

# DVD-Audio Projects

## Description of Format

DVD-A is a special variant of the main DVD specifications as laid out by the DVD Forum to cater specifically for high quality audio playback as its main focus. Additionally, it is recommended that all DVD-A projects also include a DVD-V file set to maintain compatibility with non audio capable players. This basically means that these discs will play on all existing DVD players in one form or another. One of the biggest problems with DVD is the sheer number of different players that are available, and the purpose of this document is to try & ensure the least amount of problems during production. For ease of use, we shall refer to the fileset for DVD-A discs as the **Audio\_TS** and the fileset for DVD-V compatibility zones as the **Video\_TS**.

## Preparation of Assets.

### 1-1 - Audio Files.

All Audio files should be supplied as uncompressed data files to the following specifications:

**Audio\_TS - Surround Sound** can be from 16/44.1 to 24/96 in resolution and will be compressed to MLP Lossless (1).

**Audio\_TS - Stereo** can be from 16/44.1 to 24/192 in resolution and will be compressed to MLP Lossless (2).

**Video\_TS - Surround Sound** can be from 16/48 to 24/96 in resolution (3) and will be compressed to either Dolby Digital or DTS (4).

**Video\_TS - Stereo** can be from 16/48 to 24/96 in resolution and can be left as LPCM or compressed to Dolby Digital (5).

Surround files should be phase coherent mono WAV or AIFF type files, with their channel assignment clearly marked in the filename by either number or direct indication i.e., "Surround\_Mix\_Main\_Album\_01.WAV" or "SurroundMix\_MainAlbum\_L.WAV" - please note that only alphanumeric characters should be used in filenames and underscores in place of spaces - there should be no gaps in the name at all as a general rule.

notes:

1 - MLP Lossless encoding is mandatory for streams that have a data rate exceeding 9.6Mbps, which gives a practical limit of 24/96 at 4.0 for uncompressed bitstreams. However, there is no good reason \*not\* to use MLP Lossless even on streams that fall below the threshold as the output of the process is bit-for-bit identical to the input - meaning a quality loss of zero (MLP is a lossless technology) and in addition there is a further benefit of saving space on the disc. Finally, MLP has one big advantage over LPCM with its built-in internal error protection, something lacking in LPCM systems. If it is decoding, then you can be assured that the output delivered is correct.

2 - Whilst MLP is not mandatory for any stereo stream, its use is recommended regardless for the reasons given above.

3 - The DTS system allows for encoding from a source file resolution of up to 24/96 by use of a core bitstream plus extensions. This works by taking the 24/96 source, and encoding the **core audio** at 24/48, then using an extension in the bitstream to allow for 24/96 source files. At a DTS decoder, the highest quality is always selected automatically by the decoder - if the decoder can recognize the 96KHz extension, it will decode it for you with no additional settings needed. If, on the other hand, it does not recognize the extension, it will instead decode the core stream at 24/48 resolution.

In a similar manner, a mix intended for DTS encoding can also be up to 6.1 surround, with the centre rear channel being either discrete or matrixed into Ls/Rs - all the complexity is again in the encoder, and all decoding is transparent to the user (earlier decoders might require the DTS-ES flag to be set manually, but failure to do this will not adversely affect playback as the 5.1 core will always be present no matter what.

4 - Dolby Digital & DTS are vastly different systems. Whilst all DVD players are required to be capable of decoding a Dolby Digital stream, the ability to decode a DTS bitstream is optional in Video-Only players. This is important because we have to make a decision here - in DVD-Audio, when a Video\_TS is imported for Video-Only player compatibility **each VTS can only have a maximum of 2 audio streams**. This means that we cannot create a Video\_TS with 3 Audio streams containing stereo, Dolby Digital surround and DTS surround. If all 3 of these options are required, we must use 2 VTS titles instead of 1. In most cases, this will not be a problem unless space on the disc is really tight. We recommend - emphatically - that wherever possible, a dedicated stereo mix should

be used in both Audio\_TS as well as the Video\_TS simply because the only alternative to this would be a Dolby Digital stream with a downmix as the stereo option. This, in our opinion, is a very bad idea for a number of reasons, the main ones being as follows:

A - A downmix is always a compromise, as you are collapsing 5 channels into 2. This means the chances are very high that the mix will simply not be as good as it otherwise should be because the options in Dolby Digital streams are very limited. You can reduce the levels of the centre channel by -3, -4.5 or -6dB, and you can reduce the levels of the surround channels by the same amount albeit independently of the centre channels. This does not give the producer a great deal of control.

B - A downmixed Dolby Digital file simply cannot sound as good as a dedicated stereo mix will, and by removing the uncompressed stereo option for Video-Only players and forcing them to listen to a massively compressed downmix you are effectively telling them that their choice of playback is unimportant to you. This could well lose you vital sales.

However, in our experience all surround music fans are generally capable of decoding a DTS bitstream with no difficulty at all.

For this reason, we recommend using a single VTS per title with LPCM stereo and 5.1 DTS audio options. If a Dolby Digital stream is important to your project for any reason at all, we would recommend 2 VTS per title, with VTS1 holding the surround streams and VTS2 holding the stereo streams. (A DTS stream, being optional, cannot be the sole stream in a VTS. Therefore it is better to keep the surround streams together, with the DD stream as audio stream 1 & the DTS as audio stream 2 in VTS1, and to have the LPCM on it's own as the sole stream in VTS2)

5 - In Video-Only players, the DVD specification allows for stereo LPCM Audio at resolutions from 16/48 up to 24/96.

However - only 16/48 is actually mandated, meaning that no Video-Only player is actually required to decode or pass a bitstream above this resolution. Players that cannot decode these streams will either truncate the stream down to 16-bits, or else resample it down to 48KHz, or both. We have seen players that will pass a 16/96 file, but truncate 24/96 to 16/96 and also players that will pass 24/48 but resample 24/96 to 24/48 on the fly.

## 1-2 - Encoding of Audio files

We would prefer to carry out all encoding to MLP Lossless, DTS or Dolby Digital in house. We only use licensed & approved encoders to ensure all bitstreams meet required standards for guaranteed spec compliance & logo usage on packaging etc.

Specifically, we use Steinberg's Nuendo Dolby Digital Encoder, the DTS-HD MAS system and SurCode's MLP Lossless. (Verification logs can be provided for MLP Lossless as well as DTS encodings if required).

## 2-1 - Stills and Menu Images

All menu designs and stills should be created and edited to the following specifications, taking careful note of the required TV standard and Aspect Ratio (hereafter referred to as the **PAR**), as different rules apply for 16:9 widescreen compared to 4:3 full-frame. Also, please note that all screens are divided into different segments - the background, the SubPicture Highlight (hereafter referred to as the **SPHL**) and (optionally) the button layer. The differences are basically that the background image is the full-colour image, the SPHL is the shape of the button and the button layer is the "Hot-Spot" for it. Why do we need a Hot Spot layer? simply, because in the DVD Specification a button cannot overlap another button and must also be a rectangular shape - although the actual button image itself can be any shape you like it must be "bound" within its own rectangle.

## 2-2 - Design Parameters :

### Background Images

PAL 4:3	Design & Edit at 720 x 540, Square Pixel. Full colour TIFF file, no colour management, 96 DPI (1) Safe Areas are bound between points at Left Top (72,54) and Bottom Right (648,486)
NTSC 4:3	Design & Edit at 720 x 540, Square Pixel, Full colour TIFF file, no colour management, 96 DPI (2) Safe Areas are bound between points at Left Top (72,54) and Bottom Right (648,486)
PAL 16:9	Design & Edit at 1024 x 576, Square Pixel, Full Colour TIFF file, no colour management, 96 DPI (3) Safe Areas are bound between points at Left Top (103,58) and Bottom Right (922,518)
NTSC 16:9	Design & Edit at 854 x 480, Square Pixel, Full colour TIFF file, no colour management, 96 DPI (4) Safe Areas are bound between points at Left Top (85,48) and Bottom Right (769,432)

## **SPHL Images.**

As above in all details except - and this is vitally important - these images cannot be full colour. They **must** be indexed to map a **maximum** of 4 colours (5)

## **Button Hotspot Images**

As for SPHL images, but this time indexed to just 2 colours, Black & White, where the Black rectangle defines the boundaries of the button hotspot area. These images are generally used for an SPHL where automatic generation of the hotspot image might cause confusion - IE, where the actual button shape is to be text or a complex shape comprising two or more components (IE something like % or |>, and automatic creation runs the risk of generating 2 buttons where only one is wanted)

**NB** - these are strictly optional files, although highly recommended (mainly for the reason given above, but also because their usage removes any ambiguities about exactly where each button's boundaries are located)

## Notes

1 - It is perfectly permissible to create & design in non-square pixels in the correct final display resolution. This would be at 720 x 576 with a PAR of 1.07 at 72 DPI

2 - It is perfectly permissible to create & design in non-square pixels in the correct final display resolution. This would be at 720 x 480 with a PAR of 0.9 at 72 DPI

3 - It is perfectly permissible to create & design in non-square pixels in the correct final display resolution. This would be at 720 x 576 with a PAR of 1.422 at 72 DPI

4 - It is perfectly permissible to create & design in non-square pixels in the correct final display resolution. This would be at 720 x 480 with a PAR of 1.2 at 72 DPI

5 - this is a limitation of the DVD specifications. What you effectively have to work with is a background colour (usually white) leaving you with 3 colours only for the actual button colours. These are traditionally mapped as Black (0,0,0), Red (255,0,0) and Blue (0,0,255) where Black defines the border, Blue defines the foreground and Red defines the caption. In the real world, for an example where the button image is to be the word "TEXT" in an oval on a box, Black would map the box, Blue would map the oval & Red would map the text. This is known as the "**Sonic Mode**" for mapping. Just to confuse matters, there is an alternative known as "**MEI Mode**" where the mapping is Black for the background (instead of White), Blue defines the border (instead of Black), Green defines the caption (instead of Red) and Red defines the foreground (instead of Blue).

As a rule of thumb, we would prefer it if SONIC MODE is used instead as it is far more commonly recognised by other applications making the same images suitable for Video\_TS usage as well as the Audio\_TS.

## **2-3 - Definition and Detail considerations.**

Computer monitors normally have much higher resolution than video displays and use progressive video rather than interlaced, which means that details that look great on a computer screen may not look the same on a TV display - or worse still may disappear altogether. This is especially important when dealing with text. The following guidelines can help maintain clarity & legibility when viewing the results on a TV system.

- Fonts should be at least 11-12 points, with the thinnest parts of the letters at least 2 points thick to avoid moiré\* and flickering.
- Sans Serif fonts display better than Serif fonts.
- Avoid putting more than 10 to 15 lines of text on a single screen.
- In Line Art, use lines that are at least 2 pixels wide
- When creating multi-screen menus try not to exceed more than 20 images per track in total. Segues need to be taken into account here too, as the results may be unpredictable in terms of image quality.

\* - Moiré is a kind of wavy, watery effect - almost a shimmering feel.

## **3-1 - Video Files**

Video files should be supplied as either Type 2 AVI or MOV files with the codec used specified. The higher the resolution of the source files, the better the final quality will be on the DVD. Progressive Scan (**PS**) Quicktimes (Animation Codec) are preferred, although PS Lagarith or HuffYUV2\_11 AVI files are also fine. As a bare minimum, DV AVI or DV MOV files should be used. We can also accept High Definition source files, although these would of course be downscaled and MPEG-2 encoded. Whilst it is possible to mix Aspect Ratios on a Video title, it is not recommended to do this in DVD-A/V as HDMI fed TV sets will not be able to pick up & honour the correct Aspect Ratio in the vast majority of cases (this is a peculiarity of HDMI though, as analogue systems

(SCART, Composite, Component etc) all honour the flag for the aspect ratio as authored).

One additional peculiarity specific to motion menus is that the total frame count of any menu video file should be an exact multiple of its GOP structure. Therefore, as a general rule of thumb, use a GOP of 15 frames when creating MPEG-2 files, or better still simply provide us with a file that has a frame count divisible by 15 & we will do the rest.

Frame rates should be NTSC 29.97DF or 30NDF for all video files unless a PAL version is also required. If this is the case, then separate PAL PS & NTSC PS assets should be supplied, or else PAL PS only and we will perform the standards conversion.

### **3-2 - Video File resolutions**

PAL 4:3            720x576 at 1.07 PAR, or 768x576 Square Pixel, Progressive Scan.

PAL 16:9          720x576 at 1.422 PAR, or 1024x576 Square Pixel, Progressive Scan.

NTSC 4:3          720x480 at 0.9 PAR, or 640x480 Square Pixel, Progressive Scan

NTSC 16:9        720x480 at 1.2 PAR Or 854x480 Square Pixel, Progressive Scan.

If clarification or any further help or discussion is needed, please do not hesitate to ask.