Signal Quality in Audiovisual Digital Preservation

Richard Wright, Preservation Guide, London UK

Abstract

Audiovisual files hold a signal or an image, and so have a quality dimension that has no equivalent in text files. Digital Preservation projects provide guidance for optimizing the preservation of 'significant properties', but audiovisual content also have purely technical dimensions to preserve: signal fidelity and image quality. For digital libraries, automation of signal quality control is necessary, both at time of original input and for every migration or publication thereafter – because manual checking does not scale to large collections. The paper reviews quality control automation and presents a graphical approach to tracking quality over an item's entire life cycle.

The Quality Dimension

Audiovisual content is an important type of digital content. Most traditional physical carriers of sound and moving images, from wax cylinders through to DAT tape and minidiscs, and from film through to videotape and DVDs, have pressing preservation issues. Collection managers are are faced with either migrating virtually all carriers – or simply losing the content to the combined forced of decay, damage and (in particular) obsolescence. For this reason there have been major digitization and migration programmes around the work over the last decade. The PrestoPRIME project has estimated that about three million hours of audiovisual content has been digitized [1, pp.7–8]. As with any migration, the quality of the result is very important, and the main approach to quality checking has been the subjective evaluation by human operators of the transfer equipment, or by separate checking teams.

In addition to digitisation, collections acquire new content at a rate of about 6% per year [2], which is four times higher than the rate of digitisation, meaning 12 million hours of born-digital sound and moving image content in European collections acquired in the last decade. This material should also be quality checked at ingest into a formal repository – and should be checked again on every migration to a new encoding or wrapper format, and on every production of an access copy.

Fifteen million hours of digital audiovisual content in curated collections in Europe simply swamps the available budgets for staff to perform manual quality checks. Generally in digital libraries file formats are automatically checked, but the checking is for form, not content: whether the file conforms to the standards for that file type, can be successfully opened and can yield up its embedded metadata. Files are not checked for content: whether the text makes sense. But for audiovisual files, whether the sound and images are good (which is about to be defined) is vital.

Audiovisual files hold a signal or an image, and so have a quality dimension that has no equivalent in text files. Digital Preservation projects provide guidance for optimizing the preservation of whole ranges of significant properties, but audiovisual content also has purely technical dimensions. For audio, the significant property is signal fidelity as measured by bandwidth and dynamic range. For video and film, image quality is ultimately subjective, but there has been work, over decades, on objective measurements that estimate perceived visual quality.

Quality Analysis

Automation of audiovisual quality control has been developed as part of digitization of analogue carriers. The German Institut für Rundfunktechnik (IRT) and the commercial company Cube-Tec developed the audio workstation Quadriga in 1998 which included a range of signal measurements that could be used to support automation of quality checking. At about the same time (1997), the Italian national broadcaster RAI decided to convert their radio production to a fully-digital system, including an all-digital archive of some 300k hours. This decision necessitated an intense migration of audio, running 24 hours per day with each operator controlling five simultaneous tape-to-digital transfers. This project set the pace across Europe for cost-effective mass digitization of audio, and at the same time used built-in signal measurements for quality control. The ACS Elettra workstation (no longer made) was developed for the RAI work. A third project shortly thereafter (1999) started in Vienna, to produce a digital audio Mediathek. Again, a workstation with built-in support for quality control was developed, the NOA system which, as with Cube-Tec, has been widely used in the ensuing decade. Another company has produced a stand-along tool for audio quality analysis (rather than a complete workstation): Audio Inspector from a company based in Salzburg, Austria. For video, there is signal-monitoring built into the SAMMA system for automated video digitization.

For moving images there is a different history. Film is easily damaged: every projection of a print can add scratches, or worse. Virtually all copies of any work on film have damage of some sort, and so a technology of *restoration* has developed over the decades. With digital technology it because possible to attempt automation of restoration, beginning with automatically detecting such defects as scratches, dust, shake and flicker. From these beginnings grew tools for automatically assessing whether digital files of moving images had any form of impairment of the image itself. This analysis of the image for impairments is not the checking for syntactic correctness (validation) performed by conventional digital library tools such as JHOVE. Digital moving image quality analysis tools, as with the audio tools previously mentioned, analyze the actual images for defects – while a validation tool like JHOVE can only verify that the overall file meets its specification.

The company Joanneum Research in Graz, Austria developed a leading software tool for film restoration, DIAMANT. This work, as with the Quadriga and NOA systems, goes back to before the year 2000 (to 1999). In recent years Joanneum have adapted their technology and produced software specifically addressing quality control: generating a marker or warning-flag whenever there is a suspected impairment (a disruption of any sort) in the moving image signal [3].

Such tools have an important role in digital preservation, providing their performance is improved so that false alarm rates are not a problem. If defects are missed then clearly the automation is not an adequate replacement for human checking, however much money it saves. But a more insidious problem is with too many false alarms: these have to be checked manually (at least, at present) and if there are too many then manual work checking the false alarms is just as arduous and costly as the original all-manual form of checking. One way forward for the automation is to use a second stage of computation: using really computational intensive evaluation just at the points flagged as possible disruptions in the first pass. The point of the second pass would be specifically to reduce false alarms. One might ask: why not just use the better software in the first place? The answer is: computational efficiency. High though-put software is used for the first pass, and then software which would been impossibly slow as a general checking tool is run just on the areas highlighted as problematic in the first pass.

All the methods so far mentioned produce extensive logs of issues and potential faults that have to be manually reviewed. There is a need for effective integration of signal processing technology with human checking in order to produce a really efficient method of quality control within a preservation factory approach. In the US, the National Archives (NARA) are currently reviewing digitization and quality control methods for audiovisual materials, and will produce public results in 2012.

There is a growing category of software that exists between standard digital library verification tools, and true audio and moving image quality analysis tools. These are the verification tools that have been developed specifically for audiovisual files (in fact, specifically for video). Cerify from Tektronix and Baton from Interra Systems are just two. These started as formal checking tools (verification tools) but have been extended to detect some aspects of video impairment: black screens, frozen screens, the disruption typical of failed MPEG decoding which results in visible blocks on the screen (blocking).

Digital library verification tools (eg JHOVE) tend to be open source and free. The verification tools designed for video files represent the opposite end: expensive software, or even software designed to run on dedicated processors – so the customer has to 'buy the box' to run the checking.

Quality Management

Why should files need quality control? Surely a checksum can be taken when a file is created or ingested, and providing the checksum is used properly, the file is guaranteed to be as it originally was. The problem with audiovisual content is that we keep changing out minds (or 'the industry' changes its mind, if an industry can be said to have a mind) about what kind of files we want. A decade ago the BBC only used Real Audio for online audio. Then came the Windows Media format and then AAC (a part of MPEG-4). Elsewhere there is extensive use of MP3 audio. Video has an even wider range of formats, with several in use at any one time and others fading into obsolescence. These different file types actually represent two sorts of difference: the encoding (how the bits represent the signal) and the wrapper (the file type and how it is structured. There is confusion because for some files (eg MPEG) the file type tends to determine the encoding, while for others (eg MOV, AVI) the file type can hold an incredibly wide range of encodings.

The problem with the different encodings is that most involve compression, meaning throwing away some of the frequency range or dynamic range of the original signal, irreversibly. There is an inevitable effect on quality, whether or not it is immediately apparent to the human listener or viewer. The problems are compounded from successive cycles of decode-recode, such as taking Real Audio data and recoding it as AAC.

A repository of audiovisual files should be protected from further quality problems if there were no further changes of encoding, because the quality would never need re-checking if the checksums themselves showed that files had not altered. Unfortunately, recoding is just what we commonly do to these signals, as needs change and requirements for access formats change..

An excellent quality control strategy in the digital preservation of audiovisual content would be the complete avoidance of repeated application (cascading) of compression. In the analogue world when archives were forced to make a new master, there would be an inevitable generation loss. In the digital world it should be possible to make perfect copies, but compression interferes. If a master file is compressed in one lossy way and then migrated to a different type of lossy compressed file, there is a decode–recode cycle that also produces additional loss, the digital equivalent of a generation loss. However, for cycles of lossy compression there is an invidious problem. There may be no perceptible effect until finally there is major breakdown, in contrast with the gradual losses from migrations of analogue content.

The ideal way to manage the need to produce new access copies in new encodings, is to always go back to the 'preservation master' as the source for the new encoding (eliminating cascade) – and for that master to be uncompressed or losslessly compressed (so it is an exact representation of the original signal). The problems arise when there is no uncompressed version to begin with – because the content entering the repository was already in a lossy-compressed encoding.

For such signals, life for the future could just get messier and messier. Each time a new encoding was needed, it could be managed as a new cascade, and the signal would just get worse and worse. The way to avoid this problem is to pick a 'preservation master' format and stick to it, always going back to that as the source of future encodings, even though the 'preservation master' is itself a compressed version in some lossy encoding, and so not an ideal preservation master. The problem with this strategy is that at some point the encoding of the preservation master could become obsolete – meaning the file would have to be migrated to a new encoding before becoming unusable.

The solution to the obsolete preservation master problem is to decode the original file back to an uncompressed file at ingest (into the repository) – even though that would mean using more storage space than is needed. For overall best quality at the lowest price, the best strategy is to delay this decoding-to-uncompressed as long as possible (at the risk of delaying a bit too long!) so that when it is decoded and does take up more storage space, the price of storage will have reduced as much as possible.

A Quality Management Tool

Coding is used to reduce the data rate of a signal, so that it can be transmitted using a lower bandwidth, or can be stored using less storage space (or both). However coding, unless it is *lossless*, also changes a signal, and can reduce its quality (its fidelity to the original). The reduction in data rate can be easily measured, while the reduction in quality is difficult to measure, for several reasons:

- the result depends upon the signal itself; some images or sounds may be more severely affected than others, for the same data rate reduction;
- the result is, properly, a subjective matter: people judge sound and image quality; and
- objective measures exist which are used to estimate reduction in quality, but there are several such measures and no full consensus.

This paper presents a graph for showing and tracking the effects of coding. The use of two dimensions gives quality as much significance (graphically) as data rate. The hope is that people making decisions about use of coding will have a clearer picture of the costs and benefits than is obtained from simply looking at the data rate reduction and knowing little, if anything, about the quality reduction.

The premise is that we should plot quality vs data rate for processes involving encoding a signal, to track the effects as a two-dimensional issue. Here is a simple example:

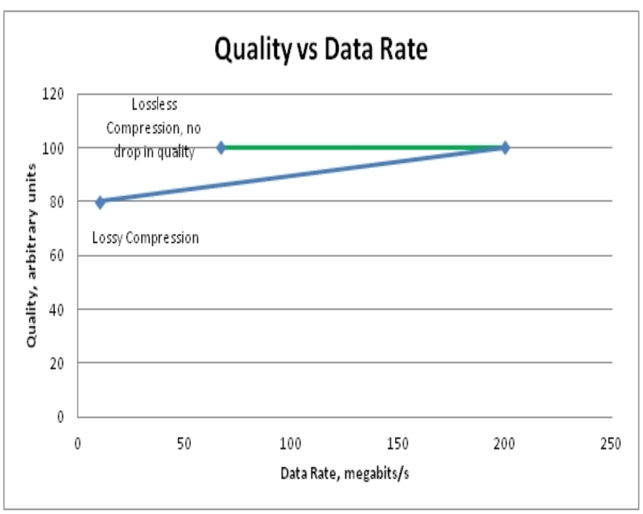

Figure 1: Quality vs Data Rate

The representation starts in the upper right, with an uncompressed signal at a data rate of 200 Mb/s, and a quality arbitrarily represented as 100. Any use of compression will result in a new data rate and a new quality, as another point on the grid.

The upper line shows lossless compression, with approximately a 3:1 reduction in data rate (down to 67 Mb/s) and NO reduction in quality (because the signal is completely unaffected). The lower line shows lossy compression, with a 20:1 reduction in data rate, and (on some kind of scale) a reduction in quality from 100 to 80.

As stated, the data rate dimension of a signal is easily measured. For quality, what is easy to measure is the power of the encoding error. If the encoded signal is subtracted from the original, the difference (the residual) should be small or the signal will be grossly distorted. The power of the residual is just the square root of the distortion as defined in Rate-Distortion theory.

It is conventional to measure the power of the residual, and compare that to the maximum possible signal power (peak power), producing PSNR – the peak signal to noise ratio. Although PSNR is widely used, there is general criticism of PSNR and many other suggested and even standardised objective measures used to estimate signal quality (eg UQI, VQM, PEVQ, SSIM and CZD; ITU-T Rec. J.246, J.247).

An obvious measure is to compare the power in the residual to the power in the original signal (rather than to peak power), and express the result in decibels.

$$\text{quality estimate} = 10 \log10 [1 - P(r) / P(o)] \qquad (1)$$

r = residual signal; o = original signal; P(r) = power in r

The plot needs a simple, constant value to represent 'no reduction in quality' and so rather than using SNR (which has a range from 0 to infinity, 'gets larger as it gets better' with no obvious maximum and is undefined for no error), we use the above formula which starts at zero for 'no reduction in quality' (and doesn't allow $P(r) \geq P(o)$). The scale may be too coarse, so 'centibels' could be used instead of decibels:

$$\text{quality estimate (centibels)} = 100 \log10 [1 - P(r)/P(o)] \qquad (2)$$

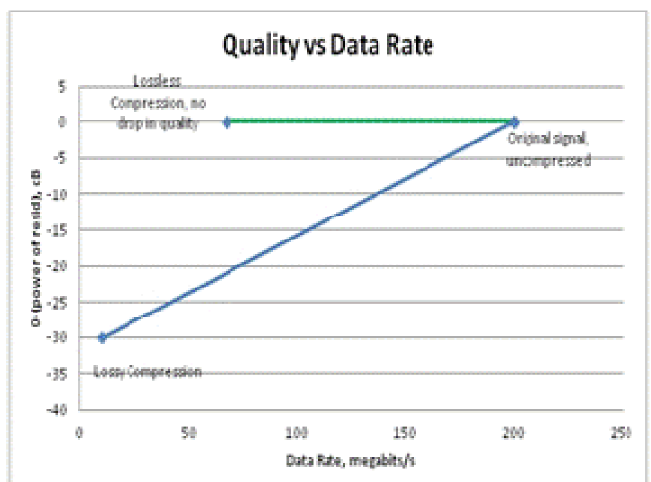

Figure 2: Power of Residual vs Data Rate

In Figure 2, it is assumed that the 20:1 lossy compression has produced a residual with a power 3 dB below the power of the original, which is very large.

The scaling of the Y-axis is not the main concern. The point is to measure the error, scaled in proportion to the original signal, and use that as the objective dimension which relates to perceived quality. No claim is made about the goodness of fit between this measurement and perceived quality. The decibel (or centibel) scale may relate to perceived quality in a non-linear way, but the relation should at least be monotonic. This metric is claimed to be better than PSNR because it compares the power in the error to the original signal, not to the theoretical peak value (which could bear no relation to actual signal power).

Using the Quality Graph

A whole range of different compression options can be plotted on the one graph, showing how each compares on data rate vs quality:

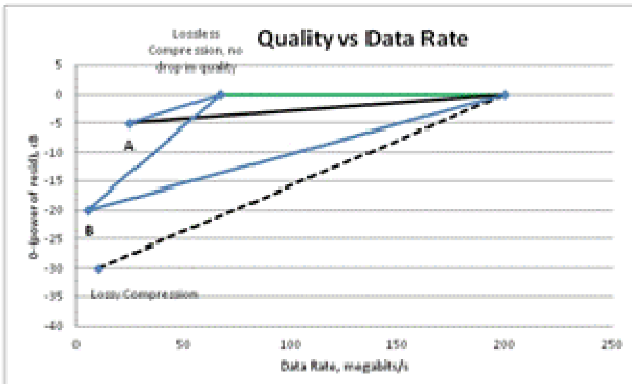

Figure 3: Reaching the same quality level from compressed and uncompressed starting points

In Figure 3, point A can be reached either in one or two stages: directly from the uncompressed signal, or via the lossless signal. The same is true of point B.

Similarly, a whole sequence of operations on one signal could be plotted, showing compression reducing data and quality, then decoding increasing the data rate but of course NOT increasing the quality (so a horizontal line, at best, for any 'motion' to the right in the graph). A second encoding would then take quality down another notch, and so forth.

Figure 4: Cascading compression – cycles of loss of quality

In Figure 4, the lossy-compressed signal at A can be decompressed back to a full data rate, but it cannot recover any lost quality, and so 'moves' horizontally. If a relatively modest compression is now applied, the quality can only drop. The signal at C has a higher data rate but lower quality than the signal at A.

Finally, how to balance data rate and quality? The cost of compression can be calculated in terms of the time and equipment it takes to produce the compression (and its complementary decoding for each use). The benefit in reduction in bandwidth and storage can also be calculated. All these numbers could in principle be converted to monetary costs, given sufficient knowledge about real costs.

Against whatever costs and savings are associated with data rate there would need to be a method to convert quality differences to monetary value. One estimate would be to equate any reduction in quality to an equivalent loss of that fraction of the value of the item. Producers and curators of content would pay more attention to signal quality if it were accepted that an item that was encoded at a data rate which lost 1 dB of quality

$$(10\log_{10}[1-P(r)/P(o)] = -1) \quad (3)$$

would have lost 20% of its monetary or heritage value!

$$(10\log_{10} 0.8 = -1) \quad (4)$$

Conclusions

This paper has reviewed the importance of quality in preservation, and the technology available for automation of quality control. It concluded with a simple, graphical approach to track the consequences of quality and data rate decisions, with the hope of demonstrating and continuously reminding anyone involved in production or archiving that encoding is not just about data rate, but also about quality. The suggested graph not only shows these two dimensions (with equal prominence for the quality dimension) but also gives a simple graphical presentation of the effects of multiple applications of encoding/decoding/recoding processes.

References

[1] R. Wright, Audiovisual Digital Preservation Status Report, PrestoPRIME Deliverable D7.1.3. (2010). www.prestoprime.org/project/public.en.html

[2] Edwin Klein. and Yola de Lusenet, Tracking the reel world: a survey of audiovisual collections in Europe, (European Commission on Preservation and Access, Amsterdam 2008). www.tape-online.net/docs/tracking_the_reel_world.pdf

[3] P. Schallauer, H. Fassold, M. Winter and W. Bailer, "Automatic quality analysis for video preservation," Proc. IEEE Int. Conf. on Image Processing, IEEE (2009).

Author Biography

Richard Wright has a PhD is signal processing from the University of Southampton. He has worked on the technical end of audio and video: speech processing , technology for speech and hearing and technology for audio and video preservation, as the research engineer in the BBC's archives(1994-2011) . Since 1995 he has led and participated in multi-partner research projects in digitization and digital preservation: the Presto projects. His current interests include preservation of social media.