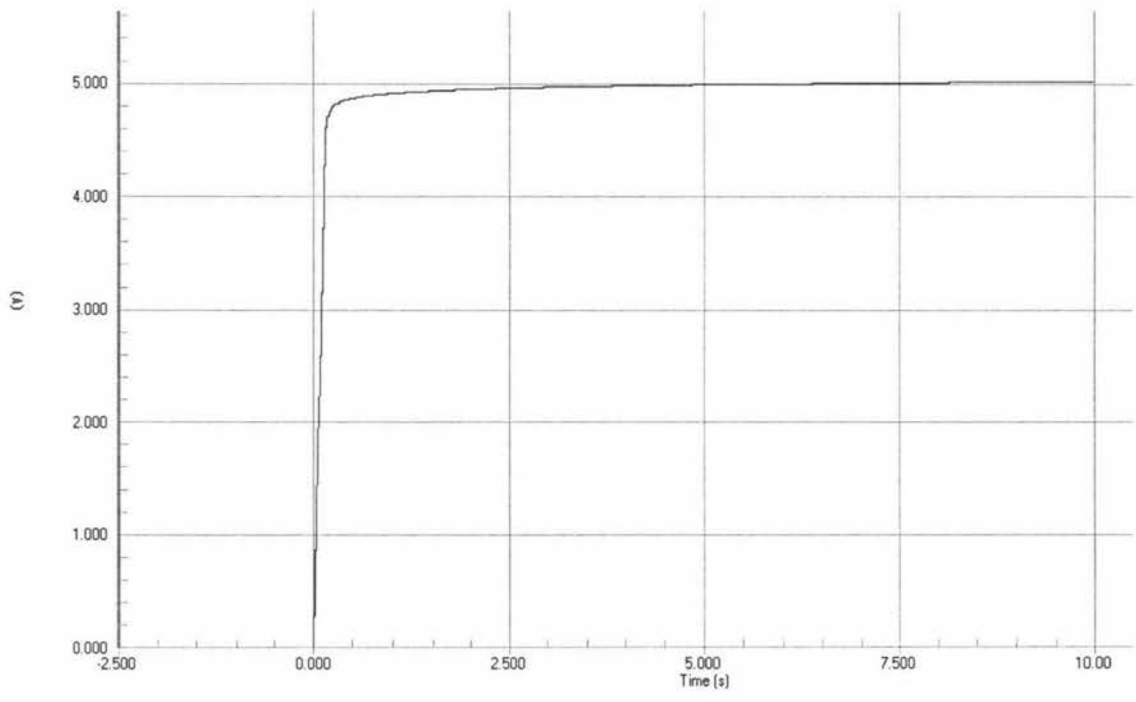


Fig.4-16 Simulation wave form of DC supply

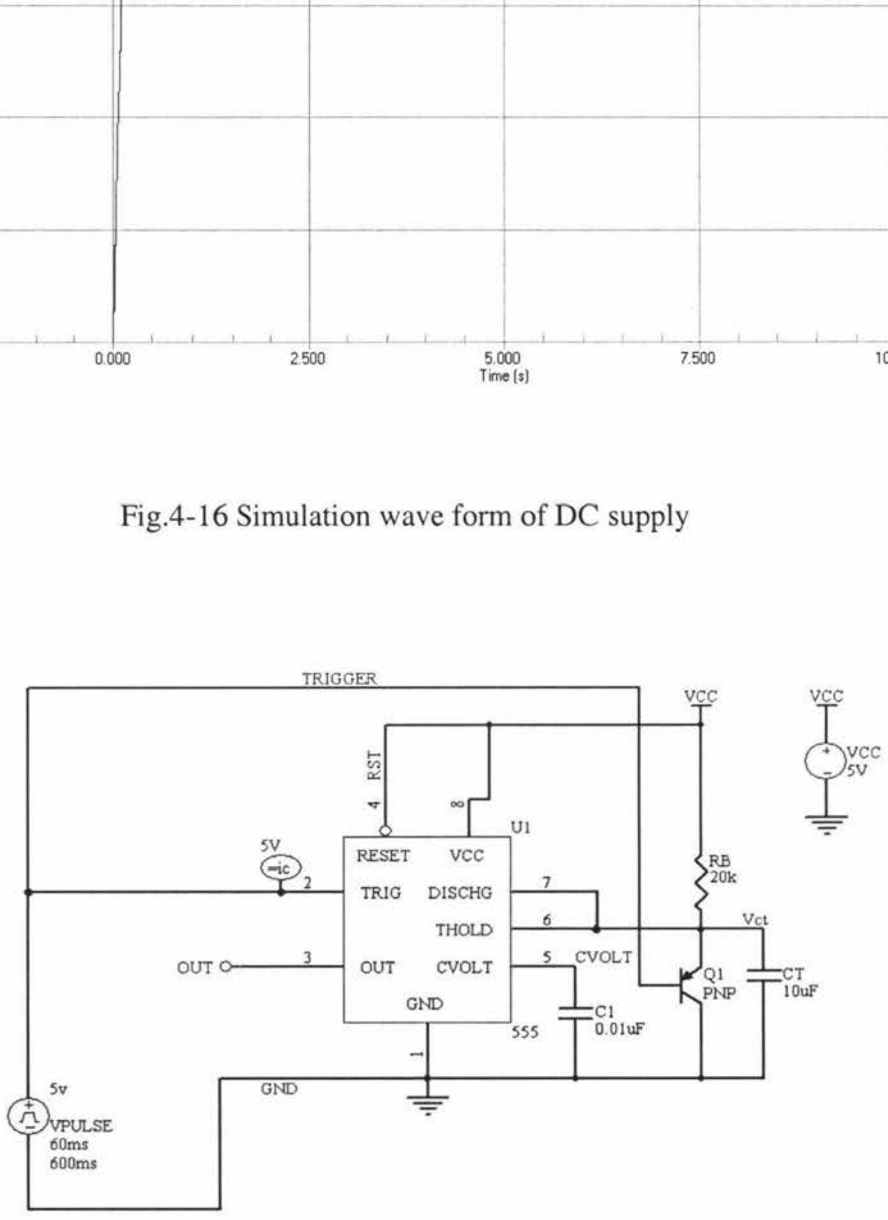


Fig.4-17 Schematic of 555 timer simulation circuit

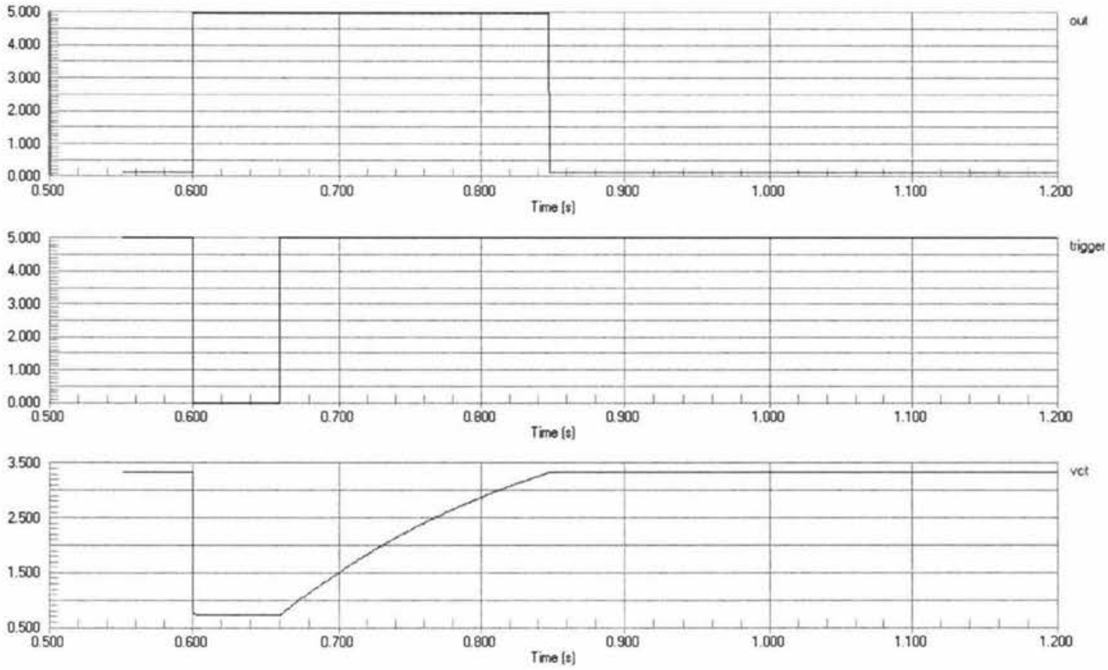


Fig.4-18 Simulation wave form of Pin out, trigger and Net of capacitor

The waveform observed from oscilloscope is very close to simulation result. The corresponding PCB layout diagrams are shown from Fig.4-19 to Fig.4-22 as follow.

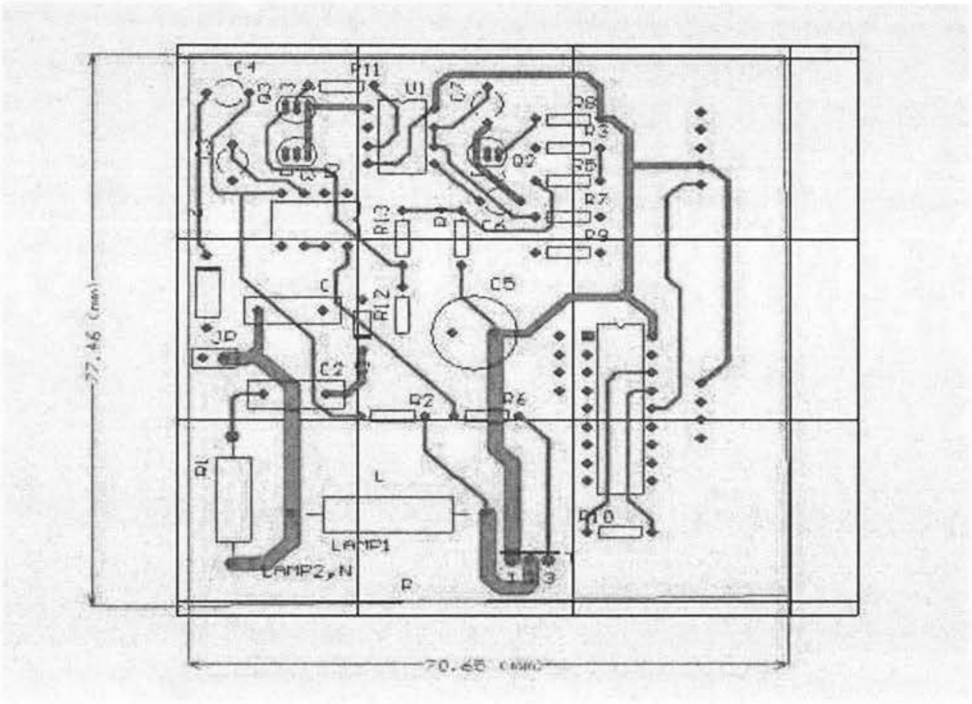


Fig.4-19 Top layer of Switch&Dimmer PCB board

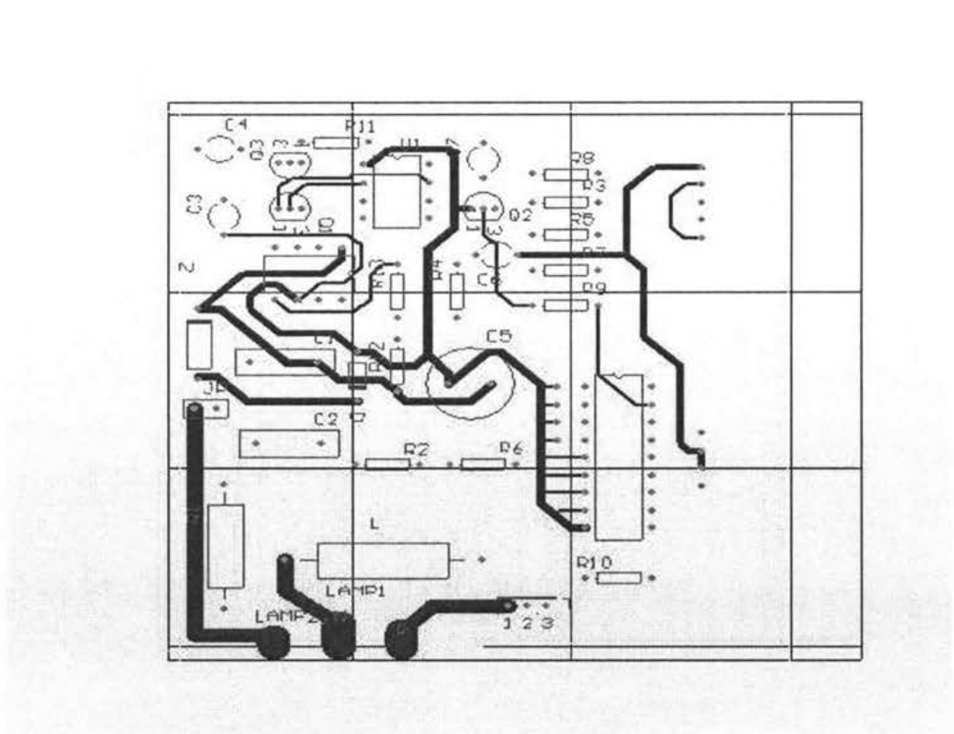


Fig.4-20 Bottom layer of Switch&Dimmer PCB board

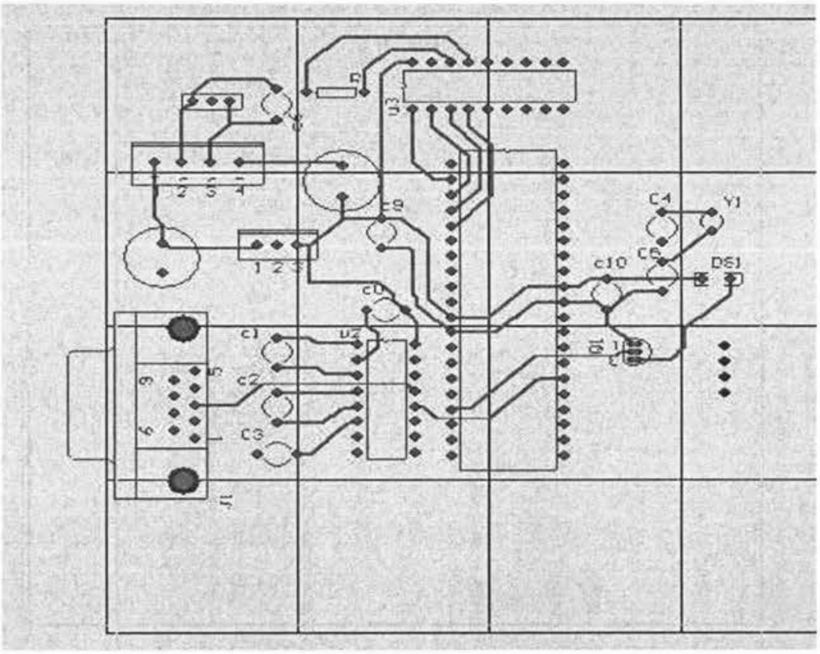


Fig.4-21 Top layer of Remote controller PCB board

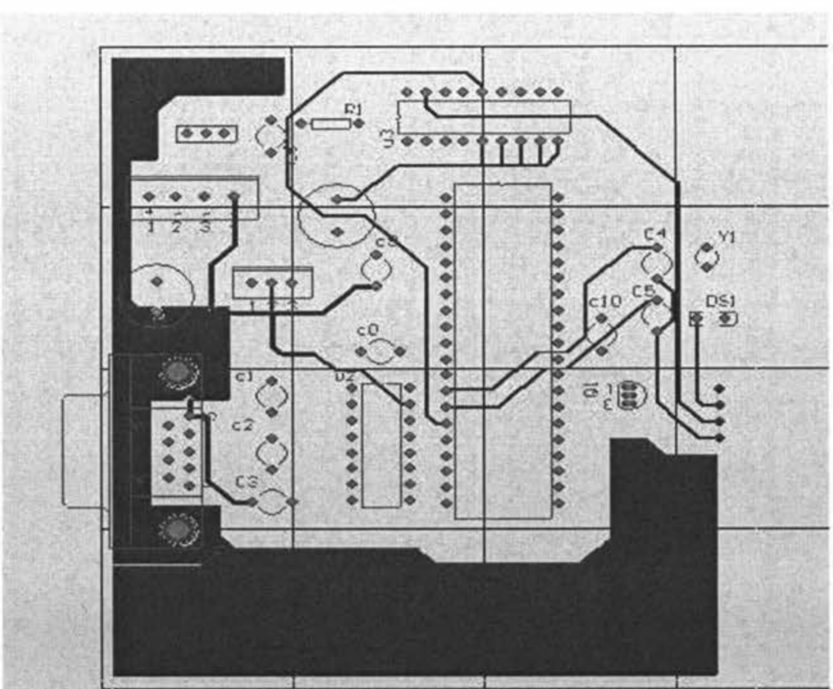


Fig.4-22 Bottom layer Remote controller PCB board

4-6-2 Programming the MCU

Microchip provides a powerful development board and corresponding software. However, the PICDEM 2 PLUS demo board does not support the Pic16F627A microcontroller directly (please refer to the article "micro icd2 header" and the datasheet number is DS51292G). So far, I write the assembler program to PIC18F452. During the edit program, simulation is used to debug. Although PIC18F452 does not meet any problem that PIC16F627A does, some interrupt flags cannot be set by man-maid in PIC18F452 while simulating step by step. So firstly, I edit program in PIC16F627A and choose debug method as simulation step by step. Until no bugs are in this program, then, I edit program in PIC18F452 and finally write in the chip.

Some points should be noticed after successfully writing program. First point is that ASCII code is used to communicate between PC and microcontroller. When I first transfer some binary data from RS232 to the PIC18f452 by using VISA in LabVIEW 6.0, I got a fixed waveform in RC-5 protocol which is 11000000110000 no matter what I transfer. Probably the serial communication of PC is not good enough because the reputation of serial communication in window 2000 operate system. Then I use the oscilloscope to check the waveform in pin 12 of SP232 in Qwikbreadboard, one of microchip development board. At same time, I use HHD, serial monitor software, to monitor any data communication between PC and development board. When I write to RS232 port a train of strings : 001110101, the monitor just display data written to port as 30, 30,31,31,31,30,31,30,31 in sequence. Consequently, I find the VISA write ASCII code to the MCU rather than binary code. This is reason why I just got a fixed waveform in RC-5 protocol that is 11000000110000 and explains the waveform in pin 12 of SP232.

The second point is how to use characters for command generation in LabVIEW. Each character written from VISA in LabVIEW to MCU is translated to ASCII code and takes 7-bit, so two characters are represented by 14-bit, just about 2-byte. However, from the ASCII code table, 34 items cannot be matched in valid character. Unfortunately, some IR commands are located in this range. So 14-bit binary command should be communicated with the MCU by character mode instead of two characters. As a result some subroutine program in MCU should be used to translate every ASCII code to corresponding characters (0 or 1).

So far, when I write some commands in LabVIEW to MCU through RS232, the oscilloscope can observe RC-5 or Sony protocol waveform, and TV does correct reaction as response of these commands. The program code can be referred in Appendix A.

4-6-3 Labview program

The LabVIEW is a kind of graphical programming language and there are some palettes of Dragon natural speaking. So it is not difficult to write a program. Different protocols for communication such as serial, General Purpose Interface Bus (GPIB), VXI, etc. are available to engineers and scientists. However, they are all inherently different, rendering each protocol incompatible with any other. Industry has been working to make a standardized system that can operate all of these devices. This standard has become known as Virtual Instrument Software Architecture (VISA). We will use VISA within LabVIEW to communicate via the serial port. VISA can serve to greatly simplify the making of many programs due to its versatility across communication protocols. However the names that LabVIEW supplies for the ports are not labeled very well, they are manageable; ASRL1::INSTR is equivalent to COM1; ASRL2::INSTR is equivalent to COM2 and so forth. What is remarkable in this diagram is the SubVI file of the Signal Generation Voice Menu and a case statement of Changing control value corresponding to the voice menu. All voice commands are stored in the SubVI file of Signal Generation Voice Menu in advance. Once a valid voice command is recognized, the corresponding command string is transferred to the write buffer of the VISA. The content of the write buffer will be kept until empty string is forced to clean the write buffer. So in this situation a command string is transferred to the serial port continuously and repeatedly. This is not desired. The solution is to empty the write buffer once the command string is transferred to serial port.

As a example, the diagram of a program that can write command to the buffer of VISA is shown in the Fig.4-23 and Fig.4-24 below. What we need to do is to type the string in the write buffer, then click the run button. The content of string then is transferred to the serial port.

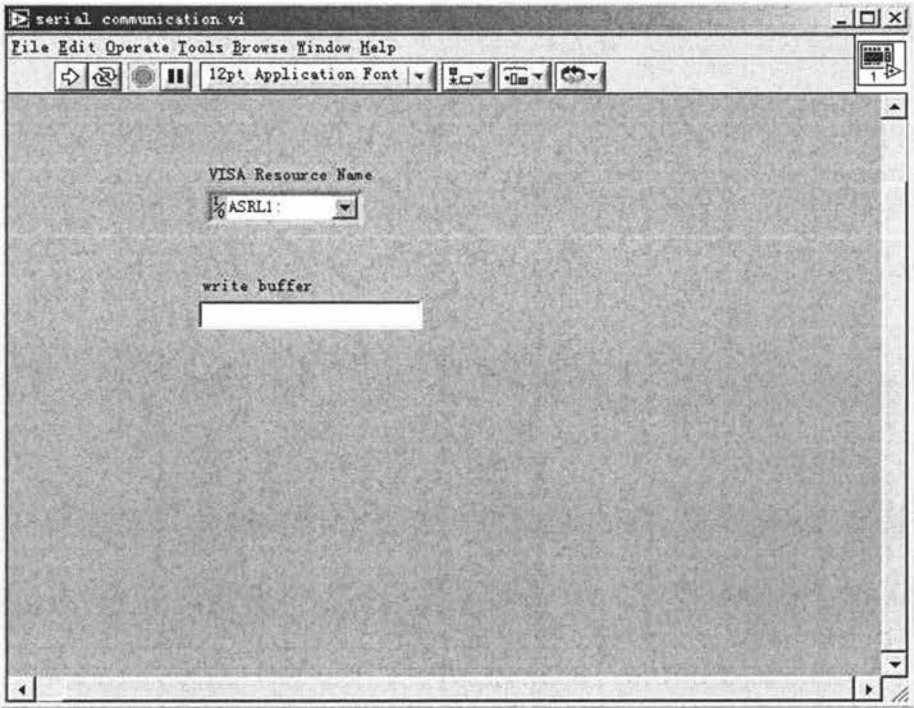


Fig.4-23 The Front panel of Serial communication.vi

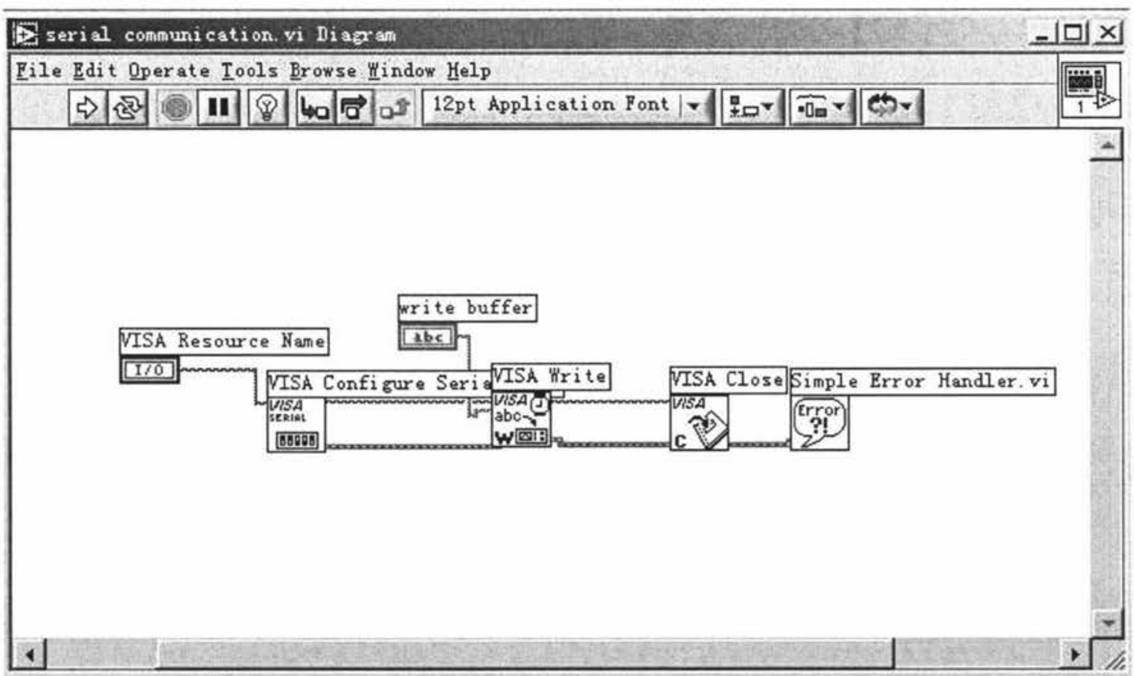


Fig.4-24 The diagram of Serial communication.vi

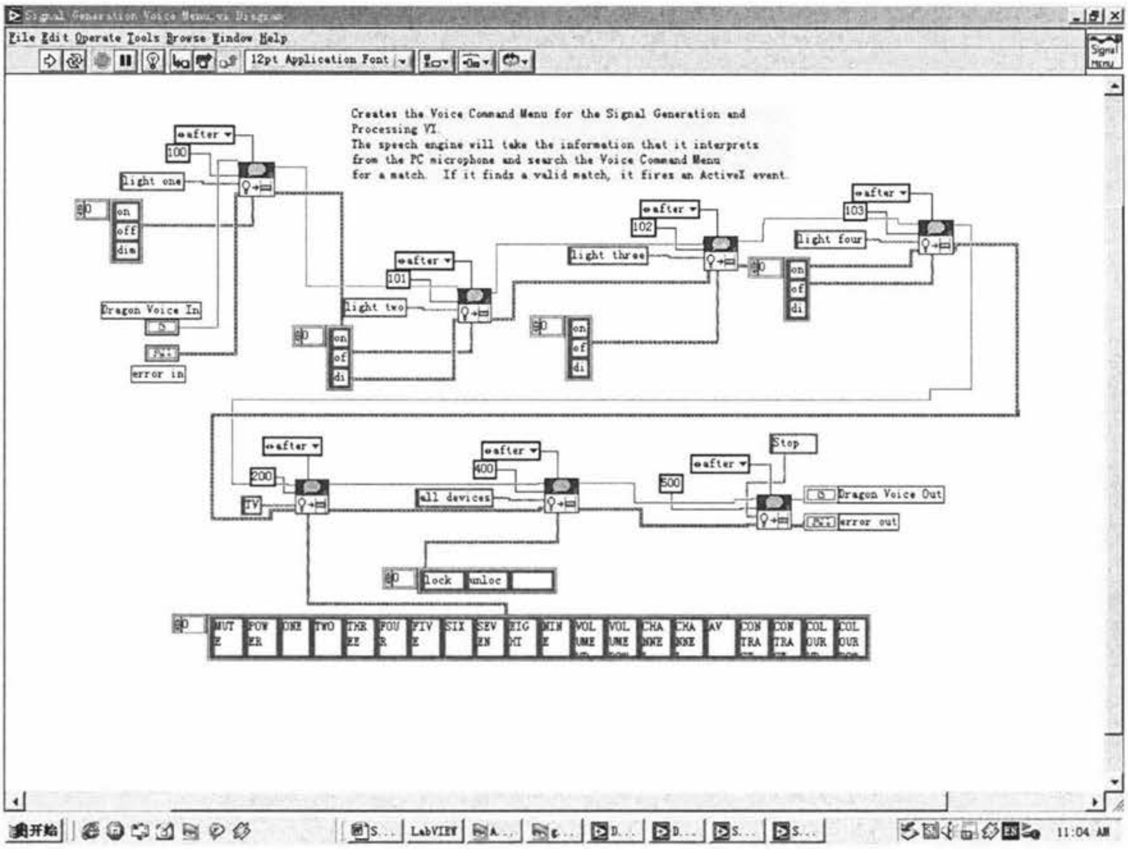


Fig.4-25 The diagram of Signal generation voice menu.vi

Any new voice command can be added into the diagram shown in Fig.4-25 above and then adjusts the control value case. So far, just television commands are build and both RC-5 Phillp and SIRC protocols are available now. In next step, commands of CD player, tape/cassette recorder and so on can be controlled by voice and we can choose the protocol the device used. The all LabVIEW program diagrams can be referred in Appendix B.

4-7 Antenna Design

4-7-1 Antenna characteristics

Before starting designing the antenna, some technological terms and parameters should be described.

Directivity:

The directivity D is an antenna parameter used to assess the ability of radiating electromagnetic power to a desired direction. It is given by a ratio of the maximum radiation intensity (power per unit solid angle) $U(\theta,\phi)_{\max}$ to the average radiation intensity U_{av} (average over a sphere). We generally are interested in the far field rather than near field because the far field is the radiative field and more value for antenna. So at a certain distance from the antenna the directivity can be expressed as the ratio of the maximum to the average Poynting vector [46]. Thus,

$$D = \frac{U(\theta,\phi)_{\max}}{U_{av}} = \frac{S(\theta,\phi)_{\max}}{S_{av}} = \frac{4\pi}{\iint \frac{S(\theta,\phi)}{S(\theta,\phi)_{\max}} d\Omega} \tag{4-1}$$

Power Gain:

An antenna that radiates poorly has low "gain". Antenna gain is a measure of how strongly the antenna radiates compared to a reference lossless isotropic source. It depends on both it directivity and efficiency. If the efficiency is not 100%, k, a factor, is used to calculate the power gain according to the directivity. Thus,

$$G_p = kD \tag{4-2}$$

On other hand to understand, compared to the directivity D which is referred to the radiative power, the power gain is referred to the power fed to the antenna. Due to the ohmic losses in the antenna, radiative power is less than the fed power. In transmitting, these losses are not radiated but heat the antenna structure.

Radiation Pattern:

Radiation pattern is to describe how the field intensity varies with the change of direction in space. The radiation pattern is drawn in three-dimension. In the spherical coordinates, it is the function of two spherical coordinates θ, ϕ. In order to avoid the difficulty to draw radiation pattern, we can individually draw the pattern only θ or ϕ is variable. The pattern has its main-lobe and the side-lobes. So two terms are used to measure the antenna's character [47].

The width of main-lobe:

This parameter is used to identify how acute the field intensity is in the radiation area. It is defined as the angle between two half-power points (-3db points) in the main-lobe. For strong direction antenna, the width of main-lobe is small.

The side-lobe level

In the undesired directions, the radiation intensity is expected to be much lower than that of main-lobe. So the side-lobe level is defined to assess the attenuation of side-lobe referred to the main-lobe. For non-isotropic antenna, in other word a desired directions antenna, the lower side-lobe level, the better character is.

The radiation resistance:

When antenna is fed by source as working in transmitting mode or collects and delivers electromagnetic power to the terminating as work in receiving mode, the impedance matching should be considered. Antenna impedance can be divided into three parts, a radiation resistance R_r, a nonradiative or loss resistance R_L and a image part X_A. The radiation resistor is not real resistor but equivalent resistor to measure the radiation power in line with the circuit theory. Assuming that the current pass through the radiation resistor is equal to the maximum current in antenna, then the power dissipated in radiation resistor is equal to the radiation power of antenna. So in a good design, the antenna has a large radiation resistance and small loss resistance and, at same time, the antenna impedance is so close to the terminating impedance.

4-7-2 Several kinds of antenna

The thin linear antenna is widely used for transmitter or receiver devices in the radio frequency range. In thin linear antenna, the half-wavelength dipole antenna is most important because it can provide a valid directional pattern and property of impedance. The simplest short dipole antenna is the "Whip". This is a quarter wavelength wire that stands above a ground plane. The typical picture of wire antenna is shown in Fig.4-26

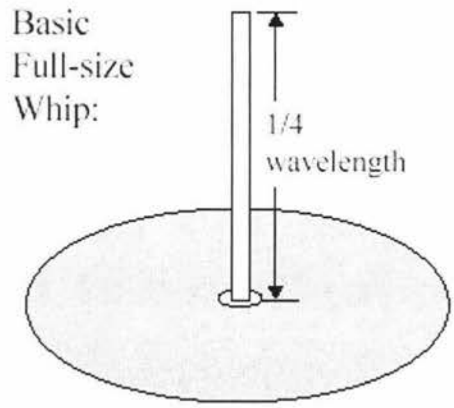


Fig.4-26 Quarter wavelength wire antenna [53]

According to the method images, we can replace the ideal ground plane to another quarter wavelength wire being symmetrical with the ground plane. So the quarter wavelength wire antenna has the same electromagnetic field with half wavelength antenna in above hemisphere. The function of direction pattern is given as follow

$$F(\theta) = \frac{\cos[(\pi / 2)\cos\theta]}{\sin\theta} \tag{4-3}$$

This pattern is only slightly more directional than the pattern of an infinitesimal or short dipole in which the direction factor is given by $\sin\theta$. The main lobe (or beam width) between half-power point of half wavelength antenna is 78 degree as compared to 90 degree for the short dipole. When length of antenna increasing to full-wave, the main lobe between half-power point is decreased to 47 degree. It means more directional. However in smart house project, the demand to direction of antenna is not strict. The more important point is the size of antenna. Due to the limitation of device, the antenna should be small in size. The wavelength of 434MHz is about 0.69m, the quarter wavelength wire antenna is about 17 centimeter. So it is impossible to use quarter wavelength wire antenna in this project. Some kinds of antenna can be used because of the compact in size such as the PCB antenna, helical antenna, chip antenna and loop antenna etc. However, in general, the compact antenna has low efficiency to propagate electromagnetic power in space.

PCB whip antenna

PCB whip antenna can be made as a trace on a Print Circuit Board. This is very practical at frequencies over 800 MHz. At lower frequencies, a full size whip may be too long, even when wrapped around a few corners. So the PCB material with higher relative dielectric constant ε_r is proposed to use to reduce the antenna size. But on other hand, the gain of antenna will be decreased when dielectric constant increases. A summary of the main electrical characteristics of three commonly used board materials for radio products is given in the table below. The cost increases with increased dielectric tolerance and reduced loss.

Board material	Dielectric constant	Loss tangent
FR4	4	0.01
Rogers R04003	3.38	0.0027
Rogers RT/Duroid	2.2	0.0009

Table 4-5 Comparison with three commonly used PCB material characteristics

An important characteristic of PCB whip antenna is the innately low power gain G_P Compared to dipole antenna, PCB whip antenna has 8db to 12db lower power gain. It means the radiation range, or distance is about a quart to a half of that in dipole antenna. Another characteristic of PCB whip antenna is that the radiation resistance is quite small. In general, the terminating impedance is $50\,\Omega$ or $75\,\Omega$. So the impedance matching network is needed to transform impedance, which increases the difficulty of design.

PCB loop antenna

Loop antenna is entirely different from a whip, in which both ends of the antenna are terminated. In this case, the end that is opposite the transmitter (or receiver) is grounded. For electric dipole antenna, the far field only has two contributions E_θ and H_ϕ. But for small loop antenna, the far field only has E_ϕ and H_θ to contribute it. By comparison the field of electric dipole and small loop antenna, the fundamental difference is that they are in time-phase quadrature [49]. One advantage is that a capacitor can now be used to tune and match the antenna, instead of a coil. Another advantage is that the loop antenna is not easily detuned by hand effects, although the impedance may still vary. The loop can be made small, is not ground plane dependant, and requires no more space than a short whip. For these reasons, loops are very common in portable devices. However, there are some disadvantages to restrict utilization. Small loop antennas have a reputation for poor gain. A small loop will have a very narrow bandwidth. This makes tuning extremely critical. Tuning is often done with a variable capacitor, which adds to the cost of time and some special instruments such the network analyzer and vector voltmeter. If the loop is large enough, it may be practical to use a non-variable capacitor. This requires careful adjustment in engineering stages, to ensure that it is properly tuned with a standard value capacitor. One example of loop antenna given by Kent Smith [53] covers a 12 by 35mm area on the end of a board and is shown in Fig.4-27. It is tuned to 433.9 MHz with a variable capacitor.

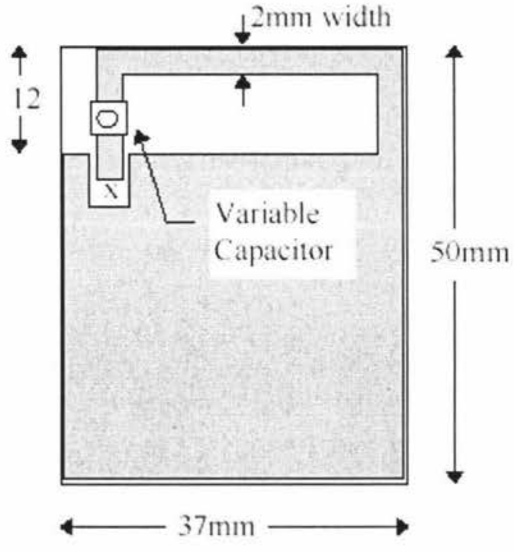


Fig.4-27 Structure diagram of 434Mhz PCB loop antenna [53]

This antenna is very omnidirectional, but had a power gain of only -18 dBd compared to the dipole antenna. This gain is too low to accept and a larger loop should have improved gain. The actual reason why the power gain is so low is that the radiation resistance is too little. The formula of resistance for small loop is given as follows:

$$R_r = 31171\left(\frac{A}{\lambda^2}\right)^2 \approx 31200\left(\frac{A}{\lambda^2}\right)^2 \qquad (\Omega) \qquad (4\text{-}4)$$

where the A is the area of loop covered, λ is the wavelength. Although on the PCB board, the wavelength is shorter than that of in free space in same frequency, which can increases the radiation resistance, the radiation resistance is still so little. Due to without proper instrument and the size limitation, the PCB loop antenna is given up to be a reasonable design.

Helical antenna

The helical antenna was invented by Kraus in 1946. It has a basic three-dimensional geometric form. A helical wire on a uniform cylinder becomes a straight wire when unwound by rolling the cylinder on a flat surface. A dimension of helical antenna is shown in Fig.4-28.

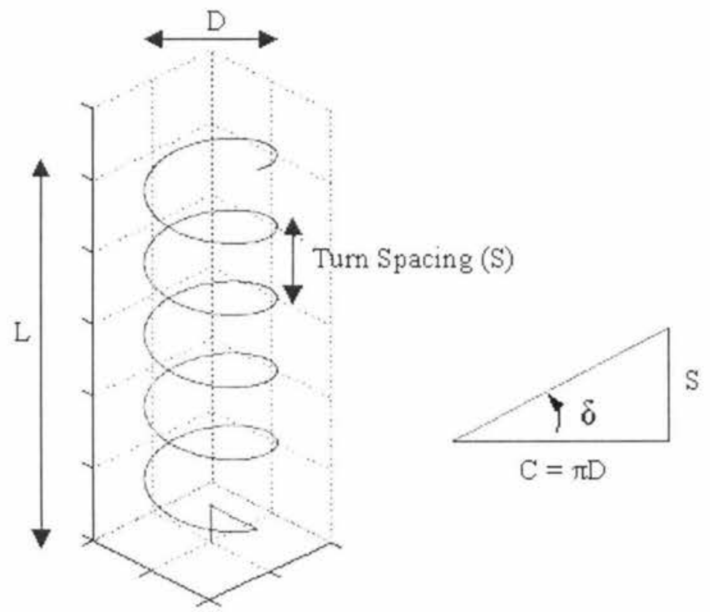


Fig.4-28 Dimensions of a helical antenna

The following symbols will be used to describe the helical antenna shown above. D is the diameter of helix, C is the circumference of helix, S is the spacing between turns, L is the axial length and δ is the pitch angle. The helical antenna works as the two main radiation modes: axial (or beam) mode, normal mode. Axial mode means the maximum radiation direction is parallel with the axial dimension of helix, but in normal mode the maximum radiation direction is perpendicular to axis of helix. Both two modes is dependant on the ratio of diameter to the wavelength. When the ratio is less or equal to

0.18, that is $D/\lambda \leq 0.18$, the radiation mode is in normal mode, the pattern is a circle in all perpendicular planes to axis of helix. The pattern of normal mode is shown in Fig.4-20 (b). When the ratio is in the range from 0.25 to 0.4, that is $0.25 \leq D/\lambda \leq 0.4$, the radiation mode is axial mode, then the pattern is shown in Fig.4-29 (a). As the ratio continuing increase, the pattern changes to the approximate cone shown in Fig.4-29 (c).

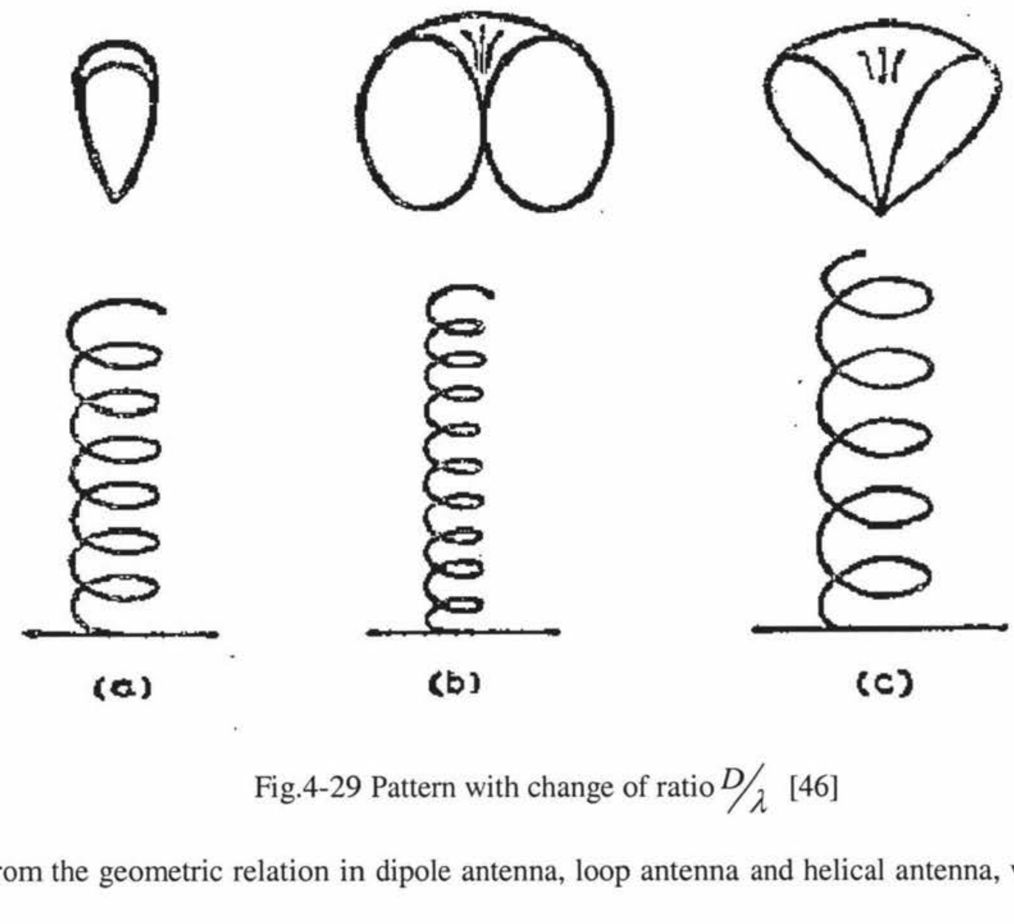


Fig.4-29 Pattern with change of ratio D/λ [46]

From the geometric relation in dipole antenna, loop antenna and helical antenna, we can find that when the pitch angle of helical antenna is zero degree, the helix becomes the loop, when the pitch angle is right angle, the helix straightens out into a linear antenna. So the loop and straight antenna is the limited case of the helical antenna. Because the phase velocity of electromagnetic wave propagated via the axis of helical antenna is lower than that of straight linear dipole antenna, so helical antenna has relatively large radiation resistance. This is the advantage to improve the efficiency of radiation.. In order to reduce the antenna's size the normal mode is chosen in the design. HFSS 9.0 is used to design the helical antenna applied in this project.

HFSS v9.0 is the product of Ansoft company. This soteware is very fit to design the high frequency non-source components such as antenna and filter. The helix antenna solution designed by HFSS v9.0 is shown as follow.

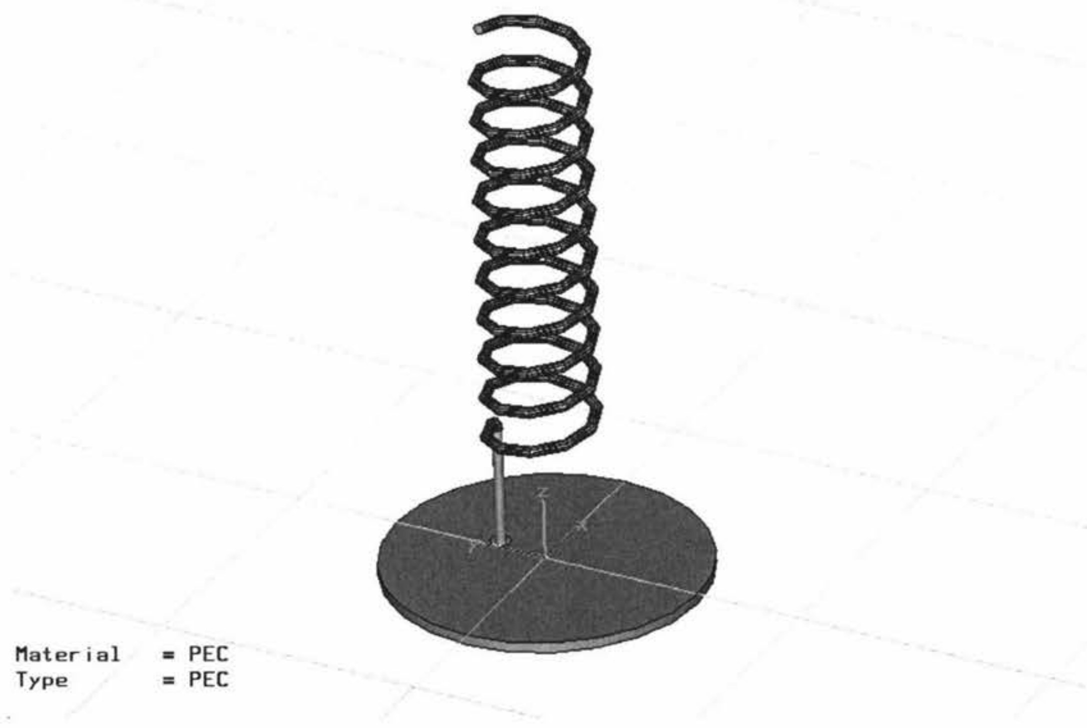


Fig.4-30 The geometry of antenna

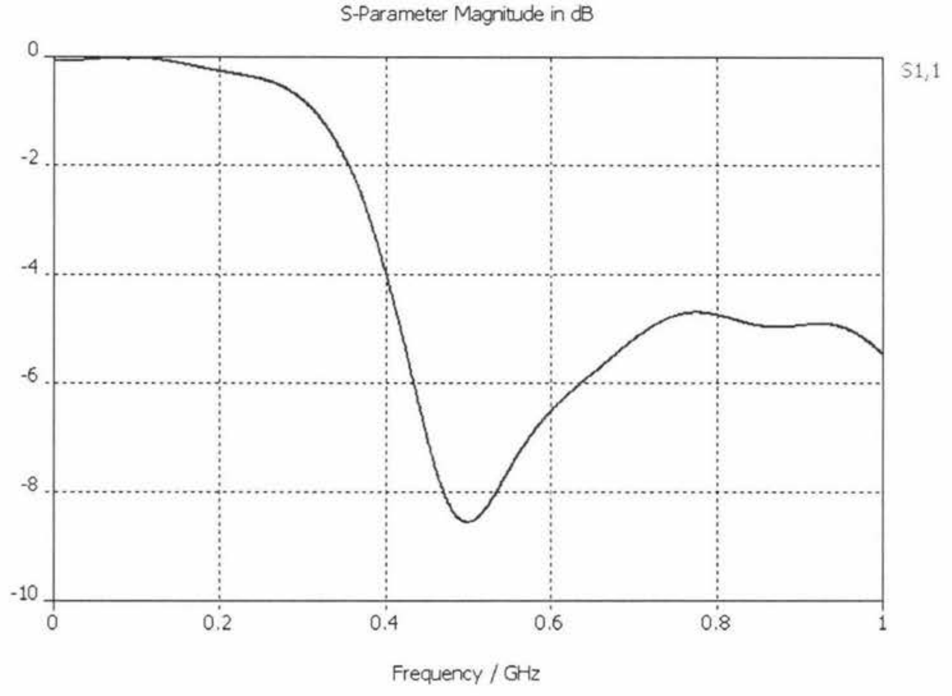


Fig.4-31 The scattering parameter of S11 in dB

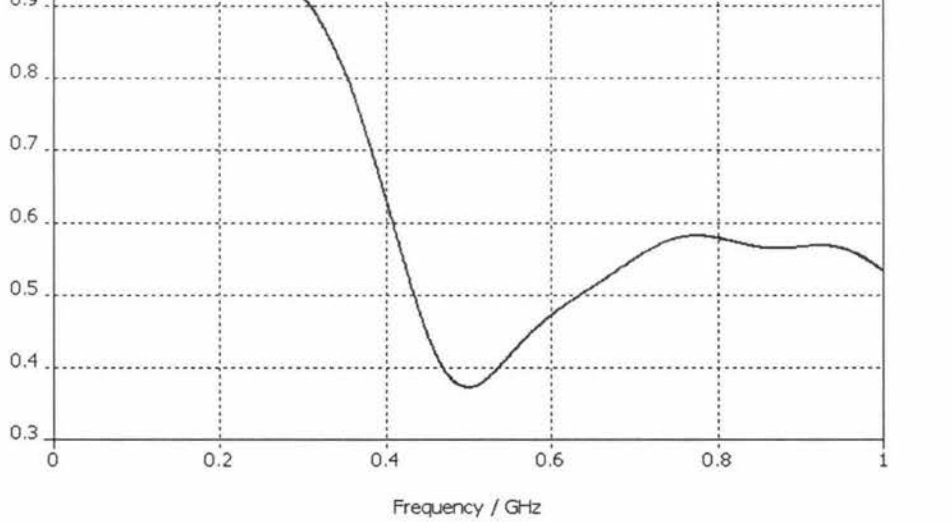


Fig.4-32 The scattering parameter of S11 in magnitude

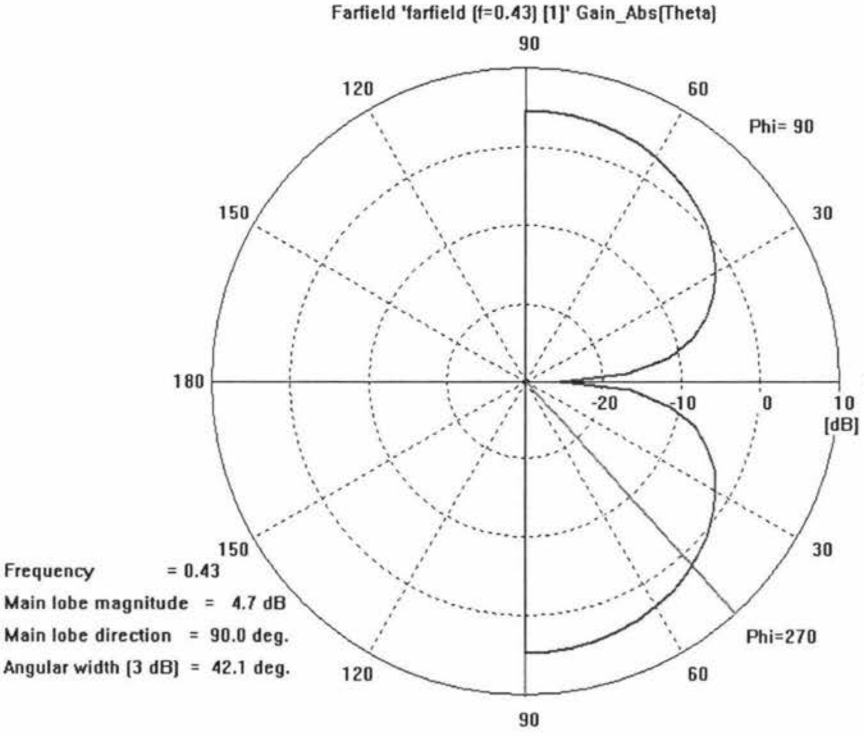


Fig.4-33 The E-plane pattern

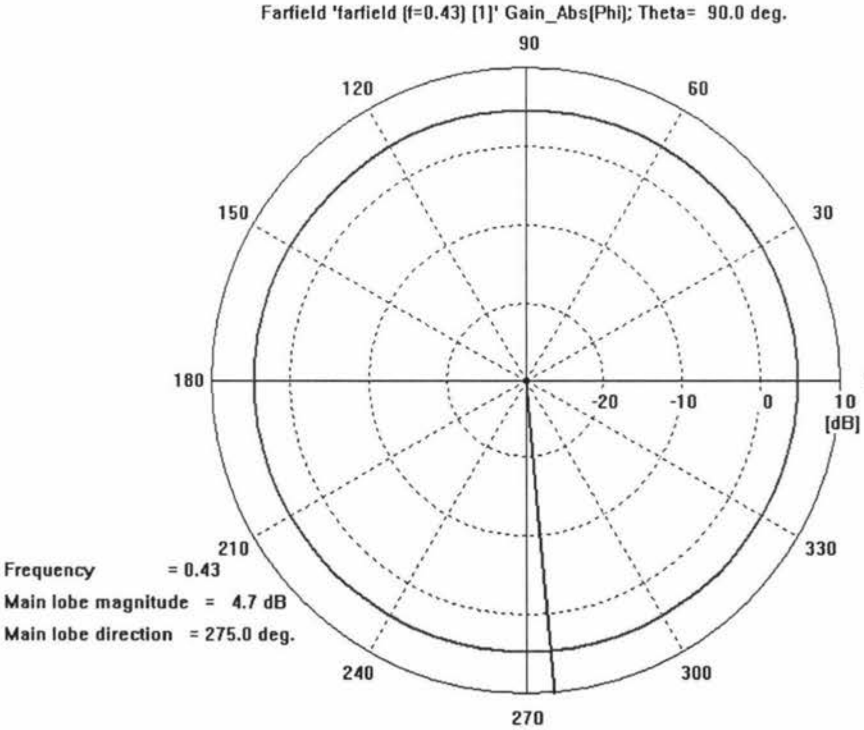


Fig.4-34 The H-plane pattern

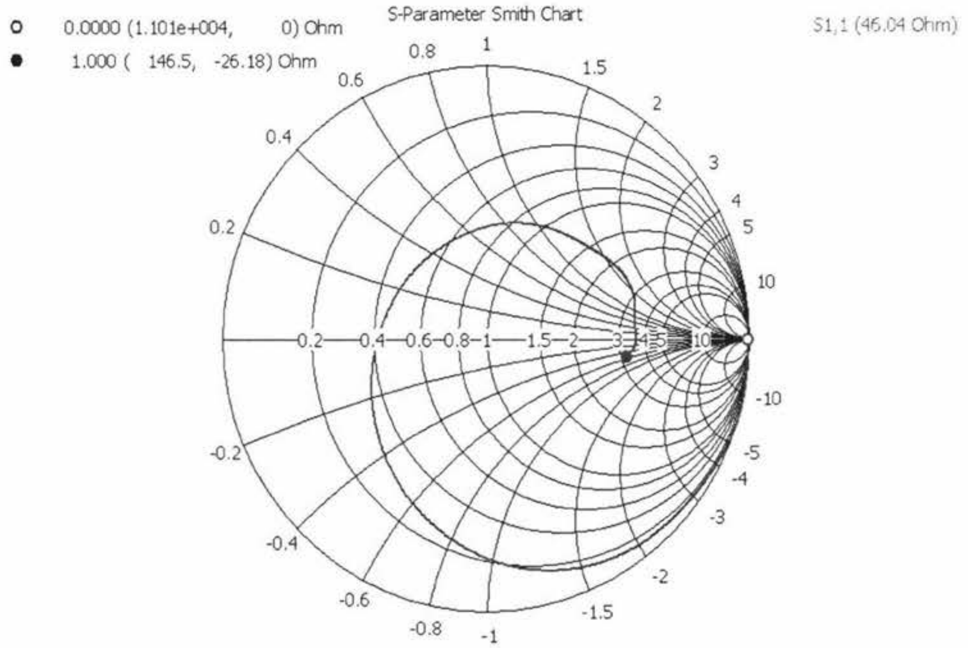


Fig.4-35 Scattering parameter of S11 in smith chart

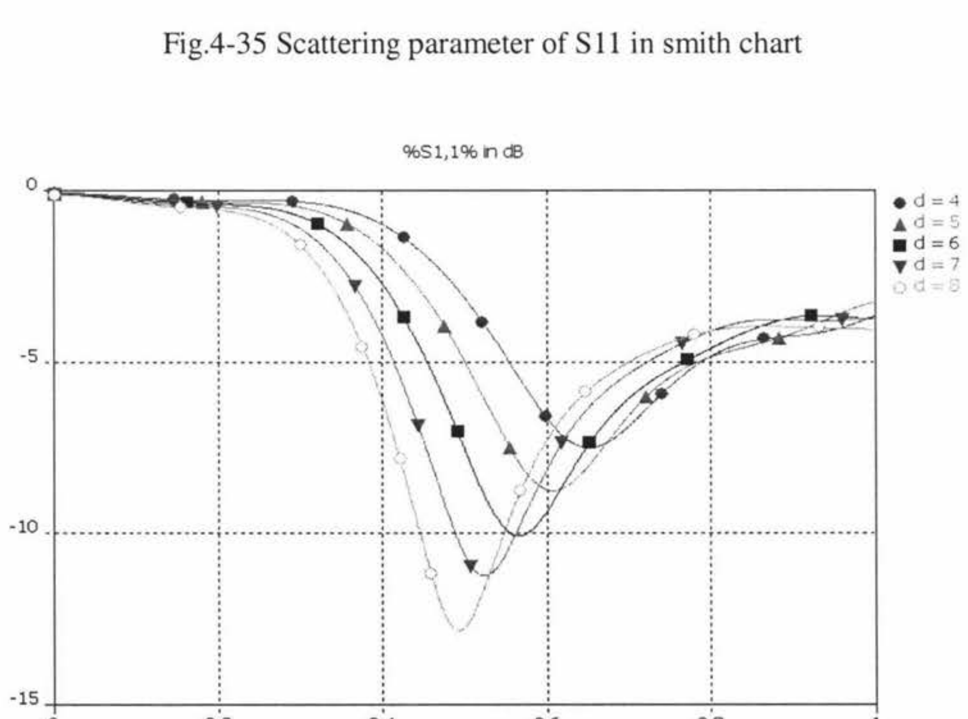


Fig.4-36 Relation of S11 and frequency for different spacing between turns

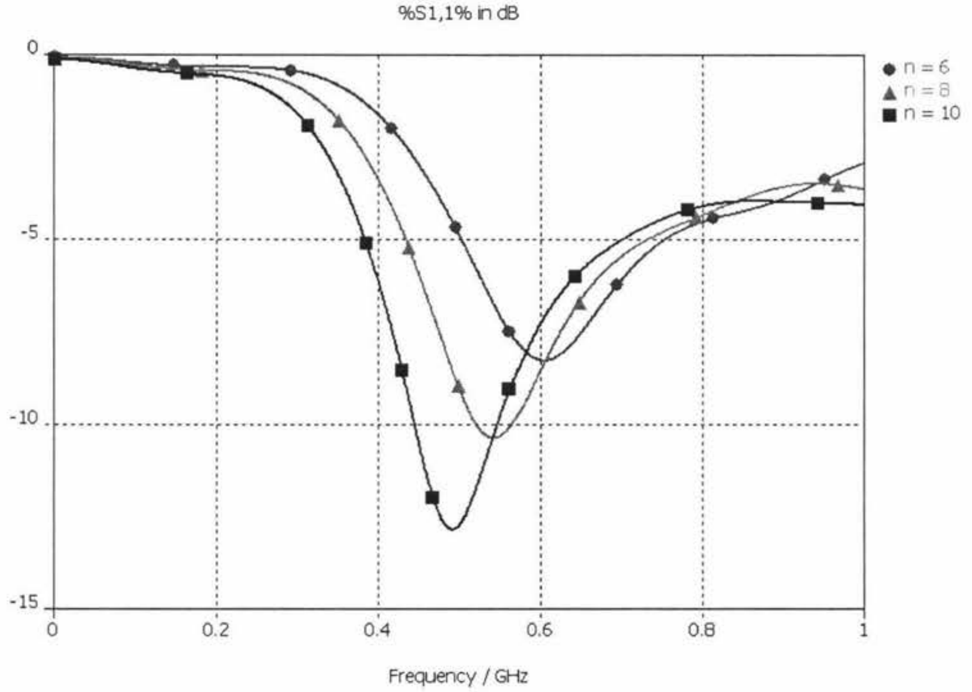


Fig.4-37 Relation of S11 and frequency for different turns of helix

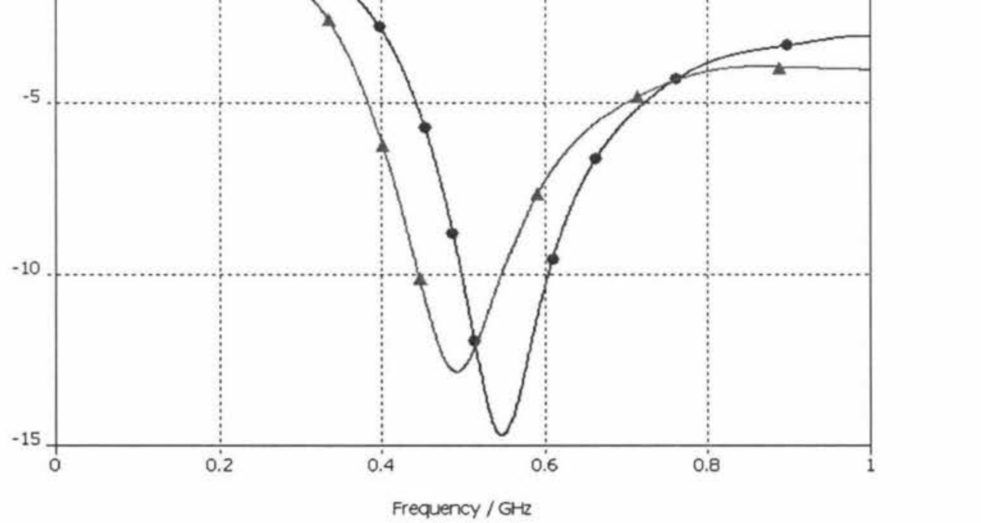


Fig.4-38 Relation of S11 and frequency for different radius of helix

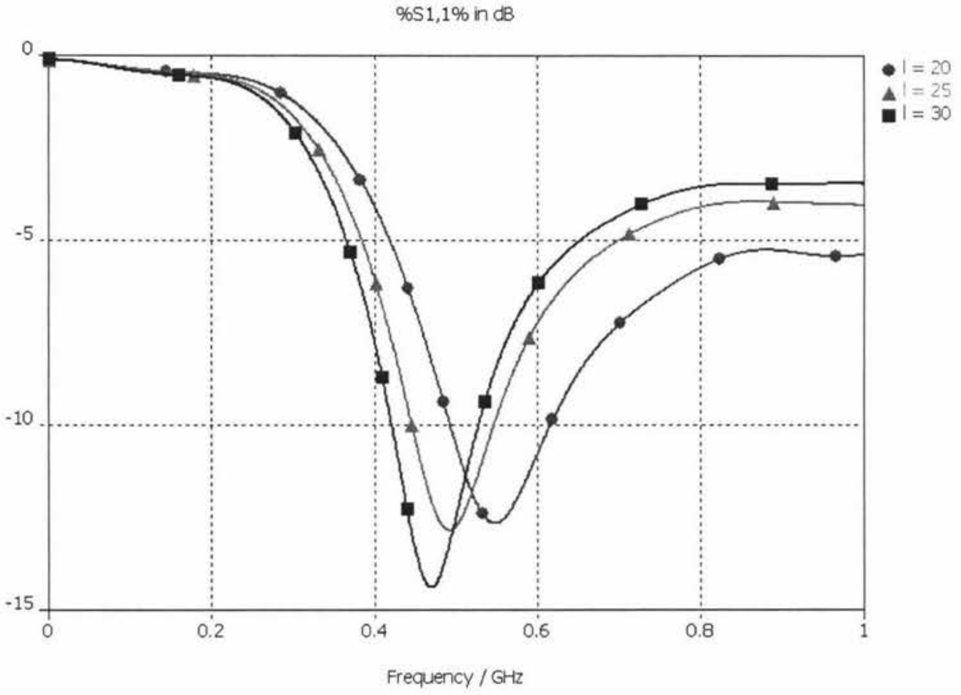


Fig.4-39 Relation of S11 and frequency for different L which is the length of straight line part of helix can been seen in figure geometry of antenna in Fig.4-30

As the most important parameter during the design of antenna, scattering parameter of S11 reveals the relation in term with incident and reflected power wave in the one-port network. Obviously, the small value of S11 is desired which means most part of incident power is transferred to the radiation power but reflect back. Under the condition of optimizing all parameter of helical antenna, S11 can be reduced to –14dB. In this situation, another important parameter VSWR(voltage stand wave ratio) is less than 1.5. Because the input resistance of helical antenna is about 46ohm, good matching between antenna and RWS-434 improves the efficiency of radiation. So the helical antenna designed by HFSS is satisfied the desired demands.

Chapter 5 Conclusions

5-1 Conclusion

Results from section 4-3 to section 4-7 show many qualities about the program to LabVIEW, program to the MCU, PCB design and antenna design. The two more important points in the LabVIEW program are the "Signal Generation Voice Menu.Vi" and the case to change the control values corresponding to the voice commands. All commands are stored in the file called Signal Generation Voice Menu.Vi. So far just television set commands are available. The users according to their habits can edit these commands. In order to conveniently use it, any abbreviated commands can be used.

For example, I replaced the "TV channel one" to "TV one". However, the users must use those commands that are easily remembered by themselves. A good hint is to use a special structure for each appliance, for instance, leading the name of the appliance such as "TV one" or "DVD one".

After adding new commands of different appliances, the users need to edit the case statement of Change the control values corresponding to the voice commands depending on what the commands are. Some popular brands of appliance companies provide the command list for their products and it can be obtain from the relevant website in the Internet. The author has some appliances' commands such TV, CD player and DVD of Phillips and Sony products available.

In the section 4-3 and 4-6-2, the microcontroller and programming to MCU are discussed. The programming results are quite good. During the process of manual generation of commands by VISA in LabVIEW, the oscilloscope was used to observe the waveforms in the development board. The waveforms show that the assembler programming can generate the correct command string in different IR protocols. At the same time, the commands are effective to control a Phillips TV. However, when I use the Dragon NaturallySpeaking to catch the commands of what has been said, sometimes, no response happens until the commands are repeated once or more. This situation will be improved after recognizing several phases or sentences. The probably reason is that Dragon NaturallySpeaking needs some time to recognize the identity of the speaker. So in the start stage of each usage, the accuracy of recognition is not good enough.

Although short of the valid instruments to check the performance of the antenna, to check the transmitter and receiver separately by changing the distance and position in different rooms is a useful exercise. The waveforms checked by real oscilloscope show that the transmitter and receiver can communicate in a real house environment. So I expect the speech driven remote controller will be installed in the ceiling of the room in where the TV, CD or house theater are placed.

5-2 Further work

In this moment, the remote controller board works well whereas the switch and dimmer board cannot trigger the triac although the transmitter and receiver work well. Therefore, more the troubleshooting is needed. In addition to the light control, we still need to control some other appliances such as an oven, air conditioner etc. The method can be to build home automation network HAN as discussed in section 1-2. We can use the radio as the media, and the further work is to adjust the control part of appliances in order to match the demands of HAN.

So far, we just use the headset microphone to catch the speech commands from the real acoustic environment. However, this mode is not convenient. In the future, I need to connect the Real-time implementation of a Microphone Array beam-former and check the performance of part of the speech recognition software. In a real acoustic environment, it is not enough that only one speaker can be recognized. So different people's training files generated by Dragon Naturally Speaking absolutely need to be stored and loaded.

As the part of whole project, personal identity recognition can be used to help not only load the different people's training files but also control appliances. For example, if the host's personal identity is set and shows that 18 °C is his or her favorite environment temperature in summer. Once the sensor recognizes the identity of host after he or she enter house, the air conditioner should switch on and change the current setting. We need to integrate all parts of project and control the house at the system level.

References

[1] Mozer, M.C, "An Intelligent Environment Must Be Adaptive", IEEE Intelligent System March-April 1999,Vol.14, Issue 2

[2] Wacks, K.; "Home systems standards: achievements and challenges" IEEE Communications Magazine, Volume: 40 Issue: 4, Page(s): 152 –159, April 2002

[3] Turcan, E.; Graham, R.L.; Hederen, J.; "Peering the smart homes", First International Conference on Peer-to-Peer Computing, 2001. Proceedings, Page(s): 103 -104 27-29 Aug. 2001

[4] EIA Home Automation System (CEBus) Standard EIA-600, Feb. 5th,1995

[5] EIA Home Automation standard (LonWorks) EIA-709.1-A.

[6] EIBA Brussel, EIBA Handbook Series, Release 3.0, 1999.

[7] Zhaohui Ye; Yindong Ji; Shiyuan Yang, "Home automation network supporting plug-and-play", IEEE Transactions on Consumer Electronics, vol50, Issue1, Feb 2004

[8] European committee for electrotechnonlogy standardization: CENELEC TC205WG5

[9] Kauffmann. P., Misson. M. "Study of a home automation network radio variant: the case of Teledomotis", Proceedings of Fifth IEEE Symposium on Computers and Communications, 2000, France

[10] John G. Proakis Digital communication 3rd edition

[11] S. G. Glisic ; P. A. Leppanen, "Wireless Communications: TDMA versus CDMA", Kluwer Academic Publishers, 2000.

[12] Nan Xie; Leung. H, "An analog multi-user spread spectrum technique for wireless home automation", IEEE Transactions on Consumer Electronics, Volume: 48 , Issue: 4 , Nov 2002

[13] Hakem. N; Misson. N, “Study of the encapsulation of two medium access methods for a wireless home automation network”, Proceeding of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2002, Volume: 1, pp15-18, Sept 2002

[14] B. Widrow, P.E. Mantley, L.J. Griffiths, B.B. Goode, "Adaptive Antenna Systems," *Proceedings of the IEEE*, vol. 55, December 1967.

[15] J. Capon, "High Resolution Frequency-Wavenumber Spectrum Analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408-1418, August 1969.

[16] R.O. Schmidt, “Multiple Emitter and Signal Parameter Estimation,” *Proceedings, RADC Spectral Estimation Workshop*, pp. 243-258, October 1979.

[17] R.N. McDonough, "Application of the Maximum-Likelihood Method and the Maximum-Entropy Method to Array Processing," in *Nonlinear Methods of Spectral Analysis*. S. Haykin, Ed., Springer-Verlag, New York, 1983.

[18] P.W. Howells, "Intermediate Frequency Sidelobe Canceler," U.S. Patent 3202990, August 24, 1965

[19] L.J. Griffiths and C.W. Jim, “An Alternative Approach to Linearly Constrained Adaptive Beamforming,” *IEEE Transactions on Antennas and Propagation*, vol. AP-30, pp. 27-34, January 1982.

[20] M.J. Levin, "Maximum-Likelihood Array Processing," in *Seismic Discrimination Semi-Annual Technical Summary Report*, M.I.T. Lincoln Laboratory, Lexington, MA, Technical Report DDC 455743, December 1964.

[21] O.L. Frost, III, "An Algorithm for Linearly-Constrained Adaptive Array Processing," *Proceedings, IEEE*, vol. 60, pp. 926-935, August 1972.

[22] Jan, E.; Flanagan, J.; "Microphone arrays for speech processing", Proceedings of IEEE International Symposium on Signals, Systems, and Electronics, 1995. ISSSE '95 Page(s): 373 –376, 25-27 Oct 1995

[23] Flanagan, J.L., "Auto-directive sound capture: towards smarter conference rooms" IEEE Intelligent Systems, Volume: 14, Issue: 2, Pages: 16 – 19, March-April 1999

[24] Goodwin, M.M.; Elko, G.W.; "Constant beamwidth beamforming" IEEE International Conference on Acoustics, Speech, and Signal Processing 1993. ICASSP-93, Volume: 1, Pages:169 – 172, 27-30 April 1993

[25] Schmidt, R. O. "Multiple emitter location and signal parameter estimation", IEEE Transaction on Antennas and Propagation.,vol.3,issue 3,Mar 1986

[26] I.N.EL-Behery and R.h.MacPhie, "Maximum likelihood estimation of source parameter from time-sampled outputs of a linear array", Journal of Acoust. Soc. Am.,vol.62, no.1, July 1977.

[27] Omologo, M.; Svaizer, P.; "Use of the crosspower-spectrum phase in acoustic event location", IEEE Transactions on Speech and Audio Processing, Volume: 5, Issue: 3, Pages: 288 – 292, May 1997

[28] M. Omologo, P. Svaizer, "Acoustic Event Localization using a crosspower-spectrum phase based technique", Proc. ICASSP94, 1994

[29] Nishiura. T.; Yamada. T.; Nakamura. S.; Shikano. K, "Localization of multiple

sound sources based on a CSP analysis with a microphone array", Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2000. ICASSP 2000, Volume: 2, 5-9 June 2000

[30] Guangshu, Hu; "Digital signal processing", Tsinghua Unversity press, 1st edition, 1997

[31] Pirz, F.; "Design of a wideband, constant beamwidth, array microphone for use in the near-field" IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'79, Volume: 4, Pages: 318 – 321, Apr 1979

[32] Ward, D.B.; Kennedy, R.A.; Williamson, R.C.; "FIR filter design for frequency invariant beamformers" IEEE Signal Processing Letters, Volume: 3, Issue: 3, Pages: 69 -71, March 1996

[33] Jan, E.; Svaizer, P.; Flanagan, J.L.; "Matched-filter processing of microphone array for spatial volume selectivity" IEEE International Symposium on Circuits and Systems, 1995, ISCAS '95, Volume: 2 , 28, Pages:1460 – 1463, April-3 May 1995

[34] Rabinkin, D.; Renomeron, R.; Flanagan, J.; Macomber, D.F.; "Optimal truncation time for matched filter array processing", Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '98. Volume: 6, Page(s): 3629 –3632, 12-15 May 1998

[35] Ea-Ee Jan; Flanagan, J.; "Sound capture from spatial volumes: matched-filter processing of microphone arrays having randomly-distributed sensors" Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing, 1996, ICASSP-96. Volume: 2, Pages: 917 – 920, 7-10 May 1996

[36] M.Omologo, P.Svaizer, "Acoustic source location in noisy and reverberant

environment using CSP analysis", Proc. ICASSP96, 1996

[37] P. Svaizer, M. Matassoni and M. Omologo, "Acoustic source location three-dimensional space using CSP", Proc.ICASSP97, 1997

[38] La Bouquin-Jeannes. R.; Azirani. A.A.; Faucon. G., "Enhancement of speech degraded by coherent and incoherent noise using a cross-spectral estimator" Speech and Audio Processing, IEEE Transactions on , Volume: 5, Issue: 5 , Sept. 1997

[39] Hang HU, "Speech signal processing", Herbin Institute of Technology press, May 2000

[40] Nishiura, T.; Nakamura, S.; Shikano, K, "Talker localization in a real acoustic environment based on DoA estimation and statistical sound source identification", Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002.. (ICASSP '02)., Volume: 1 , 13-17 May 2002

[41] Nishiura T; Yamada T; Nakamura. S; Shikano,K, "Speech enhencement by multiple beamforming with reflection signal equalization", Proc.ICASSP2001, May 2001

[42] Wilson, K.W.; Darrell, T, "Audio-video array source localization for intelligent environments", Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002. (ICASSP '02)., vol.2, 2002

[43] Kaynak, M.N.; Qi Zhi; Cheok, A.D.; Sengupta, K.; Ko Chi Chung, "Audio-visual modeling for bimodal speech recognition", IEEE International Conference on Systems, Man, and Cybernetics, 2001, Volume: 1, 7-10 Oct. 2001

[44] Gary W. Johnson, "LabVIEW Graphical Programming", 2nd edition, McGraw-Hill, 1997

[45] Evangelos Petroutsos, "ActiveX Development with Visual Basic 5", 1st edition,

Ventana communication Group, 1997

[46] John D.Kraus, "Antennas", 2^{nd} edition, McGraw-Hill, 1988

[47] D.K.Cheng, "Electromagnetic field and wave" (translated to Chinese by Yaotong, Zhao and Binghong, li), Shanghai Jiao Tong University press, 1984

[48] http://www.xs4all.nl/~sbp/knowledge/ir/rc5.htm

http://www.xs4all.nl/~sbp/knowledge/ir/sirc.htm

[49] http://www.Microchip.com

[50] http://www.rentron.com/Files/rf.pdf

http://www.rentron.com/Files/ht-12e.pdf

http://www.rentron.com/Files/ht-12d.pdf

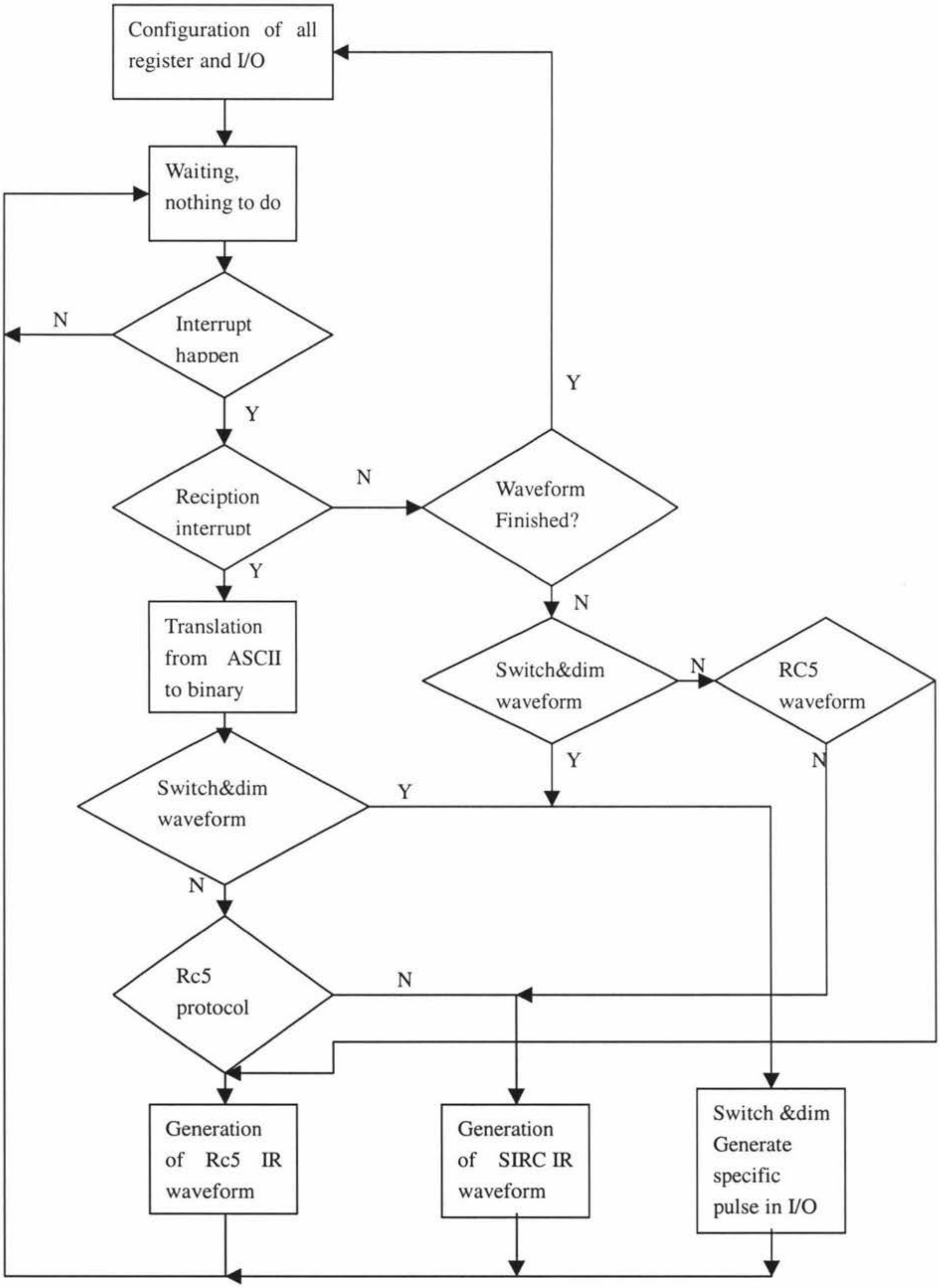
[51] http://www.aaai.org/AITopics/html/rooms.html

[52] http://oxygen.lcs.mit.edu/Overview.html

[53] Kent Smith, "ANTENNAS FOR LOW POWER APPLICATIONS",

http://www.web-ee.com/primers/files/antenna.pdf

Appendix A



Configuration of all register and I/O
Waiting, nothing to do
N
Interrupt happen
Y
Y
N
Reciption interrupt
Waveform Finished?
Y
N
Translation from ASCII to binary
Switch&dim waveform
N
RC5 waveform
Y
N
Switch&dim waveform
Y
N
Rc5 protocol
N
Generation of Rc5 IR waveform
Generation of SIRC IR waveform
Switch &dim Generate specific pulse in I/O

```
;This program is for remote controller which can generate IR protocol code to control TV, CD player etc.

    LIST P=18F452
    #include <P18F452.INC>    ; File contains addresses for register and bit names

;************************************************************
; reset and interrupt vectors

    org   0x00000             ; Reset Vector Address
    goto  Start

    org   0x00008             ; interrupt Vector Address
    goto  ISR                 ; goto Interrupt Service Routine

;************************************************************
; program code starts here
        byte0    equ    0x0020    ;store the data from serial port
        byte1    equ    0x0021    ;store the data from serial port
        byte2    equ    0x0022
        byte3    equ    0x0023    ;rc5   start parameter
        byte4    equ    0x0024    ;sony start parameter
        byte5    equ    0x0025    ;12-bit command including add and command
        byte6    equ    0x0026    ;high-lower check bit
        byte7    equ    0x0027    ;sony protocol bit
        byte8    equ    0x0028    ;rc5 protocol bit
        byte9    equ    0x0029    ;sharp protocol bit
        byte10   equ    0x002a    ;second control byte for sony
        byte11   equ    0x002b    ;second control bytr for rc5
        byte12   equ    0x002c    ;check bit for rc5 and sony's first 6-bit,whether it is finished
        byte13   equ    0x002d    ;check bit for sony' start-bit,whether it is finisned
        byte14   equ    0x002e    ;check bit for dimmer
        byte15   equ    0x002f    ;check bit for switch_on_off
        byte16   equ    0x0030    ;control the dimmer time
        byte17   equ    0x0031    ;protocol rc5 two start bits both are 1
        byte18   equ    0x0032    ;check bit whether rc5 two start bits is finished
        byte19   equ    0x0033
        byte20   equ    0x0034
        byte21   equ    0x0035
        byte22   equ    0x0036
```

```
        byte23    equ     0x0037
        byte24    equ     0x0038


    org         0x00020

Start
;           clrf        PORTB                   ; clear all bits of PORTB
;           BCF         STATUS,IRP
            ;   bsf         STATUS,RP0
        movlw   B'10000000'
        movwf   TRISC                   ; Set rC7 as serial input and rC2 as pwm out
        bsf     PORTC,0             ;Rc0 maintain high until dimmer is working
        call    InitializeUSART             ; configure serial receiver module
        call    pwm_generater
        bsf     PIE1,TMR1IE
        movlw   B'00000001'
        movwf   byte2                   ; the content of byte2 is 1
        movlw   B'00000111'
        movwf   byte5                   ; the content of byte5 is 7, six bit is used to generate waveform for all
ir protocol
        movlw   B'00000000'
        movwf   byte6                   ;high-low bit
        movlw   B'00000000'
        movwf   byte7                   ;sony protocol bit, it will be set when sony is identified
        movlw   B'00000000'
        movwf   byte8                   ;rc5 protocol bit,it will be set when rc5 is identified
        movlw   B'00000000'
        movwf   byte9                   ;sharp protocol bit,it will be set when sharp is identified
        movlw   B'00000111'
        movwf   byte10                  ; the content of byte10 is 7
        movlw   B'00000111'
        movwf   byte11                  ; the content of byte11 is 7 both 2 byte 6 bit is used to generate
waveform
        movlw   B'00000000'
        movwf   byte12                  ; the initialised content of byte12 is 0,it will be set when first byte is
                                          finished to generate
                                        ; waveform
        movlw   B'00000000'
        movwf   byte13                  ;it will be set when start bit of sony is finished to generate waveform
        movlw   B'00000000'
```

```
        MOVWF    byte14                 ;if byte14 is 1,dimmer works
        movlw    B'00000000'
        movwf    byte15                 ;if byte15 is 1 ,switch works
        movlw    B'00000111'
        movwf    byte16
        movlw    B'00000011'
        movwf    byte17                 ;two start_bit
        movlw    B'00000000'
        movwf    byte18                 ;if rc5 two start bit finished, byte18 is set 1
        movlw    B'00001001'
        movwf    byte22                 ;transfer 8-bit to first byte
        movlw    B'00000000'
        movwf    byte23                 ;if transfering of first byte is finished, is set 1,otherwise 0. initial is 0
        movlw    B'00000000'
        movwf    byte24                 ;if transfering of two byte is finished, is set 1,otherwise 0. initial is 0

Main goto  Main                 ; do nothing loop

;************************************************************
; Service   interrupt

ISR
; Save context (WREG and STATUS) if required.
        btfss    PIR1,RCIF              ; check whether serial reception cause interrupt?
        goto     other_intservice
        goto     for_translation
for_translation
        call     translation
        btfss    byte24,0               ;check if the procedure of transfer is finished
        goto     EndISR                 ;the procedure of transfer is NOT finished
        nop
identify

        movlw    B'11000000'
        andwf    byte0,0                ; get the protocol bit7-6
        btfss    STATUS,Z
        goto     continue_check
        goto     protocol_rc5
continue_check
        bcf      STATUS,C
        rlcf     WREG,0
```

```
        rlcf      WREG,0
        rlcf      WREG,0
        decfsz    WREG,0
        goto      check1
        goto      protocol_sony
check1
        decfsz  WREG,0
        goto      light_dimmer
;*******************************************************
sharp_protocol
        movwf     byte2
        decfsz   byte2,1
        goto      EndISR

;*******************************************************
light_dimmer
        bcf       STATUS,C
        RLCF        byte0,1
        RLCF        byte0,1
        RLCF        byte0,1
        BTFSS       STATUS,C
        GOTO         SWITCH_ON_OFF
        GOTO         DIMMING
SWITCH_ON_OFF
        bcf       byte15,0
        call     time_6.5ms
        CLRF       TRISA
        MOVF       byte1,W
        MOVWF     PORTA
        bcf       PORTC,0
      bsf     T1CON,TMR1ON
        goto      EndISR

time_6.5ms
        MOVLW       B'00000000'
        MOVWF       T1CON                          ;timing at least 50ms for switch
        MOVLW     B'00000000'
        MOVWF     TMR1L
        MOVLW     B'11000000'
        MOVWF     TMR1H
        RETURN
```

```
DIMMING
        bsf     byte14,0
        call    time_6.5ms
        CLRF    TRISA
        movf    byte1,W
        movwf   PORTA
        bcf     PORTC,0
      bsf   T1CON,TMR1ON
        goto    EndISR

time_120ms
        MOVLW   B'00110000'
        MOVWF   T1CON
        MOVLW   B'10000011'
        MOVWF   TMR1L
        MOVLW   B'01101101'
        MOVWF   TMR1H
        BSF     byte15,0
        bcf     byte14,0
        return

;*************************************************************
dimmer

        BTFSS   byte15,0
        goto    dimmercheck
        goto    SWITCH_ON_OFF
dimmercheck
        btfss   byte14,0
        goto    switch_service
        goto    dimmer_service
switch_service
        bcf     PIR1,TMR1IF
        BSF     PORTC,0
        GOTO    Start
dimmer_service
        bcf     PIR1,TMR1IF
        BSF     PORTC,0
        call    time_120ms
        bsf     T1CON,TMR1ON
```

```
            goto      EndISR
;************************************************************
protocol_rc5
        bcf     STATUS,C
        rlcf    byte0,1
        rlcf    byte0,1
        bsf     byte8,0
        call    Tmr1_rc5
rc5
        BTFSS       byte2,0             ;CHECK WHETHER THE CONTENT OF BYTE2 IS 1
        GOTO        continue_timing
        btfss       byte18,0
        goto        rc5_twostartbit
        goto        loop1
rc5_twostartbit
        decfsz      byte17,1
        goto        logic1_rc5
        bsf         byte18,0
        bcf         PORTC,2
loop1
        bcf         STATUS,C
        btfss       byte12,0
        goto        first_bit
        goto        start_second_byte_rc5
first_bit
        decfsz      byte5,1
        goto        first_byte_timing
        goto        start_second_byte_rc5
first_byte_timing
        BCF         STATUS,C
        rlcf        byte0,1
bit_check_rc5
        btfss       STATUS,C
        goto        logic0_rc5
        NOP
logic1_rc5
        bcf         byte2,0
        bsf         byte6,0
        goto        starttiming_rc5
logic0_rc5
        bcf         byte2,0
```

```
        bcf       byte6,0
        goto      starttiming_rc5
continue_timing
        bsf       byte2,0
        BTFSS     byte6,0
        GOTO      lowoutput
        GOTO      highoutput
start_second_byte_rc5
        bsf       byte12,0
        bcf       STATUS,C
        rlcf      byte1,1
        decfsz    byte11,1
        goto      bit_check_rc5
        bcf       PORTC,2
        goto      Start
starttiming_rc5
        btfsc  byte6,0
        goto   lowoutput
        goto   highoutput
highoutput
        CALL    pwm_generater_rc5
        BSF     T2CON,TMR2ON        ;START TO GENERATE 38KHZ WAVeFORM
        call   Tmr1_rc5
        BSF     T1CON,TMR1ON        ;timing 900us
        goto   EndISR               ; return from ISR
lowoutput
        BCF     PORTC,2
        CLRF    CCP1CON             ;FORCE THE RC2 AS LOW OUTPUT
        bcf     T2CON,TMR2ON        ; STOP TO GENERATE 38KHZ WAVEFORM
        call  Tmr1_rc5
        BSF     T1CON,TMR1ON
        goto   EndISR

Tmr1_rc5
        MOVLW   B'00000000'
        MOVWF   T1CON           ; TIMER0 IS 16-BIT TIMER WITH 1:1 PRESCALE
        MOVLW   B'00110101'  ;at 10mhz
        MOVWF   TMR1L
        MOVLW   B'11110111'  ;at 10mhz
        MOVWF   TMR1H           ;GENERATES 900us timing
        return
```

```
;***********************************************************************
RcvError
     bcf     RCSTA,CREN        ; Clear receiver status
     bsf     RCSTA,CREN
       Bsf        PORTC,RC4
     goto    EndISR            ; go to end of ISR, restore context, return
;***********************************************************************
translation
    bcf          PIR1,RCIF          ;this subroutine will check whether the data received is correct
    movlw        B'00000110'   ; Mask out unwanted bits
    andwf RCSTA,W       ; Check for errors
    btfss  STATUS,Z; Was either error status bit set?
    goto         RcvError
    movf         RCREG,W          ;NO error then transfering data
    movwf        byte19
    btfss       byte23,0
    goto         transfer_first_byte
    goto         transfer_second_byte
transfer_first_byte
    decfsz      byte22,1
    goto         transfering_first_byte
    goto         check_transfer_second_byte
transfering_first_byte
    movlw        B'00110000'
    XORwf        byte19,0
    call        ascii_identity_first
    return
check_transfer_second_byte
    bsf          byte23,0
    movlw        B'00001000'
    movwf        byte22
transfer_second_byte
    decfsz       byte22,1
    goto         transfering_second_byte
    call         transfering_second_byte
    bsf          byte24,0
    return
transfering_second_byte
    movlw        B'00110000'
    XORwf        byte19,0
    call         ascii_identity_second
```

```
ascii_identity_first
    btfss   STATUS,Z            ; byte19 is ascii code 30 which means 0
    goto    character1_1
    goto    character0_1
character1_1
    bsf     STATUS,C
    RLCF    byte0,1
    return
character0_1
    bcf     STATUS,C
    RLCF    byte0,1
    return
ascii_identity_second
    btfss   STATUS,Z            ; byte19 is ascii code 30 which means 0
    goto    character1_2
    goto    character0_2
character1_2
    bsf     STATUS,C
    RLCF    byte1,1
    return
character0_2
    bcf     STATUS,C
    RLCF    byte1,1
    return
;*****************************************************************************
protocol_sony
        bcf     STATUS,C
        bsf     byte7,0         ;sony protocol
        rlcf    byte0,1
        rlcf    byte0,1
        goto    highoutput_sony_2.4ms
highoutput_sony_2.4ms
        call    tmr1_sony_2.4ms
        CALL    pwm_generater_sony
        BSF     T2CON,TMR2ON    ;START TO GENERATE 38KHZ WAVeFORM
        BSF     T1CON,TMR1ON
        goto    EndISR          ; return from ISR
highoutput_sony_1.2ms
        call    tmr1_sony_1.2ms
        CALL    pwm_generater_sony
        BSF     T2CON,TMR2ON    ;START TO GENERATE 38KHZ WAVeFORM
```

```
        BSF      T1CON,TMR1ON
        goto     EndISR                 ; return from ISR
highoutput_sony_0.6ms
        call     tmr1_sony_0.6ms
        CALL     pwm_generater_sony
        BSF      T2CON,TMR2ON           ;START TO GENERATE 38KHZ WAVeFORM
        BSF      T1CON,TMR1ON
        goto     EndISR                 ; return from ISR
lowoutput_sony_0.6ms
        bcf      T2CON,TMR2ON           ; STOP TO GENERATE 38KHZ WAVEFORM
        call     tmr1_sony_0.6ms
        BSF      T1CON,TMR1ON
        bcf      PORTC,2
        CLRF     CCP1CON                ;FORCE THE RC2 AS LOW OUTPUT
        goto     EndISR
;*************************************************************
tmr1_sony_1.2ms
          MOVLW    B'00000000'
          MOVWF    T1CON                 ; TIMER1 IS 16-BIT TIMER WITH NO PRESCALE
          MOVLW    B'01000111'           ;at 10mhz
          MOVWF    TMR1L
          MOVLW    B'11110100'           ;at 10mhz
          MOVWF    TMR1H                 ;TIMER1 WORK AS 1.2mS TIMING
          return
;**************************************************************
tmr1_sony_2.4ms
          MOVLW    B'00000000'
          MOVWF    T1CON                 ; TIMER1 IS 16-BIT TIMER WITH NO PRESCALE
          MOVLW    B'10001111'           ;at 10mhz
          MOVWF    TMR1L
          MOVLW    B'11101000'           ;at 10mhz
          MOVWF    TMR1H                 ;TIMER1 WORK AS 2.4mS TIMING
          return
;**************************************************************
tmr1_sony_0.6ms
          MOVLW    B'00000000'
          MOVWF    T1CON                 ; TIMER1 IS 16-BIT TIMER WITH NO PRESCALE
          MOVLW    B'00101100'           ;at 10mhz
          MOVWF    TMR1L
          MOVLW    B'11111010'           ;at 10mhz
          MOVWF    TMR1H                 ;TIMER1 WORK AS 600US TIMING
```

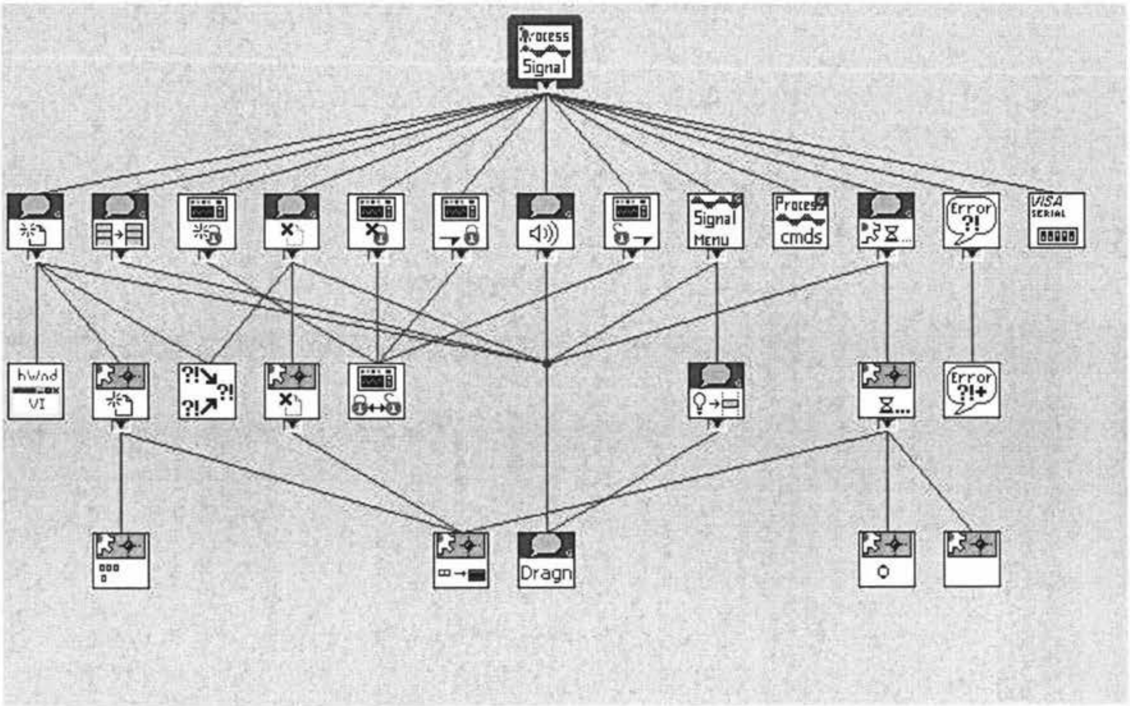
```
        return
;****************************************************************
EndISR
      ; Restore context if saved.
      retfie                    ; Return, enables GIE
;************************************************************
InitializeUSART

      movlw    B'01000000'          ; Make BAUD RATE IS 9600 at 10mhz
      movwf       SPBRG
      clrf    RCSTA       ; continuous receive
                    ; initialise asynchronous port as receiver with 8bit
      BSF      TXSTA,BRGH
      bcf     TXSTA,SYNC                    ; asynchrnous receive
      bsf     RCSTA,SPEN                    ; receive enable
      BSF      RCSTA,CREN
      bsf     PIE1,RCIE          ; Enable receive interrupt
      bsf     INTCON,PEIE            ; Enable peripheral interrupts
      bsf     INTCON,GIE             ; Enable Global interrupts
      return
;************************************************************
other_intservice
        btfss     PIR1,TMR1IF
        goto      EndISR
timr1service
        BCF       PORTC,2
        bcf      PORTC,2
        bcf      PIR1,TMR1IF
        bcf      T1CON,TMR1ON
        bcf      T2CON,TMR2ON                ;stop 38khz waveform
        btfss    byte7,0
        goto      continue_check_protocol_bit1
        goto      sony
continue_check_protocol_bit1
        btfss    byte8,0
        goto      continue_check_protocol_bit2
        goto      rc5
continue_check_protocol_bit2
        btfss    byte9,0
        goto      dimmer
        goto      sharp
```

```
;*********************************************************
sony
        btfss   byte6,0             ; check whether 0.6ms low output is finish
        goto    unfinished_low_output
        goto    bit_check
unfinished_low_output
        bsf     byte6,0
        goto    lowoutput_sony_0.6ms

bit_check

        bcf     STATUS,C
        bcf     byte6,0
        btfss   byte12,0
        goto    first_bit_sony
        goto    second_bit_sony

first_bit_sony
        decfsz  byte10,1
        goto    loop2
second_bit_sony
        goto    start_second_byte
loop2
        RLCF    byte0,1
        btfss   STATUS,C
        goto    logic0
logic1
        goto    highoutput_sony_1.2ms
logic0
        goto    highoutput_sony_0.6ms

start_second_byte
        bsf     byte12,0
        bcf     STATUS,C

        decfsz  byte11,1
        goto    loop3
        goto    Start
loop3
        rlcf    byte1,1
        btfss   STATUS,C
```

```
        GOTO      logic0
        goto      logic1
;**********************************************************************
sharp
        goto      EndISR
;************************************************************
;This is used to configure the pwm mode and generate a wave form which duty cycle is able to change

pwm_generater_rc5
       movlw   B'01000100'     ;at 10mhz
       movwf   PR2              ;the value in pr2 is equal to 68
       movlw   B'00010111'     ; the value of   CCP1R1L for the duty is 33%
       MOVWF   CCPR1L
       goto    pwm_generater
pwm_generater_sony
       movlw   B'00111110'     ;at 10mhz
       movwf   PR2              ;the value in pr2 is equal to 62
       movlw   B'00010110'     ; the value of   CCP1R1L for the duty is 33%
       MOVWF   CCPR1L
pwm_generater
       CLRF    CCP1CON
       BSF     CCP1CON,CCP1M2
       BSF     CCP1CON,CCP1M3    ;configure the ccp1 as pwm mode,38khz with 33% duty cycle
       MOVLW   B'00000000'
       MOVWF   T2CON             ;configure the timer2 as 8-bit timer with 1:1 prescale
       CLRF    TMR2
       BCF     PIE1,TMR2IE
       RETURN
END
```

Appendix B

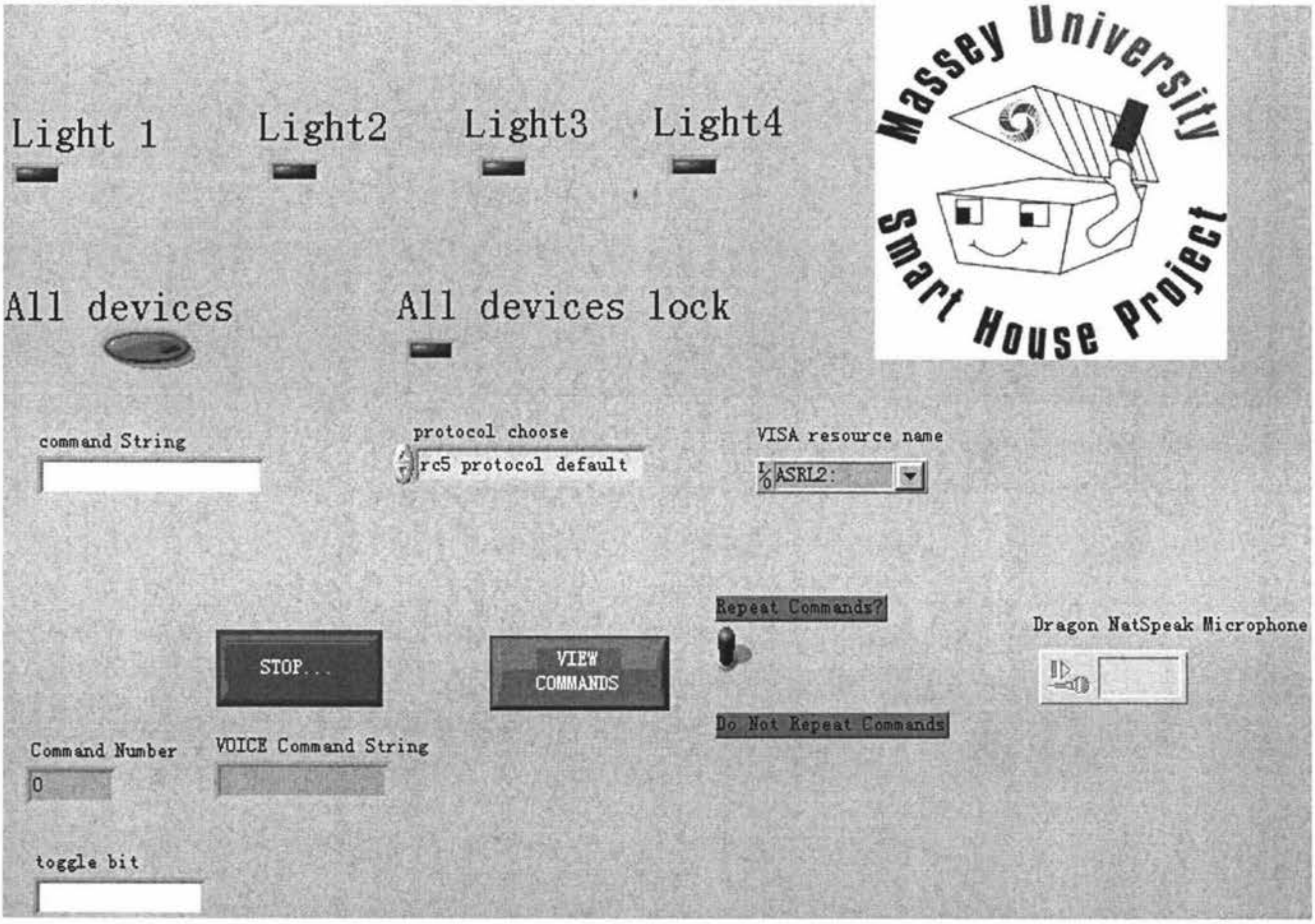
The hierarchy of LabVIEW program is shown in below.



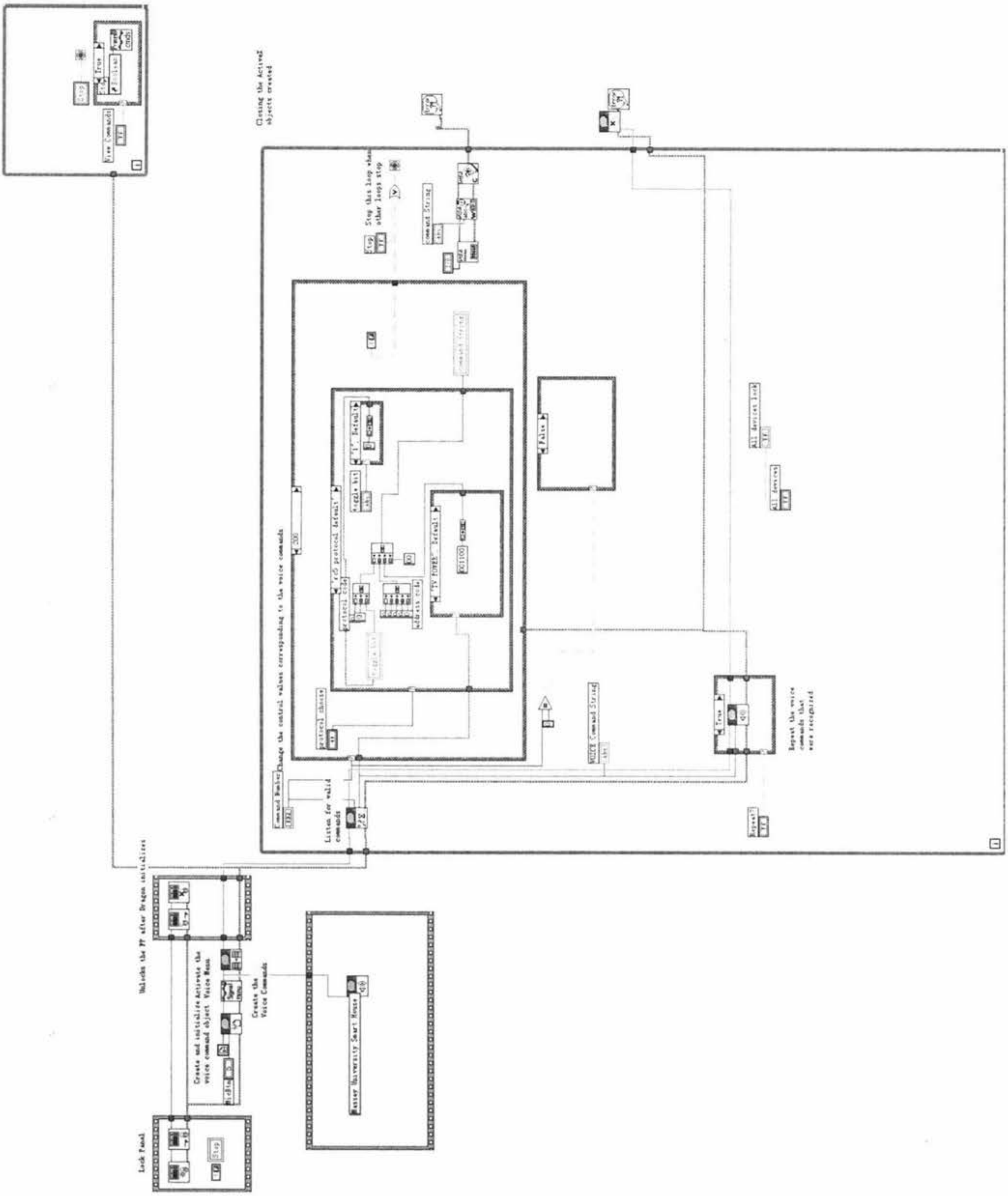
Dragon Signal Generation -2.vi
This VI illustrates adding Voice recognition and speech feedback to a user interface. This example without voice recognition ships with LabVIEW.

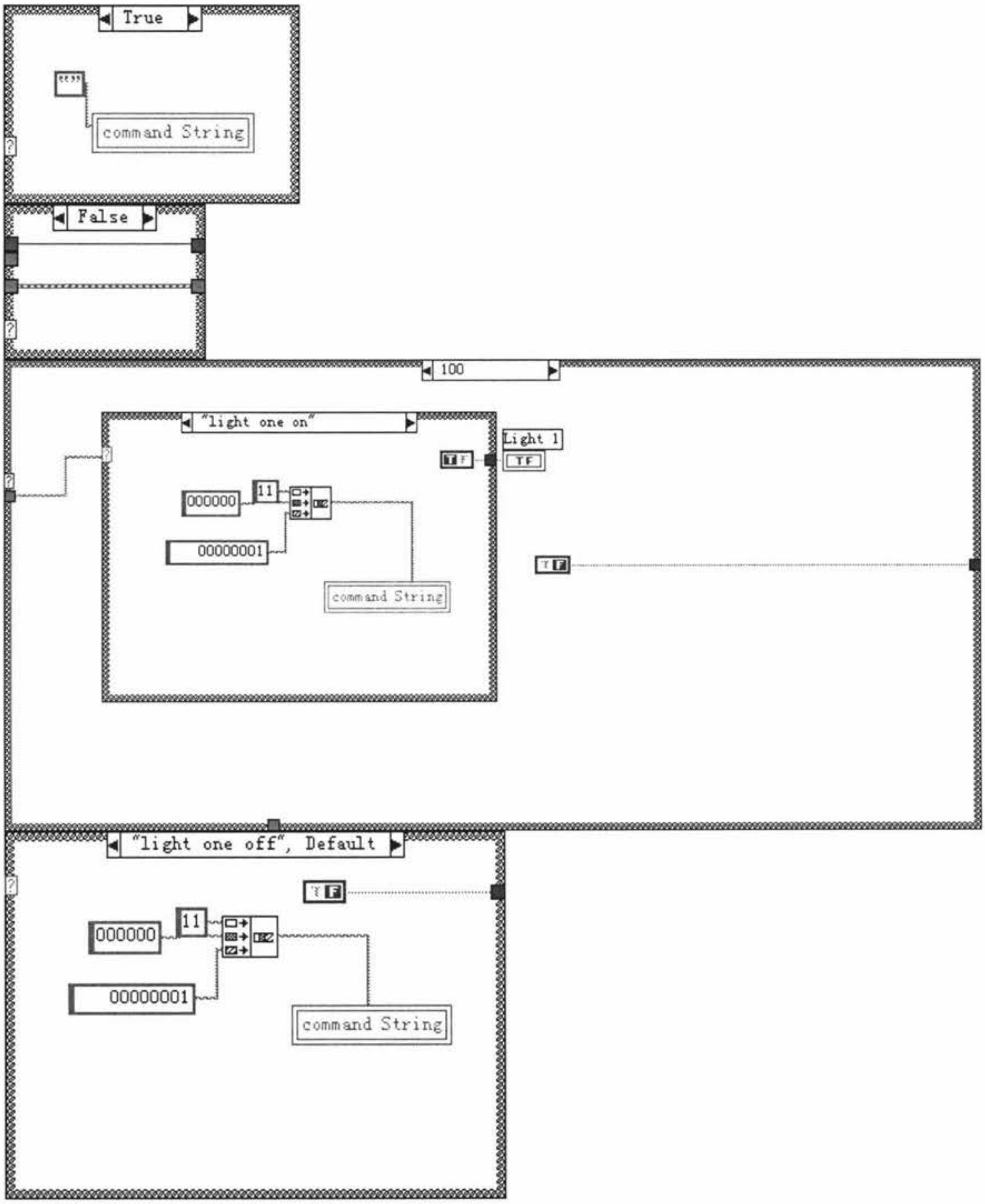
Connector Pane

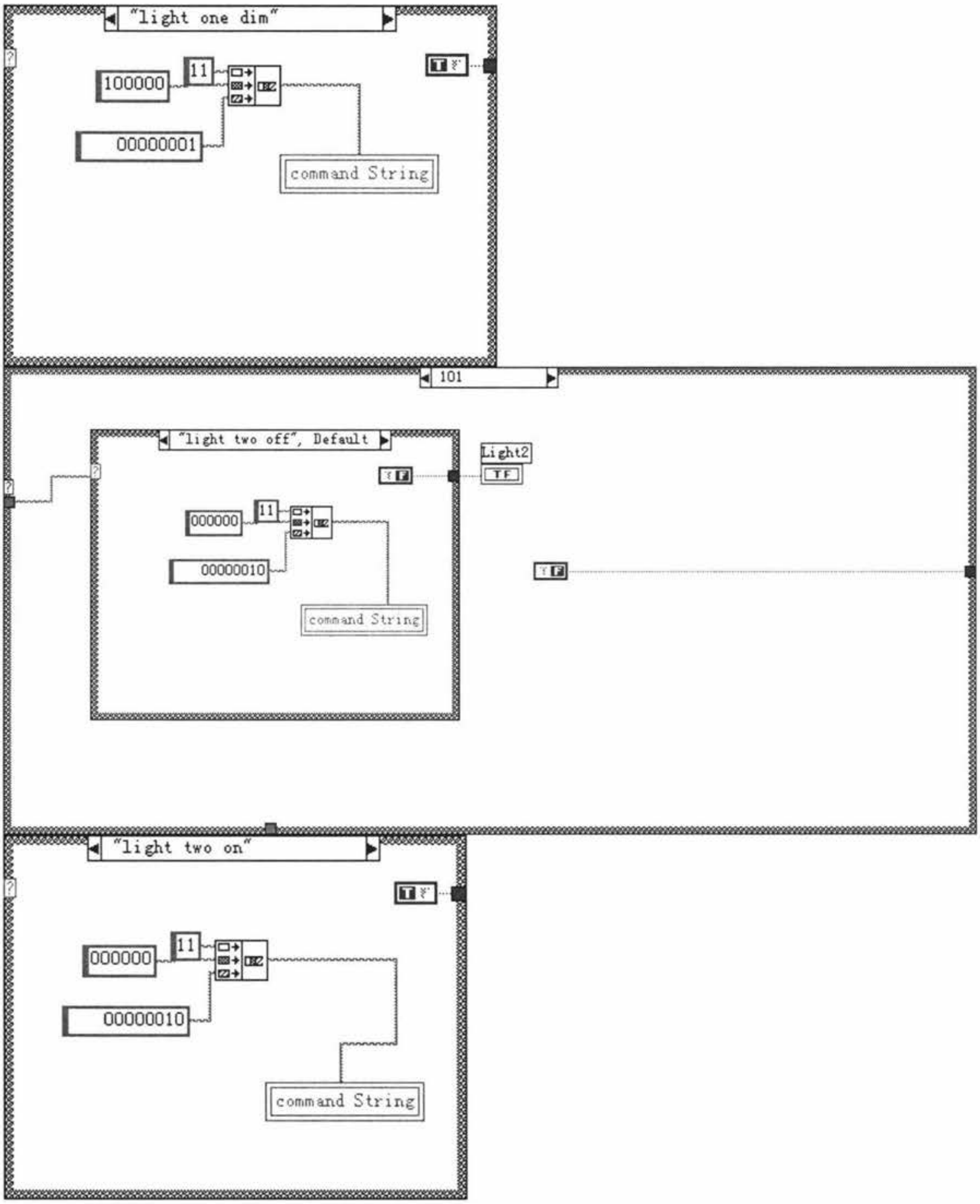
Front Panel

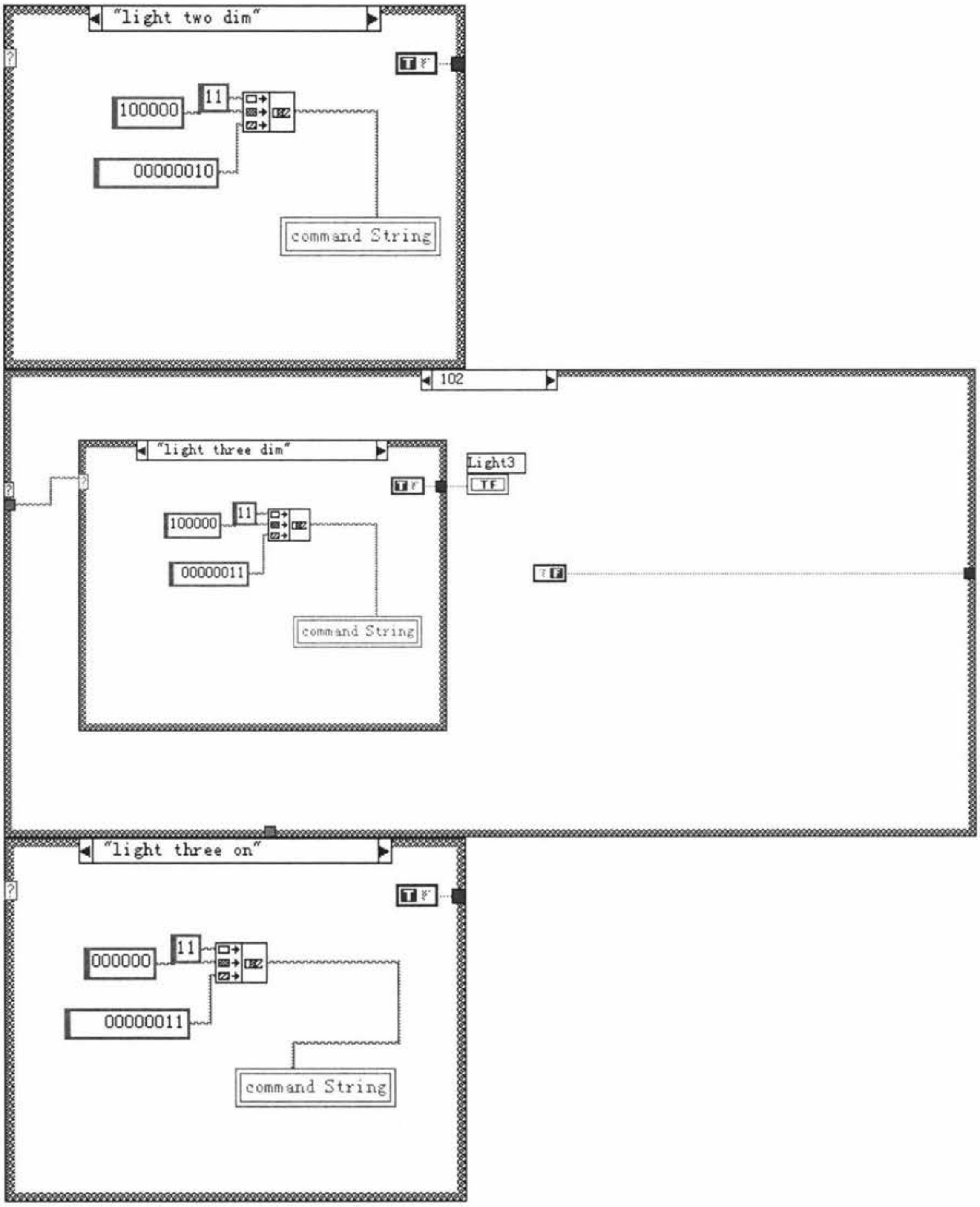


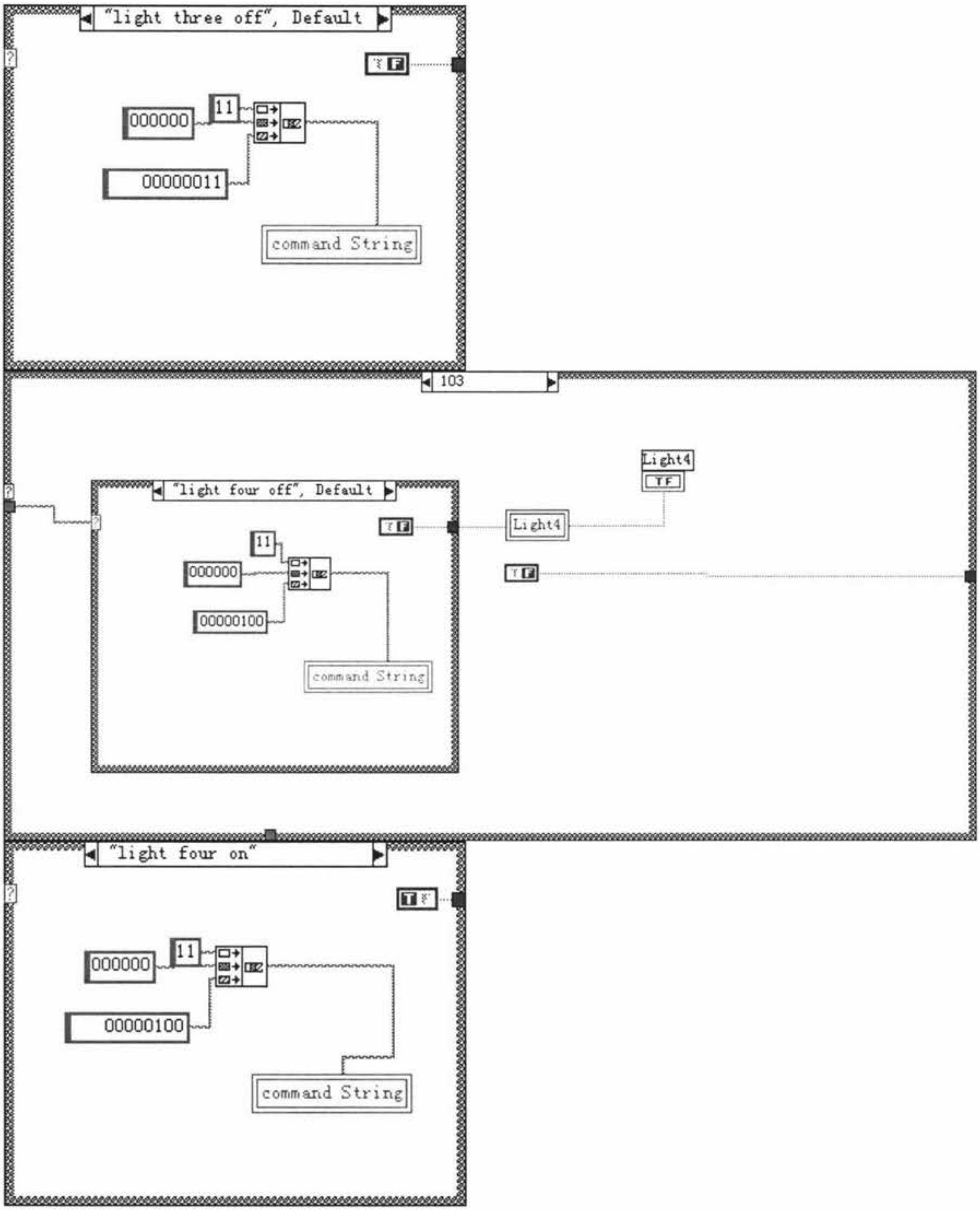
Block Diagram

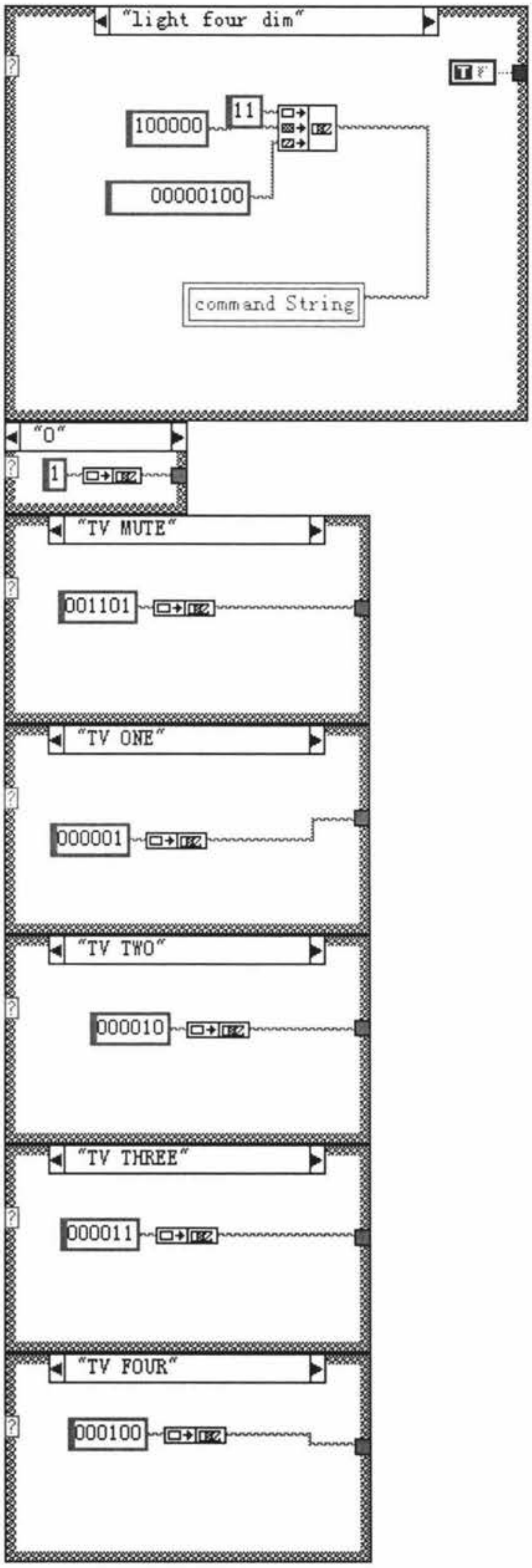

View Commands
Stop
Closing the ActiveX objects created
Stop this loop when other loops stop
Command String
Change the control values corresponding to the voice commands
protocol choose
VOICE Command String
All devices Lock
All devices
Repeat the voice commands that were recognized
Repeat?
Command Number
Listen for valid commands
Unlocks the FP after Dragon initializes
Create and initialize Activate the voice command object Voice Menu
Create the Voice Commands
Massey University Smart House
Lock Panel

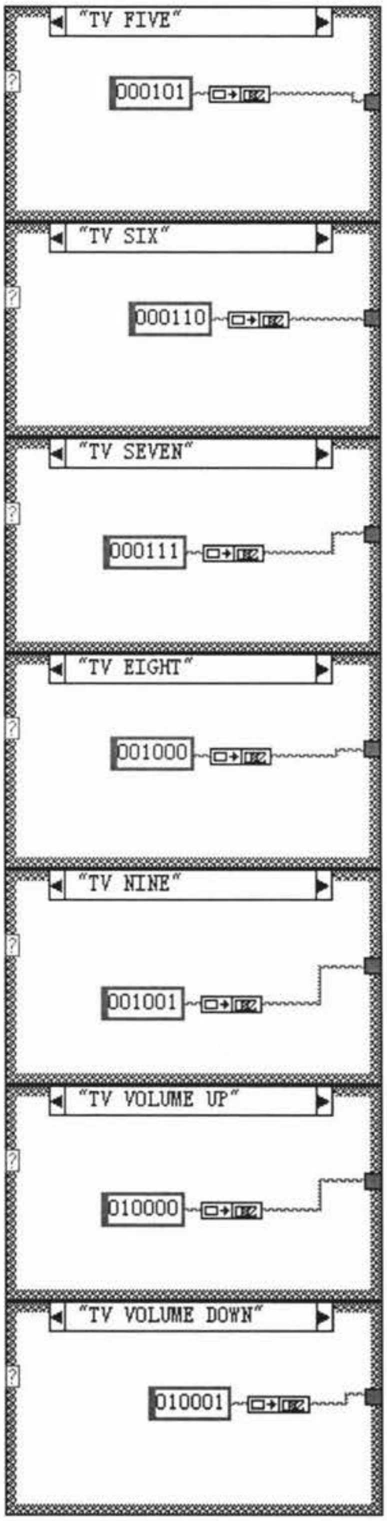

True
command String
False
100
"light one on"
Light 1
000000
11
00000001
command String
"light one off", Default
000000
11
00000001
command String

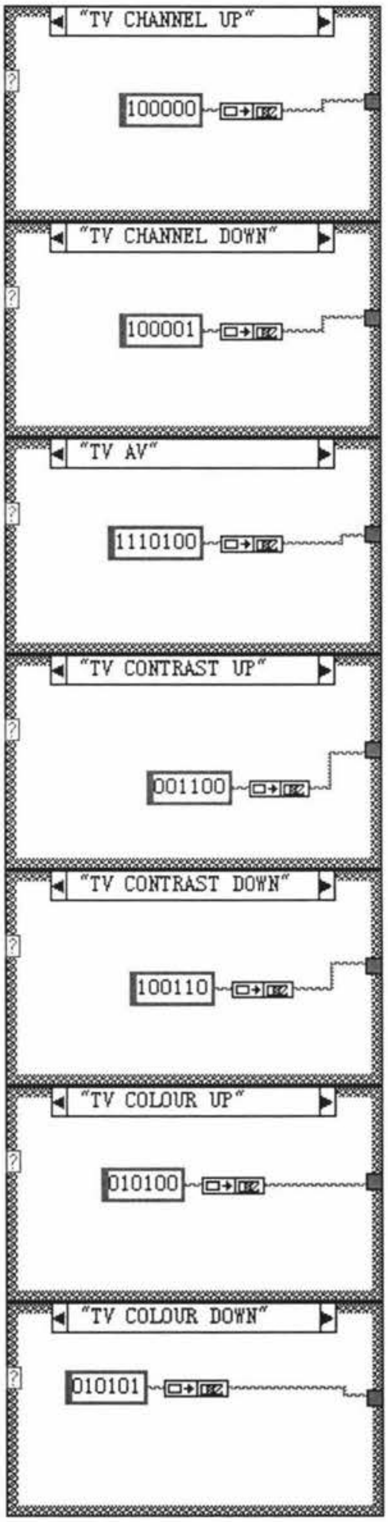

"light one dim"
100000
11
00000001
command String
101
"light two off", Default
Light2
000000
11
00000010
command String
"light two on"
000000
11
00000010
command String

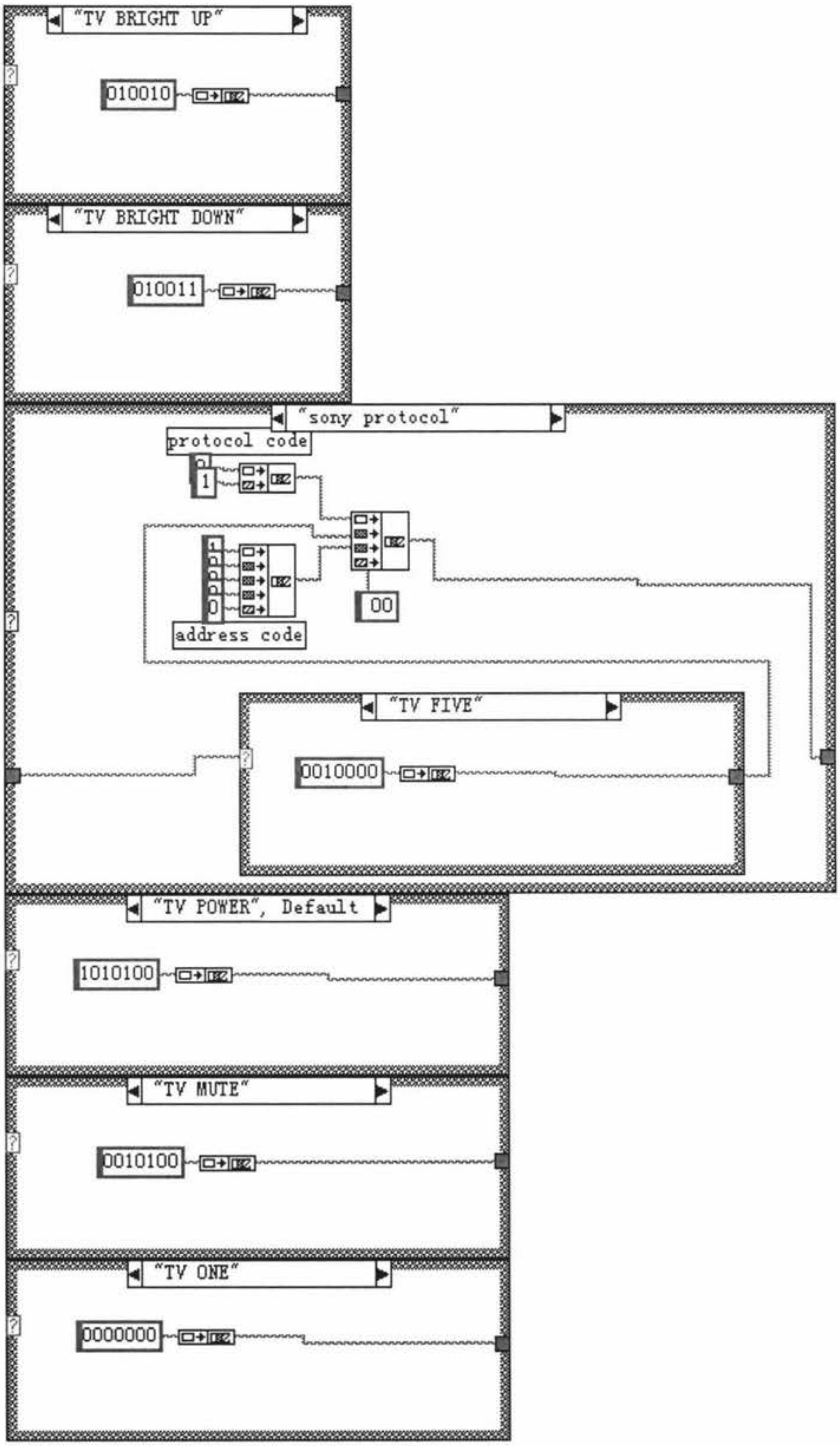

"light two dim"
100000
11
00000010
command String
102
"light three dim"
Light3
100000
11
00000011
command String
"light three on"
000000
11
00000011
command String

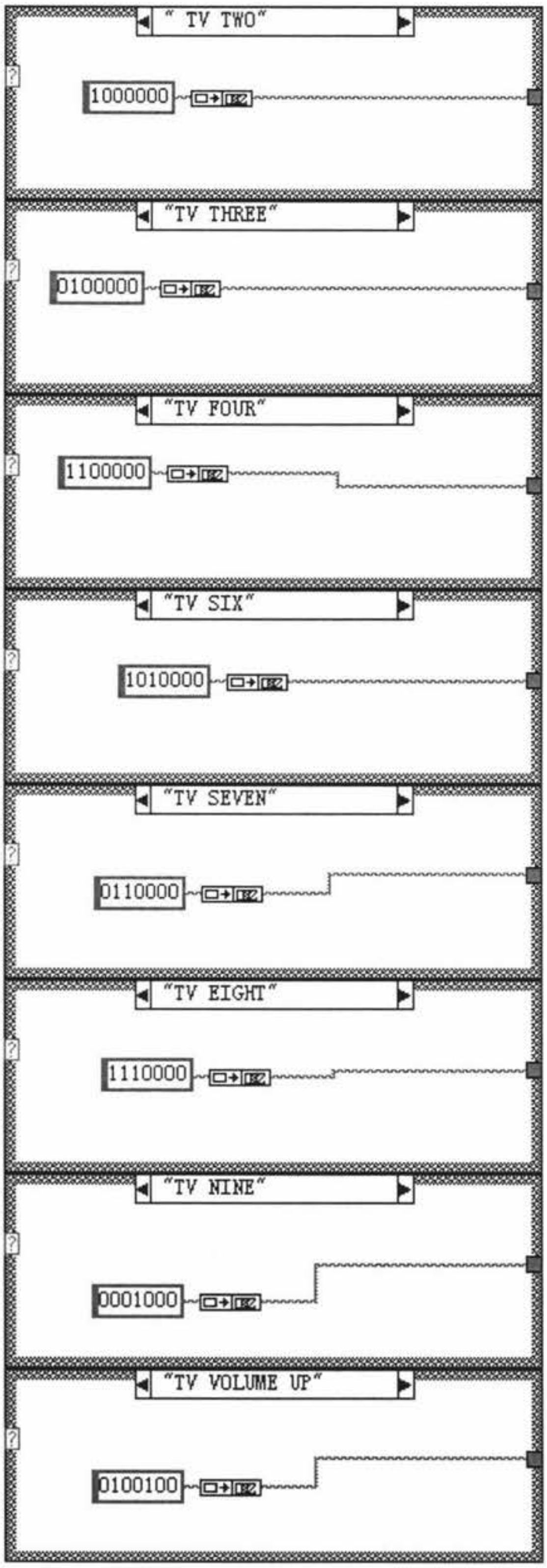

"light three off", Default
11
000000
00000011
command String
103
Light4
"light four off", Default
11
000000
00000100
command String
Light4
"light four on"
11
000000
00000100
command String

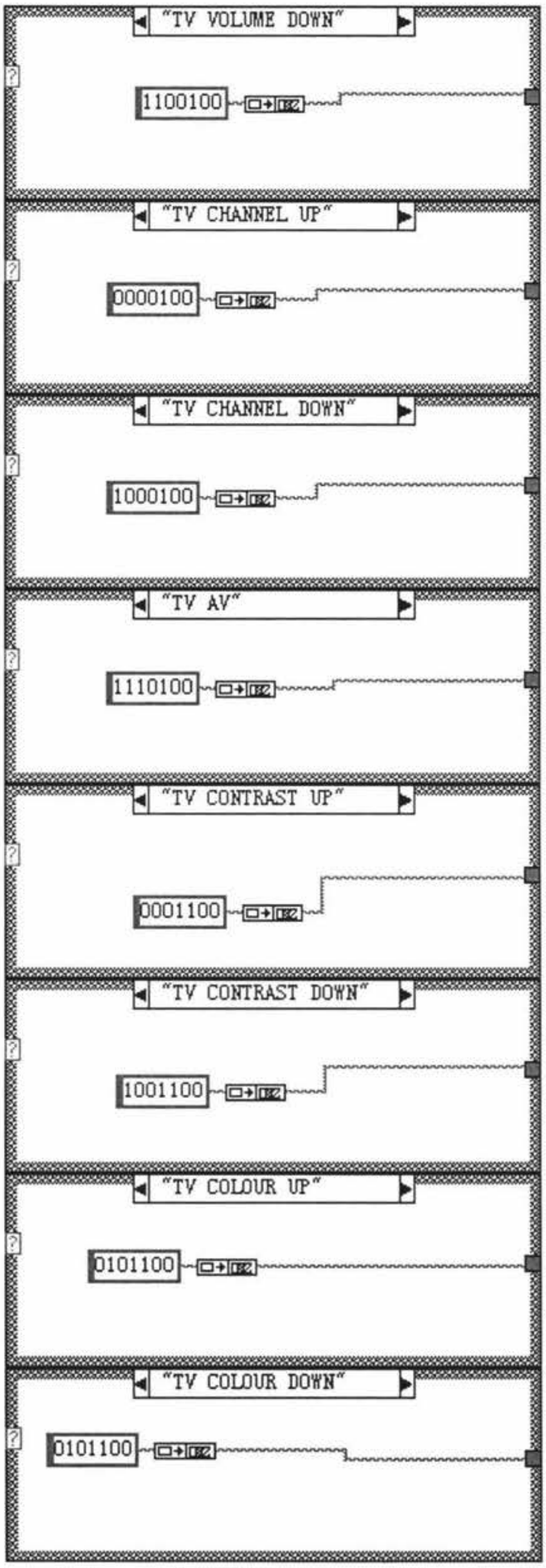

"light four dim"
100000
11
00000100
command String
"0"
1
"TV MUTE"
001101
"TV ONE"
000001
"TV TWO"
000010
"TV THREE"
000011
"TV FOUR"
000100

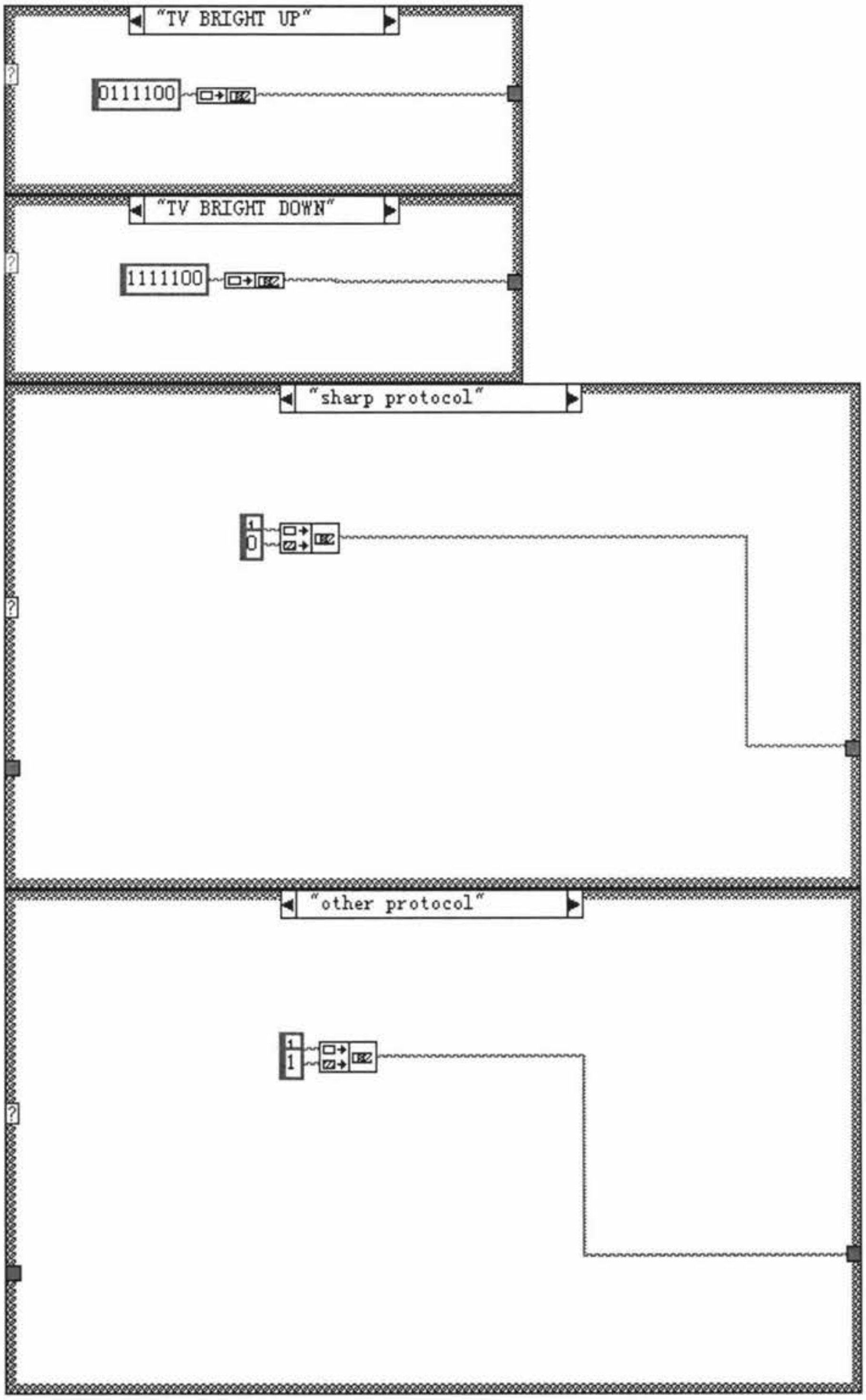

"TV FIVE"
000101
"TV SIX"
000110
"TV SEVEN"
000111
"TV EIGHT"
001000
"TV NINE"
001001
"TV VOLUME UP"
010000
"TV VOLUME DOWN"
010001

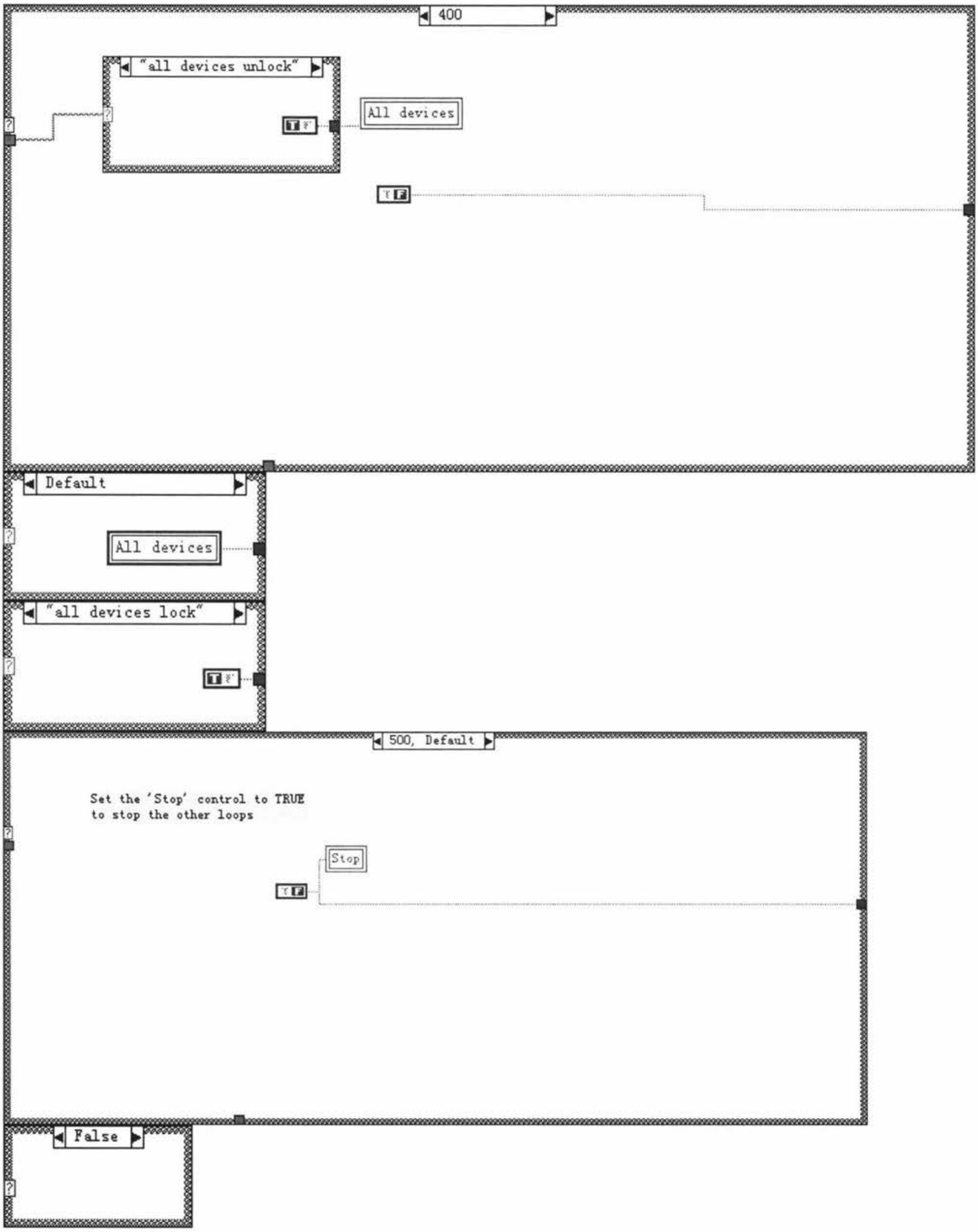

"TV CHANNEL UP"
100000
"TV CHANNEL DOWN"
100001
"TV AV"
1110100
"TV CONTRAST UP"
001100
"TV CONTRAST DOWN"
100110
"TV COLOUR UP"
010100
"TV COLOUR DOWN"
010101


"TV BRIGHT UP"
010010
"TV BRIGHT DOWN"
010011
"sony protocol"
protocol code
1
address code
1
0
00
"TV FIVE"
0010000
"TV POWER", Default
1010100
"TV MUTE"
0010100
"TV ONE"
0000000


"TV TWO"
1000000
"TV THREE"
0100000
"TV FOUR"
1100000
"TV SIX"
1010000
"TV SEVEN"
0110000
"TV EIGHT"
1110000
"TV NINE"
0001000
"TV VOLUME UP"
0100100


"TV VOLUME DOWN"
1100100
"TV CHANNEL UP"
0000100
"TV CHANNEL DOWN"
1000100
"TV AV"
1110100
"TV CONTRAST UP"
0001100
"TV CONTRAST DOWN"
1001100
"TV COLOUR UP"
0101100
"TV COLOUR DOWN"
0101100


"TV BRIGHT UP"
0111100
"TV BRIGHT DOWN"
1111100
"sharp protocol"
1
0
"other protocol"
1
1


400
"all devices unlock"
All devices
Default
All devices
"all devices lock"
500, Default
Set the 'Stop' control to TRUE
to stop the other loops
Stop
False