Technical Details of the Proposed Base-Line Format of the Computer Industry Coalition on Advanced Television Service (CICATS)[1]

July 10, 1996

Overview

The CICATS[2] technical proposal for the US digital TV standard is briefly this:

- **Adopt ACATS Low Levels**: That the FCC adopt all ACATS proposals for modulation, error correction[3], data packetization, and compression for the new digital TV channels.
- **No Video Format**: That the FCC not specify a video data format.

In other words, adopt all low-level ACATS standardization proposals, where low-level means all levels except the video data level, which is not to be standardized by the FCC.

CICATS understands that the FCC may find it impossible to honor the second point above (No Video Format), in which case we propose an alternative second point:

- **One Required Video Format (Alternative)**: That the FCC specify a single 480-line (nominal), progressive-scan video format with square pixel spacing, utilizing a base-layer technology concept. Others could be implemented but only one would be required.

As will be explained below, CICATS actually couches this alternative as follows:

- **One Required Video Format (Alternative):** That the FCC specify the CICATS Reference Decoder.

This is as opposed to the ACATS proposal of 18 video formats that do not use a base-layer concept and that include interlaced formats. The CICATS alternative

---

[1] Prepared by Dr Alvy Ray Smith, Graphics Fellow, Microsoft Corp., Redmond, WA, with technical assistance from Gary Demos of DemoGraFX and Apple and Steve Gabriel, Dr James Blinn, and Dr Brian Guenter of Microsoft.

[2] See attached Glossary for expansion of all acronyms and definitions of technical terms.

[3] There is one caveat: The error correction capability of the digital channel, when used for non-video data, may need to be increased. See further discussion below.

proposal is cost-effective for consumers, immediately gives them higher resolution video, ensures smooth and true interoperability with computers, and is ready for improvements—such as even higher resolutions—as digital component costs drop.

On Point 1: Adopt ACATS Low Levels

The ACATS proposal (formerly the Grand Alliance proposal) has much to recommend it. At its most fundamental level, it proposes a completely digital standard, and we applaud this. Eight or nine years ago, when ACATS began its deliberations, analog systems were still being contemplated.

ACATS has invested much effort in testing the radio frequency and transmission system[4] of its proposal. We recommend, in particular, that its vestigial sideband (VSB) modulation subsystem be adopted. This is the subsystem that takes a digital stream of bits and modulates the radio frequency transmission carrier with it. The most time-consuming tests of the ACATS proposal were those that eliminated all other contenders for this level (the physical level) of standardization.

CICATS further supports these ACATS protocols and technologies: Trellis coding, Reed-Solomon error correction with interleaving, and Dolby AC-3 audio.

ACATS also proposes the use of the MPEG2 transport stream system for packetization of the bit stream. CICATS endorses this packetization level proposal from ACATS, to within the caveat on error-correction of non-video data already mentioned and further discussed below.

CICATS proposes that MPEG2 error correction be adopted for video data, as per the ACATS proposal.

On the Error-Correction Caveat to Point 1

The most general view of a digital channel, and one which CICATS highly encourages, is that it is a communications medium for arbitrary digital data. In this view, video data is simply one type of digital data that can be carried in the new digital channels.

CICATS proposes adoption of the ACATS error correction technology for video data, but the error-correction level may be too low for use by the new digital TV channels for transmission of many types of non-video digital data. The ACATS error correction level is sufficient for transmission of pictorial data (e.g., video) but not for general data. For example, if a new channel were used to download a program, then loss of data generally could not be tolerated. This is especially true in the headers for the transmitted data.

[4] The "RF/Transmission System" of the ACATS proposal.

CICATS proposes that the FCC endorse the ACATS position on the low-level protocol but with the addition of a mandate to determine what needs to be done to the standard to increase its error-correction capabilities. This could take the form of an FCC specified committee to report on the problem and suggested solutions to it within one year from propagation of the initial digital TV standard. We believe it is important to officially pursue the problem with a time limit.

It is important to understand, however, that by delaying this study, a solution to the problem may be precluded. It may be that the structure of MPEG2 transport headers do not permit a sufficiently robust solution, and that a reasonable solution would require changes to the MPEG2 standard. If this standard has already been mandated and placed into use, then it cannot be changed.

Point 2: No Video Format

The computer industry is well aware of the astonishing rate of change in its underlying digital technology. It prospers because of it. It knows that it is difficult, if not impossible, to predict order-of-magnitude phenomena that happen regularly every five years in the high-technology sector (the Moore's Law phenomenon, see Glossary). So the CICATS position is that the FCC specify the fundamental protocols for its new digital TV channels but refrain from overly restricting what forms the data in those channels may take.

The most recent success in high technology, based on a similar strategy, is the Internet. Only two years ago it would have been impossible to predict today's businesses and world-wide impact of the data applications built atop the fundamental protocols of the Internet. We believe the inventiveness that industry would bring to bear on the new digital TV communications channels, to be mandated by the FCC, would be phenomenal and do not want to see any formats—specified now in the infancy of the digital world—hinder that innovation.

Hence the CICATS proposal is that the FCC not specify any video formats in particular but let market forces and economic considerations dictate how the channels get used. It is conceivable that the TV and PC industries would decide to proceed with the ACATS video formats, but we believe that the standards that would actually arise would be considerably different.

We can support the ACATS standard in many areas, as described above: VSB modulation, trellis coding, Reed-Solomon error correction with interleaving, MPEG2 transport and packet protocol, Dolby AC-3 audio and, most importantly, MPEG2 video compression. In the best of all possible worlds, we would revisit the adoption of MPEG2 itself (the real meaning of no video format), but in this proposal we are accepting the ACATS choice. Where we diverge is that we see no need for the artificial restriction of the flexibility inherent in MPEG2 to a fixed set of resolutions.

Table 3 of section 5.1.2 in Annex A of the ATSC standard[5] is the heart of our problem. It constrains the vertical and horizontal size values to specific numbers instead of simply allowing anything under a maximum value, as MPEG2 does. ACATS tries to put display constraints *in the transmission standard*, where they don't belong. When these constraints are removed, then any aspect ratio image can be sent through the channel. It would then be up to the receiver to display what it can by either pan-and-scan or letterboxing, or a combination of the two. This would satisfy the objections of the film industry, since there would be no difficulty sending 1.85:1 or 2.37:1 or any of the other film aspect ratios as long as the horizontal and vertical sizes stayed below defined limits.

We address below the possibility that the FCC is under so much pressure from video broadcast manufacturers that it is unwilling to trust the market pressures as we propose. Hence we provide our alternative second point.

Alternative Point 2: One Required Video Format

Should the FCC feel it needs to specify a video data format, then CICATS proposes that it be as simple and extensible as possible, while displaying an immediate quality advantage over the current analog TV broadcast familiar to all US citizens. Furthermore, we propose a single format that is fully compatible with modern computer technology, without the quality or cost compromises intrinsic to the ACATS proposal.

The format we propose is a base-layer concept (see the Glossary), which means by definition that it is extensible to higher resolutions and better quality. We differ from the ACATS proposal in that we do not require that higher resolutions be standardized at this time. We feel that it is premature to do so, and that the market does not yet support higher resolutions such as the interlaced 1920x1080 ACATS format. We do believe, however, that in a few years, the inexorable Moore's Law phenomenon will make the higher resolutions cost effective, and want a base-layer technology in place so that logical extensions of it to higher resolutions are possible when the costs drop sufficiently—should the market then demand it.

In order not to slow down the FCC decision process, we propose a single format that is easily within today's capabilities and has been demonstrated so.

Furthermore, we are very desirous that PCs and TVs become tightly coupled devices, sharing data and even computing on each other's data—true

[5] Doc. A/53.

interoperability, not simply side-by-side display. Any video data format should encourage this possibility, not discourage it or make it costly.

Hence CICATS proposes a video format, by way of a *Reference Decoder* (see Glossary), that

- Has 480 lines (nominal) of vertical resolution
- Is progressively scanned
- Has square pixel spacing
- Employs a base-layer design for future, logical, cost-effective extensions to higher resolutions.

(See resolution, progressive, interlace, square pixel spacing, and layering in the Glossary.)

The 480 lines is numerically the same as today's analog TV standard but, because the lines are progressively scanned rather than interlaced, the quality level is immediately higher than today's analog TV. It takes about 768 lines of interlaced video to equal the same perceived quality level of 480 progressive lines.

The base-layer concept guarantees that no digital TV set will ever go dark as newer and better extension layers are built atop it—a set bought honoring the initial standard will always continue to work even as the extension layers are added in the future. Any future TV broadcast at higher resolution would still display at the base-layer resolution on any initial set. This is because the higher resolution would consist of the base layer resolution plus an enhancement layer that is added to it to form the higher resolution. The base layer would simply be stripped off by the older set and displayed. The newer set would combine the two layers and display the higher resolution picture.

One of the advantages of the CICATS proposal is that 4-5 times less memory is required in a CICATS receiver than in an ACATS receiver that must decode all 18 ACATS formats, including high-resolution formats that are prohibitively expensive for most people today. They would have to buy a receiver with all the memory despite the fact that they needed far less. And they would have to pay for it. We believe this is an unnecessary burden on the consumer. Our base-layer concept gives higher-resolution to the consumer when the costs for it have dropped substantially in about 5 years—while increasing the quality of TV even at the base layer initially offered.

All computer displays are progressively scanned, so CICATS believes it is essential to rid the national standard of old-fashioned interlaced scanning formats. Unfortunately, the ACATS proposal includes several interlaced formats. If even one is allowed then all receivers (including all PCs) would have to provide for the conversions. These conversions are unnecessary and are difficult to do with quality. Cost pressures would dictate that they would be done at low quality.

Furthermore, all computer displays have square pixel spacing (see Glossary) and therefore so does the CICATS video format. The ACATS proposal includes several formats with non-square pixel spacing. Again, if even one is allowed then all receivers (including all PCs) would have to provide for the conversion.

Two further specifications are required to fully describe the CICATS single video format: horizontal resolution and temporal resolution (frame rate). The format should have:

- **Spatial Resolution**: A ***spatial base layer*** with horizontal resolution determined by the CICATS requirement for square pixel spacing. For example, a TV set with an aspect ratio of 4:3 and 480 lines of vertical resolution would have 640 pixels horizontally.
- **Temporal Resolution**: A ***temporal base layer*** supporting 24, 36, and 72 Hz frame rates.

The notion of a temporal base layer is a new one to these FCC-related discussions and needs some explanation. For example, it might appear that we are proposing three video formats here, one each at 24, 36, and 72 Hz. This is not the case and here's why:

In case of three separate formats, the broadcaster selects one of the three to transmit and the receiver detects which one is sent and converts, if necessary, to its local frame rate. Frame rate conversions are the most difficult, of all the conversions implied by the ACATS proposal, to do with quality at a low price.

In the case of a temporal base layer, all sets would implement the base layer (by definition of a base layer), hence all three frame rates would be implemented. Regardless of transmitted frame rate, a set receiving the proposed temporal base layer signal would operate at 72 Hz frame rate. It would select and decode the appropriate MPEG2 frames (I, P, and B frames in MPEG2 terminology) to form the 72 Hz display. The base layer technology makes this simple to do. It is a *selection* process rather than a *conversion* process.

It is important to note that the CICATS temporal base layer does not support 30 Hz or 60Hz. 30 Hz is a relic of interlaced scanning so is not needed in the progressively scanned future. The PC market has determined that 60 Hz is insufficient so it is not included in the CICATS temporal base layer.

But CICATS, again, understands that the FCC might have to support 60 Hz under pressure from the old analog world. In this case, we propose an alternative to the temporal base layer:

- **Temporal Resolution (Alternative)**: 24, 60, and 72 Hz frame rates. Not a temporal base layer.

This alternative does extend the CICATS proposal to three video formats, but the three differ only in frame rate. Although we offer this alternative, we want the FCC to understand that it implies conversion hardware and more memory in the receiver, hence more cost to the consumer. Furthermore, the conversions between 60 and 72 Hz are particularly prone to poor quality. Nevertheless, 60 Hz display displayed on 60 Hz sets and 72 Hz displayed on 72 Hz sets would suffer no quality loss.

Spatial Resolution: Reference Decoder

One way to look at the CICATS proposal is that it severs the decision to go digital from the decision to go high-resolution (or "high-definition"). We believe that going digital is the fundamental revolutionary step. We want to concentrate on doing it right. We believe that adding high resolution is straightforward if the groundwork is in place. We encourage adoption of a posture that allows this to happen when Moore's Law makes it more economically feasible than now, in about 5 years. We re-emphasize, however, that the CICATS base layer alone has higher perceived resolution than today's TV.

The preferred way to think of the CICATS proposal is in terms of a reference decoder. The CICATS Reference Decoder has a memory capable of supporting 1024 horizontal by 512 vertical pixels. This plus the requirement for square pixel spacing implies that the Reference Decoder is capable of decoding any resolution up to and including 1024x512. The following table shows several examples supported by the Reference Decoder on TV displays of various aspect ratios:

Aspect	Horizontal	Vertical	Remarks
1.33:1 (4:3)	640	480	Current TV format
1.78:1	854	480	Approximately the ACATS 16:9 format
1.85:1	944	512	Most popular Hollywood format
2:1	960	480	Acceptable to Hollywood
2:1	1024	512	Acceptable to Hollywood
2.37:1	1024	432	Popular widescreen Hollywood format

Rather than propose a single video format, CICATS proposes that the FCC mandate the Reference Decoder. Then the choice of horizontal resolution becomes a secondary choice. This choice would be left to industry—that is, to market demand.

The CICATS Reference Decoder is a way of specifying a class of video formats acceptable to the computer industry. It is a hardware specification to the same degree that an ACATS video format is a hardware specification. That is, it puts requirements on the hardware but does not specify the implementation that satisfies them. Following are some example uses of the Reference Decoder.

There are arguments for the choice to transmit 640x480 pixels: It is consistent with today's capabilities. Progressively-scanned 640x480 systems have already been demonstrated. The costs and demands for this resolution are well known. The aspect ratio of 4:3 is the current one for which CRT (cathode-ray tube) technology is already well-suited and cost effective. Computer displays are as comfortable with this format as are TV displays. The cost of a converter for one of today's analog sets to receive the new digital signal is minimal for this resolution. CICATS believes this format to be the one most likely to appeal pricewise to consumers now, encouraging them to convert to the new digital standard and thereby release the old analog spectrum.

There are, however, good arguments for other choices within the set allowed by the the Reference Decoder. Consider, for example, 1024 by 512 pixels, the maximum allowed by the Reference Decoder (base layer only). The vertical resolution would be higher than today's analog TV because 512 is greater than 480, but more importantly because progressive 512 lines is equivalent to about 780 interlaced lines. And the horizontal resolution (on a TV set with aspect ratio 2:1) would be very much higher than today's analog TV, as well as spread out much wider. 2:1 aspect is considered desirable by Hollywood. Enhancement later (with an enhancement layer added to the base layer) to a nominal 2048x1024 resolution would be straightforward.

But there are serious counterarguments against the 1024 by 512 choice. The most serious is that displays for such an aspect ratio have not been demonstrated. Even if they were, they would probably be exorbitantly expensive at this time. So the same argument we levy against the expensive ACATS array of formats holds against this format too: Only the wealthy would be able to afford it at first. Sets that displayed in the old 4:3 aspect would either have to letter-box the wide aspect ratio, or pan-and-scan in it, or both (MPEG2 supports all of these choices). Both of these are familiar practices in widescreen films broadcast on TV today. All sets would implement the Reference Decoder but only those capable of 2:1 aspect would get full benefit of the signal.

Notice that a format with approximately 16:9 aspect could be chosen within the parameters of the CICATS Reference Decoder. This is one of the ACATS proposed aspect ratios. This aspect ratio has some of the same problems as just discussed for the 2:1 aspect ratio. In particular, sets to display at that ratio are too expensive for the average consumer. It is not an interesting aspect ratio for Hollywood. On the other hand, CRTs of that aspect have been demonstrated.

Pan-and-scan or letterboxing would be required for satisfactory display on sets of smaller aspect ratio, as discussed above for the 2:1 ratio.

In any case, the new digital TV sets would implement the Reference Decoder. They would need 4-5 times less memory than the equivalent ACATS-compliant set so would be optimally cost effective for consumers—*and at no loss in quality* implied by the conversions required between the 18 ACATS formats at the receiver. The CICATS proposal would be cheaper and better. Over time the cost differential between the two types of sets would diminish (with Moore's Law again) but in the meantime, US consumers would have paid many billions of dollars for unnecessary conversion and suffered unnecessary loss of quality as well.

Spatial Resolution: Enhancement Layers

Since the submissions to the FCC so far have confused the change to digital with the change to high resolution, it is important that the CICATS proposal not be interpreted as sacrificing the push toward higher resolution. Higher resolution is desirable to all members of CICATS and to all consumers. We argue that most consumers will not be interested *so long as the price tag is in the thousands of dollars per set,* as it is today. We believe that high resolution will certainly arrive, as costs drop with Moore's Law, and hence are proposing a base-layer / enhancement-layer technology that paves the way (see layering in the Glossary).

We propose that the FCC suggest—as a *recommended practice*—how enhancement layers might be used atop the standardized base layer to reach higher resolutions immediately, if certain market sectors wish to pursue it now. We emphasize that this should not be mandated.

More important than a specific enhancement layer is a *process* for adding enhancement layers and how they work. Following is one of many possibilities[6]:

Suppose that the 1024x512 resolution discussed above has been chosen with the framework of the Reference Decoder. Suppose further that memory prices have dropped substantially so that now a memory of 2048x1024 pixels is as cheap as 1024x512 pixels is today. Moore's Law tells us this will happen in a few years. At that time the PC and TV industries might decide—or the FCC might decide—that it is time to add an enhancement layer to the base layer already in place in the national digital TV standard. An example enhancement layer that would work in this case is this:

[6] See comments of DemoGraFX, Appendix K, for details on several of these possibilities. The one used here is Resolution Enhancement Mode 4.

The base layer is expanded by 3/2 to get a resolution of 1536x768 pixels. The difference is sent in an enhancement layer. 3/2 is a straightforward multiplication factor for digital images[7]. Notice that 768 lines, progressively scanned, is perceptually equivalent to about 1180 lines interlaced, greater than the 1080 lines interlaced in the ACATS highest proposed resolution.

CICATS has determined, from the work of Gary Demos of DemoGraFX, that all of this fits into a given 6-megaHz digital TV channel using MPEG2 compression technology.

Temporal Resolution: Temporal Base Layer

This is an example of another new idea that goes beyond simply digitizing the analog video world as currently understood (the idea of a reference decoder being the other). The notion is to apply the base-layer / enhancement-layer concept in the time dimension as well as in the space dimension. The receiver of a digital TV signal would select among the various entities provided by the MPEG2 bitstream.

DemoGraFX[8] has proposed a temporal base layer that supports either 24 Hz or 36 Hz frame rates and a temporal enhancement layer that goes up to 72 Hz. Since CICATS proposes 24, 36, and 72 Hz frame rates, it is convenient to think of the DemoGraFX temporal base layer and enhancement layer to 72 Hz as a single CICATS temporal base layer.

As explained above, regardless of transmitted frame rate, a set receiving the CICATS temporal base layer signal would operate at 72 Hz frame rate. The Reference Decoder, which incorporates the CICATS temporal base layer, would select and decode the appropriate MPEG2 frames (I, P, and B) to form the 72 Hz display.

This technology has been demonstrated by Gary Demos of DemoGraFX and formally reported to the SMPTE (Society of Motion Picture and Television Engineers)[9]. Displays capable of 72 Hz are commonplace today in computers.

Just as the Reference Decoder supports many more spatial formats than will actually be used, it also supports more temporal formats than will probably be used. For example, the 36 Hz format might not typically be used, but it comes "for free" with the CICATS Reference Decoder (just as the other spatial formats do).

---

[7] op cit. (DemoGraFX)

[8] See comments of DemoGraFX, Appendix J.

[9] See comments of DemoGraFX, Appendix I.

Temporal Alternative: 3 Frame Rates

CICATS understands that there is immense pressure from the existing analog TV industry to maintain the current frame rate of 60 Hz. CICATS would prefer to see new and legacy material at this frame rate converted, at high quality, at the head-end before transmission over the new digital channels. However, we are willing to expand our one simple format (expressed in terms of the Reference Decoder) into three formats that differ only in frame rates: 24, 60, and 72 Hz. This does not have near the elegance of the temporal base layer proposal but it does allow continued use of 60 Hz.

Use of three formats requires interconversion. Interconversions have been our argument against the ACATS proposal because they typically imply loss of quality in order to meet realistic consumer price expectations. However, the very difficult conversions between interlace and progressive scan are not asked for in this alternative 3-format proposal, and spatial conversions are not required either. Only temporal conversions are required.

The most tricky conversions are between 60 and 72 Hz because they are relatively so near one another. Conversions tend to show artifacts related to the 12 Hz difference, a very visible frequency for the human eye.

Furthermore, conversions imply more machinery and hence higher costs.

Finally, the three-format scenario is not a base layer concept so there is no clear enhancement path to higher frame rates in the future.

Nevertheless CICATS would compromise to three video formats differing only in frame rate so long as the FCC understood the interconversion problem and its consequences.

Summary

CICATS proposes that the FCC mandate the ACATS low-level protocols for the new national digital TV broadcast channels.

CICATS proposes that the FCC not mandate video data formats for the new digital channels.

CICATS proposes that the FCC institute a study group chartered to return a finding within one year on how to improve the low-level digital TV protocols by several orders of magnitude to accommodate error-free transmission of non-video or non-pictorial data. It is understood that a satisfactory solution might be precluded by proceeding with the low-level protocols before this study is made.

CICATS recognizes certain political realities that may cause the FCC trouble in pursuing the proposals above. The following alternative proposals are aimed at alleviating these problems, so long as it is understood that CICATS believes them to be inferior positions.

CICATS alternatively proposes that the FCC mandate the CICATS Reference Decoder that handles up to 1024 pixels horizontally and up to 512 lines vertically (without enhancement), is progressively scanned exclusively, has square pixel spacing exclusively, is a spatial and temporal base-layer technology, and supports frame rates of 24 Hz, 36 Hz, and 72 Hz (without enhancement).

CICATS alternatively proposes that the alternative immediately above be further modified only by substitution of three frame rates for the temporal base layer concept, the three rates being 24 Hz, 60 Hz, and 72 Hz. It is understood that by so doing conversions are required in the receiving sets, implying cost and quality penalties (but far less than those associated with the 18-format ACATS proposal).

CICATS further proposes that the FCC suggest, as recommended practice, how the CICATS spatial base layer might be enhanced to higher resolutions. These enhancements are not to be mandated at this time.

CICATS proposes that the FCC recommend that old analog content be used only on old analog channels, or else be converted at high quality at the transmission head-end to the new digital signal for use on the new digital TV channels.

Glossary[10]

ACATS: Advisory Committee on Advanced Television Service, to the FCC.

Aspect ratio: The ratio of the width of a picture to its height. Standard (current) TV has an aspect ratio of 4:3 ("4 to 3") = 1.333. The ACATS proposal mixes 4:3 with 16:9 aspect ratios. 16:9 = 1.777 is a strange aspect ratio that is wider than current TV but is not a Hollywood compatible aspect ratio. Hollywood films are most often in 1.85 ("academy") aspect or in 2.37 ("scope") for very wide-screen films. Hollywood would apparently be content with a 2:1 aspect ratio, but not with 16:9.

Base Layer: See layering.

CICATS: Computer Industry Coalition on Advanced Television Service, representing 10 leading personal computer companies (hardware and software).

FCC: The Federal Communications Commission.

Frame rate: The number of video pictures displayed per second. The goal is to seem continuous. Film's frame rate is 24 frames per second, where each frame is repeated 2 (or sometimes 3) times by a film projector to give the equivalent frame rate of 48 frames per second (or sometimes 72). The word "Hertz" is used often to abbreviate "frames per second". The highest ACATS frame rate is 60 Hz

[10] Underlined words refer to defined terms elsewhere in the Glossary.

("60 Hertz" or 60 frames per second), whereas computer consumers rejected 60 years ago in favor of 70 or more frames per second to avoid objectionable flicker. (Looking at a TV or PC screen out of one's peripheral vision reveals the flicker.) 72 Hz is an attractive frame rate because it is computer friendly and an easy multiple of film rate (film is a major source of all TV content).

Hertz (Hz): One Hertz is short for one cycle per second, or one frame per second. Frequencies were formerly expressed in cycles per second—for example, a radio station might broadcast at 98.1 on the radio dial, meaning at 98.1 megacycles per second. Today this would be expressed as 98.1 megaHz, in honor of electromagnetic pioneer Heinrich Hertz. In a related usage, the "width" of a TV channel is measured in Hz—6 megaHz per channel.

Interlace: Current analog TV scans each frame by first drawing every other horizontal scanline across the face of the TV set, then starting over at the top and drawing all the skipped in-between scanlines. The first set, called a "field", is said to be interlaced with the second set, or second field. Interlaced scanning is opposed to <u>progressive</u> scanning.

Layering: A layered system is a logical system of related frame sizes, rates, and resolutions (as opposed to a grab-bag of unrelated formats as in the <u>ACATS</u> proposal). A layered system has a "base layer" that must be honored plus "enhancement layers" that may be added to the base layer to make it higher resolution. A good example of a layering scheme is that used by Kodak's PhotoCD. Snapshots are taken to a photo house from which they are returned in digital form on a CD, Kodak's PhotoCD. Each of the snapshots will appear on the CD in several resolutions. The base resolution is 768x512 (approximately video resolution), but the CD also contains enhancement layers that are added to the base resolution to make it into 1536x1024 pixels or 3072x2048 pixels. So one CD contains at least these three resolutions. Similarly, a layered TV channel could contain several resolutions simultaneously so long as they were layered logically. <u>ACATS</u> misuses the term "layering" to simply mean a TV picture is layered atop a string of digital bits, which is layered atop a radio frequency modulation technique. Theirs is a much more generic use of the term than the <u>CICATS</u> (or Kodak) use.

Moore's Law: The "law" that says computers get twice as fast every 18 months. In general, anything digital gets twice as good every 1.5 years. For example, memory doubles or the processor gets twice as fast—for a fixed cost—every 1.5 years. To understand how stunningly fast this is, let's restate it as 10 times faster every 5 years (that's the same as 2 times faster every 1.5 years). During the 8 years that ACATS has been working on its proposal, personal computers have increased in speed and memory by a factor of 50 to 100 times (at the same cost). At the beginning of the ACATS process, PCs weren't powerful enough for TV, but now they are. There is good reason to believe that Moore's Law will continue to operate for another 15 years - thus for another improvement of 1000 times over what we have today! This incredible digital revolution is what makes CICATS encourage the FCC not to freeze any digital standards now that it could better make 5-10 years from. We are simply incapable of predicting what an "order of magnitude" (10x) change means conceptually. Any standards made now will look foolish 5 years from now, so only the minimum should be done now. (Two years ago there was no Netscape, and Microsoft was not an Internet company. Things change *very* fast in the digital world. The old analog modes of thinking do not work.)

Order of Magnitude: One power of 10. So 100 is two orders of magnitude larger than 1. The term "order of magnitude" means more than simply a larger number, however. It implies a conceptual change as well. Moore's Law says that computers get faster by an order of magnitude every 5 years (at a fixed cost), but more importantly it also means that we require a different level of understanding every 5 years.

Pixel: Short for picture element. In the digital world, a picture is represented by an array of tiny samples or picture elements - so many per line and so many lines. (Pixels, by the way, are single points, not little squares or rectangles as popularly described. We are careful to say "square pixel spacing", not "square pixels".)

Progressive: Current PC screens draw each scanline in order from top to bottom. They are said to be "progressively scanned". This is opposed to interlaced scanning.

Reference Decoder: CICATS proposes a layered video format scheme by way of a reference decoder, which is a specification of the decoder of the new digital TV signal—separate from the display of that signal. This concept permits a degree of freedom not formerly present in these FCC-related discussions. For example, instead of specifying a specific horizontal resolution, which depends highly on the capabilities of a particular display, the CICATS Reference Decoder says only that the format must have 480 progressive lines (nominally) and square pixel spacing. So, if the display device has electronics and width enough to handle a 2:1 aspect ratio, then the Reference Decoder will honor a width of 960 pixels (assuming the industry has agreed to broadcast this signal). If the display device can only handle 4:3 aspect ratio (as the affordable ones today do), then the Reference Decoder would dictate a horizontal resolution of 640 pixels. The same decoder circuit, at the same parts cost, would handle either situation. The Reference Decoder is not hardware. It is a way of specifying a class of acceptable video formats rather than a single video format. Within this class there is no conversion required, but wide signals (wide aspect ratio) would have to be either letterboxed or pan-and-scanned to a display with smaller aspect ratio.

Resolution: The number of pixels per line and the number of lines - equivalently, the number of horizontal pixels and the number of vertical pixels.

Thus a resolution might be given as 2048x1024 pixels, meaning 1024 scanlines with 2048 pixels on each scanline - equivalently, a rectangular array of 2048 times 1024 pixels (about 2 million pixels). Standard (current) TV has 480 lines vertical resolution and about 700 horizontal. But it is interlaced, which brings down its effective vertical resolution to about 320 scanlines. The ACATS "high-definition" format of 1080 vertical lines is really about 700 lines since the 1080 lines are interlaced at the 60 Hz frame rate.

Spectrum: Simply all the channels used for TV, cable TV, AM radio, FM radio, ham radio, and so forth - treated as a single entity. The full electromagnetic spectrum is vast, including X rays, heat, and even ordinary light. The FCC has dominion over only the "radio frequency"—those uses listed in the first sentence of this paragraph. That is, the radio frequency spectrum is a subset of the full electromagnetic spectrum. One TV channel, whether old analog or new digital, is a slice of the radio frequency spectrum.

Square pixel spacing: This just means that the horizontal spacing between pixels is the same as the vertical spacing between pixels (or between scanlines). Although the conversion from non-square pixel spacing of many of the ACATS formats to square pixel spacing is straightforward, there are about 200 million PCs in existence that assume square pixel spacing and do not have the software for doing the conversion.

Exhibit C

Cost Comparison of ACATS and CICATS Set-top Converters, Receivers, and PC Decoders[1]

Introduction

Adoption of the ACATS standard portends a transition to digital TV (DTV) that will be enormously costly to consumers. In order for a consumer to receive any of the digital channels when these go on the air in 1998, or if the NTSC channels go dark at the ten-year mark as some propose, the only way for a consumer to use a legacy NTSC receiver to view DTV broadcasts will be to purchase a set-top converter box. These devices will not be inexpensive.

The cost of ACATS converters will be substantially greater than the cost of a converter that would be required under the CICATS proposal.

While the precise engineering and design specifications for DTV converters has not been established, it is possible to develop broad-guage estimates of the price levels that consumers can expect to confront at various stages of the migration period. An analogous product, already on the market, is a digital satellite service (DSS) television receiver, such as the DSS decoder.[2] This device receives a compressed digital signal from the satellite, decodes the signal and converts it into analog NTSC form suitable for display on a conventional NTSC receiver or monitor. Hence, the current price of a DSS receiver provides a real-world starting point for our cost model.

The modulation scheme used by DSS is similar to that being proposed by ACATS, providing approximately 20 Mb/s per satellite transponder. Four or five MPEG-2 MP@ML SDTV streams are multiplexed together on one of these digital data channels.

---

[1] This analysis was prepared for CICATS by Steve Gabriel, Architect, Graphics and Video Systems, Microsoft Corporation.

[2] DSS receivers are currently selling in stores at a retail price of approximately $500. However, if that retail price is in any way subsidized by the DSS provider or equipment manufacturer in order to encourage subscribership to the satellite service, the effective retail price of the unit would be considerably higher.

The ACATS system can also do this in one of its modes of operation. Even though DSS brings only low-resolution, SDTV programming into the home, the multiplexing of the data brings the complexity of the demodulation and control section (front end) of the receiver to about that needed by an ACATS converter. The MPEG-2 decoder is also similar to that required by the ACATS standard, with one very important difference: An ACATS HDTV decoder needs to process data five times faster and uses five times more memory than is required for DSS. For complex processes, these requirements translate almost linearly to the cost of the devices involved. Hence, we assume that the full-function ACATS HDTV-capable MPEG-2 decoder section will cost approximately five times as much as the decoder used in a SDTV DSS receiver.

Description of the cost model

The cost model breaks the DSS receiver into three sections:

(1) Packaging and power supply hardware;

(2) Demodulator and control circuits; and

(3) the MPEG-2 decoder.

The cost of the VLSI (Very Large Scale Integration) electronics that are employed in the digital converter units is expected to decrease steadily over the next 10 to 15 years, according to industry trends that are encompassed in a rule-of-thumb known as "Moore's law." Moore's Law states that the price of computation and memory halves every 18 to 24 months. Taking the more conservative end of this range, we assume here that costs will drop by 50% every 24 months.

The first section — Packaging and power supply hardware — does not involve VLSI electronics and is thus not subject to Moore's Law. In our model, we assume that the cost of this section remains essentially constant over time. These costs are, of course, sensitive to overall production volume; hence, as demand (output) grows, small decreases in the packaging and power supply costs can be expected. Digital satellite receivers are already being manufactured in reasonable volume (2 million so far, another 4 million projected), so using present day prices for this section is not unreasonable.

Sections (2) and (3) — the Demodulator and control circuits and the MPEG-2 decoder — use VLSI electronics and will thus respond to Moore's law.

Cost Comparison of ACATS and CICATS Converters

Structure of the cost model

The cost model contains the following variables:

PDSS	the present price of a DSS receiver
HW	the present fraction of the price of the DSS receiver that is not subject to Moore's law cost reductions, *e.g.*, case, power supply, circuit boards, etc.
VLSI	the remaining fraction that is subject to Moore's law, the VLSI electronics. HW plus VLSI add to one.
MPEG	the fraction of the VLSI electronics devoted to the MPEG-2 decoder
Half_life	Moore's law scaling factor, the number of years for price to drop in half.

The CICATS base-line converter unit

Our prediction of the price of a DSS set-top converter is given by the formula

PDSS_future(t) = PDSS * [HW + VLSI * 2 ^ (-(T-1996)/Half_life)].

where T is the year. The exponent of 2 is the negative of the number of years into the future divided by the estimated half life of the price of the VLSI.

The CICATS base-line SDTV decoder is substantially equivalent in complexity to a DSS decoder, so the future price of a DSS receiver is a good predictor of the future price of a CICATS base-line SDTV decoder.

The ACATS set-top converter unit

To predict the price of an ACATS set-top converter, we use the formula:

PACATS(T) =
PDSS * [HW + VLSI * ((1-MPEG) + 5 * MPEG) * (2^(-(T-1996)/Half_life))]

Here we break the VLSI section into the non-MPEG part (which is essentially the same for both DSS and ACATS — the "1-MPEG" term), and the portion that is devoted to MPEG decoding, which is five times more expensive for full ACATS capability.

PC add-in converter/decoder cards

We can also estimate the prices of add-in cards for PCs, or of the electronics for the digital tuner and decoder of an ACATS HDTV set, by simply deleting the HW portion of the price. We assume that the *entire* collection of electronics responds to Moore's law. This very generous assumption probably underestimates the price.

Cost model results

For our model, we selected 15% as the fraction of the set-top box that does not respond to technology scaling. This may be low. For the fraction of VLSI devoted to MPEG, we used 50%. With a halving time of 2 years, this results in the following table comparing CICATS base-line converter costs to that for full ACATS-capable units. The first two price columns are the selling prices of set-top converters for existing NTSC sets. The next two columns are the incremental costs for a PC decoder card or the VLSI electronics inside the TV set.

Assumptions

HW	15%
MPEG	50%
1996 Price of DSS or CICATS	$500
Half life	2 years

Table 1

Cost Comparison of ACATS and CICATS
Set-top Converters, Receivers, and PC Decoders*

	Set-top Box		**TV Internals or PC Decoder**	
	CICATS	**ACATS**	**CICATS**	**ACATS**
1996	$500	$1,350	$425	$1,275
1997	$376	$977	$301	$902
1998	$288	$713	$213	$638
1999	$225	$526	$150	$451
2000	$181	$394	$106	$319
2001	$150	$300	$75	$225
2002	$128	$234	$53	$159
2003	$113	$188	$38	$113
2004	$102	$155	$27	$80
2005	$94	$131	$19	$56
2006	$88	$115	$13	$40
2007	$84	$103	$9	$28

---

* Prices shown are in constant 1996 dollars.