Customizing the MT9092/94 Handsfree Microprogram

ISSUE 1 August 1993

Introduction

Trying to obtain a written record of a particular piece of information while cradling a telephone handset between your shoulder and chin is an activity known to everyone. Being chained to within four or five feet of the telephone while what you need is just beyond reach is another reason that handsfree telephony has been developed. This physical freedom, which handsfree telephony provides, is also of great benefit to the handicapped. For marketing purposes the addition of handsfree to the line-up allows a family of phones to be differentiated by features. Although not usually admitted, in many cases the actual use of handsfree is secondary to the prestige of having it available.

Historically handsfree has provided the advantages just presented, but in most cases has not been an absolute necessity. However, an application is emerging where handsfree telephony may become mandatory. Mobile car-phones are a case where its use may be legislated, in some national or regional jurisdictions, for safety.

Whatever the reasons, handsfree functionality is becoming more widespread. It is now up to the set designer to implement the most cost effective solution while still maintaining the level of subjective performance which the end-user expects.

This application notes discusses, in a generic manner, the current architectures for implementing handsfree telephony. In particular this application note is intended to extend the information available, within the MT9092/94 data sheets, regarding the implementation of Mitel's solution.

Types of Solutions

Two architectures exist for handsfree implementations.

1. Handsfree using echo-cancellation (Full-duplex operation)

Full duplex, echo-cancellation technology has historically been too expensive for other than shared resource implementations. This method maintains the natural flow of speech since normal conversations are full-duplex. State of the art cancellation algorithms are reducing the cost of implementation but this remains high for most applications. For this reason this application note will concentrate on switched loss (half-duplex) algorithms.

2. Switched loss algorithms (Half-duplex operation):

Switched loss algorithms are simpler than echo cancellation and are, therefore, more economical to implement. With this technology, echoes are not suppressed so there is a need to switch loss into the loop to maintain stability. Two main methods apply:

a/ Switch on an absolute threshold level

b/ Switch on the relative level difference between transmit and receive directions.

The half-duplex nature of these methods means that the switching transitions could interrupt the natural

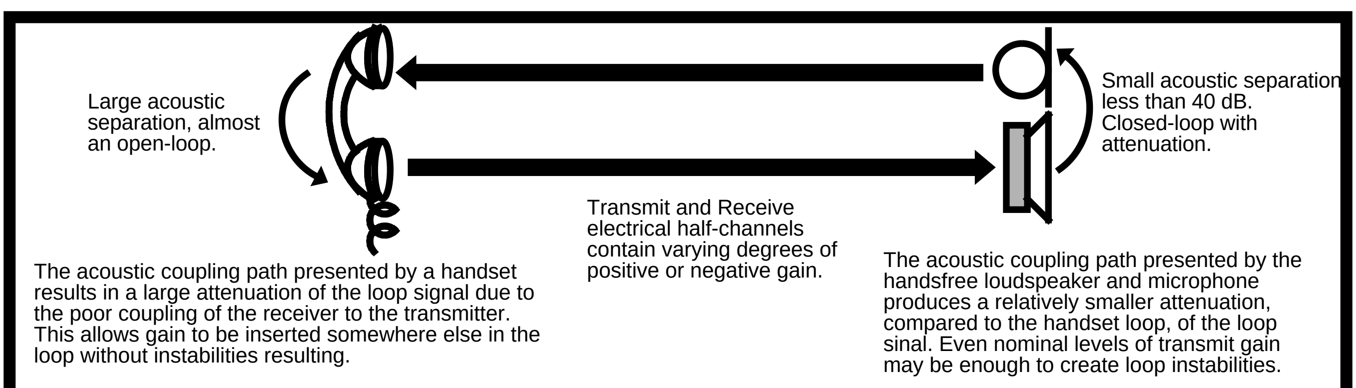


Figure 1 - Handset vs. Handsfree Operation

flow of full-duplex conversation. The proximity to which an algorithm approaches real conversational needs depends upon its complexity. Method a/ (absolute level switching) is usually to crude to allow a natural conversational flow. Method b/ (relative level switching) is the better method and, depending on the subtlety of the algorithm, can support full-duplex speech in a half-duplex environment without noticeable distress to the users.

Handset Versus Handsfree Operation

In the definition of handset versus handsfree operation the concept of an open loop versus a closed loop is useful. The complete loop consists of the electrical half-channels of the transmit and receive paths combined with the acoustic loop-backs provided by the speaker to microphone and handset acoustic coupling. This is illustrated in Figure 1.

As illustrated, handset operation effectively opens the loop by providing a large attenuation of the loop signal. This is accomplished by the close proximity of the receiver to the user's ear during use. This proximity allows the actual received acoustic power to be very small. This accomplishes two things. First, the small acoustic power emanating from the receiver does not couple back to the transmitter within the handset. Second, any radiated signal is effectively cut off by the users head and therefore does not find a path back to the transmitter.

This physical nature of a handset conversation acts to insert attenuation into the complete transmit/ receive loop guaranteeing total gain below 0 dB and thus maintaining stability.

Handsfree operation entails the removal of the acoustic separation inherent in handset operation allowing the total loop gain to approach and perhaps exceed 0dB. This causes ringing (more commonly called feed-back). Introducing a switched loss into the transmit or receive path allows the active direction its full path gain while maintaining loop stability by inserting loss into the inactive direction.

Problems Encountered with Switched Loss Algorithms

The implementation of a switched loss handsfree is a balance between many opposing factors. To complicate the situation even more many of these factors do not remain consistent throughout the duration of a conversation.

The environment in which the handsfree operates is probably the most variable factor encountered. For example room acoustics vary greatly from one setting to another with the worst case being a highly reverberant room composed of hard, flat surfaced walls. Within a specific environment the echo patterns will change as the user, or users, move about.

The presence of background noise sources such as chronic fan or motor noises, impulsive sound sources such as keyboard use or items falling or being moved about, and imitative noises such as radio/television all work to degrade the performance of an algorithm which must decide between these noises and the actual conversation.

The coupling of received speech from the loudspeaker to the transmit microphone, known as acoustic separation, will vary according to the physical implementation of each telephone design. Within a given phone design, the component of acoustic separation resulting from energy transfer within the set will remain relatively constant while the component produced externally may change. This change may result from the surface on which the set resides as well as by other items placed in close proximity which may reflect voice energy back to the set. Therefore, an algorithm must be able to operate with a minimal amount of separation.

A factor not related to the environment at all is the individual user's subjective opinion. What may seem a natural switching flow for one user may be unacceptable to another. As well, languages possess characteristics which affect the detection and identification of speech. As these differ from one language to another, the algorithm must be able to span these differences.

Finally the transmission medium can vary in the amount of echo being reflected back from the trunk, as well as noise energy generated within the network or being transmitted from the far end background. Of these echo is the more important because of the gain which the echo component (near-end plus far-end) may encounter during its round trip.

In designing the algorithm care must be taken not to degrade the performance of one parameter while optimizing another. Some examples of this are:

Setting the transmit and receive thresholds at a level where speech from a quiet talker will trigger the switching decision but not low enough that background noise effects this decision. At the same time a fair balance between the two directions is required or one side will have a distinct advantage over the other in terms of control of the loop.

In separating speech energy from noise energy of the same level, the main criteria used is the variation over time. Voice signals are assumed to be non-uniform over time. Chronic noise by definition is fairly constant over time and therefore can be integrated and removed from the total signal. Impulse noise cannot be removed from the composite signal requiring that the algorithm ignore it while still reacting to speech. By allowing the switching process to proceed slowly, compared to the actual decision process, an algorithm can be made fairly insensitive to impulse noise but still speech responsive. In this case a switching transition may be initiated by impulse noise but will not proceed to far along the slope before the impulse is gone and the algorithm settles back where it belongs.

Acoustic reverberation within the listening area will cause the received signal applied to the loudspeaker to appear at the transmit microphone at a fairly high level and smeared over time (echo). It is therefore necessary to ignore room reverberation as much as possible but to still remain responsive to transmit speech. In most cases the reverberant signal will be significant for the first 10 to 20 mSec after the speaker signal disappears with decreasing echo for the next 200 mSec. By setting the attack time of the algorithm properly most of the echo signal can be ignored while still responding properly to transmitted speech energy. If the attack time is too long then loss of the first (or more) syllable(s) in the transmit direction will occur.

When implementing a switched loss algorithm decisions must be made regarding the threshold levels at which the transmit and receive paths will be deemed active and what proportion of this activity is actual speech energy and what proportion noise energy. Further, what level of speech energy in the non-active path greater than that of the active path will be cause for a switching decision change (hysteresis)?

To control erratic or excessive switching it is necessary to distinguish between low speech energy due to inter-syllable and inter-word pauses from that occurring at the conclusion of a speech segment. The attack and decay rates and hold-over or pause time, designed into an algorithm, perform this function.

The final implementation of a switched loss algorithm is very dependent upon the physical application as well as the user's subjective evaluation of performance. There is no single "best" solution which fits all application environments. The CCITT has produced a Recommendation (P.34) which should be used as a basis for implementation. A handsfree designed on these parameters can be relied upon to give satisfactory performance under most conditions. However this is a recommendation only and is therefore open to adjustment as required. In many situations the ability to deviate from P.34 requires acoustic laboratory resources beyond the reach of many designers. The optimum implementation is therefore an algorithm which operates by default with parameters in line with CCITT P.34 but which offers the ability to easily adjust those parameters which might tailor the algorithm to its specific environment. The remainder of this applications note deals with Mitel's solution to this requirement.

Mitel's Solutions

Mitel offers a fully digital implementation of a switched loss algorithm in its digital telephones. In particular the MT9092 and MT9094 incorporate a Digital Signal Processor implementing a sophisticated half-duplex switching algorithm for handsfree telephony operation.

Gain switching is performed in continuous 1.5dB increments and operates in a complimentary fashion. That is, with the transmit path at maximum gain the receive path is fully attenuated and visa versa. This implies that there is a mid position where both transmit and receive paths are attenuated equally during transition. This is known as the idle state.

Of the 64 possible attenuator states the algorithm will rest in only one of the three stable states; full receive, full transmit and idle. The algorithm determines which path should maintain control of the loop based upon the relative levels of the transmit and receive audio signals after detection and removal of noise energy. If the algorithm determines that neither the transmit or the receive path has valid speech energy then the idle state will be sought. The present state of the algorithm plus the result of the Tx vs. Rx decision will determine which transition the algorithm will take toward its next stable state.

By default the algorithm uses parameters generally conforming to CCITT Recommendation P.34. This is the state of operation as described in the MT9092/94 data sheets. In most circumstances these default parameters are adequate and will probably not be in need of adjustment.

The MT9092/94 device data sheets provide no information related to altering the switching algorithm parametrics. However many of these parameters may be re-programmed under micro-processor control to tailor the algorithm to its specific environment.

Table 1 lists the handsfree parameters which may be customized and the effect each parameter has upon the algorithm's operation.

Detailed descriptions for each of these parameters are provided at the end of this application note.

Reprogramming the 16-Bit Handsfree Parameters

The MT9092/94 utilizes a serial microcontroller interface compatible to Intel MCS-51 (mode 0) specifications. As outlined in the device's data sheet, access to normal read/write registers is via a two-byte operation consisting of a Command/Address Byte followed by the data byte.

The handsfree parameter addresses 22h to 2Dh are 16 bit registers requiring that this procedure be extended to a three-byte operation. Addresses 1Dh and 20h are single byte registers and remain a two-byte transfer.

To set up for a 3-byte transfer, Address 1Fhex should be written with 80 hex.

To perform a 3-byte transfer the Command/Address byte must be altered and the Chip Select ($\overline{CS}$) signal extended. In detail, the Command/Address byte data bit D_7 (normally written "0") must be set to a logic "1" and the $\overline{CS}$ signal must remain active for the duration of a 3-byte transfer. The byte sequence is COMMAND/ADDRESS followed by DATA BYTE1 (D_0-D_7) followed by DATA BYTE2 (D_8-D_{15}) (see Figure 2). This is true for both read and write functions. After completion of the 3-byte transfers Address 1Fhex should be returned to 00Hex.

Switching Algorithm Flow

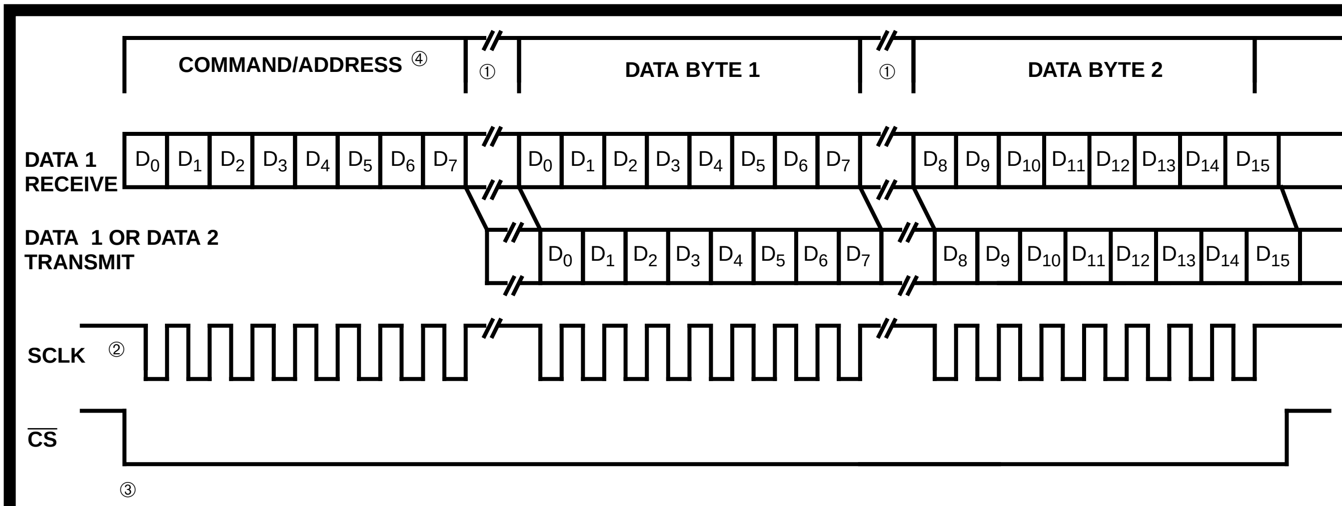
Since the complete switching algorithm is implemented as a stored program there are no actual hardware blocks performing the required handsfree functions. In order to visualize the algorithmic operation Figure 3 has been included. This hardware block diagram does not actually exist but is used as an analogy of the ROM code implementation.

In this analogy the receive and transmit PCM are first high pass filtered, to remove low frequency environment noise, rectified and then peak detected. These peak signals are then processed in the speech detectors to identify the presence of speech signals from chronic background noise. As well the two peak signals are directly compared to determine which direction merits control of the loop.

If both speech detectors indicate only background noise then neither direction will control the loop,

Register Name	Address (hex)	Purpose
Receive Gain Control	1D	As described in MT9092/94 data sheets. The Rx attenuator uses this programmed value as a maximum receive setting.
Transmit Audio Gain	20	As described in MT9092/94 data sheets. The Tx attenuator uses this programmed value as a maximum transmit setting.
Tx Speech Detector Threshold	22	Adjusts the sensitivity of the algorithm to transmit energy.
Rx Speech Detector Threshold	23	Adjusts the sensitivity of the algorithm to receive energy.
Idle State Level	24	Controls the depth of the loop attenuation during idle state.
Comparator Decrement	25	Decay rate adjustment. Algorithmic adjustment to echo from reverberation (speaker to microphone from room acoustics) and transhybrid reflections (impedance mismatching). Can be thought of as a comparator insensitivity control.
Ramp-out	26	Controls the decay or slope of the ramp back to idle.
Hold-over	27	Delay until ramp-out is invoked. Also called Hang-over time.
Tx HPF Gain	28	Transmit values for gain (G) in Equation 1.
Rx HPF Gain	29	Receive values for gain (G) in Equation 1.
Filter Co-efficient A1	2A	Filter coefficients for Equation 1.
Filter Co-efficient A2	2B	Filter coefficients for Equation 1.
Filter Co-efficient B1	2C	Filter coefficients for Equation 1.
Filter Co-efficient B2	2D	Filter coefficients for Equation 1.

Table 1



① Delays due to MCS-51 internal timing which are transparent.

② Received data latched on the falling edge of SCLK. Transmit data output on the falling edge of SCLK.

③ The falling edge of $\overline{CS}$ indicates that a COMMAND/ADDRESS byte will be transmitted from the microprocessor. The subsequent bytes are always data followed by $\overline{CS}$ returning high.

④ The COMMAND/ADDRESS byte contains:
1 bit - Read/$\overline{Write}$
6 bits - Addressing Data
1 bit - 16 bit access, write Logic "1"

D_7 ⟷ D_0

1	A_5	A_4	A_3	A_2	A_1	A_0	R/$\overline{W}$

Figure 2 - 16-bit Access Relative Timing

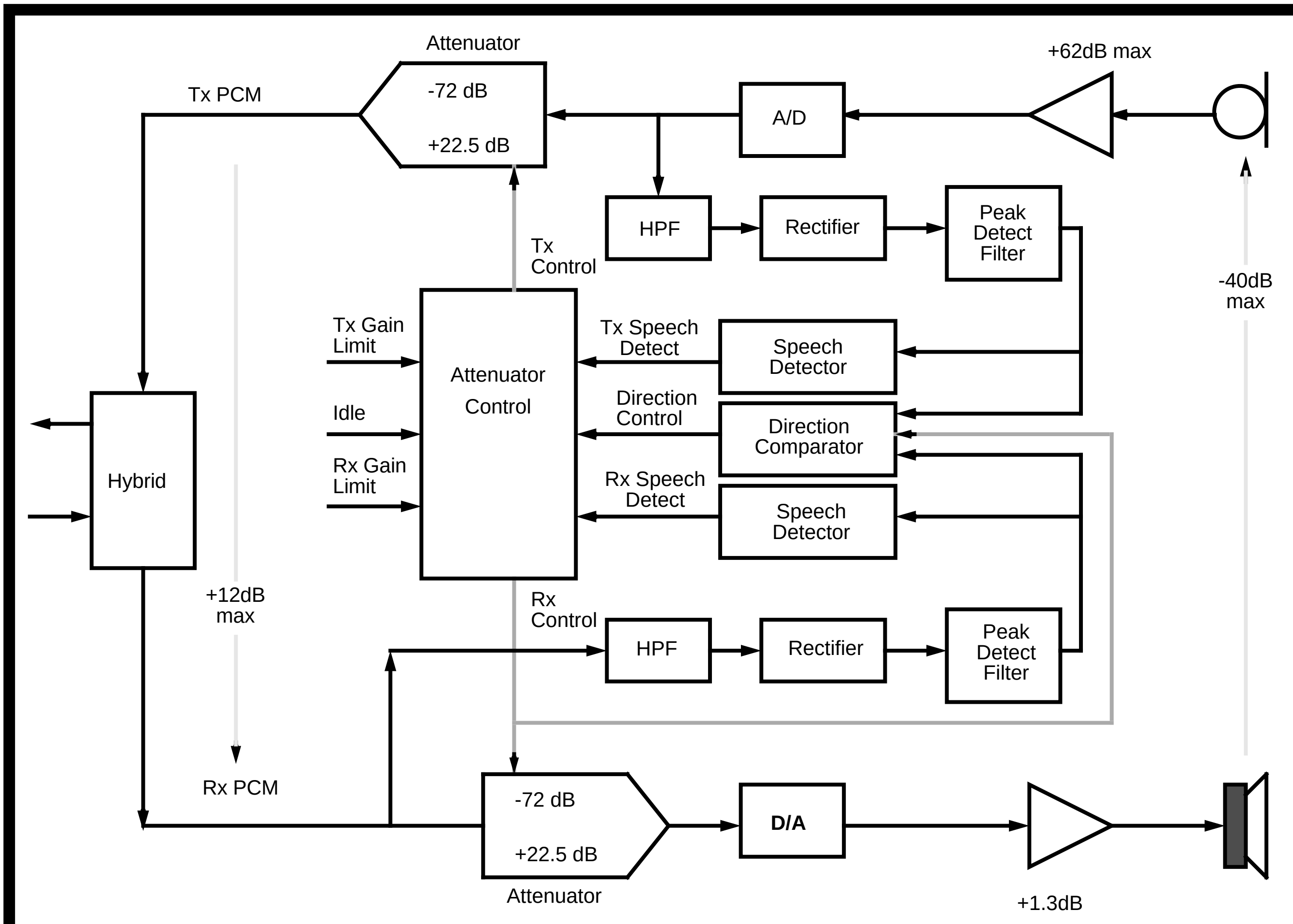


Figure 3 - Software Functional Block Diagram

regardless of the direction comparator output, and the idle state will be sought. The result of this processing is then used to control the transmit and receive attenuators in a complimentary mode. The transmit, receive and idle state levels, programmed via microport register, are also included in the attenuator control.

Attenuators

The attenuators are programmable for maximum transmit and receive gain as well as for the idle state level.

Receive gain setting:

$$-72.0dB <= RxGain <= +22.5dB$$

(i.e., no restrictions)

Transmit gain setting:

Any positive gain programmed into the Transmit attenuator will be reflected as decreased noise performance for the transmit path. Therefore, the following is recommended;

$$-72dB <= TxGain <= 0dB$$

Idle level setting:

The idle level should not be programmed too far below the maximum transmit or receive gain level. The advantage of switching from idle to full state will be lost as the idle state level approaches the full attenuation level. Therefore, the following is recommended:

$$(2 \text{ X IDLE Level - RxGain Level}) >= 0$$

$$(2 \text{ X IDLE Level - TxGain Level}) >= 0$$

Attenuator Flow Diagram

The attenuators are controlled according to the algorithm shown in Figure 4.

Parametric Programming Details

Filter Programming

Before entering the peak detectors the transmit and receive audio signals are filtered. The biquadratic filter function is defined as:

Equation 1

$$Y(z)/X(z)=G[1+A_1Z^{-1}+A_2Z^{-2}]/[1-B_1Z^{-1}-B_2Z^{-2}]$$

The filter defaults to a second order high-pass function with a -3dB point at 400 Hz and a passband gain of -12dB.

Using the coefficients at addresses 28h to 2Dh the filter response can be altered according to the above equation where:

The decimal coefficient x 2^{12} (4096) gives a decimal number which, when converted to hexadecimal, is stored in the respective register. Note that for negative numbers the decimal coefficient is first subtracted from 65536_{10} before conversion to hexadecimal. The default coefficients and calculations are given below.

G: +0.2054 decimal(-12dB passband)
Actual default = 0350hex

$$0.2054 \times 4096 = 841.3184_{10} = 349_{16}$$

A_1: -1.9911 decimal
Actual default = E024hex

$$65536 - (1.9911 \times 4096) = 57380.4544_{10} = E024_{16}$$

A2: +1.0000 decimal
Actual default = 1000hex

$$1.000 \times 4096 = 4096_{10} = 1000_{16}$$

B_1: +1.6067 decimal
Actual default = 19C0hex

$$1.6067 \times 4096 = 6581.0432_{10} = 19B5_{16}$$

B_2: -0.6722 decimal
Actual default = F540hex

$$65536 - (0.6722 \times 4096) = 62782.6688_{10} = F540_{16}$$

Round-off errors account for slight differences between the calculated coefficient and the default value. Note that finite word lengths within the filter will degrade performance for gain (G) greater than -12dB. A further caution is that any new filter coefficients should be simulated with 16-bit precision to ensure stability of the design.

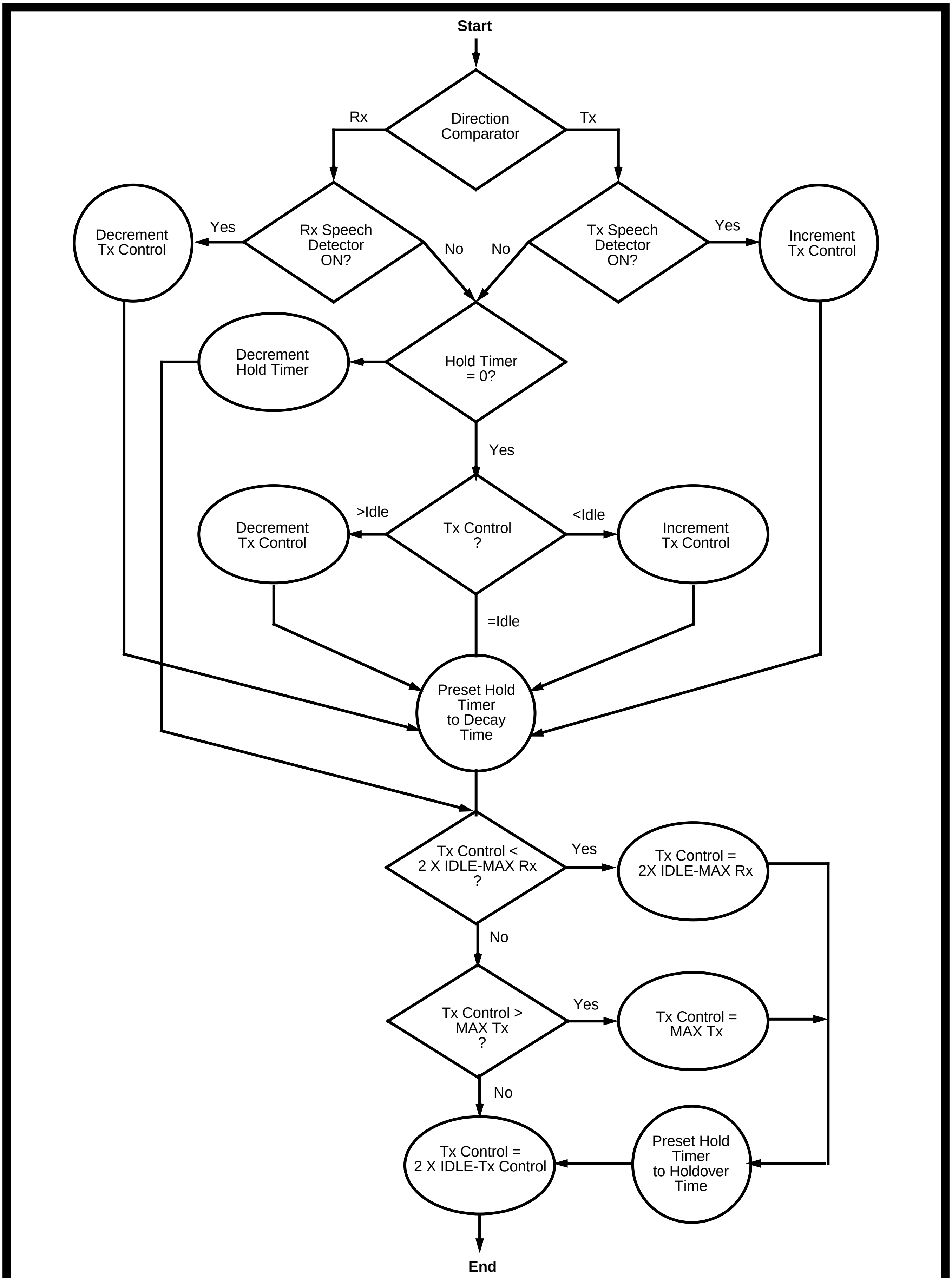


Figure 4 - Switching Algorithm

Receive Gain Control Register
Address = 1Dh Write/Read verify

Transmit Gain Control Register
Address = 20h Write/Read verify

These are 8-bit registers operating as described in the MT9092/94 data sheets. The algorithm uses these programmed values as the maximum settings which the attenuators will use.

Idle State Level
Address = 24h Write/Read verify
Default 0026hex

The idle state level is programmed in the same manner as the gains for transmit and receive. Only the 6 least significant bits (D_0 - D_5)are used to form a value shown as $Idle_{10}$ in the following equation. Note that the hex register value is converted to a decimal value. Alternatively, the table in the Receive Gain Control Register (Address 1Dh) of the MT9092/94 data sheets may be used. This is a sixteen bit register and bits D_6 - D_{15} should be set to logic "0" at all times.

$$Idle\ Level\ (dB) = (1.5\ dB \times idle_{10})\ -72dB$$

Transmit Speech Detector Threshold
Address = 22h Write/Read Verify
Default 01C0hex

Receive Speech Detector Threshold
Address = 23h Write/Read Verify
Default 00E0hex

These are 16-bit registers containing the threshold levels used by the speech detector comparators. The hexadecimal value stored in this register, after conversion to decimal, can be related to the backplane level in dBm0 by the following equation:

$$20\ LOG\ (Threshold_{10})-87=dBm0$$

For the MT9092/94 CODEC dBm0 is related to dBv by the overload decision levels. For example:

In the A/D path for µ-Law the 3.17dBm0 level is set at:

$5.79\ V_{p-p}$ or $2.05V_{RMS}$.
$20_{LOG} \times 2.05 = 6.23dBv$.

Therefore, 0dBm0 = 3.06dBv.

Comparator Decrement
Address = 25h Write/Read Verify

This 16-bit register controls the full-scale decay rate of the comparator low pass filters.

There are only four recommended values even though this is a sixteen bit register.

These are:

Hex Value	Full Scale Decay
0001	72 mSec
0002	52 mSec
0003	42 mSec
0004	35 mSec (Default)

Ramp-out Register
Address = 26h Write/Read Verify
Default 00A0hex

The slope of the ramp from either of the transmit or receive states back to idle state is controlled by the value programmed into this register. As the attenuators pass through each 1.5dB gain setting they pause for a time interval programmed according to:

$$Time\ Interval\ per\ 1.5dB\ Step= (Value_{10})x(0.5mSec)$$

Hold-over Register
Address = 27h Write/Read Verify
Default 0190hex

The amount of time the attenuators hold at full (Tx or Rx) state before beginning to ramp down to idle is controlled by the value programmed into this register. The Hold time is programmed according to:

$$Hold\ Time\ =(value_{10})\ x\ (0.5mSec)$$