

Understanding digital video compression

The managed delivery of video-based services is not without challenges: the requisite bit rates are high and the quality of service is critical. All the same, it is these requirements that play to network operators' favor, allowing them to drive the advantages of a managed network.

✦ GILES WILSON

Video-based entertainment and communication services are rapidly increasing in importance for operators of all network types. User demand for these services continues to increase, driving requirements for greater bandwidth. Likewise, a maturing user base is beginning to demand better-quality experiences from suppliers.

This article details the technical basis for digital video compression and how the quality and costs of video-based services are directly affected by the practical performance of video-compression technologies.

To compete effectively in the marketplace for customer subscriptions, more and more network operators are compelled to provide a full-service offering. This includes high-speed data access, telephony, and video-based services. Furthermore, given the rise of "over-the-top" service delivery from internet-based service providers, network operators must add clear value over and above the basic furnishing of bandwidth. They can do this by exploiting the advantages of the managed-network infrastructure. Video-based services, including IPTV for fixed-line operators and mobile TV for

wireless operators, play a key role in this context.

Direct-to-home delivery of digitally compressed video programming has been available for 15 years, but constraints on bandwidth for delivering content have largely limited deployment to dedicated satellite- and cable-delivery systems. However, in the past few years, several factors have changed to enable telecom operators to compete in this market. Improvements in access network technologies have increased the size of the last mile "pipe" to end-users. At the same time, the arrival of more modern and powerful video-compression technologies has significantly reduced the bandwidths required for video services. For example, the H.264/AVC video standard, which is the codec of choice for new deployments in IPTV and mobile TV, is capable of compressing video services to half the bandwidth required using the legacy MPEG-2 codec for traditional satellite and cable deployments.

Structure of a video encoder

Figure 1 shows, in simplified form, the key algorithmic components on which all modern block-transform-based digital video compression algorithms are designed, including H.264/AVC. Each video frame is divided into an array of rectangular blocks of pixels, called *macro blocks*, which are processed through the system in order to compress the images.

BOX B

Over the top

Digital industry term which describes third-party home entertainment services that are delivered on top of a broadband network without affiliation with the broadband service provider.

Redundancy identification and removal

The redundancy identification and removal (RIR) block represents processes associated with identifying and removing redundant information in video. The RIR block outputs a *predictor* and a *residual* pixel array. A decoder uses the predictor to create a best estimate of the data required by the macro block. The residual represents an error signal that must be added to the predicted macro block (**Figure 2**).

Prediction types fall into two main categories: interframe and intraframe prediction. Interframe prediction targets the removal of redundant information that exists between the frames of a video sequence. Based on nearby (in time) frames of video, the technique predicts the content of a coming frame. This kind of prediction relies on the knowledge that, in any sequence of video frames, large sections of the picture (such as the background) change very little between frames. Also, many areas within consecutive frames of video have similar content that has been displaced spatially – moving objects, for instance. The predictions derived from this process, known as motion vectors, represent the relative two-dimensional position of the most similar block in a previous frame.

Intraframe prediction targets the removal of redundant information that exists within a single frame of video, such as areas of similar color, brightness, and texture. This process selects

- ✦ neighboring blocks that are most similar to the current block; and
- ✦ the correct algorithmic procedure for making the prediction.

Transform and quantization

The residual blocks from the RIR process are submitted to a transform-and-quantization (TQ) process in order to

BOX A Terms and abbreviations

AVC	Advanced video coding	MPEG	Moving Picture Experts Group
CAPEX	Capital expenditure	OPEX	Operating expense
H.264	Video compression standard equivalent to MPEG-4 Part 10 or MPEG-4 AVC	RIR	Redundancy identification and removal
IPTV	IP television	TQ	Transform and quantization
		VoD	Video on demand

reduce their dynamic range and produce a more efficient representation. In short, a discrete block-based transform is applied to the block in order to compact energy into lesser coefficients. These coefficients are then quantized in order to reduce their dynamic range.

Entropy encoding

The entropy-encoding block is responsible for converting the predictors and quantized residuals into a more compact representation. Several techniques are employed, including Huffman encoding, variable-length code tables, and arithmetic coding.

Bit-rate control

It is usually necessary to exercise control over the video compression in order to achieve a desired outcome, such as constant bit-rate operation or constant quality output. The primary way of doing this is to adjust the degree of quantization applied within the TQ block of algorithm. Increasing the size of the quantization step reduces the number of bits produced by the TQ block; reducing the size of the quantization step increases the number of bits.

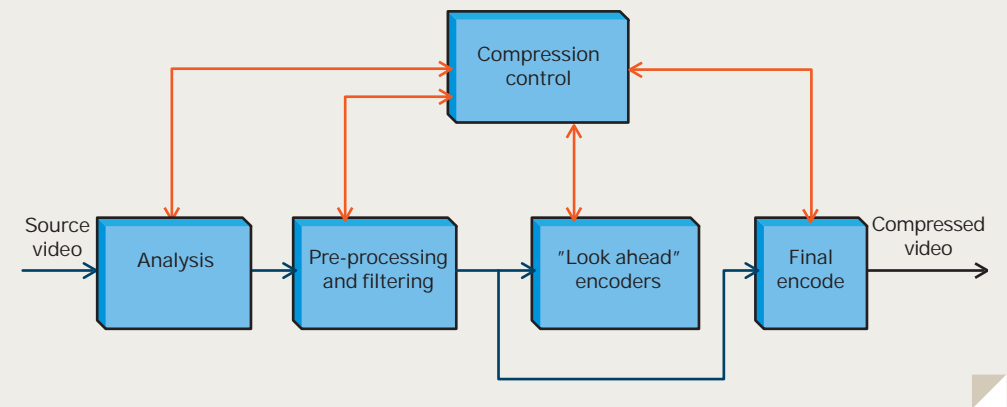
Characteristics of video encoding

Changing the size of the quantization step in the TQ block can effectively reduce bit rates, but doing so also affects video quality. This is because quantization and subsequent reconstruction is a lossy process. Therefore, increasing the coarseness of the quantization by increasing the size of the step reduces video quality. This is illustrated by the typical quality-per-bit-rate curve (Figure 3).

The shape and position of the curve (quality versus bit rate) is not solely dependent on the performance of the compression algorithm. Indeed, the characteristics of the curve are extremely dependent on video content. More specifically, they are dictated by the complexity of a given sequence or piece of video content.

Video complexity is a measure of the amount of natural redundancy in a piece of content. Video that has low motion content contains a lot of temporal redundancy and is thus deemed to be of low complexity. Conversely, video that has a lot of motion content as well as inconsistent motion has low temporal redundancy and is deemed to be of

FIGURE 1 Simplified view of the key components on which block-transform-based digital video compression algorithms are designed.



high complexity. Pictures that contain a lot of detailed textures have low spatial redundancy and are therefore deemed complex, and so on.

The level of video complexity affects the quality bit-rate curve in several ways (Figure 4). A lower level of complexity gives better performance during the RIR stage of the encoding process. This, in turn, means that fewer residuals are produced and less quantization is required to achieve a given bit rate. The quality-versus-bit-rate curve is thus higher and, since the data from residuals is a lesser component of the total, the knee on the curve occurs at a lower bit rate.

The reverse is true for high complexity video. A high degree of complexity reduces performance during the RIR stage of the encoding process, giving

more residuals. Coarser quantization is thus required to achieve a given bit rate, which results in poorer quality at that bit rate. This shifts the curve down. In addition, the breakdown point on the curve occurs at a higher bit rate. In summary,

- ✦ quality increases with bit rate;
- ✦ the relationship between quality and bit rate is dictated by video complexity;
- ✦ complex video has greater motion, greater spatial detail or texture, or both; and
- ✦ for a given algorithm, complex video compressed at a given bit rate results in poorer quality than a similarly encoded piece of video with lower complexity.

Relative performance in encoder implementations

Thus far, we have discussed the fundamental building blocks of the

FIGURE 2 The redundancy identification and removal (RIR) block.

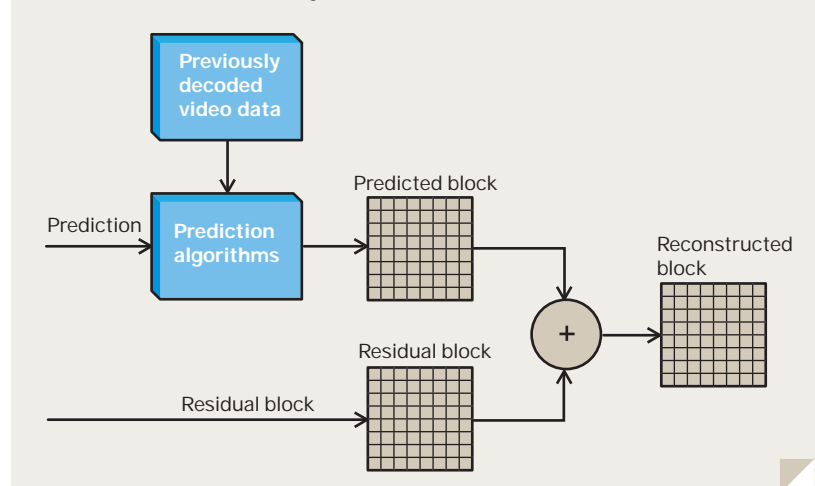
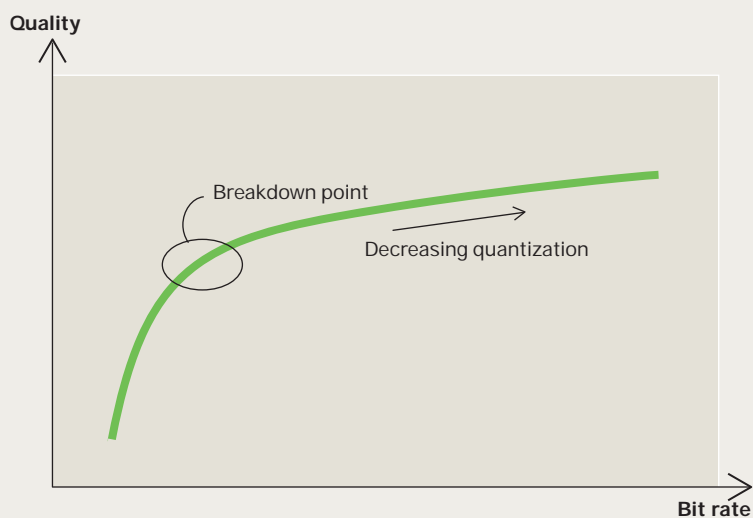


FIGURE 3 Typical quality-per-bit-rate curve.



❖ video-compression algorithm and the impact it has on compressed video quality and bit rate. We have also explained how the quality-versus-bit-rate characteristics of the compression algorithm are dependent on content complexity, which is largely a function of where the majority of compression takes place in the algorithm – in the lossy quantization block or the lossless RIR block.

However, this is not the end of the story concerning encoding performance. The specifications for video encoding have been authored to define a

- ❖ toolset of techniques that can be used for compression;
- ❖ syntax for describing the operations in a bit stream; and
- ❖ fixed and proscribed methodology for decoding this syntax.

Notwithstanding, the specifications do

not tell us how to do the encoding or, in particular, how to set about the task of identifying redundant information.

Given the large number of options available to the encoder, the complexity of calculating each of them, and the intractable interactions and relationships between decisions made for each piece of content and each macro block in a frame, it is not economically feasible to design a single encoder that can try all possibilities in an exhaustive fashion and then select the best one. Designers of encoders must thus compromise in order to achieve a good balance of cost and performance (Figure 5). How well this is achieved is a function of numerous factors including the

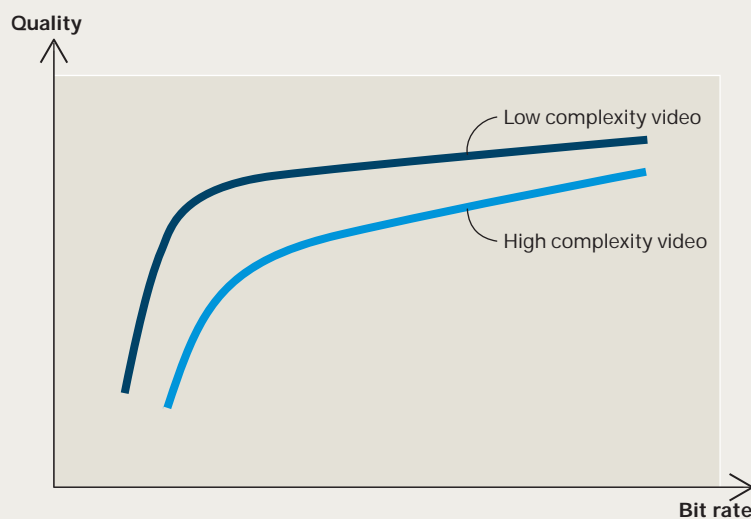
- ❖ amount of processing applied;
- ❖ efficiency with which the processing is applied;
- ❖ technology used in the implementation; and
- ❖ algorithmic knowledge.

Encoders with average performance put constraints on number of choices as well as limiting parameters, such as the range of motion search. They also use simplified methods for determining the best predictions and other shortcuts that reduce calculations. They might also employ relatively simple rate-control strategies – that is, the algorithms that control the variations in quantization in order to control bit rate and quality.

High-performance encoders, on the other hand, such as those used in the traditional broadcasting markets, are more efficiently designed on custom platforms in order to achieve greater computational throughput, allowing more choices to be considered and with less limitations on scope. More complex measures are employed to determine how the choices are used. Complex rate-control strategies are implemented to allow accurate control of bit rate and to obtain consistent quality.

As a result, one sees a dramatic difference in performance between different encoders. The high-performance hardware-based professional encoders, for example, outperform more compromised and naïve software encoders by as much as 30 to 40 percent – that is, they achieve the same quality while using up to 40 percent fewer bits per second of bandwidth.

FIGURE 4 Impact of video complexity on the quality bit-rate curve.



Benefits of high-performance encoding

Operators who are setting out to deliver video over fixed-line or mobile networks or even the internet stand to benefit from investing in high-performance video-encoding platforms. In particular, they will have lower operating expenses (OPEX) and better video quality, which translates into greater user satisfaction.

Where linear TV channels in an IPTV network are concerned, the improvements in compression efficiency can have numerous benefits. One can achieve competitive broadcast quality at lower bit rates, which in turn, means less core network distribution and greater reach on the access network.

For video-on-demand (VoD) services, the reduction in video bit rates has a positive impact on the expenses (CAPEX and OPEX) associated with server infrastructures. Moreover, operators can serve more users from each VoD pump and reduce the storage space required per asset.

Where mobile TV is concerned, many of these same savings apply. Mobile TV operators may choose to use the higher performance either to improve the customer experience through better video quality or to achieve specific quality levels at substantially lower bit rates. In the latter case, one can reduce server and storage capacity without reducing the number of streams. In addition, the reduction in bit rate per stream extends the number of simultaneous users who can be supported in each cell.

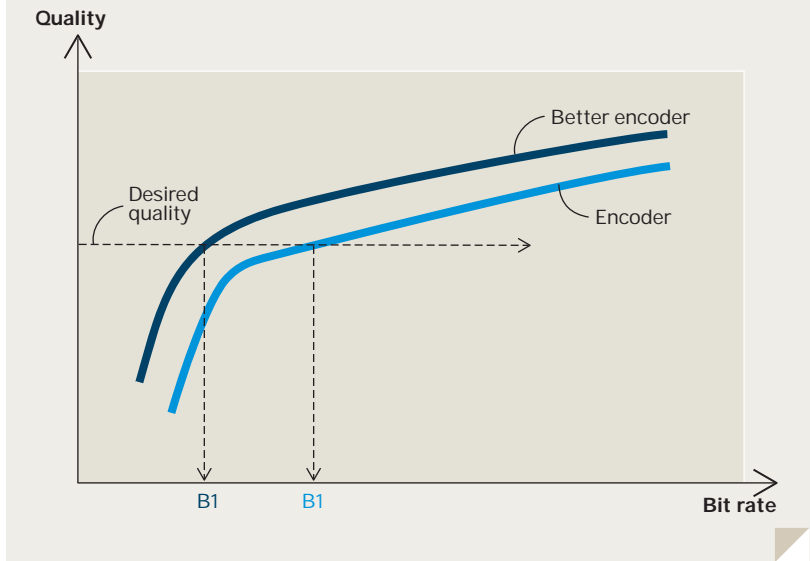
When delivering internet-based video services, the reduction in bit rates for a given level of quality again translates into lower server and storage costs. Likewise, one can lower the operating expenses associated with the volume of data served per month via the content-distribution network. If one employs video download instead of streaming, the lower bit rates and hence smaller asset sizes reduce users' download times.

TANDBERG Television video compression solutions

Video encoding solutions from TANDBERG Television, such as the EN8090 High Definition H.264 encoder, lead the market in terms of compression performance (Figure 6). The video algo-

BOX C
TANDBERG Television, part of the Ericsson Group. With a track record of industry firsts and highly innovative products and solutions for the most demanding of applications, TANDBERG Television is acknowledged as the world's leading provider of video-encoding solutions to the professional broadcasting industry.

FIGURE 5 Designing encoders to achieve a good balance of cost and performance.



rithm implementations in these encoders are specifically designed to optimize the perceived quality of the output video at all bit rates. This is achieved through a combination of sophisticated video-manipulation techniques, subtle algorithmic intelligence, and raw processing power.

Before it is encoded, the video is analyzed in detail to determine complexity information such as which regions contain the most detail and which sequences contain the most motion. Complex adaptive spatial and temporal filtering is then applied to remove anomalies, such as camera sensor noise, analog distortions and digital noise. Unlike conventional video-restoration filters, these

pre-processing filters condition the video for the encoding stage by removing non-visible artifacts that might reduce the efficiency of the subsequent encoding stages.

After the pre-processing stages, but prior to final encoding, the video stream is run through several temporary "look ahead" encoding passes. These do not form part of the final stream, but are instead used to generate metrics (for example, how the various approaches to encoding the video should perform and where most difficulties are encountered). These, in turn, are fed into the final encoding stage where they are used to optimize the algorithm in real time.

The final encoding stage makes

FIGURE 6 The EN8090 High Definition H.264 encoder developed by TANDBERG Television.



❖❖ use of all information garnered from look-ahead encoding passes and the video pre-processing and analysis stages to produce the final bitstream. This stage employs many high-power and computationally intensive techniques to find the most efficient tools for compressing each portion of video, and optimizing the bit rate and visual quality. Intelligent algorithmic components are also employed to guide the optimization process, in order to ensure that the best visual results are obtained. This includes techniques such as hiding artifacts in areas of the video where the human visual system is least likely to notice them and directing maximal quality to regions where a viewer will generally notice distortions.

Conclusion

The performance of video-compression encoders is characterized by the quality-versus-bit-rate curve on different types of video content. The process of compression is a complex and computationally intensive task with a large degree of latitude for optimization and efficiency.

High-performance professional broadcast encoders significantly outperform less optimized solutions. Superior video performance can have a material and positive impact on the business of delivering video-based services over all types of networks. This applies to linear TV, video on demand, as well as video download services.

Video-compression solutions from TANDBERG Television are built on proprietary IPR developed over many years in the professional broadcast space. They employ many sophisticated techniques both in the video-compression algorithm itself and in the associated video processing, resulting in optimal compression performance across all types of content. ❖

Giles Wilson



❖ is a key member of TANDBERG Television's CTO group and serves as senior VP of technology, overseeing the company's strategy and vision for integrated digital video technology for the creation, management and delivery of high-definition, IP, on-demand and interactive television. Previously he was chief architect at TANDBERG Television, leading the development of solutions based on advanced video-compression technologies. His work with AVC has resulted in TANDBERG Television's award-winning solutions for VC-1 and MPEG-4 AVC high-definition and standard-definition encoding technologies. These world-first solutions have cemented his reputation as one of the foremost engineering experts in the digital media industry. Giles graduated from the Queen's University of Belfast in 1994 with a Ph.D. in robotics.