# A Detection System of Infant's Cry, Using Fuzzy Classification: from Theory to Practice

Mohammad Kia[1], Shabnam Kia[2], Kamaledin Mousavi Mashhadi[1]
[1] Department of Electrical Engineering
Iran University of Science and Technology
Tehran, Iran
[2] Department of Engineering
Islamic Azad University, Science and Research Branch
Tehran, Iran
m_kia@hotmail.com, kia_shabnam@yahoo.com, sk_mousavi@iust.ac.ir

***ABSTRACT:*** *Nowadays, it is inconvenient for housekeeper parents to constantly watch over their newborn baby while doing their work or chores. This paper proposes a simple voice recognition system which can be applied practically for designing a device with capability to detect a baby's cry and informing the parents automatically.*

*There are a lot of similar projects and experiments which have been performed recently, but most of them are about recognizing and classifying different types of crying (like for hunger, etc.) and have used complex methods of implementation such as neural network. But in this paper our aim is to merely detect infant's crying, and our solution is to use a fuzzy classifier which is easy to implement and fast to execute.*

*The overall algorithm is to evaluate the resemblance of the infant's voice signal with the data stored in a database, which is already prepared by recording some cry and laughter samples, using an automatic fuzzy classifier system which can lead to detection of cry or laughter.*

*This algorithm can serve as a reliable foundation on which the future creation of a portable real-time, automatic voice detection device can be based. It is a pretty formidable task to implement complex algorithms, such as neural networks, on common available microcontrollers, however we proposed a much simpler algorithm which enables us to develop a real-time and low cost device.*

*To evaluate the algorithm, we have created a database of sample cry and laughter signals and developed a sample Matlab program for carrying out the real-time frequency-domain calculations and a sample visual program in Labview programming environment for interfacing with user.*

**Keywords:** Infant's Crying Detection, Riemann Sum of Signal, Fuzzy Classification, Alarm Calls

**Received:** 23 June 2012, Revised 3 August 2012, Accepted 12 August 2012

## 1. Introduction

Nowadays it is a difficult for housekeeper mothers to take care of their newborn baby while doing house chores. There are some monitoring devices (including a wireless camera and a LCD) to solve this problem, but their expensive price and the fact that the mother should always watch the baby are the negative points of such solution. Our idea is to present a simple algorithm applicable to create a low cost device with the ability to detect infant's crying and call the parents.

Other researches in this area have been mostly about different methods of voice recognition and categorization such as: analysis of acoustic features of infant cry for classification purposes [1], emotion detection from infant facial expressions and cries [2], infant cry classification to normal and pathological cry [3], detection of asphyxia from infant cry [4], etc. [5, 6, 7, 8] in which they have rather used complex features and classifiers like neural networks. Our approach is to design a detection system (not recognition) which uses fuzzy classifier that is much simpler and can lead to create an easy structure low cost device. A similar project has been performed which uses a Cry Unit as primary element for cry analysis and also an evolutionary-neural approach [9].

In this project a program is implemented to detect an infant's crying or laughing. It can detect baby's cry while ignoring other sounds like human speech, music, etc. The project consists of several phases; among them is development of a unique digital signal processing algorithm, software development, quality assurance tests and practical implementation.

As mentioned before, the purpose of this project is to detect the sound of crying infants and make an alarm call to the parents; so the two objectives of this project will be as follows:

1. Detecting the crying or laughing voice of the infant.

2. Making a phone call to default number.

For the first purpose, we recorded several signals of infant laughing and crying at first, and then we designed a classification algorithm based on the recorded signals. At last a telephone system was implemented to make a phone call to the parents.

Being familiar with the overall project, we will describe it in detailed sections.

It should be noted that in the computer based application, parts of recording and playback sound (crying and laughing), and the alarm system including modem connection are implemented by Labview software; whereas the parts of signal features and processing, fuzzy theory and its rules are implemented by Matlab application.

Block diagram of the overall project is as follows in Figure 1.

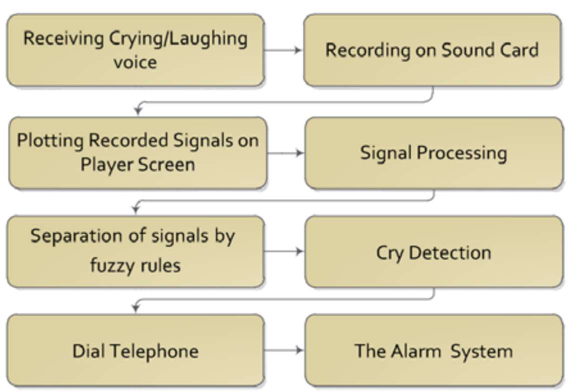



Figure 1. Block diagram of the overall project

## 2. Procedure and System Structure

### 2.1 Signal Preparation

In this study several samples of different infants' laughter and cry have been collected in a database. The instances have been taken from various infants. In fact, by considering just one particular infant, it may only response in that infant audio frequency system. So we can claim that numerous examples of other will acceptably answer. This database is used to train the classifier. The signals are digitally recorded and saved. Due to higher quality, recording has been performed in Mono 16bit mode of recorder section in Labview; and the signals are recorded entirely in circumstances of noise-free places. These signals form the training database. In the main process of operating, the program is running and signals are being recorded; the procedure is running in real-time mode and until a crying voice is detected the program does nothing.

### 2.2 Signal Processing

Now that the sounds were recorded and saved, we turn to the second part, signal processing, which is generally done by MATLAB software [10].

#### 2.2.1 Investigation about Features of Crying and Laughing Signals

To train the classifier a feature is needed groups through which the classifier can use to divide signals into two groups, laughing and crying [11]. In this section we discuss about this criterion.

Actions taken in this area are listed below:

a) Taking fast Fourier transform of signals (F.F.T.).

b) Dividing frequency domain.

At first, because the infant's crying and laughing maximum signal frequency with noise recorded does not exceed 20 KHz, it is sufficient to analyze it between the low frequency range of X1, considered 0 Hz, and high frequency range of X2 considered 20 KHz. Then this frequency range is divided into eight sections with the same size called levels. Thus each level of frequency is 2500Hz; and then the fast Fourier transform of the signals is taken at this range.

c) Calculating Riemann sum of F.F.T. of the signal in different levels.

Since Fast Fourier Transform has been taken, the frequency domain signal is pointy [12]. For example there are 2500 points in the second level (which shows 2500 different frequencies with unit differences, such as2500, 2501, 2502 ... 4999), each of which has a weight or coefficient. Here we determine the signal average coefficient in each level by adding the F.F.T coefficients of these 2500 points and dividing it by dt = 2500.

d) Obtaining criteria in order to separate signals from each other, crying and laughing.

The main thing that draws our attention to the shapes in Figure 2, is the markedly differs of the Riemann sum of the signal, laughing and crying. As it can be seen that the Riemann sum of signal laugh only at low frequencies is considerable and at high frequencies value is not; however, the Riemann sum of signal cry at low frequencies and at high frequencies has a significant amount. So here we fell upon the criteria to distinguish signal cry from signal laughter.

#### 2.2.2 Fuzzy Theory

We have proposed a criterion to separate the signals into two groups. Now there is a need for a system to learn how to recognize crying. Indeed we need a classifier. Here we used fuzzy rules to design the classifier.

Since the advent of modern science until the late nineteenth century, the uncertainty as a negative phenomenon that should be avoided was considered. Gradually this manner changed at the beginning of the twentieth century with the advent of probable mechanics.

Up to insofar, Lotfi Asgarizadeh developed the new theory of uncertainty in 1965 which was distinct from the probability theory. Asgarizadeh had a passion to solve complex systems through simple modeling. His various scientific and practical experiences indicate this fact that traditional methods of mathematical were not proper for this way of modeling.

His famous essay introduced the concept of fuzzy sets with non-radical and non-clear boundaries. In other words, in a classic set with radical boundaries each element is either a member of the set, or not inside the set [13].

Despite the radical boundaries of classical sets, fuzzy sets of boundaries are not very radical and clear. In other words, the condition of membership or non-membership of an element in a set does not depend on the full membership or full non-membership in the set. An element may have more degree of membership in a set, or less degree of membership than the other element.

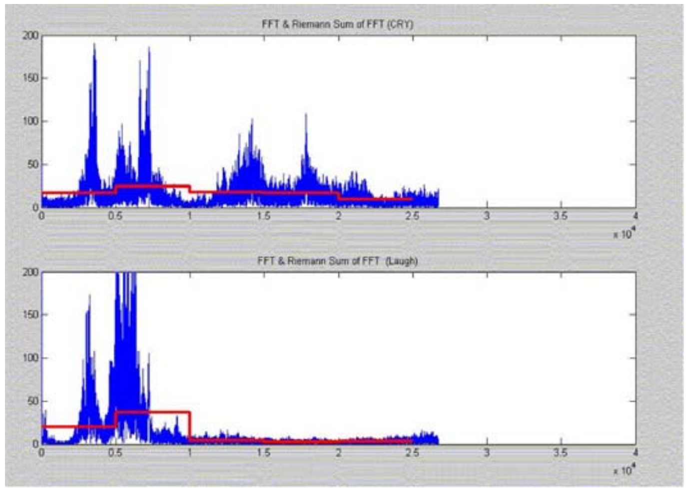


Figure 2. F.F.T. and Riemann sum of F.F.T. for two crying and laughing samples

One of the incentives underlying the introduction of fuzzy sets is imprecise and vague concepts. Because an individual member or element in a fuzzy set may be associated with uncertainty, membership of the elements is according to their degree of membership. Each set of fuzzy is identifiable with its own uniquely membership function and each member within a set, has a degree of membership. The membership degree is generally between zero and one [14].

**2.2.3 System Training**

Seeking a criterion for distinguishing the signal cry of laughter, we realize that the difference is in the sum of Riemann of the signal. It is time now we can draw a pattern according to this difference that would help separation of any other cry and laugh signals. So the first we go to draw the pattern of cry and laugh which is called membership functions in terms of fuzzy. Here we allocate Gaussian functions to the cry and laugh membership functions; because they are easy to implement and have many applications in fuzzy theory. To draw Gaussian functions we need to know two parameters; the first parameter is the center of the Gaussian and the other one is its standard deviation.

At first, to draw membership function of the laugh and cry signals in each level, we need a symbolic signal of laugh and cry such that if it is considered as the base signal of cry or laugh, if a membership function is drawn of on it, for any test signal, a cry or laugh, it could be used.

To achieve this goal, of all cry signals have recorded, half are randomly chosen as samples and their Riemann sum is calculated in each level. Then average Riemann sum is calculated for each signal and saved. Same is done for laugh signals. Now to determine membership functions we need the average Riemann sum (exp. for cry signal) to be considered as the center for the Gaussian function with a particular standard deviation. The SD should be a way to take in all the Riemann sums of other signals. For this we calculate the most distant Riemann sum between cry signals and average Riemann sum, and consider it as the SD; which here is around 3.6. And we do same for laugh signals and consider the utmost calculated value as SD which is again about 3.6. In accordance with the empirical studies have been conducted, this SD is good for analysis of infants' crying and laughing signals.

Membership functions are now being drawing to a symbolic model is given in Figure 3. In this section we have three parts of input, output, and fuzzy logic; which inputs are the crying and laughing membership functions. Two test signals are at each level; for example there is a three level seen in the Figure 3.

At each level, mf1 represents membership function to cry and laugh membership function is represented as mf2.

Each of the Output functions below is a probability density function between 0-1. Each level will have an output that is expressed as a similarity percentage of its membership function to the Gaussian function of that identic level.

If a similarity percentage of the membership function is related to the cry for more than 50% the system detects a cry state. And about laugh is the same, if a similarity percentage of the membership function is related to the laugh for more than 50% the

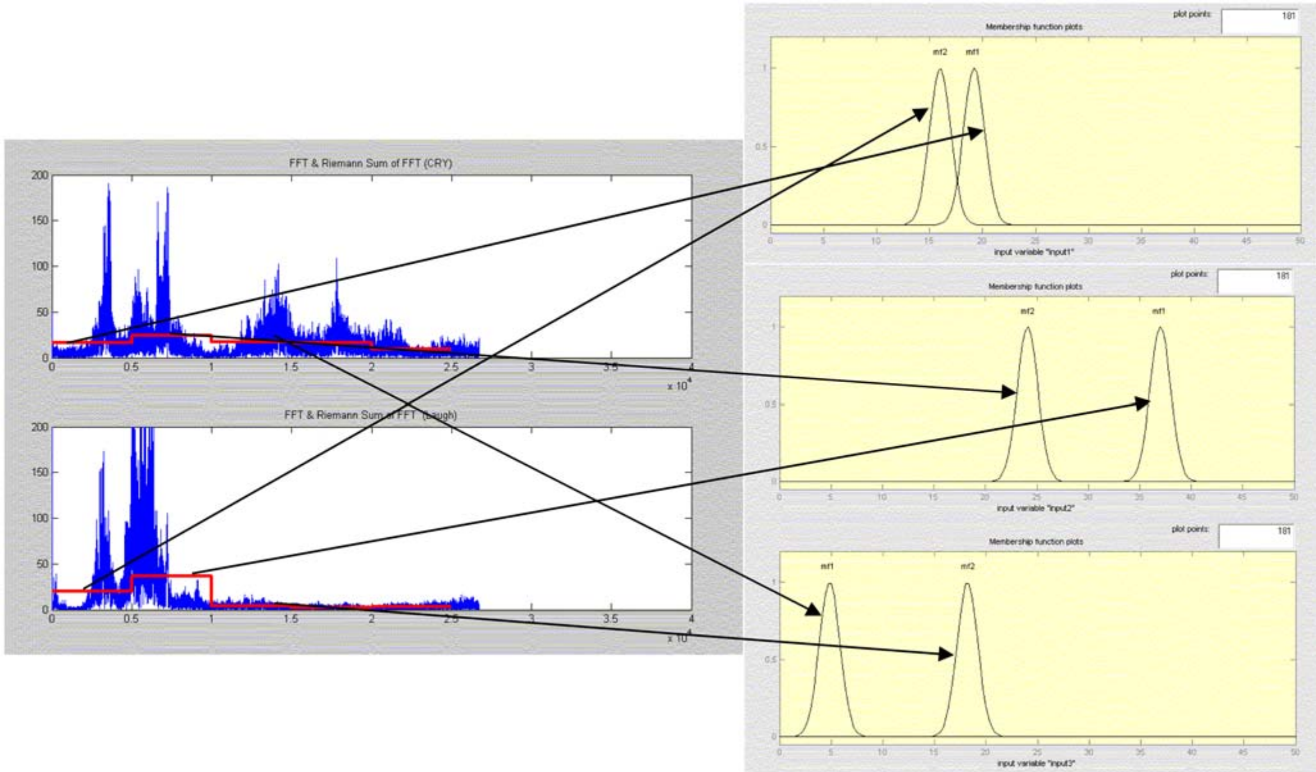



Figure 3. Input membership functions

system detects a laughing state. Figure 4 shows the output corresponds to the cry signal and for the laugh signal is seen in Figure 5.

For any typical signal given to the system, at first its fast Fourier transform is taken, then it is divided into 8 frequency levels and in each level Riemann sum of F.F.T. is calculated. Then on each Riemann sum in each level, a Gaussian function is created for that level which is compared to the cry membership function and the similarity percentage is computed and for more than 50% value of similarity, along with reporting this percentage, the specified alarm is raised.

Following after sections described the fuzzy input and output we shift to the fuzzy rules that will explain it below. The Matlab fuzzy rules in fuzzy space between the “*Or*” and “*And* ”, the option “*And* ” selected [15]. Namely the fuzzy AND law of separation has been used to cry of laughter. This mode of operation of these rules is that if for the all eight input, the option MF1 is selected (means cry), output1 corresponds to the cry will be set in mode MF1 and the output2 related to the laugh, will be none. If all of the eight levels detect cry, by AND law operation of these together, the output will be cry; i.e. the cry output will be true and the false output will be false. So cry state is diagnosed by system.

Because of the used fuzzy AND law, in each eight level, one thing must to be diagnosed, cry or laugh, so we can certainly claim that the output is cry or smile. This feature is the fuzzy AND law.

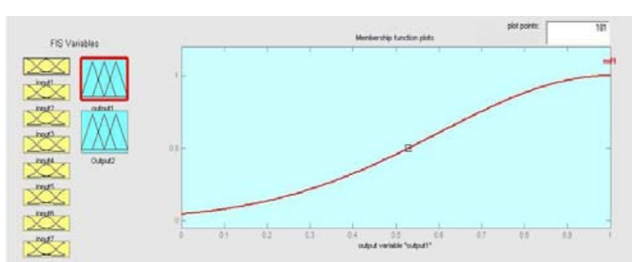



Figure 4. Cry output figure

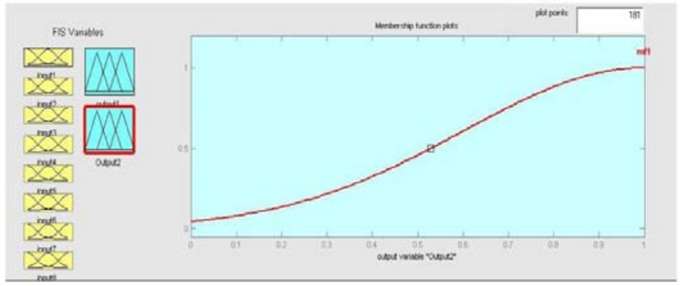


Figure 5. Laugh output figure

**2.3 The Alarm System and Telephone Calls**

Now that we have achieved an algorithm to detect the infant's crying, we reach to the final step which is making a phone call to the parents and alarming them about their child. The program is implemented using a simple dial-up modem.

If the similarity of the test signal to the cry membership function is over 50%, the cry output will be true and multiple beep system goes alarming. If the similarity percentage of the signal to the laugh membership function is up more than 50%, the system will detect laughing and will not beep. Figure 6 shows a sample of a diagnosed cry signal.

This can be any other desired signal inputted into the system, and as a result of the similarity percentage compared to the crying and laughing membership functions, which for example leads to diagnosing cry, the system will automatically beep.

In this project we used a dial-up modem to call the desired number. For using the modem there is need to do some software implementations. Here this preparation is done by Labview application which includes 4 stages [16, 17].

1) Initializing: serial port is initialized.

2) Configuration: configuration of the serial port will be fetched from ROM using "?" command [18, 19].

3) Dialing: tone dialing will begin by sending "*atdt* " command and the number will be sent in a string of ASCII codes. The call connection is established and the desired number rings [18, 19].

4) Redialing: the modem response is checked in a while block and if the phone is not picked up or the number is busy, the operation is repeated.

In the end, this system has the ability that after the above-mentioned laws and programs, if crying is diagnosed, it records the time and date of crying to give to the infant's guardian.

**3. Performance Analisis**

At this part of the essay we have designed the tables by which we discover the percentage of system error. For this reason our data are divided into two categories: training and testing.

This means that we allocate 70 samples of 100 cry signals to train data and 30 to test data. And among 120 laughing signals we allocate 80 to train data and 40 to test data.

Then we consider 250 experimental states from different combinations of 70 cry sample and 80 laugh sample signals. In each state we have 8 levels for crying and 8 levels for laughing. Each number written for a level in a cry signal is considered as the average Riemann sum of the 70 samples in the identic level. And each number written for a level in a laugh signal is considered as the average Riemann sum of the 80 samples in that identic level. This table is for the training data.

Now we create two membership functions according to these averages in each level; one for cry and one for laugh. Then we calculate Gaussian function of the test data in each level and compare it to the average Gaussian function of the train data and find out the similarity percentage.

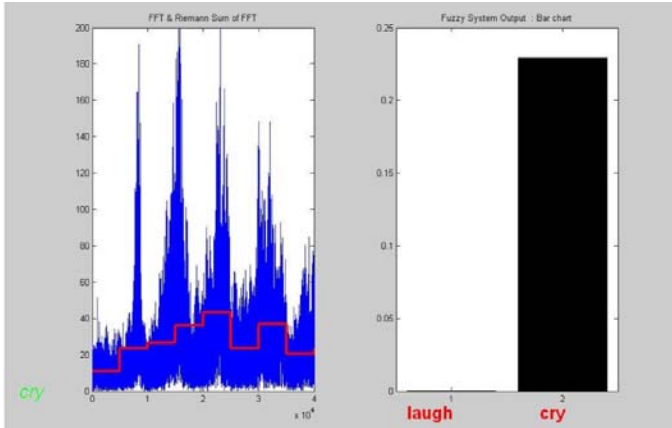



Figure 6. The test signal of the cry

So we find 8 similarity percentages for each signal corresponded to the 8 levels. Then we calculate the average of these 8 values and the result shows the similarity to the cry or laugh. If the similarity percentage is more than 50% to cry, and crying is diagnosed, this value is written in the "Correct" field of another table and its complement "*X* " as the error of diagnosing is written in the "*Error* " field.

X = 1 – similarity percentage to cry table 1 is consisted of 2 samples of various states showing the similarity of the test signals to the train signals.

In each state this process is repeated for 3 cry signals and 2laugh signals. Finally for this 25 states average value of the correct field is calculated.

| | cry1 | cry2 | cry3 | Cry _ Ave | laugh1 | laugh2 | Laugh _ Ave |
|---|---|---|---|---|---|---|---|
| correct | 0.63 | 0.9 | 0.9 | 0.81 | 0.94 | 0.73 | 0.835 |
| error | 0.37 | 0.1 | 0.1 | 0.19 | 0.06 | 0.27 | 0.165 |
| correct | 0.68 | 0.56 | 0.58 | 0.606667 | 0.68 | 0.76 | 0.72 |
| error | 0.32 | 0.44 | 0.42 | 0.393333 | 0.32 | 0.24 | 0.28 |

Table 1. The Similarity of the Test Signals to the Train Signals

This shows the accuracy of the system in diagnosing correct crying. Then average of the error field is calculated and this shows the error of the system in detecting cry. Next the same operation is performed for the laugh signals and the average of the error and correct fields are calculated.

Table 2 shows two sample states which in each state, first row shows average Riemann sum of the cry signals, and the second row is for laugh signals. Right side of the table shows the number of averaged signals.

Eventually by averaging of the above values, percentage of accuracy and error of the system in each case, is determined which is presented in table 3.

By increasing the number of samples and filtering noises the error rate will be reduced.

The error rate may stem from the uncertainty of the fuzzy theory for diagnosing the crying or laughter. Or it can be a result of the uncertainty of the registered SD.

**4. Practical Implementation**

In order to implement the presented method in a practical device, there are some important issues pertaining to electronic circuits and programming algorithms to consider. The overall procedure is that the signal is captured and filtered by an analog circuit, and then a microcontroller takes discrete samples of the signal and executes the detection algorithm. If the signal is detected to be a cry signal, the microcontroller sends a text message or makes a phone call to a default number using a GSM modem; otherwise it will be continuing the sampling procedure.

| | Level1 | Level2 | Level3 | Level4 | Level5 | Level6 | Level7 | Level8 | Number of signals |
|---|---|---|---|---|---|---|---|---|---|
| Cry Average | 18.4525 | 24.32601 | 18.70791 | 15.26803 | 11.00236 | 7.428829 | 10.72809 | 22.40279 | 2, 3, 4, 5, 6, 7, 8 |
| Laugh Average | 9.497366667 | 25.7233 | 12.22437 | 8.7621 | 3.9603 | 4.320467 | 4.209133 | 3.1688 | 2, 3, 4 |
| Cry Average | 18.4525 | 24.32601 | 18.70791 | 15.26803 | 11.00236 | 7.428829 | 10.72809 | 22.40279 | 2, 3, 4, 5, 6, 7, 8 |
| Laugh Average | 9.497366667 | 25.7233 | 12.22437 | 8.7621 | 3.9603 | 4.320467 | 4.209133 | 3.1688 | 2, 3, 4 |

Table 2. Average of the Levels in Two Sample States

| | | | |
|---|---|---|---|
| Accuracy percentage identified for the cry signal | %71.6 | Error percentage for detection of cry signal | %28.5333 |
| Accuracy percentage identified for the signal of laughter | %67.72 | Error percentage for detection of laughter signal | %32.48 |
| Accuracy percentage identified for the system | %69.66 | Error percentage for the detection system | %30.5067 |

Table 3. Accuracy and Error of the System

To implement such a device, it is required to use an appropriate microcontroller with DSP capabilities. In this project we have considered Microchip dsPIC33F family for they have a powerful DSP core and they are not very expensive. Employing this controller and some low-cost electronic elements could lead to construction of a low-cost device, affordable for many families. The only expensive element in this set is the GSM modem whose price could also decline in mass production.

### 4.1 The Input Electronic Circuit

First of all, an electronic circuit is needed in order to capture the infant's voice through a microphone. The first part of this circuit is shown in Figure 7. This circuit primes the voltage signal requisite for the signal processing performed by the microcontroller.

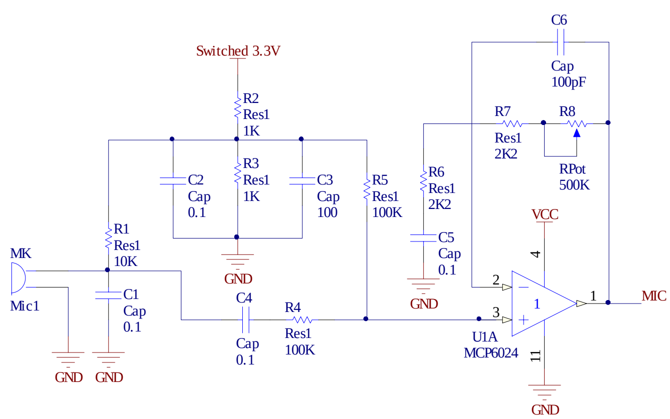



Figure 7. Microphone input pre-amplifier [20]

However, the output signal of this circuit needs to be filtered to ensure the correctness and precision of the FFT calculation. Therefore we could use the circuit presented in the Fig.8 as an anti-aliasing filter.

Also an audio codec can be exerted for a higher-end audio application. The input to the audio codec will be the output of the microphone pre-amplifier. The codec must interact with the application program running on the microcontroller device. Commands from the application program control the codec operating parameters (such as communication protocol, sampling rate, volume control, level control, filter settings, etc.). Command information could be exchanged over the Inter-Integrated Circuit™ (I2C) module on the device. The codec converts the incoming audio signal to a digital signal for the Digital Converter Interface (DCI) module of the microcontroller device.

However we will not discuss the audio codec in this project for the ADC module of the microcontroller device could sufficiently satisfy our need for sampling.

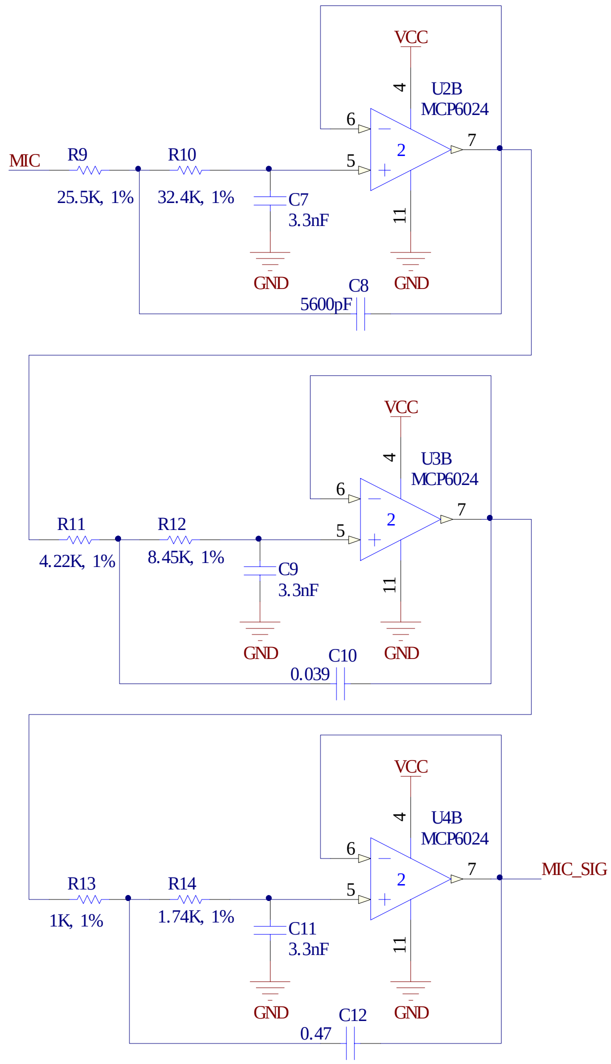



Figure 8. Anti-Aliasing Filter [20]

Second of all, the output of the filter should connect to the analog input channel of the microcontroller. A suitable model of the FJ family for this purpose could be dsPIC33FJ256GP506. The 64pin DSP microcontroller available TQFP package lets us make a compact device. The 12-bit ADC module has up to 500 ksps conversion rate with auto-scanning ability which is perfect for this application.

**4.2 Microcontroller Based Discrete Time Signal Processing**

Now that we have discussed the input hardware requirements, we should concentrate on the discrete time signal processing performed by the microcontroller.

First and foremost, a timer overflow interrupt service routine must be implemented in order to sample the input signal at a rate of 8 ksps. The samples should be stored in an array so as to be prepared for FFT transformation and further classification using fuzzy rules. Because the classifier needs the last 'n' seconds of the input signal, the input signal in the microcontroller RAM must be updated at the rate of 8 KHz. To achieve that, an array must be declared with the length of '$n \times 8000$' and updated with the definite sample rate. The only important point in details here would be the proper usage of pointers to maintain the program performance at its utmost level. In order to update the array, instead of shifting the whole array, a pointer could be used to replace the oldest sample with the recent sample.

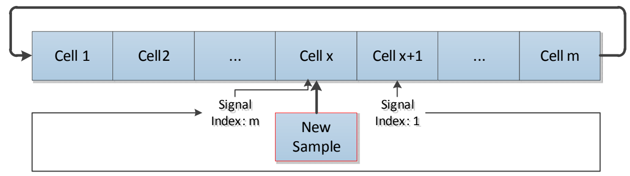


Figure 9. Signal array

Technically an approximate 5 seconds input signal is enough for the classifier to distinguish cry from laughter or silence. For further convenience, we choose to store $2^{16} = 65536$ samples in the input array which indicates around 8 seconds of the infant's voice. An external SRAM will be probably needed for the input signal is an extensive array; a 256 KB SRAM should be abundant.

Now that enough samples of the input signals have been collected, the FFT transformation needs to be taken of the input signal. As mentioned before, dsPIC33F series are appropriate digital signal processors. They could calculate the FFT of a 256 samples signal in around 500 µs [21]. Therefore they are applicable in live DSP applications like this project. DSPIC compilers have apt libraries for digital signal processing which could lead to a convenient and optimized programming. For example MikroElektronika MikroC pro has a library to calculate FFT of a 512 samples signal to which we refer to implement our detection algorithm.

The important issue here is how to calculate the FFT of a 65536 samples input signal using a FFT calculator function which calculates up to 512 samples at a time. Among all theoretical methods to solve this problem, we present a simple but effective solution. For taking the FFT of the main signal, we break it into 128 blocks with a length of 512 samples. Then we calculate the FFT transform for each block and finally we calculate the sum of all 128 FFT results. The final result presents the FFT of the main signal. Although is not free from error, however the error rate is negligible and the overall algorithm is acceptable. Figures 10.a and 10.b depict this allege.

Consecutively, we must calculate the Riemann sum of the spectrum. This could be easily and speedily accomplished. The output of the FFT function represents both sides of the amplitude axis. Therefore we need only half of the FFT data, i.e. 256 items. Now we must calculate the 8 levels. Because the sampling frequency is 8 KHz, the maximum measured frequency by FFT would be 4 KHz.

Since this maximum value is less than 20 KHz, we utilize all of the data for Riemann sum calculation. This summation which could be performed quickly, will prime the 8 features for the fuzzy classifier.

The next step is to implement the fuzzy classifier. At this step we firstly define the membership functions. Just like before, here we have also 8 inputs, each possess 2 Gaussian membership functions. The related Gaussian parameters are provided in table 4.

| *Input* | *MF* 1 | | *MF* 2 | |
|---|---|---|---|---|
| *number* | *c* | *ó* | *c* | *ó* |
| 1 | 19.2 | 1 | 16 | 1 |
| 2 | 37 | 1 | 24.1 | 1 |
| 3 | 4.85 | 1 | 18.2 | 1 |
| 4 | 2.12 | 1 | 16.2 | 1 |
| 5 | 3.75 | 1 | 9.12 | 1 |
| 6 | 4.64 | 1 | 8.19 | 1 |
| 7 | 3.84 | 1 | 11.2 | 1 |
| 8 | 3.7 | 1 | 21.1 | 1 |

Table 4. Input Gaussian membership functions' parameters

If computation of Gaussian function would take so much time by the processor, they could be easily substituted with their

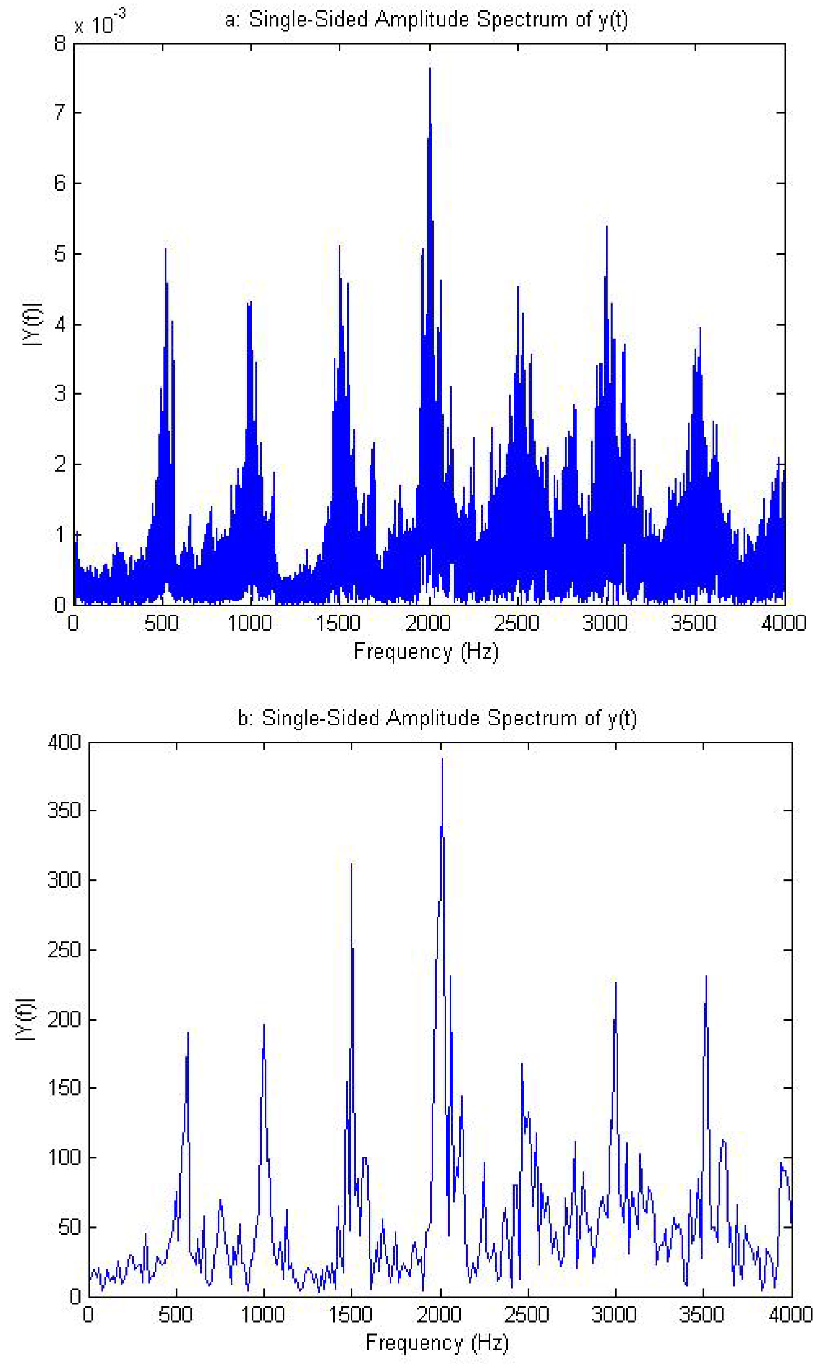



Figure 10a. FFT calculated originally. b: FFT calculated by presented method

triangular equivalent in order to speed up the performance. Table 5 represents the triangular equivalent parameters for input membership functions.

Although these parameters have been tested by authors, they may need to be readjusted for other platforms according to the different parameters of the platform. Nevertheless this algorithm could be pliable with any platform with merely some changes in parameters.

Afterwards, we get to discuss the fuzzy rules. As mentioned before, the connection method is 'AND' with a weight of '1'. The

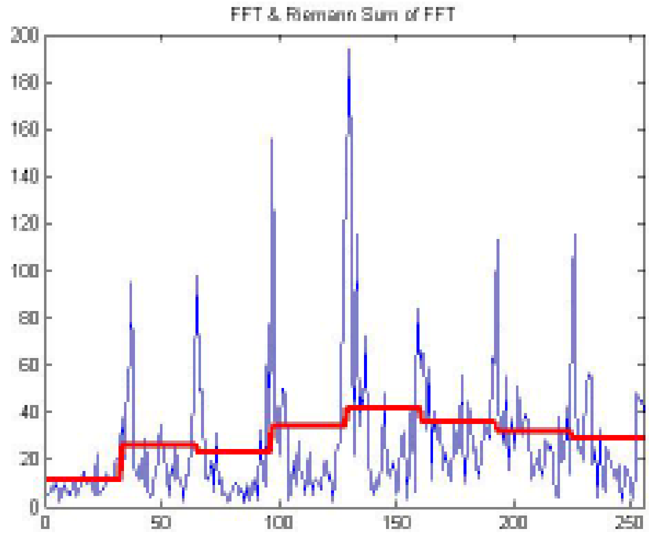


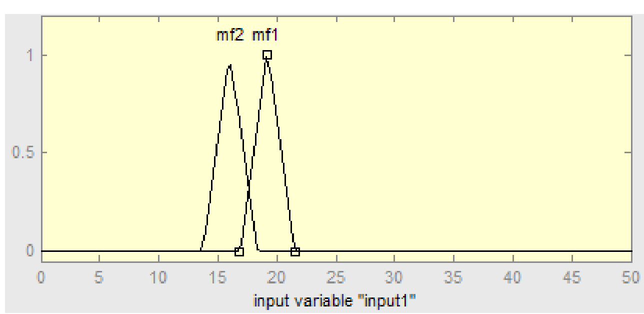

Figure 11. FFT and Riemann sum of FFT

| *Input number* | *MF* 1 | | | *MF* 2 | | |
|---|---|---|---|---|---|---|
| | *a* | *b* | *c* | *a* | *b* | *c* |
| 1 | 16.85 | 19.2 | 21.55 | 13.65 | 16 | 18.35 |
| 2 | 34.65 | 37 | 39.35 | 21.75 | 24.1 | 26.45 |
| 3 | 2.5 | 4.85 | 7.21 | 15.85 | 18.2 | 20.55 |
| 4 | -0.23 | 2.12 | 4.48 | 13.85 | 16.2 | 18.55 |
| 5 | 1.3 | 3.75 | 6.11 | 6.77 | 9.12 | 11.47 |
| 6 | 2.29 | 4.64 | 6 | 5.835 | 8.19 | 10.54 |
| 7 | 1.49 | 3.84 | 6.2 | 8.845 | 11.2 | 13.55 |
| 8 | 1.35 | 3.7 | 6.1 | 18.75 | 21.1 | 23.45 |

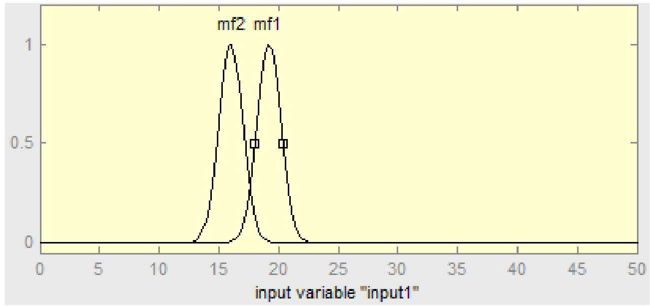
Table 5. Input triangular membership functions' parameters

'*AND*' method is '*product*', the output membership functions are like Figure 4 and Figure 5, and the defuzzification strategy is to compute the smallest (absolute) value of maximum.

Because the output transfer function is monotonically increasing and the connection method is 'AND', we could change fuzzy rules and defuzzification method in a way to decrease calculations without gaining any error. The idea is to calculate the value of MF1 for each input, then compute the average of the results and name it AVE1. Then do the same for MF2 and calculate AVE2. If the AVE2 is greater than AVE1, the signal is a cry signal and consequently the signal is a laugh signal if it is otherwise.




a. Triangular membership functions




b. Gaussian membership functions

Figure 12. Comparison of different types of membership functions

$$AVE\ 1 = \frac{\sum_{i=1}^{8} MF\ 1\ (input\ i)}{8} \tag{1}$$

$$AVE\ 1 = \frac{\sum_{i=1}^{8} MF\ 2\ (input\ i)}{8} \tag{2}$$

$$AVE\ 2 > AVE\ 1 \rightarrow Cry \tag{3}$$

The signal processing procedure should be performed in the main loop of the microcontroller program. If the signal is detected to be a cry signal, then a subroutine should be executed to make a phone call to the parents. This could be easily achieved by embedding a GSM modem, with a RS232 or UART interface such as Siemens tc35i modem which is not very expensive, in the device.

## 5. Conclusion

In this paper we have proposed a simple but effective method for detecting an infant's crying. We implemented our method using MATLAB and Labview softwares and have evaluated the results and error rates which support the properness of the algorithm. At last we have presented the essential information for practical development of a physical device to detect an infant's crying and alerting the parents. The proposed method could be implemented as a low-cost real-time operating detection device.

## References

[1] Analysis of acoustic features of infant cry for classification purposes, (2011). Messaoud, A., Tadj, C. 24th Canadian Conference on Electrical and Computer Engineering (CCECE).

[2] Emotion Detection From Infant Facial Expressions And Cries, (2006) Pal, P., Iyer, A. N., Yantorno, R. E. IEEE International Conference on Acoustics, *Speech and Signal Processing*, ICASSP 2006 Proceedings.

[3] Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks (2003) Garcia, J. O., Reyes Garcia, C. A. (2003). *In*: Proceedings of the International Joint Conference on Neural Networks.

[4] Detection of asphyxia from infant cry using support vector machine and multilayer perceptron integrated with Orthogonal Least Square, (2012). Sahak, R., Mansor, W., Khuan, L. Y., Zabidi, A., Yassin, A. I. M., *IEEE-EMBS International Conference on Biomedical and Health Informatics* (BHI).

[5] A System for the Processing of Infant Cry to Recognize Pathologies in Recently Born Infants with Neural Networks (2004). Reyes-Galaviz, O. F. Reyes-Garcia, C. A. 9th Conference on Speech and Computer St. Petersburg, Russia.

[6] Optimized Support Vector Machine for classifying infant cries with asphyxia using Orthogonal Least Square (2010). Sahak, R., Lee, Y. K., Mansor, W., Yassin, A. I. M., Zabidi, A. *International Conference on Computer Applications and Industrial Electronics* (ICCAIE).

[7] Wang, X., Nagashima, T., Fukuta, K., Okada, Y., Sawai, M. Tanaka, H., Uozumi, T. (2010). Statistical method for classifying cries of baby based on pattern recognition of power spectrum, *International Journal of Biometrics*, 2 (2) 113 – 123.

[8] Messaoud, A., Tadj, C. A cry-based babies identification system; ICISP'10 Proceedings of the 4th International Conference on Image and Signal Processing.

[9] Validation of the Cry Unit as Primary Element for Cry Analysis Using an Evolutionary-Neural Approach (2008).

[10] Reyes-Galaviz, O. F., Cano-Ortiz, S. D., Reyes-Garcia, C. A. ( 2008). Mexican International Conference on Computer Science. ENC '08.

[11] Dowell, A. Sound Recognition for the Hearing-Impaired; Department of Electronics, University of York, York, YO10 5DD, UK

[12] Muruganantham, J., Amarnath, R., Jawahar, K. V., Kalyanasundaram, C. (2003). Methods for Classification of Phonocardiogram, TENCON. Conference on Convergent Technologies for Asia-Pacific Region.

[13] Neil, O'. Advanced Engineering Mathematics; Part Five; Fourier analysis and Boundary Value Problems

[14] Klir, G., Yuan, B., Hall, P. (1995). Fuzzy Sets and Fuzzy Logic: Theory and Applications.

[15] Wang, L. X. a Course in Fuzzy Systems and Control, Prentice-Hall International, Inc. Reyes-Galaviz, O. F., Reyes-Garcia, C. A. A System for the Processing of Infant Cry to Recognize Pathologies in Recently Born Infants with Neural Networks; SPECOM'2004: 9th Conference; Speech and Computer; St. Petersburg, Russia; September 20-22.

[16] Puli, R. (2010). Managing Dialup connection, www.CodeProject.com, 1 Dec.

[17] Naughter, PJ. (2000). CRasMonitor, 1 (41), www.CodeProject.com; 10 Mar.

[18] Conexant; Commands for Host-Processed Modems; Reference manual; Doc. No. 100498E; February 9, (2004).

[19] www.zoltrix.com, Accessed at November 18, (2011).

[20] MPLAB Starter Kit for dsPIC® Digital Signal Controllers User's Guide. www.microchip.com, Accessed on 2 September (2012).

[21] dsPIC® DSC DSP Library. www.microchip.com Accessed on 5 September (2012).