To Normalise or To Manage the Multitude? Determining Workflows and Output Specifications for the Transfer of a Heterogeneous Collection of DV-based Video Cassettes

Brecht Declercq, VIAA, Belgium
Peter Bubestinger, AV-RD, Austria
Gaël Fernandez-Lorenzo, ECPAD, France
DOI: https://doi.org/10.35320/ij.v0i50.97

Abstract

This article discusses the preparation of workflows and output specifications for a transfer project of a collection of DV-based video cassettes (DV, DVCAM, DVCPRO), by VIAA, the Flemish Institute for Archiving. Two hard to predict characteristics of the collection determined the implementation of this project: the quality of the signal (the extent to which signal loss had occurred on the cassettes) and the precise technical characteristics of the recordings. Given these unknown factors, three major decisions had to be made for these cassettes, all of which together determined the work of an external transfer service provider: 1) the necessary steps in the workflow, 2) the output used of the player(s) and 3) the desired output format(s). In making these decisions, the confrontation between two good practices from the audiovisual archive world was crucial: on the one hand, keeping the technical characteristics of the source signal unchanged and, on the other hand, limiting the number of file formats of master files in the archive, in order to increase file format manageability. This article first outlines the project and its mission. Further on it discusses and arguments the technical choices made, culminating in a workflow proposal for the transfer of the cassettes, that brings together the advantages of two different approaches.

1. Project context and mission

The Flemish Institute for Archiving (VIAA) digitizes, stores and provides access to audiovisual material, photos, documents, etc., for so-called content partners from the cultural, heritage and media sectors in Flanders. 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 does not own a collection, but acts as a service provider to a growing group of currently 151 Flemish media and cultural organisations (content partners). These partners include broadcasters (national public and commercial as well as regional), cultural heritage institutions, governmental bodies, city archives and performing arts organisations.

The partners take part in VIAA's transfer projects, depending on the carrier formats they have in their collections. VIAA plans, coordinates and finances the transfer projects for them, by inventorying the collections, drafting the specifications for the transfer process, selecting a specialised transfer service provider and coordinating the project logistics of carriers and files. After the transfer, 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.

As a part of its mission to digitise the Flemish audiovisual heritage, VIAA prepared a state-of-the-art transfer-to-file project of DV-format based video cassettes in 2018-2019. The estimated volume of carriers is around 20.500, coming from all kinds of content partners, in total around 65 (cf. fig. 1).

CARRIER TYPE	DV (cassette)	¼" DVCAM (cassette)	¼" D7 DVCPRO 25 (cassette)	¼" D7 DVCPRO 50 (cassette)	¼" D7 DVCPRO HD (cassette)	TOTAL
CULTURAL HERITAGE	406	736	220	0	0	1362
BROADCASTERS	4321	9623	132	0	0	14076
CITY ARCHIVES	405	93	29	0	0	527
PERFORMING ARTS	4004	373	3	1	0	4381
GOVERNMENT BODIES	205	31	0	0	0	236
TOTAL	9341	10856	384	1	0	20582

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

One of the most important specifications for VIAA to determine was about a suited output format (container, codec and file format specifications), taking into account the properties of the digital signal on the cassettes, requirements of digital sustainability of the file formats, codecs and specifications, requirements of data efficiency in the transfer-to-file process and last but not least the requirement to deliver also a *showable* result, i.e. images that were ready to use on VIAA's access platforms.

2. Technical properties influencing the transfer workflow

Before considering file formats, codecs and specifications (and their subsequent workflows steps) 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[1].
- **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 were cleared up, two partially interfering choices are to be made in the workflow:

- **To normalise or to keep-as-is?** Whether the intra- and inter-cassette variations should be normalised (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).

1 Commonly known as FireWire or Sony i.link.

- **To use the IEEE 1394 output (data transfer) or the SDI output (signal transfer)?** Which one of the player's output should be used: both outputs have advantages and disadvantages.

2.1 Signal quality

In order to mediate the unpredictability of both factors they were analysed further. The first one, a low signal quality on DV-based video cassettes, can have several causes. The most common causes of a bad signal quality or problems occurring during their read-out are stickiness or dirt on the tapes themselves, tape demagnetisation causing bit errors, reading head clogging and reading head misalignment. All of these can at least partially be eliminated, except for the second one. For this one the effects can only be limited.

2.1.1. Dirty tape, sticky tape and 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:

- Undefined tape gaps, which are in fact blank signal levels: real tape gaps appear between recordings, but 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.

2.1.2. Reading head misalignment errors

As any tape-based material, DV-based cassette formats are also subject to reading head misalignment errors. It is therefore necessary to apply a correct alignment of the heads onto the tape path to do a proper transfer. If the original recording device had a misaligned reading head angle, players with an actually correct reading head angle will show an increased number of reading errors. In such a case, the player's reading head angle must be 'misaligned' accordingly for these tapes (and adjusted back to its proper angle afterwards, of course).

2.1.3. Bit errors due to tape demagnetisation

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 this might depend 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 captured via the IEEE 1394 output however, other things may happen, like losing audio/video synchronicity after these error positions, as well as causing an application that transcodes the DV file later on in the process 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 as a prerequisite for triggering notification or 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 the IEEE 1394 data output), 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 unrateable 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 restoration-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 data 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.

The risk for a crash of the capturing application exists, but it is related to the extent that the application interprets the binary data coming from the tape. If it writes the received bites to a file as they are, error manifestations may only happen when a later process (e.g. of transcoding) tries to interpret the data. In practice many tapes can be transferred without a crash of the application and only when trying to transcode or play a position where the DV-bitstream was malformed, problems occur. This even counts for intra-cassette format changes. If the capture application just dumps the stream as it is, it keeps the intra-cassette format changes. Only when interpreting the data problems on how to deal with these changes might come up. It might be possible though to automatically split the file upon detected format changes in the data.

2.2 Signal diversity

In this project, 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):

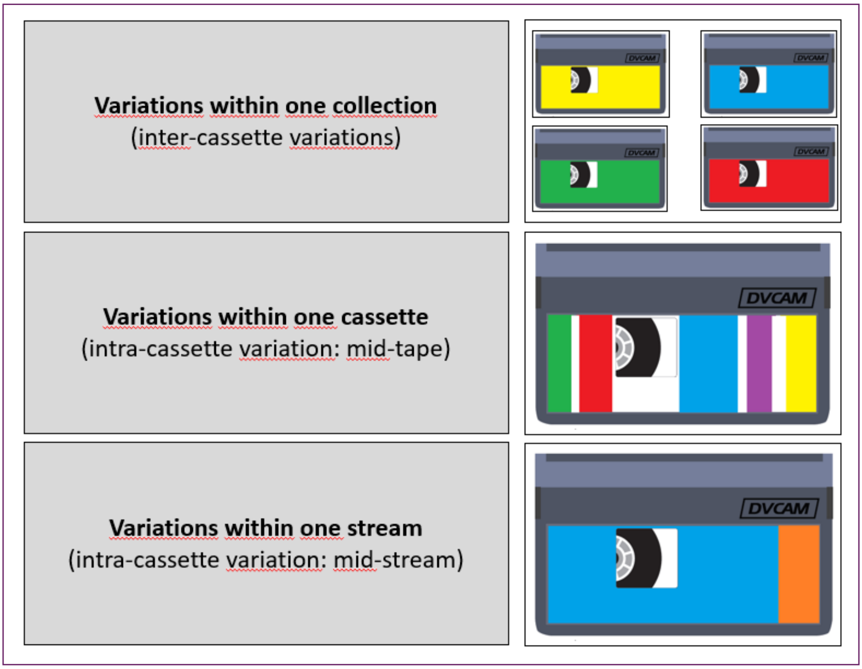

Fig. 2: types of variations within one collection, within one cassette and within one stream.

We identified the most important properties for the images as:

- **Pixel resolution and frame rate**, both important properties of the television standard: for SD these are 720 x 576px at 25 fps (for PAL) and 720 x 480px at ca. 29,97 fps for NTSC.
- **Chroma subsampling:** 4:2:0 (standard for DV and DVCAM in PAL), 4:1:1 (standard for DVCPRO 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.

Image					
Format	**Pixel resolution**	**Frame rate**	**Chroma sub-sampling**	**Scan type**	**Display Aspect Ratio**
PAL	720 x 576px	25 fps	4:2:0 (PAL only)	**Interlaced**	**4:3**
NTSC	720 x 480px	ca. 29.97 fps	4:1:1 4:2:2	Progressive	16:9

Fig. 3: most important image properties in the collection to be transferred. The combination expected to occur most is indicated in purple.

For the sound, the following most important properties were identified:

- **Frequency, bitrate, number of channels:** 32 kHz and 12 bit for 4 channels, or 48 kHz and 16 bit for 2 channels.

Sound		
Sample rate	**Bit depth**	**Channels**
32 kHz	12 bit	4
48 kHz	16 bit	2

Fig. 4: most important sound properties in the collection to be transferred. The combination expected to occur most is indicated in purple.

3. Normalisations to be considered

The combination of the most important variable properties in sound and image – further on called significant technical properties – lead us to a total of **40 theoretically possible combinations** of settings for one audiovisual stream. The combinations expected to be most common in this project are indicated in green in the two tables above, but all the others may possibly occur. This is the point where preservation theory seems to contradict itself: on the one hand we'd like to preserve unaltered signals as much as possible, on the other hand we'd like to keep the number of master file formats in a digital format not larger than necessary. In order to be able to judge whether the heterogeneity in technical characteristics may be limited, we must consider whether a possible normalisation to one common characteristic would entail changes in the image and sound, and whether these changes, if any, are acceptable.

3.1 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 from 4:1:1 (NTSC) to 4:2:0 (PAL)[2]

3.1.1. Normalising display aspect ratio (DAR)

The display aspect ratio (DAR) of images stored on DV-based cassettes will – in this project – 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, normalising (i.c. 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.

3.2 Normalising 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.

3.3 Normalising audio resolution to 48kHz and 16bits

DV also allows to record 32 kHz / 12 bits audio, offering 2 additional channels, but this is a rather uncommon audio resolution. If the choice is made to normalise 32 kHz / 12 bits (one common audio resolution for a mixed collection), 48 kHz / 16 bits may be a good option, since that is the more common resolution for the audio in DV files, it is well supported across different domains (professional and consumer) and tools (hardware and software) and it also happens to match the SDI resolution. Obviously any change made (e.g. resample audio) should be documented properly.

4. Workflow recommendations

After identifying the most important variable video and audio properties and how they could be normalised, we have considered whether these normalisations were opportune, considering explicitly the possible alterations of the video and audio signal they would cause. The conclusions of these considerations are listed below. As a more general recom-

2 This only applies to DV and DVCAM, as DVCPRO is always 4:1:1 and DVCPRO50 is always 4:2:2.

mendation, considering the complexity of possible issues and to some extent unpredictable results of the transfer, we concluded that testing the capture- and transcoding pipeline thoroughly on every scenario of signal quality and of signal diversity is crucial.

4.1 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 version created with the IEEE 1394 output is the archival master file, while the one created via the SDI output can be considered as a normalised, restored copy. This so-called 'parallel capture' method can be chosen to save time (cfr. 4.3).

Taking into consideration only aspects of signal quality, a possible transfer workflow could then look as follows:

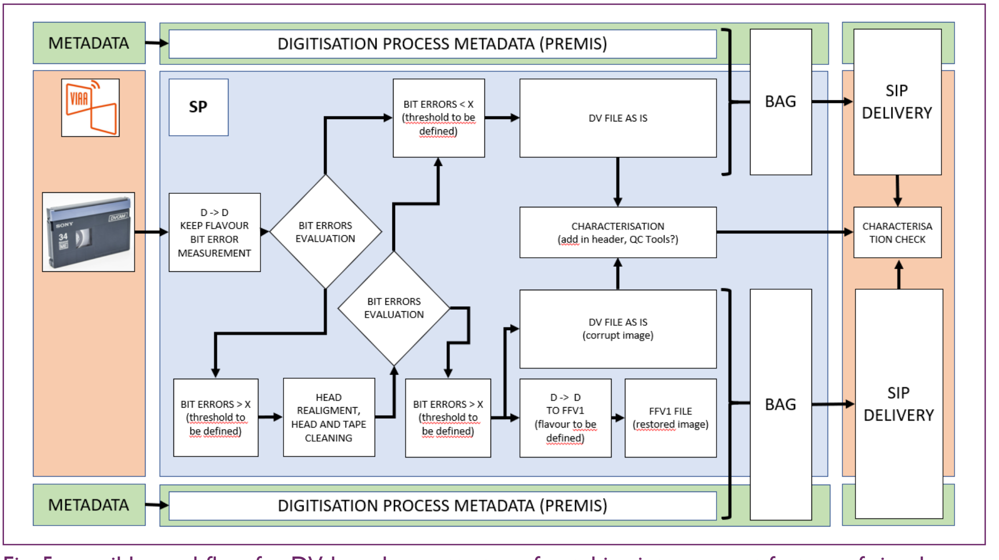

Fig. 5: possible workflow for DV-based cassette transfer, taking into account factors of signal quality only.

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 as a fallback option.

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

4.2 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.

4.2.1. 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 heterogeneous. For homogenous ('clean') collections, it is more likely that their native DV-streams can be recorded and kept as preservation master, whereas for heterogeneous ('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:

- A 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.

4.2.2. Intra-cassette variations

4.2.2.1. 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 behaviour 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 behaviour will stay with the file as long as it exists. The following options should be considered:

1. 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. To normalise all recordings to a common set of tech-properties afterwards. This is non-trivial, as the actual behaviour 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 normalising.

4.2.2.2. 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.

4.2.2.3. 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.

4.2.3. Properties to normalise or not to normalise

As mentioned above (cfr. 2.2 Signal diversity), the following properties could potentially be normalised. For all of these, we discuss whether this normalisation 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 or uncompressed, an additional generation loss can be avoided.

4.2.3.1. General normalisation of NTSC based streams to PAL

As described above (cfr. 2.2.1), a general normalisation 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.
- **normalising chroma subsampling:** all 4:2:0 or 4:2:2? See paragraph 4.2.3.3 for details.

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'.

4.2.3.2. Normalisation of the display aspect ratio

As explained under 3.2.2, normalising the display aspect ratio (DAR), is not necessary and therefore not recommended.

4.2.3.3. Normalisation of the chroma subsampling

Changing the chroma subsampling will require interpolation of colour 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 normalised to 4:2:2, regardless of the source material. In this case the colour information of any DV source (except DVCPRO50) will be interpolated. This conversion is done by the player. This means that normalising the chroma subsampling is an irreversible signal alteration with possible visible effects and is therefore not recommended.

4.2.3.4. Normalisation of the audio resolution to 48 kHz, 16 bit

As mentioned, 48 kHz and 16 bits is most common for DV cassettes, but 32 kHz and 12 bits is also possible. The normalisation 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

However, this normalisation also has a few disadvantages:

- This normalisation means an irreversible sample rate conversion. Since 48 is not a multiple of 32, sample interpolation needs to be done (whereas resampling from e.g. 96 to 48kHz is done, every 2nd sample is simply dropped). 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, as an optional part of the normalisation: not only the number of samples, but also the samples itself must be modified in an irreversible way: the 12-bit values would need to be stretched to their 16-bit representation as they're not just zero-padded with 4 bits. 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.

Normalising the audio resolution is an irreversible signal alteration, but unnoticeable if done properly. It can be applied if necessary in order to reduce the technical diversity in the archive, but as mentioned earlier, it is optional.

4.3 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 'normalised' 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 applied when transferring via the SDI is also a disadvantage from a heritage-theory perspective (see above), as an error concealment is automatically applied whereas when the data stream is captured directly as-is via the IEEE 1394 output, it is possible to display the images with unconcealed dropouts.
- Image and audio quality may also depend on player.

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 or 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 normalising 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, the above mentioned 'DV Analyzer' could be used. To do this in only one tape-transfer step, a 'parallel-capture' setup would be required.[3]

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, normalised 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.

4.4 Documenting aspect ratio and field order information

Whatever the method chosen, the following technical metadata shall be properly stored in the resulting video files - this means in a machine-readable, standard-conform and supported way (e.g. in the container) so that not only the information is preserved, but also that a player can automatically read and interpret them so that the image is displayed correctly.

As these three tech-properties are very basic and common, it can be expected that any modern video-container, and application, will support handling them by default (cfr. 3.2.2. Normalising Display Aspect Ratio).

- Scan type (interlaced or progressive)
- Field order (if interlaced)
- Display- and Storage Aspect Ratio (DAR/SAR) information

3 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.

5. Conclusion

5.1 Recommendations in normalisation and output use

This conclusion summarizes the recommendations considering normalisation of the most important 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 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 normalise only the audio resolution (frequency and bit depth). Normalising the audio resolution is irreversible, but it's the only normalisation that does not constitute a significant alteration of the signal and at the same time tempers the effects of file format heterogeneity. Normalisations 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. Normalising the display aspect ratio has no significant advantage in tempering the negative effects of file format heterogeneity. Considering the levelled definitions of 'significant properties' given by Grace and Montague (2008), all the considered technical properties would have significance level 10 (essential and unchanged), whereas the audio resolution (frequency and sample rate) could be considered level 07 to level 09 (essential – some variation allowed).

This leaves us with the following possible output formats:

Image			Sound	
Pixel resolution, frames, chroma subsampling	**Scan type**	**Display Aspect Ratio**	**Original samplerate, 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, 4 channels
			48 kHz, 16bit, 2 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	

Image			Sound	
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. 6: table with all possible output formats and their normalised sound specification.

5.2 General workflow proposal

Based on the signal quality evaluation and on the signal diversity evaluation we have developed the following proposal for a transfer workflow, in which the diversity of technical properties of the files is limited.

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 normalisation is in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalised 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 (one file via the IEEE 1394 output[4] and one file via the SDI output[5]) 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 normalisation is in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalised 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 (one file via the IEEE 1394 output and one file via the SDI output) 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 normalisation is in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalised to 48 kHz and the bit depth to 16 bit. keep it in one file. The service provider should check whether the player normalises automatically. If not, normalisation 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 normalisation, this number of files is doubled to 2 essence files (one file via the IEEE 1394 output and one file via the SDI output) 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 normalisations are:

- in the sound domain: if recorded in 32 kHz, 12 bit, the frequency should be normalised to 48 kHz and the bit depth to 16 bit.
- in case of mid-stream changes, in order to keep the stream in his entirety, the transfer should follow the specifications of largest part of the stream. 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 normalised FFV1 version should be made via the SDI output.

4 VIAA has chosen here to go for an unwrapped .dv file.
5 VIAA has chosen here to go for an FFV1 file with LPCM sound, wrapped in an MKV container.

This results in the following possible workflow scheme:

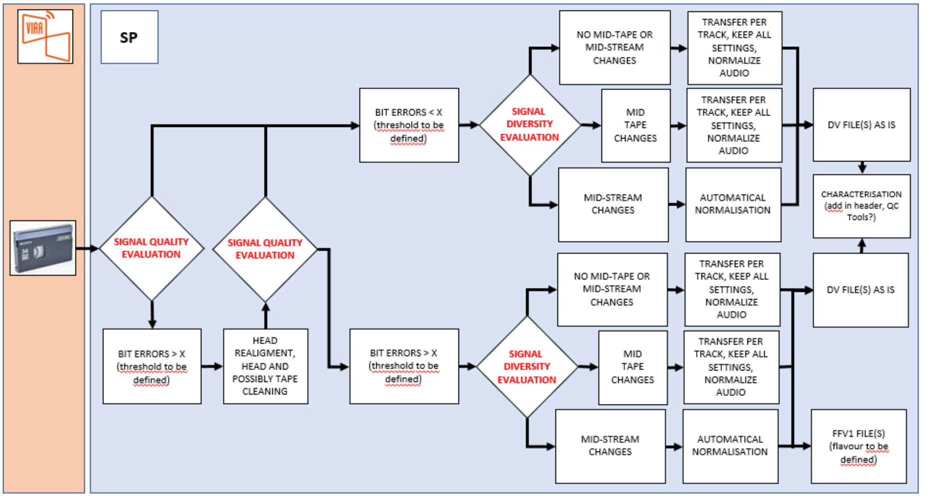

Fig. 7: proposed workflow scheme for the transfer of DV-based video cassettes.

Whether this workflow scheme will work in practice has to be considered further. Especially determining a bit error rate threshold value might prove to be hard, as this not only has to take into account an average bit error rate, but also the occurrence of peak values and possibly even at which point in the stream these peak values occur. Also the behaviour of the playback machines used will be crucial: discarding, blanking or interpolation of bit errors might cause audio/video synchronicity issues. Because of this, the file resulting from the transfer via the SDI output may be retained more often than anticipated together with the file resulting from the transfer via the IEEE 1394 output, to serve as a displayable and reusable video file.

5.3 Overall conclusion

In this article we have considered the preparation of workflows and output specifications for a transfer project of a collection of DV-based video cassettes (DV, DVCAM, DVCPRO), by VIAA, the Flemish Institute for Archiving. The quality of the signal (the extent to which signal loss had occurred on the cassettes) and the precise technical characteristics of the recordings in this collection are unknown and hard to predict and this has appeared to be a circumstance with far-reaching consequences when taking the most important decisions in the design of a transfer workflow.

We have therefore studied how the quality of the signal could be respected as much as possible, whether and how the diversity in technical properties of the signal could be accommodated without violating essential preservation principles and – subsequently – which outputs of the player should be used to achieve these goals. In making these decisions, the confrontation between two good practices from the audiovisual archive world was crucial: on the one hand, keeping the technical characteristics of the source signal unchanged and, on the other hand, limiting the number of file formats of master files in the archive, in order to increase file format manageability.

The study of the technical properties has shown that their normalisation always implies an unacceptable change of the signal. All these properties can therefore be labelled as significant properties. The only technical feature for which a normalisation is allowed as an option to reduce the technical diversity of the files resulting from the transfer project, is the normalisation of the audio resolution (frequency and bitrate).

This argument - together with the fact that the data stream may contain useful data that can only be read via the IEEE 1394 output - leads to the recommendation to certainly use this IEEE 1394 output for the transfer. However, in order to simultaneously include the benefits of the automatic error concealment that is only possible with a transfer via the SDI output, such a transfer via the SDI can also be considered. The file that results from the transfer via the IEEE 1394 output must be considered as the master file, whereas the file that results from the transfer via the SDI must be considered as a normalized, restored copy.

This article therefore recognizes the benefits of both approaches and refuses to choose only one of the two methods. By assigning the status of archive master to the files from the transfer via the IEEE 1394 output, we give clear priority to that method, but at the same time we do not neglect the importance of the files from the SDI output.

6. References

Grace, S. and Montague, L. (2008) Framework for the definition of significant properties. InSPECT Project Document. London, National Archives, https://www.kdl.kcl.ac.uk/fileadmin/documents/digifutures/materials/preservation/DF09_prsrv_knight-definingSigProperties.pdf, last retrieved June 10th, 2019.

7. Annex I: norm per DV based cassette subtype

This part lists which technical properties are to be considered *normal* or *most common* in this collection, depending on which DV-type (DV, DVCPro, etc). HDV is not expected in this collection.

DV Norm	**Image**					**Sound**		
	Pixel reso-lution	**Bits per com-ponent**	**Sub-sampling**	**Scan type**	**Display Aspect Ratio**	**Samplerate**	**Bit depth**	**Channels**
DV	PAL SD (720 x 576 px)	8 bpc	4:2:0	Interlaced (BFF)	4:3 16:9	32 kHz 48 kHz	12 bit 16 bit	4 2
DVCAM			4:2:0			48 kHz	16 bit	=2*
DVCPRO			4:1:1					=2*
DVCPRO50			4:2:2					=4*

* For DVCAM, DVCPRO and DVCPRO50 the number of channels is fixed.

8. 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	Audio	Matroska	FFV1
Constant Framerate (CFR)	Codec: Uncompressed, Linear PCM (LPCM)	Version >= 4	Version: FFV1.3
Colorspace: YUV	Resolution: 48 kHz / 16 bits	SegmentUID: present	GOP size: 1
Scan type: must be defined. Interlaced or Progressive	Channels: 2 or 4	SeekHead: present	SliceCRC: enabled
Field order: must be defined. Top- or Bottom-Field-First			Slices: 24(for SD), 64 (for HD)
DAR: must be defined. Valid options: 4:3 or 16:9			Coder: Range Coder
Valid subsampling options: 4:2:0 (PAL) 4:1:1 (DVCPRO or NTSC) 4:2:2 (DVCPRO50)			Context: small