

Efficient Dual-Tone Multifrequency Detection Using the Nonuniform Discrete Fourier Transform

Matthew D. Felder, James C. Mason, and Brian L. Evans, *Senior Member, IEEE*

Abstract—The International Telecommunication Union (ITU) recommendations for dual-tone multifrequency (DTMF) signaling are not met by conventional DTMF detectors. We present an efficient DTMF detection algorithm based on the nonuniform discrete Fourier transform that meets all of the ITU recommendations. The key innovations are the use of two sliding windows and development of sophisticated timing tests. Our algorithm requires no buffering of input samples. To perform DTMF detection on n telephone channels, our algorithm requires approximately n MIPS on a digital signal processor (DSP), $75 + 30n$ words of data memory, and 1000 words of program memory. Using the new algorithm, a single fixed-point DSP can perform ITU-compliant DTMF detection on the 24 telephone channels of a T1 time-division multiplexed telecommunications line.

Index Terms—Modified Goertzel algorithm, multichannel DTMF detection, sliding window, spectral estimation, touchtone dialing.

I. INTRODUCTION

DUAL-TONE multifrequency (DTMF) signaling is used in telephone dialing, voice mail, and electronic banking systems. A DTMF signal corresponds to one of sixteen touchtone digits (0–9, A–D, #, *) and consists of a low-frequency tone and a high-frequency tone. Four low-frequency tones and four high-frequency tones are possible. The DTMF standard was initially developed by Bellcore and recently redefined by the International Telecommunication Union (ITU) [1].

DTMF detection amounts to detecting two sinusoids in noise subject to constraints on frequency resolution, time duration, and signal power. The ITU recommendations, as shown in Table I, place stringent constraints on DTMF detector performance in both time and frequency. In this letter, we present an efficient DTMF detection algorithm based on the nonuniform discrete Fourier transform (NDFT) that meets all of the ITU recommendations. The key innovations are the use of two sliding windows and development of sophisticated timing tests. The detector requires no buffering of input data,

Manuscript received May 28, 1997. The technology described herein is protected by a U.S. patent application entitled "Efficient Digital ITU-Compliant, Zero-Buffering DTMF Detection Using the Non-Uniform Discrete Fourier Transform" and available for licensing from The University of Texas at Austin. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. G. W. Wornell.

M. D. Felder is with the Department of Electrical and Computer Engineering, The University of Texas, Austin, TX 78712-1084 USA, and with the Motorola Engineering Rotation Program, Austin, TX 78704 USA.

J. C. Mason is with the Department of Electrical and Computer Engineering, The University of Texas, Austin, TX 78704 USA, and with the Wireless Subscriber Division, Austin, TX 78721 USA.

B. L. Evans is with the Department of Electrical and Computer Engineering, The University of Texas, Austin, TX 78712-1084.

Publisher Item Identifier S 1070-9908(98)05132-3.

and is simple enough to decode 24 telephone channels of a time-division multiplexed T1 telecommunications line using a single programmable fixed-point digital signal processor (DSP).

II. BACKGROUND

Previous DTMF detection algorithms [2], [3] have been largely based on the discrete Fourier transform (DFT). Given a sequence of N samples, the DFT uniformly samples the discrete-time Fourier transform [2] of the sequence at N evenly spaced frequencies, $\omega = 2\pi k/N$ where $k = 0, \dots, N-1$. Each frequency bin has a width (resolution) of $2\pi/N$. Each bin is centered at an integer multiple of $2\pi/N$, which does not correspond exactly to DTMF frequencies. Although many values of N can meet all of the Bellcore frequency resolution specifications [3], no single value of N can meet all of the ITU frequency resolution recommendations [4].

Instead of computing all N DFT coefficients, detection of DTMF frequencies often uses a bank of eight Goertzel filters [3], [4], i.e., one filter per DTMF frequency. The Goertzel filter is typically implemented as a second-order infinite impulse response (IIR) bandpass filter [2] with the transfer function

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

The N th output sample of the Goertzel filter is the k th DFT coefficient. The Goertzel filter requires $2N$ real multiplications/additions and three words of memory for the denominator section, and the numerator section is only computed on the N th input sample. The filter can be realized without input buffering because each sample can be processed when received [3]. By setting k to yield an exact DTMF frequency of interest f_i , i.e., $k = N f_i / f_s$, where f_s is the sampling rate (8000 Hz), we implement the NDFT [5] to detect energy at exact DTMF frequencies [6].

A DTMF detector applies frequency detection to windows of samples. Rectangular windowing produces sidelobes in the frequency spectrum. The sidelobes smear energy present in the mainlobe of the original signal [2]. For a rectangular window, the largest sidelobe is 13 dB down from the mainlobe [2], which permits us to meet the ITU recommendations. Using another window may reduce the impact of sidelobes, but at the expense of increasing computational complexity and decreasing frequency resolution because the mainlobe widens. We use a rectangular window to provide the narrowest mainlobe and minimize computational complexity.

TABLE I
INTERNATIONAL TELECOMMUNICATION UNION RECOMMENDATIONS FOR DTMF SIGNALING

<i>Signal Frequencies</i>	Low Group	697, 770, 852, 941 Hz
	High Group	1209, 1336, 1477, 1633 Hz
<i>Frequency Tolerance</i>	Operation	≤1.5% of above frequencies
	Non-operation	≥3.5% of above frequencies
<i>Signal Duration</i>	Operation	40 ms minimum
	Non-operation	23 ms maximum
<i>Signal Exceptions</i>	Pause Duration	40 ms maximum
	Signal Interruption	10 ms minimum
<i>Twist</i>	Forward	8 dB
	Reverse	4 dB
<i>Signal Strength</i>	Signal-to-Noise Ratio	15 dB minimum
	Signal Power	-26 dBm minimum

TABLE II
PERCENT ENERGY LOSS DUE TO FREQUENCY OFFSET ERRORS FOR THE MINIMUM AND MAXIMUM FREQUENCY OF THE LOW AND HIGH BANDS (WORST CASE). STRIKE-THROUGHS ARE UNACCEPTABLE VALUES FOR THE ITU RECOMMENDATIONS AS DESCRIBED IN SECTION III

N	Low Frequency Group				High Frequency Group			
	697 Hz		941 Hz		1209 Hz		1633 Hz	
	1.5%	3.5%	1.5%	3.5%	1.5%	3.5%	1.5%	3.5%
106	6.54%	29.1%	12.3%	46.8%	18.7%	67.8%	30.5%	91.4%
165	15.3%	58.3%	25.9%	83.4%	39.4%	97.8%	60.5%	97.8%
212	24.2%	79.8%	39.4%	97.8%	56.9%	98.8%	81.3%	98.8%

III. MEETING ITU FREQUENCY SPECIFICATIONS

A DTMF frequency is valid if it is within the band of ±1.5% of the nominal frequencies listed in Table I and invalid if it is outside the band of ±3.5% of the nominal frequencies. The ITU timing specifications require detection of DTMF signals of at least 40 ms duration. At a sampling rate of 8 kHz, 40 ms corresponds to 320 samples, which is an upper bound on the window size. A smaller frame size is necessary to meet all of the ITU timing specifications.

The ITU timing specifications are most easily met with a window length (frame size) of 106 samples [6]. At a sampling rate of 8000 Hz, a 106-sample frame size corresponds to a 13.3 ms window, which guarantees that a 40 ms DTMF tone would fill at least two complete frames. Unfortunately, this frame size does not meet the ITU frequency resolution specifications for all frequencies. For example, the detector is unable to accept a low-group tone at either 697 Hz + 1.5% or 697 Hz - 1.5% and reject a high-group tone at either 1633 Hz + 3.5% or 1633 Hz - 3.5%.

A frame size larger than 106 samples for the low group is necessary to meet the frequency specifications. A larger frame size, however, would make it more difficult to meet the ITU timing specifications, as described in the next section. A larger frame size would also increase the selectivity of the filter, thereby making it more likely to reject frequencies with 1.5% error.

Table II details the tradeoffs for selecting the frame size for the NDFT. In Table II, each entry is the percentage energy

lost due to difference between the received frequency and a DTMF frequency. Given any DTMF signal, the sum of DTMF energy losses (low and high group) with simultaneous 1.5% frequency errors should not exceed the DTMF energy loss with either frequency having 3.5% error. Otherwise, the detector would not be able to reject all 3.5% errors and accept all 1.5% errors. The values with “strike-throughs” in Table II violate this condition.

No single frame size meets both the high-group and low-group frequency specifications, so we choose to use a different frame size for each group. A frame size of 212 for the low-group frequencies and 106 for the high-group frequencies meets the ITU frequency specifications.

IV. MEETING ITU TIMING SPECIFICATIONS

Previously reported DTMF detectors do not meet the ITU tone interruption or the minimum tone duration recommendation. For example, one detector [6] uses sampling windows of 13.3 ms in duration. After each window, the detected signal is compared to the last and second-to-last values. If the result of the new window is the same as the last, but different from the second-to-last, then a new valid DTMF signal has been found. A 10 ms tone interruption generates zero, one, or two invalid frames, so this method might incorrectly detect two separate DTMF tones. A DTMF tone that lasts less than 23 ms should not be detected. However, in the method [6] described in the previous paragraph, a 20 ms signal centered between

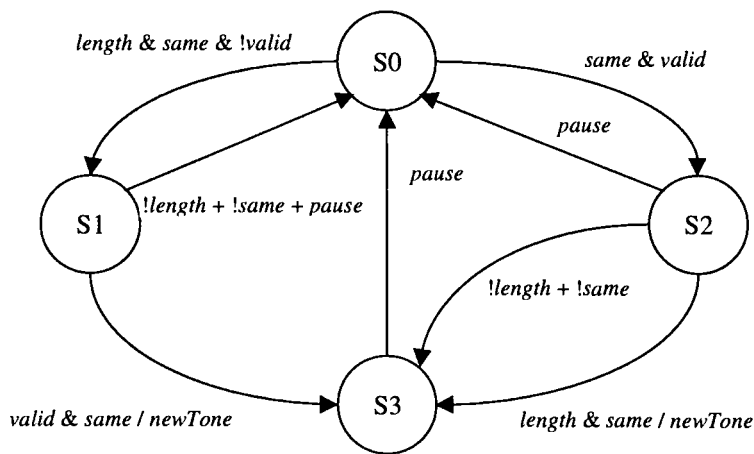


Fig. 1. FSM for the DTMF detector. The FSM specifies the control logic to meet the ITU signal timing specifications. The notation a/b on a transition arc means to emit signal b when condition a is true. Only one transition is taken every 106 samples.

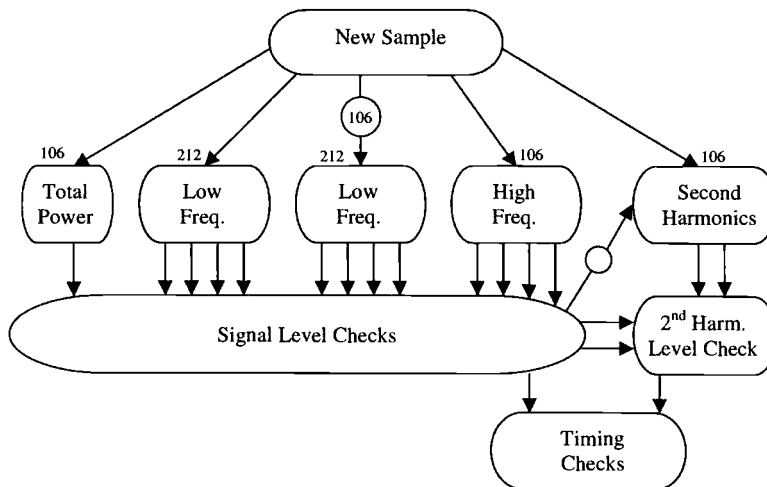


Fig. 2. Simplified dataflow graph of the new DTMF detector. The DTMF detector processes one sample at a time. The number outside of a block indicates the number of samples processed before the block is reset to its initial values. The circled 106 means that the “Low Freq.” block below it is reset 106 samples after the other “Low Freq.” block is reset. The three checks are only performed after every 106 samples.

two frames might be accepted if the signal contains enough DTMF energy.

We meet the ITU tone duration and signal interruption recommendations with two enhancements. First, the tone duration is considered sufficient when the DTMF signal spans two entire frames. We enforce this by requiring that some percentage of the DTMF power be present in a frame either before or after two valid frames. Second, the tone interruption recommendation is met by using the finite state machine (FSM) in Fig. 1. The FSM is controlled by four values determined after each frame of 106 samples: *valid*, *same*, *length*, and *pause*. *Valid* is true if a DTMF signal is present in the two most recent frames of 106 samples and passes a series of energy level checks. *Same* is true if the decoded digit is the same as the digit for the previous frame. *Length* is true when the two most recent frames of 106 samples have DTMF energies within a percentage difference. *Pause* is true if the DTMF energy is less than a given threshold of the total energy for the two most recent frames.

The FSM starts in state S0 and waits for a valid DTMF signal. If all conditions are met, the *newTone* signal is emitted

on the transition to state S3. From state S3, the FSM awaits a pause signal to return to state S0. This FSM meets the signal interruption and pause duration recommendations while maximizing the detector’s noise immunity.

V. DESIGN, IMPLEMENTATION, AND TESTING

Fig. 2 describes the detector. After every 106 samples, the detector outputs whether or not a valid DTMF signal is present and the DTMF digit. To detect the strength of each low-group frequency, we use a pair of Goertzel filters with a frame size of 212 samples. One filter is offset from the other by 106 samples, which is equivalent to using a sliding window of 106 samples. Each high-group frequency requires one Goertzel filter with a frame size of 106 samples.

We perform the following postprocessing steps every 106 samples. First, we find the strongest frequency in the high group and low group. Then, we compare signal levels [3] and check the ratio of the DTMF power to the total power against a threshold [6]. One new feature is that we adjust the threshold according to the DTMF frequencies that were

detected. Finally, we use two Goertzel filters to find the second harmonic energies of the dominant detected low-group and high-group tones. To guard against detecting DTMF tones in speech, DTMF detectors [3], [6] calculate second harmonics of the eight possible tones using eight Goertzel filters, yet only two of the filter outputs are compared each frame. We use the two dominant tones found during the last 106-sample frame and compute their second harmonics during the current frame. The FSM in Fig. 1 uses the results of the signal level checks and enforces the ITU timing constraints.

The detector processes each input sample using 14 Goertzel filters. Each filter requires eight instructions/sample, two words of read/write data memory, and one word for a constant. The total signal power calculation requires three instructions/sample and one word of read/write data memory. At the end of each 106-sample frame, 116 instructions are required to perform the postprocessing: 48 instructions to calculate the output power at eight of the Goertzel filters, 60 instructions to perform the signal analysis described in Section IV, and eight instructions for the FSM. When processing one telephone channel sampled at 8000 samples/s, the detector requires approximately one MIP, 75 words of read/write memory, 75 words of constants, and 1000 words of program memory. The MIP rating assumes a conventional DSP architecture that has parallel move operations and a single-cycle multiply and accumulate. Decoding 24 channels of a T1 line requires about 24 MIPS, 800 words of read/write data memory (by reusing memory locations), and 1000 words of program memory, which can be met by fixed-point DSP's such as the Motorola 56000 and the Texas Instruments TMS320C50.

We develop a DTMF detector testbed using Ptolemy 0.7 [7]. We built a generator to produce DTMF signals with the following parameters: frequency deviation, tone duration, pause duration, signal interruption, twist scaling, second harmonic energy, Gaussian noise, and attenuation. By varying these parameters, we validated that our detector meets all of the ITU recommendations. Our detector's guard time is 31.8 ms. The detector achieves 100% detection at 13 dB SNR and higher, and 95% detection at 12 dB SNR. The power sensitivity is -33 dBm. The detector accepts DTMF signals with twist between 8.3 and -4.3 dB with 100% detection, and completely rejects twist greater than 8.9 dB and reverse twist greater than 4.7 dB. The detector passed all of the frequency deviation tests by allowing simultaneous 1.5% errors on both frequencies, while rejecting 3.5% error on either frequency. The Bellcore standard requires a guard time (the minimal DTMF signal length that can be reliably detected) of 40 ms or less, lowest SNR for reliable detection of 23 dB or lower, power sensitivity of -25 dBm or lower, forward twist of 8 dB or more, and reverse twist of 4 dB or more.

Since the Bellcore DTMF standard is a subset of the ITU standard, an ITU-compliant DTMF detector must also be Bellcore-compliant. Bellcore produces a set of test tapes that simulate one million customer dialing attempts to a central

office and establishes a maximum number of acceptable false detections. The Bellcore DTMF test tapes showed that the talk-off performance was exceptional: the detector had less than half of the acceptable number of false detections when invalid DTMF signals (such as speech) were input.

VI. CONCLUSION

We develop and implement an efficient DTMF detector that meets all of the ITU recommendations. Our contributions include the use of two sliding windows and sophisticated timing specification checks. Two sliding windows are necessary because no single window length can meet ITU specifications. Since the low-group DTMF frequencies require a larger window size than the high-group DTMF frequencies, we choose the low-group window length (212) to be twice the high-group window length (106) to meet ITU specifications and eliminate the need to buffer input data. The sliding windows compute the NDFT using a modified Goertzel algorithm, but other more computationally efficient frequency detectors such as [8] could have been adapted. Our detector postprocesses the outputs of the sliding window NDFT's to enforce the timing specifications by means of a finite state machine (FSM). The FSM resolves the signal interruption problems in other published DTMF detectors and improves the detector's noise immunity. When applied to 24 channels of a T1 line, our DTMF detector requires about 24 DSP MIPS, 800 words of data memory, and 1000 words of program memory, so it can be implemented on a single fixed-point digital signal processor.

ACKNOWLEDGMENT

The authors would like to thank Prof. F. A. Sakarya, Yildiz Technical University, Istanbul, Turkey, for mentioning the open research issue of developing ITU-compliant DTMF detectors.

REFERENCES

- [1] ITU Blue Book, *Recommendation Q.24: Multi-Frequency Push-Button Signal Reception*, Geneva, Switzerland, 1989.
- [2] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [3] P. Mock, "Add DTMF generation and decoding to DSP microprocessor designs," *Electron. Data News*, vol. 30, pp. 205–220, Mar. 21, 1985.
- [4] G. Arslan, B. L. Evans, F. A. Sakarya, and J. L. Pino, "Performance evaluation and real-time implementation of subspace, adaptive, and DFT algorithms for multi-tone detection," in *Proc. IEEE Int. Conf. Telecommunications*, Apr. 1996, pp. 884–887.
- [5] S. Bagchi and S. K. Mitra, "The nonuniform discrete Fourier transform and its applications in filter design: Part I—1-D," *IEEE Trans. Circuits Syst. II*, vol. 43, pp. 422–433, June 1996.
- [6] S. L. Gay, J. Hartung, and G. L. Smith, "Algorithms for multi-channel DTMF detection for the WEDSP32 family," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Apr. 1992, pp. 909–912.
- [7] "Ptolemy 0.7 User's Manual," Dept. Electr. Eng. Comput. Sci., Univ. Calif., Berkeley, CA, 1997, HYPERLINK <http://ptolemy.eecs.berkeley.edu>.
- [8] J. A. R. Macias and A. G. Exposito, "Efficient moving-window DFT algorithms," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 256–260, Feb. 1998.