OPERATIONAL CHARACTERISTICS OF MODERN SET-TOP TERMINALS

James O. Farmer

Scientific-Atlanta, Inc.

INTRODUCTION

A brief history of set-top converters is presented, along with notation of the techniques commonly employed in modern terminals. The digital architecture of a modern terminal is shown. Key RF characteristics are presented, followed by cursory exploration of one class of scrambling techniques. Finally, some information concerning the compatibility between scrambling and stereo is presented. Most of the material is intended to be generic, but where particular techniques are referred to the system described is that used by the Scientific-Atlanta Series 8500 set-top terminal.

EARLY CONVERTERS

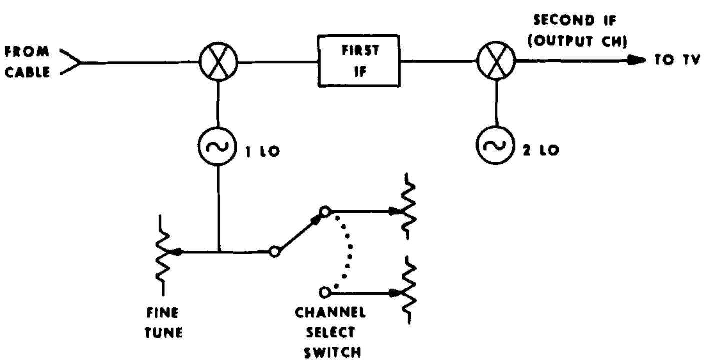


FIG 1 SET-TOP CONVERTER

Since set-top converters were first developed the architecture employed has been as shown in figure 1. Signals from the cable were applied through a low pass filter (not shown) to a balanced mixer. Here they were mixed with signals from a local oscillator whose frequency was higher than that of the incoming signal. The resultant first IF was at some conveniently high frequency, usually in the lower portion of the UHF spectrum (which officially extends from 300 MHz to 3 GHz). This high IF frequency was chosen in preference to the much lower IF used in TV receivers, in order to simplify design of the input filter and in order to prevent the possibility of L.O. radiation in-band.

The first IF signal, after amplification and filtering, was mixed with a fixed tuned second local oscillator to produce a second IF frequency which was the same as the frequency of channel 2, 3, or 4. In recent times, many people have stopped using channel 2 as an output because that is the second harmonic of the citizens band.

The first local oscillator was frequency controlled by a varactor diode. The varactor received its bias from one of a number of potentiometers, one for each channel. Because of the instability of the two local oscillators and the potentiometer voltages, a fine tuning control also had to be provided. This fine tuning voltage was added to the control voltage selected by the channel select switch.

Even today this architecture is utilized in economy lines of set-top converters. However, it exhibits several shortcomings that have driven manufacturers to develop new architectures. For example, the voltage to the varactor must be controlled very carefully. This requires good regulation. In the case of a wired remote control unit, the variability of contact resistance precludes use of plug-in cords.. Adjusting the fine tuning is difficult for some subscribers. Although descrambling has been added to this architecture, the security available is marginal, because descrambler authorization must be mechanically coupled with the channel select switch.

These shortcomings have prompted manufacturers to develop new architectures for set-top converters. The differences are sufficiently great that a new name for the box has been coined: the set-top terminal.

ENTER THE TERMINAL

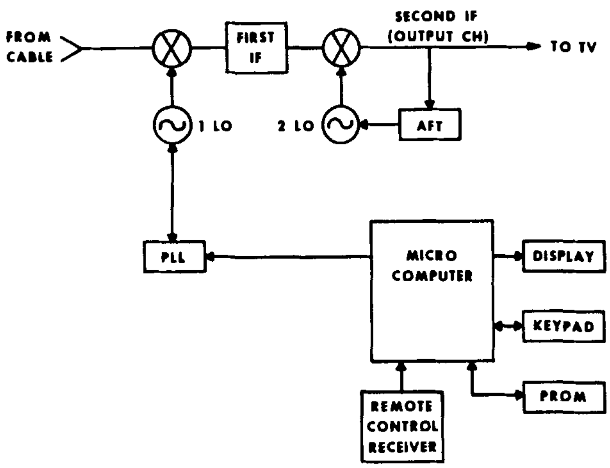


FIG 2 SET-TOP TERMINAL

Figure 2 shows the basic architecture of the new terminals, which are microprocessor controlled. The basic RF path remains much the same as before, except that first IF frequencies are going up for reasons that will be explained shortly. Better set-top terminals have replaced the manual fine tuning with automatic fine tuning to relieve the subscriber of this burden. Some set-tops have also added adjacent carrier traps, as explained below.

As the maximum number of channels went from 35 to 54 and more, most manufacturers abandoned potentiometer tuning of the first local oscillator. Phaselocked loop (PLL) tuning is now almost universal in terminals. In a phaselocked system, the local oscillator frequency is divided to a low frequency, which is then compared to a reference frequency. If the two differ even in phase (the integral of frequency), a correction voltage is sent to the local oscillator, forcing it to the correct frequency. By changing the division ratio ("modulus" of the counter), the local oscillator is forced to tune to the desired frequency. Long time stability is as good as that of the reference frequency, which is crystal controlled.

Thus, each time the tuned channel is changed, a new tuning word must be supplied to the PLL. This is generally supplied from a single chip microcomputer, which performs several other functions. Lighted channel displays (usually LED) are utilized, and are controlled by the microcomputer. The keypad used for channel entry and other functions is also scanned by the microcomputer. Wireless remote control using infrared signaling is a common option. The microprocessor accepts the modulated signal from the remote control receiver and interprets it.

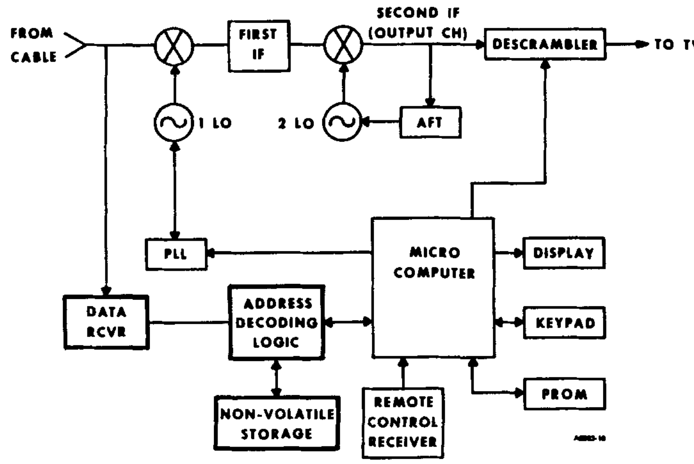


FIG 3 ADDRESSABLE TERMINAL

Figure 3 shows the same set-top terminal except that now we have added descrambling and addressability functions. No additional RF connections are necessary for descrambling, as virtually all current scrambling systems utilize in-band timing to permit an unlimited number of scrambled channels at no incremental cost per channel. The descrambler must be authorized from the microcomputer in such a way as to preclude tampering to force the descrambler on.

A data carrier on the cable is frequency shift keyed (FSK) with data representing a box address and information intended for that box, or with information intended for all boxes (a global command). The data receiver is lightly coupled to the incoming cable. This receiver is constructed similarly to a single channel FM radio through the discriminator, except that a crystal oscillator is employed for long term stability and special techniques are used to reduce local oscillator emission. After the discriminator, a low pass filter and comparator are used to convert the demodulated signal to logic levels. The logic level encoded signal from the receiver is supplied to the address decoding logic. This circuit, under microcomputer control, determines whether or not the incoming information is intended for this terminal. If so, the updated information transmitted is stored in non-volatile memory. The non-volatile memory is an electrically alterable read-only memory, which is able to retain its memory even under power-off conditions. This is important because neither the subscriber nor the cable operator should be bothered with having to initiate a memory update after every power loss.

ADDRESSABLE DIGITAL ARCHITECTURE

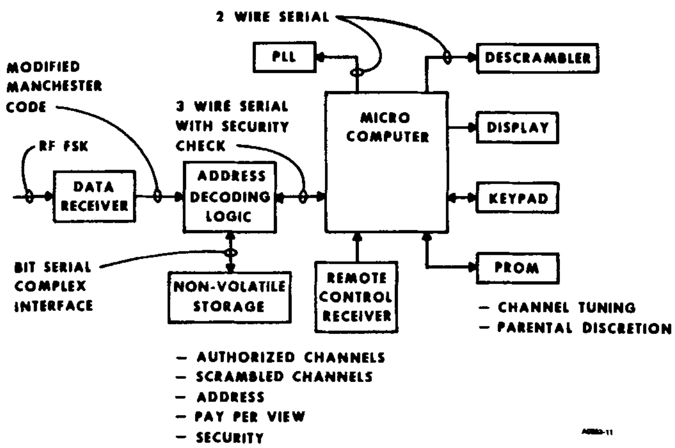


FIG 4 ADDRESSABLE DIGITAL ARCHITECTURE

The digital architecture of an addressable terminal is shown in Figure 4. At the headend the data for all terminals is formated into a modified Manchester code, in which information is carried as a logic transition for each clock pulse. If a 1 is to be transmitted an extra transition is inserted between clock pulses. A longer period with no transition indicates a start of message pulse. This format is inherently self-clocking and exhibits high tolerance for both noise and timing errors.

The Manchester encoded data is frequency shift keyed (FSK - a digital form of FM) onto a carrier. At the terminal, this carrier is demodulated in a receiver not too unlike an FM radio, whose output is applied to a custom IC in the address decoding logic. This chip translates the Manchester encoded information into parallel bytes, which are shifted into a microcomputer for interpretation.

Error detection techniques are utilized in order to prevent the reception of erroneous data. When the decoding logic interprets that it has received valid information (e.g., an updated list of authorized channels), this information is loaded into non-volatile memory. If it affects the current status of the terminal (e.g., channel tuned to) then the terminal immediately makes a change in its status.

Pay-per-view functions are also stored in non-volatile memory, in the form of a table containing pre-authorized channel and program numbers. Each pay-per-view program is pre-assigned a channel number and within that channel, a program number. When a subscriber requests a certain program, it is entered into the table in his non-volatile memory, via the headend computer. This entry is made at the time of the request, which may be hours, weeks or even months before the event. When the event begins, a global command is sent with the channel and program numbers contained within it. Upon receipt of this command, each terminal will search its non-volatile memory. If a match is found, the program is authorized. At the end of the program, another command cancels the authorization. A comprehensive set of commands generated at the headend permit the non-volatile memory in each terminal to be maintained with current requests only, and allow old and cancelled requests to be purged.

A series of commands and responses are provided to ensure that a subscriber or third party will encounter excessive difficulty in using an unauthorized terminal, while ensuring that an authorized terminal stays active.

Serial communications are utilized between different elements within the terminal. The microcomputer shown interfaces with the address decoding logic through a non-standard three wire serial interface with security check. The logic is such that no line can be tied high or low to "trick" the system. Interface between the address decoding logic and the non-volatile memory is via a very complex serial format that would be difficult for a pirate to duplicate. Again, both logic 1 and 0 levels are required for a successful transaction.

Other features of the digital architecture are shown in figure 4. The microcomputer reads commands inserted via the keypad or remote control receiver. A programmable read-only memory (PROM) stores channel tuning data to permit customization of channel lineups, and parental discretion information. All communications between integrated circuits is via serial ports, and valid communications require a series of 1s and 0s. This eliminates the possibility of tying a line high or low to force unauthorized operation.

REMOTE CONTROL

A mandatory option today is wireless remote control. An earlier generation of TV remote control, and one or two set-top units, utilized ultrasonic remote control. However, ultrasonic remote control suffered from blockage and false triggering due to high frequency noises such as keys jingling. Also, ultrasonic waves tended to be very non-discriminating in where they went, sometimes activating a TV in the wrong room. Most modern remote control systems utilize infrared remote control.

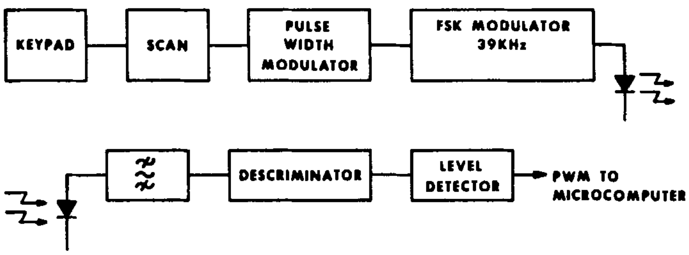


FIG 5 REMOTE CONTROL SYSTEM

Figure 5 shows a block diagram of the transmitter and receiver system. A circuit in the transmitter IC scans the remote transmitter keypad. When a key is depressed a five bit code representing the key is generated. This code is pulse width modulated (PWM) meaning that a series of pulses is generated, a long pulse representing a logic 0 and short pulse representing a logic 1. This train of pulses then frequency modulates a 39KHz carrier. This carrier then on/off (AM) modulates an infrared light emitting diode.

In the receiver, a photo transistor converts the light pulses back to the FM'd carrier. This carrier is then demodulated by a discrimator, and the resultant PWM pulses are shaped by a level detector. The pulses are then supplied to the microcomputer on an interrupt input. The microprocessor decodes the pulses to determine which key was depressed.

Since many TV receivers today utilize infrared remote control, a clear danger exists that interference will develop between the terminal and TV remote controls. Unfortunately, no clearinghouse has existed where one can obtain technical parameters of all present and planned remote control systems, so some chance exits that one will discover a conflict sooner or later. However, many TV remote control systems do exhibit some common elements. A system was developed in cooperation with a large manufacturer of remote controls, which is reasonably safe from overlap with a TV system.

With reference again to figure 5, many remote control systems utilize an FSK carrier at about 39 KHz. By changing the carrier frequency to 42KHz, some discrimination is obtained. Further security is obtained by transmitting both true and complement representation of the button pressed, with the order of the bits different from that known to be used in TV receivers. None of these counter measures will guarantee that interference will never occur, but they do move the odds in our favor.

TUNING TECHNIQUES

An almost universal attribute of set-top terminals today is control of the first local oscillator frequency by a phaselocked loop (PLL). A typical terminal which tunes from 54 to 450 MHz will have its first local oscillator frequency range from 668 to 1058 MHz (an IF of 608-614 MHz is used, as this band is reserved for radio astronomy, precluding licensing of UHF transmitters at channel 37). To control this oscillator to an accuracy of 10 KHz or so, its frequency is usually divided by 256 then compared to a reference in a CMOS phaselocked loop. In some cases, the phaselocked loop is integral to the microcomputer. Tuning is accomplished by changing the modulus of the counter fed by the prescaler. The tuning word required is often stored in PROM, to permit customization of the channel lineup for each cable system. The PLL reference is generated by a crystal oscillator, which also serves as the clock for the microcomputer.

The PLL technique is excellent for providing high stability on a multiplicity of frequencies. On the other hand, PLLs have some drawbacks. As one attempts to tune in successively smaller steps, he discovers that two undesirable things happen. First, he must store more information about each tuned frequency (i.e., the number of bits in the variable modulus counter increases) raising costs. Secondly, he must set the loop bandwith lower, resulting in longer acquisition time (e.g., from one channel to another) and potential problems with uncorrected low frequency noise.

Thus, the loop designer is forced to trade off tuning increment for storage requirement and acquisition time. We have found that a good trade-off exists for a 1 MHz tuning increment, but this leaves the problem of accomodating HRC systems, in which the frequency of each picture carrier is 0.25 MHz deviant from the nearest 1 MHz increment of a normally configured system. (A fixed 0.25 MHz offset, e.g., 55.25 MHz, is taken care of in selection of the IF frequency. But then what do we do in HRC where this is 0.25 MHz offset goes away?)

Also, how may we handle the case of a system which has offset a few tens or hundreds of KHz on one or more channels to avoid a troublesome aviation frequency ? In the days of pot-tuned set-tops, we would reajust a few pots, or let the poor subscriber fine tune every time he came to an offset channel. With PLL's and the desire for no fine tuning, we must solve this problem another way. One answer is to enclose the second local oscillator within an automatic fine tuning (AFT) loop. This technique applies the output picture carrier to a discriminator whose output is a voltage proportional to frequency error. This voltage is amplified and applied to the second local oscillator to control its frequency.

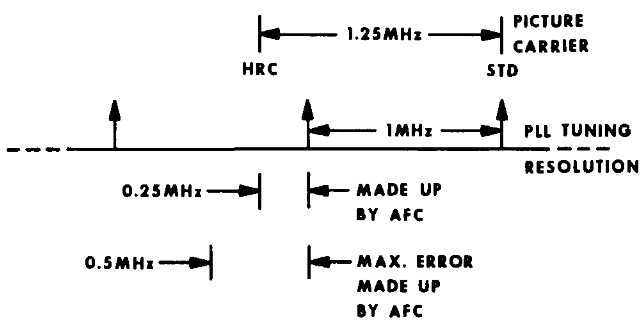


FIG 6 ACCOMMODATING HRC

Thus, the PLL tunes the first local oscillator to the nearest MHz, ensuring that the signal at the first IF is within ± 0.5 MHz of nominal. The AFC is then designed with adequate acquisition range to ensure that it will take out the remaining frequency error. Perhaps figure 6 will help illustrate. This is portion of a spectrum plot of two cable systems - one having standard carriers (on the right), and the other having HRC carriers, spaced 1.25 MHz lower in frequency. Shown below these two frequencies are the equivalent STT tuning frequencies. For the standard case we can tune to the "exact" frequency. However, in the HRC case we find ourselves with a 0.25 MHz error we can't get rid of due to the 1 MHz tuning resolution. We then turn to the AFC circuit, which is more than able to make up for this error in the conversion to output channel.

Before leaving this subject we should observe that implementation of such an AFT loop is not a job for the faint of heart. Should the second local oscillator drift or be pulled too far, then the AFT may lock to the on-channel or lower adjacent sound carriers. Should the frequency pulling range become too low, then the unit could drift to the point of not locking on frequency. And, of course, AFT discriminators have calibration errors associated with them. All of these conditions must be controlled in design and production if the AFT techniques is to be successful.

KEY RF CHARACTERISTICS

No discussion of set-top terminals would be complete without a discussion of RF specifications and their relationship to system performance. Rather than plod through the long list of specifications that are normally a part of any converter discussion, we choose to concentrate here on a few things that either have developed as issues in recent years or that continue to confound many systems engineers as to their importance or effect on system performance.

1. Frequency Response

We all know that a terminal should exhibit flat frequency response across the channel in question. But what about on adjacent channels? Ideally, the terminal should have zero response at the adjacent channel, but this would yield a more expensive terminal without offsetting technical merit. Most older generation converters relied on the passband response of the first IF to provide the entire converter response. This yields little attenuation to the adjacent channels. Many modern TV sets are designed with adequate adjacent channel rejection, so that this response is acceptable. However, some sets have insufficient adjacent channel rejection, resulting in complaints of picture quality when CATV service is installed.

To ease the adjacent channel problem some manufacturers of modern set-top terminals have added adjacent carrier traps to the output circuitry of the set-top. A typical response is shown in figure 7, which shows 12dB trap depth. A trap depth of 8 to 12 dB has been found adequate to improve the performance of marginal TV receivers, without adding excessive complexity or group delay to the terminal.

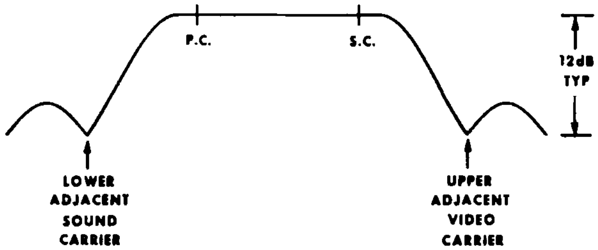


FIG 7 BANDPASS CHARACTERISTIC

Since nothing is free, however, specified frequency response may suffer slightly. The FCC specifies frequency response as from 0.75 MHz to 5MHz above the lower channel boundary (P.C. -0.5 to P.C. + 3.75 MHz). Most manufacturers have specified frequency over a broader range, to cover the sound carrier. One can see from figure 7 that the lower adjacent trap, which is 1 MHz from the FCC minimum frequency, has the potential to pull down the response in the specified region. However, the FCC frequency response can still be met or exceeded with the adjacent traps. Experience has shown that having the traps is much more important than getting the last tenth of a decibel in response flatness.

2. Backtalk

Here is an interesting specification that is relatively unknown within the industry. Under certain conditions a signal may be developed on the terminal input due to the presence of untuned signals. This spurious signal may appear on the input of another terminal, producing a beat. Until CATV systems went to extended bandwiths, this problem could be eliminated by proper choice of IF frequency. However, when frequencies above 330 MHz came into use, selection of an IF frequency to avoid this problem became impractical.

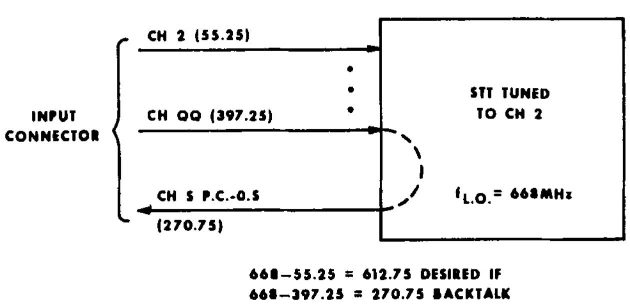


FIG 8 BACKTALK SOURCE

Figure 8 illustrates generation of backtalk by example. We have a terminal tuned to channel 2, 55.25 MHz. Among other incoming signals is channel QQ, at 397.25 MHz. Because the terminal is tuned to channel 2, the first local oscillator is oscillating at 668 MHz (668-55.25 = 612.75). Channel QQ energy is also converted by the local oscillator, to 270.75 MHz. Now this energy is ideally dissipated in losses in the mixer and IF filter, but due to incomplete balance in the mixer, some of the energy appears at the input connector. This signal is 0.5 MHz below channel S picture carrier.

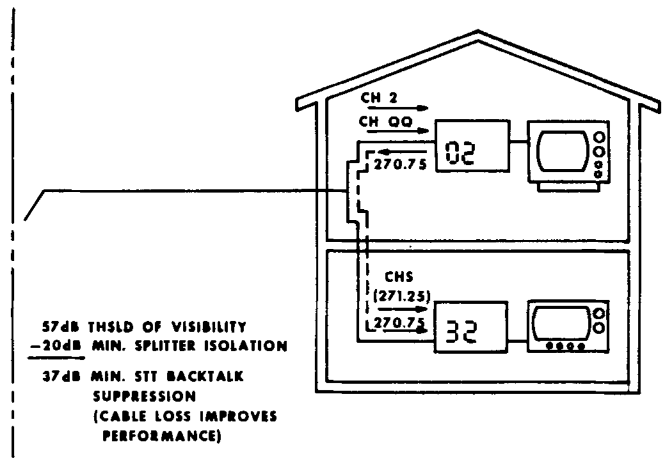
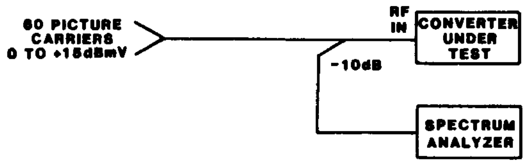


FIG 9 BACKTALK INTERFERENCE

To see how the backtalk signal can affect another TV set, see figure 9. The backtalk is generated in the terminal tuned to Channel 2 and propagates back toward a splitter. Due to imperfect isolation in the splitter, a portion of the backtalk arrives at the second set, tuned to Channel 32 (S). Here the backtalk signal shows up as a 0.5 MHz beat. Also shown on Figure 5.3 is the error budget we have used in determining what a backtalk specification should be. We assume that a backtalk signal should be -57dB from a viewed signal to be invisible. We further make the conservative assumption of 20dB isolation at the tap to arrive at a backtalk specification of -37dB from an incident carrier. We don't make allowance for frequency offset imrovement, so that we don't get trapped with an offset channel that doesn't work.



1. MEASURE SIGNAL LEVEL ENTERING CONVERTER.
2. MEASURE ANY SIGNALS FROM DIRECTIONAL COUPLER OTHER THAN INPUT SIGNALS.
3. SPURIOUS SIGNALS MUST BE -37dB FROM INPUT LEVEL, TAKING INTO ACCOUNT THE ACTUAL LOSS OF THE DIRECTIONAL COUPLER.

FIG. 10 BACKTALK MEASUREMENT

Backtalk may be measured by connnecting a 10dB directional coupler to a set-top in a direction so as to transfer backtalk energy to a spectrum analyzer. This is shown in figure 10. To allow backtalk to be seen, unmodulated carriers must be used. Calibration of this set-up is a bit tedious and should not be attempted except by someone who knows the technique well.

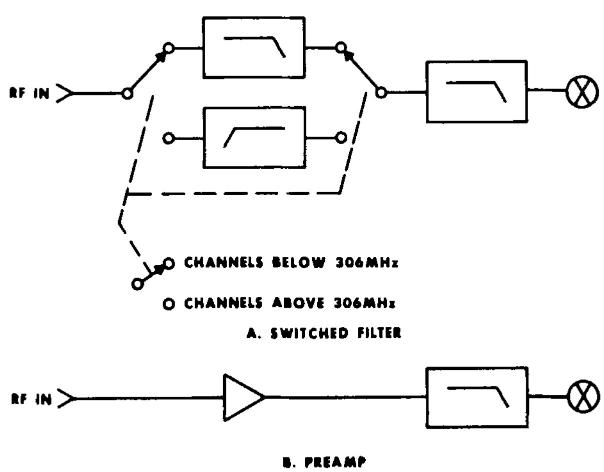


FIG 11 BACKTALK COUNTERMEASURES

If countermeasures are not taken against backtalk, normal circuit techniques will yield attenuation in the mid 20's. At least two viable countermeasures are available as shown in figure 11. In A we show the application of switched low-and high-pass filters. When tuning below 306 MHz in this case, the low pass section is switched in, attenuating the higher frequencies that cause backtalk. When the terminal is tuned above 306 MHz, we can show that backtalk doesn't occur, but we switch in a highpass filter anyway. This reduces mixed loading, reducing intermodulation distortion.

The second countermeasure available is a pre-amplifier used ahead of the mixer, as shown in figure 11B. This works because the amplifier exhibits gain in the forward direction and loss in the reverse direction. Superficially, this seems to be the way to go because the amplifier can also improve broadband input return loss and can improve noise figure. Unfortunately, these benefits come at a price, as illustrated in figure 12.

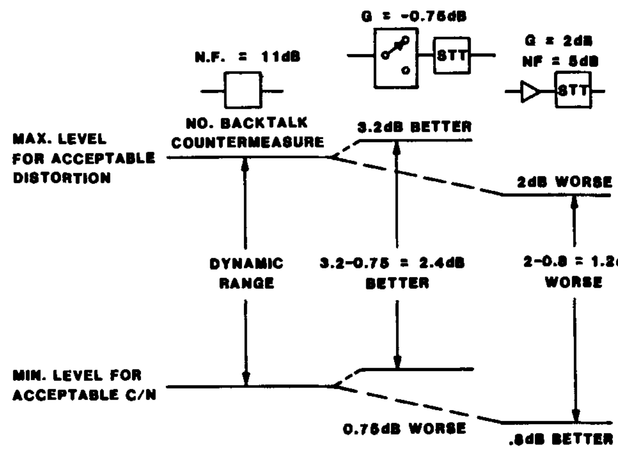


FIG. 12 EFFECT ON DYNAMIC RANGE

Figure 12 is divided into three parts. To the left is represented the dynamic range of a particular converter having neither backtalk countermeasure applied. Dynamic range is simply the window of input level which yields an acceptable picture. On the low end it is limited by noise figure, and on the high end by distortion. For the present purpose, we won't argue about the definition of where the picture becomes unacceptable, because we merely want to explore the change in dynamic range with backtalk countermeasures.

Returning to figure 12 we have on the left portion the dynamic range of a conventional converter. In the center we have added a switched filter having 0.75 dB of flat loss. To the right, we have added an amplifier to the basic converter. We compare what happens to the dynamic range of our basic converter when we add either the switched filter or the amplifier (but not both). We'll base our comparison on a 440 MHz system.

When we added the switched filter the signal hitting the mixer dropped by the filter's 0.75dB loss, so our minimum acceptable carrier level got worse by 0.75dB. However, we improved our situation on the high side. Empirically, we have determined that as more signals hit a mixer, the worst case composit triple beat increases by about

24.5LOG(NUMBER OF SIGNALS)(1)

A fully loaded 440 MHz system has 60 carriers on it. When the backtalk filter is in the lowpass position it will allow perhaps 38 carriers to the mixer, decreasing worst case CTB by about 4.86dB. Since CTB goes about 2:1 with level, this means we can increase signal level by half this, or 2.43dB before we reach maximum level. To this we add 0.75dB loss of the filter, permitting about 3.2dB increase in signal level. Thus, with the filter we improve high end acceptable level by 3.2dB while giving up 0.75dB on the low end. We now find that the terminal is 2.4dB more forgiving on input level as a result of adding the switched backtalk filter.

Now consider what would hapen if we had added the amplifier rather than the switched filter. We'll assume gains and noise figure as shown on figure 12. We calculate an overall noise figure of 10.2dB, an improvement of 0.8dB over the converter alone. This means that we can reduce the level at the low end by .8dB. On the other hand, the 2dB gain means that we must drop the maximum signal level by this much to avoid overloading the mixer (we assume that the amplifier doesn't add distortion). Thus, after adding the amplifier, the converter is 1.2dB less forgiving of signal level errors that it was previously. The difference in dynamic range between the switched filter and the amplifier is 3.6dB.

3. Noise Figure

We observed above a circumstance in which we improved dynamic range but at the expense of noise figure. Let us take a closer look at noise figure and the real effect it has on CATV system performance. The reader will recall that, by definition, noise figure (NF) is the excess noise introduced by an amplifier or other circuit element. Excess compared to what ? Every real resistance, including the 75 ohm source resistance that our distribution systems look like, generates noise as a result of random electron movement. But real electronic circuits exhibit even more noise than this, and we call the measure of extra noise generated noise figure. If an amplifier were perfect in this respect, it would have a NF of 0dB. A set-top with an 11dB NF contributes 11dB more noise than does an ideal terminal.

If we know the NF of a set-top and the incoming signal level, we can compute the resultant carrier to excess noise (C/N) ratio. If further we know the incoming C/N, we can compute the C/N ratio out of the terminal. If we also know the TV NF we can compute the C/N at the detector of the TV, the actual C/N seen by the subscriber.

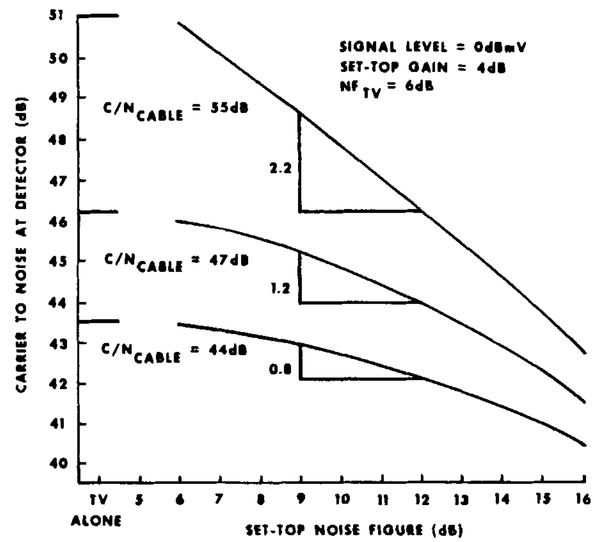


FIG 13 CARRIER TO NOISE RATIO vs. SET-TOP NOISE FIGURE

We have done all of this for several conditions, and show the result in figure 13. Here we plot C/N seen by the subscriber, v.s. set-top N.F., for three different cable C/N's. Other conditions assumed are shown. We now focus on cable C/N=44dB, representative of the end of a modern system. If the TV is directly connected to the cable, then the subscriber sees a 43.6dB C/N. If a set-top having a 9dB NF is inserted, the C/N becomes 43dB, while a 12dB NF terminal produces a 42.2dB C/N. Thus, improving the NF by 3dB in this case resulted in a C/N improvement of only 0.8dB seen by the subscriber.

From these curves, we can see that as noise figure is improved, pictures will improve, but not as fast as one

(1) CTB increases with the number of channels through two mechanisms: increased mixer voltage loading which goes as 20LOG(NUMBER OF CHANNELS), and additionally increases due to the larger number of CTB products on each channel. As the number of channels is increased, the channel of greatest CTB changes.

might hope. The reason is that we are partially limited by cable C/N and partially by TV N.F.

4. Input Return Loss

Finally, let's look for a moment at the effect of input return loss to the terminal. Return loss is defined as the ratio of incident signal to reflectetd signal. Compare waves on a pond which strike a piece of wood and bounce back. The ratio of the incoming (incident) wave to the reflected wave amplitude is the return loss. A return loss of 0 dB means all the signal is reflected (hence none is transmitted). An infinite return loss means that none of the signal is reflected, hence it must all be transmitted into the set-top.

What happens when signal is reflected from an input to a set-top ? It bounces back to the other end of the drop. If everything were perfect it would either be dissipated in the coupler resistor or the output impedance of the last amplifier, and nothing would happen. But taps and amplifiers also have finite return losses, so some of the signal is reflected again toward the set-top terminal. Or, due to limited isolation, the reflected signal may find itself on another drop, enroute to another set-top. Since the signal, wherever it arrives, will arrive later than the direct signal, it will produce that well known phenomenon known as a ghost.

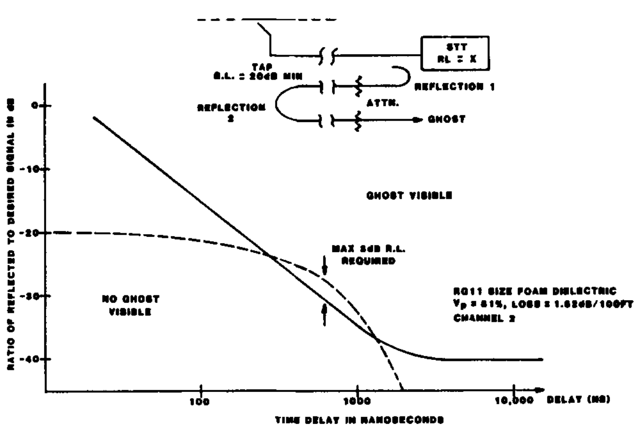


FIG 14 EFFECT OF ECHO

To analyze the severity of the ghost, we must determine the amplitude of the echo as it arrives at the set-top, plus its delay. To help, we invoke the standard Mertz curve, which defines the threshold of visibility of a ghost, as a function of its amplitude and delay. Figure 14 shows a popular incarnation of the Mertz curve. To utilize this information, Mr. R. Pidgeon of our staff has superimposed a curve showing round-trip attenuation and delay of an echo. The curve shown assumes a 20dB return loss at the last splitter (or tap). If we are analyzing the effect on a second set-top, we take this to represent a 20dB tap-to-tap isolation. Thus, Mr. Pidgeon has plotted delay vs amplitude for a particular cable, whose length is a parameter which shows only indirectly.

The curve shown is for an RG11 size cable having a loss of 1.62dB per hundred feet and a velocity factor of 81%. The cable characteristic plot invades the "ghost visible" area of the Mertz curve for a small region around 600nS delay. We can force the curve below the "visible" threshold by reducing the reflection by 3 dB. This may be achieved by using a terminal with a 3dB minimum return loss. Smaller sizes of cable will generally require even lower return loss to render the ghost invisible. Thus, we can see that a minimal set-top return loss will render ghosts invisible. Typical worst case return loss for current terminals range from about 6 to 8dB.

Another potential problem is that the reflection can affect frequency response, but calculation shows the effect to be negligible.

SCRAMBLING

No portion of a set-top terminal receives more attention than does the scrambling system. Virtually all of the popular scrambling systems in use today are based on rendering sync information indistinguishable from video. This is done variously by suppressing the sync pulses (horizontal and/or vertical) by attenuating the RF envelope, or shifting the baseband level prior to modulation. Descrambling is achieved by reversing the process. Synchronizing signals are inserted either within the vertical interval or as AM information on the sound carrier.

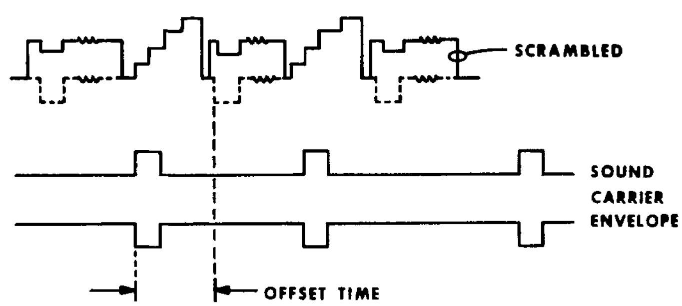


FIG 15 OFFSET SCRAMBLING TIMING

Figure 15 illustrates one such system. Shown at the top is a video waveform, which represents the bottom of a modulated RF envelope, this being an RF system. As shown, an attenuator is switched into the video IF path, reducing the amplitude by nominally 6dB during the horizontal and vertical blanking intervals. In this system pulses are placed on the sound carrier for synchronization, but they lead the horizontal blanking interval which they will decode, by a discrete length of time. Furthermore, this offset time is varied randomly from one field to the next, to make recovery by common pirate decoder boxes difficult. Extra pulses buried within a field are used to communicate timing information, which sets up a variable delay in a crystal controlled counter. For a pirate to duplicate this circuitry, he would need a rather large number of standard ICs. The manufacturer can accomplish the same end more economically because he has the volume to develop a custom IC.

Use of offset timing was developed as an improved security measure, but it has been found to offer further advantages. For example, when the video sync is restored by amplifying the RF, the sound carrier is also amplified, again by 6dB nominal. If the sound carrier has a 6dB pulse already at this time, then effectively the sound carrier is amplified 12dB from its nominal level during sync times. The sound carrier is thus more likely to crosstalk into the picture channel, creating ringing around the leading edge of sync. This effect has been observed, but the severity with different equipment is unknown. Also, with stereo this presents a problem because AM to PM conversion can cause errors in pilot amplitude and phase.

A particularly vexing problem with early realizations of scrambling systems using sound carrier synchronization was that the sound carrier had to be accurately tuned to recover pulses. Tuned radio frequency (TRF) receivers were used, and filtering was difficult due to the frequencies used.

A patented improvement makes use of the intercarrier technique to develop a 4.5 MHz sound (and sync) IF by mixing picture and sound carriers. This eliminates the critical tuning requirement of the earlier TRF approaches. By doing the filtering at 4.5 MHz, the filtering can be improved dramatically.

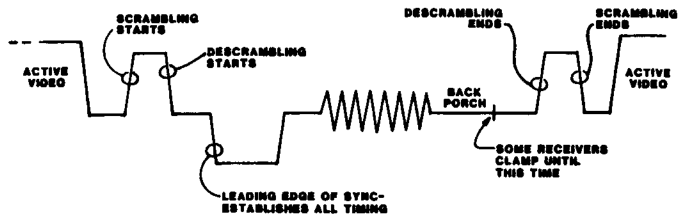


FIG. 16 DESCRAMBLED VIDEO

A compromise that must be made in descrambling is that the transition time of the descrambler must be controlled. Figure 16 shows this. The scrambling and descrambling processes must start and end in the order shown, without invading viewable picture area, clamp time or the leading edge of sync. When descrambling starts, the transition must be fast to complete itself before start of sync. On the other hand, if it is too fast, components beyond the bandwidth of the receiver are generated. In this case the receiver will generate Gibbs ringing which can overlap the leading edge of sync and cause timing errors.

Also, the scrambler must control its transition times to prevent ringing, though this is generally not as critical.

The scrambler must also follow a rather onerous set of standards for the sound carrier pulses. If the pulse transition times are too slow then timing errors can develop. If too fast, then spectrum overlap with on-or adjacent-channel video can occur.

A final requirement on the scrambler is that it should ensure that energy content on the video doesn't fall around the sound carrier, as this would result in timing errors. This is controlled by the insertion of a phase equalized lowpass filter in the scrambler baseband video loop. Characteristics of this filter should be checked when evaluating scramblers.

STEREO

A leading question in the industry today concerns the compatibility between scrambling systems and the stereo TV system. The format is nearly identical to the FM stereo format that has been with us for 25 years or so. Differences between the two can be summarized briefly.

1. TV stereo transmits a pilot at 15.734KHz (locked to the horizontal rate), rather than at 19KHz as in FM stereo.

2. TV stereo transmits a secondary audio program (SAP), FMd onto a carrier at 5 times the horizontal rate (78.67KHz). FM stereo may transmit an SCA on a 57KHz subcarrier.

3. TV stereo also transmits a low grade "professional channel" by FMing a subcarrier at 6.5 times the horizontal rate. This is intended to be used by a TV station for low rate data (e.g. telemetry from transmitter to studio) or intercom quality voice.

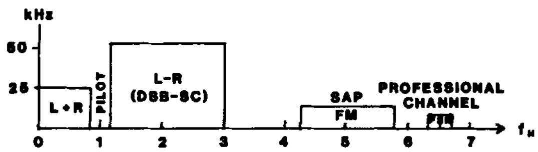


FIG. 17 MULTICHANNEL SOUND BASEBAND FORMAT

Refer to figure 17 for a spectrum of the proposed TV stereo baseband format (before modulation). As in FM stereo, a sum (L+R) signal is transmitted at its normal baseband frequency. Above this, the difference signal (L-R), is transmitted. In order to shift its spectrum, it is double sideband suppressed carrier (DSB-SC) modulated onto a carrier at twice the horizontal rate ($2f_h$). In order to recover the L-R signal the suppressed carrier must first be recovered. To do this, a pilot is transmitted at the horizontal rate. This pilot is doubled (usually in a phaselocked loop) and synchronously detects the L-R signal. The figure shows the SAP and professional channels, but let us cease our digression prior to talking about them. We have adequate background for the present purpose if we note that phase errors in the pilot result in phase errors in the L-R signal, and this causes separation errors.

Interaction between a scrambling system and stereo basically breaks down into two problems: what we do to them and what they do to us. We'll take the problems in reverse order.

What They Do To Us

The problem here is that we put amplitude modulation on signals that have been FMd with a deviation of $\pm$25KHz, and now we find ourselves facing a deviation nearly three times as great! What is bothering us is that in the signal path to recover our AM for synthronization, we have necessarily placed bandpass filters. Now bandpass filters by definition are not flat outside their passband, and they may not be totally flat within their passband. Thus, as the sound carrier's frequency is changed by the imposed frequency modulation, it may ride up and down a filter response, turning FM into AM (FM/AM). Filters that worked allright with 25KHz FM may have a problem with nearly 75KHz deviation. This could be conjectured to be a particularly onerous problem for systems that use sinewave modulation in a linear feedback loop, as the FM/AM component, now increased, is directly transferred to the video signal. Switched systems tend to be somewhat less susceptible as a result of their employment of a threshold at which a decision is made. Of course FM/AM can distort the decision point, resulting in the eventual collapse of switched systems as a result of switching time jitter.

What We Do To Them

Look back to figure 17 momentarily and observe the pilot signal, having a deviation of only 5KHz. Remember that the pilot is used to detect presence of a stereo telecast, switching in the stereo decoder. Now recall that many common RF scrambling schemes put AM on the sound carrier locked to the horizontal rate. Just as FM can get converted to AM, AM can get converted to FM. The mechanisms for doing this are legion, and include limiter errors and asymmetrically tuned bandpass filters. The more AM we put on the sound carrier the more spurious FM will get generated.

When this FM sound carrier, now contaminated with spurious FM modulation is detected, the resultant signal will have excess energy at f_h. If enough energy is present, the stereo pilot detector may falsely trigger on a monaural signal. Also, the artifact energy generated by the scrambling processs is at a random phase with respect to the pilot during a stereo broadcast. This causes a phase shift in the pilot carrier (now contaminated), which in turn results in a phase shift of the reinserted carrier. Finally, this results in loss of separation.

In summary, it is essential that AM/FM conversion be controlled to permit good quality audio on a scrambled stereo telecast. A portion of the responsibility for controlling AM/FM resides within the TV receiver or stereo adaptor. Obviously, the more AM that exists on the sound carrier, the more severe will be the problem. Recall that, by the time the receiver gets the audio signal, the signal has been amplitude modulated not only by the scrambling process, but by the descrambling process as well. The greater the scrambling depth the greater AM on the sound carrier. Also, if timing pulses are put on the sound carrier coincident with the horizontal blanking interval, then the pulse amplitude seen by the TV is the sum of the scrambling and descrambling pulse amplitudes. A particularly worrisome case is an on-time descrambling system having extra scrambling depth (e.g. 10dB suppression). With 6dB timing pulses, the audio carrier has pulses of 16dB seen by the TV audio system. This could give rise to considerable AM/FM conversion.

As of this writing, the above caveats remain speculative. Quantitatively the severity of the effects discussed is not known. Early tests tend to support conjectures made above, but more testing is needed before definitive statements can be made.

CONCLUSION

Even though this has been a cursory examination of the current state of the art in set-top terminals, the author has had to beg forgiveness for an excessively long paper. Many subjects were omitted or given only limited coverage. The role of data communications to and within addressable terminals merits much more attention (a slightly less cursory examination published previously by the author is available in the Winter '83 Addressability supplement published by CED, or from the writer.) Headend computer support systems are an integral part of addressability technology. One could write volumes concerning tradeoffs between security, cost, reliability and ease of use of addressable systems, including the role of various legal activities. Even the relatively comfortable area of RF specification needs much more work to determine the true optimum in cost vs performance tradeoffs.

ACKNOWLEDGEMENTS

Many people have contributed to the information pool that the author has drawn on in preparing this paper. Mr. R. Pidgeon, principal engineer, developed the analysis concept shown for evaluating echoes. Mr. A. Best, principal engineer, has been a prime mover for improved RF specification and stereo compatibility. Mssrs. G. Mobley and A. Kozushin developed many of the RF concepts. Mr. D. Foster, technician extraordinaire, made tedious measurements from which conclusions were drawn concerning CTB vs channel loading. Concepts of digital architecture and scrambling were contributed by M. Roth, R. Banker, C. Plonsky, J. Lappington and others.