Design of a smartphone with a Digital Signal Processor

Lecluse, Joep

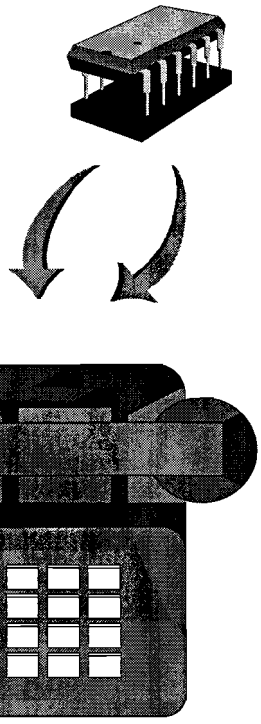
Award date:
1996

Eindhoven University of Technology

Faculty of Electrical Engineering
Section of Digital Information Systems

Master's Thesis:

Design of a Smartphone with a Digital Signal Processor



Author: Joep Lecluse
Id. No.: 268628

Coaches : K.J. Delst / drs.H. Braams
Supervisor : Prof.ir.M.P.J. Stevens
Period : August '95 - April '96

Summary

At the company Pijnenburg Electronic Products one is interested whether the application of a Digital Signal Processor (DSP) in a telephone called Avenue is meaningful and possible. The DSP should be able to handle answering machine functions, modem functions for the consultation of a telephone directory, and a full-duplex handsfree speakerphone function.

The Avenue is an advanced screenphone with a lot of special functions, like a telephone index, text telephony, cost counting and handsfree calling.

Because the telephone works with low frequency components, a Signal Processor should be capable of handling these functions.

First I examined the complexity of the different functions. These functions are DTMF decoding, voice compression and decompression and echo cancellation.

Next I searched for General Purpose DSPs, their specifications and their prices. During this search, I discovered another kind of DSP called a Digital Tapeless Answering Device (DTAD). This device contains a DSP and already implemented Digital Signal Processing algorithms. You only have to send commands like "record message" to it, using a microcontroller. A number of companies are developing DTADs and they are becoming more powerful with a lot of special functions like voice recognition, etc.

In order to obtain more information about DSPs, I visited a company called CME in Veenendaal, which helps companies obtaining information about Digital Signal Processing. They introduced the third possibility to me, the creation of an ASIC. With the aid of their computers and software, one is able to make his own Custom Digital Signal Processor.

But, the question still remained, where to get the software? Creating our own software would take a long development time. Therefore I searched for companies who offer *of-the-shelf* processing algorithms concerning our application. The companies I came up with offer some good solutions, but are expensive.

The conclusion for this moment, considering the prices for General Purpose DSPs and the third party software is that we should best use a DTAD from the DSP-Group, or the MSP58C85 from Texas Instruments. They are relatively cheap and there are possibilities for the implementation of Pijnenburg's CAS-algorithm within the DTAD.

Table Of Contents

1. Introduction

At the company Pijnenburg Electronic Products one is interested whether the application of a Digital Signal Processor (DSP) in two telecommunication projects is meaningful and possible. The application concerns a telephone with a number of functions which at present are being realized with separate standard components. For a new required function: a narrowband dual-tone detection, a new chip is being developed. At a second generation of the product, the functions could possibly be implemented more efficiently together in a single DSP. The DSP could, if possible, be included as a core in an ASIC.

This report gives an overview of the functions of the smartphone that need to be implemented on the DSP. Three different approaches are described to solve the DSP implementation problem, i.e. general purpose DSPs, Digital Tapeless Answering Devices (DTAD), and ASICs. Some useable general purpose DSPs for this project are listed and a description of the companies who are developing software for the digital telephony solutions are given. Also some DTADs will be highlighted. In the chapter ASIC a description is given about the pros and cons of ASICs and a company called CME will be discussed. This company renders assistance by the creation of ASICs.

2. Problem definition

During my final project I should examine whether the application of a Digital Signal Processor in an intelligent telephone (smartphone) is meaningful and possible. This telephone should in the first instance have at it's disposal the following functions:

- Answering machine with digital recording of incoming and outgoing messages.
- Full duplex hands-free speakerphone
- 2400/1200 Baud modem for the consultation of a telephone-directory.

In a later version it should be possible to add functions like **C**ustomer Premises Equipment **A**lerting **S**ignal (CAS) detection and **C**alling **L**ine **I**dentification (CLI).

The DSP should be applied in the product for:

- DTMF detection on the basis of filtering
- Acoustic echo-cancellation for a hands-free full duplex speaker phone operation
- Compression en de-compression of speech signals
- Realization of a hybrid for electric echo-cancellation
- Tone-detection for the implementation of a modem
- Call Progress Monitoring

The final project consists of the following tasks:

1. A systematic analyses of the required DSP properties and the necessary peripherals.
2. Specification and structure description of the DSP software
3. Examination into the usefulness of the existing development tools for DSPs in relation to this application
4. Construction of a functional prototype or an evaluation system on the basis of a standard DSP-structure.

3. Digital Signal Processing

With Digital Signal Processing the intended result is to process signals in such a way that application useful information can be separated from non-relevant information. Depending on the application, this information can be processed in different ways. Examples of signal processing are filtering and reconstruction of noise distorted signals, the improvement of the contrast in pictures, the recognition of handwritten characters, the generation of sound, or the compression of audio signals for efficient storage.

An essential characteristic of digital signal processing is that the signals to be processed are sampled in amplitude as well as in time. In practice this means that signal first have to be converted from the continuous, analog domain to the discrete, digital domain (A/D-conversion, sampling) and back (D/A-conversion).

Figure 1 shows a simple system configuration. Both A/D and D/A conversion is not always needed. In the case of signal generation for instance, you don't have to perform the A/D part and when signal detection is performed only A/D in needed.

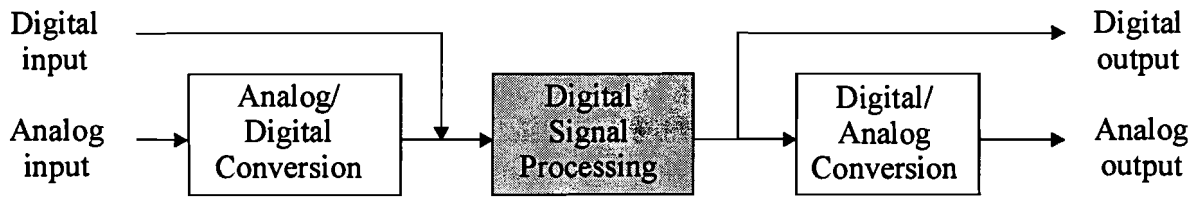


Figure 1: Simple system configuration

Which algorithms are being applied? The functionality of a signal processing system is determined by the algorithms it executes. Over the past ten years a lot of algorithms were developed. Here is a list of often used algorithms.

Finite Impulse Response filtering (FIR)

The output is a linear combination of a finite series of samples of the input signal. The number of previous samples stored in the filter determines the order of the filter. The filtering takes place in time-domain. An important property of FIR filters is that they are unconditionally stable.

Infinite Impulse Response filtering (IIR)

In contrast with FIR filtering the output of a IIR filter is defined by the samples of both the input and the output signal. An IIR filter has the advantage that is has a lower order than a comparable FIR filter, and therefore needs less memory elements. The main disadvantage is that this type of filter is not unconditionally stable.

Fast Fourier Transformation (FFT)

The FFT transforms a series of samples from the time-domain to the frequency-domain. It is an example of a block operation. The FFT gives information about the amplitude and phase of the frequency components contained in the signal. For the transformation from the frequency- to the time-domain, **Inverse Fast Fourier Transform** (IFFT) is used.

Auto- and Cross Correlation

Correlation is used to give a relation between two series of samples of the same signal (auto-correlation) or between a series of samples of two different signals (cross-correlation). Auto-correlation can be used to separate the periodical components from the stochastic noisy components. Cross-correlation can be used to determine the phase-shift of two signals.

Discrete Fourier Transformation (DFT)

The DFT can be compared to the FFT. An important difference is the order of the number of multiplication's and additions. The FFT is used when one processor is present, the DFT is used in parallel implementations.

4. Digital Signal Processors

Digital Signal Processors, or DSPs are integrated circuits specialized to accept and process a range of real-time signals like audio and speech communication, medical instrument sensors, and compact-disc data and video. They can be considered as micro-processors which are optimized for the fast processing of digital signal algorithms. Most of these algorithms are characterized by the large number of additions and multiplication's. Besides that, these operations must be repeated many times in a loop. The reason why a DSP can process these operations much faster than ordinary micro-processors is that they are equipped with a hardware-multiplier. One multiplication can be performed in one single clock-cycle. Some DSPs have a MAC (multiply/add) unit which increases the performance because they can simultaneously perform an addition and a multiplication. The special-purpose design of the DSP is the key to its effectiveness. DSPs available today share five important characteristics:

- Multiple bus architecture, which permits a large number of simultaneous inputs and outputs.
- Extensive pipelining to execute parts of several instructions in a single cycle.
- Dedicated hardware multipliers capable of true multiplication rather than consecutive additions.
- Special instructions that combine several operations - multiplication, addition and memory read/write - into a single cycle.
- Fast instruction cycles, currently 10 to 20 nanoseconds and likely to drop.

What do these characteristics mean to system designers and to end users? For designers, they mean more performance, greater flexibility, better power management and lower cost. For end users, they mean added features and system response that look and feel like the world they know.

With software, a DSP can be programmed to be a high speed modem, a high fidelity sound and music synthesizer or a telephone answering machine. Not only can a DSP perform a variety of different functions, it can do so concurrently. The operating system enables several independent real-time tasks to operate simultaneously with no loss in performance and no intervention by the host CPU.

The programmability feature of DSP has several distinct advantages:

1. **Upgradeability.** Each time a new standard is released for any advance in technology the DSP can be upgraded by simply loading the new software. DSP-based techniques protects hardware investment against obsolescence.

2. **Reliability.** DSPs increase communication reliability. Communications handshaking is more successful, and discrimination of incoming calls based on ring, fax tones, caller ID, etc. is more accurate.

In the 1970s, digital signal processing first began to emerge as a powerful solution to the problem of analyzing real-world phenomena. At the same time, discrete implementations of DSP solutions were constructed of linear-circuit analog front-ends, A/Ds and other peripheral interface circuits, large blocks of random glue logic, PALS and PROMS, memory blocks, and a processing element of some sort (a minicomputer or an attached processor). In short, 1970s DSP solutions represented an aggregate of all the discrete hard- and software technology required to input, condition, process, and perhaps output real-world signals.

In the 1980, as DSP technology was used to solve an increasing number of applications, performance demands and cost pressures brought the advent of the DSP. The DSP was ideal for handling the kinds of high-speed I/O and math intensive processing demands associated with signal processing applications. As is true with all computer chips, higher volume production and learning curves have lowered their cost over time. The very same chip carrying a $500 price tag in 1982 now costs $5 in single unit quantities and as little as $3 in high volume. It is anticipated DSPs will have as great an effect on the future of products as the microprocessor computer chip has in recent years.

By the early 1990s, the digital signal processor had evolved considerably and today, the DSP is considered to be the main driver behind most of the world's emerging technologies, including telecommunications and multimedia. As DSP costs continue to drop, in real dollars and in $/MIPS, focus has begun to shift back to the big picture where the processor is but one element in a total digital signal processing solution.

What are the options? Today's typical home office consists of a number of individual pieces of software and hardware: A fax/modem package takes care of faxing and sending data. An answering machine takes messages. The telephone provides standard features such as speed-dial, a speakerphone, and other programming functions. Sound cards provide voice annotation, special effects, and enhanced sound. What these individual packages won't do, however, is manage all these communication activities and allow them to interact with one another. The alternative is a DSP multifunction communications board.

5. Description of the smartphone (Avenue)

The Avenue screenphone is an advanced telephone device with a lot of special functions. The build in screen always shows the action currently being executed and displays what you should do to achieve the desired result. Figure 2 shows a photograph of the Avenue.

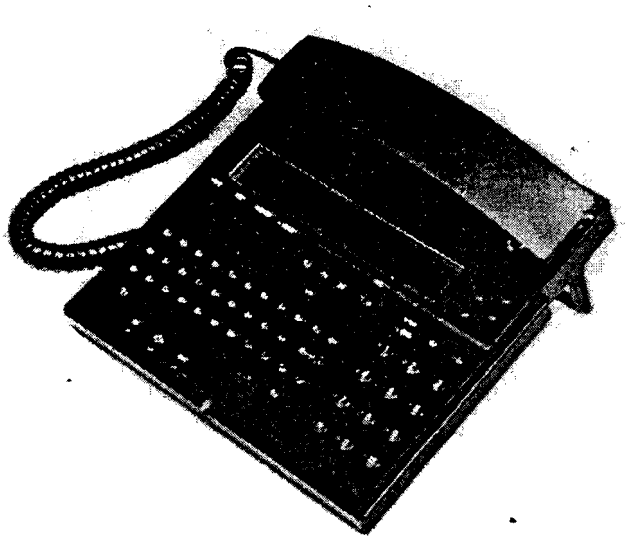


Figure 2: External view of the Avenue

The Avenue has the following special functions:

- Telephone index

The telephone index has the possibility to store more than 300 telephone numbers. Each card in the telephone-index can contain a name, address, zipcode, city and telephone number. The Avenue automatically orders them alphabetically.

- Electronic telephone directory

This function enables the automatic search for telephone numbers. The Avenue makes contact with an electronic databank and downloads the telephone number, address, zipcode and city of the person you want to call. This data can be copied and stored in the telephone index of the Avenue. The connection is made with a 1200 baud modem.

- Log-book

The log-book function in the Avenue - when switched on - registers the number, time and date of all incoming and outgoing calls. In the log-book the 50 last calls are being registered. When the log-book becomes full, the first call will be written over the last one.

- Cost counting

The Avenue has a build-in cost counter. With this function you can easily keep track of the costs of your telephone conversations. This is very useful when the telephone is being used by more than one person or when you want to register the costs for business reasons. Also a the time of a conversation will be displayed.

- Text-telephone

With the Avenue written messages can be sent, either to another Avenue or a text-telephone. This function is needed when you want to communicate with deaf or hard-of-hearing persons. The characters are transmitted as combinations of DTMF signals.

- *-service

The Avenue is prepared to use a number of (new) PTT Telecom services. These services are:

- * 21 : This is a switch-through function. It allows you to redirect all incoming calls to another telephone.
- Calling Line Identification: This feature will be introduced in 1996. It shows the telephone number of the person who is calling while the phone is ringing. With the Avenue it is possible to block this function in order to call anonymous.
- Switching Conversation: This feature will also be introduced in 1996 and allows you to handle two telephone conversations at once. When you are in conference, you will hear a tone as a sign that a third person tries to reach you. By pressing a key, you can switch between the persons, putting the present person on hold.

- Handsfree calling

With this function you can use the telephone without having to lift the horn of the hook. The current version of the Avenue has a half-duplex handsfree function which means that only one person can speak at once. The telephone decides which party is talking and switches this signal to the line. In a newer version, this function will be replaced by a full-duplex handsfree function allowing both parties to talk at the same time. It is a complex function, but the introduction of a DSP in the Avenue can give a solution.

- Listen-in function

It is also possible to listen to a conversation through the internal loudspeaker without also transmitting the sound in the room. This function can be used to allow other people to just listen to a conversation.

6. Overview of the functions and their specifications

As described in chapter 2, the smartphone should at first be able to handle the following functions:

- Answering machine
- Full duplex hands-free speakerphone

These functions contain certain sub-functions which could either be implemented analog with standard components or digitally with the aid of a DSP. These sub-functions are:

- DTMF detection/generation
- Voice compression/decompression
- Echo cancelling (acoustic an electric)

The following sections give an overview of these sub-functions.

6.1 DTMF

There are two standard dialing conventions used in telephone systems throughout the world. The most common, and by far the oldest is known as pulse or loop-disconnect dialing. Dual-Tone Multi-frequency (DTMF) is a relatively new all-electronic method which is rapidly replacing the older electromechanical system.

The use of the DTMF signaling scheme within telecommunications systems has become widespread over the past few years. It is replacing the older type of pulse oriented dialing methods in telephones worldwide, and also finds application in a number of other equipment types, such as personal computer (PC) telephone peripherals, remote signaling schemes etc. The DTMF system has been adopted as the universal standard through the CCITT (Comite Consultatif International de Telephonie et de Telegraphie) which is a committee of the International Telecommunication Union (ITU), now part of the United Nations.

DTMF signaling consists in representing the dialing information in a form of a superposition of two sinusoidal waveforms with frequencies chosen from a set of eight standardized frequencies. These signals have to be recognized in the exchange in order to determine the number dialed by a caller. Also a remote telephone can be controlled by sending DTMF signals, e.g. to listen to recorded messages on an answering machine. This task is realized with DTMF receivers. DTMF receivers can be designed using analog or digital filters. Figure 3 gives an illustration of the DTMF frequencies. Pressing any key causes an electronic circuit

to generate a tone which is a summation of the two individual frequencies related to the row and column of that key.

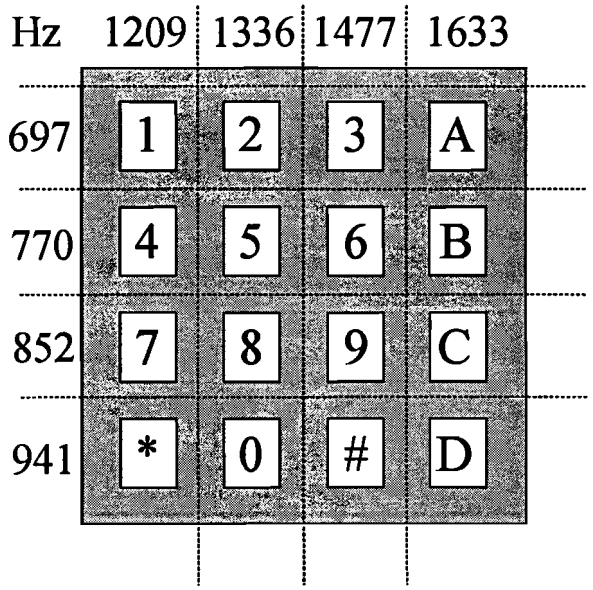


Figure 3: Illustration of DTMF coding

The frequencies used in DTMF dialing have been carefully selected so that any DTMF decoding circuit will not confuse them with other tones that may occur on the line. As the tone generation does not involve a disconnect of the telephone circuit, DTMF tones may be sent down the line during a call just by pressing any key on the keypad. When this method is used as a form of low speed data transmission, it is important that speech is not accidently interpreted as a DTMF tone. In order to reduce the risk of this happening, tones must be present continuously for a minimum period , With an interdigit period of similar length. Parameters guaranteed by the DTMF sender and receiver are specified in CCITT Q.24 Recommendation. These requirements are as follows:

- frequencies of receiving signals
  - low frequency group: 697, 770, 852, 941 Hz,
  - high frequency group: 1209, 1336, 1477, 1633 Hz,
- frequency tolerance: ± 1.8 %,
- level of receiving signals for which the receiver works correctly: 0..-30 dBm,
- twist: 5 dB,
- time parameters:
  - minimum duration of signal to be recognized: 60 ms,
  - maximum duration of signal not to be recognized: 30 ms,
  - break between signals of two consecutive digits at least: 60 ms,
  - maximum break in the signal not to be recognized: 10 ms,
  - maximum speed of signaling: 120 ms per digit.

6.1.1 Digital reception of DTMF signals

The whole set of samples to be analyzed in the digital DTMF receiver can be divided in blocks of N samples. Additionally to the analyses of the signal spectrum, the duration of respective signals and separating breaks must be tested. Signal detection can be interpreted as filtering with a nonuniform filter bank. This filter can be based on a proper discrete Fourier transformation (DFT) algorithm.

Proper selection of the number N is very important. If the set of N analyzed samples corresponds to a relatively long sample acquisition time T, the correct analysis of duration of the signal (or break) would not be possible. If this time were selected much shorter, DTMF signals of duration equal to or longer than 40 ms would be correctly recognized. That fact implies the time T to be as short as possible what means the small number N. An other reason to make the number N as small as possible is the is the reduction of processor computing time, what allows to process many other functions. Decreasing the number of samples decreases, however, signal-to-noise (S/N) ratio. Another disadvantageous effect is the degradation of resolution of the receiver and increasing of the error of the center frequency of the filter passbands. Detection of the DTMF signal might even be impossible if the number N were too small because attenuation characteristics which should be different for all filters would become to be the same (for example, if N = 10 then for the sampling rate f_s = 8 kHz all filters for the low frequency group have the same center frequency equal to 800 Hz). As one can see, the number N should be optimized making a compromise between all these conflicting criteria. The time T should be no longer than ≈ 20 ms because of duration of the signal or break to be recognized. It gives about 160 samples for sampling rate f_s = 8 kHz. However, one can notice that all DTMF frequencies are below 2 kHz, what allows, according to the Nyquist theorem, to decrease the sampling rate to 4 kHz and, thus, the number of block samples to N = 80.

6.1.2 Description of the receiver program

A DTMF detection program should take N samples making simultaneously the computation of the DFT, and then compute the energies of the eight DTMF frequencies and their second harmonics. The second harmonics allow to discriminate between DTMF tones and speech or other disturbing signals which can contain DTMF frequencies. In the next step, the two frequencies with the biggest energies should be searched and checked if they are above the threshold (-30dBm).

6.1.3 Tone generation method

DTMF tones can be obtained by adding two different waves generated by a tone generator. So the problem to be solved is how to construct a tone generator algorithm. There a several solutions to this problem. One is to store some values of a sinusoidal waveform in a memory. The signal can then be generated by reading out this memory at a specific speed. This algorithm is very simple but it uses memory space. For each of the 8 different sine-waves you need a different table. It can be reduced by just storing a quarter of the signal.

Another way to generate a sinusoidal waveform is more complex, but uses less memory space. The algorithm is derived from the following two goniometric formulas:

- $\sin(x + \Delta) = \sin(x) * \cos(\Delta) + \sin(\Delta) * \cos(x)$
- $\cos(x + \Delta) = \cos(x) * \cos(\Delta) - \sin(x) * \sin(\Delta)$

When building the wave you must consider a constant step size Δ. This means that $\cos(\Delta)$ and $\sin(\Delta)$ are also constants. Lets call these constants:

- $a := \sin(\Delta)$
- $b := \cos(\Delta)$

The next point to calculate is $\sin(x + 2\Delta)$. This is equal to $\sin((x + \Delta) + \Delta)$. You can see that each new point is a function of the constants and the previous calculated points. The function of interest is the $\sin(x + \Delta)$ function, but because this functions contains a $\cos(x)$ term, also the $\cos(x + \Delta)$ needs to be calculated.

The formulas can now be rewritten to:

x[n] := b * x[n-1] + a * y[n-1]
y[n] := -a * x[n-1] + b * y[n-1]

or in matrix form :

$$\begin{bmatrix} x \\ y \end{bmatrix}_n = \begin{bmatrix} b & a \\ -a & b \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}_{n-1}$$

with $x[n] = \sin(x + n\Delta)$ and $y[n] = \cos(x + n\Delta)$.

For the stepsize Δ the following holds:

$$\frac{\Delta}{2\pi} = \frac{f_{generate}}{f_{sample}} \Leftrightarrow \Delta = 2\pi . \frac{f_{generate}}{f_{sample}}$$

f_{sample} = 8 kHz and $f_{generate}$ is the frequency you want to generate.

In Appendix A the C-listing of the algorithm to generate DTMF-tones is listed, given the two frequencies. The program is tested and the DTMF-tones were recognized by the telephone.

6.2 Compression

The technique which enables the digital implementation of an answering machine is the DSP implementation of compression-algorithms. Voice coding is used to compress voice information for efficient storage. The Human ear can detect sound from a minimum of 5 Hz to a maximum of 20 kHz. The majority of audio information is located in the lower half of the 20 kHz band. Human speech rolls of around 7 kHz, and telephone lines provide about 4 kHz of bandwidth. Many voice coding software products focus on telephone bandwidth audio compression.

The function of a speech codec is to convert an analogue speech signal into a digital form, and to perform the complementary function of converting a digital signal back to analogue. The speech codec is an important part. It defines the basic speech quality for the whole system, since no amount of sophistication elsewhere in the system can compensate for degradation introduced in the codec. The users perception of the system is strongly influenced by the quality of this component.

A telephone signal has a bandwidth of 3.4 kHz, and requires 96 kbit/s assuming 8 kHz sampling and 12-bit linear quantization. It is possible, however, to reduce the bit rate of a telephone signal to 64 kbit/s by replacing the 12-bit linear quantization by 8-bit non-linear quantization (using either an A-law or μ-law non-linear quantization scheme). Since speech is structured, its inherent redundancy can be exploited to reduce this rate still further. This can be done with the aid of a DSP. Advances in speech coding techniques have lead to greatly improved speech quality at low bit rates.

Speech coding algorithms can be classified into the following three types:

- waveform coding
- vocoding
- hybrid coding

The most basic waveform coders do not attempt to exploit any knowledge of the speech production process in the encoding of the input signal. The aim of waveform coders, as the name implies, is to reduce the original waveform as accurately as possible. As these coders are not speech specific they can cater for many non-speech signals, background noise and multiple speakers without difficulty. The penalty of a relatively high bit rate, however, must be paid for this 'acoustic robustness'.

In contrast, vocoders (voice + coders) make no attempt to reproduce the original waveform but instead derive a set of parameters at the encoder which can be used to control a speech production model at the decoder. The parameter set for the speech production model is relatively small and they can be efficiently quantized for storage; hence vocoders operate at very low bit rates.

Hybrid coders combine features from both waveform coders and vocoders to form sophisticated coding schemes which provide good quality, efficient speech coding. In common with vocoders, hybrid coders make use of a speech production model. However, unlike the very simple representation of the excitation signal used by vocoders, a refined representation of the excitation is employed. Hybrid coders operate at a medium bit rate somewhere between vocoders and waveform coders.

The following sections give a review of the different speech coding techniques. The last section describes a comparison of the speech quality and complexity of different speech coding techniques.

6.2.1 Waveform coding

The simplest and best known waveform encoding technique is pulse code modulation (PCM). PCM is a method of digitizing or quantizing an analog waveform. As in any analog-to-digital (A/D) conversion, the quantization process produces an estimate of the signal sample, possibly introducing an error into the digital representation because of the finite number of bits available to represent the value. In theory, this error can be made insignificant by representing the estimate with a large number of bits (high precision). In practice, however, there must be tradeoff between the amount of error and the size of the data representation. The goal is to quantize the data in the smallest number of bits that results in a tolerable error. In the case of speech signals, a linear quantization with 12 or 13 bits is the minimum required to produce a digital representation of the full range of speech signals accurately.

The number of bits required is reduced to eight in the CCITT recommendation **G.711** by exploiting a nonlinear characteristic of human hearing. The human ear is more sensitive to quantization noise in small signals than to noise in large signals. **G.711** applies a non-uniform (logarithmic) quantization function to adjust the data size in proportion to the input signal. Thus, smaller signals are approximated with greater precision. Two quantization functions, or encoding laws, are defined by **G.711**: μ-Law and A-Law. In most cases, the United States and Japan use μ-Law, whereas Europe uses A-Law.

Rather than taking the logarithm of the linear input directly, which can be difficult, μ-Law PCM matches a logarithmic curve with a piecewise linear approximation. Eight straight line segments along the curve produce a close approximation to the logarithmic function. Each of these lines is called a segment. A sample value is represented by its segment and its position within the segment.

The CCITT recommendation provides a μ-Law conversion table. This table has several regular characteristics, however, so that the conversion can be implemented without storing the entire table. The PCM value is in signed-magnitude format, so the conversion table for negative numbers is the same as for positive number except for the sign. In addition, adding 33 to the segment endpoints produces boundaries at even powers of two.

The format of the μ-Law PCM 8-bit word consists of three parts. The most significant bit (MSB) is the sign bit, the next three bits contain the segment number, and the last four bits indicate the position within the segment. All bits of the number are inverted from their actual values to increase the density of 1s, (because speech is typically low-energy) a property that can be used by error-correcting circuitry on transmission lines.

A-Law PCM uses the same approach as μ-Law PCM in approximating the logarithmic curve using eight line segments. In A-Law conversion, however, the segment endpoints are at even powers of two, rather than offset by 33. On output, only the even bits of the encoded number are inverted in A-Law. Otherwise, the conversion is the same. The format of the 8-bit A-Law PCM word is the same as the μ-Law format.

The bit rate required by waveform coders for speech encoding can be reduced by exploiting the correlation between adjacent samples, for example by encoding the difference between successive samples rather than the samples themselves. One such scheme is known as differential pulse code modulation (DPCM). By adapting the quantizer step-size of a DPCM coder according to the short-term speech power, the speech coding technique known as adaptive differential pulse code modulation (ADPCM) is obtained.

ADPCM is an important speech coding technique and as a result several CCITT ADPCM speech coding standards exist. Perhaps the most important CCITT ADPCM standard is Recommendation **G.721**, as this Recommendation also forms the core of other CCITT ADPCM standards. Recommendation **G.721** fully defines the 32 kbit/s fixed rate CCITT speech coding algorithm.

The CCITT Recommendation **G.726** extends the **G.721** ADPCM to include 40, 24 and 16 kbps, as well as 32 kbps. **G.726** at 40 kbps performs comparable to **G.711**.

6.2.2 Vocoding

The most simple model of speech production used by vocoders is illustrated in Figure 4. For voiced speech (such as the vowel sound 'a') the voiced excitation signal is modeled by a train of unipolar, unit amplitude impulses at the required fundamental frequency (the equivalent perceived frequency is known as the pitch frequency). For unvoiced sounds (such as the 'f') the unvoiced excitation is modeled as the output from a pseudo-random noise generator. The voiced/unvoiced switch selects the appropriate excitation and the gain term controls the level of the excitation. The spectral shaping of the excitation signal by the vocal tract is modeled by a time-varying spectral shaper.

Linear predictive coding (LPC) analyses may be used to derive the coefficients of a time varying linear digital filter which models the spectral shaping of the vocal tract. For speech coding the parameters of this filter are updated typically at intervals of between 20 and 30 ms, with 20 ms being the most common update period and 30 ms normally reserved for codecs operating at bit rates of 4.8 kbit/s and below. The aim of LPC analysis is to extract the set of parameters from the speech signal which specifies the filter transfer function giving the best spectral match to the speech being encoded. An all-pole filter p (usually in the range 10 to 16) is used to model the spectral shaping of the vocal tract.

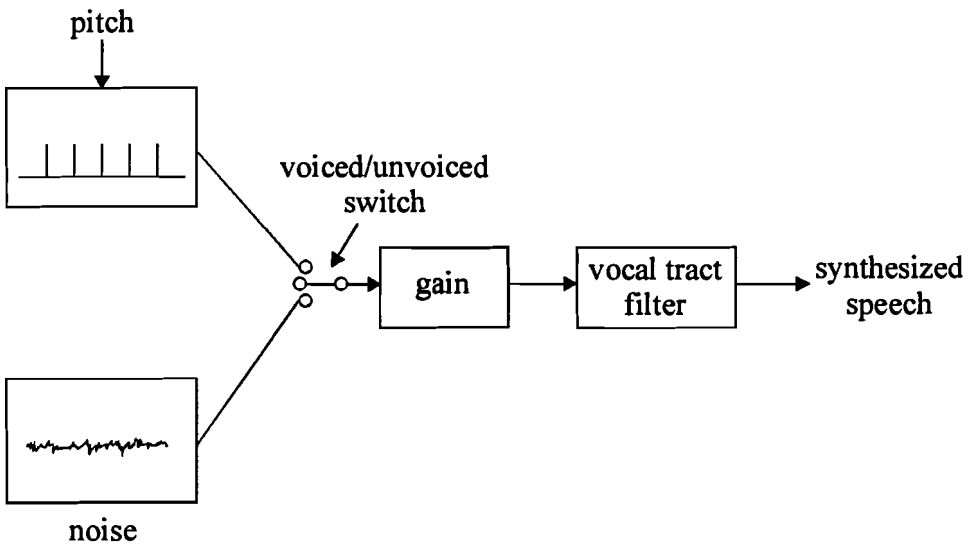


Figure 4: Speech production model.

The excitation is usually modeled in vocoders as shown in Figure 4, i.e. a series of unipolar pulses spaced at the pitch frequency for voiced speech and noise for unvoiced speech. This excitation model is poor for several reasons, including the following:

- speech does not fall neatly into the two categories of voiced and unvoiced;
- pitch is not constant in voiced speech but subject to 'micro-variations' - the constant pitch of the excitation over several voiced excitation periods contributes to the synthetic sound of vocoders.

As vocoders are so strongly based on the simple speech production model of Figure 4, they perform very poor with high levels of background noise, multiple speakers and non-speech signals.

6.2.3 Hybrid coding

Hybrid coders combine features from both waveform coders and vocoders to form sophisticated coding schemes which provide good quality, efficient speech coding. The most common hybrid coders use the LPC synthesizer model as the speech production model. An understanding of the operation of LPC-based hybrid coders can be obtained by considering how the all-zero inverse LPC filter, i.e. the inverse of the all-pole LPC filter, could be used to form a speech coder.

When a speech signal is filtered through the inverse LPC filter the short term correlations of the speech signal are effectively removed leaving a noise-like waveform, known as the residual signal. If the speech is voiced the residual waveform will contain periodic spikes at the pitch frequency and if it is unvoiced the residual waveform has almost no distinguishable structure. If the residual signal is not quantized and is filtered by the unquantized all-pole LPC synthesis filter, the original speech signal is reproduced at the output of the synthesizer (Figure 5). Even if the LPC filter is quantized, provided the quantization is relatively accurate, the reproduced speech is almost indistinguishable from the original. With this in mind, it is clear that provided the residual signal can be quantized accurately and effectively, a high quality practical speech codec can be obtained.

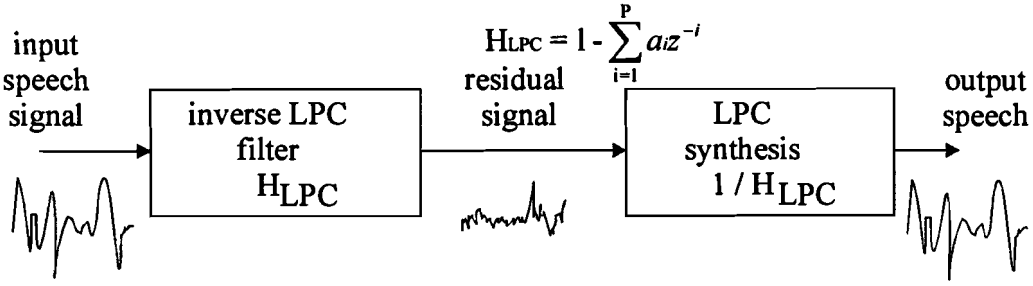


Figure 5: The formation of the residual signal.

The most obvious way to construct a codec based on this principle is to treat the residual as a waveform and quantize it directly - this form of coding is known as adaptive predictive coding (APC). However, as the residual is noise-like there is almost no correlation between adjacent samples, so none of the differential encoding schemes can be effectively employed and a relative high bit rate is required.

For an LPC vocoder the LPC synthesizer model is used only at the decoder. However, for an analysis-by-synthesis LPC-based coder the LPC synthesizer model is employed at both the encoder and decoder. At the encoder the LPC coefficients of the synthesizer model are obtained directly from the input speech signal and the excitation signal is derived using a closed loop analysis-by-synthesis technique (Figure 6). Analysis-by-synthesis is the use of synthesis as an integral part of the analysis process. The objective of the encoder analysis-by-synthesis process is to derive an excitation signal, such that the difference between the input and synthesized signals in minimized according to some suitable criterion. The LPC coefficients and the closed-loop derived excitation signal are stored as quantized values.

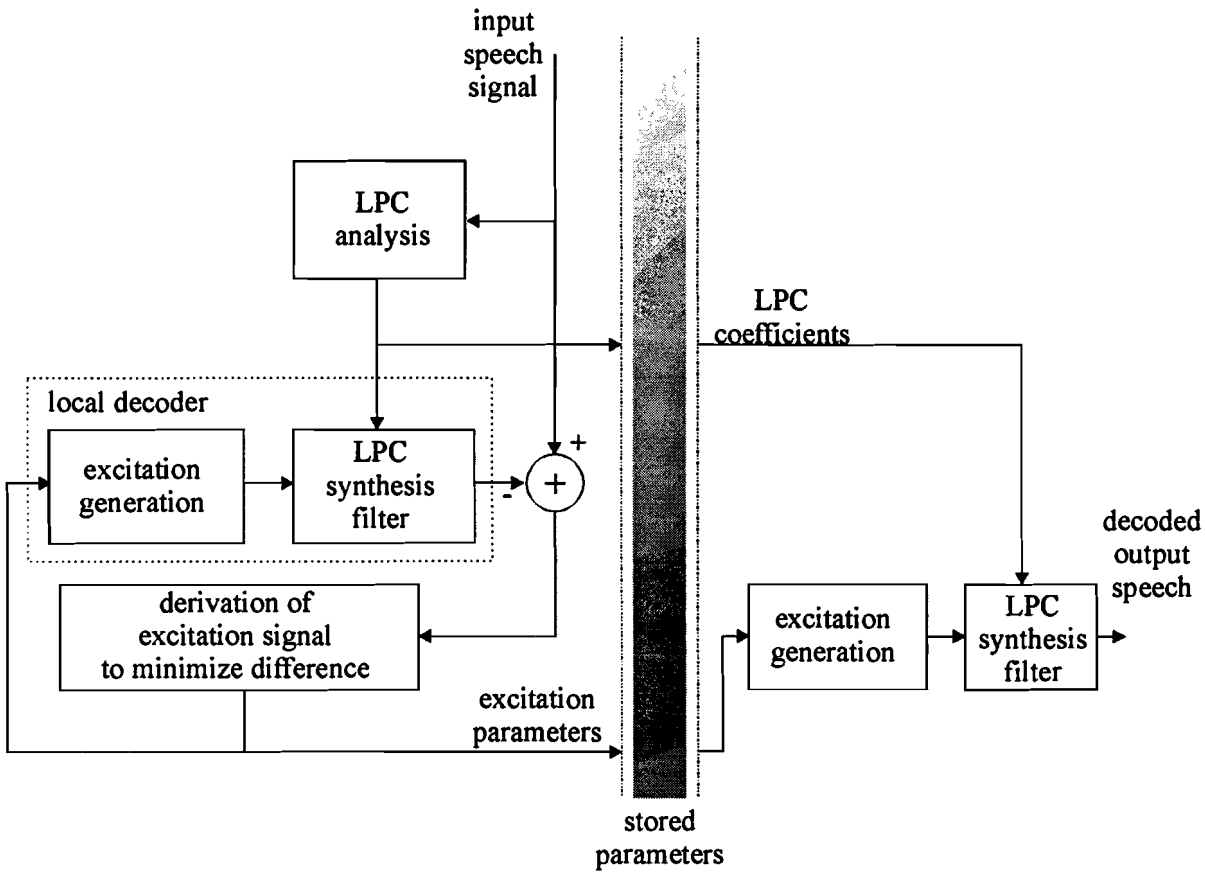


Figure 6: Analysis-by-synthesis hybrid coding scheme.

A brief description of the very popular low bit rate analysis-by-synthesis LPC coders is presented in the next few sub-sections.

6.2.3.1 Multipulse excited LPC

In a multipulse LPC (MPLPC) decoder a series of non-uniformly spaced pulses with different amplitudes is used to excite the filter. Unlike LPC vocoding, no distinction is made between voiced and unvoiced speech: the same type of excitation waveform is used for all speech segments.

At the MPLPC encoder the method of deriving the multipulse excitation involves an analysis-by-synthesis procedure. The excitation analysis procedure requires the partitioning of the input speech into small blocks (20-80 samples), and a search for the pulse positions and amplitudes which minimize the error, between the input and the synthesized speech, over the block. A one pass solution to finding the optimum positions and amplitudes is a highly non-linear problem and is thus extremely complex. Sub-optimal methods have thus been developed which find the pulse positions and amplitudes one at a time.

The MPLPC coder can be extended to include long-term prediction and perceptual weighting. The long-term prediction included in the encoding process takes advantage of the long-term correlations in speech which arise primarily as a result of pitch related correlations in voiced speech. With the inclusion of long-term prediction, fewer pulses are required per pitch period to obtain the same speech quality. The long-term prediction parameters can be accurately quantized with relatively few bits.

Perceptual weighting of the error is included at the encoder so that the pulse positions and amplitudes are chosen in a way which minimizes perceived distortion rather than just minimizing the mean squared error. Multipulse codecs can operate successfully over a wide range of bit rates, e.g. from 8 kbit/s to 16 kbit/s.

6.2.3.2 Regular Pulse Excited LPC

The regular pulse excited (RPE) LPC coder is a variant of the multipulse coder. The excitation signal pulses are spaced uniformly. The 13 kbit/s full-rate codec algorithm adopted as the standard for the forthcoming pan-European digital cellular mobile telephone service is essentially a regular pulse-excited LPC algorithm. A new set of long-term prediction parameters (gain and delay) and excitation parameters (RPE index and excitation pulse amplitudes) are derived every 5 ms using an analysis-by-synthesis procedure.

6.2.3.3 Code excited linear prediction

The code-excited linear prediction (CELP), or stochastic excited linear prediction (SELP) coder is a coder based on an analysis-by-synthesis technique. The difference between multipulse and CELP coders is the excitation function - the pulses of multipulse are replaced by an 'innovation sequence'. At the decoder of a CELP codec each block of reconstructed speech samples is produced by filtering the selected innovation sequence through the long-term filter and then the LPC vocal-tract filter. At the encoder there is a codebook which

contains many different innovations. The encoder selects the 'optimum' innovation sequence by filtering each sequence in the codebook in turn; the sequence which results in the minimum weighted mean squared error between the input speech signal and the synthesized signal is chosen. An index number is stored to identify the selected innovation sequence along with a gain term which sets the energy of the excitation signal. As with multipulse coders no distinction is made between voiced and unvoiced speech. The innovation sequence length is typically 40 samples and the long-term predictor analysis, or now more commonly, the adaptive codebook analysis, is also typically performed every 40 samples.

The CELP codecs specified for the US and Japanese digital cellular telephony standards were both developed by Motorola and are known as vector-sum excited LPC (VSELP) codecs. The major difference between VSELP and conventional CELP codecs is the structure of the fixed codebook(s) (two fixed codebooks are used for the 8 kbit/s US digital standard and one codebook for the 6.7 kbit/s Japanese standard.

One very important recent CELP codec development from AT&T Bell Laboratories is the low delay CELP (LD-CELP) speech coding algorithm which has been selected as the CCITT 16 kbit/s speech codec standard. The most significant feature of this codec is that it has a delay of less than 2 ms. The main disadvantage of this coder is the high complexity. Unlike previous CCITT Recommendations for speech coding algorithms, the Recommendation for the 16 kbit/s algorithm (Recommendation **G.728**) is not a 'bit-exact' specification. Thus to meet the requirements of the **G.728** Recommendation, implementations of the algorithm may use either floating- or fixed-point arithmetic

6.2.3.4 Frequency domain algorithms

Frequency domain coders are most often used for the high quality encoding of wide bandwidth (approximately 7 kHz) audio signals. As for the low bit rate encoding of telephony signals, the standards for these audio codecs are important to ensure correct interworking and compatibility between different manufacturers' implementations. One of the most important standards is the CCITT Recommendation **G.722** which specifies an algorithm for the encoding of 7 kHz audio signals within 64 kbit/s.

A newer frequency domain technique is known as multiband excitation (MBE) coding. This coding technique is based on a more sophisticated model of the speech production process than that shown in Figure 4. In the MBE model this speech signal is split into 20 ms speech frames and the spectral shaping of the vocal tract for each speech frame is modelled by means of a number of non-overlapping, variable width frequency bands.

6.2.3.5 TrueSpeech 6.3/5.3/4.8

TrueSpeech is a family of speech compression and decompression algorithms and software developed by DSP Group. It is an enabling technology for personal computers and future personal communications devices.

With the high compression ratios ranging from 15:1 to 27:1, TrueSpeech improves the efficient storage and communications transmission of digital voice information. It also facilitates the integration of personal computer and the telephone. TrueSpeech can be utilized in many products and applications.

The TrueSpeech 6.3/5.3/4.8 kbit/s speech compression algorithm is an analysis-by-synthesis vocoder. It is based on Multipulse Maximum Likelihood Quantization (MP-MLQ) Linear Predictive Coding (LPC) technique and provides excellent speech quality at very low data rates with reasonable computational complexity. TrueSpeech 6.3/5.3 has been designated as **G.723** by the ITU.

The encoder operates on frames of 30 ms. Each frame is divided into 4 subframes. For every subframe, a 10th order LPC filter is computed from the original speech samples. The LPC coefficients of the last subframe is quantized using Predictive Split Vector Quantizer. A Formant Perceptual Weighting filter is constructed from the LPC coefficients to filter the input speech in each frame. For every 2 subframes, the open loop pitch is computed using the perceptual weighted speech signals. The speech is then processed on a subframe basis. A closed loop 3rd order pitch predictor is computed for each subframe and its contribution is subtracted from the error signals. The random excitation is approximated using multipulse excitation. Its quantization and coding are performed simultaneously using a maximum likelihood quantization approach. All properly quantized parameters are packed as code blocks to be sent to the receiver.

The decoder operation is also performed on a frame by frame basis. First, the decoder receives the quantized LPC and the estimated pitch indices. After these parameters are decoded, it constructs the LPC synthesis filter. Then, for every subframe, both the pitch predictor and the random excitation are decoded. By passing the excitation signals through the synthesis filter, the reconstructed speech can be obtained. The adaptive post filter, which consists of formant and pitch post filters, is applied to the synthesized speech to improve the quality of the final output speech. The data rates can be switched among 6.3, 5.3 or 4.8 kbit/s. The 6.3 kbit/s algorithm produces higher quality than the 4.8 kbit/s version, but requires higher computation. This algorithm requires an 8 kHz sampling rate using 16 bit samples.

6.2.4 Comparison of the different coding techniques

There is no universally accepted measurement of complexity for speech coding algorithms, as neither the number of instructions per second (MIPS) nor the number of operations (MOPS) by themselves adequately define the complexity of an algorithm. While the two are closely clearly related, unfortunately the MIPS or MOPS complexity do not tell the whole story with regard to power consumption. To measure the complexity of an algorithm in such a way that it relates more directly to power consumption, the amount of memory required by an implementation must also be taken into account. While measurements of speech coding algorithm complexity have been employed based on *ad hoc* combinations of MIPS/MOPS and memory requirements, no universally formula has been established. Without a meaningful widely accepted complexity measurement available, it is difficult to present precise complexity comparison for different coding techniques. Therefore complexity comparisons in broad terms only are presented in Table 1 for various speech coding techniques. All comparisons are relative to PCM, which has been assigned the arbitrary complexity of 1. The TrueSpeech algorithms are compared to eachother and listed in Table 2.

While the speech quality of a speech coder can be defined in terms of mean opinion scores (MOS), and, provided the same test conditions and references are used, meaningful comparison of speech quality between coders can be obtained in terms of MOS, only a broad classification has been attempted in Table 1. The three speech quality categories used in Table 1 are:

- **toll quality** (quality of a long distance PSTN connection)
- **communications quality** (speaker identity maintained, good intelligibility, but distinguishable loss in quality compared with toll quality)
- **synthetic quality** (speech sounds synthetic in nature, speaker identity information largely lost)

Table 1: Complexity and speech quality comparison for several speech coding techniques.

speech coding technique	*Bit rate (kbit/s)*	*Speech quality*	*Complexity*
PCM (G.711)	64	Toll	1
ADPCM (G.721)	32	Toll	10
CELP (G.728)	16	Toll	450
RPE-LPC (GSM)	13	Communications	100
VSELP (US cellular std)	8	Communications	250
CELP	4.8	Synthetic/ Communications	400
IMBE	4.15	Synthetic/ Communications	150
LPC10 (military std)	2.4	Synthetic	100

Table 2: Comparison of the TrueSpeech 8.5, 6.3, 5.3 and 4.8 kbit/s algorithms.

		Compression			Program[3] (words)	
TrueSpeech	Quality	Ratio 1[1]	Ratio[2]	MIPS[3]	ROM	RAM
8.5	Good	15:1	21:1	9-9.5	4-6.5	1-2k
6.3	Best	20:1	29:1	16-22	16k[4]	2.2k[4]
5.3	Best	24:1	34:1	15-20	16k[4]	2.2k[4]
4.8	Better	27:1	38:1	14-18	+1k	2.2k[4]

[1] 16 Bit samples, no silence compression [2] 16 bit speech samples with silence compression (30%)
[3] 16 bit fixed point DSP, combined encoder/decoder [4] Estimates based on Motorola 56156

Table 1 illustrates that in general to increase speech quality or decrease coder bit rate results in increasing complexity of the coder. However, while the complexity of new speech coding algorithms generally exceeds that of previous algorithms, the continued increase in the available processing power from new digital signal processing devices is keeping pace with the demand.

6.3 Echo Cancellation

In the Smartphone, two kinds of echoes appear, which distort the desired signal, acoustic- and electrical echo. They are different types of echoes and will be described separately in the following sections.

6.3.1 Hybrid Echo Canceller

Most voiceband telephone connections involve several connections through the telephone network. The 2-wire subscriber line available at most sites is generally converted to a 4-wire signal at the telephone central office. The signal must be converted back to a 2-wire signal at the far-end subscriber line. The 2-to-4-wire interface is implemented with a circuit called a hybrid.

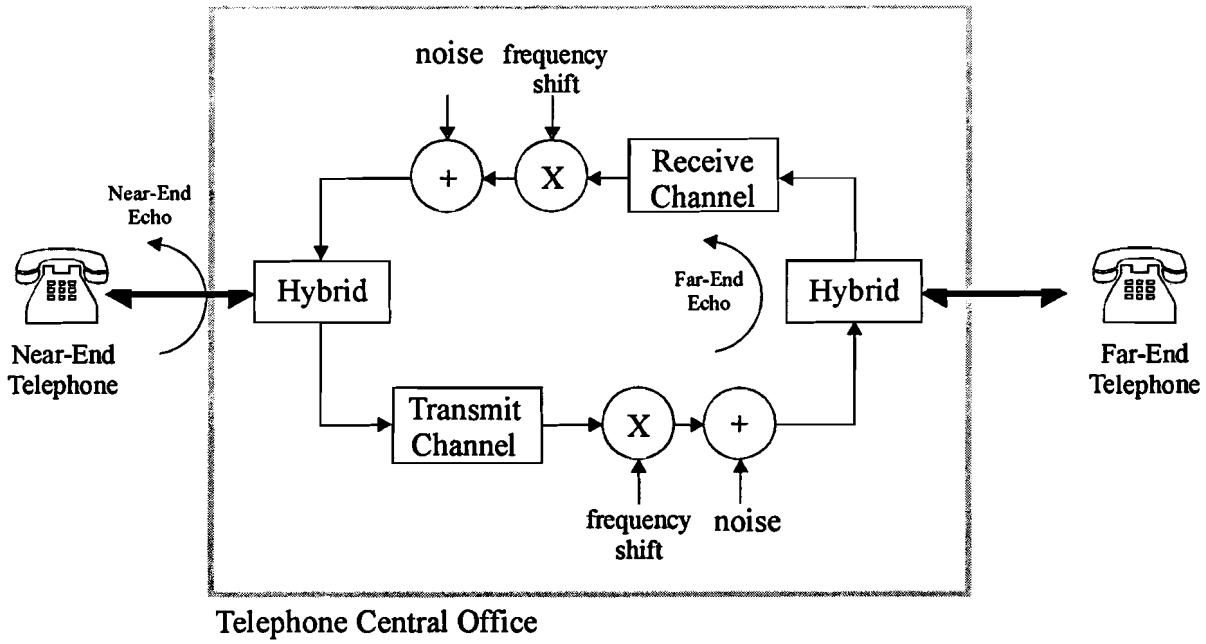


Figure 7: Telephone Channel Block Diagram

The hybrid intentionally inserts impedance mismatches to prevent oscillations on the 4-wire trunk line. The mismatch forces a portion of the transmitted signal to be reflected or echoed back to the transmitter. The telephone system and sources of echo are shown in Figure 7. There are two types of echoes in a typical voiceband connection. The first echo is the reflection from the near-end hybrid, and the second echo is from the far-end hybrid.

Communication over the existing telephone wires requires the simultaneous transmission in opposite directions (full-duplex). The public telephone network typically only provides a 2-wire connection to each customer premise (this is because of the high cost of copper wire). Full-duplex transmission on a 2-wire circuit can be accomplished by using hybrids. Hybrids form an imperfect isolation between the transmitter and receiver on either sides though. A part of the transmitted signal leaks back to the receiver. This third type of echo is called the local-end echo and is created by the local-end hybrid (see Figure 8).

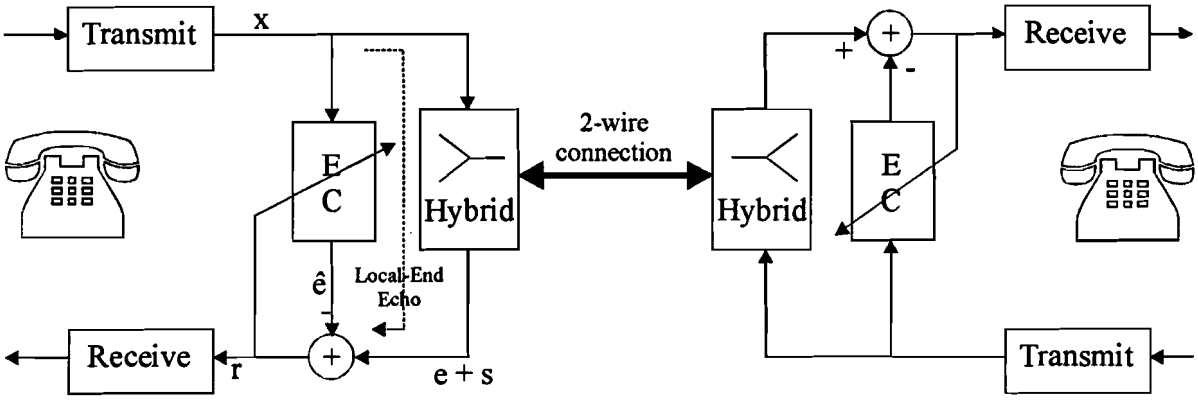


Figure 8: Echo cancellation for full duplex transmission

The local-end echo is about 10 dB, while the signal *s* from the other side is attenuated by about 40 dB. This means that this echo is about 30 dB stronger than the received signal, while this relation should be about -20 dB, in order to be able to hear or detect the desired signal.

Echo cancellation is an application of adaptive filtering technology to control the echo in the telephone network and local-end hybrid (together called *e* in Figure 8). It is accomplished by subtracting an estimate ($\hat{e}$) of the echo return signal from the actual received signal ($s + e$). The estimated echo is generated by feeding the transmitted signal (x) into an adaptive filter whose transfer function tries to model the telephone channel. The filter coefficients are determined during a training sequence prior to full duplex communications.

The Least Mean Square (LMS) algorithm is a robust and easy implementable algorithm, and is therefore used very often. In this algorithm, tap weights are assumed to have an arbitrary

initial setup and are moved in the direction of optimum value when the mean squared error is minimized.

6.3.2 Acoustic Echo Cancelling

When communicating in a hands-free mode, two types of impairment will be present in a hands-free system: ambient noise from background sources and acoustic echo in the signal picked up by the microphone. Acoustic echo is thus a crucial feature for the successful hands-free operation of a telephone.

Historically, hands-free systems have employed gain-switching techniques to attenuate the feedback path. Such techniques limit the communication in a hands-free device to half-duplex. More recently, the focus has shifted to echo control techniques based on adaptive system identification.

When calling in hands-free mode the output of the loudspeaker echoes back to the build-in microphone (see Figure 9). This phenomena is called acoustic echo. The returned signal will be amplified and echoed again. Eventually a loud noise comes from the loudspeaker and the hands-free mode crashes. The acoustic echo path, which varies in time, can have a length of some hundreds of milliseconds.

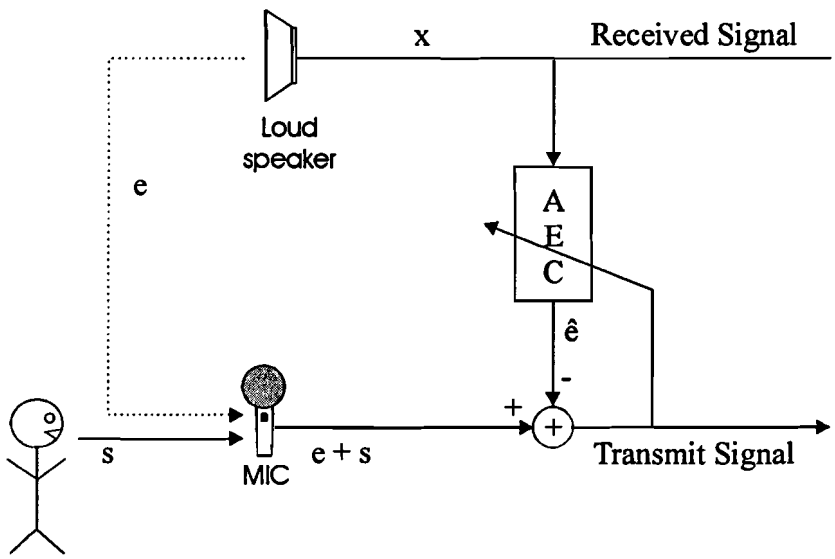


Figure 9: Acoustic Echo Path

To overcome this problem, an acoustic echo canceller is needed. This device uses a model of the acoustic echo path and generates an estimate $\hat{e}$ of the echo signal e. This signal is then subtracted from the received signal $(e + s)$. This way, the acoustic echo canceller surpresses the echo of the received signal x.

7. Answering Machines

Traditionally, answering machines use magnetic tape as the message storage medium. Such machines consist of either one or two tape recorders. One tape is used to store Incoming Messages (ICM). The other tape stores the Outgoing Message (OGM). More advanced machines store the OGM on a semiconductor device. These machines usually use adaptive delta modulation to compress the OGM, and store the result on a DRAM device.

The advantage of this classical structure is the simplicity and the maturity of the design. The cost of tape drives is very low, and the control functions can be done with a 4-bit micro-controller. The main disadvantages of such machines are the low reliability of the mechanical parts, and the wear of the magnetic tape. Another disadvantage is the serial access imposed by the tape. This does not allow selective erasure of messages, and implementation of mailboxes.

Digital answering machines solve all the above problems. The ICM is stored on a random access memory, with no mechanical parts. Selective erasure and mailboxes can easily be implemented by use of a directory of messages.

A conventional answering machine consists of five main blocks: Telephone Line Interface, tape recorders, Man Machine Interface (MMI), DTMF decoder, and a microcontroller (see Figure 10).

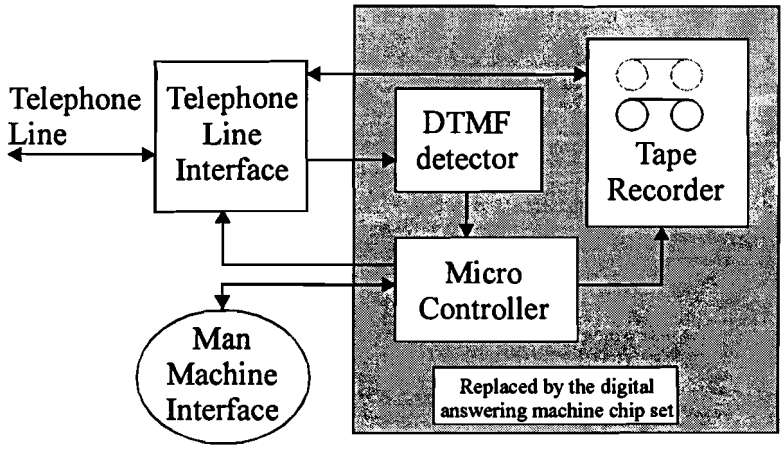


Figure 10: Tape Recorder Based Answering Machine

The telephone line interface performs functions such as line isolation, audio gain control, and ring detection. The MMI block consists of a keypad, switches, and a display. It provides local control and status (e.g. wait for a call, play a message, show number of messages). The DTMF decoder detects DTMF signaling, which are used for remote control operation. The micro-controller performs the control functions. It activates the various parts of the machine according to user commands and line status. Ring detection, for example, start a message acceptance sequence.

The digital answering machine has development in the last few years to an essential part of nowadays communication. It receives messages and compresses them for efficient storage. Next, this data is stored in memory-chips. The main advantage of the digital answering machine is the greater reliability with regard to the mechanical tape and the possibility to implement voice-mail features like selective deleting of messages and the creation of multiple mailboxes.

Two parameters determine the cost of a digital answering machine: the cost of the storage memory and number of devices in the system. To reduce the cost of the storage memory, two approaches can be combined. One is to compress the audio before recording it and the other is to use ARAM devices. Compression algorithms reduce the bit rate of the sampled voice so the system can store more audio on the same memory device. The use of ARAMs further reduces the system cost. ARAMs are DRAM devices with defective locations and slow access time. The cost per bit of such devices is much lower than that of conventional DRAMs. Error correction allows storing the compressed data on ARAM devices.

8. Three different strategies

During my search for useful Digital Signal Processors, I discovered three possibilities that are suitable for this application. The first one was the use of general purpose DSPs, a solution which I at first considered to be the only possible solution. But soon I discovered another possibility, called DTAD (Digital Tapeless Answering Device). This device contains a DSP and ROM and utilizes advanced software algorithms to provide new features like Echo Cancellation and a Digital Answering Machine, which can be used in the smartphone.

Finally the question arose if it is possible to create a custom chip (ASIC), containing all the digital signal processing power needed for this device. For this question I visited a company called Center for Micro Electronics (CME).

The following three chapters give a description of these possible solutions and lists some of the companies and devices which are currently available. Chapter 12 describes the peripherals which need to be connected to these devices.

9. General Purpose DSPs

As described in Chapter 4, Digital Signal Processors are integrated circuits specialized to accept and process a range of real-time signals. There are a great number of different vendors of general purpose DSPs and each has a large collection of different types. The 16-bit fixed-point devices continue to dominate low-cost applications. As new variations of each DSP μP hit the market, each family fills a greater variety of applications. The largest markets for these 16-bit speedsters include telecommunications and digital cellular telephones. The next sections describe some of the commonly used DSPs and a give an indication about the minimal processing power required in this application.

9.1 Choosing a DSP

Finding the right DSP for a specific application is not as easy as it seems. You can compare the problem to finding the way to a small village. At first one must know at which direction he has to go. This direction is a first criteria for the selection. Next some main cities must be chosen which lead towards the area you should end up. Finally you can select the right village.

In order to select the right chip for an application, first a direction should be chosen. This direction is DSP. Next the search continues to the companies who deliver DSPs which are able to handle the functions you want to implement. After a choice of company has been made a list of all family-members can be generated in order to make the best possible selection.

Finding the best DSP is not possible. First of all, there are too many DSPs to compare with each other. Second, When you think you found the best DSP for your application, better DSPs arrive at the market and the search can start all over again. Therefore it is better to find the DSP which is capable enough for your application.

9.2 Analyses of the required DSP properties

In order to be able to choose a suitable DSP, you must have information about the worst case situations in your application. This means that information regarding memory usage and processing power should be gathered. This not only requires knowledge about the different types of processes in the smartphone. It is also very important to know which of them must be executed concurrently. When all concurrent processes are investigated, you have the information about the maximum processing power (expressed in MIPS) and memory usage.

In the smart-phone, the following functions are going to be implemented:

- Voice-compression and de-compression
- DTMF detection
- Acoustic echo-cancellation
- Electric echo-cancellation
- Caller ID
- Tone detection for modem-implementation

According to these functions the following criteria are important for the selection of the DSP:

- Internal Memory
- Number of Mips/Mops

The global flow-diagram of the smartphone (see
Figure 11) shows three different transitions the state can make, when the phone rings: Answering Machine, Normal Telephone Mode and Speakerphone Mode. When dialing yourself another state arises, the Modem State.

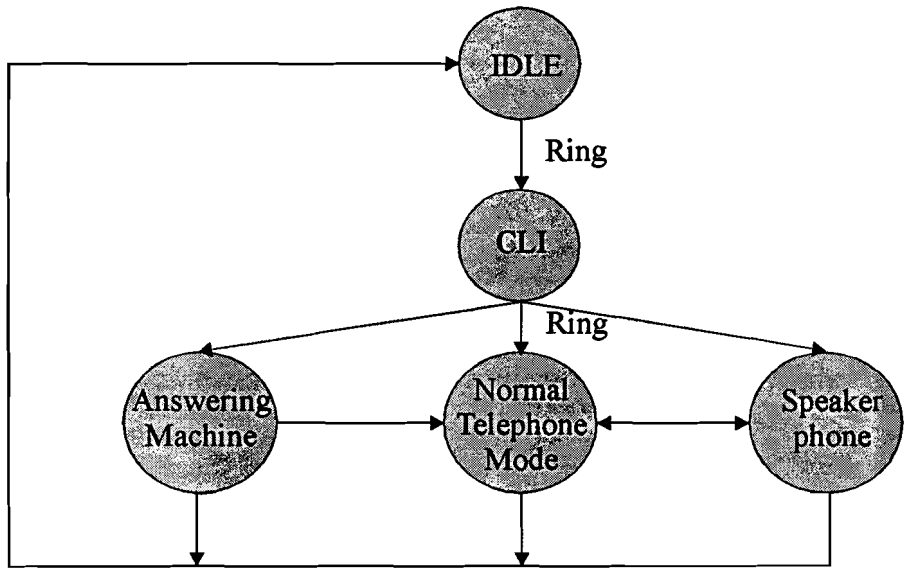


Figure 11: Global flow-diagram of smartphone

The Normal Telephone state contains no complex processing algorithms. The Speakerphone state only contains the Acoustic Echo Cancelling routine. This is indeed a complex algorithm, but since this is the only algorithm, no state diagram is given. The Modem state only has tone-detection and tone-generation, which is also not very complex.

The answering machine has concurrent voice-compression and DTMF decoding. This and the Speakerphone algorithm form the most complex functions of the DSP. Figure 12 shows the flow-diagram of the Answering Machine mode.

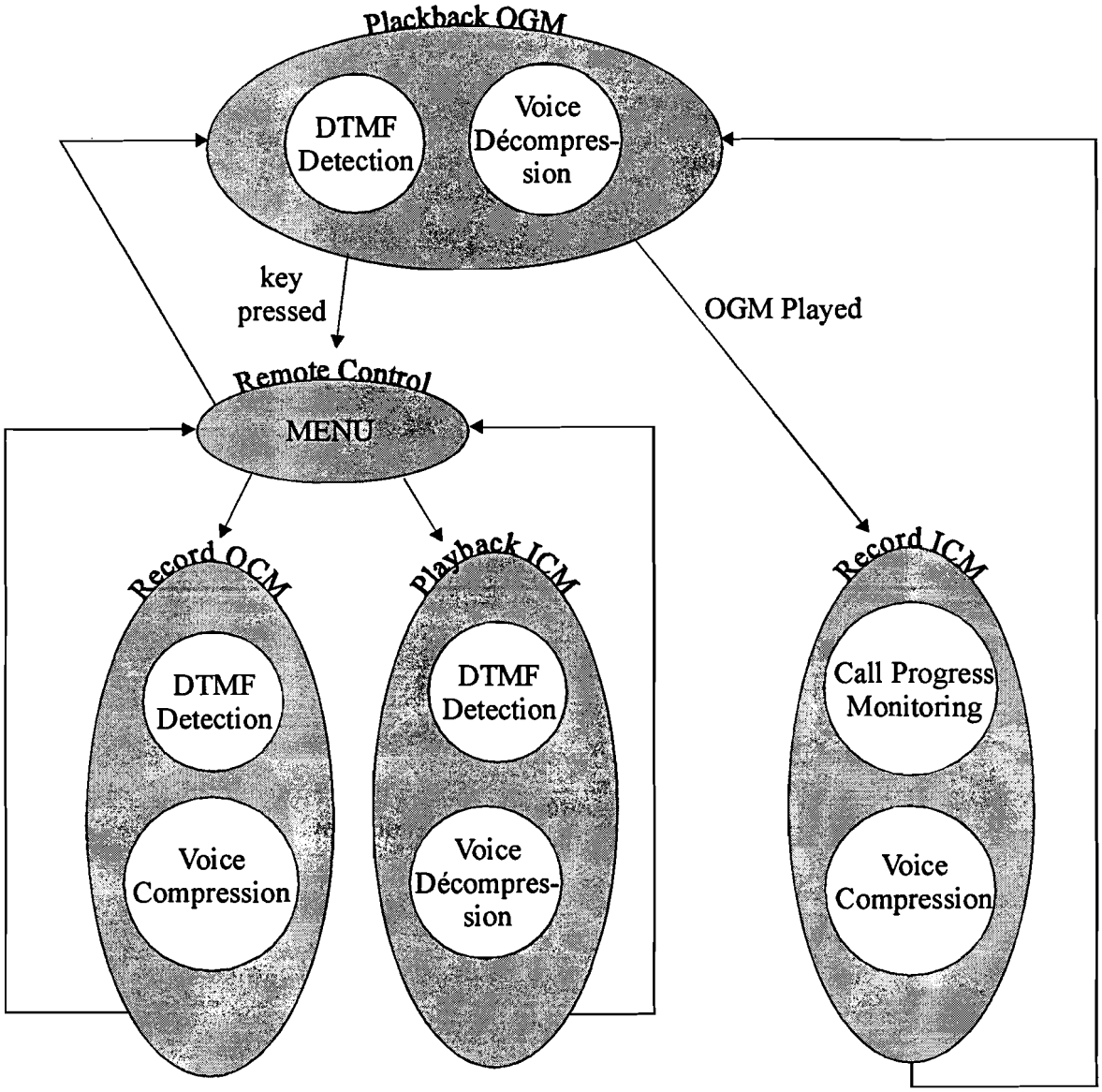


Figure 12: Answering Machine State Diagram

9.3 List of general purpose DSP chips

The following sections provide up-to-date information on the general purpose DSP chips available today. This overview is a summary of information gathered from EDNs DSP-CHIP DIRECTORY [11] and on-line information from the internet. Although new processors are being developed, it can be used as a starting point. For detailed technical information it is best to read the EDN [11].

When comparing different DSP architectures, you can see how DSP manufacturers choose to optimize their products. In many DSP applications, it's important to move instructions and data around inside the CPU as quickly as possible. To accomplish this, some devices such as Motorola's DSP96002, have taken the Harvard architecture to the extreme by using as many as eight internal buses, while others, such as AT&T DSP32C, have limited the number of internal buses to two, thereby optimizing device cost. In addition, some DSPs place an emphasis fast context switching and, therefore, include multiple register sets; and others still depend on an accumulator based architecture. The following sections describe the general purpose DSPs from Analog Devices, Texas Instruments, Motorola and AT&T. Because the fixed-point processors are fast enough and the cheapest around, only these types will be discussed.

9.3.1 Analog Devices

Analog Devices has four family members that exhibit the most distinct features. These family member are:

- ADSP-2100: Family Base Architecture-contains the computational units, address generators, and program sequencer
- ADSP-2101-contains the base architecture, plus on-chip memory (program, data, and boot-memory), a programmable timers and enhanced interrupts
- ADSP-2111-contains the features of the ADSP-2101, plus a host interface port (HIP)
- ADSP-21msp50-contains the features of the ADSP-2101, plus a host interface port (HIP) and a voice-band analog front-end

Other family members are variations on these DSPs. For example the ADSP-2105 is based on the ADSP-2101, but it has less on-chip memory and only one serial port; the ADSP-2171 is similar to the ADSP-2111, except it has functional enhancements (low power operation and expanded instruction set). All the microcomputers of the ADSP-2100 family are code-compatible, although some modifications for interrupt vectors, peripherals, and control registers may be necessary. Table 3 gives an overview of de specifications of the Analog Devices DSPs.

Table 3: Family functional differences

Model	Mips	Internal Program RAM/ROM (24 bits)	Internal Data RAM (16 bits)	Host Interf Port	Serial Ports	On Chip AD /DA	Current mA	Special feature	Price $/amount
2101	20	2k / -	1k	-	2	-	70	-	26 / 1000
2105	10	1k / -	0.5k	-	1	-	55	-	9.90 / 1000
2111	20	2k / -	1k	√	2	-	70	-	38 / 1000
2115	20	1k / -	0.5k	-	2	-	70	-	13.6 / 1000
2161	16	- / 8k	0.5	-	2	-	42	-	15.81 / 10.000
2163	16	- / 4k	0.5	-	2	-	42	-	9 / 10.000
2165	20	1k / 12k	4k	-	2	-	60	-	25 / 10.000
2171	33	2k / 8k	2k	8/16b	2	-	76	-	23 / 10.000
2181	33	16k / -	16k	-	2	-	76	16 bit DMA	38 / 1000
21msp58	-	2k / -	2k	√	2	16bit	95	-	26 / 10.000

* all these processors have boot program memory, programmable timers, 3 external interrupts, low power modes, and a total addressable space of 16k.

9.3.2 Texas Instruments

The Texas Instruments TMS320 family of digital signal processors offer a wide range of fixed- and floating-point solutions. High-performance single and multiple DSPs, available for as little as $3, are viable for even the most cost-sensitive applications, like consumer and automotive products.

The TMS320 family currently comprises six generations of DSPs and features a variety of price-performance alternatives. The TMS320C1x, 'C2x, 'C5x generations are fixed-point DSPs, while the 'C3x and 'C4x are floating-point generations. Spin-offs based on the core CPU of each generation offer a variety of memory and peripheral configurations to address specific application needs. The newer higher performance, code-compatible generations provide an easy product upgrade path while preserving software investment. All TMS320 DSPs are supported by products and services from independent vendors and consultants, known as third parties.

The TMS320C1x has a large collection of application code supporting a large range of applications. The chip requires careful assembly-language programming, but costs have dropped to the point that C1xs are competitive with many µPs and µCs. In volume, C10s cost less than $3 apiece. You can use the 16-bit external data bus as an I/O port to connect external peripherals such as ADCs or DACs. Separate I/O signals let you use up to eight separate I/O

ports on the bus. On the C17 companding hardware compands (compresses and expands) data for serial or parallel mode. It handles both the A- and μ-Law forms.

The TMS320C2x is TI's second generation 16 bit, fixed-point DSP family. Source-code-compatible with the first generation C1x, the C2x improved on the C1x limitations. External memory addressing is 16 bits (64k words each for code and data), and it has 24 new instructions. Program ROM is 8 kbytes and DMA simplifies external access to internal memory. All versions have a synchronous serial port and a timer.

The TMS320C5x, source code compatible with the C2x, is TI's highest performance 16 bit-fixed-point DSP family. The C5x runs at up to 100 MHz. The TMS320C52 DSP CPU is offered as an application-specific-processor (ASP) core for a TI gate array, the TEC320C52. The power-down mode minimizes power by shutting down the CPU (IDLE 1 instr) or the CPU and the peripherals (IDLE 2). Pulling down an external pin can always force the chip to go into power-down mode. An interrupt brings the chip up to normal run conditions.

Table 4 gives an overview of the different fixed-point types and their specifications.

Table 4: TI's family functional differences

Model	Mips	Internal RAM/ROM (16 bits) (bytes)	Timers	Parallel Ports 16 bits	Serial Ports	Current mA	Price $/amount
320C10	6	288/3k	-	8	-	94	$4/1000
320C14	6	512/8k	2	16	-	50	$8/1000
320C15	6	512/8k	-	8	-	85	$5.6/1000
320C16	8	512/16k	-	8	-	100	$6/1000
320C17	5	512/8k	1	6	2	65	$5/10.000
320C25	12.5	1088/8k	-	16	1	185	$11.6/1000
320C26	12.5	3x1k/512	1	16	1	220	$10.5/1000
	MHz						
320C50	40/57/80	20k/4k	1	-	2	60/67/94	$32/$42/$48
320C51	40/57/80	4k/16k	1	-	2	60/67/94	$21/$23/$25
320C52	40/57/80	2k/8k	1	-	1	60/67/94	$13/$14/$17
320C53	40/57/80	8k/32k	1	-	2	60/67/94	$31/$34/$39
320LC56	57/80	14k/64k	1	host	2	57	$41
320LC57	57/80	14k/64k	1	host	2	80	$54

9.3.3 Motorola

The DSP56000 targets audio and other applications that take advantage of the larger data size. The chip has no on-chip program ROM, except for a small boot ROM on some versions. However, the DSP56000 can access external memory each cycle without any penalties. The processor combines 16 bit addressing with 24-bit words. It supports three 64 kbyte address spaces, one each for X, Y, and program memory.

Motorola's 16-bit DSP561xx processors are build on the basic DSP56000 architecture, adding a codec for A/D and D/A voice conversions. The DSP56156 suits the European cellular standards; the DSP56166 suits the American and Japanese VSELP standards (see section 6.2.1). The DSP561xx has on-chip program RAM and dual-ported data RAM. Each has it own address and data bus.

Table 5 gives an overview of the fixed-point DSPs from Motorola.

Table 5: Motorola's fixed-point DSPs

Model	MHz	Prog/Data RAM 24 bits	Prog boot ROM 24 bits	Host-processor	Special	Current mA	Price $/amount
56002	40/66	512/2x256	64	8b parallel*	serial communication interface a)	90	$29/$39/ -
56L002	40	512/2x256	64	8b parallel*	serial communication interface a)	50	$34/-
56004	50/66	512/2x256	32	serial	serial audio interface a)	90	$25/$30/ -
56007	50	6400/3200	52	serial	serial audio interface a)	90	$38/-
		16 bits	16 bits				
56156	60/40	4k/4k	128	8b parallel	codec/timer	188	$46/$44/ 1000
56166	60	4k/8k	128	8b parallel	codec/timer	100	$51/1000

* with DMA support
a) include PLL, power-saving modes, a timer, and a data ROMs containing sine, A-Law and μ-Law tables.

9.3.4 AT&T

The DSP16xx family has multiple on-chip memories, a small instruction cache, dual address generators, and multiple buses to move data in parallel. The DSP defines two 64k-word address spaces, one for program/coefficients and one for data. Both connect to the same dual-ported RAM. Table 6 gives an overview of the 16-bit fixed-point DSPs of AT&T.

Table 6: Fixed-point DSPs of AT&T

Model	MHz	Internal RAM (16 bits)	Internal ROM (16 bits)	timers	JTAG	Current mA	Special features	Price $/amount
16A	40	up to 3k bytes	up to 32 kbyte	-	-	-	serial & parallel	$15/10.000
1604/06	40	1/2k*	16/24k	2	yes	100	3x 8 bit & serial	$10/$15/-
1605	40	1k*	16k	2	yes	100	serial & DRAM controller	$10/-
1610	40	4k-or 8k*	512 boot	1	yes	130	2 serial	$57/-
1611	100	124k*	2k boot	1	yes	85	2 serial & 8 bit host	$61/-
1617	50	4k*	24k	-	yes	100	2 serial/ mask prog. clock & 8bit host	$39/-
1618	50	4k* a)	16k	-	yes	75	2 serial/ mask prog. clock & 8 bit host	$49/-
1616X30	50	4k* a)	12k	-	yes	100	2 serial/ mask prog. clock & 8 bit host	$25/-
1627	70	6k*	32k	-	yes	98	2 serial/ mask prog. clock & 8 bit host	$45/10.000

* Dual ported RAM
a) or Flash memory

10. DTAD

Digital Tapeless Answering Devices (DTADs) are single chip microcontrollers for digital answering machines. Only two other devices, an Audio quality DRAM (ARAM) and a codec are needed to implement all the digital part of such an answering machine. The device performs voice compression and decompression, DTMF detection, voice synthesis, and control functions of digital answering machines.

The main advantage of this chip is that it includes all the complex digital signal processing software. You only have to write the control software of the microcontroller attached to it. The disadvantage is that the chip can not be extended with additional functions in the internal ROM. This means that self-developed algorithms like CAS detection have to be added to the core in consultation with the DTAD vendor.

The following sections describe some of the DTAD chips available today.

10.1 DSP-Group

DSP Group, Inc. develops and markets Digital Signal Processors and digital speech products for the PC, multimedia, consumer and communications markets. Headquartered in Santa Clara, California, the Company offers TrueSpeech®, its speechcompression software, the PineDSPCore and OakDSPCore designs and families of speech and telephony processors. DSP Group's speech and telephony processors are designed for digital telephone answering devices, consumer telephones, digital recorders, laptops, personal computers and computer telephony equipment. DSP Group's technologies and products have been adopted by industry leaders including AT&T, Microsoft, Motorola, Intel, LSI Logic, VLSI Technology, Sony, Sharp, Panasonic, Philips, Siemens, GEC Plessey, Matra, Alcatel, British Telecom, France Telecom, U.S. Robotics, and others.

In 1994, DSP Group emerged as a world leader in the growing market for digital signal processors for the Telephone Answering Device (TAD) market. Product sales in this category increased 271% in 1994 and the Company believes that it now ships more than 40% of the world's digital TAD speech processors. Through its digital TAD speech processor business, DSP Group works with many major manufacturers, including among others, Alcatel, AT&T, Bell South, British Telecom, Double Kingdom, France Telecom, GE, HB, HPF, Matra, Murata, Northwestern Bell, Panasonic, Philips, PhoneMate, Radio Shack, Sagem, Samsung, Sanyo, Sharp, Smoothline, Sony and Siemens. In 1994, the DSP Group introduced several new digital TAD speech processors in the D6300 series. These products utilize advanced software algorithms to provide new features and capabilities that make telephones and

answering devices easier to use and more productive. The D6000 series of Digital TADs are DSP subsystems that include all of the necessary functions for the design of an all digital answering machine. These functions include TrueSpeech® compression for speech recording and playback of messages and voice prompts, DTMF and tone generation, DTMF and tone detection with near-end echo cancellation, time and date stamping of messages, and memory management. New technologies for the product family, introduced in 1995, include True FULL Duplex SpeakerPhone™ and noise robust Speech Recognition for truly "hands free" operation. The current series of products allow 25-27 minutes of recording time in 4Mbit of audio grade DRAM (ARAM). The new generation of products also include those which interface directly with Flash memory devices. The Flash memory based products allow storage of messages, with no battery back-up, and enable the technology of hand held portable recording devices. As the price difference between digital and analog telephony devices diminishes, these new capabilities will drive consumer and business demand for DSP-empowered answering devices for telephones, fax machines and cordless phones.

FEATURES OF THE D6000 FAMILY

- TrueSpeech high-quality speech compression allowing 15 minutes for each 4 Mbit memory Flash or ARAM (Audio grade DRAM)
- Flexible storage of digitized incoming messages (ICM), memos, and outgoing messages (OGM)
- TrueSpeech natural-sound voice prompting, for Day/Time stamp and voice instructions
- Reliable DTMF and call progress detection
- Supports "offset playback" for jumps within a message
- FlexiSpeech variable speed playback
- TrueSpeaker Digital Speakerphone with Acoustical Echo Cancellation

Figure 13 shows a typical application with the D6000 family DTAD.

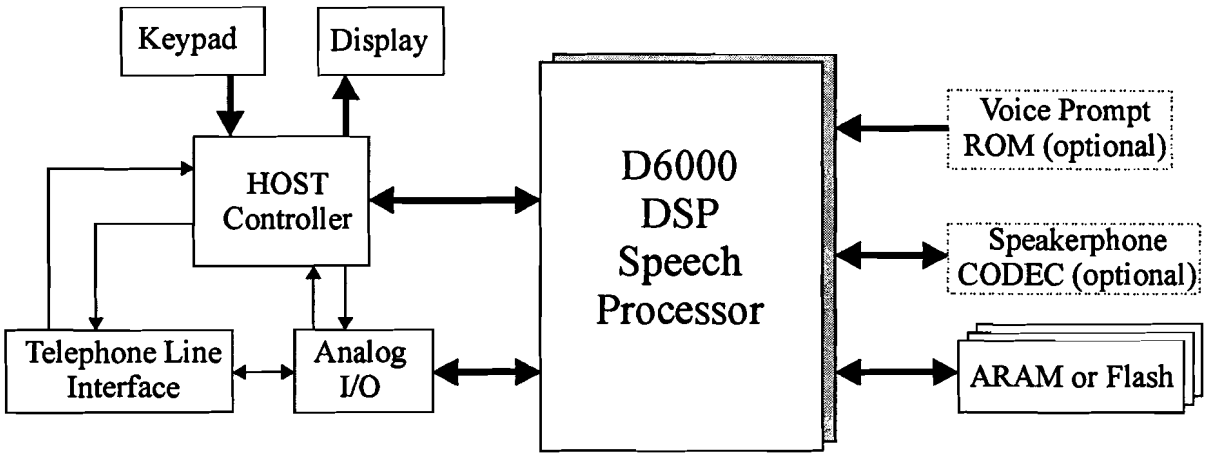


Figure 13: Typical Application

10.1.1 D6000 Family Overview

The DSP-Group has developed a large collection of DTADs. Table 7 gives an overview of the interesting DTADs and their special features. The Samsung Flash devices listed in this Table are of the types: KM29N040, KM29N1600T, and KM29N16TS.

As a result of my questions about the different types of DTADs, a representative from the DSP-Group visited Pijnenburg. He showed us which kinds of DTADs they are currently developing and indicated that they are interested in Pijnenburg's CAS algorithm. Their future plans exists of developing:

- Caller ID
- Built-in Codec
- Speaker Independent Voice Recognition
- DTAD + Fax Modem chip

The price of the DTADs lie around $12 (depending on the amount of features) for quantities of 200k/year. At the moment I only know the price of two of them. The D6471A costs $15 and the D6305B, $7.5.

Table 7: DSP-Group D6000 Family Matrix

Function	6455 A	6471 A	6331 A	6351 A	6365 A	6305 B	6375 A	6371 A	6385 A	6386 A
TrueSpeech (15-17 min in 4Mbit)			X	X	X	X	X	X	X	X
Super Low Rate TrueSpeech (24-27 in Mbit)	X	X								
DTMF Generation	X	X	X	X	X	X	X	X	X	X
DTMF Detection with echo cancellation	X	X	X	X	X	X	X	X	X	X
DTMF A,B,C,D Tone Detection		X								
Tone Generation and Call Progress detection	X	X	X	X	X	X	X	X	X	X
Caller ID		X								
Digital Volume Control	X	X	X	X	X				X	X
Programmable Signaling Sub-System Sensitivity	X	X	X	X	X				X	X
Flexible storage of OGM and ICM	X	X	X	X	X	X	X	X	X	X
Natural Sound Voice Prompt - Date Time stamp	X	X	X	X	X	X	X	X	X	X
Support for "offset playback"	X	X	X	X	X	X	X	X	X	X
Pseudo Full Duplex SpeakerPhone							X	X		
True Full Duplex Speakerphone (0 dB attenuation)	X	X		X	X				X	X
FlexiSpeech - Playback speed from 50% to 200%		X	X	X			X	X	X	X
Speaker Dependent Speech Recognition										X
8bit Host Interface	X	X	X	X	X	X	X	X	X	X
Voice data transfer to/from host - Data pump		X					X	X	X	X
Selectable Slave/Master Codec Mode - DECT support		X							X	X
Selectable A-Law or μ-Law support		X								
Selectable Sampling Rate (7.2 or 8 kHz)		X								
Support for 4 Mbit and 16Mbit ARAM	X				X	X	X	X	X	X
Flash memory via Host								X		
Direct AMD/Intel Flash Interface			X							
Direct Samsung Flash Interface		X		X						
Samples available	now	Mar 96	tbd	Jan-96	Jan-96	now	now	now	now	now
Production	now	Jun-96	tbd	Apr-96	Mar-96	now	now	now	now	now
Note:D6471 will be made available without Flexispeech (D6451) and without Flexispeech + Full Duplex (D6401). D6351, D6471 are pin and command compatible.										

10.1.2 D6000 Evaluation System

The D6000 evaluation system is comprised of three separate circuit boards. The mother board, the D6000HC, includes a host microcontroller (8051), audio interface, power supply circuitry, built in microphone and a RS232 interface to a PC. This board provides connectors to attach the other two circuit boards. A PC based package is used for control of the board and a demonstration of features of a D6000 product.

The second board is the D6xxx Evaluation board. This board includes a D6xxx family digital signal processor and, one analog I/O device (codec), optional speakerphone codec, one 4-megabit ARAM, and a Voice Prompt EPROM. This board plugs into the D6000HC board or can be plugged into a developers system for use in engineering development of a digital telephone answering machine.

The third board of this system is the optional D6000TL telephone line interface board. This board permits the evaluation of all chipset functions which work with the telephone line. This board is particularly useful in providing a demonstration of the Full-Duplex Speakerphone feature.

We received this evaluation board with the D6455B DTAD. The quality of the recorded sound was very good, in spite of the fact that this version compresses the data to 2.8 kbps. The Full duplex function was also tested and seemed to work good.

10.2 National Semiconductor

10.2.1 NS32AM163

National Semiconductor first developed the NS32AM163 voice processor. This is an integrated embedded processor tuned for Digital Answering Machines. It integrates an integer CPU core, a Digital Signal Processor and system control functions on a single integrated circuit. The AM163 was designed as a low cost solution for tapeless DTADs. Only two additional devices, a Codec and ARAM, are needed to implement the digital part of such answering machines. This system can either be used as a stand-alone DTAD or as a voice module in microcontroller based systems. The AM163 performs voice processing, and can handle system related functions such as phone line monitoring, display and keypad control.

The Codec, a TP3054 (or compatible), consists of standard PCM encoder-decoder and filter devices. It serves as the interface between the processor and the Data Access Arrangement (DAA), that connects the DTAD to the telephone line. The DAA includes line protection devices, a line transformer, dialing and audio circuitry, a hook and control switches, and detectors.

The AM163 has the following three operating modes:

- On-chip ROM mode
  This mode is ideal for mass-produced DTAD applications that must be implemented with the lowest possible chip count. In this mode the AM163 provides up to 32 kB of on-chip program ROM, 2.1 kB of on-chip RAM, one 8-bit dedicated output port and two 8-bit general purpose output ports.
- Off-chip ROM mode
  This mode is designed to be used in custom DAM systems that will be produced in smaller quantities. This mode allows attachment of an external ROM to the AM163 to provide a fast way of changing the application program. One on-chip general purpose I/O port is provided, as well as support for adding two external ports. On-chip ROM and RAM are still accessible.
- Development mode
  In this mode, external memory and I/O devices can be connected to the AM163, and internal processing can be monitored for application development and debugging. On-chip ROM and RAM are still accessible.

10.2.2 NSAM265

In august 1995, National Semiconductor announced it first RISC-based digital speech processor with a Flash memory interface. The NSAM265 is the first member of National's CompactSPEECH™ family, and is the first product implementation of National's CompactRISC™ technology. It is being used in Sharp's new CL-500 and CL-550 cordless phone with answering system. These are the first digital answering machines to use Flash memory.

The NSAM265 offers the alternative to scarce ARAMs and it provides a direct interface to Flash memory. This way it don't requires a battery back up to keep the messages from being lost in the case of a power failure. It supports ARAM or standard 4Mbit and 8Mbit byte-wide Flash-devices, which allow up to 15 minutes of recording on a 4Mbit Flash device.

This device implements voice compression and decompression, tone detection and generation, message storage management, on-chip speech synthesis for time and day stamp and support for user-defined voice prompts in various languages. Echo cancellation techniques are implemented to support improved DTMF tone detection during message playback.

The price of the device is about $10 for quantities of 10.000 units.

10.2.3 NSAM266

The NSAM266 is the latest version of National Semiconductor's DTAD chips. This device has all of the functions of the NSAM265, but it also features an interface to serial Flash and Caller-ID support, based on the Bellcore standard used in the USA.

The NSAM266 compresses data at 4.8 kbit/s which allows a recording time of 15 minutes on 4Mbit memory. The interface to the external microcontroller is called MICROWIRE and consists of 3 wires and 3 control lines.

Demonstration Boards show the capabilities of the chipset. With the use of a DTAD-board and a DAA-board, both available from National Semiconductor, it is possible to demonstrate a working answering machine. A microphone, loudspeaker and power supply are the only things a designer needs to add.

Customization of CompactSPEECH is also possible. For development and prototyping of customized firmware On -Time-Programmable (OTP) part of the NSAM266 are available. This programming is done with a standard programmer and an additional adapter card.

10.3 Texas Instruments Mixed Signal Processor

Texas Instruments offers a two chip solution for Digital Tapeless Answering Systems. This mixed-signal, DSP-based chipset from Texas Instruments can incorporate voice mail features such as selective saving or deleting messages, skipping through messages or speeding up playback. The solution is provided by the MSP58C80 processor and the MSP58C20 analog front end. This solution is based on a TI TMS320C25 core to provide a low cost DSP solution.

Recently, TI launched its MSP80C85 processor. The following sections describe these two chips.

10.3.1 MSP58C80

The MSP58C20 is a delta-sigma converter which provides analog to digital conversion for the MSP58C80.

The MSP58C80 has advanced built-in DSP algorithms, including a 4.8kbit/s MCELP vocoder which can achieve close to 15 minutes recording on a 4Mbit ARAM. This DTAD must be driven by an external host controller, a powerful command set is used as software interface to the MSP58C80 for realization of a DTAD system. Listed below are the main features:

Main Features:

- Selectable compression MCELP 4.8 kbp/s or 7.2 kbp/s.
- ARAM support
- Real time clock
- Two 8-bit I/O ports
- Three 8-bits ADCs for event monitoring
- User programmable single/dual tone generator
- Call Progress Tone detect for US, Germany, France & UK.
- DTMF detect
- Ringing detect
- Day/ Time Stamp
- support up to 8 mailboxes
- Power Down Mode

The MSP58C81 is a derivative of the MSP58C80. The difference between the MSP58C81 and the MSP58C80 are listed below:

- The MSP58C81 I/O interface has more pins than the MSP58C80. This provides the extra I/O pins necessary to implement a software-driven external memory interface, without losing general-purpose I/O capability.
- The MSP58C81 does not have a hardware-driven external memory interface.
- The MSP85C81 serial port can supply hardware-generated CLKX and FSX signals.

10.3.2 MSP58C85

The MSP58C85 is the heart of a cost-effective digital telephone answering device (DTAD). It integrates a T320C2xLP digital signal processor core with DTAD-optimized ROM, RAM, and interface logic. With the addition of software, ARAM, and a combo, such as the DTAD3054, it makes a complete DTAD subsystem.

The MSP58C85 is very flexible. It can interface to 1 or 2 ARAMs/DRAMs, each of which can range in size from 256 kbit to 16 Mbit. It can operate with or without a host microprocessor. It can interface with Flash memories as large as 8 Mbit, without external glue logic, as well as interfacing with external ROM/SRAM/EPROM. It also incorporates an 8 bit analog-to-digital converter and very flexible power-down and low-power modes.

TI provides a complete selection of DTAD software with the MSP58C85, including 2.4 kbps and 4.8 kbps vocoders, DTMF/CPTD detection (Call Progress Tone Detection), and DTAD executives for the standalone and Host microprocessor modes of operation.

The MSP58C85 has a rich instruction set, flexible on-chip interfaces, and fast execution speed. Listed below are some of the potential applications that can be used in the smartphone:

- Digital Tapeless Answering Machine
- Voice mail
- Speech vocoding
- Speech recognition
- Echo cancellation
- Modems
- Adaptive equalizers
- DTMF decoding
- CLI
- FAX
- Speakerphone

Description

The MSP58C85 has a DSP core, which performs both signal processing and control functions. To support high-speed processing, it has 16 kwords of internal ROM, and 1312 words of internal RAM. It has a flexible clock control system, which allows the power consumption to be reduced during the execution of routines that do not require high-speed processing (see Figure 14).

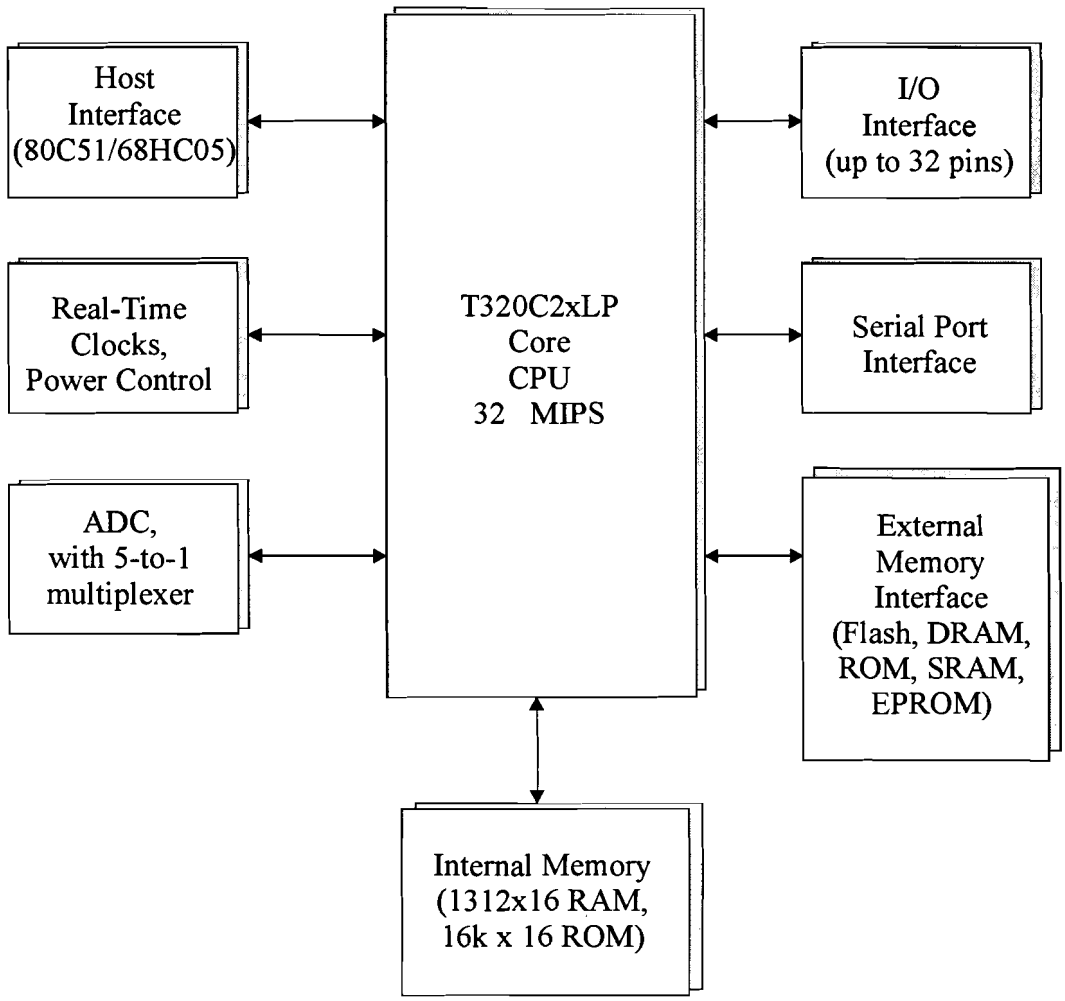


Figure 14: MSP58C85 Simplified Block Diagram

The external memory interface supports SRAM, ROM, EPROM, DRAM Flash, and memory mapped peripherals, allowing a wide range of system configurations with no external glue logic. The I/O interface provides up to 32 general-purpose I/O pins. The Host interface provides an easy-to-use parallel interface to a Host microprocessor, if one is needed.

The serial interface supports several analog interface chips for speech I/O. The successive-approximation ADC can be used for monitoring telephone line voltage, and other non-speech signals. The real-time clocks allow day/time stamping of messages, and the power control permits minimized power consumption. The Host interface is designed to interface to industry standard microprocessors, such as the 70C4X/370 family or the 80C51 or 68HC05 type.

Software packages

A software package is included with the MSP58C85 to facilitate the implementation of a digital telephone answering device. The software includes 2.4 kbps LPCII, as well as 4.8 kbps and 7.4 kbps MCELP speech analyses and synthesis routines. There is also an LPC synthesis routine that allows day/time stamp and voice-menuing functions to be implemented with pre-recorded, edited speech. The other routines included in the software packages are VAD (Voice Activity Detection), DTMF detection and generation, CPTD, autodisconnect, ring detection, RTC (real-time clock) handling, near-end echo cancellation, ARAM management, and an executive program that the consumer can tailor for specific hardware and user interfaces.

10.4 AT&T

AT&T also markets Digital Tapeless Answering Devices. Table 8 shows a feature support road map.

Table 8: Feature Support Road Map

		DSPhone™ Digital Speakerphone Solutions	
	ETAD	ECSP + ETAD	AECSP + ETAD
ARAM Solutions Grade A,B,C,D x1, x4, x8 4Mbit	LC30	LD30 +16 Mbit ARAM support	
FlashTAD™ **Flash Solutions** AMD/Intel Style Nor Flash Technology 4 Mbit & 8 Mbit	LE30		
Samsung/Toshiba Nand Flash Technology 4Mbit & 8 Mbit	LJ30	LH30	LF30

These devices are comparable to the ones described before. They also need ARAM or Flash, a controller and a Codec to form an all digital answering machine. These devices use either 6.8 kbps/5.44 kbps CELP+ or 6.1 kbps/5.0 kbps/4.2 kbps/3.4 kbps RCELP speech

compression with silence compression, allowing for rates down to 2.7 kbps. It give a storage time of 17 minutes with CELP+ and up to 25 minutes with RCELP in 4 Mbit.

The ECSP (Echo cancelling Speakerphone) supports near full-duplex speakerphone applications with a Hybrid Echo Canceller. The AECSP (Acoustic Echo Cancelling Speakerphone) supports full-duplex applications with both Acoustic and a Hybrid Echo Canceller.

All products incorporate DTMF detection and generation as well as call progress detection and tone generators. Also Day/Time Stamping and Powerdown Modes are supported.

11. ASIC

Besides the use of a general purpose DSP or a DTAD chip, there is another way of implementing digital functions. You can create your own custom processor. In this way it is possible to combine a DSP core with a set of peripherals, memory devices and your own proprietary logic to build a custom, integrated "system on a chip." Unfortunately this is not only difficult, but expensive. You may have to qualify for a vendor's core-based development program for which you typically have to pay over $125,000 and commit to building 100,000 units. The cost gets even higher if the ASIC vendor does your design work. The up-side though, is that you can reduce the cost of each system you build. Because Pijnenburg has a core-based development program at its disposal, these costs can drop, but it will still be expensive because of the long development trajectory.

A custom processor can reduce system cost in several ways. For one thing, the smaller process geometry's now available allow higher integration in silicon and reduce the number of board level components. Also, a custom processor may cost less than an off-the-shelf version, because you can optimize the custom processor's feature set for the target application.

Furthermore, higher integration most likely helps reduce power consumption, which lets you use a smaller power supply. Higher integration also adds the benefit of increased reliability. For example, you can reduce to one component a board-level design that otherwise would require 20 components with 16 to 64 pins requiring 350 connections. This reduction would yield a several-orders-of-magnitude increase in product reliability.

There are plenty of companies to turn to for help when you want to create your own processor. In Table 9 some of those companies are listed. Developing a custom processor is no small feat, though. To begin, you have to sort through a wide variety of cores and peripherals. After that comes the different task of connecting all the pieces.

In addition you must consider emulator and hardware-debugging support for cores. Core-based devices have neither standard pinouts nor standard functionality: The more integrated a design is, the fewer the pins that have to come out of it. This fact makes it impossible to develop an emulator that uses a probe to replace the target processor. So you must consider the emulators and debuggers for custom-processor designs.

As you begin to assemble your custom processor, note that most of the pieces, or modules are reusable, and you can interchange them among designs. The ability to reuse pieces reduces your design time. Also, integrating system functions onto one piece of silicon helps to maintain the confidentiality of your design: Competitors must reverse-engineer your device to determine your "system" components.

Table 9: List of companies who assist in designing ASICs.

ASIC/FPGA vendors offering cores	Core-only vendors	Emulators/debugger vendors
Altera Corp San Jose, CA (408) 894-7000	**DSP Group** Santa Clara, CA (408) 986-4315	**Applied Microsystems** Redmond, WA (206) 882-5326
LSI logic Corp Milpitas, CA (408) 433-8000	**Mentor Graphics** Wilsonville, OR (503) 685-7000	**Embedded Support Tools** Canton, MA (617) 828-5588
Motorola Inc Austin, TX (512) 891-2000	**SUN Microsystems** Mountain View, CA (408) 779-8119	**Huntsville Microsystems** Huntsville, AL (205) 881-6005
Texas Instruments Inc Denver, CO (800) 477-8924, ext 4500	**3Soft Corp** San Jose, CA (408)467-0410	**Metalink Corp** Chandler, AZ (602) 926-0797
VLSI technology San Jose, CA (408) 434-3000	**Alta Group** Foster City, CA (415) 574-5800	**Orion Instruments** Sunnyvale, CA (408) 747-0440
SGS-Thomson Phoenix, AZ (602) 867-6200	**Western Design Center** Mesa, AZ (602) 962-4545	**Quickturn Design** Mountain View, CA (415) 967-3300

Design reuse is a primary advantage of using hardware description languages to describe a core or peripheral. After describing a design in VHDL (Very High Description Language) or Verilog, you can - with the push of a button - target designs to those improvements. Designers can ensure that their code will migrate from one technology to another by using appropriate coding techniques.

Designers should not be concerned with the target technology, and their code should not incorporate technology-specific macros, such as adders and multipliers. Instead, the synthesizer should optimize for a technology, and designers should focus on implementing on a synchronous design. Avoid code that generates feedback loops, latches, tristate buffers, and asynchrone sets and resets, which makes it difficult to perform synthesis and timing analysis and reduce code portability.

The next two sections give an overview of the cores of two different core-vendors, the DSP Group and SGS-Thomson.

11.1 DSP cores from the DSP Group

DSP Group developed its own 16-bit fixed-point DSP core architectures, the Oak and Pine DSPCores, and licenses them to leading electronics and semiconductor manufacturers including Adaptec, Asahi, DSP Communications, GEC Plessy Semiconductors, Integrated Circuit Systems, Kenwood, Siemens AG, Silicon Systems, TEMIC(Daimler-Benz), Xicor, LSI Logic, and VLSI. Once again, using a core adds benefits over standard products. Oak and Pine cores support eight off-core, on-chip user-defined registers that appear in the data-register fields of all relevant instructions. In terms of the programming model these registers are part of the register set. You can build on-chip computation units, such as an FFT block (Fast Fourier Transforms) or a FIR (Finite Impulse Response) filter, around these registers. Loading these registers via an appropriate instruction directly loads the computation unit. At the end of the computation, you can load internal core registers with the resultant data in one cycle. You can also use these registers as interfaces to dedicated hardware, such as timers, serial/parallel ports, or a host port.

The DSP group provides emulation/evaluation boards that include general-purpose Oak and Pine emulation processors and a wire wrap area. These processors bring internal buses to pins on the chip, making all internal circuitry and timing easily accessible. A special on-chip emulation module hooks into the inner core. The core processors obtain information through these hooks and store it in registers that then transfer the information to the outside world. You can use an on-chip serial port on final-silicon versions to obtain this information.

11.2 DSP core from SGS-Thomson

SGS-Thomson's contribution to 16-bit, fixed point DSP's is the D950 core. The key to the core's performance is parallelism, which lets the D950 core perform multicycle functions simultaneously with other processors. The company supports this core with an ASIC-development program and a variety of peripherals. Although the core is inflexible, you can use its coprocessor interface to add processing capability, such as a complex multiplication cell. When the D950 core detects a dedicated coprocessor instruction, the core asserts the valid-coprocessor-instruction (VCI) signal to the coprocessor. The coprocessor than goes to the instruction address and executes the instruction.

The core contains an emulation-and-test unit for JTAG compatibility, SGS-Thomson provides a JTAG-emulation board with a graphical, high level source debugger that connects to your PC through a serial link with a JTAG test-access-port (TAP) controller. To further support development efforts, SGS-Thomson provides the ST18951 test/emulator chip. The device contains interrupt and DMA controllers, a TAP, a bus-switch unit, and a coprocessor. You can use this chip to debug your code in parallel with chip development.

A company that renders assistance by the making of DSP based cores is the CME (Center for Micro-Electronics) in the Netherlands. The next section describes this company.

11.3 Center for Micro-Electronics (CME)

The design of ASICs is executed at an ever increasing level (behavioral-, system-level) because the complexity of the electronic circuits has increased enormous. Within the JESSI AC75 project, CAD software has been developed, the **DSP Station** of Mentor Graphics, for the design and analysis of complex, digital signal processing applications. The tools of the DSP Station admit to examine very effectively, the Design Space (area, speed, and power consumption) without having to be an expert in the processing of megabytes of VHDL files.

CME services

As a part of the JESSI AC75 Project, the CME has setup a DSP Support Center. The purpose of this center is to make the tools of Mentor Graphics' DSP Station available to middle-sized and small companies (MSC). This way the JESSI results of the AC75 project will pass to the MSC industry. Small concerns can create very advanced designs with these tools without the need for very high (risky) investments. The use of this system costs $50 per hour and $100 when you want technical assistance.

Design Flow

Figure 15 shows the design flow, which is now available at the CME. This covers the complete path from behavior-specification to ASIC layout or assembly code for commercial DSP processors.

Behavioral-level Specification

Filter Architect lets you design the most accepted filter structures with all classical IIR (Infinite Impulse Response) and FIR (Finite Impulse Response) approximation algorithms. For system-simulation, a large collection of telecommunication-blocks is available for mobile radio and wireless communication. With this a complete system can be configured and tested very fast. Further more, all digital signal processing algorithms can directly be specified in DFL (Data Flow Language). This is an applicative language in which parallel and bit-true behavior can be defined in a simple way.

Circuit Analysis

The analysis tools of the DSP Station are unique in their way. Next to the well known time-simulations (on high level as well as on bit-true level) and Discrete/Fast Fourier Transformations, the frequency can also be simulated directly. Further more, coefficients can be optimized for instance for finite wordsizes and fixed-point implementations. Noise and Templates calculates signal-to-noise ratios and possible overflow conditions for every vertex in the circuit. This admits very efficient dimensioning of the circuit. Finally, (worst case) limit cycles can be found for fixed-point conditions, on small scale (as a result of quantization-effects) as well as on large scale (by overflow oscillations).

Hardware Mapping

The DSP Station contains a number of true architectural synthesis tools for the translation of the (behavioral)-specification to a specific target architecture. Mistral 1 creates a bitserial architecture for medium-speed applications. Mistral 2 synthesizes a micro-programmed bit-parallel architecture, specially designed for medium-speed and intense decision algorithms. The result of this architecture-synthesis is a VHDL, EDIF, GN or N netlist, which can be used for logical synthesis.

It is also possible to translate the algorithm to assembly code for an existing general purpose DSP-processor of Motorola (M56 family) and Texas Instruments (C30 family). The quality of this assembly code is comparable with manually generated code, because the parallelism, completely specified in DFL, can be fully utilized.

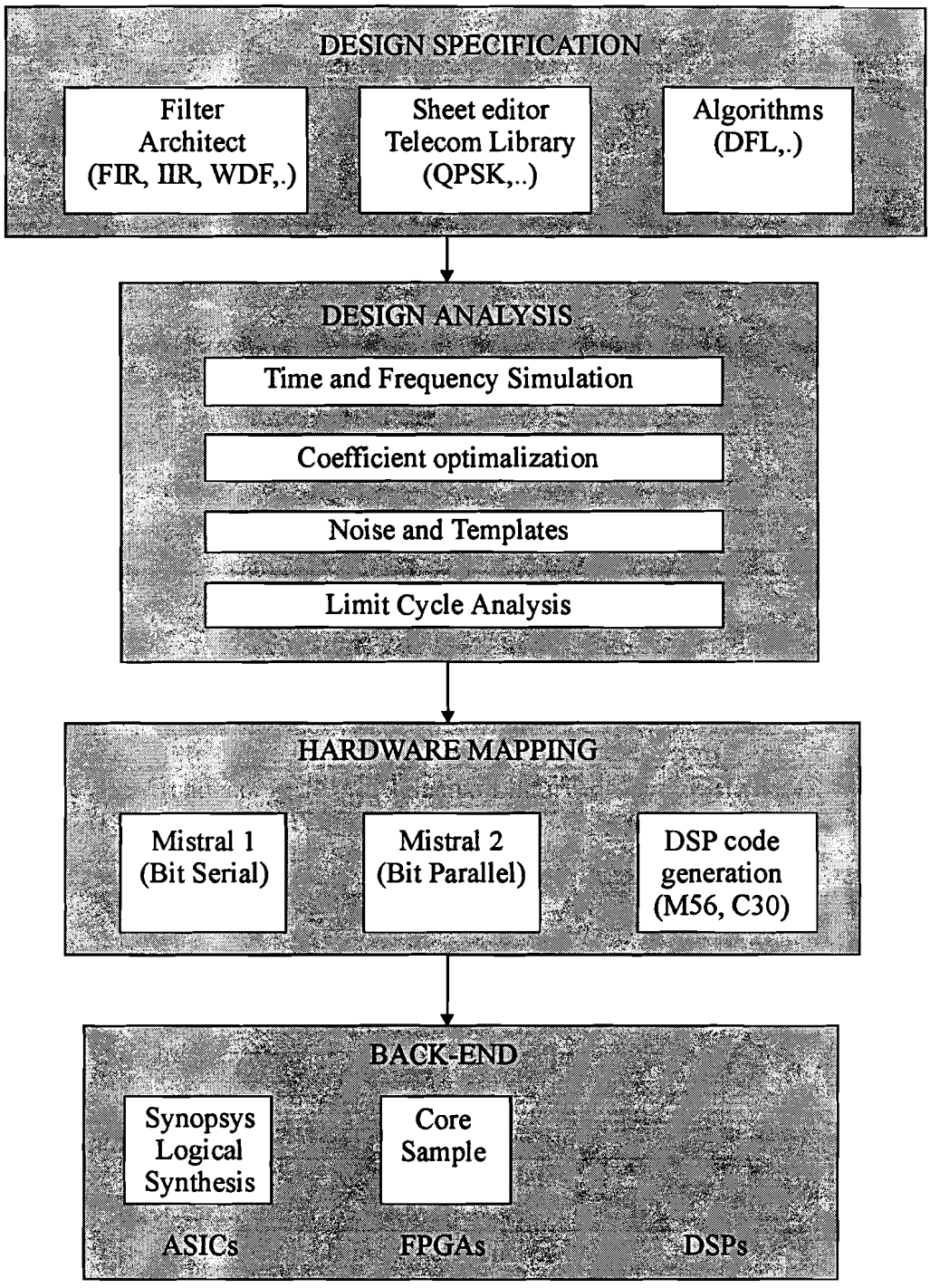


Figure 15: Design Flow

VHDL Synthesis - Back-End

The generated VHDL code, of the Hardware Mapping section, can be further processed within the company Pijnenburg itself. Here the necessary hard- and software to process Synopsys VHDL is present.

Also the back-end of the designflow is supported by the ChipShop production facility. These is a path available through the Fraunhofer Institute in Germany. They can perform logical synthesis with Synopsys and generate layouts. After that, the chips can be processed, using a Multi-Project-Wafer manufacture-path of Nordic VLSI in the AMS 0.8 µm CMOS double metal process.

12. Overview of the peripherals and their specifications

The digital part has to interface to the peripherals. This chapter describes these peripherals and their specifications.

12.1 CODECs

The DSP signals are interfaced to the analog world via CODECs (coder/decoder) chips or linear analog-to-digital (A/D) converters and digital-to-analog (D/A) converters. CODEC chips contain all the necessary A/D, D/A, sampling and filtering circuitry for a bi-directional analog/digital interface. The CODECs with on-chip filtering are sometimes called CODEC / filter combo chips, or combo chips for short.

The CODEC channel used in telephones are bandlimited to pass only frequencies between 200 Hz and 3400 Hz. Some CODECs also incorporate companding (audio compression / expanding) circuitry for either of the two companding standards: A-law and μ-Law (see section 6.2.1).

12.2 Controller

The controller of the system is different for the three possible solutions described before. The next sections explain these differences.

12.2.1 DTAD

When using a DTAD, all the speech processing tasks are done by the DSP Speech Subsystem. This allows the use of a very low cost microcontroller to be used for basic control of the system. The Host need only send high level commands to perform functions such as *Record Message* or *Delete Message* and the operation will be performed by the DTAD which will report the status of the operation to the Host. All memory interface and management will be taken care of by the DTAD requiring the Host to only handle high level system functions. A summary of the functions performed by the DTAD and Host controller are shown in Figure 16.

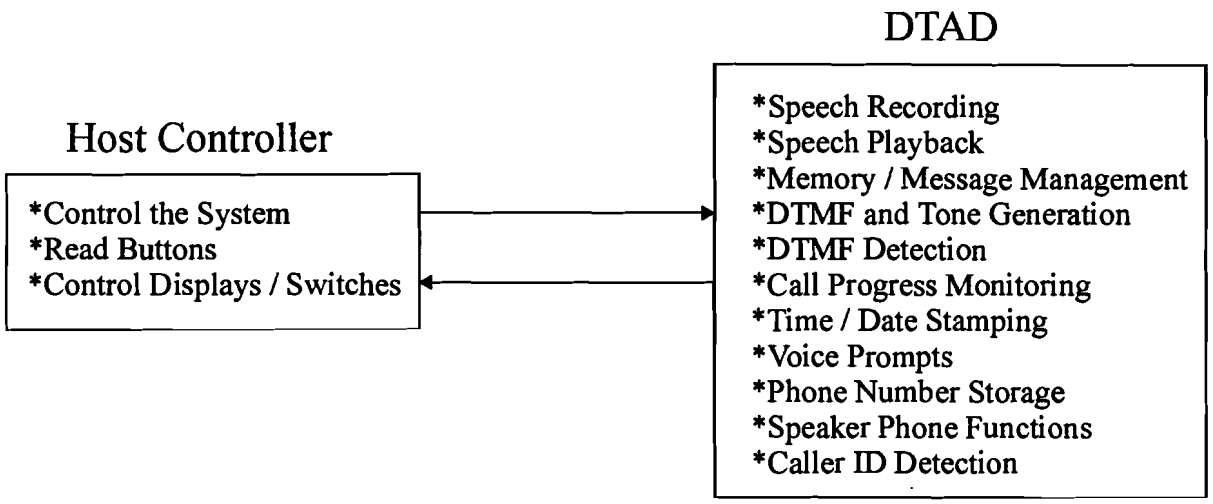


Figure 16: System Functions

The hardware interface between the DTAD and the Host Controller is a simple one requiring only an 8 bit parallel port and 4 handshake lines. The Host writes high level commands to the DTAD and the DTAD will respond with Status information. Once a Command is issued the DTAD will use the ACK pin to indicate that the Status is available to be read.

12.2.2 General Purpose DSP

There are two ways to use a General Purpose DSP in this application:

1. Stand-Alone DSP
2. Co-Processor

In the first case both the Digital Signal Processing routines and the control software run on the DSP. As we have seen in Chapter 9, not all DSPs have enough ports to be able to control all the peripherals. Also, controlling functions use up processing power, so that bigger and more expensive DSPs are needed. The advantage is though that no extra controller is needed which consumes power and adds an extra component to the circuit board.

In the second case a separate microcontroller is needed. This device controls the peripherals and tells the DSP which Signal Processing task to execute. The main advantage of this configuration is that you can use less powerful DSPs, which saves total system cost.

When large portions of memory are needed or Flash memory is used for storage of the messages, this memory is connected to the microcontroller in stead of the DSP. During playback or recording of a message, the DSP generates a data flow of about 8 kbps, depending on the compression algorithm used. The microcontroller should be able to read the data, add some error correction bits to it, and store it at the same time. This means that a contoller should be used which is capable of transporting at least 17 kbps, which equals 2.2 kbyte/s.

12.2.3 ASIC

When creating an ASIC, you can integrate all functions on one chip. This way, there is no need for external peripherals like Codecs and microcontrollers. Only external memory, containing the control program and storage ARAM or Flash needs to be added.

12.3 Flash

The answering machine should be able to maintain the incoming and outgoing messages without using a battery back-up. This is the reason why Flash memory must be used in stead of ARAM within this application. Also there is a great shortage of ARAM devices at the moment.

Flash memories have revolutionized the way system designers store digital control code and data. Flash memory technology offers a unique combination of features - non-volatility, in-system reprogrammability, and high density - that makes it the fastest growing IC memory type today. There are two kinds of Flash: NOR- and NAND Flash. NAND offers a smaller memory cell, more efficient operation and lower manufacturing costs than NOR architecture Flash.

Flash memory has a slower programming speed, but this is not an issue in this application because digital TADs record messages sequentially at a bit rate of less than 4 kB per second.

Another disadvantage of Flash is the number of times it can be erased and programmed. That is why it is best to write and erase each block more or less the same number of times to ensure that all blocks have the same lifetime. Flash is build up in blocks of pages, each containing a number of bytes. A block is the smallest unit that can be erased. Most Flash

devices can be programmed and erased for about a 100,000 times. Consider the following extensive usage of all the blocks in a 4 Mbit Flash device:

1. Record 15 minutes of time (until the memory is full)
2. Playback 15 minutes (all the recorded messages)
3. Delete all messages

Assuming the Flash device is used in this manner 24 times a day, its expected lifetime is:

- Flash Lifetime = 100,000 / (24 * 365) = 11.4 years.

An 8 Mbit device, used in the same way, will last for over 20 years.

At this moment, the companies Samsung, AMD, Toshiba, National Semiconductor and Intel are producing Flash memories. The following sections describe the Flash devices from AMD and National Semiconductor, which can be used in combinations with DTAD devices.

12.3.1 AMD

AMD has developed a single-power-supply architecture for Flash memory and incorporates this architecture into its 5.0 Volt-only Am29Fxxx family and 3.0 Volt-only Am29LVxxx family of Flash memory devices. The Am29Fxxx 5.0 Volt-only family is the current industry standard for Flash architecture. The Am29Lxxx 3.0 Volt-only family provides Flash memory devices that are ideal for battery-powered system designs. AMD's Flash memory architecture includes features as fast access times, sector organization, simplified program and erase algorithms, and high endurance. AMD's proprietary Negative Gate Erase technology yields single-power-supply Flash memory devices with negligible impact on die size when compared to dual-power-supply Flash memory devices.

The Am29Fxxx 5.0 Volt-only Flash Memory

Features	Benefits
5.0V only single-power-supply	• Eliminates need for DC /DC converter
Embedded Algorithms	• Reduce time-to market • Simplifies program/erase software • Reduces program/erase software
100,000 program/erase cycle endurance (guaranteed minimum)	• Provides lowest failure rate on every program/erase cycle • Provides high reliability
Fast access time	• Increase system performance • Reduces or eliminates wait states

The Am29Fxxx 3.0 Volt-only Flash Memory

Features	Benefits
3.0V only single-power-supply	• Eliminates need for DC/DC converter
2.7V to 3.6V extended operating range	• Allows longer battery life in portable battery-power-systems
Embedded Algorithms	• Reduce time-to-market • Simplifies program/erase software • Reduces program/erase software
100,000 program/erase cycle endurance (guaranteed minimum)	• Provides lowest failure rate on every program/erase cycle • Provides high reliability
Fast Access time	• Increase system performance • Reduces or eliminates wait states

AMD's FLASH Embedded Algorithms

Flash offers a cost effective single transistor cell architecture that programs similar to an EPROM using hot electron injection and erases similar to an EEPROM using Fowler-Nordheim tunneling. First generation Flash devices required extensive firmware to simply program or erase the device. Embedded Algorithms not only simplify the program and erase

algorithm implementation, but offer superior reliability with a guaranteed minimum of 100,000 write cycles of endurance. Write endurance is defined as the number of times that a Flash device can be programmed and erased without failure. AMD's Embedded Program™ and Embedded Erase™ Algorithms offer simplified program and erase by including software in the silicon. The AM29LV800 offers access times of 90 ns, 120 ns, and 150 ns. A Sector of this device is typically erased and verified within 1 second.

12.3.2 National Semiconductor

National Semiconductor offers two serial Flash memory devices: the NMS29A040 (4 Mbit) and NMS29A080 (8 Mbit).

The NMS29A040 is organized as 128 blocks of 128 pages, each containing 32 bytes. A block is the smallest unit that can be erased, and is 4 kbytes in size. Not all 128 blocks are available for recording. Up to 10 blocks may contain bad bits, and one block is write-once and holds the locations of these unusable blocks.

The NMS29A080 is organized as 256 blocks of 128 pages, each containing 32 bytes. Up to 20 blocks may contain bad bits, and two blocks are write-once and hold the locations of the unusable blocks. This serial Flash may be erased up to 100,000 times.

12.4 Telephone line interface

The Telephone Line Interface, also called DAA (Data Access Arrangement) provides the final connection between the smartphone and the telephone network. Depending on the network in use, the DAA may provide an analog interface to the public-switched telephone network (PSTN), use a digital interface to the ISDN or connect to a cellular transmitter. It serves as an interface between the smartphone and the telephone network.

The DAA includes line protection devices, a line transformer, dialing and audio circuitry, a hook and control switches, and detectors.

13. Overview of the Algorithms and Vendors

As a result of the DSP integration, software becomes a vital part in future telephones. The functions are executed by software programs like modem emulation, DTMF coding and decoding, line (hybrid) echo cancelling, acoustic echo cancelling, and call-progress control. For the implementation of an answering machine, functions like speech coding en decoding should be executed. The system may also include advanced features, such as text-to-speech conversion and voice-recognition.

Many of these software blocks are available in libraries. Some libraries come directly from the DSP chip manufacturer. You can also obtain software from DSP development-tool vendors, such as Spectron and Spectrum Signal Processing, and from independent software vendors, such as VoCal Technology, GAO and DSP Software Engineering Inc. Table 10 lists some of the representative computer-telephony manufacturers. The following sections give an overview of the software specifications and prices from some of these companies.

Table 10: Representative computer-telephony manufacturers

Product Type	**Manufacturers**
Algorithm vendors	DSP Software Engineering GAO Research & Consulting HotHaus Technologies VoCal Technologies
Application software	Active Voice Applied Voice Technology Delrina
DSP tools and libraries	Analogic Corp Spectron Microsystems Spectrum Signal Processing
DSP/data-pump chips	Analog Devices AT&T Microelectronics Cirrus Logic Motorola Texas Instruments
Telephony boards	Analogic Ariel Dialogic

13.1 GAO

GAO Research & Consulting Ltd. (together with its subsidiary Trisoft Technology Corporation) is a supplier of digital signal processing algorithms, software and hardware modules, and custom software and hardware design services. Its customers are primarily telecom and telephony original equipment manufacturers (OEMs). Customers of GAO's products and services are from all over the world, including Japan, USA, Canada, Belgium, South Africa, Germany, Singapore, Korea and Switzerland. GAO was founded and has been managed by Dr. Frank X.Y. Gao, an internationally well-known DSP expert.

13.1.1 Algorithms and Software Products

The highly optimized DSP algorithms and software modules provided by GAO can serve as easy-to-use and off-the-shelf building components of OEM products which use the general-purpose programmable fixed- or floating-point digital signal processors of Analog Devices (see section 9.3.1). An OEM can purchase or license these algorithms and software modules to design a product, or can request GAO to custom design a product based on these algorithms and software components. At this moment GAO is offering the following software modules for implementation on the ADSP fixed point DSPs:

Modem (ITU Standards)

V.21 (300 bps)
V.22 (1200 bps)
V.22bis (2400 bps)
V.32 (9600 bps)
V.32bis (14.400 bps)
V.34 (28.800 bps)
V.42 (Error Correction)
AT Command Set

Telephony

Acoustic Echo Canceller (**AEC**)
Automatic Gain Control (**AGC**)
Line Echo Canceller (**LEC**)
DTMF Tone Generation
and detection
Caller ID
Call Progress Monitoring (**CP**)

Speech

Time Scale Modification
Voice Activity Detection
Sample Rate Conversion

Speech Codecs

ITU G.728 (16 kbps)
RPELTP (13 kbps)
G.721 (32 kbps)
G.711 (64 kbps)
G.722 (64 kbps at 16 kHz)

Table 11 gives up-to date price information about some of the interesting algorithms.

Table 11: License Fee and Cumulative Unit Production Royalty For Telephony Functions (For Production Volume of At Least 5,000)

Product	License Fee	1-5k	5k-20k	20k-100k	100k-1M	1M+
V.22	12,000.00	0.95	0.50	0.20	0.10	0.05
V.22bis	19,000.00	0.95	0.50	0.30	0.20	0.10
AEC	15,000.00	2.5	1.95	0.85	0.65	0.25
AGC	9,000.00	2.5	1.95	0.85	0.65	0.25
LEC	14,000.00	2.5	1.95	0.85	0.65	0.25
Caller ID	16,000.00	2.5	1.95	0.85	0.65	0.25
DTMF	16,000.00	2.5	1.95	0.85	0.65	0.25
CP	16,000.00	2.5	1.95	0.85	0.65	0.25
PSM*	12,000.00	2.5	0.95	0.55	0.45	0.15

* Playback Speed Modification

Table 12 gives an overview of how many MIPS and memory each algorithm uses.

Table 12: Specifications of the Algorithms

Function	Average MIPS	Data Memory	Program Memory
V.22	5.0	0.8k	1.5k
V.22bis	6.0	1.0k	1.5k
DTMF decoder	0.7	350	500
DTMF generator	0.7	160	140
Caller ID	1.0	0.2k	0.5k
Call Progress Detector	0.69	60	200
LEC	5	400	700
PSM	2.5	300	1,000
G.728 Encoder	20	4k	6k
G.728 Decoder	12	3k	4k
AGC	0.1	80	440
Acoustic Echo Cancel	*	*	*

* The memory and MIPS requirements of an acoustic echo canceller highly depend on the maximum duration of all possible echoes. The longer the duration of an echo, the more MIPS and memory it requires. Assume the maximum echo duration is 80ms. Echo with this duration may be generated by a room with 4 meters by 4 meters hard wall. To cancel such an echo, this algorithm needs about 16 MIPS, 1k words of data memory, and 1.5k-words of program memory.

13.1.2 GAO Soft Phones

GAO Soft Phones are software driven and DSP-based phones which integrate speakerphone, caller id, answering machine, fax, & modem. This system is designed to be licensed to telephony and telecommunication equipment manufacturers, who would like to manufacture a digital telephone or a family of digital telephones with advanced features such as acoustic echo cancellation, which allows high-quality speakerphone function. Major components of the basic configuration include, an ADSP2105/1, SRAM, codecs, LCD, keypad, and PCB (see Figure 17).

These products are called Soft Phones because they share one basic configuration and their functions are all implemented in software. The products can be easily upgraded by downloading the software of the digital signal processor or replacing a ROM containing the software.

The products of GAO have been well tested under various test conditions and in practical systems. Whenever possible, industry standard tests, such as MITEL DTMF Detection test tape, ITU Acoustical Echo Cancellation Test Standard G.167, are used. The products meet or exceed the requirements.

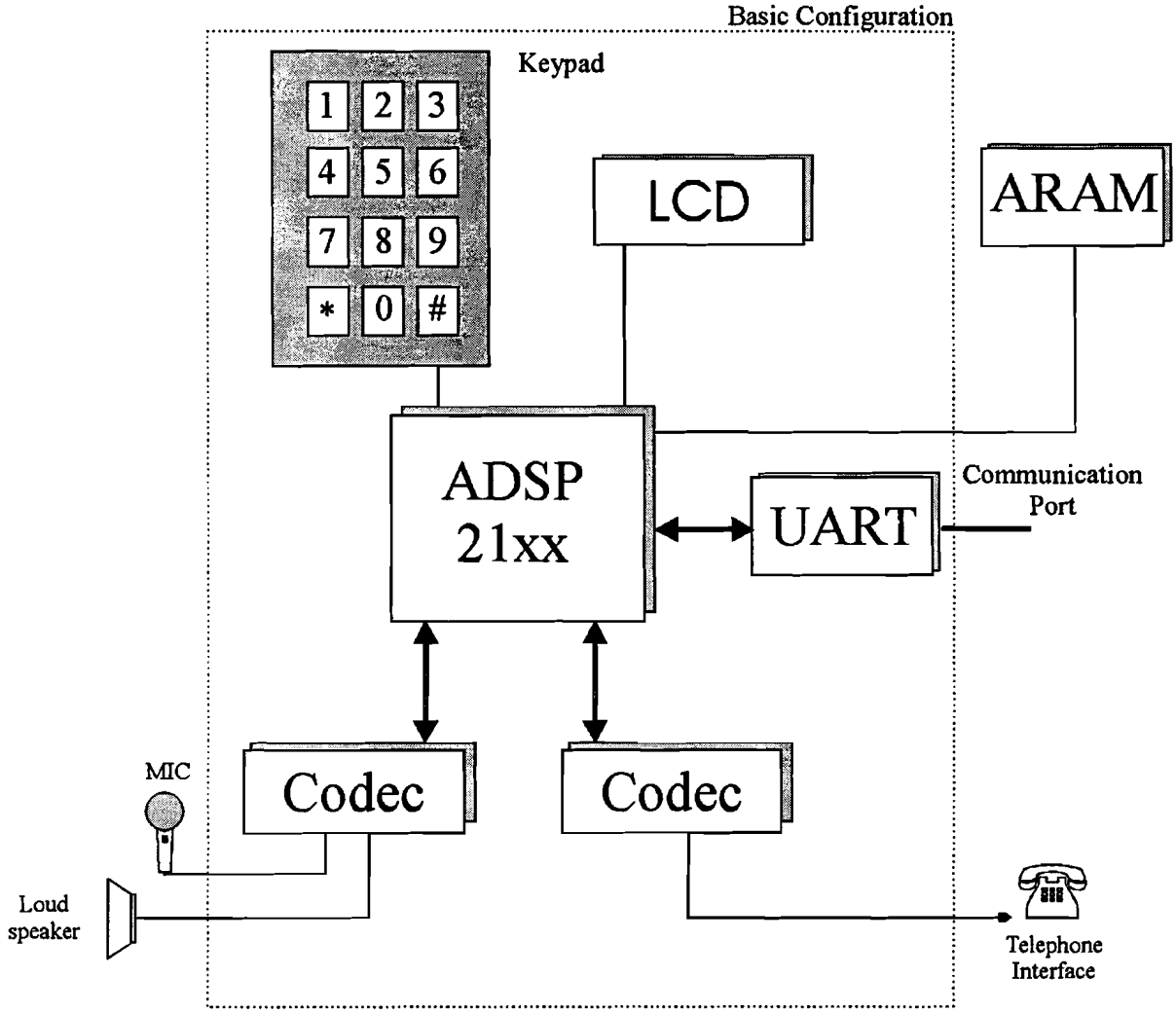


Figure 17: Hardware Configuration

For OEM purposes, the board can be configured or modified into two platforms: a basic configuration, and the basic configuration with ARAM (low-cost DRAM). Table 13 gives an overview of the products based on the Basic Configuration.

Table 13: products based on a Basic Configuration

Hardware Configuration	Component Cost (US)	Products
Basic Configuration	$22	1. Digital Speakerphone 2. Digital Speakerphone and Answering Machine (Messages stored in a PC) 3. Digital Speakerphone and Modem 4. Digital Speakerphone, Modem, Fax and Answering Machine (Messages stored in a PC)
Basic Configuration + ARAM	$22 + ARAM Cost	5. Digital Speakerphone and Answering Machine 6. Digital Speakerphone, Modem, Fax, and Answering Machine

Note:

Products 1 and 5 are standalone products which do not need to send data to a PC or receive data from a PC. The communication port may be used for software upgrading. The port is optional to these products since software upgrading can be done by replacing a ROM manually.

The difference between products 1, 2, 3, and 4 and between products 5 and 6 lie only in software. In the case of products 1 to 4, an OEM can build one hardware platform and market it under four different products. On the other hand, a customer initially purchase a product with less functions. When he wants more functions, he can upgrade his product by simply replacing its software in a ROM or computer diskette purchased from the manufacturer.

Products 1 to 6 listed above employ a combination of the following software:

Digital Speaker phone

Acoustic Echo Canceller
Caller Identification
DTMF Tone generation
Call Progress Monitoring
Analog/Digital Automatic Gain Control
Line Echo Canceller

Modem Software

V.21, V.22, V.22bis, V.32, V.32bis, and AT Command Set (V.34 later)

Answering Machine Software
Playback Speed Modification
Speech Compression
Digital Automatic Gain Control
DTMF Tone Detection

Control Software
LCD, Keypad, and Other Functions

Optional Software
Voice Disguising
Home Security Functions

Every manufacturer has its own requirements. Based on the designs described above, GAO can make custom modifications to fit specific needs. GAO has developed engineering and production prototypes which can demonstrate all of the functions of the six products discussed above. The user interface (control software) is not very sophisticated since the software highly depends specific requirements of an OEM. GAO works closely with OEMs to define and implement the user interface functions. It takes a couple of months.

13.2 DSP Software Engineering, Inc.

DSP Software Engineering, Inc. (DSPE) specializes in the development of complex digital signal processing algorithms. DSPSE specializes in the development of standard algorithms, providing an implementation of the standard algorithm that runs on popular DSP chips. DSPSE thoroughly tests each implementation for conformance with the standard. By using DSPSE's standard building blocks, you save development time and test time.

DSPE is a world leader in DTAD applications. They currently service 60% of the world wide DTAD solutions. At the moment they are working on a variety of programs based upon Texas Instruments MPSA58C8x series DTAD-ICs (see section 10.3). By combining software components, integrated solutions, and expert consulting, DSPE provides a complete solution to DSP applications.

13.2.1 DSPE Software Components

DSPE software components are DSP communication algorithms implemented and optimized for TI fixed and floating point DSPs (TMS320C5x and TMS320C3x). These algorithms are application building blocks from which DSP products and systems are built. DSPE's solutions offer fast, cost effective DSP market access for developers with or without DSP expertise. The software components are application libraries with minimized integration interfaces. The components are processor specific, platform independent, and interface with arrays of samples and/or bits via pointers. The advanced DSP functionality of each component is accessed by one or two function calls. This "black box" approach provides advanced DSP technology with minimal integration effort. The engineers at DSPSE have extensive experience in telecommunications and software technology. The same engineers that design the algorithm implementations also provide quality customer support.

All software component products are priced with a one-time license fee. License fees are paid at the time of delivery and range from $3,000 to $75,000. Production rights are purchased with royalty payments. Right for the first 1000 units of production are provided at no cost.

Table 14 gives a fixed-point software component product summary.

Table 14: Fixed-point software components for TMS320C5x

Function	MIPS	Data (16 bits)	program (16 bits)
USFS 1015 LPC10e 2.4kbps	8.7	4.8k	7.4k
ITU G.728 LD-CELP 16kbps	36.9	2.2k	10.5k
ITU V.22bis Modem 2.4, 1.2 kbps	9	6k	20k
Telco Call Progress Detector	2.2	0	512
Telco DTMF Generator / Detector	3	56	819
Bellcore Caller ID	8.8	4.2k	11k
DSPE Line Hybrid Echo Canceller	2.7	0	256

13.2.2 SoftTAD Software

DSPE's Telephone Answering Device software (SoftTAD) is an integrated portable solution targeted for the consumer market. SoftTAD provides high quality voice message storage and retrieval as well as standard telephone signalling software integrated with an answering machine application. This software has been fielded by major Telephone Answering Machine manufacturers in the US and European consumer markets.

The design methodology and software architecture used in the development of SoftTAD enables quick porting to other consumer platforms or applications. The application can be easily modified without changing the real-time signal processing software.

13.3 HotHaus Technologies Inc.

HotHaus is a supplier of DSP software for telecommunications applications. They work exclusively with Texas Instruments fixed point devices, and have an array of software primarily for the C5x and C25 families. Also supported are the C203, C54x, and in some cases the older C17.

HotHaus offers a complete system architecture designed specifically for telecommunications devices, called HausWare. HausWare combined with some specific algorithms will likely be all the DSP software required. HausWare virtualizes the hardware interface to the lines, provides a BIOS to interface to the specifics of your hardware, includes a scheduler to control task execution on the DSP, a device manager to control interfacing between the host application and the virtual device drivers, and a well documented Application Programmers Interface (API) to the host application. HotHaus will perform the BIOS port, so that coding will be restricted to writing the host application.

The following configuration is proposed by HotHaus for our product.

Algorithm	Price ($US)	MIPS/MEM	Availability
HausWare - includes scheduler, device manager, switchboard, media player (a device driver for a speaker/microphone), BIOS	$5,000		Now
BIOS Port to Pijnenburg Hardware	N/C	-	3 weeks
PSTN Driver - includes DTMF detection/generation, Call progress monitoring and Caller ID services	$5,000	3.5 / 4.5k	Now
V.22bis - includes fallback to V.22	$15,000	4.5 / 5k	Now
Acoustic Echo Cancellation	$75,000	15 / 2k	6 weeks
G.721 ADPCM (32 kbps) speech compression	$7,500	5 tx, 8 rx / 2k	Now
VR CELP speech compression			
8 kbps	$45,000	11 /18k	May 1996
4.8 kbps	$45,000	12 / 18k	May 1996

HausWare is an innovative approach to DSP firmware for advanced telecommunications products. It is a new kind of product, more than an operating system and more than just a collection of telecommunications utilities. It is a complete *DSP software system architecture* designed specifically for DSP based telecommunications products. Being purpose-built for telecommunications means HausWare is very MIPS and Memory efficient, and also provides all the services required in telecommunications products, from low level line control and call progress monitoring to host access via well defined high level standard API calls.

Figure 18 shows the HausWare architecture.

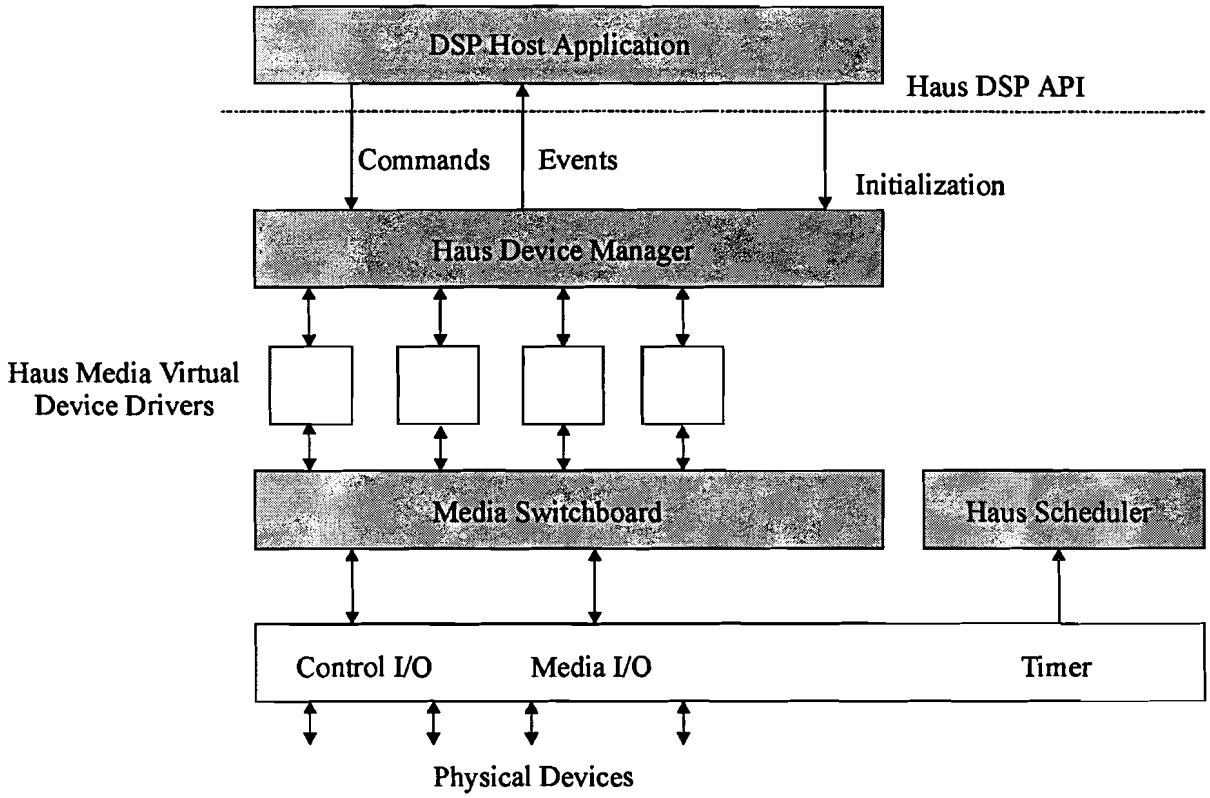


Figure 18: HausWare Architecture

HausWare can be used virtually without modification in feature phones, digital telephone answering devices, and a host of other similar products. It applies high level software engineering concepts to the real time DSP software environment. High level application programmers will be familiar with its model of virtual device drivers with defined attributes that are controlled via the HotHaus API with typical API commands such as *create, open, close, read, write, etc.* With a modular design it is easily extensible to add additional types of telecommunications device drivers.

The HausWare architecture accepts commands from a host DSP application (which would typically be responsible for communicating with a microcontroller) and control the operation. Physical devices include channel interfaces, voice and audio interfaces (speaker/microphone), and various other devices such as analog/digital, recording equipment, and analog debug ports.

HauseWare includes an array of standard services to operate on any given channel, including echo cancellation, DTMF, call progress, etc. Optional services include modems, caller id, speakerphone, etc.

The design process decomposes into the following primary tasks:

1. design the hardware
2. port the BIOS to the hardware
3. write the DSP host function
4. write the Host microcontroller application software.

14. Conclusions

DTAD

DTADs are ready-to-go processors which save development time and research cost. Also they are relatively cheap. This is why they are the most interesting chips for this application. The main disadvantage is that they are not expandable in an easy way. Newer software routines can only be added in consultation with the DTAD vendor.

General Purpose DSPs

There are a lot of different general purpose DSPs and a lot of different vendors. There are two kinds of DSPs: fixed-point and floating-point. Floating-point DSPs are more accurate and exclude the appearance of limit cycles due to rounding and overflow. They are however very expensive and therefore not suitable for this application. The fixed-point DSPs are much cheaper and it is possible with the right software to overcome the problem of limit cycles. It takes a deeper insight in digital signal processing techniques.

The main advantage of using general purpose DSPs is that you can write all software yourself, and buy only those algorithms that are complex, like Speech Compression and Echo Cancellation. This way, you can add new functionality to the smartphone at any moment. The disadvantages are that it takes a lot of time to develop an algorithm, the chips are at the moment more expensive than DTADs and the complex algorithms are to expensive yet.

ASICs

When you create an ASICs you have the possibility to combine a DSP core with a set of peripherals, memory devices and your own proprietary logic. This way, the number of chips in the device decreases, and so you can reduce the cost of each system you build. Unfortunately it is difficult and expensive because of the long development trajectory, because you have to develop the architecture first, then write the digital signal processing software and finally write the control software.

The Microcontroller

The microcontroller controls the peripherals and the functionality of the DSP. When using Flash or large portions of memory, coupled to the controller in stead of the DSP, the data

transport of the compressed messages passes through the microcontroller. This gives a lower bound on the speed of this controller. When using compression algorithms of about 8 kbps, the controller should be able to transport about 2.2 kbyte/s.

Flash Memory

Because the answering machine should be able to maintain the messages without using a battery back-up, and due to the fact that there is a shortage of ARAM devices, Flash memory offers a perfect solution. It has a slower programming speed, but since DTADs record messages at a bit rate of less than 4 kB per second, this is not an issue in this application.

Flash can be erased and programmed for about 100,000 times, which assures a minimum lifetime of over 11 years for a 4 Mbit device, when used extensively.

Flash has a different architecture than normal memory devices. It contains blocks of pages, and only a complete block can be erased. Because the erasing algorithm is not easy, current Flash devices have Embedded Erase and Programming Algorithms.

DTMF generation

As we have seen, the DSP will perform the DTMF decoding. The generation of DTMF signals though, must be executed by another chip, because when a power failure occurs, the DSP will no longer function and dialing will no longer be possible. This extra chip, can operate on the power generated by the telephone network.

15. Recommendations

Considering the prices for General Purpose DSPs and the third party software, the conclusion is that we should best use a DTAD from the DSP-Group, or the MSP58C85 from Texas Instruments. They are relatively cheap and there are possibilities for the implementation of newer features like CAS detection.

When this project continues, the next student could be given the assignment to develop a completely working answering machine with full-duplex handsfree capabilities, using one of the DTADs described before. After that, software algorithms should be developed to complete the smartphone, like CAS-detection, Modem-function and CLI.

16. Company Addresses

HotHaus Technology, Inc.
Suite 1B - 7218 Progress Way, Delta, British Columbia, Canada.
Tel: (604) 946-0060, fax: (604) 946-5811, email: info@hothaus.wimsey.com

DSP-Group, Inc.
3120 Scott Boulevard, Santa Clara, CA 95054, USA.
Tel: (408) 986-4300, fax: (408) 986-4490, email: mgrozenski@dspg.com, Mary Grozenski (Sales Administrator).

European Office:
DSP-Semiconductor, Givat Shmuel, Israel, tel: (972) 3 531-3241
Nogatech Ltd, Givat Shmuel, Israel, tel: (972) 3 531-3253

Ensigma Ltd.
Turing House, Station Road, Chepstow, Gwent NP6 5PB, United Kingdom, tel: (44) 291 625-422, fax: (44) 291 620-301.

GAO Research & Consulting Ltd.
Primary Office: 55 Nugget Avenue, Unit 204, Scarborough, Ontario, Canada M1S 3L1, tel: (416) 292-0038, fax: (416) 292-2364, e-mail: gao@io.org.

Secondary Office and Mailing Address: 85 Dundalk Drive, Scarborough, Toronto, Ontario, Canada M1P 4V1

AIM High Technology.
85 Dundalk Drive, Scarborough, Toronto, Ontario, Canada M1P 4V1, tel: (416) 292-0038, fax: (416) 292-2364, email: aim@io.org.

Spectrum Signal Processing.
9 Village Circle, # 450, Westlake, TX 76262, tel (817)430-5840, fax (817) 459-7331.

Arcobel Industrial Electronics bv. (LSI products).
Postbus 344, 5340 AH OSS, The Netherlands, tel: 04120-42433, fax: 04120-30635.

DSP Software Engineering, Inc.
Peter D. Robinson, 175 Middlesex Turnpike, Bedfod, MA 01730 USA, tel: (617) 275-3733, fax: (617) 275-4323.

SGS-Thompson Microelectronics.
1310 Electronics dr, Carrollton, TX 75006, tel: (214) 466-6000, fax: (214) 466-7602.

VoCal Technologies.
1576 Sweet Home Road, Buffalo NY 14228, tel: (716) 688-4675, fax: (716) 636-3630, email: victord@vocal.com.

Spectron Microsystems.
Elma Carey, European Sales Manager. 20 Upper Merrion Street, Dublin, 2 Ireland. Tel: 353-1-6761291, fax: 353-1-6762125, e-mail: elma@spectron.com.

CME, centre for Micro-Electronics
J. Steensma and M. Bloemendaal, Vendelier 71, Postbus 1001, 3900 BA Veenendaal, tel: 0318-580200, fax: 0318-580234, e-mail: dsp@cme.nl.

17. Internet addresses

www-pages: http://

Global information sites:		www.dspnet.com
	www.cera2.com	
Texas Instruments:		www.ti.com
	www.marshall.com	
GAO:	www.io.org/~gao	
SGS-Thomson:	www.inmos.co.uk	
DSP-Group:	www.dspg.com	

ftp adresses: ftp://

Texas Instruments:		ftp.ti.com
Analog Devices	ftp.analog.com	

gopher-sites: gopher://

News groups:

dsp related:	comp.dsp

BBS telephone numbers:

Analog Devices:	(617) 461-4258
Texas Instruments:	(713) 274-2323

18. Literature

[1]: Varn, D.K and Newman, W.
ANSWERING MACHINE DEMONSTRATES BENEFITS OF SYSTEM-LEVEL DESIGN.
EDN, the design magazine of electronic industrie, April 14 1994, p. 107-112.

[2]: Leonard, J.P.
EDNs DSP-CHIP DIRECTORY.
EDN, the design magazine of electronic industry, June 9 1994, p. 75-135.

[3]: Quinntell, R.A.
PCs AND TELEPHONY START TO MERGE.
EDN, design feature, August 3, 1995, p. 34-44

[4]: Leibson, S.H.
DSP DEVELOPMENT SOFTWARE.
EDN, November 8, 1990, p. 157-165.

[5]: Ing. A.C.P. van Meer en ir. G. Verkroost.
DIGITALE SIGNAALBEWERKING
Een zes daagse cursus in november 1990 van het PATO.

[6]: Sandbank, A. et al.
THE NS32AM163. A VERSATILE DIGITAL ANSWERING MACHINE.
In: Proceedings of the Fourth International Cpnference on Signal Processing Applications and Technology, Santa Clara, CA, USA, Sept. 28-Oct. 1, 1993.
DSP Associates, Newton, MA, USA. Vol. 1, p. 293-9.

[7]: Leenhardt, O.
ST18XSTAR. THE FIRST STATIC ANSWERING/RECORDING SINGLE DSP CHIP.
In: Proceedings of the Fourth International Conference on Signal Processing Applications and Technology, Santa Clara, CA, USA, Sept. 28-Oct. 1, 1993.
DSP Associates, Newton, MA, USA. Vol. 1, p. 269-73.

[8]: Berger, S. et al.
AN APPLICATION SPECIFIC DSP FOR SPEECH APPLICATIONS.
IEEE Transactions on Consumer Electronics, Nov. 1993, Vol. 39, Iss.: 4, p. 733-9.

[9]: Intrator, I. and Falik, O.
A SINGLE CHIP CONTROLLER FOR DIGITAL ANSWERING MACHINES.
IEEE Transactions on Consumer Electronics, Feb. 1993, Vol. 39, Iss.: 1, p. 45- 48.

[10]: Westall, F.A. and Ip, S.F.A.
DIGITAL SIGNAL PROCESSING IN TELECOMMUNICATIONS.
Published by Chapman & Hall. First Edition 1993.

[11]: M. Levy and J. P. Leonard..
EDNs DSP-CHIP DIRECTORY.
EDN, the design magazine of electronic industry, May 11, 1995, p. 40-95.

[12]: Databook, Zilog.
TELEPHONY ANSWERING DEVICE CONTROLLERS
Product Specifications Databook. 1993.

[13]: ir.A.W.M.van den Enden, ir.N.A.M. Verhoeckx
DIGITALE SIGNAAL BEWERKING
Delta Press BV, Overberg, The Netherlands.

[14]: L. H. Sibul
ADAPTIVE SIGNAL PROCESSING
IEEE Press , the institute of Electrical and Electronics Engineers, Inc., New York.

[15]: J.L. Lacoume, et.al.
SIGNAL PROCESSING IV, THEORIES AND APPLICATIONS
Proceedings of EUSIPCO-88, Fourth European Signal Processing Conference, Grenoble, France, September 5-8, 1988.

Appendix A.

```
#include <stdio.h>
#include <math.h>

void generate_DTMF(double lowfreq, highfreq)
CONST generatefreq = 8000                          /* same as samplefreq */
        duration = 0.1                             /* 100 ms DTMF tone */
        twist = 1.778                              /* 5 dB */
{

   double a_lowfreq, b_lowfreq, a_highfreq, b_highfreq;
   double x_lowfreq, y_lowfreq, x_highfreq, y_highfreq;
   double old_x_lowfreq, old_x_highfreq), DTMF_sample;
   int number_of_samples, n;

   a_lowfreq = SIN (2*M_PI * (lowfreq/generatefreq));   /* calculate the constants
   b_lowfreq = COS (2*M_PI * (lowfreq/generatefreq));      which are used in the
   a_highfreq = SIN (2*M_PI * (highfreq/generatefreq));    generation of the sine
   b_highfreq = COS (2*M_PI * (highfreq/generatefreq));    and cosine waves      */

   x_lowfreq = x_highfreq = 0;                      /* define the starting point
   y_lowfreq = y_highfreq = 1;                         of the 2 waves          */
   TO_DAC (0);                                      /* output to digital to analog conv */
   number_of_samples = generatefreq * duration;

   for (n = 1; n <= number_of_samples; n++)
   {
      old_x_lowfreq =  x_lowfreq;                   /* store previous value of sin wave */
      old_x_highfreq = x_highfreq;
            /* calculate the new values of the 2 sine          and cosine waves */
      x_lowfreq = (b_lowfreq * old_x_lowfreq) + (a_lowfreq * y_lowfreq);
      y_lowfreq = (b_lowfreq * y_lowfreq) - (a_lowfreq * old_x_lowfreq);
      x_highfreq = (b_highfreq * old_x_highfreq) + (a_highfreq * y_highfreq);
      y_highfreq = (b_highfreq * y_highfreq) - (a_highfreq * old_x_highfreq);

      DTMF_sample = x_lowfreq + (twist * x_highfreq);   /* add the two samples, using the twist */
      TO_DAC (DTMF_sample);                             /* output DTMF sample to DAC */

      /* Wait for (T_generate-duration) or interrupt set to appear every T_generate seconds */
   }
}
```

C-listing of the DTMF-generation algorithm.