

Adaptive Differential Pulse Code Modulation using speech compression algorithm

J. Venkateswara Rao

(Professor, Department of Electronics and Communication Engineering, Vignan Institute of Technology and Science, Hyderabad, Telangana, India)

Abstract: The aim of this project, was to compress audio data in such a way so that the quality was preserved but was given at a lower bit rate, utilizing a PIC32 microcontroller. The speech compression algorithm used is known as adaptive differential pulse code modulation or ADPCM. The ADPCM algorithm can be broken down into two major components, the encoding process and the decoding process. In an effort to give audio capabilities to a microcontroller, a C implementation of a simplified ADPCM algorithm was developed and programmed onto the PIC32. During the testing phase, the C implementation for the PIC32 was compared to a working MATLAB implementation of the same algorithm, in order to confirm the numerical data was the same throughout the compression process. Once both the encoding and decoding processes produced identical outputs in both C and MATLAB, the code was put onto the microcontroller. The result was that audio compression was successful; the spectral content of the raw speech data and compressed speech data are the same and the entire process only used 1/5 of the CPU of the PIC32.

Key Word: Digital Signal Processing, Adaptive Pulse Code Modulation, Microcontroller.

Date of Submission: 12-06-2020

Date of Acceptance: 29-06-2020

I. Introduction

Being able to play recorded speech out of an electronic device is an extremely fascinating concept and makes the product itself that much more interesting. Especially in areas of interest such as robotics, adding the ability of speech playback is a great add on to any project. There are various ways to add the capability of speech playback to a device. One such way is to utilize adaptive differential pulse code modulation or ADPCM. Because consecutive speech samples are typically close together in value, the algorithm allows you to predict what the next speech sample is and adjust its encoding and decoding of the sample accordingly. This algorithm allows speech data to be compressed by only encoding the difference between the actual audio sample and a predicted audio sample, so it is given at a lower bit rate than the original data, making it easier to house this capability on a device such as a PIC32. The only other ways to add such capabilities to a project would be to use a special audio chip or processor and integrate it with your device somehow. By being able to handle the audio playback in software, without the need for additional hardware, you reduce the overall complexity of a device and don't lose any functionality.

This project is based on this need to have the ability of speech playback on a small device, such as a PIC32 microcontroller, without the need for additional hardware. This simplified ADPCM algorithm is based on guessing what the next speech sample is and adjusting its compression of the sample accordingly. This is actually the fundamental idea of the algorithm, and because it makes sound data much smaller it proves to be a much better choice when adding audio playback capabilities to any project.

II. Implementation

a). ADPCM Algorithm

The ADPCM algorithm has two major parts, the encoding process and the decoding process. The algorithm starts with the encoding part of the audio compression. The overall idea of the encoder is that it takes in a 16-bit speech sample and returns a 4-bit value which will be used to reconstruct the speech sample later on. Essentially, it takes a derivative of the speech sample. At a lower level, the 16-bit speech sample is passed into the encoder. After this, the values of the predicted audio sample and quantizer step size index from the end of the previous iteration of the encoding function are restored. Then the quantizer step size index is used to determine the actual quantizer step size for this iteration of the process. This is followed by finding the difference between the speech sample that is being encoded and the predicted sample. If the difference happens to be negative, then the absolute value of the difference is found and used instead. Next the difference is quantized into a 4-bit ADPCM code using the quantizer step size. This step is followed by the new ADPCM

code being inverse quantized, or numerically integrated, into a predicted difference value, once again using the quantizer step size. Now to find the predicted sample to be used for the next iteration, the new predicted difference value is added to the old predicted sample value. In this version of the algorithm there is a value overflow check, to ensure all predicted sample values are 16-bit signed values. If the predicted sample value falls below -32768, it will cap that value at -32768; if the predicted value rises above 32767 then the value is reset to 32767, thus ensuring the predicted samples never leave the 16-bit signed range. Then the new quantizer step size index is found by using the ADPCM code as an index in a look up table containing various values. This value is added to the current index which gives us the new index for the next iteration. Both the new predicted sample and step size index are then saved and the 4-bit ADPCM code is output from the encoder.

The decoder process is not only simpler than the encoding process but also similar in some ways. The high level idea of the decoding process is that takes the 4-bit code from the encoder and outputs a 16-bit new speech sample. In more detail, the 4-bit code is passed into the decoder. The previous values of the quantizer step size index and the predicted sample are once again gotten from the previous loop of the process. As with the encoder, the quantizer step size is gotten from a table look up using the step size index. Then the 4-bit ADPCM code input is inverse quantized, or numerically integrated, into another predicted difference value. This value is then added to the old predicted speech sample value to get the new speech sample value. After this, there is a bounds check on the value to make sure it stays a 16-bit signed variable. Then the new step size index is found by adding the value from the table of index changes to the current index. Finally the new predicted sample and step size index are saved and the new 16-bit sample is output.

As mentioned earlier the decoding process is similar to the encoding but not exactly the same. The one major difference is the encoder does both quantization and inverse quantization while the decoder only does inverse quantization.

b). MATLAB Baseline

Before beginning to attempt to put the ADPCM algorithm onto a PIC32, I needed to find a way to not only understand how the algorithm worked but hear it for myself. With the suggestion from Professor Bruce Land, I decided to use MATLAB as a way to have a functioning version of Microchip's ADPCM algorithm. Having a working implementation of the algorithm in MATLAB served to be a great basis and tool of comparison throughout this development process.

First I used to the MATLAB code [1] to hear what the compressed audio should sound like running at 16kHz sampling frequency. Then after running the encoder and decoder functions, I converted my audio file into numerical data and output it to a text file. These numbers would serve as the speech samples needed for the ADPCM algorithm on the PIC and would be stored into an array.

c). C Implementation

After having a working implementation of Microchip's ADPCM algorithm in MATLAB, my next step in working on this project was to get the algorithm working in C. This was done to make sure that all the numerical values being passed were the same before attempting to put it on the PIC32. I began to implement just the encoding function in the appendix of the documentation from Microchip in C. I decided I would start with just the encoding function and once it was providing the same output as the MATLAB function, I would move on. After programming the encoding function, I took the floating point values from my MATLAB output text file and put those values into an array. However, the entire audio file was entirely too big to put into an array, so I chopped the down to the first 25000 entries. This also meant that my encoding function would be slightly different than the version in the documentation. Microchip's implementation takes a signed long integer as the input while mine takes floating point values as the input. However, the values still get stored into integers so the outcome is still the same. I used print statements and for loops to test random segments of the data in the audio floating point array. Then I would run the encoding function in MATLAB for those same segments of data, and compare the numbers. After testing and comparing numerous sets of data between the C implementation and MATLAB and receiving the same values, I moved to working on the decoding function. I followed the same process to test the decoder. After programming the function from the appendix in the documentation, I began to test the decoder with the outputs from my encoder. Because I was sure that my encoder worked from my testing I knew that whatever outputs my decoder had, should be correct. I once again compared the outputs of random sets of data from the MATLAB function and the C function and confirmed they were the same. Finally, I ran both codes in their entirety and compared the outputs and the numbers all matched up.

Working On The PIC32

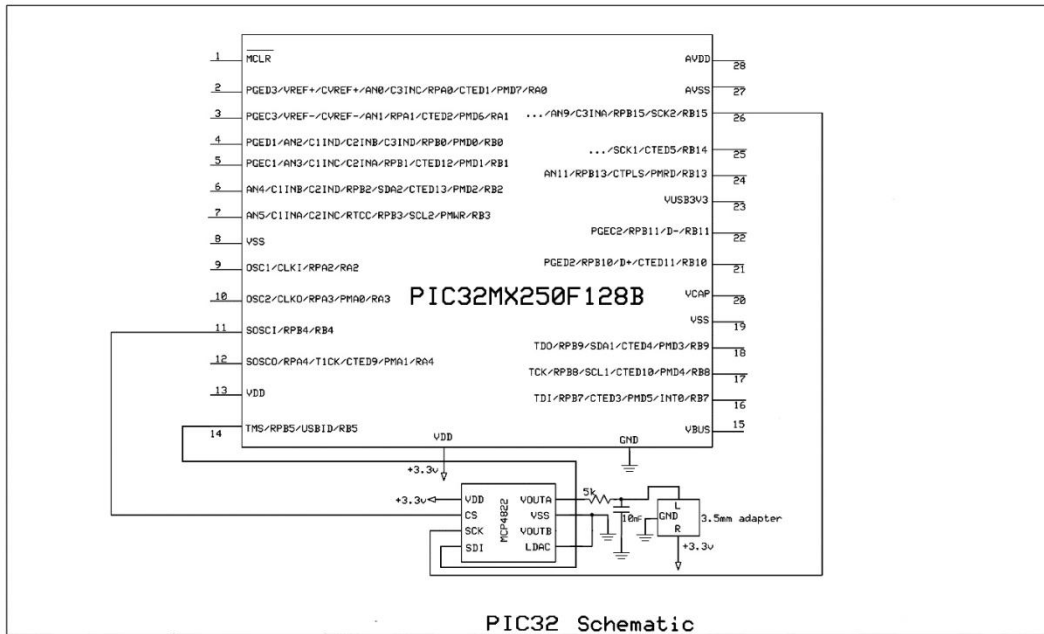


Figure 1: PIC32 Schematic.

The final part of the development part of this project was to move my C code to the PIC32. Even though I knew my ADPCM algorithm worked numerically, porting the code over to the PIC and having audio come out would be no easy feat. I started out by taking some of my old lab 2 code from ECE 4760 and using that as a starting point. Because lab 2 dealt with outputting sound through a 12-bit DAC, it was the perfect code to reference for help. First I added my functions from my standalone C code to my code that would go on the PIC32. Then I moved my array of floating point speech samples to the code. After this was done I began to write code to set up a 12-bit DAC. I needed a DAC in order to be able to output the compressed speech through speakers, to prove that the algorithm does indeed work. Otherwise, there would be no true indication of functionality or not. The first part in setting up the DAC was to define the channels for the DAC, which was promptly followed by setting up the SPI channel and SPI clock divider for the DAC to use. Next, I had to set up lines of code for the DAC to be able to output data through the speakers. Because I had code that used the same concept from ECE 4760, I copied the DAC output lines of code from that class into this code. Not only did this save me time from writing it, I already knew that it would function correctly. However, because the value from the decoder was a 32-bit long value, and the relevant information from the decoder was placed in the 16 most significant bits of this 32-bit value, I needed to shift the value 20 bits in order for the value to be passed through the 12-bit DAC. After setting up the DAC, I needed to set up the timer on the PIC. Because the PIC32 was running at 40MHz I needed to setup the timer, Timer 2, to overflow ever 2500 cycles. This would be needed to ensure that I was running this algorithm at a 16kHz sampling rate, the same as the MATLAB code. After setting up the timer, I moved onto creating the Interrupt Service Routine. The ISR is where the ADPCM algorithm occurs in this project. Every time Timer 2 overflows, the ISR encodes, decodes and outputs to the DAC which causes the speech to be heard. However, in order for the Protothreads threading library to run, there has to be at least one thread in the code. I simply made an empty thread and let it run in a round robin scheduler. After all of the code was written, I then began wiring up the DAC according to the datasheet which was simple and straightforward. Once all of this was done, I downloaded my code to the PIC and wired a speaker audio jack to the DAC and successfully heard the anticipated speech audio.

III. Result

After 6

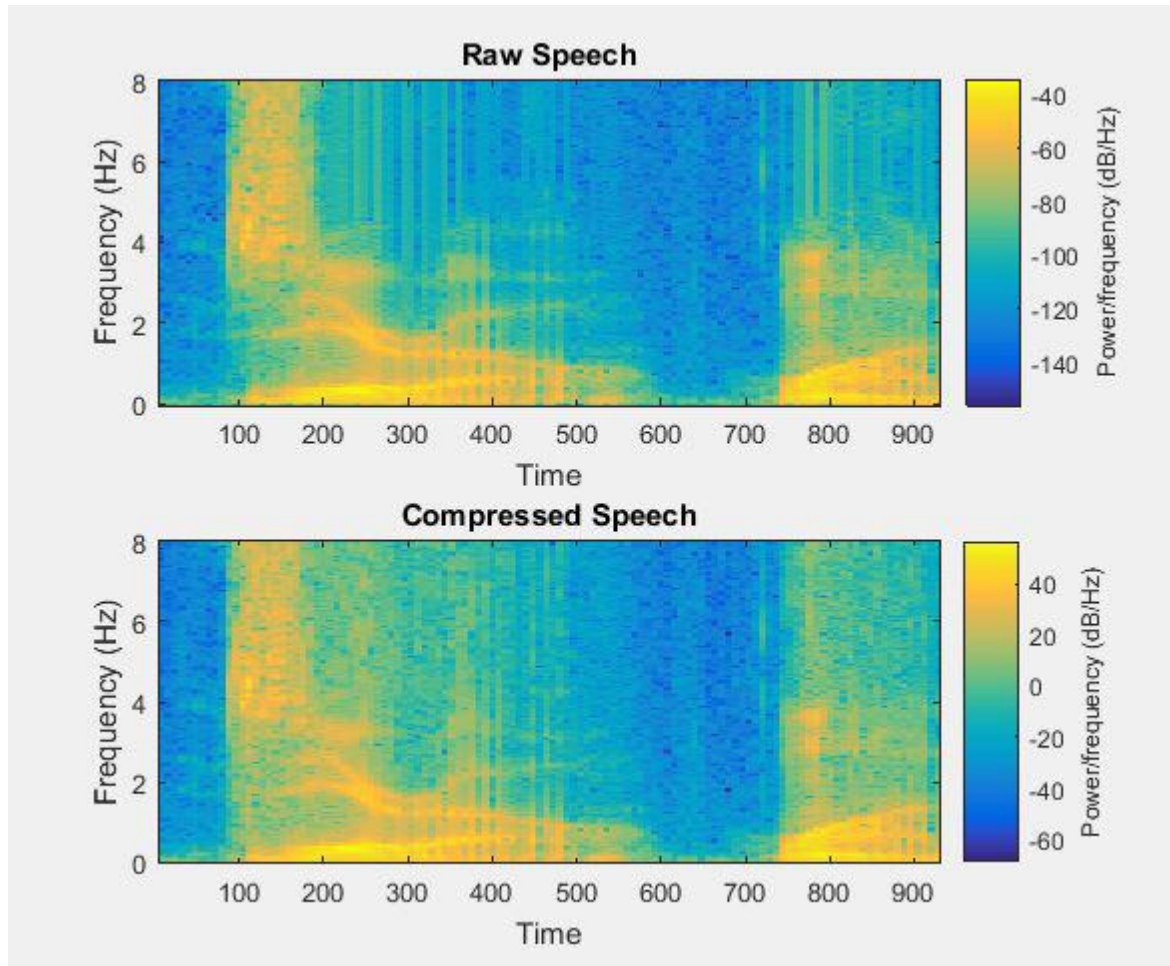


Figure 2: Raw Vs. Compressed Audio Data

Figure 2 contains the spectrograms for the raw audio data and the compressed audio data. As can be seen by comparing the two, the major features of the spectral content are almost identical. There is a bit more high frequencies that can be seen in the compressed spectrogram, as well as a bit of noise. Nevertheless, this quantitatively shows just how close the uncompressed and compressed data actually are. As far as performance on the PIC32 is concerned, the entire ADPCM process took anywhere from 540-581 cycles to complete. Because the timer would overflow at every 2500 cycles, I was roughly using about 1/5 of the CPU on the PIC.

Although the project was successful, there were many issues along the way. The first was when I first attempted the project. I immediately began trying to program the algorithm on the PIC and wiring up the DAC. During these attempts at just going straight to programming the device I did not hear anything coming out of the DAC. This confused me because I had idea

whether I was having issues with the algorithm logic or was it a hardware issue. This caused me to end up getting rid of that entire version of that code. Next was trying to get MATLAB to not output it's numerical data in scientific notation. Because I would need strictly floating point values when working in C and would not want to have to write a particularly complicated conversion from scientific notation function, I had to search and find out about the `formatSpec` function in MATLAB. This gave all my floating point values in a way that I could easily translate into a C array. The third problem I had was loading the array of values onto the PIC. Because the entire audio file used with MATLAB was over 90000 speech samples, I had to trim the number of values down to 25000 when I was working with the C code in Code Blocks. However, once I tried to move that over the PIC, I did not have enough space on the device to hold that many values. This made me cut the number of values down again to 15000 but this still wasn't enough allow it to fit on the PIC32. In order to get around this running out of space issue, I declared the array of floating point values to be `const` so the compiler on the PIC would place the array in Flash instead of RAM. This particular solution was recommended by Professor Land. I also had an issue with hearing strange noises come from the speakers. This was due to the fact that I was attempting to run my algorithm in my main code, but after moving the algorithm to an ISR I began to hear intelligible

speech. A fifth issue was after getting the speech to successfully be heard coming from the PIC, there was quite a lot of noise that was heard through the speakers. To solve this issue I used a 3 kHz low pass filter on the output of the DAC which greatly helped the quality of sound being heard.

IV. Conclusion

This project was designed to implement speech and audio playback using a PIC32 microcontroller using adaptive differential pulse code modulation or ADPCM. ADPCM is a signal encoding process that takes audio data in and produces digital signals as an output. By only recording the differences between a sample and a predicted sample, the predictor can adjust itself appropriately which allows for signals to be produced at lower bit rates than when utilizing standard pulse code modulation.

This implementation is based on Microchip's documentation of a simplified ADPCM algorithm, designed to work on any PIC micro device. Before any actual development was done, a MATLAB implementation of Microchip's algorithm served as not only a basis for sound output but also for comparing the speech sample values as they were passed through the ADPCM encoder and decoder. Once confirming the MATLAB implementation worked, the development of the ADPCM algorithm in standard C began, using Code Blocks as the IDE. This allowed for a comparison of the numerical outputs of the C implementation of Microchip's ADPCM algorithm to the MATLAB implementation. After comparing both the outputs of the MATLAB and C codes with various sets of input data and confirming the correct functionality of the C program, the next step was porting the code to the PIC32.

Before being able to produce any sound output from the PIC32, a few peripherals needed to be set up. A 12-bit DAC, which communicated using SPI, was needed to convert the digital signals to analog signals so the audio could be played through speakers. A timer also needed to be set up, so the audio sampling rate could happen at 16kHz. This sampling frequency was found during the initial testing phases of the MATLAB code. The algorithm works successfully on the PIC32 microcontroller, as it does produce intelligible speech. When comparing the spectrograms of the raw speech data and the compressed audio data, the major features of the speech were unchanged. However, there was a small amount of noise in the compressed speech. It takes the PIC32 anywhere between 540 and 581 cycles to complete the entire ADPCM algorithm.

References

- [1]. <https://www.mathworks.com/matlabcentral/fileexchange/6480-adpcm-encoder-and-decoder?focused=5056015&tab=function>
- [2]. Adaptive differential pulse-code modulation http://en.wikipedia.org/wiki/Adaptive_differential_pulsecode_modulation
- [3]. Pulse-code modulation http://en.wikipedia.org/wiki/Pulse-code_modulation
- [4]. DPCM <https://supportforums.cisco.com/docs/DOC-1423> [4] AVR336: ADPCM Decoder http://atmel.com/dyn/resources/prod_documents/doc2572.pdf
- [5]. Using the ADPCM Algorithm in Dialogic® Voice Processing http://www.dialogic.com/~media/products/docs/appnotes/10532_Dialogic_ADPCM_Algorithm_an.pdf
- [6]. Palnitkar, S. (2006). A Guide to Digital Design and Synthesis (2nd ed). India: Dorling Kindersley Pvt. Ltd.

J. Venkateswara Rao. "Adaptive Differential Pulse Code Modulation using speech compression algorithm." *IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)*, vol. 10, no. 1, 2020, pp. 15-19.