DISTORTION EFFECTS OF SPEECH CODERS ON DTMF SIGNALS

İsmail Uzun[1] Rifat Edizkan[2]
e-mail: iuzu@ekom.com.tr *e-mail: redizkan@ogu.edu.tr*
[1]*EKOM Telecommunication Technologies, TÜBİTAK-MAM Teknoloji Serbest Bölgesi, 41470, Gebze, Kocaeli.*
[2]*Dept. of Electrical-Electronics Eng., Eskişehir Osmangazi University 26480, Eskişehir.*

Key words: DTMF signals, speech coders, distortion effects

ABSTRACT

In this study, the distortion effects of speech coders on Dual-Tone Multi-Frequency (DTMF) digits are investigated. Speech coders used in Voice over IP employ block processing and predictive techniques, thus causing distortion on DTMF signals. G.729 and iLBC are used in this study to measure distortions on DTMF signals. Three distortions are measured which are power level distortion, frequency deviation distortion and Signal-to-Noise (SNR) of the frequency spectrum.

I. INTRODUCTION

DTMF detection is very important in telephony communication and applications. In Public Switched Telephony Network (PSTN) signaling, data is transmitted in log-PCM format (either a-law or u-law) with 8 kHz sampling rate. The PCM compression is a waveform coding technique with 64Kb/s compression rate and does not affect narrow band signaling detection such as DTMF or line tone detection.

The use of Voice over IP (VoIP) applications is also gaining popularity as one of the mediums of communication through the Internet. Since PSTN systems are used widely throughout the world, support of PSTN signaling for IP networks and VoIP applications is needed, such as media gateways, IVR and voice mail applications. Low bit rate coders are needed for VoIP applications due to bandwidth limitations which use block processing and predictive techniques to increase the compression rate [1]. These coders decrease the bandwidth requirement for voice traffic, but cause distortion on the DTMF signal waveform. This distortion affects DTMF detection performance and degrades the system quality, a situation which may not be acceptable for telephony applications.

To overcome DTMF detection problems in VoIP, DTMF signals are transmitted through the out-of-band transmission technique, which is specified by RFC2833 [2]. In this protocol, the RTP payload defining the name of the DTMF signal is carried instead of regular audio packets. However, the out-of-band transmission of DTMF signals causes some synchronization problems with speech packets in transmit direction, as well as experiencing detection problems on the receiver side, if high packet losses occur during an active session.

In this study, we analyze the distortion effects of two low-bit-rate speech coders on DTMF signals: G.729 (Annex A) [3] and Internet Low Bit Rate Codec (iLBC) [4], [5] which are widely used in VoIP.

The performance of G.729 speech for narrow-band signals was investigated in detail [6] based on the measurements specified in the ITU-T Recommendation G.720 [7]. In this study, we have analyzed how low-bit-rate coders, G.729, iLBC20 and iLBC30, affect DTMF signals and compared the results to give practical ideas for the developers in VoIP industry. To investigate distortion in the frequency domain, we considered only the frequency standards specified in ITU-T Q.23 [8] and Q.24 [9]. Duration and harmonic distortions are not considered in our study.

Three distortion criteria are used in our distortion analyzes: average power level distortion, average frequency deviation distortion, and Signal-to-Noise ratio in the frequency domain.

II. DUAL-TONE MULTI-FREQUENCY

DTMF signaling is used in telephony systems, such as telephone signaling, call centers and voice-mail applications [10].

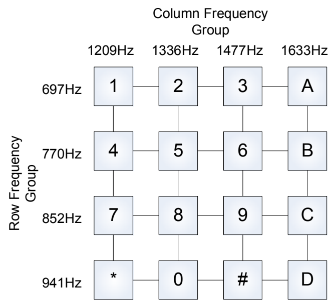

Figure 1. DTMF frequency matrix

DTMF signals consist of two sinusoidal signals: high and low frequency sinusoids. They are described with a 4x4 DTMFs. The DTMFs have two frequency components with each belong to a row and column frequency as shown in Figure 1.

DTMF signals should be generated and detected according to International Telecommunication Union (ITU) recommendations Q.24 [4] and Q,25 [5] and should satisfy the following requirements as shown in Table 1.

Table 1. DTMF requirements

DTMF Duration	Min. 40ms
Pause Duration	Min. 40ms
Signal Interruption	Min. 10ms
Power Level Difference	Max./Min. +4/-8 dB
Frequency Tolerance $\lvert\Delta f\rvert$	Max. +/- %1.5

Speech coders only reshape the frequency spectrum of DTMF signals. They do not change DTMF durations affectingly. Therefore, in this study, we only measured frequency distortion in DTMF signals. DTMF duration problems may occur if there are packet losses in the network and can be compensated easily by increasing the number of DTMF packets sent.

III. SPEECH CODERS

G729 and iLBC are widely used in VoIP. G.729 is an ITU standard speech coder with a compression rate of 8Kb/s [11]. G.729 implements the Conjugate-Structure Algebraic Code Excited Linear Prediction algorithm [6]. G.729A (Annex A) is a simplified version of G.729 and is most frequently used in Internet applications due to its requiring low processing power. [6].

iLBC is a narrow-band speech coder. It works with 20ms and 30ms speech frames, and its compression rate for these frame durations are 13.33Kb/s and 15Kb/s, respectively [4]. It was mainly designed for speech coding in Internet applications. The iLBC gives a good performance in the case of high packet losses and is becoming very popular in VoIP industry.

IV. DISTORTION MEASUREMENT

ITU-T Recommendation G.720 defines the evaluation methodology for the characterization of low-bit-rate speech coders for non-voice signals, including coding delay, phase and amplitude jitter, noise etc. [7]. In this study, our aim is to evaluate coding effects based on three basic frequency dependent measurements and make a comparison between the speech coders commonly used in telecommunication. These three different comparisons may be very helpful for developers in improving DTMF detection.

DTMF detectors detect frequencies within the range defined in ITU standards. Therefore, it is essential to understand how speech coders affect DTMF signals if they are generated within standard limits.

In this study, three different distortion measurements are used:

- Average Power Level Distortion
- Average Frequency Deviation Distortion
- Signal-to-Noise ratio in frequency domain

Spectral power level changes and frequency deviation are very important in DTMF detection. For a DTMF detector, the level and deviation must be within predefined limits. Therefore, these are the most important distortion measurements for comparison of speech coders. SNR in frequency domain is used as a third distortion measurement method which may be useful in defining speech coder distortion characteristics and enhancement techniques in order to improve DTMF detection performance for narrow-band speech coders.

In the experimental study DTMF digits are sampled with 8 kHz sampling rate. To be compliant with VoIP applications, 2560 samples are used for each DTMF digit. Note that VoIP applications send DTMFs with 2500-3000 samples to reduce packet loss errors.

In the experiments, the DTMF signal is encoded and then decoded with the speech coders, G.729A, iLBC-20 and iLBC-30.

In distortion analysis, we compared the reference and distorted signals in the frequency domain. While analyzing the signals, we divided them into 11 frames (11*256=2560) and Hamming windowing is applied to each frame. Following this, each frame is analyzed with FFT.

After FFT analysis, distortion measurement analysis is applied over all the frames for one digit and a calculated mean value is used as a distortion value for each criterion.

AVERAGE SIGNAL POWER LEVEL DISTORTION

Power level changes in the signal spectrum affect the DTMF detection performance. According to the ITU-T standards, frequency components of a DTMF signal should satisfy the following equation:

$$-8dB \leq \frac{Col.Frequency\,Energy}{Row\,Frequency\,Energy} \leq 4dB \quad (1)$$

Speech coders degrade DTMF detection performance if they cause distortion in the signal spectrum. The power level distortion caused by G729A is shown in Figure 2.

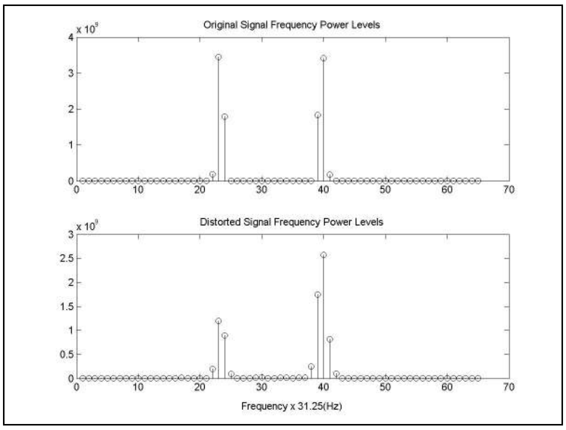

Figure 2. Power level distortion on DTMF digits for G729A speech coder.

As shown in Figure, the power level of row frequency component which has lower frequencies between 697Hz and 941Hz is distorted much more than column frequency component which has higher frequencies between 1209Hz and 1633Hz.

Table 2. Average power level distortion values for DTMF signal

	Coders		
Digits	iLBC-20	iLBC-30	G.729A
1	-0.5870	-0.8228	-1.1116
2	-1.3616	-1.4880	-1.5113
3	-0.4836	-0.1398	-1.5195
A	-0.1350	-0.2082	-0.8948
4	-1.1627	-1.2417	-1.4655
5	-0.4046	-0.4277	-1.4996
6	-0.4696	-0.6344	-1.7674
B	-1.3407	-1.9758	-0.3364
7	-0.0577	0.0283	-1.0693
8	-0.2007	-0.2510	-1.3380
9	-0.0815	-0.1611	-1.3602
C	1.3745	0.5062	0.2914
*	-0.3858	-0.2573	-0.5597
0	-0.2200	-0.1838	-1.0090
#	-0.7632	-1.0637	-1.4243
D	0.4807	0.1025	-0.6218

The average signal power level distortion is defined as:

$$L_{av} = \frac{1}{M}\sum_{j=1}^{M}\left(\frac{X^{j}_{col}}{X^{j}_{row}} - \frac{Y^{j}_{col}}{Y^{j}_{row}}\right) \quad (2)$$

Here, X^{j}_{col}, X^{j}_{row} and Y^{j}_{col}, Y^{j}_{row} denote the spectral power at column and row frequency components of the jth frame of the reference and the distorted signal, respectively. In (Eq. 2), M denotes the number of frames.

The average signal power distortion for each DTMF signal at -10dBm level is shown in Table 2. In the table, negative values reveal that column frequency power is less than row frequency power. As shown in table, G729A speech coder affects power level of DTMF signals more than iLBC coders.

To see the power level distortion at the different signal levels, the signal level is varied to between 0 to -30dBm with a 5dBm increment. Distortions that occurred in the speech coders for the DTMF digit ‘1’ are shown Figure 3.

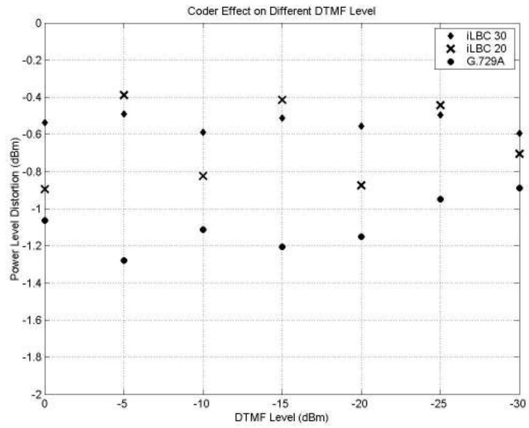

Figure 3. Effects of speech coders on frequency power levels for different levels

It is easy to see that the effect of iLBC coding with 30ms frame size is at about -0.5 dBm for all levels and does not deviate too much. However, the effect of an iLBC coder with a 20ms frame size changes according to signal levels. One important observation is that the G729A coder affects the signal less when the signal level is low. This gives an idea to the developers that G.729A speech coder can be used for lower level signals instead of iLBC to increase the DTMF detection performance during in-band transmission. However, iLBC is stable in relation to signal level changes. This feature could be employed as another decision criterion for the systems that have variable signal levels.

FREQUENCY DEVIATION DISTORTION

Frequency deviation gives the deviation of the measured frequency from the reference frequency in percentage. Frequency deviation is calculated by using the following equation:

$$Deviation(\%) = \frac{(f_x - f_y)}{f_y} x100 \qquad (3)$$

In Eq.(3), f_x and f_y denote the distorted and reference signal frequency, respectively.

Table 3 gives frequency deviation in the reference signal, as well as the signal reconstructed from the selected coders. The frequency deviation in the original signal is actually a result of the spectral analyzes method. In our study, we analyze the signal spectrum with 256-point FFT. The deviations measured in the reference signal are given in the second column of Table 3.

Table 3. Deviation on DTMF signals

	Deviation (%) Row/Column			
Dig.	Original	iLBC 20ms	iLBC 30ms	G.729
1	0.10	0.14	0.08	0.18
	-0.06	-0.08	-0.06	-0.06
2	0.10	-0.06	0.13	0.19
	-0.05	-0.07	-0.08	-0.09
3	0.10	0.12	0.11	0.13
	0.05	0.06	0.07	0.05
A	0.10	0.15	0.05	0.14
	0.04	0.05	0.08	0.13
4	-0.10	-0.10	-0.07	-0.09
	-0.06	0.05	-0.09	-0.08
5	-0.10	-0.06	-0.04	-0.05
	-0.05	0.08	-0.07	-0.07
6	-0.10	0.06	-0.07	-0.08
	0.05	0.08	0.12	0.08
B	-0.09	-0.08	-0.05	-0.13
	0.04	0.04	0.11	0.05
7	0.08	0.11	0.15	0.08
	-0.06	0.03	-0.10	-0.05
8	0.08	0.08	0.11	0.11
	-0.06	-0.06	-0.10	-0.09
9	0.08	0.11	0.14	0.12
	0.05	0.05	0.09	0.06
C	0.08	0.10	0.05	0.06
	0.05	0.05	0.08	0.00
*	0.08	**0.25**	**0.19**	**0.20**
	-0.06	-0.11	-0.08	-0.09
0	0.07	0.11	0.08	0.05
	-0.06	-0.6	0.06	-0.09
#	0.07	0.12	0.12	0.13
	0.05	0.04	0.06	0.13
D	0.07	0.09	0.19	0.13
	0.04	0.03	0.13	0.05

The distortions on frequency deviation for the selected speech coders are very low and there is not much difference between these and those of other coders. The largest deviation is obtained for DTMF '*' that is around 0.2%.

In our study, it is seen that the frequency deviation of row frequencies is lower than that of column frequencies.

SNR IN FREQUENCY DOMAIN

Distortion in the frequency domain can be measured with the signal-to-noise ratio (SNR).

$$SNR(dB) = 10\log_{10}\frac{\sigma_x^2}{\sigma_d^2} \qquad (4)$$

Here, σ_x^2 and σ_d^2 denote standard deviation of the reference and distorted signal spectrum, respectively.

Figure 4 shows spectral components in the range of 0-1875 Hz for the reference and distorted signals. Although the spectrum has been analyzed with 256 point FFT, the figure gives only the first 60 spectral components instead of 128.

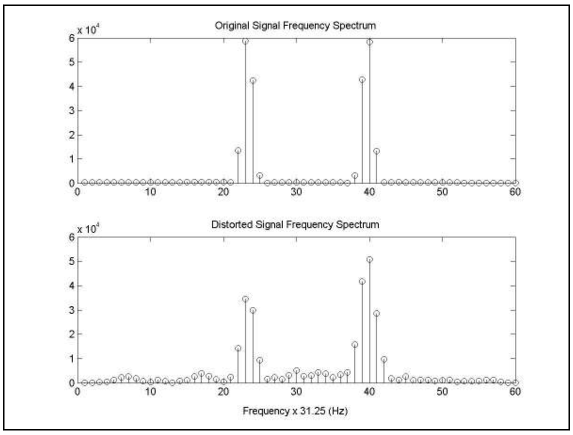

Figure 4. Effects of G729A speech coder on DTMF '1' frequency spectrum

In this study, the distortion measurement of spectral components is obtained from whole frequency band, i.e. from 128(256/2) spectral components.

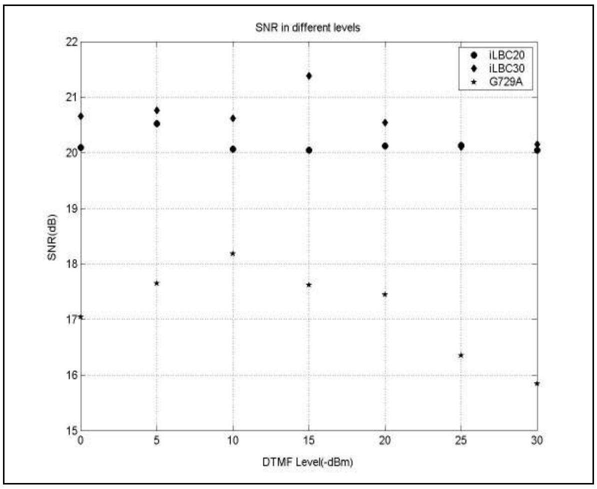

Figure 5. SNR in frequency spectrum for DTMF '1' in different levels

Distortions caused by different coders for the DTMF digit '1' are given in Figure 5. The distortions are obtained for the signal levels between 0 dBm to -30 dBm with -5 dBm step size.

G729A generally distorts the frequency spectrum much less than iLBC. It can also be observed that iLBC with a 20ms frame length distorts the spectrum much less than iLBC with a 30 ms frame length.

V. CONCLUSION

In this study, the affects of the speech coders G.729 and iLBC (with both 20ms and 30ms frame sizes) on DTMF signals are investigated. Three criteria are used in our comparisons: Average Power Level Distortion, Average Frequency Deviation Distortion, and Signal-to-Noise Ratio. All of the measurements are obtained from the frequency spectrum of DTMF signals.

The results of distortion measurements show that DTMF signals are less affected by iLBC. It is advisable to use iLBC with 30 ms frame length in DTMF encoding. For G.729A, low signal levels are preferable for transmitting the DTMF signals.

REFERENCES

1. K. Sayood, Introduction to Data Compression, Morgan Kaufmann Publishers Inc., Second Edition, 2000.
2. RFC2833 - RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals, H. Schulzrinne, May 2000.
3. R. Salami, C. Laflamme, B. Bessette, J.P. Adoul, "ITU-T G.729 Annex A: Reduced complexity 8 kbits/s CS-ACELP codec for digital simultaneous voice and data", IEEE Communications Magazine, vol. 35, pp. 56-63, September 1997.
4. RFC3951 - Internet Low Bit Rate Codec (iLBC). S. Andersen, A. Duric, H. Astrom, R. Hagen, W. Kleijn, J. Linden. December 2004.
5. S. Andersen, W. Kleijn, R. Hagen, J. Linden, M. Murthi, and J. Skoglund, "iLBC-a linear predictive coder with robustness to packet losses," in Proc. IEEE Speech Coding Workshop, pp. 23–25, 2002.
6. Ferraz De Campos Neto, S Karapetian W. "Performance of ITU-T G.729 speech codec with 8 kbps CS-ACELP nonvoice narrowband signals", Communications Magazine. IEEE 1997 and 35 (9): 82~91.
7. ITU-T Rec. G.720, "Characterization of Low-Rate Digital Voice Coder Performance with Non-Voice Signals," Geneva, Switzerland, July 1995.
8. ITU-T Recommendation Q.23: Technical Features of Push-Button Telephone Sets.
9. ITU-T Recommendation Q.24: Multi-Frequency Push-Button Signal Reception.
10. 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," IEEE International Conference on Telecommunications, pp, 884-887, April 1996.
11. ITU-T Recommendation. G.729, "Coding of Speech at 8kbits/s using Conjugate Algebraic Code-Excited Linear Pre-diction (CS-ACELP)", 1996.