THE QUEEN'S AWARDS FOR ENTERPRISE: INNOVATION 2011

Research & Development

White Paper

WHP 206

February 2012

On the concatenation of video processing within news journalism production

S Thompson

BRITISH BROADCASTING CORPORATION

White Paper WHP 206

On the concatenation of video processing within news journalism production

S Thompson

Abstract

In a genre which focuses on immediacy, the move to high definition with its requisite high datarates and processing loads is proving particularly challenging if picture quality is to be maintained. This paper presents research undertaken alongside BBC Journalism into the encoding, compositing and scaling processes required by the news production workflow and their effects on the high definition picture quality delivered to the home.

In this paper the complex processes used within the news production life-cycle are modelled e.g. in-field editing, low datarate contribution circuits, graphics and the transmission of content back to the field for re-purposing. The onwards playout through a broadcaster's transmission systems are also discussed.

This paper focuses on the criteria to be considered in making a choice of high definition compression codec for use in a journalism production chain, the requirements for initial acquisition and the problems caused when some elements of the system are pre-determined. Finally we show that, in some circumstances, using a standard definition workflow gives better results.

The content of this paper was originally presented as a poster at IBC 2011.

Additional key words: Picture Quality, High Definition, HD, Standard Definition, SD, Peak Signal to Noise Ratio, PSNR, SNG, Satellite News Gathering, H.264, XDCAM, AVC-Intra, Journalism, Workflow

White Papers are distributed freely on request.

Authorisation of the Chief Scientist or General Manager is required for publication.

BBC Research & Development
White Paper WHP 206

On the Concatenation of Video Processing Within News Journalism Production.

Simon Thompson

1 Introduction.

This paper originated in work undertaken by BBC Research and Development (BBC R&D) to support BBC Journalism in an OJEU[1] procurement exercise for a new editing system for their London Broadcasting House base. The non-linear editor at the centre of the procured system came from Quantel[2]. The author worked with BBC Journalism to create a test plan to allow objective checking of the available codecs and configuration settings in tendered equipment and to ensure that a number of likely scenarios were catered for. This base edit codec may differ from the edit codecs used in the field, which are usually chosen to match the camera in use for speed of editing.

The content of this paper was originally presented as a poster at IBC 2011.

2 Our Correspondent in Benghazi - An example of today's SD Journalism Workflow

To help us understand more fully the difficulties faced in delivering news footage, let us explore a current, real-world, standard definition example workflow. Some of the BBC camera teams working in the Libyan desert used Beta SX-based cameras. The material recorded in Beta SX format was transcoded to DV25 via the camera's SDI output. The DV25 was ingested to a laptop-based non-linear editor (NLE) via Firewire. The NLE was used to edit the DV25 material without further transcoding in the field.

The resulting content was fed back to BBC Television Centre (TVC) in London via two distinct routes (budget dependent). Route 1 used a satellite uplink dish based in Benghazi. Camera crews would take their NLE laptops to the uplink feedpoint and utilise a conversion device to generate a realtime SDI signal which was fed to the satellite network. Route 2 utilised a Quicklink network device and sent the material to TVC London via a Broadband Global Access Network (BGAN), which took 30-40 minutes on average to send a 2 minute and 30 second piece. The Quicklink device transcodes the DV25 to H.264 or VP7 encoding to maximise throughput whilst minimising datarate.

If required, archive material can be received in the field from TVC. Normal practice whilst in bureaux is to use the BBC's Jupiter Media Exchange system (JEX) file transmitter. Archive footage is transcoded and sent via the Internet to a local client, and then converted to DV25 for use in the NLE. This is known as "round tripping". Once a finished piece has been received back at base, it can be versioned for a specific news bulletin and have additional graphics inserted.

3 Working in the compressed digital domain - what is the worst-case for Journalism?

In order to understand the absolute necessity to control the video processing undertaken at various points within a workflow, let us explore a worst-case scenario (this is shown in figure 1). This is

[1] Official Journal of the European Union
[2] http://www.quantel.com/

appropriate to both UK reporting and reporting from abroad.

1. Camera acquisition – after the usual camera processing, content is recorded onto media using a datarate reduction encoder.

2. In-field editing – the in-field editor will create a rough cut and may insert material requested from base. This will normally utilise a NLE on a commercial off-the-shelf computer. Proprietary software may be required to allow the files produced by the camera to be edited. To speed the process the field edit codec is matched to the camera acquisition codec whenever possible.

3. Contribution to base – the contribution to base may be a low latency satellite link (DVB-DSNG), an Internet file transfer protocol (FTP) link, physical camera media card or tape.

4. "Round-tripping" circuit to field – This is a non real-time link to the field editor to allow him to insert archive material.

5. Base editing – this is a larger, networked editing system running on custom hardware, normally using an open standard codec capable of ingesting material from the contribution link, camera media, archive etc. It will be capable of adding graphics overlays.

6. Playout storage – this storage device holds the finished video piece until the newsroom automation system requests it to play the file to air. It is not used for live events.

7. Direct-to-home (DTH) encoder – this encoder encodes the video in a way suitable for watching on a domestic receiver, *e.g.* H.264 for HD delivery via DVB-T2

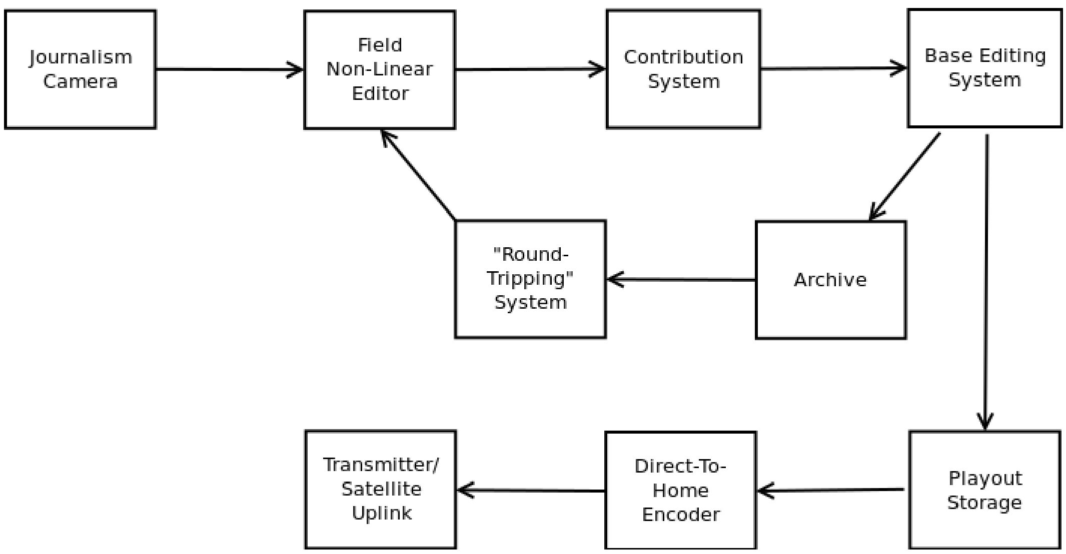

Figure 1: An example journalism workflow.

The transfers between each stage may take the form of an uncompressed video signal, a direct file copy or a file transcode and copy.

4 Simulations and Results

In order to assess the performance of each base edit codec, BBC R&D measured the objective performance of the end-to-end production and transmission chain using a selection of BBC and EBU test material (EBU Crowd Run, EBU Park Joy, EBU Ducks Take-off, EBU Into Tree, EBU Old Town Cross, BBC Football, BBC Bleak House, BBC Jools Holland).

4.1 Assumptions

1. Editing in the field will utilise the same codec as the recording made in the camera without a transcode.

2. Editing at base may not use the same codec as the in-field editor but will if practicable.

3. Wrapper file format changes are done without transcoding the A/V file content.

4.2 Methodology

The author uses Peak Signal-to-Noise Ratio (PSNR)[1] to objectively assess the quality of image output from each encoder or process. In order to best represent the quality of image seen by the audience, the author always presents PSNR results as they would be seen by the audience, *i.e.* with a final stage of direct-to-home (DTH) video encoding included. The PSNR is calculated using the open-source "calcpsnr" software. The DTH encoder is the same as used for "BBC One HD" in the UK with similar settings. The production and and contribution codecs under test are provided by proprietary hardware encoders. PSNR measurements are taken between the original source and the output of the DTH encoder, *i.e.* across the whole production and distribution chain.

These tests are designed to allow BBC R&D to measure the impact of upstream production codecs on the emission encoding toolset. Acquisition and production codecs utilise far higher datarates than DTH encoders as they are designed to have minimal impact on the video signal to allow further downstream processing. Neither are they subject to transmission datarate constraints with typical datarates in the range of 5 to 20 times that of the DTH encoder. Therefore, we can say that the dominant encoding toolset in the chain is in the DTH encoder. As the DTH encoder settings are constant, this dominance is identical in all tests and PSNR loss measurements will correlate well with subjective performance measurements as production codecs are changed, see [2].

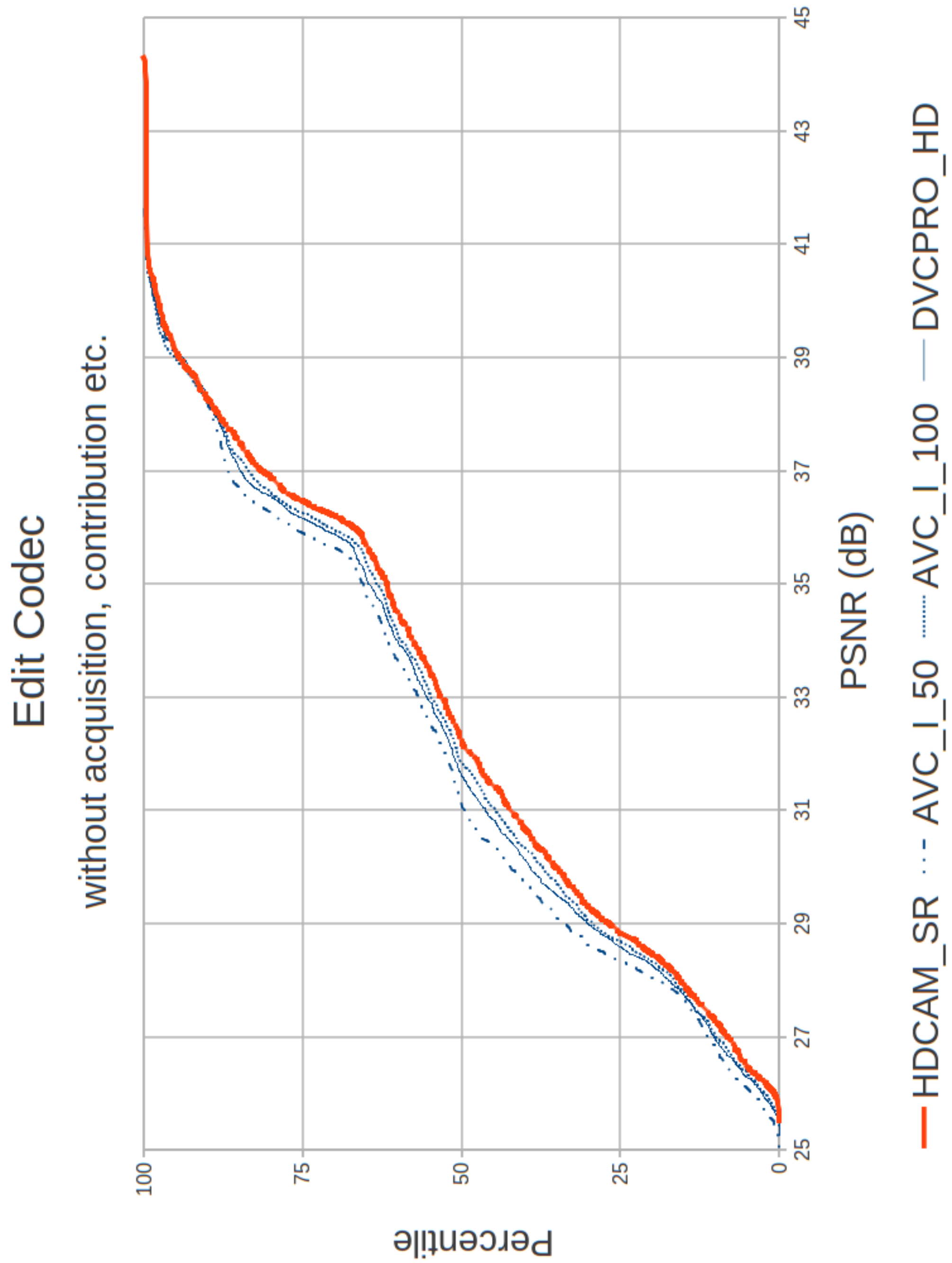

Figure 2: Edit codec PSNR.

Results are presented in cumulative distribution frequency graphs as they clearly show the system performance across a range of complexities. On the graphs, the abscissa represents the PSNR in decibels and the ordinate the percentile of frames with less than that PSNR. For example, if the trace crosses the 35 dB point at the 62^{nd} percentile, 62% of frames have a PSNR of 35 dB or worse. An example is shown in figure 2.

Previous work by BBC R&D has shown that the variance in PSNR caused by using a 1440x1080 codec over a 1920x1080 codec is negligible with current cameras as the full spectrum is not used, see [3]. Therefore, PSNR can still be used as an indication of the chain's performance even when the resolution of the production codecs differs.

A reference trace is provided utilising the HDCAM-SR original test material to highlight the impact of the production chain on the image quality. HDCAM-SR is taken to be the best possible result that can be obtained from the production process as it uses a very high datarate MPEG4 codec (440 Mbps studio profile, approximately 4 times that of a typical production codec).

Due to the nature of a journalistic organisation's work and their requirement for immediacy, source picture quality is sometimes compromised due to complexity of pictures, low datarate file transfer, the nature of the camera used, location or environmental conditions, etc. We therefore concentrated on the 20^{th} to 60^{th} percentile of frames within the test material. This represents the majority of footage shown, accepting that under certain circumstances for the most challenging material, the final quality delivered to the home may be less than ideal. The results for the least challenging frames are discarded as they are usually of sufficient quality.

4.3 Codecs Do Not Offer the Same Level of Performance

It can be seen in figure 2 that the edit codecs under test, with no acquisition or contribution codec, do not perform equally well. Objectively, AVC-I 100 degrades video quality by 0.30 dB on average over the range of interest (20% - 60%) compared to the reference. This outperforms both AVC-I 50 (0.83 dB) and DVCPRO-HD (0.49 dB). The difference between the best-performing codec and the worst averages 0.53 dB.

However, the more interesting result is to learn how the various codecs perform when concatenated with other codecs.

4.4 Some Combinations of Camera and Base Edit Codec are Better Than Others.

In an ideal World, we would engineer the production chain so that it has negligible impact on picture quality. This is not financially practical. Instead our goal is to introduce an impairment that is just perceptible by non-experts for the majority of content. This equates to a PSNR loss of around 0.5 dB compared to the HDCAM-SR reference.

Using just the EBU test clips, it can be seen that:

1. From figure 3, when AVC-Intra 50 is used as an acquisition codec, for best results AVC-Intra 50 must also be used as the edit format with direct file transfer. However, the other edit codecs perform almost as well. The degradation in objective video quality compared to HDCAM-SR using AVC-Intra 50 averages 0.73 dB over the range of interest.

2. From figure 4, the XDCAM-HD422 acquisition codec with an AVC-Intra 100 edit codec performs best with a degradation averaging 0.47 dB compared to the HDCAM-SR reference over the range of interest. This outperforms the worst combination of acquisition and edit codecs (AVC-Intra 50) by an average of 0.22 dB. XDCAM-HD422 outperforms the other acquisition codecs.

3. From figure 5, the XDCAM-EX35 acquisition codec with an AVC-Intra 100 edit codec performs best with a degradation averaging 0.77 dB compared to the HDCAM-SR reference over the range of interest. This outperforms the worst combination of acquisition and edit codecs (AVC-Intra 50) by an average of 0.13 dB.

4. The losses for all combinations of acquisition and edit codecs are greatest for the most easy to encode frames, where any loss is less significant.

5. The differences in PSNR performance of the base edit codecs are less marked when combined with acquisition codec than standalone, so other factors may influence the choice of base edit codec.

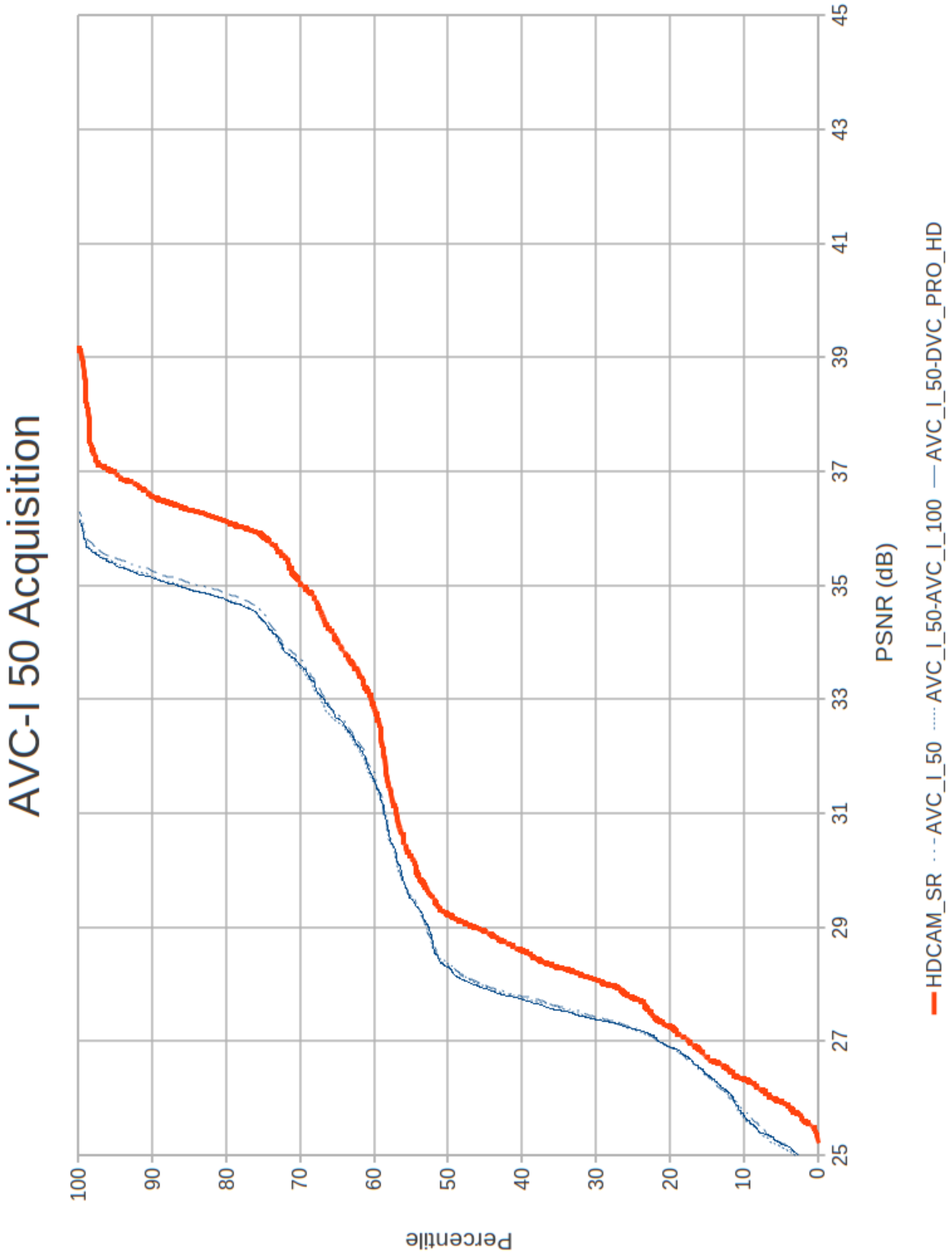

Figure 3: AVC-I 50 Acquisition Codec with AVC-I 50, AVC-I 100 and DVCPRO-HD Base Edit.

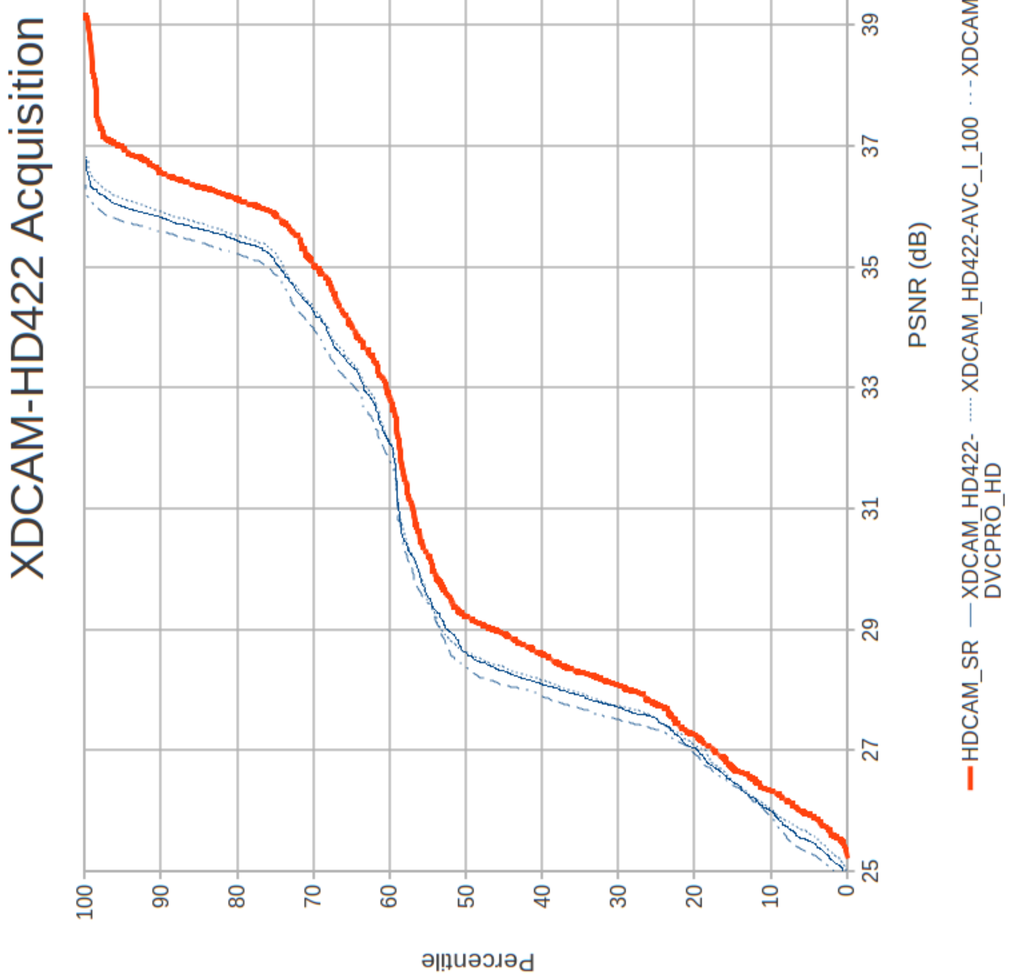

Figure 4: XDCAM-HD422 Acquisition Codec with AVC-I 50, AVC-I 100 and DVCPRO-HD Base Edit.

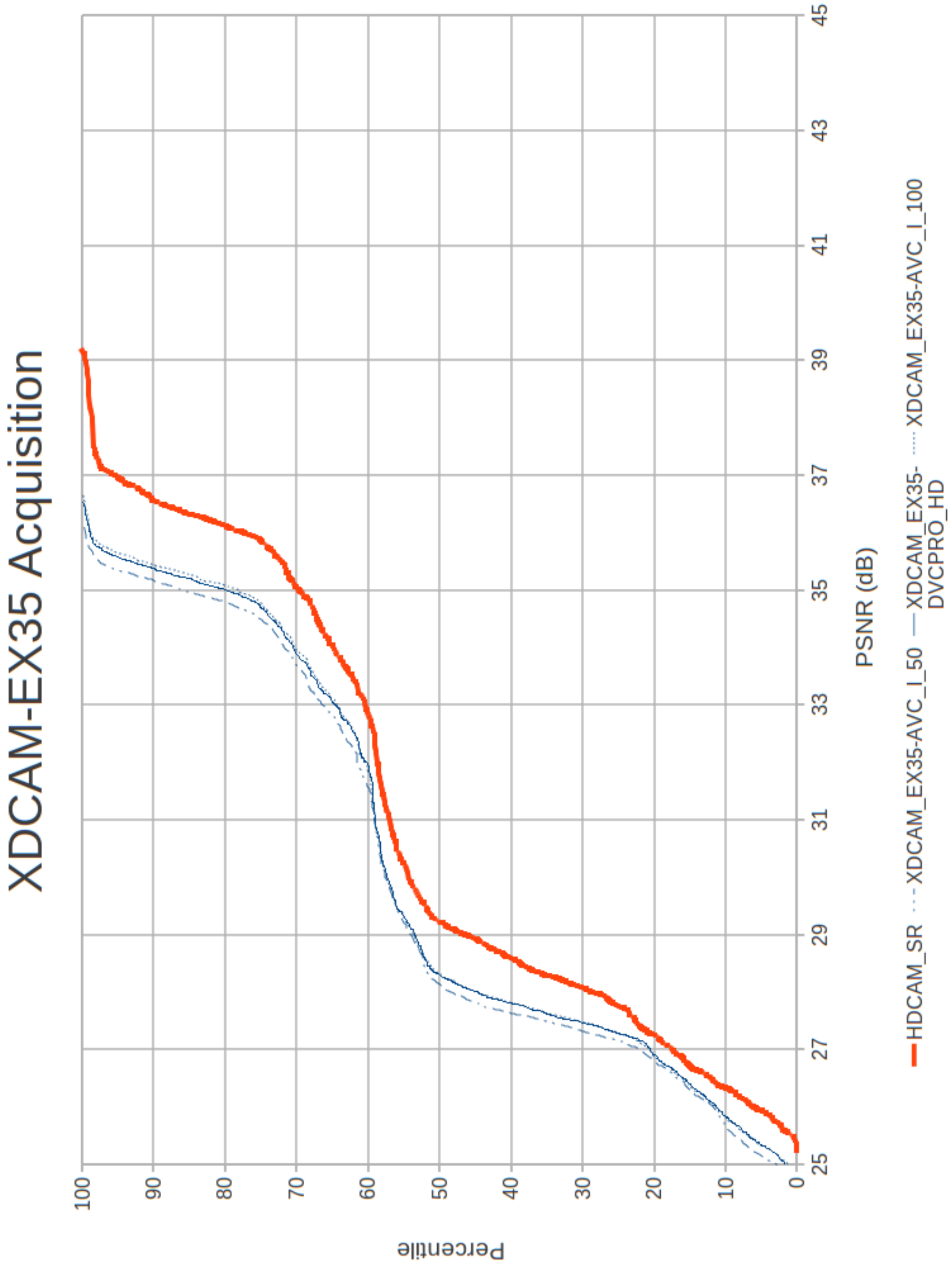

Figure 5: XDCAM-EX35 Acquisition Codec with AVC-I 50, AVC-I 100 and DVCPRO-HD Base Edit.

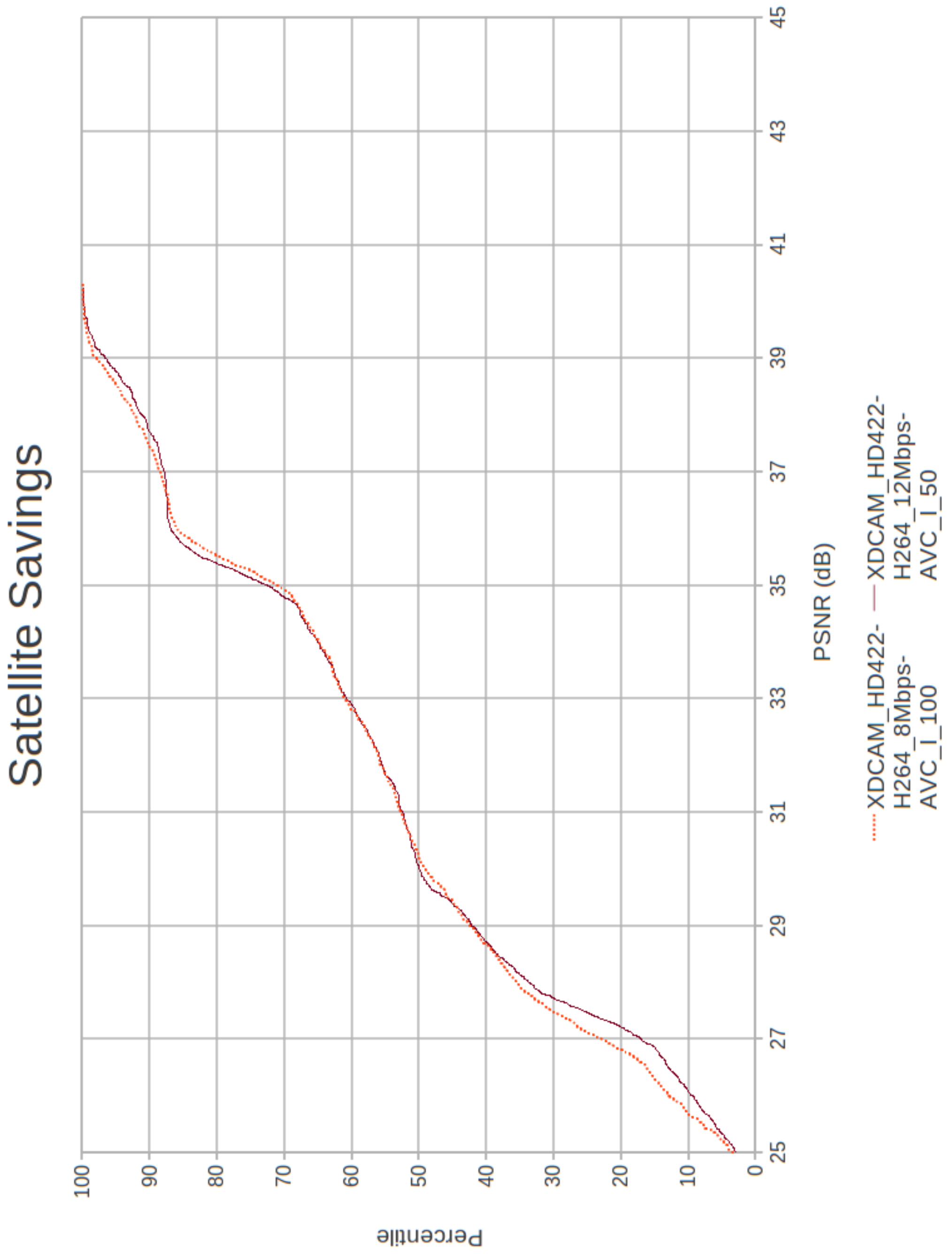

Figure 6: PSNR Showing Potential Savings in Satellite Bandwidth.

4.5 The Choice of Edit Codec affects the Choice of Contribution Codec.

A great deal of content will be delivered to base via a satellite contribution codec or similar. The author considered whether, by using a higher datarate edit codec within the newsroom, significant financial savings might be made by reducing the satellite capacity used to bring video back from shooting assignments around the World.

Figure 6 shows two results for a combination of acquisition, contribution and edit codecs. It can be seen in figure 6 that video acquired on an XDCAM-HD422 camera, contributed to base using H.264 at 12 Mbps and edited using AVC-Intra 50 has a similar PSNR to the same camera codec contributed at 8 Mbps and edited at 100 Mbps, *i.e.* a satellite datarate saving of 4 Mbps.

Similar results were seen for XDCAM-EX35 and AVC-Intra 50. The size of any savings will likely depend on the satellite circuit costs and the scale of the base edit system.

4.6 Chrominance — The Application of Graphics Overlays

Graphic overlays are an important feature used extensively in journalistic reporting. These will often be rendered in the base edit codec rather than as uncompressed graphics files (*e.g.* PNG). An "alpha" channel is not always available so the "key" signal must be derived from the graphic itself. Sub-sampling of chrominance and coding noise makes it harder to derive a clean "key".

At the time of undertaking the tests, BBC R&D's HD edit system did not support the candidate AVC-I based edit codecs. To investigate quality impairments to chroma-keying of graphics, representative codecs from the available range were chosen instead:

- DVC-PRO HD (1440x1920 4:2:2) – an implementation of a codec offered by the Quantel system. The horizontal sub-sampling is also similar to Quantel's AVC-Intra 50 codec.
- XDCAM-EX35 (1920x1080 4:2:0) – similar to Quantel's AVC-Intra 50 in terms of coding artefacts and vertical chrominance resolution, but of higher horizontal resolution.
- XDCAM-HD422 (1920x1080 4:2:2) – similar to Quantel's AVC-Intra 100.

Visual inspection showed that only the 1920x1080 4:2:2 codec performed well. The 4:2:0 codec showed issues in correctly identifying the chroma key area whilst the horizontally sub-sampled codec was not able to represent diagonal straight lines correctly within the editor, leading to a jagged appearance of such lines. More formal investigation would be required to confirm this result.

5 Trawling the Archive - "Round-tripping"

Within the BBC, the process of sending data from base to the field for editing is known as "round-tripping". The majority of this material will be from the archive and the transmission will use a low-datarate data link. At present, most of the content is standard definition but the ratio of high definition content will increase in due course. The JEX system uses an open source H.264 encoder to compress the video prior to sending at a maximum datarate of 4 Mbps (datarate is dependant on the location of the journalist). In the HD broadcasting system, we have the option of sending data as HD, sending the data as SD and upconverting to HD prior to editing and sending the data as SD (*i.e.* without scaling to HD).

For each option, the settings used by JEX and its video scaler require tuning to maximise picture quality. Figure 7 shows that when using the SVT/EBU test sequences below 2 Mbps (*i.e.* when linking to an area of the world whose infrastructure can't support the higher datarates), it is possible to use an SD link and scale the video back to HD and have a higher PSNR than video sent as HD. Whilst PSNR may not be an exact measure of picture quality in this scenario (due to the presence of filtering), these results are a good match to the author's subjective picture quality assessment.

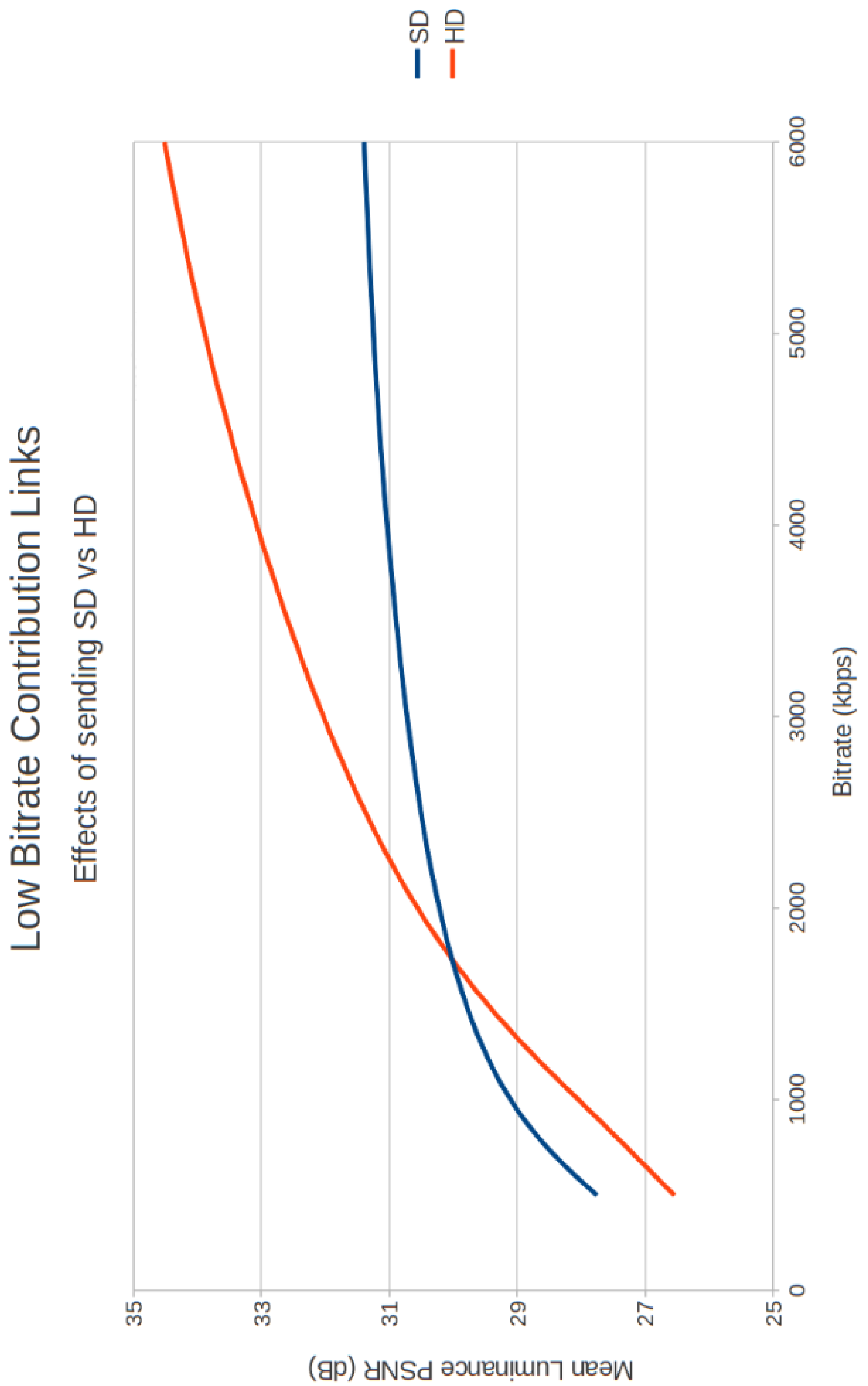

Figure 7: PSNR for HD vs. SD - Extrapolated.

6 Conclusions

- The performance difference between base edit codecs is less marked when combined with acquisition codecs than when used standalone.
- Some combinations of acquisition, edit and distribution codec offer marginal performance, especially in the presence of contribution encoders. Prior to choosing a workflow, these combinations should be tested to ensure a chosen combination is adequate. The best combinations use XDCAM-HD422 acquisition codecs with AVC-Intra 100 and DVCPRO-HD edit codecs.
- Expenditure at one point in a workflow can bring cost savings at another (*e.g.* satellite datarate).
- In designing a workflow the number of video processes undertaken should be minimised to ensure best picture quality and fastest turn-around. For example, it makes sense to minimise transcodes by using editing and archiving platforms that can natively support the same codec family as the camera.
- In designing a workflow the architect must consider all aspects of production, not just codec compression performance, *e.g.* the ability to insert overlays.
- Due to the nature of journalism, not all parts of the workflow will be ideal. Unusual acquisition formats will need to be incorporated into the workflow *e.g.* content from other news agencies and round-tripping, so the quality may be compromised to break the story sooner.

References

[1] *Objective perceptual multimedia video quality measurement in the presence of a full reference*, International Telecommunications Union Telecommunication Standardization Sector - Pre-published Recommendation J.247, August 2008. [Online]. Available: http://www.itu.int/rec/T-REC-J.247/en

[2] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *IET Electronics Letters*, vol. 44, no. 13, pp. 800 – 801, June 2008.

[3] A. Roberts, *Circles of Confusion.* L Ancienne Rue, 17A, Grand-Saconnex, Geneva, Switzerland CH-1218: The European Broadcasting Union, 2009.

7 Acknowledgements

The author would like to acknowledge the work of Justen Dyche (BBC Journalism) in helping compile the graphics overlay sequences used in this paper and Phil Goodwin in providing data related to the BBC's operations in Libya.