Call Progress, DTMF Tones, and Tone Detection

HP TeleShare's voice mode firmware has the ability to detect a number of tones used commonly in telephone communications, including DTMF tones and call progress tones like busy, ringback (the ringing sound you hear when you call someone), and dial tone.

DTMF Tones

Dual-tone multifrequency (DTMF) tones are made up of two separate tones, as the name suggests, and can be accurately generated using easily understood principles. The DTMF standard specifies two sets of distinct tones, called row frequencies and column frequencies (see Fig. 1). The row frequencies correspond to the horizontal rows on a standard telephone touchpad. The column frequencies correspond to the vertical columns on the touchpad, plus an additional column to the right of the last touchpad column.

This makes eight separate frequencies, which combine for a total of sixteen DTMF tones (see Fig. 2).

Generation of a DTMF tone is accomplished by creating a sinusoid for each of the two frequencies, row and column, and then adding the results. In a digital implementation, the sinusoids are computed and added on a sample-by-sample basis. HP TeleShare uses a five-coefficient Taylor series approximation for the sinusoid generation. The sinusoid samples are updated and added at 8 kHz, or every 125 microseconds, and the sum of the sinusoid samples is used as the current DTMF sample.

Tone Detection

Tone detection is accomplished through the use of a 512-point fast Fourier transform (FFT), which is implemented in the ADSP2101 C-language run-time library. The FFT, when given a set of samples of an input signal over some time interval, returns the frequency spectrum of the signal during the interval. This can be done in almost real time with a DSP, making it very useful for detecting incoming tones. The following important rules and relationships should be noted concerning sample rate, input points, output points, time, frequency, and the FFT in general:

- The FFT requires complex (real and imaginary) data for input (two arrays).
- The imaginary input array may be filled with zeros if unused.
- The output data is complex (two arrays).
- The frequency spectrum returned covers half of the sampling frequency.
- Only the first half of output data is used, and the other half is a mirror image.
- The output frequency resolution is equal to (sampling rate)/(number of input points).

Using an 8-kHz sampling rate and 512 points causes the FFT to return a spectrum from 0 to 4 kHz, with 512 complex output points. The second 256 output points can be ignored since they are the mirror image of the first 256. The output will have a resolution of 15.625 Hz per point, using the formula above. These output points will be referred to as bins since they include spectral data on either side of each point.

HP TeleShare calculates magnitude-squared values for each bin by squaring the real and imaginary values at each point and adding them. The magnitude-squared

	Column Frequencies			
	1209 Hz	1336 Hz	1477 Hz	1633 Hz
697 Hz	1	2	3	A
770 Hz	4	5	6	B
852 Hz	7	8	9	C
941 Hz	*	0	#	D

Fig. 1. Dual-tone-multifrequency digits and the frequencies associated with them.

	Tone	Frequency (Hz)	512-Point FFT Index
Call Progress Tones	Dial Tone	350 + 440	22.4 + 28.16
	Busy	480 + 620	30.72 + 39.68
	Ringback	440 + 480	28.16 + 30.72
16 DTMF Tones	DTMF 1	697 + 1209	44.61 + 77.38
	DTMF 2	697 + 1336	44.61 + 85.5
	DTMF 3	697 + 1477	44.61 + 94.53
	DTMF 4	770 + 1209	49.28 + 77.38
	DTMF 5	770 + 1336	49.28 + 85.5
	DTMF 6	770 + 1477	49.28 + 94.53
	DTMF 7	852 + 1209	54.53 + 77.38
	DTMF 8	852 + 1336	54.53 + 85.5
	DTMF 9	852 + 1477	54.53 + 94.53
	DTMF *	941 + 1209	60.22 + 77.38
	DTMF 0	941 + 1336	60.22 + 85.5
	DTMF #	941 + 1477	60.22 + 94.53
	DTMF A	697 + 1633	44.61 + 104.51
	DTMF B	770 + 1633	49.28 + 104.51
	DTMF C	852 + 1633	54.53 + 104.51
	DTMF D	941 + 1633	60.22 + 104.51

Fig. 2. Call progress and DTMF tones recognized by HP TeleShare.

values correspond roughly to the power of the signal in each bin. Once the powers are known for each bin in the spectrum, they can be analyzed to see if any DTMF or call progress tones are present.

As an example, suppose the telephone has been taken off-hook in preparation for dialing and HP TeleShare is configured to check for dial tone. 512 samples of the input signal would be stored in the real input array, while the imaginary array is filled with zeros. Next, the FFT function is called, returning the real and imaginary arrays. The magnitude-squared values of the first 256 bins are computed using the two output arrays. The two frequencies that make up a dial tone are 350 and 440 Hz (see Fig. 2), so FFT indexes (or bin numbers) must be computed for these frequencies:

$$350/15.625 = 22.4 \qquad 440/15.625 = 28.16$$

An effective method of checking for the existence of a particular frequency is to compare the power present at that frequency with the total power of the spectrum. This is done quite easily with magnitude-squared values since they represent power in each bin already. Total power is simply the sum of all the magnitude-squared values for the first 256 FFT return values. Divide this into the power of the frequency being checked for, and the result is the percentage of total power for that frequency. For example, when checking for 350 Hz, compute the sum of the power values for bins 22 and 23 since the real index (22.4) falls between them, and then divide by the total power. The result is the percentage of the total power present around 350 Hz. The same can be done for 440 Hz, using bins 28 and 29.

Once the percentage of total power is calculated, a comparison can be made to see if the power in each frequency meets *match criteria*. The HP TeleShare firmware typically uses 35% of total power as a match condition. In other words, if the power present at the desired frequencies is 35% or more of the total power, dial tone has been detected. Otherwise, no dial tone is found.

The number of bins used in the comparison and the match criteria can be fine-tuned for a particular application. The match criteria can include other tests and can be relaxed or tightened as needed. The number of bins used can be influenced by the total number of points in the FFT and by a preprocessing tool that does windowing. Windowing is used to create a finite-length sequence from a continuous sequence. It is basically a digital filter that truncates an infinite-length input sequence while preserving its frequency characteristics. Since we are grabbing finite pieces (sequences) of data, we need to window the data.