An Autoregressive Model Based Approach for DTMF Detection

A.V.N.S Anjani and *B.Yegnanarayana*
Department of Computer Science and Engineering
Indian Institute of Technology, Madras, 600 036, India
email: {anjani@hynek,yegna@svalpha2}.iitm.ernet.in

Abstract

In this paper, the problem of designing a DTMF detector that meets international specifications is discussed. Methods based on explicit frequency estimation are briefly reviewed. A method for DTMF detection based on spectral estimation using autoregressive modeling is proposed. The method meets all the specifications. The issue of robustness of the method in presence of noise and speech interference is also discussed. Results are given to demonstrate the efficacy of the method.

1 Introduction

Dual Tone Multi Frequency (DTMF) is the generic name for push button signalling. The conventional pulse signalling has been replaced by DTMF signalling as DTMF offers fast dialing and the dialing rate is independent of the number dialed [1]. Generation of DTMF signals is straightforward. As the name suggests, digital samples of two sine waves are generated, scaled and added together to produce a DTMF signal. Fig.1 shows the DTMF frequencies for the digits 0 to 9. The standards for DTMF generation are more or less same through out the world. The frequencies are the same while the power levels and tolerances are similar.

Decoding a DTMF signal involves extracting the two tones in the signal and determining from their values, the intended DTMF digit. Most of the DSP based DTMF detection algorithms do not meet all specifications. Only certain approaches based on explicit frequency estimation might work. The DTMF detection should be robust to speech interference as well as channel effects for it to be useful for commercial applications. In this paper we propose a method for DTMF detection based on spectrum estimation using autoregressive (AR) modeling. The method meets all the specifications and is also robust to speech interference. This paper is organized as follows: In section 2, methods for spectrum estimation are briefly reviewed. In section 3, the proposed method for frequency detection using autoregressive modeling is presented. The results obtained from the experimental studies conducted are given in section 4.

HF / LF	1209	1336	1477	1633
697	1	2	3	A
770	4	5	6	B
852	7	8	9	C
941	*	0	#	

Figure 1: DTMF keypad

2 Methods for spectrum estimation

The classical methods of spectrum estimation such as Discrete Fourier transform (DFT), operate on windowed data or windowed autocorrelation function (ACF) estimates. Windowing makes the implicit assumption that the values outside the window are zero which is a very unrealistic assumption. A smeared spectrum estimate is the consequence of such a windowing. Use of a priori information may permit the selection of an approximate model for the process that generated the data samples. One of the advantages of spectral estimation using modeling approach is the high resolution achievable with these techniques over conventional techniques. Moreover, we can make realistic assumptions concerning the nature of the measured process outside the measured interval. Thus the need for windowing functions can be eliminated together with their distorting effects.

3 Autoregressive model based method for frequency estimation

Some of the considerations involved in choosing a model are the ability of the model to represent spectral peaks, valleys and roll-offs. For spectra with sharp peaks it is necessary to employ a model that has poles. Autoregressive(AR) model is suitable for such purpose. AR model is good for spectral estimation when the spectrum is characterized by sharp peaks. The use of AR model for a sinusoidal data is shown to be valid in the context of spectral estimation as the spectrum of a sinusoid is characterized by a sharp peak [2].
A process is termed an autoregression (AR) if the sequence $x[n]$ is a linear regression on itself with some error $u[n]$.

$$x[n] = -\sum_{k=1}^{p} a_k x[n-k] + u[n] \tag{1}$$

represents an AR model of order p. The reason for high resolution property of the AR spectral estimate is the implicit extension of the measured autocorrelation sequence. The AR spectra do not exhibit side lobes due to windowing. Accurate estimates of the AR parameters can be found by solving a set of linear equations. When the AR modeling assumption is valid, AR spectrum estimators have less bias and low variability than the conventional Fourier- based spectrum estimators. Other names by which the AR spectrum estimator is known are the maximum entropy spectrum estimator (MESE) and the linear prediction (LP) spectral estimator. We have approached the problem of DTMF detection from LP spectrum estimation. In linear prediction, the unobserved sample $x[n]$ of a signal is predicted based on the past p samples. The optimal predictor, one that minimizes the mean square error is given by

$$\hat{x}[n] = -\sum_{k=1}^{p} a_k x[n-k] \tag{2}$$

The prediction coefficients $a_1, a_2, ...a_n$ are chosen so as to minimize the energy of the prediction error $e[n]$ where

$$E = \sum_n e^2[n] = \sum_n (x[n] - \hat{x}[n])^2 \tag{3}$$

3.1 Choice of model order

Choice of model order in LP spectral estimate is critical. Too low an order results in smoothed spectrum estimate, while too large an order causes spurious peaks. For a signal of M real sinusoids the minimum choice of p, i.e., the order of LP analysis is 2M.

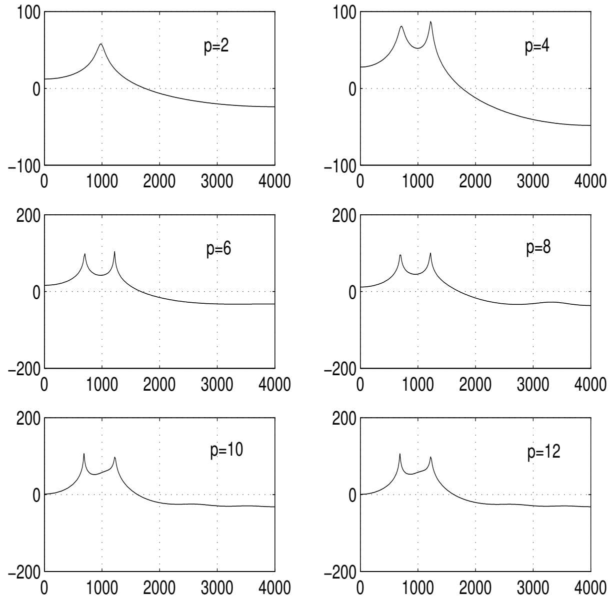

Figure 2: LP spectrum of sum of sinusoid for varying LP order

3.2 Effect of noise on LP spectral estimate

The effect of noise is to produce a smoothed spectrum estimate. The resolution decreases with increasing signal-to-noise ratio (SNR). A solution for this problem is to have a high model order. The resolution increases with increasing model order in presence of noise. But too high an order results in spurious peaks. So a trade-off is expected. Fig.2 shows the LP spectrum of a signal which is a sum of two sinusoids of frequencies 697 Hz and 1209 Hz, for order p varying from 2 to 10.

3.3 Steps in the proposed method

The set of specifications for DTMF detection are shown in [3]. These specifications indicate that the detection algorithm should detect digits over a wide range of power levels. The frequency offset is the more important aspect. Even if the power level requirements are met, the tone outside the offset range should be rejected. In our algorithm while doing the frequency estimation, we are assuming that the valid input DTMF frequencies will always be within $\pm 2\%$ range of the standard value. An estimate of the tone duration is also required to check whether the tone sustains for some specified minimum duration. The implementation details are as follows:

- The signal is blocked into frames of N samples with adjacent frames being separated by M samples.

- Let TD_{min} be the minimum duration for which the DTMF digit tone should sustain for it to be declared to be valid. If $T_{frame} > TD_{min}/2$, and if DTMF digit is present in the frame, one is not sure how much of the frame is actually filled with the tones. Choice of $T_{frame} \leq TD_{min}/2$ ensures that at least one frame is filled with the tones for every valid digit dialed. since $TD_{min} = 25$ m-sec and sampling frequency is 8 kHz, the DTMF signal for a digit should last for at least 200 samples. We considered a frame size of 12.5 msec(100 samples) for our studies.

- The adjacent frames are separated by a duration of T_{shift}. Smaller T_{shift} leads to a better temporal resolution but at the cost of higher computational complexity. Therefore, there is a trade off. We have chosen a frame shift of 10 samples taking both onset point accuracy and computational complexity into account.

- The LP coefficients are obtained for each frame. To get a good frequency resolution, the order of the model should be high. As the model order increases, the accuracy in the frequency estimation increases. For our studies, we have used the model order as 12.

  Let us consider the polynomial $A(z)$

$$A(z) = \sum_{k=0}^{p} a_k.z^{-k} \tag{4}$$

  where a_ks are LP coefficients with $a_0 = 1$. The roots of the polynomial lie inside the unit circle in z-domain. The angle of these roots are calculated to get the resonance frequencies of the LP spectrum. A complex conjugate root pair accounts for one peak in the LP spectrum. Problems of submerging of weak peaks in the roll off spectrum of a closely placed strong peak are not encountered in our approach. Phase spectrum is not sensitive to the strength of the peaks, therefore weak peaks even if placed near to the strong peaks, are not lost.

Three checks were made in our method to detect the DTMF digit. They are:

1. If the frequencies obtained lie in the DTMF band, the condition of offset limit is checked. The tolerance limit of $\pm 2\%$ is used in our implementation.

2. The frame index where the offset condition is satisfied is marked as "Onset" of the tone. At this stage, the condition of minimum tone duration is imposed to check that the frequencies with in the offset range are present for a minimum duration of 25 msec. A 7 point median smoothing is done on the estimated frequencies to remove the outliers.

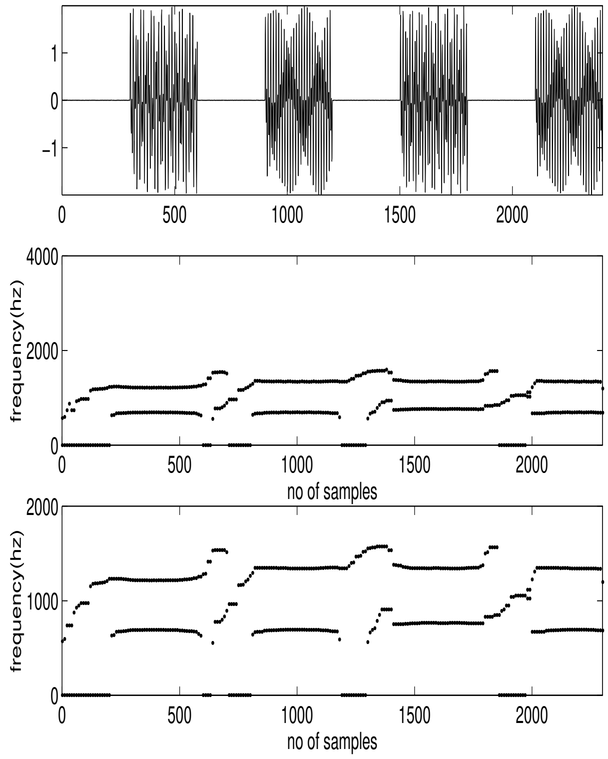

Figure 3: (a) signal (b) frequencies in DTMF range and (c) median smoothed plot of (b)

3. If the valid DTMF frequency pair is detected, the power levels of both high and low DTMF frequencies present are computed and compared with the standard values and if they lie in the specified range, the presence of DTMF tones is reported along with their values.

4 Experimental results

Fig. 3 shows (a) signal containing DTMF tones (b) the frequencies in the DTMF range obtained using our method and (c) median smoothed plot in b.
This plot shows the consistency in the frequencies when DTMF tones are present.
Table 1 shows the results for two digits '1' and '5' for the three cases of
1.Clean signal
2. Signal corrupted by noise(random) for different signal-to-noise ratios (SNR)
3. Signal corrupted by speech for different signal-to-speech ratios (SSR)
Digits 1 and 5 are chosen as they appear next to each other and their frequencies are nearer. This choice helps in analyzing the effectiveness of the method. Fig.4 shows a plot of DTMF frequency estimation error with varying model order

Robustness to noise:
The effect of noise is to flatten the spectrum. This can be compensated by choosing a higher order of LP

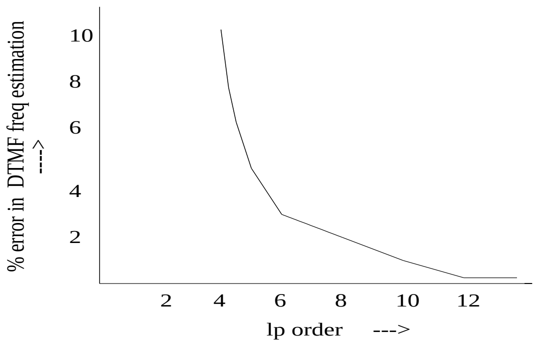

Figure 4: DTMF frequency estimation error for varying model order

model. The model order of 12 chosen for the implementation gives good performance for signal to noise (SNR) ratio of up to 0 DB.

Robustness to speech

Even though the DTMF frequencies are chosen and specifications are designed such that non-DTMF signals are rejected to the maximum extent, a short segment of speech may pass some of the validation tests. Our method has shown good performance even in presence of speech. The reason for it's effectiveness could be following:

- Even though the speech signal might contain one of the DTMF frequencies it is unlikely that any two of the valid DTMF frequencies will be present simultaneously. One of the reasons behind smoothing is to remove such spurious frequencies.
- Even if one of the frequency components of the speech signal lies in DTMF band, it is unlikely that the frequency will sustain with the required accuracy for the whole TD_{min} duration.

Tables 1 and 2 show the performance of our approach for different signal-to-noise (SNR) ratios and signal-to-speech (SSR) ratios. The results show that the detection performance strictly passes all the specifications for SNR and SSR varying up to 0 dB.

Tone	Freq (Hz)	Clean	SNR (20dB)	SNR (10dB)	SNR (0dB)
1	LF	693	691	686	692
	HF	1213	1213	1220	1215
5	LF	752	755	751	758
	HF	1344	1340	1344	1348

Table 1

Tone	Freq (Hz)	Clean	SSR (20dB)	SSR (10dB)	SSR (0dB)
1	LF	693	693	688	675
	HF	1213	1213	1216	1223
5	LF	752	752	754	771
	HF	1344	1344	134	1349

Table 2

Tables 1 & 2. Results of DTMF tone detection for different noise conditions
Parameter values : Frame size : 100(samples)
Frame shift : 10 (samples)
Offset range : $\pm 2\%$
LP order : 12

References

[1] A. D. Inc., *Digital Signal Processing Applications with ADSP 21xx family.* 1992.

[2] S. M. Kay, *Modern Spectral Estimation.* 1996.

[3] K. R. S. Mohan and B. Ramamurthi, "A DSP-based DTMF detector for universal application," in *National Conference on Communications*, (Kharagpur), pp. 192–198, Jan. 1999.

[4] A. Oppenhium and R. Schafer, *Digital Signal Processing.* New Delhi: Prientice Hall, 1991.

[5] Y. Chan, J. M. M. Lavoie, and J. B. Plant, "A parameter estimation approach to estimation of frequencies of sinusoids," in *IEEE Transactions on acoustics, speech and signal processing*, pp. 214–219, Apr. 1981.