Efficient Dual-Tone Multi-Frequency Detection Using the Non-Uniform 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 multi-frequency (DTMF) signaling are not met by conventional DTMF detectors. We present an efficient DTMF detection algorithm based on the non-uniform 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), $120 + 30\,n$ 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.

The authors are with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-1084 (e-mail: felder@ece.utexas.edu, mason@ece.utexas.edu, and bevans@ece.utexas.edu).

M. D. Felder is also with the Motorola Engineering Rotation Program, 505 Barton Springs Rd., Suite 404, Austin, TX 78704. J. C. Mason is also with the Wireless Subscriber Division, 3501 Ed Bluestein Boulevard, Austin, TX 78721.

The technology described herein is protected by a US 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.

I. INTRODUCTION

Dual-tone multi-frequency (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 later refined 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 1, place stringent constraints on the DTMF detector performance in both time and frequency. In this letter, we present an efficient DTMF detection algorithm based on the non-uniform 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, 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 $k = 0, ..., N - 1$. Each frequency bin has a width (resolution) of $2\pi/N$. Each bin is centered at an integer multiple of $2\pi/N$, which is an exact DTMF frequency only for an appropriate large N (e.g. 8000). Although many values of N can meet the Bellcore frequency resolution and time duration specifications [3], no single value of N can meet the equivalent ITU specifications [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 IIR bandpass filter [2] with the transfer function

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

The Nth output sample of the Goertzel filter is the kth DFT coefficient. For the denominator section, the Goertzel filter requires N real multiplications, $2N$ real additions, and 3 words of memory. The numerator section is only computed on the Nth input sample. The filter can be realized without input buffering because each sample can be processed when it is 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.

III. MEETING ITU FREQUENCY SPECIFICATIONS

A DTMF frequency is valid if it is within the band of ±1.5% of a nominal frequency listed in Table 1 and invalid if it is outside the bands 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.

ITU time duration 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 all of 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 2 details the tradeoffs for selecting the frame size for the NDFT. In Table 2, each entry is the percentage energy lost from the difference between the received frequency and a DTMF frequency due in part to the sinc shape of the mainlobe of the magnitude frequency response of the rectangular window [2]. Given any received 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 2 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 minimum tone duration recommendations. For example, one detector [6] uses a window length of 13.3 ms. After each window, the detected signal is compared to the last and second-to-the-last values. If the result of the new window is the same as the last, but different from the second-to-the-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. The same detector has difficulty rejecting DTMF tones that last less than 23 ms. A 20 ms signal centered between two frames might be accepted if the signal contains enough DTMF energy.

We meet the ITU minimum tone duration and tone 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 Figure 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 ITU specifications are met, then 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

Figure 2 describes the detector's operation. 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 post-processing 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 Figure 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 Goertzel filter requires 8 instructions/sample, 2 words of read/write data memory, and 1 word for a constant. The total signal power calculation requires 3 instructions/sample and 1 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 8 instructions for the FSM. When processing one telephone channel sampled at 8000 samples/s, the detector requires approximately 1 MIP, 75 words of read/write memory, 75 words of constants, and 1000 words of program memory. The MIP rating assumes a conventional digital signal processor architecture that has parallel move operations and a single-cycle multiply and accumulate. Decoding 24 channels of a T1 line requires about 24 MIPS, 840 words of read/write data memory (by reusing memory locations), and 1000 words of program memory, which can be met by fixed-point DSPs 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. We varied these parameters to validate that our detector is ITU-compliant. Our detector’s guard time is 31.8 ms (guard time is the minimal DTMF signal length that can be reliably detected). 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 dB 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. These ratings surpass the ITU specifications in Table 1. Note that the Bellcore standard requires a guard time of 45 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: our 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 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. We show that no single window length can meet ITU specifications and that the low-group DTMF frequencies require a larger window size than the high-group DTMF frequencies to meet ITU specifications. We use two low-group windows of 212 samples each and a high-group window of 106 samples. We arrange the two low-group windows to be offset, overlapping windows to gain the advantages of sliding windows without having to buffer data. To compute the NDFTs for frequency detection, we use the modified Goertzel algorithm, which we realize without buffering. A faster frequency detection technique, such as a variant of [8], could have been used instead. After every 106 input samples, our detector enforces timing specifications by means of a finite state machine (FSM). The FSM resolves the signal interruption problems in other published DTMF detectors and improves noise immunity. When applied to 24 channels of a T1 line, our DTMF detector requires about 24 DSP MIPS, 840 words of data memory, and 1000 words of program memory, so it can be implemented on a single inexpensive fixed-point DSP.

ACKNOWLEDGEMENTS

We would like to thank Prof. F. Ayhan Sakarya at the Yildiz Technical University in 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, 1989.

[2] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, 1989.

[3] P. Mock, “Add DTMF Generation and Decoding to DSP Microprocessor Designs,” *Electronic 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,” *Proc. IEEE Int. Conf. on Telecommunications*, pp. 884-887, Apr. 1996.

[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 and Systems-II: Analog and Digital Signal Processing*, vol. 43, no. 6, pp. 422-433, Jun. 1996.

[6] S. L. Gay, J. Hartung, and G. L. Smith, “Algorithms for Multi-channel DTMF Detection for the WEDSP32 Family,” *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 909-912, Apr. 1992.

[7] “Ptolemy 0.7 User’s Manual,” Dept. of EECS, University of California at Berkeley, Berkeley, CA 94720-1770, 1997, http://ptolemy.eecs.berkeley.edu/.

[8] J. A. R. Macias and A. G. Exposito, “Efficient Moving-Window DFT Algorithms,” *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 45, no. 2, pp. 256-260, Feb. 1998.

Table and Figure Captions

Table 1. International Telecommunication Union recommendations for DTMF signaling.

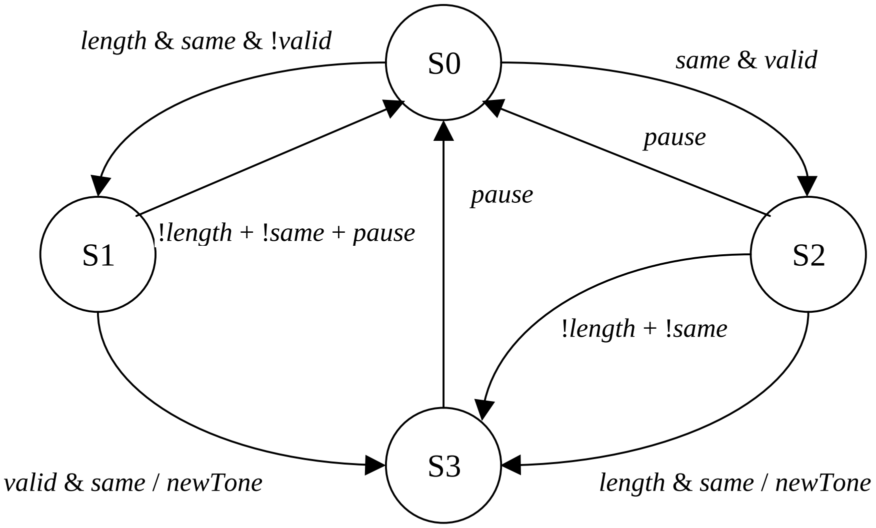
Table 2. Percent energy loss due to frequency offset errors for the minimum and maximum frequency of the low and high bands (worst case). Strike-throughs are unacceptable values for the ITU recommendations as described in Section III.

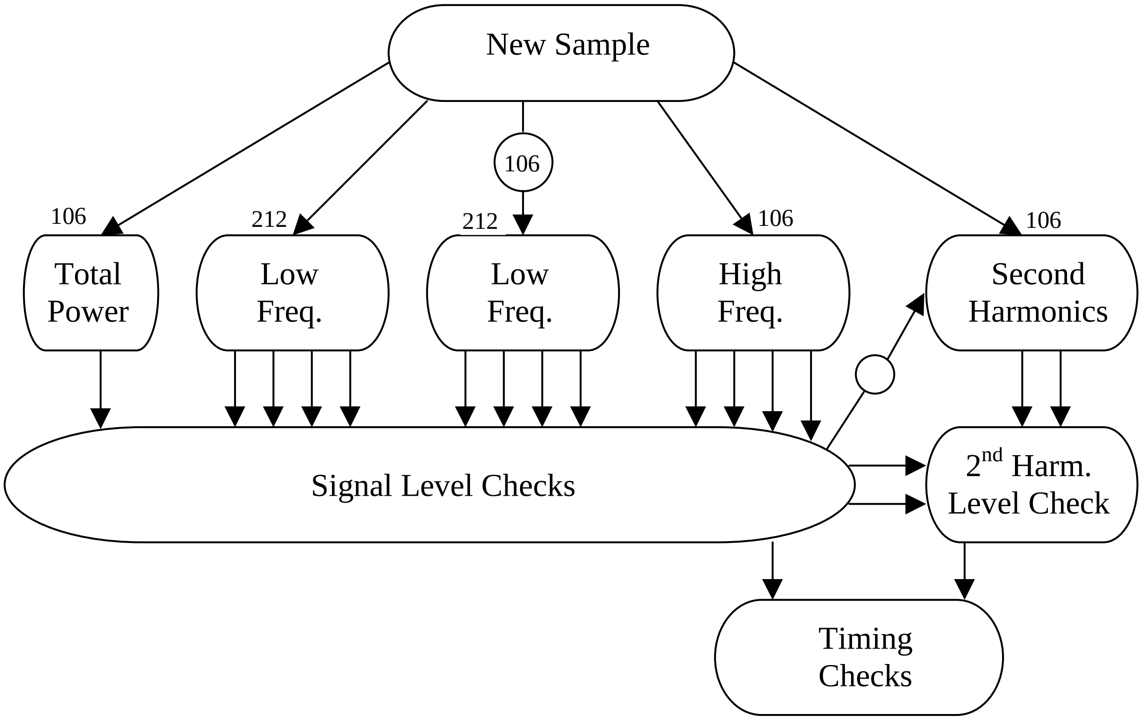
Figure 1. A finite state machine (FSM) for the DTMF detector. The FSM specifies the control logic to meet the International Telecommunication Union signal timing specifications. The notation *a* / *b* on a transition arc means to emit signal *b* when condition *a* is true. After every 106 input samples to the detector, the FSM makes at most one transition.

Figure 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.

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

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%


length & same & !valid
S0
same & valid
pause
pause
S1
!length + !same + pause
S2
!length + !same
S3
valid & same / newTone
length & same / newTone


New Sample
106
106
212
212
106
106
Total Power
Low Freq.
Low Freq.
High Freq.
Second Harmonics
Signal Level Checks
2nd Harm. Level Check
Timing Checks