

Considerations when choosing a video logging system and site-specific configuration details.

mediaproxy Pty Ltd, Australia

The choice of compression strategy is a fundamental requirement to the successful implementation of a logging solution, and will affect the perceptions of various users. To understand the variables involved is not 'rocket science' – but does require some consideration to realise that a piece of string is only 'so long', and there is a limit of what can be squeezed from any particular strategy.

The primary factors that define a logging system on the technical level are:

1. **Acceptable quality**
2. Reliability
3. Ease of Use
4. Capital Cost
5. Operating Costs

This document addresses the subjective values of 'Acceptable Quality'

This is a wide open statement – the most common logging requirement is referred to as 'VHS quality' – as this is most often what is being replaced - although both higher and lower standards are desirable in different logging situations.

The first principal to grasp is that traditional consumer level 'VHS' is an uncompressed, analogue recording format – with the quality limitation being a side-effect of the VHS recording process.

In reality, a good VHS recording is presenting about 250 lines of vertical resolution – in an analogue recording format, while the source material was at 500 lines or greater ^(interlaced).

By its nature, linear analogue recording applies the same process to the whole video program in the same manner, and any loss of quality, or degradation is progressive, based on the quality of the signal chain, and quality of maintenance - from a relatively sharp, clear image...eventually leading to a fuzziness that is unacceptable, or unstable in the viewing environment.

This is important, as the progressive degradation is potentially more 'acceptable' to many viewers, as they see a parallel in their domestic viewing expectations.

The digital equivalent of 'VHS quality' is a nebulous concept, as the processes invoked in making a digital recording are applied in a completely different way. In mediaproxy logging solutions, we typically record the digital streams at around 250 lines – as a compromise between cost of equipment and acceptable quality. To increase the resolution applies the square law – so double the resolution, requires four times the encoding resources. This is an important factor to consider when assigning the number of logging channels to each chassis.

The outcome is that those traditional thresholds of analogue viewability are realised in harsh 'digital' steps – both in resolution & detail, and other variables over the length of the recording. These could be spatial blocking of the image, motion 'cogging' or other playback artefacts.

However the benefits of digital recording and playback access will usually outweigh the trade-offs between progressive analogue and harsh digital thresholds. Keep in mind that even the original full-resolution video will not scale up well to a desktop computer monitor – that is often running at nearly twice the original video resolution. (Figure 1 - Capture – Compress – Display)

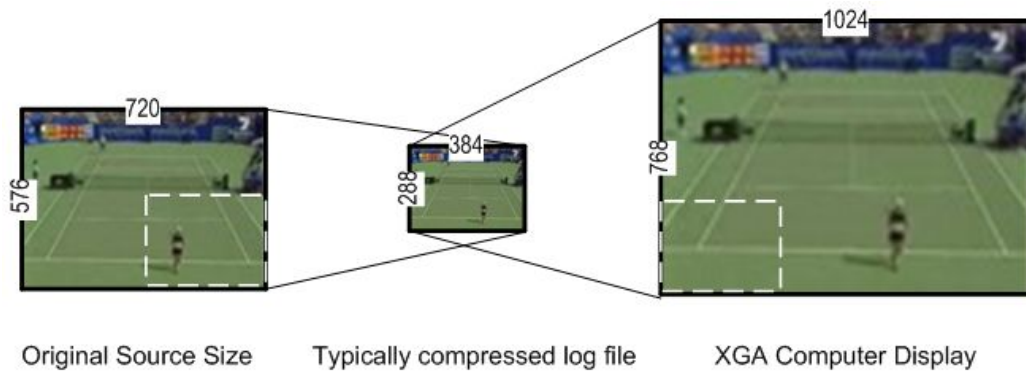


Figure 1 - Capture – Compress – Display

Temporal vs Spatial Compression

As mentioned earlier, analogue recording is a linear process applied to the whole of the media in the same way. Digital video compression works in an entirely different way – significantly dependent on the actual content and quality of the source material.

For this reason – consideration number one – is that for digital compression, the source content should be the best quality possible – stable image, no flickering, no noise etc, when possible. If these criteria can't be met, the compression process will spend an inordinate amount of time working on these source artefacts – rather than the actual video or audio content!

How It Works...

To explain compression simply - at the first level – each video frame is examined for content and complexity, and the encoding algorithms process that **Spatial** information to generate the smallest equivalent output with the most detail in the available period of one frame. If this strategy was applied in isolation, it would be referred to as 'I-frame only' compression – Intra-coded frames only.

This spatial understanding also has impact when the video is rendered out to a 'hard copy' tape recording. If the source is compressed to 320x240 (VCD resolution) and then expanded on playback to desktop size for VHS output – it is desirable to use lower resolution screen display modes for that output – e.g. VGA (640x480) is an ideal candidate that realises the original resolution without blowing the pixels up disproportionately. Higher desktop resolutions will work fine, but add nothing to the perceived output quality.

However, as this does not provide the most efficient compression on *typical* video content, a second dimension of compression is applied – temporal compression.



Spatial vs Temporal Compression

Figure 2 - Compression Models

Temporal (*across time*) compression takes the detail from the first pass of spatial data, and rationalises it across a number of frames (often referred to as the GOP – Group Of Pictures) to identify repeating elements in those adjacent video frames which may be discarded, and subsequently reconstructed in real time at the moment of playback.

This provides significant improvements in compression efficiency, and allows less digital data to carry substantially the same video content.

The output of this process is a number of I-frames at the GOP boundaries, and intermediate B- (Bi-directional) or P- (Predictive) frames - these terms are borrowed from MPEG, but apply in principal to most modern compression strategies.

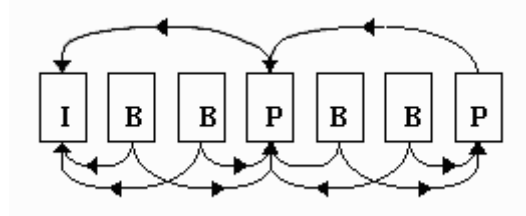


Figure 3 - GOP frame structure – example only

When a scene-change / cut occurs, the encoder forces a new I-frame in order to capture the new scene information – and moves forward from in a new GOP with ‘difference’ frames until the GOP expires, or a new I-frame is demanded..

Quality (aka the inevitable trade-off)

Now that we know we can ‘have our cake and eat it too’... there is a gotcha.

I mentioned in the earlier paragraph that this compression can only be achieved in real-time (one frame per frame) with the available resources. In some cases this can be mitigated with dual- or multi-pass encoding, however our logging requirement for real time recording from live sources – necessitates live encoding. No second chance.

This is where the encoder must be instructed to dynamically allocate spatial (detail) or temporal (motion and smoothness) algorithms to achieve the desired result. As luck would have it – no single setting will work for all source material, unless you have an enormous amount of CPU power and bandwidth available.

So we need to identify an ‘acceptable’ trade-off capable of meeting most expectations and source conditions.

Encoding bit-rate

OK – we have compressed the video data from our original 240Mbps source (the approximate payload of standard definition video streams, or a little less from typical analogue sources with added noise and artefacts). The problem is that we don’t have enough disk storage to record, or bandwidth to distribute multiple users at 240Mbps... (Somewhat more than your average 100Mbps LAN with 30-40Mbps sustained throughput).

Thus we have to nominate a practical bandwidth for the media storage, and a second rate for any streaming requirements. This is where the real compromise is occurring.

For typical compliance logging requirements –we’re going to need somewhere between 150Kbps and 500Kbps for the recorded content. With the material encoded at quarter-frame (e.g. 320x240) – this renders quality suitable to ‘confirm’ that material went to air in the correct form. Using the higher end of this scale provides much better quality – almost good enough to watch that episode you missed last night.

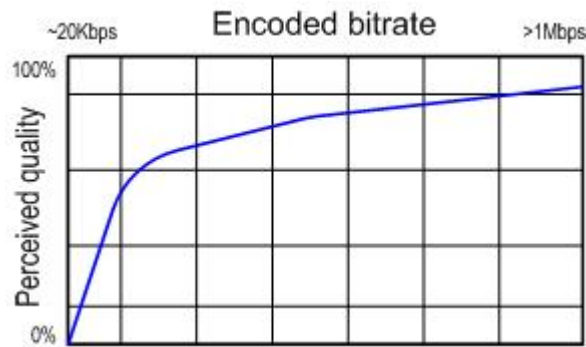


Figure 4 – Perceived Quality vs encoded bit-rate

Moving upwards from 500Kbps will incrementally improve the quality, but not as aggressively as the move from 150 up to 500. These improvements will mainly occur in the edges and detail (remember spatial compression?)

For *near* broadcast reception quality – consider 2Mbps at half-resolution, or 1.2Mbps at full resolution... remember those gotchas! The piece of string is only “so” long, and there s a good chance that the encoding CPU load will move in to demand more power as the quality increases.

Encoding frame-rate

Where possible – our recording should reflect the same frame rate as the source material – but it’s not really important! Why?

...because the ‘lossy’ compression of the video content is going to mask or distribute the original content (temporal compression) related to the nominated encoding profile.

If you feel that motion looks acceptable at 20fps – no-one is going to argue. The encoder put the same video content into less frames by distributing the temporal content – just as it would have at 30fps. The most noticeable visible artefact at 20fps will be flicker - when related to any adjacent 25/50Hz or 30/60Hz displays.

Mediaproxy proposals are often made at 15 - 20 fps – not always at the source frame rate (25 or 30fps) – as the compression is working in a different paradigm to the video display... as long as you are comfortable with the output image for the intended purpose – 20fps is fine for compliance logging. (P.S. your PC playback monitor isn’t running at 25 or 30fps either).

Increasing the encoded frame rate – increases the CPU load, and must be considered as a major factor when loading a single chassis with multiple channels.

But what if I want to increase edge detail in the logged recordings... i.e. higher bit rate or resolution?

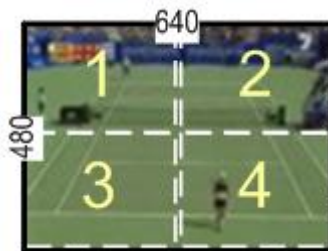


Figure 5 - Encoding at higher resolution

You can increase the encoding resolution – but remember we discussed this earlier... double the resolution (320x240 → 640x480) requires FOUR times the encoding power. This will typically turn each 4-channel logging chassis into a single channel – due to the increase in required processing power per channel.

Time Code display

For the same reasons noted above – burnt-in time code is only relatively useful – as the time display may render on-the-frame, but the actual video behind the time display could be several frames earlier – depending on how the temporal compression is applied at that moment!

Choice of codecs

Once we get this far, the choice of codecs – MPEG, Windows Media etc – is a sidebar. Each codec family has benefits and limitations – specifically targeted at different types of content and target audience.

It should be noted here- that different codecs have ‘sweet spots’ at different resolutions, bit-rates and frame rates. Often these can only be determined in use with the actual source material and customer expectations.

In most applications, mediaproxy utilise Windows Media 9 for recorded programs, as they offer a marginally better video image at the same bit rate than other codecs – however at a price of greater CPU load. For this reason, where CPU capacity is limited, we may propose MPEG4 codecs on low-resolution streams – due to the lower CPU impact at the encoder and player.

Hardware vs Software Encoders

There are really good reasons to use hardware encoders – less CPU load and better encoding in many cases. However there are many features of software encoders that make them eminently suitable for the continuous recording mode of a logging system.

If you have a preference for either strategy, discuss this with us, as it is relatively straightforward to adopt different components into the mediaproxy logging platform.

It is also worth keeping in mind that other server-side activities will require processor and memory resources... caption decoding, text insertion, audio monitoring along with general system housekeeping.

Conclusion

The use of low-bit rate logging is targeted at specific market requirements - legal compliance recording, long term retention of the captured content, and intuitive desktop retrieval of that media.

These systems are not intended to address high-resolution technical analysis of the content – they are implemented to provide sufficient quality to satisfy legal and contractual evidence requirements. There is always an option to return to the original media to verify technical standards in the content.

Full-resolution short-GOP ‘broadcast’ servers are available at a premium price.

Low-bit rate logging addresses a specific application – we believe mediaproxy offer more options in that segment than other vendors.

-end-