Algorithms for Multi-channel DTMF Detection for the WE DSP32 Family

Steven L. Grant
Missouri University of Science and Technology, sgrant@mst.edu

J. Hartung

G. L. Smith

D6.19

ALGORITHMS FOR MULTI-CHANNEL DTMF DETECTION FOR THE *WE®* DSP32 FAMILY

Steven L. Gay, John Hartung, and Geoffrey L. Smith
AT&T Bell Laboratories
Holmdel, New Jersey 07733

ABSTRACT

This paper describes two DTMF detection algorithms which are highly efficient in the use of both real-time and memory. The first algorithm is based on linear prediction (LP) and can be used to implement up to 32 DTMF detectors on a single 25 MHz NMOS *WE®* DSP32[1]. Using a 50 MHz *WE®* CMOS DSP32C[2], up to 45 detectors can be implemented. The second algorithm is based on a slight modification of the Goertzel algorithm[3] [4] (referred to here as the MG based detector) and can be used to implement up to 16 DTMF detectors on a single DSP32 and up to 32 detectors on a DSP32C. In each of these implementations no external memory is used. Thus, for DSP32 implementations, the small 40 pin DIP package can be used. While the LP based algorithm is the more efficient of the two, the MG based detector performs better in the presence of speech.

INTRODUCTION

The economical detection of DTMF (dual-tone multiple frequency) signals is of critical importance in developing cost-effective telecommunications equipment today. While many single-chip DTMF detectors currently exist, a multiple channel implementation is more appropriate in environments that have a concentration of many lines. Examples include: T-1 facilities in the United States (24 channels) and CEPT facilities in Europe (32 channels). In addition, a digital signal processor (DSP) implementation is often more desirable in applications such as switches where a single hardware resource may be shared among many channels and be used to perform many different signal processing functions at different times.

In section 2 the DTMF receiver requirements are discussed. Sections 3 and 4 describe the LP and MG based algorithms, respectively. The performance of the two algorithms is reported in section 5. In section 6 the multi-channel hardware interface to the DSP32 serial port is described and finally, conclusions and future work are presented in section 7.

DTMF REQUIREMENTS

Figure 1 shows the matrix of frequencies used to encode the 16 DTMF symbols. Each symbol is represented by the sum of the two frequencies that intersect the digit. The row frequencies are in a low band, below 1 kHz, and the column frequencies are in a high band, between 1 kHz and 2 kHz. The digits are displayed as they would appear on a telephone's 4x4 matrix key-pad (on standard telephone sets, the fourth column is omitted). DTMF receivers are required to detect frequencies with a tolerance of ± 1.5 % as valid tones. Tones that are offset by ± 3.5 % or greater, must not be detected. This requirement not to detect tones is necessary to inhibit the detector from falsely detecting speech and other signals as valid DTMF digits. The receiver is required to work with a worst-case signal-to-noise ratio (SNR) of 15 dB and with an attenuation of 26 dB.

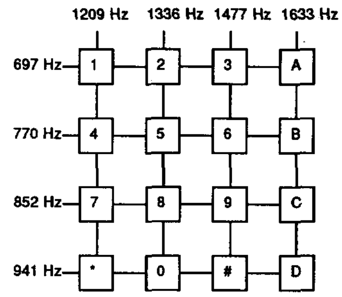

Figure 1. 4x4 Matrix Telephone Keypad

Another requirement is the ability to detect DTMF signals when the two tones are received at different levels. The high-band tone may be received at a lower level than the low-band tone due to the attenuation characteristics of the telephone network. This level difference is called twist, and the situation described is called normal twist. Reverse twist occurs when the low-band tone is received at a lower level than the high-band tone. The receiver must operate with a maximum of 8 dB normal twist and 4 dB reverse twist.

In addition to frequency, noise, and twist requirements, the DTMF signal must meet timing requirements for duration and spacing of digit tones. Digits are required to be transmitted at a rate of less than ten per second. A minimum spacing of 50 ms between tones is required, and the tones must be present for a minimum of 40 ms. Any tone-detection scheme used to implement a DTMF receiver must have a significant time resolution to verify correct digit timing.

A final requirement for the receiver is that it operate in the presence of speech without incorrectly identifying the speech signal as a valid DTMF symbol. This is referred to as digit simulation. Although this requirement is not stated in strict numerical terms, standard recordings such as the Mitel DTMF test tape[5] contain speech segments that are used to test the receiver's digit simulation performance.

THE LP BASED ALGORITHM

The basic approach of the LP based algorithm is to split the incoming signal into two bands so that each tone of the DTMF pair may be analyzed separately. Then, in each band, the roots of the second order prediction polynomial,

$$1-a_1z^{-1}-a_2z^{-2}, \qquad (1)$$

are periodically observed. If a tone is present, the roots of the polynomial should appear on the unit circle at the

CH2673-2/89/0000-1134 $1.00 © 1989 IEEE

corresponding angle. If speech is present, it is unlikely that the resulting prediction polynomial will have roots close to the unit circle at angles corresponding to DTMF tones in both the high and low band simultaneously. This results in a low probability of digit simulation by speech.

The signal from the channel is sub-sampled at 4khz and fourth order bandpass filters are used to separate the low and high group of tones. Within each group the covariance matrix,

$$\mathbf{R}=\begin{bmatrix} r_{1,1} & r_{1,2} \\ r_{1,2} & r_{2,2} \end{bmatrix}$$

and the cross-correlation vector

$$\mathbf{r}=\begin{bmatrix} r_{0,1} \\ r_{0,2} \end{bmatrix}$$

are calculated recursively, using:

$$\mathbf{R}=\lambda\mathbf{R}+\mathbf{y}_n\mathbf{y}_n^T$$

$$\mathbf{r}=\lambda\mathbf{r}+y_n\mathbf{y}_{n-1}$$

where y_n is the present input sample, $\mathbf{y}_n$ is the vector $[y_n, y_{n-1}]^T$ and λ is the "forgetting factor", a number slightly less than one. Every 13.33 ms (the frame rate) the Normal equations, $\mathbf{Ra}=\mathbf{r}$ (where, $\mathbf{a}=[a_1, a_2]^T$), are solved for $\mathbf{a}$ and the roots of the prediction polynomial are found and tested to see if they are within an acceptable range (in magnitude and angle) of a DTMF tone pair. A DTMF signal is present for at least 40ms, so the frame length of 13.33ms will guarantee two full frames of DTMF signal if a valid tone pair is present.

The resolution of the LP algorithm is a function of the number of samples in the estimate of $\mathbf{R}$ and $\mathbf{r}$. At least two frames of data are needed to meet the requirements of section two. To accomplish this, the LP detector is implemented using three modes: a search mode, a verification mode, and a tracking mode. In the search mode $\mathbf{R}$ and $\mathbf{r}$ are set to zero at the beginning of each frame and the apertures of the decision regions are wide. If a single frame of DTMF signal is detected using these wide aperture then the detector transitions to the verification mode. In the verification mode $\mathbf{R}$ and $\mathbf{r}$ are not set to zero at the beginning of the frame and narrow apertures are used. If the same DTMF pair is not detected using the narrow apertures then the detector transitions back to the search mode, otherwise a valid tone-pair is declared and the tracking mode is entered. In the tracking mode the detector simply waits for the DTMF tone to fail the narrow aperture test for the detected tone-pair and a transition back to the search mode occurs.

An efficient method for finding and testing the roots of (1) from $\mathbf{R}$ and $\mathbf{r}$ is now derived. Solving the Normal equations for $\mathbf{a}$ we find that

$$a_1=\alpha_1/(\det R) \tag{2a}$$

and

$$a_2=\alpha_2/(\det R) \tag{2b}$$

where

$$\det R = r_{1,1}r_{1,2}-r_{1,2}^2 \tag{3}$$

is the determinant of R,

$$\alpha_1=r_{2,2}r_{0,1}-r_{1,2}r_{0,2}, \tag{4a}$$

and

$$\alpha_2=r_{1,1}r_{0,2}-r_{0,1}r_{1,2}. \tag{4b}$$

Using the binomial equation it is easy to see that the roots of the prediction polynomial in (1) are at

$$z=\left[a_1+(a_1^2+4a_2)^{1/2}\right]/2 \tag{5}$$

and

$$z=\left[a_1-(a_1^2+4a_2)^{1/2}\right]/2$$

For these to form a complex conjugate pair corresponding to a sinusoid of nonzero frequency, a_2 must be negative with magnitude greater than $a_1^2/4$. Therefore, equation 5 can be expressed as

$$z=\left[a_1+j(-(a_1^2+4a_2))^{1/2}\right]/2$$

The magnitude of which is

$$|z|=\sqrt{-a_2} \tag{6}$$

and the angle is just

$$\theta=Tan^{-1}\left[\frac{\sqrt{-(a_1^2+4a_2)}}{a_1}\right]$$

The decision regions are bracketed by angle and magnitude thresholds for each DTMF frequency. From equations 6 and 2b it is easy to see that the magnitude threshold test can be implemented as the requirement that

$$-\alpha_2 \leq M^2(\det R) \tag{7}$$

where M is the minimum acceptable magnitude. The angle threshold tests for each frequency can be expressed as the requirement

$$f_{i,L}\leq\theta\leq f_{i,H} \tag{8}$$

Where $f_{i,L}$ and $f_{i,H}$ are the low and high thresholds of the i^{th} DTMF frequency, f_i, respectively. For $0<f_i<1000$ Hz, it can be shown that (8) can be written

$$T_{i,L}\alpha_1^2+4(\det R)\alpha_2\leq 0\leq T_{i,H}\alpha_1^2+4(\det R)\alpha_2 \tag{9}$$

Where $T_{i,L}=(1+\tan^2(f_{i,L}))$ and $T_{i,H}=(1+\tan^2(f_{i,H}))$. Similarly, for 1000 Hz $<f_i<$ 2000 Hz (8) becomes

$$T_{i,L}\alpha_1^2+4(\det R)\alpha_2\geq 0\geq T_{i,H}\alpha_1^2+4(\det R)\alpha_2 \tag{10}$$

Therefore, to find and test the roots of (1) α_1, α_2 and det R are calculated. Then, for the lower tone group, tests 7 and 9 are performed or, for the upper tone group, tests 7 and 10 are performed. Where the values $T_{i,L}$ and $T_{i,H}$ for each DTMF frequency and M^2 have been precomputed and stored in tables in memory.

THE MG BASED ALGORITHM

The general approach of the MG based DTMF detector is to examine the energy of the received signal at the eight DTMF frequencies to determine whether a valid DTMF tone pair was received. In addition, as described below, the second harmonics of these frequencies and the total energy of the received signal are used to assist in guarding against digit simulation.

An FFT can be used to calculate the energies of N evenly spaced frequencies. To achieve the required frequency resolution to detect the 8 dtmf frequencies within the ± 1.5 % frequency deviation requirement, a 256 point FFT would be needed (assuming an 8 kHz sample rate). Since only 8 frequencies are of interest, it is more efficient to use the DFT directly

$$X(k)=\sum_{n=0}^{N-1}x(n)e^{-j2\pi kn/N}$$

and calculate only those $X(k)$s closest to the frequencies of interest.

These $X(k)$'s can be more efficiently calculated using the Goertzel algorithm which can be thought of as a matched filter for each DFT frequency. The transfer function of the filter is

$$H_k(z)=\frac{1-e^{j2\pi k/N}z^{(-1)}}{1-2\cos(\frac{2\pi k}{N})z^{(-1)}+z^{(-2)}}.$$

The state variables of the filters are set to zero at the beginning of each analysis frame and at the N^{th} time instant the output of the filters are the desired $X(k)$s.

Since the Goertzel algorithm only finds coefficients for the frequencies of an N point DFT, N must be large to find $X(k)$s close to the DTMF frequencies. Resolution can be increased by evaluating the Fourier transform $X(f)$ at the exact frequencies of interest. This can be accomplished by modifying the transfer function of the matched filter to

$$H_{f_i}(z)=\frac{1-e^{j2\pi f_i/f_s}z^{(-1)}}{1-2\cos(\frac{2\pi f_i}{f_s})z^{(-1)}+z^{(-2)}}. \qquad (11)$$

where f_i is a DTMF frequency and f_s is the sampling frequency.

Using this algorithm a frame length of N=106 was found to give the required frequency resolution. This, once again, guarantees two full frames of data for the minimum 40ms DTMF pulse sampled at 8kHz.

The signal flow graph of the transfer function of (11) is shown in figure 2. The recursive part of the filter is on the left-hand side of the delay elements, and the non-recursive part is on the right-hand side. Since only the output $y_{f_i}(n)$ at time N is needed, it is only necessary to compute the non-recursive part of the filter after the last iteration of the recursive part. A further simplification in the algorithm is made by realizing that only the magnitude squared of $X(f_i)$ is needed. The non-recursive calculation of $y_{f_i}(N)$ is

$$X(f_i)=y_{f_i}(N)=S_i(N)-e^{-j2\pi f_i/f_s}S_i(N-1)$$

where $S_i(N)$ and $S_i(N-1)$ represent the values of the state variables at times N and N-1. It can be shown that

$$\left|X(f_i)=\right|^2=\left|S_i(N)\right|^2-2\cos(\frac{2\pi f_i}{f_s})S_i(N)S_i(N-1)+S_i(N-1)^2$$

This eliminates the need for complex arithmetic and it is seen that it is only necessary to store the coefficient, $2\cos(2\pi f_i/f_s)$, for each $\left|X(f_i)\right|^2$ to be evaluated.

To guard against digit simulation, the total energy is calculated and compared to the sum of the largest magnitudes in the high and low frequency bands. If a valid DTMF tone pair is being transmitted, then the two values should be equivalent. If speech is present, the total energy should be much greater.

The algorithm also exploits the rich harmonic structure of the speech signal to assist in guarding against digit simulation. If speech has energy at a DTMF frequency, f_i, it most likely has significant energy at twice that frequency, $2f_i$ while an actual DTMF signal will have very little. So, in addition to calculating the energy of the received signal at each DTMF frequency the energies at the second harmonics are also found.

At the beginning of each frame, the state variables of each of the 16 modified Goertzel filters are set to zero. Then for 13.33ms (106 samples at a sampling frequency of 8kHz), the recursive part of each filter is executed. At the end of each frame, the square of the magnitude, $X\left|(f_i)\right|^2$, for each DTMF frequency is computed. The following four tests are then performed on the 1st harmonic frequencies to determine if a valid DTMF digit has been detected.

A. *The Magnitude Test.* The largest magnitude in each band has to be greater than a threshold or the DTMF signal is less than -23dBm.

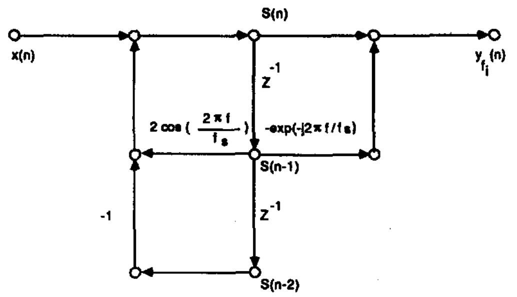

Figure 2. Flow Graph of the Transfer Function

B. *The Twist Test.* The largest magnitude in the low band is compared with the largest magnitude in the high band. This measured twist must fall within the region +4dB to -8dB.

C. *The Frequency Offset Test.* The largest magnitude in each band is compared to the magnitudes of the other frequencies in that band. The difference must be greater than a threshold in each band.

D *The Tone-to-Total Energy Test.* Let c1, c2, and c3 be three different constants, each greater than one. The energy of the low band detected tone is weighted by c1, the energy of the high band detected tone is weighted by c2, and the sum of the two detected tones energies are weighted by c3. Each one of these terms must be greater than the total energy.

E. *The Harmonic Ratio Test.* The energies of the two detected tones are compared to their corresponding 2nd harmonic energies. The difference must be greater than a threshold or the digit is being simulated by speech.

If tests A through E pass, the tone pair is decoded as an integer between 0 and 15. This value is placed in a memory location designated D(j) and is the digit detected for frame j. If any of the tests fail, then -1, representing "no detection", is placed in D(j). For a new valid digit to be declared and sent to the parallel port, D(j) must be the same for two successive frames. If it is valid for more than two successive frames, the receiver is detecting the continuation of the previously validated digit, and a new digit is not output.

PERFORMANCE

The Mitel test tape[5] was used to evaluate the performance of the algorithms. Both algorithms met all of the requirements outlined in Section 2.

Digit simulation was also tested with the Mitel tape. The tape contains 30 minutes of condensed digit simulation speech gathered from several hours of speech recordings. The MG based algorithm outperformed the LP based algorithm by a factor of 2 to 1, even though the LP based algorithm only had 10 false detections, well below Mitels required limit of 30 false detections.

HARDWARE DESCRIPTION

The *WE* DSP32 receives its sampled data input from a time-division multiplexed serial bit stream which contains M channels (where M=16 for the MG based algorithm and M=32 for the LP based algorithm). A sample timing diagram for M=32 is shown in Figure 3. The rate of the input load clock(ILD) must be (8M)kHz (one clock cycle for each channel). The input bit clock(ICK),which is the rate at which bits are input to the DSP32, should equal (64M)kHz. The signal SY, the external synchronization pulse, tells the DSP32 when the first channel is being transmitted. The DSP32 DMA hardware handles the loading of input data to a buffer in RAM.

The output of the DSP32 is connected to an external microprocessor. When a digit has been validated, it is written to the 16 bit parallel I/O data register (PDR). The lower 8 bits contain the decoded digit and the upper 8 bits contain the channel number. For the MG based algorithm both the channel number and decoded digit can be encoded in the lower 8 bits of the PDR. When the PDR is written, the parallel data full (PDF) pin goes high, which can be used to interrupt the external microprocessor.

CONCLUSIONS AND FUTURE WORK

In this paper two DTMF detection algorithms, highly efficient in both real-time and memory, were described. For the MG based algorithm up to 16 DTMF detectors can be implemented on a DSP32 and up to 32 on a DSP32C. The LP based algorithm can be used to implement up to 32 detectors on the DSP32 and 45 on the DSP32C. In each of these implementations it was assumed that only internal memory would be used. While the LP based algorithm was found to be the more efficient of the two, the MG based algorithm was found to have better digit simulation performance.

At the time of this writing the MG based algorithm is being implemented on the *WE*® DSP16 and the *WE*® DSP16A. It is expected that up to 12 DTMF detectors will be implemented on the DSP16 and up to 24 on the DSP16A.

REFERENCES

1. *WE*® *DSP32 Digital Signal Processor Information Manual,* September 1986, AT&T Technologies.
2. *WE*® *DSP32C Digital Signal Processor Advanced Data Sheet,* May 1988, AT&T Technologies.
3. Oppenheim, A. V. and Schafer, R. W., *Digital Signal Processing,* Prentice-Hall, Englewood Cliffs, NJ, 1975
4. Mock, Patrick, "Add DTMF Generation and Decoding to DSP μP Designs," EDN, March 21, 1985.
5. *Tone Receiver Test Cassette #CM7291,* Mitel Technical Data Manual, Mitel Semiconductor, 2321 Morena Blvd. Suite M, San Diego, CA 92110.

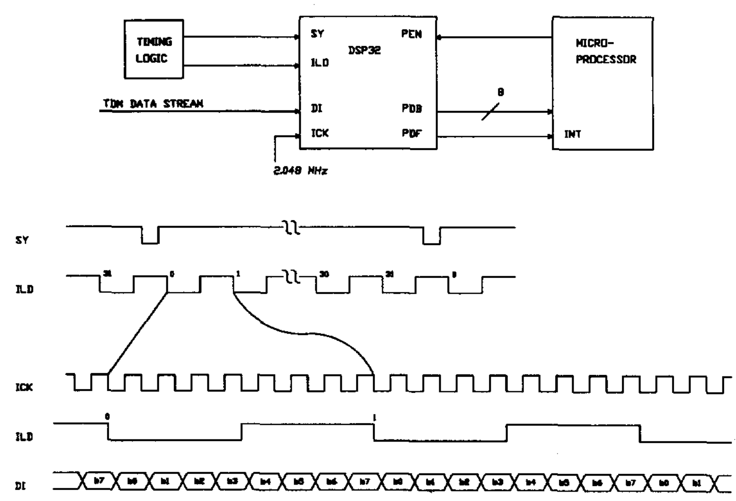

Figure 3. DSP32 TDM and Microprocessor Interfaces