ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

Control of Home Appliances through Embedded Systems Using Voice Based System

Ramya.J, Shankavi.S, Raveena.K , Saranya.P, Dr.E.Nanda Kumar

Final year, Dept. of EEE, Sri Krishna College of Technology, Coimbatore, Tamil Nadu, India

Associate Professor, Dept. of EEE, Sri Krishna College of Technology, Coimbatore, Tamil Nadu, India

ABSTRACT: In daily life, there will be many situations where it happens to return all the way from office to home to switch on or off the device, which is very difficult. That is the situation where we imagine a system, which enable us operating devices from remote locations. When not only far away from home but also when busy at home, we imagine of operating our appliances through voice either it be AC, heater, coffee pot, microwave, lamp or any other appliance. This is no more imagination now. The digital processing capabilities of microcontrollers (MCUs) allow penetration of voice control to embedded systems. Practically any new device containing a microcontroller Unit (MCU) could be controlled by voice. This paper emphasizes on controlling of lights and heat Using DSP56F805 with Integrated Simple Voice Control. Two basic algorithms that can be used for speech recognition are DTW(Dynamic Time Warping) and HMM(Hybrid Markov Models).This system uses DTW algorithm(speaker dependent) which has fewer requirements on hardware and less speech data in the "training" phase. Current system for lights and heating control can be generalized for controlling any device with a limited command set. Speaker dependence is the main restriction but “new” algorithms based on hidden Markov models (HMM) adapted to the MCU environment are going to change this situation soon as DSP56F805 has 32.5K Flash program memory so the recognition algorithm can be easily upgraded.

KEYWORDS: MCU, DTW, DSP56F805, HMM.

I. INTRODUCTION

Incorporation of embedded voice control capability is described for light and heat control .This solution can be re-purposed to other types of embedded systems that require voice control with a limited command set. However, one noticeable limitation of voice control is audio/T.V systems as it is difficult to separate the speech of the user (commands) from the other audio signals. As the digital home becomes a reality, control mechanisms for everyday appliances are growing beyond the traditional touch screens, keypads and remote control units to include voice command. Voice recognition systems can now be added easily and inexpensively to practically any new home appliance.

Digital signal processing capabilities integrated with microcontroller (MCU) functionality enable voice control capability to be added to Embedded Systems (computers that are generally a combination of hardware and software designed to perform a dedicated function and are often part of a larger system or product). MCUs are today used in thousands of electronic products and systems in which many decisions or calculations are required. Such systems monitor and control everything from spacecraft to robots and factory equipment, home appliances, security systems, automobiles, VCRs and TVs, cellular telephones and personal digital assistants -- virtually any electronic device used in our everyday lives. A new breed of digital signal processor (DSP)-controllers have emerged that provide the performance required for real-time speech processing as well as the traditional MCU features needed for control functions.

II. CAPABILITY OF DSP56F805

The pivotal feature for speech processing is the ability to parameterize speech signals in real time. The key and most time consuming operation in speech Parameterization is Fast Fourier Transformation (FFT). If we use 8 kHz sampling frequency, 12-bit quantification (storing in 16-bit), 16 ms segmentation with 50 percent overlapping, we will

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

need to process 125 frames per second. If we use common speech parameterization, e.g. cepstral coefficients, we will process one FFT and one inverse FFT on a 128-point frame (corresponding to16ms/8kHz). This computation consumes about 50,000 clock cycles on DSP56F8XX using the SDK signal processing library. If we keep 50 percent in reserve for additional computations we will need 50,000 * 1, 5 * 125 (frames) = 9,375,000 clock cycles per second, meaning 12 percent utilization for 80MHz clock frequency. The Remaining reserve can be utilized later using more sophisticated speech processing algorithms. Memory requirements create command set size restrictions. If we use speech parameterization from the above mentioned example with eight coefficients per frame, we will need 1,000 data words per second of speech. Assuming a minimal command set of about 30 words of speech; we will need 30,000 data words. DSP568XX can directly address two, 65,536 data words.

III. LIGHT AND HEAT CONTROL USING DSP56F805 WITH SIMPLE INTEGRATED VOICE CONTROL

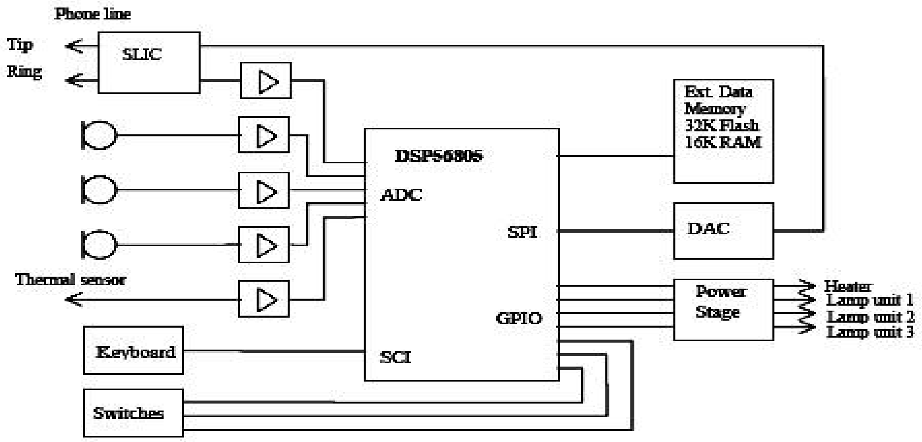

Figure 1. Light and Heating Control using DSP56F805 with Integrated Simple Voice Control

The command set is designed with respect to speaker dependence of the used algorithm. Assuming the proposed system will be used by four people, the command set could be the following: “light”, “dark”, “heat”, ”cold” (recorded four times individually for each speaker). “Temperature”, “Heat”, “0-9” are recorded by single authorized person. The speaker dependence represents an advantage when controlling devices by phone.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

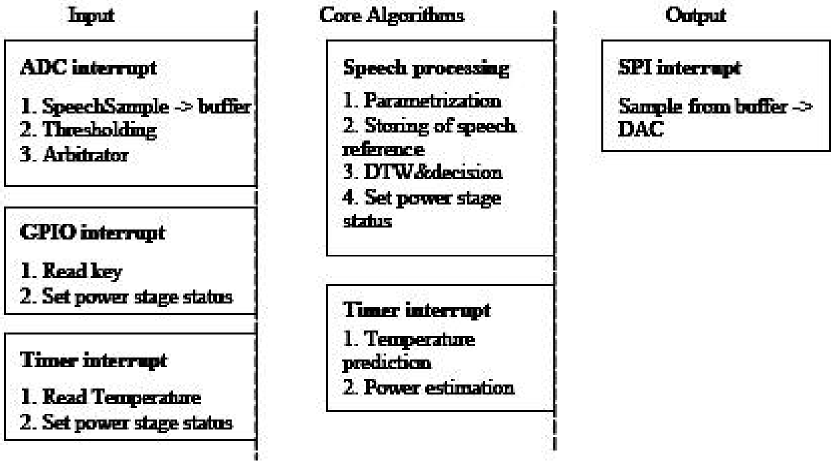

Figure 2. Software Model

As shown in the Figure 1 and 2.All controlling functions of proposed system are accessible both by voice input and manual switch, or keypad regarding to the possibility of noisy environments. Each microphone has a corresponding switch or button and a corresponding lamp unit. The arbitration process, choosing the controlled device, is solved by software. DSP continuously samples all analog to digital converter (ADC) inputs and the affiliated software process adjudicates if there is a speech on input and assigns actual device. The ADC on the DSP56F805 has two channels, each multiplexed to four pins (eight sources of analog signal can be connected in all).

One ADC pin can be optionally connected to phone line via subscriber line interface circuit (SLIC) and one pin is connected to a temperature sensor. The remaining six pins can be connected to microphones. The ADC resolution is 12-bit, and maximal sampling frequency is 800 kHz, allowing us to perform time multiplexed sampling of all eight ADC channels. To minimize memory requirements the sampling frequency, 8 kHz is recommended for all speech channels. The recognition algorithm is based on dynamical time warping (**DTW**) to harmonize requirements on memory and on recognition scores. Cepstrum and linear predictive coefficients (LPC) are recommended for speech parameterization but any other one can be imported from other recognition system as C source code. Heating can be optionally controlled by phone. The recognition process will be the same but we recommend the additional implementation of the answering function– if the speech reference is recognized, the DSP sends it back as an audio signal to confirm validity of recognition. DSP is connected to phone line via SLIC.Internal ADC is used at input side and external digital to analog converter (DAC) and is used at the output**.** DAC and DSP are connected via SPI. Modem functions are provided by software. All the software should be programmed in C language. Metrowerks C compiler and integrated debugging environment is the only software tool supported at present. Proposed layering of the software and interrupt usage is displayed on Figure 2. Individual software tasks communicate using semaphores (global status variables). Modem functions and basic DSP algorithms are components of the Software Development Kit (SDK) provided by Motorola. The functionality of the proposed application design can be demonstrated on evaluation module DSP56805EVM. This board contains all required hardware except input analog parts (microphones, preamplifiers, thermal sensors) and SLIC.The DSP56F805 evaluation module is fully supported by Metrowerks IDE.[1-5]

IV. SPEAKER DEPENDENT AND INDEPENDENT SYSTEMS

There are two types of system named ‘Speaker dependent’ and ‘Speaker independent’. Latter requires more memory, needs more speech data in the "training" phase and also more processor performance. Speaker dependency is sometimes required to improve the security of the system.[6-11]

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

The terms **speaker dependent** and **speaker independent** are given as:

4-1. Speaker dependent: Each word to be recognized is only compared with words spoken by the same speaker.
4-2. Speaker independent:
This has been given in two meanings:

1. Each word to be recognized is compared with all the words available; i.e. words spoken both by the speaker of the test word and the other speakers.
2. Each word to be recognized is compared with all the words spoken by different speakers; i.e. words spoken by the speaker of the test word are not used in the recognition process. This implies that the speaker is totally unknown to the system.[12]

Some of the recognition algorithms are DTW, HMM etc.DSP56F805 with integrated simple voice control uses DTW (Dynamic Time warping) algorithm. Let us now look in detail the DTW algorithm.[13]

V. RECOGNITION ALGORITHMS

Two basic choices for recognition algorithms include Dynamical Time Warping (DTW) and Hidden Markov Models (HMM). DTW has fewer requirements on hardware. HMM is more complex, enables better recognition scores, but needs more speech data in the "training" phase. DTW is used in the design described below which is used in this system.

5-1. Dynamic Time Warping algorithm (DTW)

Dynamic Time Warping (DTW) algorithm is normally used to compare and classify similar patterns by means of a measure of similarity. This approach is especially suitable for those errors related with time distortions .The dynamic time warping (DTW) algorithm provides a procedure to align optimally in time the test and reference patterns and to give the average distance associated with the optimal warping path. Sakoe and Chiba proposed a DTW algorithm for spoken word recognition and showed experimentally its performance.[14]

In this type of speech recognition technique the test data is converted to templates. The recognition process then consists of matching the incoming speech with stored templates. The template with the lowest distance measure from the input pattern is the recognized word. The best match (lowest distance measure) is based upon dynamic programming. This is called a Dynamic Time Warping (DTW) word recognizer.

In order to understand DTW, two concepts need to be dealt with,

- *Features:* the information in each signal has to be represented in some manner.
- *Distances:* some form of metric has been used in order to obtain a match path. There are two types:
- *Local:* a computational difference between a feature of one signal and a feature of the other.
- *Global:* the overall computational difference between an entire signal and another signal of possibly different length

Since the feature vectors could possibly have multiple elements, a means of calculating the local distance is required. The distance measure between two feature vectors is calculated using the *Euclidean* distance metric. Therefore the local distance between feature vector x of signal 1 and feature vector y of signal 2 is given by, the equation

$$d(x,y) = \sqrt{\sum_{i}(x_i - y_i)^2}$$

[15]One of the earliest approaches to isolated word speech recognition was to store a prototypical version of each word (called a template) in the vocabulary and compare incoming speech with each word, taking the closest match. This presents two problems: what form do the templates take and how are they compared to incoming signals.[16]

The simplest form for a template is a sequence of feature vectors -- that is the same form as the incoming speech. We will assume this kind of template for the remainder of this discussion. The template is a single utterance of the word selected to be typical by some process; for example, by choosing the template which best matches a cohort of training

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

utterances. Comparing the template with incoming speech might be achieved via a pair wise comparison of the feature vectors in each. The total distance between the sequences would be the sum or the mean of the individual distances between feature vectors. The problem with this approach is that if constant window spacing is used, the lengths of the input and stored sequences are unlikely to be the same. Moreover, within a word, there will be variation in the length of individual phonemes: *Cassidy* might be uttered with a long /A/ and short final /i/ or with a short /A/ and long /i/. The matching process needs to compensate for length differences and take account of the non-linear nature of the length differences within the words.
The Dynamic Time Warping algorithm achieves this goal; it finds an optimal match between two sequences of feature vectors which allows for stretched and compressed sections of the sequence.

5-2. The DTW Grid

We can arrange the two sequences of observations on the sides of a grid (figure 3) with the unknown sequence on the bottom (six observations in the example) and the stored template up the left hand side (eight observations). Both sequences start on the bottom left of the grid. Inside each cell we can place a distance measure comparing the corresponding elements of the two sequences.

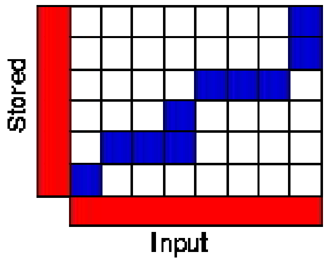

Figure 3. An example DTW grid

To find the best match between these two sequences we can find a path through the grid which minimizes the total distance between them. The path shown in blue in (Figure 3) gives an example. Here, the first and second elements of each sequence match together while the third element of the input also matches best against the second element of the stored pattern. This corresponds to a section of the stored pattern being stretched in the input. Similarly, the fourth element of the input matches both the second and third elements of the stored sequence: here a section of the stored sequence has been compressed in the input sequence. Once an overall best path has been found the total distance between the two sequences can be calculated for this stored template.

The procedure for computing this overall distance measure is to find all possible routes through the grid and for each one of these compute the overall distance. The overall distance is given as the minimum of the sum of the distances between individual elements on the path divided by the sum of the warping function. The division is to make paths of different lengths comparable.

It should be apparent that for any reasonably sized sequences, the number of possible paths through the grid will be very large. In addition, many of the distance measures could be avoided since the first element of the input is unlikely to match the last element of the template for example. The DTW algorithm is designed to exploit some observations about the likely solution to make the comparison between sequences more efficient.

5-3. Optimizations

The major optimizations to the DTW algorithm arise from observations on the nature of good paths through the grid. These are outlined and given as:

- ***Monotonic condition*:** the path will not turn back on itself, both the i and j indexes either stay the same or increase, they never decrease.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 3, March 2017

- ***Continuity condition***: The path advances one step at a time. Both i and j can only increase by 1 on each step along the path.
- ***Boundary condition***: the path starts at the bottom left and ends at the top right.
- ***Adjustment window condition***:a good path is unlikely to wander very far from the diagonal. The distance that the path is allowed to wander is the window length *r*.
- ***Slope constraint condition***: The path should not be too steep or too shallow. This prevents very short sequences matching very long ones. The condition is expressed as a ratio n/m where m is the number of steps in the x direction and m is the number in the y direction. After m steps in x you must make a step in y and vice versa.

By applying these observations we can restrict the moves that can be made from any point in the path and so restrict the number of paths that need to be considered. For example, with a slope constraint of P=1, if a path has already moved one square up it must next move either diagonally or to the right.

The power of the DTW algorithm goes beyond these observations though. Instead of finding all possible routes through the grid which satisfy these constraints, the DTW algorithm works by keeping track of the cost of the best path to each point in the grid. During the match process we have no idea which path is the lowest cost path**;** but this can be traced back when we reach the end point.

VI. CONCLUSION

Advancements in the technology have enabled appliances to understand and respond to human voice. As appliances can be controlled through voice they can also be controlled through telephone even from remote locations which is very much useful in some situations. Controlling home appliances through voice also facilitates old and physically handicapped persons to easily operate the devices without any physical strain.

Current system for lights and heating control can be generalized for controlling any device with a limited command set. Speaker dependence is the main restriction but “new” algorithms based on hidden Markov models (HMM) adapted to the MCU environment are going to change this situation soon as DSP56F805 has 32.5K Flash program memory so the recognition algorithm can be easily upgraded.

REFERENCES

1. J. Wang, Z. Zhang, B. Li, S. Lee, and R. Sherratt, "An enhanced fall detection system for elderly person monitoring using consumer home networks," IEEE Trans. Consumer Electron., vol. 60, no. 1, pp. 23-29, Feb. 2014.
2. H. Y. Tung, K. F. Tsang, H. C. Tung, K. T. Chui, and H. R. Chi, "The design of dual radio ZigBee homecare gateway for remote patient monitoring," IEEE Trans. Consumer Electron., vol. 59, no. 4, pp. 756- 764, Nov. 2013.
3. G. Qiao, G. Song, Y. Wang, J. Zhang, and W. Wang, "Autonomous network repairing of a home security system using modular selfreconfigurable robots," IEEE Trans. Consumer Electron., vol. 59, no. 3, pp. 562-570, Aug. 2013.
4. S. Huynh, D. Parry, A. C. M. Fong, and J. Tang, "Novel RFID and ontology based home localization system for misplaced objects," IEEE Trans. Consumer Electron., vol. 60, no. 3, pp. 402-410, Aug. 2014.
5. D. He, N. Kumar, and J. H. Lee, "Secure pseudonym-based near field communication protocol for the consumer internet of things," IEEE Trans. Consumer Electron., vol. 61, no. 1, pp. 56-62, Feb. 2015.
6. F. Erden and A. Cetin, "Hand gesture based remote control system using infrared sensors and a camera," IEEE Trans. Consumer Electron., vol. 60, no. 4, pp. 675-680, Nov. 2014.
7. Y. Tian, W. Wang, X. Gong, X. Que, and J. Ma, "An enhanced personal photo recommendation system by fusing contextual and textual features on mobile device," IEEE Trans. Consumer Electron., vol. 59, no. 1, pp. 220-228, Feb. 2013.
8. D. Kim, S. Ahn, S. Park, and M. Whang, "Interactive emotional lighting system using physiological signals," IEEE Trans. Consumer Electron., vol. 59, no. 4, pp. 765-771, Nov. 2013.
9. J. Jo, S. Lee, and J. Kim, "Software-defined home networking devices for multi-home visual sharing," IEEE Trans. Consumer Electron., vol. 60, no. 3, pp. 534-539, Aug. 2014.
10. D. S. Kim, S. Y. Son, and J. Lee, "Developments of the in-home display systems for residential energy monitoring," IEEE Trans. Consumer Electron., vol. 59, no. 3, pp. 492-498, Aug. 2013.
11. L. C. P. Costa, N. S. Almeida, A. G. D. Correa, R. D. Lopes, and M. K. Zuffo, "Accessible display design to control home area networks," IEEE Trans. Consumer Electron., vol. 59, no. 2, pp. 422-427, May 2013
12. S Baskar, S Pavithra, T Vanitha,” Optimized placement and routing algorithm for ISCAS-85 circuit”, Electronics and Communication Systems (ICECS), 2015 2nd International Conference on 2015/2/26, IEEE, 958-96
13. S. Baskar,”ERROR DETECTION AND CORRECTION ENHANCED DECODING OF DIFFERENCE SET CODES FOR MEMORY APPLICATION,” International Journal of Advanced Research in Computer and Communication Engineering, IJARCCE, Issue 10, 2012, 816-820.
14. S.Baskar, “Error recognition and correction enhanced decoding of hybrid codes for memory application” at Devices, Circuits and Systems (ICDCS), 2014 2nd IEEE Conference 1-6.
15. S.Baskar, “RELIABILITY ORIENTED PLACEMENT AND ROUTING ANALYSIS IN DESIGN OF LOW POWER MULTIPLIER FOR WIRELESS SENSOR NETWORKS” at International Journal of Applied Engineering Research, Volume 10,issues 44, 31384-31390.
16. S Baskar and V.R. Sarma Dhulipala, “Comparative Analysis on Fault Tolerant Techniques for Memory Cells in Wireless Sensor Devices”, Asian Journal of Research in Social Sciences and Humanities 6 (cs1), 519-528