Wave 7: DV-based video cassettes

Draft technical specifications for the transfer to files

White Paper – v1.0

Brecht Declercq, Gaël Fernandez (VIAA)
Peter Bubestinger-Steindl (AV-RD)

VIAA vzw – Kleindokkaai 9a, 9000 Gent

Document revision history

This document initiated from the internship of Gaël Fernandez Lorenzo (INA, Master en Gestion des Patrimoines Audiovisuels et Numériques) at VIAA from March to August 2018. Part of his internship goals was to deliver an advice regarding the transfer workflow and output specifications for DV-based cassettes, to prepare the drafting of VIAA's tender for the outsourced transfer of Flemish DV-based cassette collections.

At the end of his internship, it appeared that finishing this advice before the end of his internship was not achievable and more expertise was needed to make further progress. Peter Bubestinger-Steindl (AV-RD) was contracted to help finalizing this advice. In the course of September 2018 until January 2019 this document was elaborated further, restructured and finished.

Version	Date	Author	Revisions
v0.1	14.08.2018	Peter Bubestinger-Steindl	Original version
v0.2	12.09.2018	Brecht Declercq, Gaël Fernandez	Further structuring, VIAA lay-out
v0.3	28.12.2018	Brecht Declercq	Further structuring and revision
v0.4	11.01.2019	Brecht Declercq	Revisions upon suggestions by Peter Bubestinger-Steindl
v0.5	31.01.2019	Brecht Declercq	Revisions upon suggestions by Peter Bubestinger-Steindl
v0.6	12.03.2019	Brecht Declercq	Minor adjustments to fig. 4 and 6
V1.0	14.03.2019	Lobke Vanden Eynden	Numbers update page 7 and renaming fig.

Table of contentss

I. Project context

1.1 VIAA

The Flemish Institute for Archiving (VIAA) digitizes, stores and provides access to audio visual material, photos, documents, etc., together with partners from the cultural, heritage and media sectors. Its mission can be summarized as to preserve and archive the digital heritage of Flanders in a sustainable manner and to make it accessible to everyone.

VIAA was founded on 21 December 2012 by the Flemish Government. The founding of VIAA was entrusted to iMinds (a strategic research centre of the Flemish Government, focusing on ICT), but since 1 January 2016 VIAA is an independent governmental organisation, with its stakeholders represented in the Board.

Several research projects from 2008 onwards triggered the founding of VIAA, as they demonstrated that there was a great demand for a more integrated approach to digitising, digitally storing and sharing our heritage. According to a preliminary estimate from February 2013, there were around 600,000 hours of audio, video and film material in Flemish archives still to be digitised. In terms of content, it ranges from film, television and radio programs to music recordings, audiovisual art objects, home videos and recordings of all manner of events.

Next to audiovisual media, digital archiving and the various carriers pose a great challenge for other kinds of information carriers. That's why when VIAA came along it was explicitly asked to also assume a position in other domains. An example of this extended scope is the project 'News from the Great War', which digitizes and stores trench newspapers and other newspapers and magazines from the First World War and makes them accessible. Originally the acronym VIAA stood for 'Vlaams Instituut voor Audiovisuele Archivering' ('Flemish Institute for Audiovisual Archiving'). However, for the time being, the name VIAA will remain unchanged and the organisation will be referred to as the Flemish Institute for Archiving, in order to emphasize the versatility of the organisation.

VIAA does not own a collection, but acts as a service provider to an ever-larger group of Flemish media and cultural organisations. These services include:

- **digitisation**: converting the information on physical carriers into digital files,
- **archiving**: storing digital material sustainably, attaching metadata and indexing the material so that it becomes retrievable and understandable for different target groups,

- **interaction**: making the digital material available for education, scientific research and the general public, with full respect for intellectual property, ethical and other kinds of rights.

1.2 VIAA's content partners

VIAA does not own a collection, but acts as a service provider to a growing group of currently 127 Flemish media and cultural organisations (content partners). These partners include:

- the Flemish public broadcaster VRT
- regional broadcasters
- commercial broadcasters
- cultural heritage institutions recognized under the Cultural Heritage Decree
- the Flemish governmental institutions
- the Flemish city archives
- Performing arts organisations, recognized under the Arts Decree

The partners take part in VIAA's digitisation projects, depending on the carrier formats they have in their collections. VIAA plans, coordinates and finances the digitisation projects for them, by:

- inventorying the collections
- drafting the specifications for the digitisation process
- selecting a specialised digitisation service provider
- coordinating the project logistics of carriers and files

After the digitisation, the original carriers are returned and stored again by the content partners. The files are ingested on the VIAA servers and made accessible via the VIAA Media Asset Management system (MAM). Finally, they're made available to VIAA's target groups, whilst respecting IPR and other rights.

II. Wave 7 transfer project

As a part of its mission to digitise the Flemish audiovisual heritage, VIAA would like to execute a state-of-the-art transfer-to-file project of **DV-format based video cassettes**, in 2019. The estimated volume of carriers is around 20.222, coming from all kinds of content partners (cf. fig. 1).

Content partners	DV	sub rate	DVCAM	sub rate	DVCPRO 25	sub rate	DVCPRO 50	sub rate	Total
Cultural heritage	384		786		249		0		1419
Sub. Rate	4,28%		7,24%		64,84%		0,00%		7,02%
Performing arts	4321		416		3		1		4491
Sub.Rate	48,11%		3,83%		0,78%		100,00%		22,21%
Local broadcasters	2321		9623		132		0		14076
Sub.Rate	25,84%		88,64%		34,38%		0,00%		69,61%
Governemental archives	205		31		0		0		236
Sub.Rate	2,28%		0,26%		0,00%		0,00%		1,17%
Total	**8981**	**44,41%**	**10856**	**53,68%**	**384**	**1,90%**	**1**	**0,01%**	**20222**

Fig. 1: distribution of the estimated numbers of DV-cassettes in types and per sector.

An unknown number suffers from several forms of decay, urging for intervention. These carriers are kept in the collections of at least 65 content partners, of which the biggest part is from the performing arts sector.

Before the envisioned transfer-to-file project can start however, one of the aspects to prepare are the technical recommendations of the transfer to files.

III. The mission

In this part of the project preparations, VIAA would like established:

- A suited output format (container, codec and specifications), taking into account:
  - the properties of the digital signal on the cassettes
  - requirements of digital sustainability of file formats, codecs and specifications
  - requirements of data efficiency in the transfer-to-file process.
- The settings to be applied in the transfer lines in order to get to this output format, respecting the information stored on the carriers as good as possible.

In broad terms, VIAA initially chose to follow these paths:

a) In case the data is not corrupted too much – the threshold to define this is to be defined:
   - Transfer from digital to digital, without entering the analog domain.
   - Working via the IEEE 1394 output[1] of the player.
   - Transfer to an MXF OP1a file with DV as a codec.

b) In case the data is corrupted too much – the threshold to define this is to be defined:
   - Playing the cassette back with a DV-player to obtain an uncompressed digital signal in which the player automatically applies its error correction mechanisms, then capturing and encoding this in a lossless format.
   - Working via the SDI output of the player.
   - Transfer to a FFV1/PCM in MKV file.

The main showstopper is the expected but undocumented technical heterogeneity of the cassettes, for example in aspect ratio, interlaced or progressive, long or short play mode, ... And maybe worst of all: technically one DV cassette could hold a mix of (many of) these parameters. This brings us to the following options:

1) **For the cassettes with a good signal:**
   a. **No normalisation: respect the technical heterogeneity of the signal on the cassettes:** the variety in the technical properties of the cassettes is mirrored in the variety of the technical properties of the files. This is achieved by transferring all tapes as they are via the IEEE 1394 output.
   b. **Absolute normalisation: homogenise the technical heterogeneity of the signal on the cassettes:** the variety in the technical properties of the cassettes is brought back to absolute homogeneity in the technical properties of the output files. To achieve this to the fullest, this is only possible via the SDI.

[1] Commonly known as FireWire or Sony i.link.

c. **Limited normalisation: a mix of the above:** the heterogeneity in the technical properties of the files is brought back to a limited number of 'flavours' of files. This is achieved by first transferring all tapes as they are via the IEEE 1394 output and then transcoding all files that do not comply with a limited number of 'flavours' to the flavour that is the easiest to reach.

2) **For the cassettes with a bad signal**, there's only one option: trough the transfer process via the SDI capture, the technical properties are normalised to one flavour of FFV1/PCM in MKV files.

The mission of the research task in this regard was established as:

1) Define good properties to which the output formats for all the DV-transfers could be homogenised as good as possible. These properties would ideally have the features that are common for as many as possible DV-cassettes in this project. Obviously, information loss is to be prevented as much as possible and it should respond to sustainability criteria (e.g. playability by common video playback and editing tools).
2) Define a good profile for the FFV1/MKV files resulting from the normalisation process (transcoding or SDI capture).

IV. Technical properties influencing the transfer workflow

Before considering possible workflows for the transfer of DV-based video cassettes, two partially interfering unpredictable factors were identified:

- **Unpredictable signal quality:** the number of broken bits and their consequences for the quality of the sound and image is unpredictable and certain dropout compensation mechanisms are only applied if the cassette is played back via the SDI output, in certain cases it might be interesting to keep a file transferred via SDI, next to the file as transferred via the IEEE 1394 output.
- **Unpredictable diversity of the significant technical properties of a stream or full cassette:** how homogenous (or heterogenous) are the technical specifications according to which the content has been recorded and which variations occur within the total collection and possibly even within one cassette or stream?

When both of these unpredictable factors are cleared up, two partially interfering choices are to be made in the workflow:

- **To normalize or to keep-as-is?** Whether the intra- and inter-cassette variations should be normalized (signal alteration but limiting the number of different file specifications in the archive) or kept as they are (no signal alteration but increasing the different file specifications in the archive dramatically).
- **To use the IEEE 1394 output or the SDI output?** Which one of the player's output should be used: both outputs have advantages and disadvantages.

4.1 Signal quality

A low signal quality on DV-based video cassettes can have several causes. The most common causes are:

- Dirty or sticky tapes
- Reading head clogging
- Reading head misalignment (azimuth problems)
- Tape demagnetisation causing bit errors

The first three causes can at least partially be eliminated, the fourth is irrevocable. For this, only the consequences can be limited.

4.1.1 Dirty tape, sticky tape, head clogging

Dirt on the tape, or even parts of the magnetic layer itself (sticky tape), might chip off and clog the player's reading head. The rate of 'head clog' problems depend on things like tape brand, its previous storage conditions, etc. A clogged head might issue different symptoms, not all of them clearly identifiable as 'head needs cleaning'. In any way, all the occurring symptoms are identical to 'tape data not readable' in one way or another:

- Behaves like 'undefined tape' (gaps). Real tape gaps appear between recordings. If the same behaviour occurs during a recording, it might be a clogged head.
- Visual drop outs.

If errors like these suddenly occur on tapes previously known as 'clean', or if the rate of these errors increases over time, it is advised to check the player's head. Physically cleaning the DV tapes (using a tape cleaning machine) has not shown sufficient improvements, but might cause additional damage to already 'worn out' tapes. Some machines even have a bad tape-feeding mechanism that breaks the cartridge (e.g. double-hatch).

It was also observed in previous DV ingest operations, that DV cleaning machines that claimed to be able to report about the tape's condition, merely printed out 'reports' with rather arbitrary 'Error' counts. Tests were performed in real world practice situations, analysing several tapes multiple times, showing completely stochastic numbers. So, practice seems to indicate that the physical tape cleaning step before ingest could be questioned whether it's really an improvement. This does not however, apply to tapes that obviously require cleaning before putting them into a player.

4.1.2 Azimuth errors

As any tape-based material, DV is also subject to azimuth errors. It is therefore necessary to have control over correct azimuth adjustments for proper transfer. If the original recording device had a misaligned azimuth angle, players with an actually correct azimuth will suffer from reading errors. In such a case, the player's azimuth must be 'misaligned' accordingly for these tapes (and adjusted back to its proper angle afterwards, of course).

4.1.3 Bit errors

If all causes above are excluded, it is still possible that at least some of the tapes will suffer from unrecoverable data errors due to tape demagnetisation. Although it somewhat depends on the tape brand and type and its storage conditions, even well stored cartridges may suffer from data errors. These errors in the DV-bitstream may cause different erratic behaviour when interpreting the data stream. Typical effects are drop-outs in the image and audio cracks. When transcoding these erroneous bitstreams however, other things may happen, like losing audio/video synchronicity after these error positions, as well as causing a transcoding application to prematurely exit or even crash. Of course, this depends on how the transcoding tool is implemented.

DV has error detection features built-in, which can be used for error concealment. If and how error concealment is applied, is up to the hardware / software interpreting the DV-stream though. When attempting to capture and preserve the original DV-stream (via IEEE 1394), it might be good to define a certain threshold of (sequential) errors in order to abort soon enough and switch to an alternative capture method because depending on what kind of data error, it might take an unratable amount of time trying to improve the situation.

Transfer via SDI is the only way in which a specific dropout compensation mechanism is applied. This dropout compensation mechanism is to be considered as a (automatic and immediately applied) form of restoration. From a heritage-theory perspective, it deserves recommendation to preserve the signal as unaltered as possible. However, the application of the automatic dropout compensation is probably the best image restoration method that is currently available. To create a reusable image, it would be a shame not to use it.

Notwithstanding the value of the dropout compensation when transferring via the SDI output, working via the IEEE 1394 output still has a few important advantages, also for cassettes of which the signal contains many errors:

- In theory, better algorithms for dropout compensation could be applied to the unaltered signal.
- Working via the IEEE 1394 output is the only way to recuperate certain data in the data stream such as the time code signal, the number of dropouts, etc. To do this, a tool like AVP's DV Analyzer can be used.

4.2 Signal diversity

Many significant technical properties of the signal on the cassettes are unknown, but they are expected to be rather variable. The following variations can occur:

- **Inter-cassette variations:** within the whole collection of cassettes to be transferred, there may be differences in significant technical properties between the cassettes.
- **Intra-cassette variations:** within one cassette to be transferred, there may be differences in significant technical properties. Here, three scenarios can apply:
  - **mid-tape variations:** on one tape, several streams are recorded, one or more of them with different significant technical properties.
  - **mid-stream variations:** within one stream, one or more significant technical properties change during a continuous recording.

The scheme below explains the different kinds of variations, each of them has consequences for the transfer workflow (cfr. 5.1):

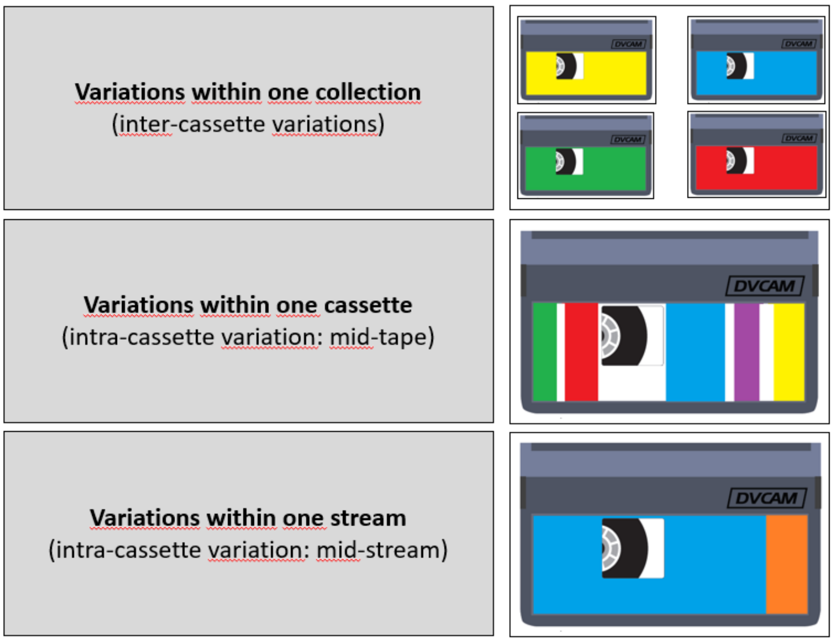

Fig. 2: types of variations within one collection, within one cassette and within one stream.

The most important properties are:

- For the image:
  - **Pixel resolution and frame rate**, both important properties of the television standard: for SD these are 760 x 576p at 25 fps (for PAL) and 760 x 480p at ~29,97 fps for NTSC.
  - **Chroma subsampling:** 4:2:0 (standard for DV and DVCAM in PAL), 4:1:1 (standard for DVCPRO25 and for DV and DVCAM in NTSC) and 4:2:2 (standard for DVCPRO50).
  - **Scan type:** interlaced or progressive.
  - **Display aspect ratio:** 4:3 or 16:9.
- For the sound:
  - **Frequency, bitrate, number of channels:** 32 kHz and 12 bit for 4 channels, or 48 kHz and 16 bit for 2 channels.

This leads to the following matrix of theoretically possible combinations of settings for one audiovisual stream:

Image					**Sound**			**Norm in this project for**
Pixel resolution	**Frame rate**	**Chroma sub-sampling**	**Scan type**	**Display Aspect Ratio**	**Frequency**	**Bitrate**	**Nr. of channels**	
720 x 576 px	25 fps	4:2:0	Interlaced	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	DV, DVCAM

			Progressive	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	
720 x 576 px	25 fps	4:1:1	Interlaced	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	DVCPRO 25
			Progressive	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	
720 x 576 px	25 fps	4:2:2	Interlaced	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	DVCPRO 50
			Progressive	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	
720 x 480 px	29,98 fps	4:1:1	Interlaced	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	
			Progressive	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	
720 x 480 px	29,98 fps	4:2:2	Interlaced	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	
			Progressive	16:9	32 kHz	12 bit	4	
				4:3	32 kHz	12 bit	4	
				16:9	48 kHz	16 bit	2	
				4:3	48 kHz	16 bit	2	

Fig. 3: all possible combinations of significant properties of the DV signal encoding per stream. The combinations expected to be most common in this project are indicated in green.

Based on the combinations above, the following normalizations can be considered.

4.2.1 General NTSC to PAL conversion

A general conversion from NTSC to PAL would affect:

- The pixel resolution: from 720 x 480 px (NTSC) to 720 x 576 px (PAL)
- The frame rate: from 30000/1001 fps (NTSC) to 25 (PAL)
- The chroma subsampling (only for DV and DVCAM, as DVCPRO25 is always 4:1:1 and DVCPRO50 is always 4:2:2): from 4:1:1 (NTSC) to 4:2:0 (PAL)

4.2.2 Normalizing display aspect ratio (DAR)

The display aspect ratio (DAR) of images stored on DV-based cassettes will most likely be 4:3 (both in PAL and in NTSC), with exceptionally also 16:9 (both in PAL and in NTSC). The storage aspect ratios (SAR), also referred to as horizontal x vertical resolution, is 720 x 576 for PAL and 720 x 480 for NTSC. For DV in PAL the pixel aspect ratio (PAR) is 5:4 and in NTSC the pixel aspect ratio is 6:4.

Since the Storage Aspect Ratio (SAR) is identical to the pixel dimensions (width x height) and therefore always defined in a video file, the Pixel Aspect Ratio (PAR) can be calculated if the DAR is known. The formula is as follows: 'PAR = DAR/SAR'. Therefore, it is only mandatory to store the DAR metadata within the resulting video file. Since the default for DV is to have a 5:4 (PAL) or 6:4 (NTSC) SAR, normalizing (= resizing) the pixel resolution is not necessary - even for anamorphic material - as it can be assumed that proper resizing, according to the DAR, is a default use-case and therefore well supported.

4.2.3 Normalizing chroma subsampling

For DV and DVCAM in PAL the chroma subsampling is normally 4:2:0. For DV and DVCAM in NTSC, and for DVCPRO 25 the chroma subsampling is normally 4:1:1. For DVCPRO 50 the chroma subsampling is 4:2:2. Via normalisation this diversity could be reduced to one of those mentioned. It should be acknowledged that subsampling normalisation is an irreversible interference in the signal, and 4:2:2 might be the preferred option in this case, because it is the highest quality of the three.

4.2.4 Normalizing audio resolution to 48kHz and 16bits

DV also allows to record 32 kHz / 12 bits audio, offering 2 additional channels. This is a rather uncommon audio resolution, whereas 48 kHz / 16 bits is very common and well supported across different domains (professional and consumer) and tools (hardware and software). This resolution also conforms to the SDI standard.

V. Workflow recommendations

As a general recommendation, considering the complexity of possible issues and to some extent unpredictable results of the transfer, it is recommended for the service provider to test their capture- and transcoding pipeline thoroughly on every scenario of signal quality and of signal diversity and submit samples of the results.

5.1 Solving signal quality issues

Considering the factors above, as soon as the dropout rate of a certain cassette ends up above a certain threshold, it deserves recommendation not to choose between the two transfer methods (via SDI or IEEE 1394), but **to keep both essence files**. The so-called 'parallel capture' method can be chosen to save time (cfr. 5.3). Taking into consideration only aspects of signal quality, a possible transfer workflow could then look as follows:

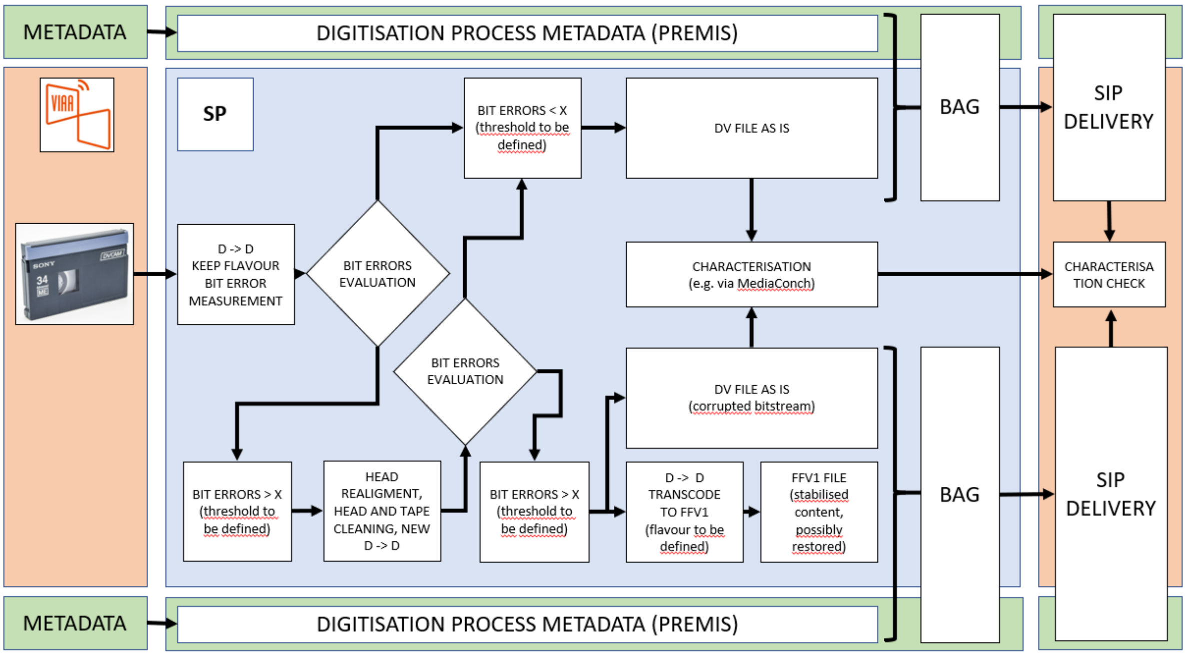

Fig. 4: possible workflow for DV-based cassette transfer,
taking into account factors of signal quality only.

This workflow has the obvious advantage that for cassettes with low error rates, not more essence than strictly needed (the capture via the IEEE 1394 output) has to be preserved. For the cassettes with higher error rates, an unaltered file stays available, the data that comes with the essence via the IEEE 1394 output can be preserved and there's an 'as good as possible' file from the SDI output available for reuse.

The exact threshold of the bit error rate can be agreed upon in collaboration with the service provider, based on a testing phase.

5.2 Solving signal diversity issues

As argued above, considering the diversity of significant technical properties of the audiovisual stream, the following variations may occur: inter-cassette variation and intra-cassette variation, with this last one subdivided into mid-tape variations and mid-stream variations. A special form of mid-tape variations are the undefined gaps on a tape. For each of these diversity issues, the paragraphs below discuss possible workflow consequences.

5.2.1 Inter-cassette variations

In order to be able to decide which tapes shall be captured natively (DV stream as-is over the IEEE 1394 output) or rather over SDI, it is good practice to separate tape collections according to their provenance or source. In this project it would most likely have to be done by the service provider.

For certain material sources, it can be assumed that they are more likely to be homogenous - meaning, they stick to a certain set of technical properties of their recordings - whereas other sources might be very heterogenous. For homogenous ('clean') collections, it is more likely that their native DV-streams can be recorded and kept as preservation master, whereas for heterogenous ('unclean') collections it might be significantly faster to ingest them over SDI.

When capturing a collection that is assumed to be 'clean', the ingest operator may still encounter situations that are out of the norm, such as:

- Significantly high data error rate
- Recordings that have tech-properties that don't conform to the defined 'norm' for its DV type.
- Other problems

In such cases, it might be good for the operator to be allowed to ingest these tapes over SDI, in order not to spend too much time trying to fix or deal with these situations.

5.2.2 Intra-cassette variations

5.2.2.1 Mid-tape variations

Recordings with different properties on the same tape could possibly cause issues during capture, depending on how the capture application deals with this case. Not all capture applications deal properly with mid-tape (and mid-stream) changes during capture. If they 'lock' onto the technical properties of the first recorded stream, unclear things may happen when individual recordings have different properties (codec, audio, etc).

Additionally, if these recordings are stored in a single video file, their behavior might be different and possibly erratic upon playback, transcoding or usage. Again, depending on which tools (hardware and software) are being used to work with these files. This behavior will stay with the file as long as it exists. The following options should be considered:

1) Possible to split the tape into each recording, allowing to maintain its properties. This would also create 'stable' files with a single set of technical properties across the whole file.
2) Normalize all recordings to a common set of tech-properties afterwards. This is non-trivial, as the actual behavior of the transcoding application must be checked beforehand, as it is often not well supported to deal with these kinds of mid-file property changes.
3) Record the tape over SDI, which can be considered as a different way of normalizing.

5.2.2.2 Mid-stream variations

This possible issue is similar to mid-tape variations (multiple recordings with different tech-properties on same tape), but not identical. Capture applications often lock on to the tech-properties initially present on the tape when the DV recording is started and store only those in the header of the video container.

For example, it is possible that the audio resolution was changed on-the-fly during the original recording. How mid-stream changes are dealt with greatly depends on the capture application, as well as the container format used for capture, because it is usually assumed that the technical properties do not change within one video file.

5.2.2.3 Mid-tape undefined gaps

Between individual recordings, the tape contains 'no' information. Depending on the capture application, as well as player- and transcoding applications, these 'undefined tape gaps' may cause e.g. audio/video synchronicity issues when dealing with the material. This applies when the DV-stream was captured directly 'as-is' and is then used or transcoded.

5.2.3 Properties to normalize or not to normalize

As mentioned above (cfr. 4.2 Signal diversity), the following significant characteristics could potentially be normalized. For all of these, we discuss whether this normalization is opportune from a heritage and / or workflow perspective.

It should be remembered that any modification to the video material on DV requires complete re-encoding of the original stream. There is no such thing as "lossless DV-to-DV conversion". Only if the target codec is lossless, an additional generation loss can be avoided.

5.2.3.1 General normalization of NTSC based streams to PAL

As described above (cfr. 4.2.1), a general normalization of NTSC to PAL would include:

- **pixel resolution compensation:** since NTSC has the same width but a smaller height, the missing 96 lines could be padded, adding 49 black lines at the top and at the bottom: 'letterboxing'. This would allow keeping the original resolution without interpolation. This is an acceptable alteration of the signal.
- **frame rate conversion:** about 5 frames will have to dropped every second to fit the ~29.97 fps of NTSC into 25 fps PAL. This is an irreversible alteration of the content.
- **normalizing chroma subsampling:** all 4:2:0, or 4:2:2? Changing the chroma subsampling will require interpolation of color values. It can be assumed that the visible impact will not be too severe, but should be avoided unless absolutely necessary. When capturing uncompressed DV over SDI, the subsampling will always be normalized to 4:2:2, regardless of the source material. In this case the color information of any DV source (except DVCPRO50) will be interpolated. This conversion is done by the player.

From a heritage perspective, if possible, keeping NTSC sources as-is is preferred for preservation. Especially because the impact of any conversion step is non-trivial, irreversible - and it can be assumed that most (if not all) applications dealing with audiovisual files will be able to handle NTSC properly.

It can further be assumed that the most likely use-case where this might be a problem is, if someone who is not experienced enough with digital video is trying to use an NTSC source material, mixed with PAL material in a PAL production. But in such a case, it is also very likely that they will manage to use the NTSC clip, but maybe their conversion/import step was not 'as good as it could have been'.

Normalizing NTSC to PAL has non-trivial and irreversible consequences and is therefore not recommended.

5.2.3.2 Normalization of the display aspect ratio

Cfr. 4.2.2 Normalizing display aspect ratio (DAR), p. 15.

Normalizing the display aspect ratio is not necessary is therefore not recommended.

5.2.3.3 Normalization of the chroma subsampling

Changing the chroma subsampling will require interpolation of color values. It can be assumed that the visible impact will not be too severe, but should be avoided unless absolutely necessary. When capturing uncompressed DV over SDI, the subsampling will always be normalized to 4:2:2, regardless of the source material. In this case the color information of any DV source (except DVCPRO50) will be interpolated. This conversion is done by the replayer.

Normalizing the chroma subsampling is an irreversible signal alteration with possible visible effects and is therefore not recommended.

5.2.3.4 Normalization of the audio resolution to 48 kHz, 16 bit

As mentioned, 48 kHz and 16bit is most common for DV cassettes, but 32 kHz and 12 bit is also possible. The normalization of these into 48 kHz and 16bit has the following advantages:

- One audio sample rate across all collections: all audio behaves the same.
- A common audio resolution vs a non-common one
- Less issues expected when working with the material

And disadvantages:

- Irreversible sample rate conversion. Since 48 is not a multiple of 32, sample interpolation needs to be done. The quality and effect of this step depends on the tool (hardware/software) being used - as well as the audio source itself.
- Dithering 12 to 16 bits: not only the number of samples, but also the samples itself must be modified in an irreversible way. The quality and effect of this step depends on the tool (hardware/software) being used - as well as the audio source itself.

If done properly, the artefacts introduced by this conversion step should be minimal to unnoticeable.

Since the audio is originally interleaved into the DV video signal, when capturing the original DV stream, it is possible to keep the original audio stream, while simultaneously writing the resampled (48 kHz / 16 bits) PCM channels to a separate audio track inside the recorded video container.

Most (if not all) applications will prefer the container's audio track over the interleaved audio inside the DV stream, therefore using the video files will behave like any other regular file. Yet, if for whatever reason, the original, unresampled audio is to be accessed, the video track can be unwrapped from the container and written to a native ".dv" file. When accessing this file, it will be a valid video file with its own audio track(s). This can easily be done using e.g. FFmpeg.

Normalizing the audio resolution is an irreversible signal alteration, but unnoticeable if done properly. It can be applied if necessary.

Provider requirements:

Digitization provider must test their capture- and transcoding pipeline and submit samples of the results. Depending on how exactly this quality test is to be done, different source audio materials should be used (e.g. classical music vs. interview speech recording).

As a possible reference to impact and differences of sample rate conversions, with different applications/algorithms, see: SRC Comparisons (96kHz to 44.1kHz) by "Infinite Wave Mastering Studio": http://src.infinitewave.ca/ (last update: 17th July 2018)

If the "additional audio track" option is chosen, one might check with a native DV capture from the provider, if that sample is written correctly, so the original DV video track can be extracted as described above.

5.3 Signal capture scenarios: IEEE 1394, SDI or both?

Looking at things from another perspective, one might ask: when is the IEEE 1394 output preferred and when the SDI output?

In principle, it is always and for all tapes recommended to use the IEEE 1394 output of the player to capture the stream and transfer it to file completely, including its DV-specific properties and original time code information.

As a possible fallback option where the original DV-stream has too many errors or where too many different recordings with changing properties (mid-stream-, mid-tape-changes) are expected (or present), capturing the uncompressed audiovisual stream over SDI is an option. The main advantages are:

- Different tech-properties are 'normalized' on the fly during playback
- Realtime error concealment, creating 'as good as possible' files for re-use
- No further issues with problematic bitstreams, as all bitstream-quirks will automatically be manifested as uncompressed-SDI essence during playback.

The main disadvantages are:

- If this is the only copy: loss of DV-specific information and the original timecode information
- The realtime error concealment is also a disadvantage from a heritage-theory perspective (see above)
- Image and audio quality depend on player's decoder

To remediate the first disadvantage, the following measures could be taken:

- To preserve the timecode via the SDI output: this depends on the player, as well as the capture application and target format being used. If a lossless/uncompressed codec is used for capturing the SDI signal, the process is equivalent to the following individual steps:

- Capture the original DV-Stream
- Transcode it to lossless/uncompressed, while normalizing its technical properties to:
  - a common resolution
  - a common framerate
  - apply error handling/concealment
  - determine gaps on the tape as an audio and video 'placeholder' (e.g. still image and mute audio)
  - convert chroma subsampling to 4:2:2

- To preserve the DV-specific information: use the above mentioned 'DV Analyzer'. To do this in only one tape-transfer step, a **parallel-capture setup** would be required.[2]

The **parallel-capture system and workflows** might be a bit more complex to set up, but this approach may save a great amount of time - as well as reducing the physical wear on the tapes. It provides the capture operator with two capture versions of one tape: the original DV-Stream (as-is) and the digitally decoded, normalized and error-corrected, etc. audio and video signal (over SDI).

The difference to just capturing the SDI signal is, that due to the availability of the DV stream capture, all technical metadata can be extracted and its information can be stored as preservation metadata, or applied to make certain decisions like re-capture or drop the SDI version because the DV stream is 'fine'.

In order to avoid hardware performance bottlenecks (which might lead to interstitial errors during the capture), it might be good (or necessary) to have two separate computers - each capturing only one signal: 1 DV, 1 SDI.

5.4 Documenting aspect ratio and field order information

Cfr. 4.2.2. Normalizing Display Aspect Ratio (p. 15)

Provider requirements:

The following technical metadata shall be properly stored in the resulting video files:
• Scan type (interlaced or progressive)
• Field order (if interlaced)
• Display- and Storage Aspect Ratio (DAR/SAR) information

[2] Some DV-players output the IEEE 1394 stream and the decoded, uncompressed SDI simultaneously. This can be used to capture both signals in one recording step, while requiring to play each tape only once (as compared to capture DV first and then SDI as fallback). This great idea is from Marion Jaks, video archivist at the Austrian Mediathek.

VI. Conclusion

6.1 Recommendations in normalisation and output use

This conclusion summarizes the recommendations considering normalization of the significant technical properties of the signal on the DV cassettes and advises on the choice for the IEEE 1394 output, the SDI output or both.

The recommendation is influenced **firstly by the assignment of signal quality** and **secondly by the signal diversity,** the reason being that a possible remediation on signal quality (by measures like reading head realignment, head and tape cleaning) can effectively improve the results of the signal diversity evaluation.

For tapes of which the signal remains below a certain threshold of bit errors, the IEEE 1394 output of the player should be used. For tapes of which the signal exceeds a certain threshold of bit errors, the SDI output should be used as a fallback, additional to the capture via the IEEE 1394 output. The height of this threshold should be determined in collaboration with the service provider during a testing phase.

Regarding differences in the significant technical properties of the signal, as occurring between cassettes (inter-cassette variation), within one tape (intra-cassette, mid-tape) or even within one stream (intra-cassette, mid-stream), the recommendation is to normalize only the audio resolution (frequency and bit depth). Normalizing the audio resolution is irreversible, but it's the only normalization does not constitute a significant alteration of the signal and at the same time that is tempering the effects of file format heterogeneity. Normalizations such as on the television signal (from NTSC to PAL) and the chroma subsampling are also irreversible, but they also hold the risk of significant alterations on the signal. Normalizing the display aspect ratio has no significant advantage in tempering the negative effects of file format heterogeneity.

This leaves us with the following possible output formats:

Image			Sound	
Pixel resolution, frames, chroma subsampling	**Scan type**	**Display Aspect Ratio**	**Original frequency, bitdepth, channels**	**Normalised frequency, bitdepth, channels**
720 x 576 px 25 fps 4:2:0	Interlaced	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
	Progressive	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	

			48 kHz, 16bit, 2 channels	48 kHz, 16bit, 4 channels
720 x 576 px 25 fps 4:1:1	Interlaced	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
	Progressive	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
720 x 576 px 25 fps 4:2:2	Interlaced	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
	Progressive	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
720 x 480 px 29,98 fps 4:1:1	Interlaced	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
	Progressive	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
720 x 480 px 29,98 fps 4:2:2	Interlaced	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
	Progressive	16:9	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	
		4:3	32 kHz, 12bit, 4 channels	48 kHz, 16bit, 4 channels
			48 kHz, 16bit, 2 channels	

Fig. 5: table with all possible output formats and their normalized sound specification.

6.2 General workflow proposal

The proposed workflow again is based on firstly the **signal quality evaluation** and secondly the **signal diversity evaluation.** For tapes of which the signal remains below a certain threshold of bit errors, the IEEE 1394 output of the player should be used. For tapes of which the signal exceeds a certain threshold of bit errors, first head realignment and reading head and even possibly tape cleaning should be tried. If these measures result in an improvement of the signal, again only the IEEE 1394 output should be used. However, if these measures do not result in an improvement of the signal, the SDI output should be used as a fallback, additional to the capture via the IEEE 1394 output. The height of this threshold should be determined in collaboration with the service provider during a testing phase.

For both kinds of tapes (below and above the bit error threshold), the transfer workflow via the IEEE 1394 output is determined by the **signal diversity evaluation**. This step may result in finding cassettes of three kinds:

- **Tapes with no mid-tape, nor mid-stream changes:** should be transferred according to the specifications of the stream(s). The only allowed normalization is in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalized to 48 kHz and the bit depth to 16 bit. This cassette will result in as many essence files (DV) as there are recordings on the cassette. For cassettes with a high number of bit errors this number of files is doubled to 2 essence files (DV + FFV1) per stream on the cassette. However, all files under the same codec should have the same specifications.
- **Tapes with mid-tape changes:** should be transferred according to the specifications of the streams. The only allowed normalization is in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalized to 48 kHz and the bit depth to 16 bit. This cassette will result in as many essence files (DV) as there are recordings on the cassette. For cassettes with a high number of bit errors this number of files is doubled to in 2 essence files (DV + FFV1) per stream on the cassette. At least two files of the same codec will have different specifications.
- **Tapes with mid-stream changes:** should be transferred according to the specifications of largest part of the stream. The only allowed normalization is in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalized to 48 kHz and the bit depth to 16 bit. keep it in one file. The service provider should check whether the player normalizes automatically. If not, normalization has to happen through an additional transfer via the SDI output. This cassette will result in as many essence files (DV) as there are recordings on the cassette. For cassettes with a high number of bit errors, or if the player doesn’t apply a correct automatic normalization, this number of files is doubled to 2 essence files (DV + FFV1) per stream on the cassette. How many different specifications will exist under one codec, depends on the presence of mid-tape changes on that same cassette.

One can conclude that each essence *stream* should be digitised according to its specifications at the time of recording. The only allowed normalization is in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalized to 48 kHz and the bit depth to 16 bit. Streams with mid-stream changes should always be transferred to short play mode. If a stream is part of a cassette with a large number of errors, also a normalized FFV1 version should be made via the SDI output.

This results in the following possible workflow scheme:

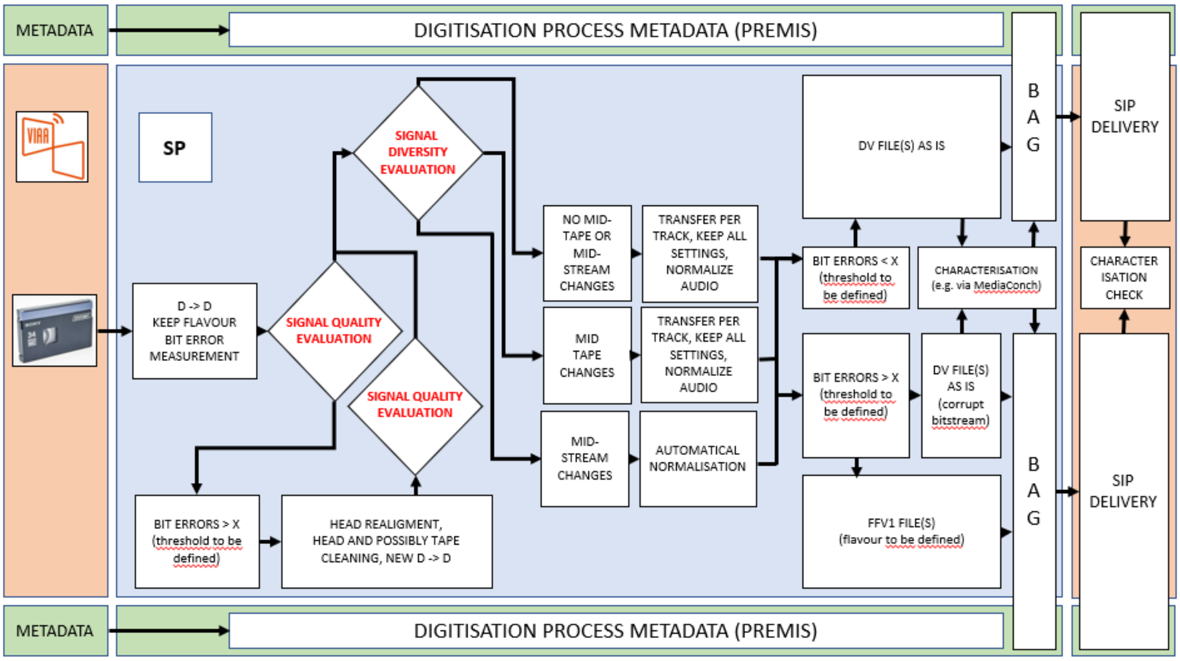

Fig. 6: proposed workflow scheme for DV-based video cassettes.

VII. Annex 1: norm per DV based cassette subtype

This part lists which technical properties are to be considered 'normal' or 'most common', depending on which DV-type (DV, HDV, DVCPro, etc).

7.1 DV

Video in PAL:

- Resolution: 720x576 px
- Subsampling: 4:2:0
- Scan type: Interlaced (BFF)
- Codec: DV
- Bits per channel: 8
- Display aspect ratio (DAR): 4:3 or 16:9

Audio:

- Codec: Uncompressed PCM
- Resolution: 48 kHz / 16 bits (linear)
- Channels: 2

7.2 HDV

Not expected in this collection.

7.3 DVCAM

Video in PAL:

- Resolution: 720x576 px
- Subsampling: 4:2:0
- Scan type: Interlaced (BFF)
- Codec: DV
- Bits per channel: 8
- Display aspect ratio (DAR): 4:3 or 16:9

Audio:

- Codec: Uncompressed PCM

- Resolution: 48 kHz / 16 bits (linear)
- Channels: always 2

7.4 DVCPRO

Video in PAL:

- Resolution: 720x576 px
- Subsampling: always 4:1:1
- Scan type: Interlaced (BFF)
- Codec: DV
- Bits per channel: 8
- Display aspect ratio (DAR): 4:3 or 16:9

Audio:

- Codec: Uncompressed PCM
- Resolution: 48 kHz / 16 bits (linear)
- Channels: always 2

7.5 DVCPRO50

Video in PAL:

- Resolution: 720x576 px
- Subsampling: 4:2:2
- Scan type: Interlaced (BFF)
- Codec: DV
- Bits per channel: 8
- Display aspect ratio (DAR): 4:3 or 16:9

Audio:

- Codec: Uncompressed PCM
- Resolution: 48 kHz / 16 bits (linear)
- Channels: always 4

VIII. Annex 2: format encoding policies for FFV1/PCM in MKV

This annex suggests a list of specifications for the FFV1/PCM in MKV files, resulting from the transfer via the SDI output.

General:

- Constant FrameRate (CFR)
- Colorspace: YUV
- Scan type defined (interlaced or progressive)
- Field order defined (top- or bottom-field first)
- Display Aspect Ratio (DAR) defined. Valid options: 4:3 and 16:9
- Valid subsampling options: 4:2:0 (PAL) – 4:1:1 (DVCPRO or NTSC) – 4:2:2 (DVCPRO 50)

Audio:

- Codec: Linear PCM (LPCM)
- Resolution: 48 kHz / 16 bits
- Channels: 2 or 4

FFV1 specific

- Version: FFV1.3
- GOP size: 1
- SliceCRC: enabled
- Slices: 24 (for SD), 64 (for HD)
- Coder: Ranger Coder
- Context: small

For FFmpeg encoding these are:

- -c:v ffv1
- -level 3
- -slices 24
- -slicecrc 1
- -coder 1
- -context 0

Matroska specific

- Matroska version >= 4
- SegmentUID: present
- SeekHead: present