New video compression standards: meeting the test challenges

Imagine...
...being able to see high quality, high definition video, on demand.
...renting a DVD with choices of standard or high definition video, with multiple camera views and other options
...receiving good quality video on a mobile device – the latest goals in soccer, movie highlights, or maybe even the whole movie.

These are increasingly the demands of consumers.

Tektronix
Enabling Innovation

As a video professional, you will want to fill this demand – with more channels, more choices and better video, including HD –but without the expense of having to establish enormous new infrastructures.

H.264/AVC, VC-1: the new video compression technologies offer the possibility of meeting these goals, but at a price – complexity. And complexity brings with it substantial engineering challenges:

- **standards compliance/interoperability:** making sure the equipment interoperates with that of other vendors, so the video will play properly for the consumer;
- **visual quality:** getting the best quality out of the available bandwidth;
- **real time performance:** making it all work in real-time without having to use over-expensive hardware.

The correct test approach, using appropriate test and analysis tools, can make the development and implementation of these new compression standards much faster and much easier, minimizing costs and maximizing revenues.

Introduction

Video as a medium is changing.

- high definition
- increasing satellite and cable operations
- the beginnings of mobile wireless video
- the myriad of consumer devices playing stored content (on DVDs and similar)
- video-on-demand
- compressed digital video for applications such as security

All of these factors are significantly increasing the demands on the video compression algorithms needed to provide a valuable and fulfilling consumer experience – that is, one that the consumer is willing to pay for.

The purposes of this technical brief are to:

- examine the trends in some of these markets
- provide an overview of the new video compression standards designed to address the needs of these markets
- consider the test challenges of advanced video compression

In doing so, this brief focuses mainly on the internationally agreed standards, rather than various proprietary – albeit popular and widespread – video compression algorithms.

Note: this technical brief focuses on video compression only, so when referring to 'MPEG-4' or 'MPEG-2', it is meant to refer to the ***video*** parts of these standards only.

Why are new video compression standards needed?

Or...why not just stick with good old MPEG-2?

Digital communications media – cable, satellite, Internet, DVDs etc. – have been providing drastically increasing capacity over recent years, but even with these increases, the capacity provided by these media is two or more orders of magnitude too small to transmit/store uncompressed high-quality video.

There is rapidly increasing demand for more channels, mobile video, higher quality video, high definition video; however, the rate of growth of these – and consumer expectation – is outstripping the rate of growth of capacity of the digital communications media.

What bandwidth does the video take?

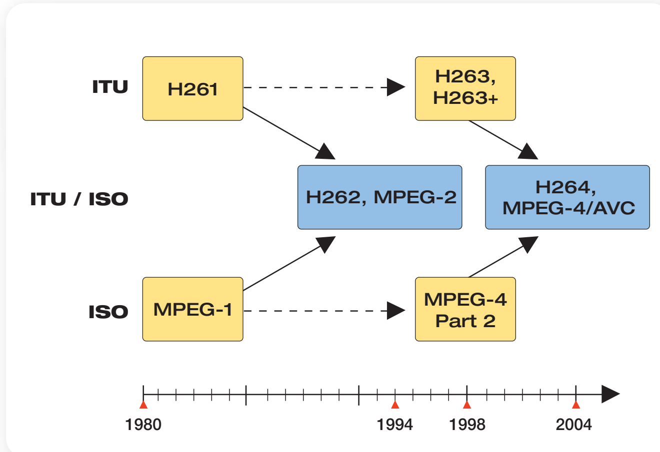
A single 'good' quality standard definition (SD) video channel (720x480 NTSC or 720x576 PAL) can be transmitted in 10-12 Mbps of cable bandwidth. However, how much better would it be to be able to transmit three or more channels of SD video in the same bandwidth? This provides more consumer choice and more revenue opportunities for the service provider.

What happens with high definition (HD) TV? Transmitting a single channel of HD at resolution of '720p' (i.e. 1280x720, progressive) takes ~2.6 times the bandwidth of a single channel of SD, and transmitting '1080i' (i.e. 1920x1088, interlaced – which is becoming the preferred choice in many broadcast areas) takes ~six times the bandwidth. But the consumer doesn't want fewer channels (and also doesn't want to pay a lot more for HD), and the content provider probably wants to offer HD *in addition to* SD programming. So without additional investment in transmission and reception technology, better compression of the video is the only solution.

Mobile video – 'coming real soon'

"Mobile video is coming along any day now"… for how many years have we been hearing this? Well, at long last, a number of companies are beginning to offer services such as football highlights, personal video conferencing, or downloads of movie trailers, and the screen size of mobile video players is increasing (not only for mobile phones, but also PDAs and other sorts of mobile video players). The number of these services will continue to increase as will the screen sizes: the only questions are how much and how rapidly. But the spectrum availability is limited, even with 3G (and the as-yet distant 4G) and WiFi hotspots. This is another pressure for better video compression.

With the set-top box/PVR ('Personal Video Recorder') hard disks becoming more common, there is an increasing requirement to store more and more video. And when it comes to movies, it is all very well getting a 2 hour SD video on a DVD, but soon SD *and* HD (1080i ?) versions will ideally be on the same DVD, which pushes even the new 'Blu-ray' DVD standard (which can store over 20GB of data).



▶ **Figure 1.** *Evolution of the video standards*

What are the 'new' video compression standards?

Video compression is not new; it started around 1980 with the H.261 video standard, which was designed for video conferencing applications at a bit rate of 64 kbps, illustrated in Figure 1. This was for small pictures, with limited frame rate, but represented what could be done at the time with the available bandwidths and processing powers.

Standards-bodies; standards evolution

Over the years, there have been two main bodies doing parallel development of video compression algorithms/standards:

- the "H" standards, developed by ITU (International Telecommunication Union)
- the "MPEG" standards, developed by the MPEG committee, then ratified by ISO (International Standards Organization).

Comparing performance: ‘new’ versus ‘old’

As expected, new video compression standards outperform the older ones. A comparison of the same sequence, shown in Figure 2, compressed with MPEG-2 and compressed with H.264/AVC, shows the improved compression ratio.

Proprietary codecs

There are a number of notable proprietary codecs from various companies, such as Microsoft (Windows Media), Apple (QuickTime), Real Networks, DivX, On2 and others. Of these codecs, VC-1 (previously called VC-9) and derived from Windows Media Video 9, has now been adopted by a number of industry standard bodies:

- SMPTE (Society of Motion Picture and Television Engineers)
- DVD forum as part of the new HD-DVD standard
- Blu-ray Disc Association for the BD-ROM standard

Therefore, based on the adoption activities above, it would be reasonable to argue that VC-1 is now becoming a ‘standard’, rather ‘proprietary’.

In addition, the performance of the encoding within a particular video compression standard also improves, as the quality of the encoders improves.

What has held back MPEG-4 from wider adoption?

MPEG-4 Part 2 can provide 'good' video compression, and give a clear advance over the earlier standards, so why has it not been adopted more fully, now that it is being overtaken – in terms of commercial exploitation – by H.264/AVC?

What is “performance”?

Performance is simply the measure of how many bits it takes to achieve a particular visual quality in a video sequence. However, it is much more complex than it appears, as the performance and “visual quality” is a subjective combination of many things, e.g.:

- sharpness of the video
- “blockiness”, or how easy it is to see the edges of the blocks
- color fidelity
- video/audio synchronization
- smoothness of motion

These visual features derive from the different types of encoding used throughout the sequence, the different frame types and the allocation of bit usage variation during the sequence. While the subjective elements are hard to measure, it is possible to quantify the encoding, bit usage etc. which generates the visual quality, and have some correlation between the visual quality and the quantifiable parameters, for example, using measurements such as PSNR (Peak Signal-to-Noise Ratio).

Definitions:
(as used in this technical brief):

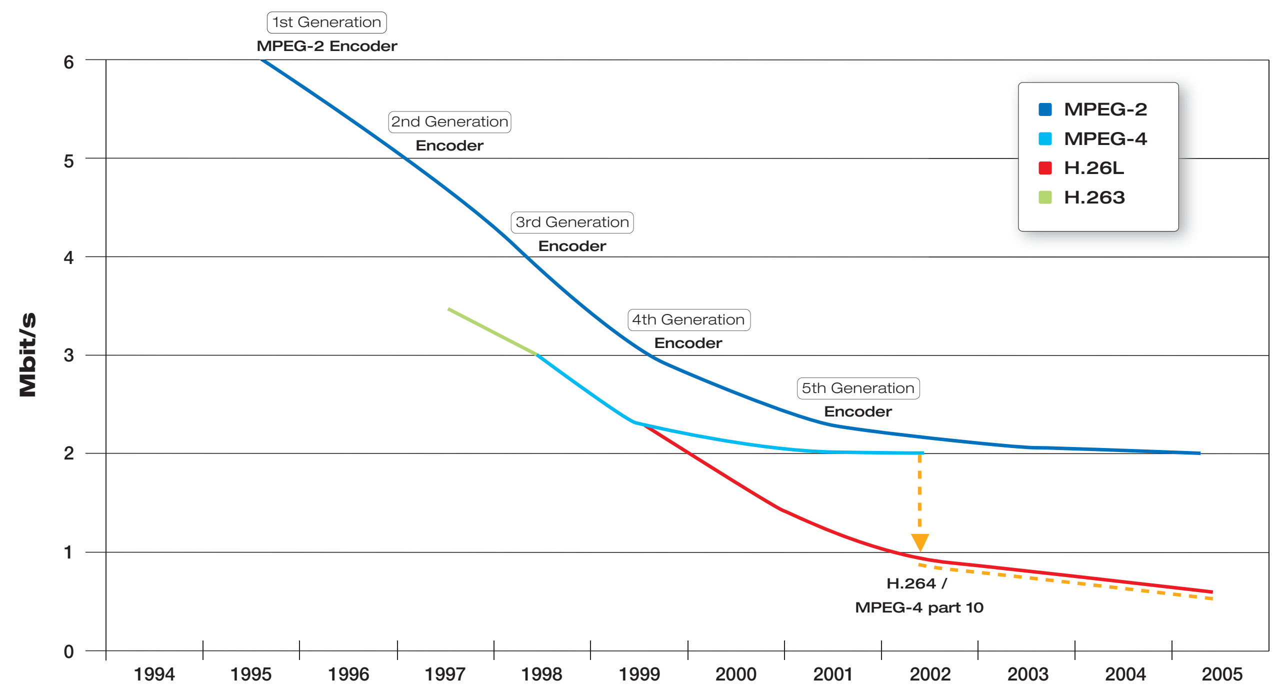
H.263 = ITU standard (‘Video Coding for Low Bit Rate Communication’)

MPEG-2 = ISO 13818-2

MPEG-4 = ISO 14496-2 (i.e. MPEG-4 Part 2 Visual)

H.264/AVC = ISO 14496-10 (i.e. MPEG-4 Part 10)

VC-1 = Microsoft VC-1 video codec (formerly called WMV9)



▶ **Figure 2.** *MPEG-4 Evolution*

There are several reasons:

- The MPEG-4 standard is actually a very large, complex standard, comprising many elements: video, synthetic structures (graphics-like), audio, systems, reference software, test bitstreams, digital rights management, and so on. This has made it a major undertaking to get the various parts compatible, and has slowed the standardization process.
- The MPEG-4 video part alone (i.e. ISO 14496-2 MPEG-4 Part 2 Visual) is a large standard and still yet has a number of issues for commercial implementation
- Over-complexity (many profiles, the vast majority of which are unused by commercial applications)
- It includes a number of technical compromises, such that the syntax is not as well structured and clear as it could be, making implementation and interoperability more difficult
- Some elements are not entirely clear and are open to interpretation
- There are some errors, such as the standard and the normative reference software and normative bitstreams that are sometimes at odds with each other

In the time it has taken to develop and standardize MPEG-4, technology has moved on – H.264/AVC undoubtedly gives better compression. One of the major issues of commercial exploitation has been the question of licensing: it took too long to complete the license arrangements for MPEG-4 Part-2.

So, although MPEG-4 Part 2 has many devotees, much more new work is now going into H.264/AVC in particular and also VC-1.

And what is pushing forward H.264/AVC and VC-1?

H.264/AVC is attractive for many reasons:

- It gives the best compression that can currently be obtained (it is 'state-of-the-art')
- It is an internationally agreed-upon standard, backed by both MPEG/ISO and the ITU
- H.264/AVC concentrates solely on the video, and as a result has been easier and quicker to standardize (it does not have to interoperate with many other parts of the same standard)
- The H.264/AVC standards body ensured that systems elements were taken care of elsewhere – for example, with a minor modification to the MPEG-2 Transport Stream it can carry incorporate H.264/AVC video as well as MPEG-2 video
- As a result of the above, H.264/AVC has become highly attractive to the whole broadcast industry, which uses MPEG-2 Transport Stream ('TS'). The approach of embedding H.264/AVC video in an MPEG-2 TS means that it is backwards compatible with the majority of the existing highly expensive infrastructure – satellite transmission, TV/cable transmission, etc. Only the video elements at the source and destination ends need to be changed to take advantage of the improved video quality/more channels/high definition possibilities that H.264/AVC offers
- Although complex in itself (see 'Complexity of H.264/AVC'), the standard has been well designed, well written, and relatively good reference software is also available. The result is that for engineers implementing H.264/AVC, although it is a very large task, it is generally clearer to implement than with MPEG-4
- many companies recognize the maturity of the standard, so there is now real commercial impetus behind the standard

There are of course down sides to H.264/AVC - see Sidebar 'Complexity of H.264/AVC'.

VC-1 provides performance similar to H.264/AVC, and potentially offers reduced processing and memory requirements. VC-1 has been adopted by a number of standards bodies, so it will find widespread use.

Crucially also for both H.264/AVC and VC-1, the issues which made an MPEG-4 license commercially unattractive were avoided – essentially, there are no 'usage' fees – so that it is feasible to use both in the vast majority of commercial applications.

Is the Codec correct, will it interoperate?

When producing a product using these video standards, standards compliance and interoperability are the first concerns to look at. If the product does not comply with the standard, then there will likely be interoperability issues – the video won't play or will look poor, or the decoder equipment will crash.

Testing your own. Testing your own decoder with your own encoder generally gives very limited test coverage – many times, an error in the encoder is mirrored in the decoder, so the video plays perfectly well in this situation. If the video does not play correctly (or at all) with a 3rd party codec, the questions are: first, who has the problem (you, the 3rd party, or perhaps both) and second, where exactly is the problem (which syntax elements, which values, which features used, which parameters – or more typically, what *combination* of elements/values/features – cause the problem).

Testing with someone else's. A very simplistic approach of extending the testing of an encoder or decoder is by simply playing lots of videos (ideally from different sources) and looking for errors, or testing for interoperability on a few other vendor's encoders and decoders. This is not a rigorous enough test to find the many errors that can occur in these complex video standards. However, using a tool that has already been used by hundreds of other developers for testing against the same standard would be a valid approach.

Complexity of H.264/AVC

In order to achieve high compression performance H.264/AVC has considerable complexity.

Two relevant examples are:

Intra-prediction

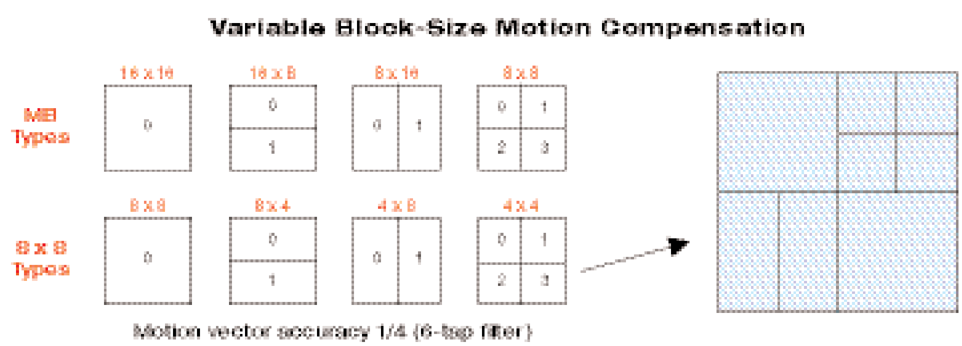
In the other video standards, in an ‘Intra’ frame, all the 16x16 Macroblocks are self-contained: they are not generated based upon Macroblocks in preceding or following frames, nor from other Macro-blocks in the same frame. However, H.264/AVC takes advantage of the fact that within a frame, Macroblocks which are located close to each other often have similar data – so that even within an ‘Intra’ frame it is possible to use prediction techniques where the value of one Macroblock is derived from the values of the video data in one of the surrounding Macroblocks.

Smaller blocks for enhanced efficiency

H.264/AVC allows the 16x16 Macroblocks to be sub-divided down to blocks as small as 4x4, to enhance the compression efficiency. In MPEG-2 and MPEG-4, there is just one type of 'Intra' Macroblock, containing compressed video which does not refer to any other Macroblock.

However, with the complexities above (and others), in H.264/AVC there are 26 types of Intra Macroblock. There are many other complex elements in H.264/AVC, such as the ‘CABAC’ Entropy coding, where bits are shared such that fractions of 1 bit effectively represent a specific syntax element.

On the converse side, substantial attempts have been made to understand and reduce the computing power required, where possible. For example, instead of a DCT a simple 4x4 transform is used, which uses only fast-to-execute shift, subtract and add operations.



Knowing all the details of a standard. No one can carry around in their memory all the information needed to guarantee compliance with the standards, and achieving compliance by going through the standards by hand and comparing it with the encoder or decoder that has been developed is slow and prone to error. The optimal approach is to automate the testing task as far as possible.

Understanding the particular standard. Although expertise in one video standard (such as MPEG-2) gives some insights into the basics of another video standard (such as H.264/AVC), because of the major differences between the different video standards, the engineer doing the work needs to have good understanding of the video standard concerned, and ideally standards-specific test tools that can aid the work dramatically.

Test coverage – is it finished? One of the major problems with the complexity of the new video standards is that there are a large number of syntax elements that can optionally be used. The problem then becomes, have you tested all the possible options, and all the boundary conditions? It may be that all interoperability testing has gone without a hitch – but this may only be testing 25% of the possible different modes that could be used in a standard. Comprehensive testing needs to ensure that it covers all the different parts of the syntax elements that could possibly apply. (Plus, you ideally want to see a real print out that demonstrates that this is the case…)

When it doesn't work. When a problem of interoperability occurs the important point is to be able to pinpoint the cause quickly – and determine whose problem it is.

When this happens, it is vital to be able to quickly be shown where the error is in the video, in the syntax, in a 'hex' view of the data, perhaps other views of the video data, and back this up with buffer and quality analysis. Ideally, you would be shown where the same error is in all the different views of the data, and be able to toggle between the views quickly.

Comparing codecs. Often, there is a requirement to compare different codecs. Perhaps this is to select a codec vendor, or to see why one codec is better than another (perhaps a competitor's codec), or to compare different versions of an in-house developed codec to see how much better the new one is. In these situations, it is vital to have detailed knowledge of the performance of the codec – and what is going on 'inside' it; how are the bits being used to obtain the quality shown (and to be able to quickly answer questions such as "why is one codec better on these types of sequences or better on Intra frames than the other?").

Comparing standards. Often, there is a requirement to compare different standards, which trigger questions like…

How much better is H.264/AVC than MPEG-2?

Why is MPEG-4 Advanced Simple Profile as good as H.264/AVC on specific video sequences?

What features and options in MPEG-4 and H.264/AVC are being used/not used, that cause this?

In order to answer these sorts of questions, it is essential to have test tools that provide detailed, consistent information on **all** the video standards: H.261; H.263 (particularly important for 3GPP mobile as H.263 is mandatory); MPEG-2; MPEG-4; H.264/AVC; and VC-1.

The test requirements for Encoders and Decoders, while related, differ in many ways:

Encoder testing and optimization

An encoder is much more complex than a decoder, as it has to not only produce a correct syntax bit stream, but also intelligently work out what is going on in the input video, in order to make the optimal choices for compression, and do this as fast as possible (perhaps in real time).

Making optimal choices about which to use out of the various encoding schemes permissible for a specific frame/slice/Macroblock/block can take a long time. This process includes searching through other parts of the frame or other frames for correlated or similar data, working out whether it uses fewer bits to repeat a Macroblock or predict the Macroblock concerned from one of the surrounding Macroblocks in the same frame or from a frame before and after. Typically, there are many iterations for every part of each frame in order to arrive at a decision as to which compression feature or combination of features gives the highest compression ratio. As might be expected, this can take a long time; for the 'new' video compression schemes, this is the trade-off - compression efficiency versus speed of encoding.

So, an encoder that encodes in real time, may be relatively simple and fast, but may give a lower compression ratio for the same video quality than one that does not work in real time. One difficulty with encoder optimization is that small improvements in efficiency are often not visible to the eye, so the video quality must be measured in order to improve the efficiency in an ordered manner.

Continuous improvement. Over time, encoders improve. This has happened with the encoders of all the recent video standards: the later generation encoders give higher compression ratios for the same video quality as the encoder becomes more intelligent at making decisions; the additional processing power required has become available, and the encoder engineering has improved.

The important issue then, is to quickly get as far along this performance improvement curve as possible – to be better than the competition – by having the detailed knowledge to be able to make trade-offs such as performance versus video quality. By being able to answer questions such as "if the motion vector search range is restricted by X amount, by how much is the quality reduced for this type of video?"

Decoder testing

A decoder is simpler than an encoder; it has to take the bit stream it is given, parse and interpret it correctly and display it correctly, within the given time for each frame.

The ideal scenario here is to be able to:

- compare bit-for-bit with the display (YUV) output of a known-good decoder
- if there are any differences, be able to get detailed information on how the known-good decoder works out the pixels to display, i.e. to be able to have full, easy to understand information on all the internal workings of the decoder.

PSNR measurement of video quality

The PSNR metric is defined as the ratio between signal power and noise power, on a decibel scale. In the context of image processing, signal power is taken to be the square of the peak image sample value and noise power is taken as the square of RMS error in the image.

$$PSNR = 10 \cdot \log (\hat{S}2 / RMS2)$$

In a number of ways, PSNR is not an ideal measurement of visual quality, as many people argue that it does not correlate well with visual perception (although of course this is extremely hard to measure).

Also, when comparing between different video streams, the PSNR measurement gives a poor representation of the relative quality of the different streams. In any event, to measure PSNR, the original, uncompressed source video, is required.

However, when comparing the same single video sequence, compressed in a number of ways (for example with different video compression standards or with different versions of the same standard) then it does provide useful data on which is better.

Other issues

Competition. The video market is very competitive: codec products must work correctly, and the video quality must compare favorably with every other product out there – if not, the product will not be a success. Many customers are highly demanding, with little patience for errors (such as interoperability problems or non-conformance) or video quality that is not state-of-the-art.

For some applications, getting it wrong – or not fully optimal – the first time can be very expensive. For example, when making a semiconductor, it is very expensive to re-make masks. Plus, of course, if the semiconductor has already been put into another device, such as a camera, or mobile phone, or set-top box – it is too late (except perhaps for software workarounds). Also, in the field of semiconductors, there is a trade-off of extra silicon area for more advanced processing, versus the cost of designing and supplying a larger chip.

Time to market. Development schedules are continually under pressure: normally, a project schedule is set at the start, when the full complexities of the development are not understood. Yet the engineers still have to stick to this timetable…

Often an organization is geared up to release a product at a certain time; perhaps to coincide with a trade show. If the product is ready at the time desired, it can generate a lot more revenue for the company. (Or conversely, if not ready, cause a lot of revenue to be lost.)

So, having tools which can save time can give a return far greater than the cost of the tools – not to mention the direct cost savings of freeing up engineering resources sooner.

Two application examples

Video quality optimization in a semi-conductor – test software helps generate revenue

A small European semiconductor company was recently trying to win some business from a major US technology company to supply them with an MPEG-4 processing chip. The European company was competing with other vendors in order to win the supply contract: the problem the European company had, was that their video was of similar quality as the competitors, and so was the price. Why should the US technology company buy from the European company?

After spending some months on this, the European company bought the Compressed Video Elementary Stream Analyzer to see if they could use it to optimize the video quality of their MPEG-4 processing chip.

The software provided internal analysis of the coding – how many bits were used, where, to achieve a certain quality level – and helped show where the optimum trade-off could be obtained between silicon area (processing requirements), complexity, speed and video quality.

Over a relatively short period of time, the European company was able to optimize their chip and improve the video quality substantially – so much so, that it was recognized as being clearly superior to the competing vendors, by the US technology company.

The result: the European company won the substantial order.

In the words of the President of the European company: “[the large US technology company] was very impressed with the video quality improvement, and it was this improvement that was crucial to winning the business - we would not have won it without”.

Mobile phone – getting it to decode the video

This involved a large company in Europe that makes mobile handsets. This particular handset manufacturer had a problem: their new 3G handset could not properly decode the video provided by a European network operator. The network operator was planning to buy the handsets for sale to their customers, so of course the handset manufacturer was keen to resolve the problem. (It involved potentially millions of dollars worth of handset sales.)

As might be expected, there was disagreement between the companies as to where the problem lay: but none of the engineers from either company really knew where the problem was – if it was the handset, or the way the video was encoded, or how the video was sent over the network.

After spending two months themselves trying to work this out - with the network operator waiting, with some frustration - the handset manufacturer bought the Compressed Video Elementary Stream Analyzer. Within a day of getting the test software, the semiconductor division of the handset manufacturer was able to show that the fault was not with the handset at all, nor the network handling of the video - the problem was with the video encoder - it did not comply with the MPEG-4 standard fully - and this was why the handset could not display the video properly.

The handset manufacturer was able to show not only that there was a problem with the video, but explicitly what the problem was and which specific parts of the syntax were in error.

Interestingly though, the network operator's response was: “we now understand that this video does not conform to the standard, but this is the video we provide - your handset must deal with it as it is”.

While not entirely content with the response from the network operator, the handset manufacturer was able to use the detailed information provided by the software to work out how to change the handset so that it would cope well with the non-standard video provided.

Discover your Next Video Test System

Now you only need to talk to one company to test emerging compressed video standards. The new MTS4EA Video Elementary Stream Analyzer is a revolutionary new software product that offers a flexible, upgradeable test solution for next- and current-generation video compression technologies. It provides complete elementary stream analysis support for MPEG-2, MPEG-4, H.264/AVC, H.261, H.263, H.263+ and 3GPP standards. (Please consult Tektronix for an up-to-date list of supported video compression standards.)

Go to www.tektronix.com/MTS4EA to find out more.
Don't forget to download your free 30-day trial software!

Contact Tektronix:

ASEAN / Australasia / Pakistan (65) 6356 3900
Austria +41 52 675 3777
Balkan, Israel, South Africa and other ISE Countries +41 52 675 3777
Belgium 07 81 60166
Brazil & South America 55 (11) 3741-8360
Canada 1 (800) 661-5625
Central East Europe, Ukraine and Baltics +41 52 675 3777
Central Europe & Greece +41 52 675 3777
Denmark 80 88 1401
Finland +41 52 675 3777
France & North Africa +33 (0) 1 69 81 81
Germany +49 (221) 94 77 400
Hong Kong (852) 2585-6688
India (91) 80-22275577
Italy +39 (02) 25086 1
Japan 81 (3) 6714-3010
Luxembourg +44 (0) 1344 392400
Mexico, Central America & Caribbean 52 (55) 56666-333
Middle East, Asia and North Africa +41 52 675 3777
The Netherlands 090 02 021797
Norway 800 16098
People's Republic of China 86 (10) 6235 1230
Poland +41 52 675 3777
Portugal 80 08 12370
Republic of Korea 82 (2) 528-5299
Russia, CIS & The Baltics 7 095 775 1064
South Africa +27 11 254 8360
Spain (+34) 901 988 054
Sweden 020 08 80371
Switzerland +41 52 675 3777
Taiwan 886 (2) 2722-9622
United Kingdom & Eire +44 (0) 1344 392400
USA 1 (800) 426-2200
USA (Export Sales) 1 (503) 627-1916
For other areas contact Tektronix, Inc. at: 1 (503) 627-7111
Last Updated 3 November 2004

For Further Information
Tektronix maintains a comprehensive, constantly expanding collection of application notes, technical briefs and other resources to help engineers working on the cutting edge of technology. Please visit **www.tektronix.com**

Reproduced courtesy of Vqual Ltd.

01/05 AAD/WOW 2AW-18398-1

Tektronix
Enabling Innovation