iTunes Video and Audio Asset Guide 5.2

 Developer

Contents

Contents

Overview

This document provides detailed delivery information for all accepted media and files for the iTunes Store, including music, music video, television, and movies. If further details are required, contact your iTunes Technical Representative.

Introduction

Quality is important to us at iTunes. We expect to receive the highest-quality assets available. Our product must meet or exceed the quality of the physical product already out in the marketplace. For example, if 5.1 surround sound or closed captions exist on the physical version of the product, those must be provided. If the physical product gives the chapters actual names (as opposed to Chapter 1, Chapter 2, and so on), then our product should have those same chapter titles. If the album is in stereo, stereo audio must be provided.

Changes Made in this Release

Date/Version	Changes Made
April 9, 2014 - Version 5.2	Added closed caption requirements for music videos. Added a best practice for MFiT content. Added requirements for TV cover art. Added guidelines to TV content considerations. Clarified SCC files for both TV and Film. Changed the version number of this specification from 5.1 to 5.2 to keep the version number in sync with the new schema version.

For a complete history of changes, see "Previous Spec Revisions" (page 33).

What's New in the iTunes Video and Audio Asset Guide 5.2?

Music Video: Closed Captions Profile

Closed captions are supported for music videos and requirements have been added.

Mastered for iTunes: Best Practice

Added a best practice for audible clipping. See "Best Practices for MFiT Content" (page 7).

TV: Cover Art Profile

Two new requirements have been added to the TV cover art profile. The cover art must not contain nudity, graphic material, or promotional material, including URLs or bugs.

TV: Closed Captions Profile

Clarified the requirements for SCC files.

- The timecode framerate can only be 29.97. This framerate is independent from your video source framerate.
- The timecode format must match the framerate format of the source video, either drop frame or non-drop frame.

See "TV Closed Captioning Profile" (page 21) for a further explanation.

TV: Content Considerations

Two new guidelines have been added to the TV content considerations section. Previews should be unique for each episode on a season and should not contain any spoilers.

Film: Closed Captions Profile

Clarified the requirements for SCC files.

- The timecode framerate can only be 29.97. This framerate is independent from your video source framerate.
- The timecode format must match the framerate format of the source video, either drop frame or non-drop frame.

See "Film Closed Captioning Profile" (page 28) for a further explanation.

Schema Version

The version number of this guide has changed from 5.1 to 5.2 to keep the version number in sync with the new schema version.

Music Audio Content Profiles

Music Audio Source Profile

The iTunes Store accepts audio with a sampling rate of 44.1Khz and 16-bit or 24-bit resolution and 96Khz with 24-bit resolution. Note that if stereo audio source exists, it must be used. See below for audio requirements specific to MFiT content.

Uncompressed audio formats supported are:

Format	Container Type	Qualified CODEC
Pulse-Code Modulation (PCM)	WAV (`.wav`)	
Apple Lossless (ALAC)	M4A (`.m4a`)	QuickTime http://www.apple.com/quicktime iTunes http://www.apple.com/itunes
	CAF (`.caf`)	iTunes Producer
Free Lossless Audio Codec (FLAC)	FLAC (`.flac`)	FLAC http://flac.sourceforge.net

All other audio formats will be rejected.

> **Important:** All audio must be generated using a CODEC qualified and approved by Apple.

Mastered For iTunes (MFiT) Source Profile

The audio for MFiT source must follow these requirements to be badged as MFiT on the Store:

- Audio must be delivered at 24-bit resolution in an approved format.
- Acceptable sample rates are 44.1, 48, 88.2, and 96 kHz.

Best Practices for MFiT Content

The following lists best practices for producing MFiT content. Apple recommends communicating these best practices to your mastering house:

- Source format must have been minimum 24-bit with a minimum sample rate of 44.1 kHz. (Up-sampling and/or bit-padding of 44.1 kHz/16-bit files is not allowed.)

- All masters must have been auditioned as encoded by the current Apple AAC encoder either with the "Master for iTunes Droplet", "RoundTripAAC" plug-in, or the Sonnox "Pro-Codec V2" plug-in that includes Apple's "iTunes +" AAC CODEC. Use these tools to set an appropriate level so that the encode doesn't show clipping.

- Although iTunes doesn't reject masters for specific numbers of clips, audible clipping caused by excessive levels to the encoder may be reason for tracks to not be badged and marketed as "Mastered for iTunes."

- The format of the masters must be 24-bit PCM at a sample rate of 44.1, 48, 88.2, or 96 kHz. Native resolution of project preferred. (ALAC or FLAC lossless compression is acceptable.)

Ringtone Source Profile

- Sampling rate of 44.1 kHz and 16-bit or 24-bit resolution and 96Khz with 24-bit resolution

- Must be lossless

- WAV, FLAC, or ALAC format

- Minimum length is 5 seconds and the maximum length is 30 seconds

See the table above for the uncompressed audio formats that are supported. All other audio formats will be rejected. Note that if stereo audio source exists, it must be used.

> **Important:** All audio must be generated using a CODEC qualified and approved by Apple.

Music Album Cover Art Profile

- JPEG with `.jpg` extension (quality unconstrained) or PNG with `.png` extension

- Color space: RGB (screen standard)

- For albums, minimum size of 1400 x 1400 pixels. 2400 x 2400 pixels recommended for best results

- For ringtones, minimum size of 800 x 800 pixels. 1400 x 1400 pixels recommended for best results

- Images must be square

- File formats: JPEG or PNG (100% quality)

- 1:1 aspect ratio

Do not increase the size of a smaller image to meet the minimum size standard. Excessively blurry or pixelated images will be rejected.

> **Important:** CMYK (print standard) images will **not** be accepted.

Music Digital Booklet Profile

- PDF format with `.pdf` extension

- Four-page minimum

- No more than 10 MB in size

- All fonts embedded

- 11 in x 8.264 in (28 cm x 21 cm)

- RGB color

- Horizontal presentation

- All images full-bleed as shown in sample pages

> **Important:** These booklets are expressly designed for the iTunes Store format, and cannot be reproductions of the liner notes with borders to increase their size.

Content Considerations

- When saving as PDF, make sure the document opens full screen with no negative space surrounding the document.

- If the digital booklet is many pages, consider using fewer images or optimizing images to achieve lower overall file size.

- Printer's marks are not allowed.

- You cannot sell or advertise other products or services. No other promotional sites are allowed.

- No links to anything outside of the booklet, except to the artist and/or label website(s).

- No time-sensitive information (for example, a promotion or dates for an upcoming tour or concert).

Jack Johnson, In Between Dreams

Dave Matthews Band, Stand Up

U2, The Complete U2

Music Video Content Profiles

Note: Chaptering is not supported for music videos.

Music Video SD Source Profile

NTSC

- MPEG-2 Program Stream Main Profile

- 4:2:0 chroma sampling

- ITU-R BT.601 color space

- 15 Mbps minimum

- Long GOP

- 640 fixed horizontal dimension

- Variable size vertical dimension depending on aspect ratio of source, maximum size of 480

- Square pixel aspect ratio (1:1)

- Native frame rate of original source:

  - 29.97 interlaced frames per second video source can be delivered either interlaced or de-interlaced properly tagged as progressive

  - 24 frames per second must be delivered progressive

  - 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail

  - Field dominance must be properly tagged (top field first, bottom field first, or progressive)

  - Telecine materials will not be accepted

  - For mixed frame rate material please contact your iTunes Technical Representative

- Interlaced content must be tagged non-progressive and field ordering must be defined in the stream.

- Crop inactive pixels and maintain fields. All edges must have active pixels for greater than 90% of the duration of the video.

PAL

- MPEG-2 Program Stream Main Profile

- 4:2:0 chroma sampling

- ITU-R BT.601 color space

- 15 Mbps minimum

- Long GOP

- 640 fixed horizontal dimension

- Variable size vertical dimension depending on aspect ratio of source, maximum size of 480

- Square pixel aspect ratio (1:1)

- Native frame rate of original source:

    - 25 interlaced frames per second sourced from video must be delivered de-interlaced and properly tagged as progressive

    - 24 and 25 frames per second sourced from film must be delivered progressive

    - 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail

    - Field dominance must be properly tagged (top field first, bottom field first, or progressive)

    - Telecine or interlaced materials will not be accepted

    - For mixed frame rate material, contact your iTunes Technical Representative

- Crop inactive pixels. All edges must have active pixels for greater than 90% of the duration of the video.

> **Important:** All video must begin and end with at least one black frame. In addition, videos can only have empty edits in the last edit of the edit list; videos with empty edits other than the last edit will be blocked.

Music Video HD Source Profile

- Apple ProRes 422 (HQ)

- VBR expected at ~220 Mbps

- 1920 x 1080 square pixel aspect ratio material

- Native frame rate of original source:

    - 29.97 interlaced frames per second for video sourced

    - 24 or 25 progressive frames per second for film sourced

    - 23.976 progressive frames for inverse telecine sourced from film

- Telecine materials will not be accepted
- HD source may be delivered matted: letterbox, pillarbox, or windowbox.
    - The HD source may be delivered in its full-frame state with metadata included to specify the crop rectangle. See "Music Video Single" in the *iTunes Package Music Specification* for details.
    - If the HD source file is not delivered matted or if there are no inactive pixels, we recommend setting all crop dimension attributes to '0' (zero).

> **Important:** All video must begin and end with at least one black frame. In addition, videos that begin with or contain empty edits will be blocked; the file can contain an empty edit in its edit list only if it is the last edit.

Music Video Audio Source Profile

If 5.1 Surround is available for a music video audio source, the audio should be delivered in 5.1 Surround in addition to providing a stereo version; otherwise the audio may be delivered in Stereo only.

Surround

- LPCM in either Big Endian or Little Endian, 16-bit or 24-bit, at least 48kHz
- Expected channels: L, R, C, LFE, Ls, Rs

Stereo

- MPEG-1 layer II stereo
- 384 kpbs
- 48Khz
- Included in the same file as the delivered video

Music Video Audio/Video Container

- Deliver all content in an MPEG-2 Program Stream file container
- The `.mpg` file extension is expected for all MPEG-2 content
- Audio must be delivered muxed with the video stream

> **Note:** Closed-captioning is currently not supported for music videos.

Music Video Closed Captioning Profile

> **Note:** Closed captioning can be sent with ProRes and MPEG-2 files.

- English text in EIA 608 format.

- Delivered in the same package with the video it references.

- In a Scenarist SCC formatted file, using `.scc` file extension.

- The timecode framerate can only be 29.97. This framerate is independent from your video source framerate.

- The timecode format must match the framerate format of the source video, either drop frame (DF) or non-drop frame (NDF).
    - Drop frame format has colons for the first two time delimiters and a semi-colon for the last time delimiter (HH:MM:SS;FF)
    - Non-drop frame format has colons for all the time delimiters (HH:MM:SS:FF)

Source Video Framerate	Description	Timecode Format	Timecode Example
29.97, 59.94i	NTSC Video	DF	HH:MM:SS;FF
25, 50i	PAL Video	NDF	HH:MM:SS:FF
24	Film	NDF	HH:MM:SS:FF
23.976	NTSC Film	NDF	HH:MM:SS:FF

- SCC files must be validated for proper sync against the associated video file using QuickTime 7 Pro.

- Captions should display and synchronize to within one second of the initial, audible dialog to be represented in text.

The timecodes of the captions are relative to the start of the program, and not the QuickTime movie's timecode track.

Currently, the iTunes Store does not support EIA 708 (ATSC closed captioning) or Teletext.

MacCaption is a tool you can use to create `.scc` files: http://www.cpcweb.com/products/. (Note that this product is not endorsed by Apple. Apple cannot and does not provide support for third-party products.)

Note: If closed caption data is available for any broadcast or web delivery system, it must be suppled to iTunes.

Music Video Screen Capture Image Profile

- Screen capture from delivered video

- JPEG with `.jpg` extension (quality unconstrained) or PNG with `.png` extension

- RGB (screen standard)

- Minimum dimensions: Must be 640 x 100 pixels.

- Variable size vertical dimension. Must be same aspect ratio as video source, with a maximum size of 480.

- Only the active pixel area may be included.

Do not increase the size of a smaller image to meet the minimum size standard. Excessively blurry or pixelated images will be rejected. Images must be taken directly from the video.

Important: CMYK (print standard) images will not be accepted.

Television Content Profiles

HD TV Source Profile

- Apple ProRes 422 (HQ)

- ITU-R BT.709 color space, file tagged correctly as 709

- VBR expected at 88-220 Mbps

- 1920 x 1080 or 1280 x 720 square pixel aspect ratio material*

- 23.976, 24, 25, 29.97 frame rates are supported

- Native frame rate of original source:

    - 29.97 interlaced frames per second video source can be delivered either interlaced or de-interlaced properly tagged as progressive

    - 24 and 25 frames per second must be delivered progressive

    - 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail

    - Field dominance must be properly tagged (top field first, bottom field first, or progressive)

    - Fields and frames may not be duplicated or eliminated to create a broadcast frame rate (for example, telecine, NTSC to PAL conversion)

    - For mixed frame rate material please contact your iTunes Technical Representative

- Interlaced content must be correctly tagged as interlaced and field ordering must be defined in the QuickTime container.

- Crop dimensions should be supplied in the metadata for content with inactive pixels due to letterbox, pillarbox, or windowbox. Refer to the *iTunes Package TV Specification* for further information.

- **Content upscaled from SD will be rejected.**

* If your mezzanine library is not stored in HD D5 or HDCam-SR, contact your iTunes Technical Representative.

SD TV Source Profile

NTSC

Apple ProRes:

- Apple ProRes 422 (HQ)

- VBR expected at 40-60 Mbps

- 720 x 480 or 720 x 486 encoded pixels; for display at either 853 x 480 for 16:9 content or 640 x 480 for 4:3 content. Properly created 720 x 486 content will have a minimum of 4 pixels of black at the top and 2 at the bottom. Crop values for top and bottom must total at least 6 pixels.

- All encoded content must include pixel aspect ratio (pasp) that defines content as either 4:3 or 16:9.

- Native frame rate of original source:

    - 29.97 frames per second video source can be delivered interlaced.

    - 24 frames per second must be delivered progressive.

    - 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail.

    - Telecine materials will not be accepted.

- Crop dimensions should be supplied in the metadata for content with inactive pixels due to letterbox, pillarbox, or windowbox. Refer to the *iTunes Package TV Specification* for further information.

- 4:3 standard definition video should not be delivered anamorphic 16:9 with matting.

MPEG-2:

- MPEG-2 Program Stream Main Profile

- 4:2:0 chroma sampling

- ITU-R BT.601 color space

- 15 Mbps minimum

- Long GOP

- 640 fixed horizontal dimension

- Variable size vertical dimension depending on aspect ratio of source, maximum size of 480

- Square pixel aspect ratio (1:1)

- Native frame rate of original source:
  - 29.97 interlaced frames per second video source can be delivered either interlaced or de-interlaced properly tagged as progressive
  - 24 frames per second must be delivered progressive
  - 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail
  - Field dominance must be properly tagged (top field first, bottom field first, or progressive)
  - Fields and frames may not be duplicated or eliminated to create a broadcast frame rate (for example, telecine, NTSC to PAL conversion)
  - For mixed frame rate material please contact your iTunes Technical Representative
- Interlaced content must be tagged non-progressive and field ordering must be defined in the stream.
- Crop inactive pixels and maintain fields. All edges must have active pixels for greater than 90% of the duration of the video.
- Content may NOT be delivered letterbox, pillarbox, or windowbox.

PAL

Apple ProRes:

- Apple ProRes 422 (HQ)
- VBR expected at 40-60 Mbps
- 720 x 576 encoded pixels; for display at either 1024 x 576 for 16:9 content or 768 x 576 for 4:3 content.
- All encoded content must include pixel aspect ratio (pasp) that defines content as either 4:3 or 16:9.
- Native frame rate of original source:
  - 25 frames per second video source can be delivered interlaced or de-interlaced and properly tagged as progressive.
  - Telecine materials will not be accepted.
- Crop dimensions should be supplied in the metadata for content with inactive pixels due to letterbox, pillarbox, or windowbox. Refer to the *iTunes Package TV Specification* for further information.
- 4:3 standard definition video should not be delivered anamorphic 16:9 with matting.

MPEG-2:

- MPEG-2 Program Stream Main Profile

- 4:2:0 chroma sampling

- ITU-R BT.601 color space

- 15 Mbps minimum

- Long GOP

- 640 fixed horizontal dimension

- Variable size vertical dimension depending on aspect ratio of source, maximum size of 480

- Square pixel aspect ratio (1:1)

- Native frame rate of original source:

    - 25 interlaced frames per second sourced from video must be delivered de-interlaced and properly tagged as progressive

    - 24 and 25 frames per second sourced from film must be delivered progressive

    - 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail

    - Field dominance must be properly tagged (top field first, bottom field first, or progressive)

    - Interlaced materials will not be accepted

    - Fields and frames may not be duplicated or eliminated to create a broadcast frame rate (for example, telecine, NTSC to PAL conversion)

    - For mixed frame rate material please contact your iTunes Technical Representative

- Crop inactive pixels. All edges must have active pixels for greater than 90% of the duration of the video.

- Content may NOT be delivered letterbox, pillarbox, or windowbox.

> **Important:** All video must begin and end with at least one black frame. In addition, videos can only have empty edits in the last edit of the edit list; videos with empty edits other than the last edit will be blocked.

TV Audio Source Profile

MPEG-2 Program Stream Container

Stereo

- MPEG-1 layer II

- 384 kpbs

- 48Khz

- Included in the same file as the delivered video

QuickTime Container

Surround

- LPCM in either Big Endian or Little Endian, 16-bit or 24-bit, at least 48kHz

- Expected channels: L, R, C, LFE, Ls, Rs

Stereo

- LPCM in either Big Endian or Little Endian, 16-bit or 24-bit, at least 48kHz

- Expected Dolby Pro Logic channels: Lt, Rt or expected stereo channels: L, R

TV Audio/Video Container

MPEG-2 Program Stream Container

- Deliver all content in an MPEG-2 Program Stream file container.

- The `.mpg` file extension is expected for all MPEG-2 content.

- Audio must be delivered muxed with the video stream.

QuickTime Container

- Deliver all content in a QuickTime `.mov` file container.

- The QuickTime `.mov` file extension is expected for all audio and video content.

- Each audio channel must have an assignment. The channel assignments must match one of the options below. For Option 1 (one track with all six channels), the order of the channel assignments can vary as noted in Option 1a, 1b, 1c, and 1d. Note that "Lt" and "Rt" are only used for Dolby matrix audio mixdown.

5.1 Surround Tracks						Stereo Tracks	
L	R	C	LFE	Ls	Rs	Lt	Rt

Option 1a	Track 1 -- six channels	Track 2	Track 3
	L R C LFE Ls Rs	Lt	Rt

Option 1b	Track 1 -- six channels	Track 2	Track 3
	L R Ls Rs C LFE	Lt	Rt

Option 1c	Track 1 -- six channels	Track 2	Track 3
	L C R Ls Rs LFE	Lt	Rt

Option 1d	Track 1 -- six channels	Track 2	Track 3
	C L R Ls Rs LFE	Lt	Rt

One track containing all Surround channels; Matrix Stereo with Lt in one track and Rt channel in another track

Option 2	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7	Track 8
	L	R	C	LFE	Ls	Rs	Lt	Rt

One track for each channel

Option 3	Track 1 -- six channels	Track 2 -- two channels
	L R C LFE Ls Rs	Lt Rt

One track containing all Surround channels; Matrix Stereo with Lt and Rt channels in one track

Option 4	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7 -- two channels
	L	R	C	LFE	Ls	Rs	Lt Rt

One track for each Surround channel; Matrix Stereo with Lt and Rt channels in one track

Option 5		Track 1	Track 2
		L	R

No Surround. Stereo with L in one track and R in another track

Option 6		Track 1 -- two channels
		L R

No Surround. Stereo with both L and R channels in one track

> **Important:** Refer to "Table 1: Audio Channel Assignment Labels" (page 41) for label descriptions and "How to Apply Audio Channel Assignments" (page 35) for instructions on applying audio channel assignments.

TV Closed Captioning Profile

Note: Closed captioning can be sent with ProRes and MPEG-2 files.

- English text in EIA 608 format.

- Delivered in the same package with the video it references.

- In a Scenarist SCC formatted file, using `.scc` file extension.

- The timecode framerate can only be 29.97. This framerate is independent from your video source framerate.

- The timecode format must match the framerate format of the source video, either drop frame (DF) or non-drop frame (NDF).
    - Drop frame format has colons for the first two time delimiters and a semi-colon for the last time delimiter (HH:MM:SS;FF)
    - Non-drop frame format has colons for all the time delimiters (HH:MM:SS:FF)

Source Video Framerate	Description	Timecode Format	Timecode Example
29.97, 59.94i	NTSC Video	DF	HH:MM:SS;FF
25, 50i	PAL Video	NDF	HH:MM:SS:FF
24	Film	NDF	HH:MM:SS:FF
23.976	NTSC Film	NDF	HH:MM:SS:FF

- SCC files must be validated for proper sync against the associated video file using QuickTime 7 Pro.

- Captions should display and synchronize to within one second of the initial, audible dialog to be represented in text.

The timecodes of the captions are relative to the start of the program, and not the QuickTime movie's timecode track.

Currently, the iTunes Store does not support EIA 708 (ATSC closed captioning) or Teletext.

MacCaption is a tool you can use to create `.scc` files: http://www.cpcweb.com/products/. (Note that this product is not endorsed by Apple. Apple cannot and does not provide support for third-party products.)

Note: If closed caption data is available for any broadcast or web delivery system, it must be suppled to iTunes.

TV Cover Art Profile

- JPEG with `.jpg` extension (quality unconstrained)

- RGB (screen standard)

- 1400 x 1400 pixels minimum size (2400 x 2400 recommended for best results)

- 72 dpi minimum resolution

- 1:1 aspect ratio

- No nudity or graphic material

- No promotional material, including URLs or bugs

Do not increase the size of a smaller image to meet the minimum size standard. Excessively blurry or pixelated images will be rejected.

Important: CMYK (print standard) images will not be accepted.

TV Content Considerations

- No bugs or logos should be visible during the body of the video.

- No tune-ins should be visible during the body of the video. Tune-ins are only acceptable at the end of the video.

- No ratings or advisories should be displayed at any time during the video.

- Network cards at the beginning and end of the video are accepted as long as they are visible less than five (5) seconds.

- Commercials or other promotional material, including URLs, are NOT accepted. For more details, please contact your iTunes Technical Representative.

- Commercial black may be a maximum of 5 seconds.

- Previews must contain content suitable for a general audience.

- Previews must not have opening or ending credits and should not start on a black frame.

- Previews should be unique for each episode on a season.

- Previews should not contain any spoilers.

- A minimum of 1 black frame at the beginning and end of each video is required.

Film Content Profiles

Film HD Source Profile

- Apple ProRes 422 (HQ)

- ITU-R BT.709 color space, file tagged correctly as 709

- VBR expected at ~220 Mbps

- 1920 x 1080 square pixel aspect ratio material

- Native frame rate of original source:

    - 29.97 interlaced frames per second for video sourced

    - 24 or 25 progressive frames per second for film sourced

    - 23.976 progressive frames for inverse telecine sourced from film

    - Telecine materials will not be accepted

- Content may be delivered matted: letterbox, pillarbox, or windowbox.

> **Important:** All video must begin and end with at least one black frame. In addition, videos can only have empty edits in the last edit of the edit list; videos with empty edits other than the last edit will be blocked.

Film SD Source Profile

NTSC

- Apple ProRes 422 (HQ)

- VBR expected at 40-60 Mbps

- 720 x 480 or 720 x 486 encoded pixels; for display at either 853 x 480 for 16:9 content or 640 x 480 for 4:3 content. Properly created 720 x 486 content will have a minimum of 4 pixels of black at the top and 2 at the bottom. Crop values for top and bottom must total at least 6 pixels.

- All encoded content must include pixel aspect ratio (pasp) that defines content as either 4:3 or 16:9.

- Native frame rate of original source:

    - 29.97 frames per second video source can be delivered interlaced

- 24 frames per second must be delivered progressive

- 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail

- Telecine materials will not be accepted

- Content may be delivered matted: letterbox, pillarbox, or windowbox.

- 4:3 standard definition video should not be delivered anamorphic 16:9 with matting. For example, the minimum post-cropped display must equal or exceed 768 x 576.

PAL

- Apple ProRes 422 (HQ)

- VBR expected at 40-60 Mbps

- 720 x 576 encoded pixels; for display at either 1024 x 576 for 16:9 content or 768 x 576 for 4:3 content

- All encoded content must include pixel aspect ratio (pasp) that defines content as either 4:3 or 16:9.

- Native frame rate of original source:

    - 24 and 25 frames per second sourced from film must be delivered progressive

    - 23.976 frames per second for inverse telecine must be delivered progressive; must not be delivered interlaced or delivery will fail

    - Telecine materials will not be accepted

- Content may be delivered matted: letterbox, pillarbox, or windowbox.

- 4:3 standard definition video should not be delivered anamorphic 16:9 with matting. For example, the minimum post-cropped display must equal or exceed 768 x 576.

25 fps interlaced PAL films are NOT supported.

Important: All video must begin and end with at least one black frame. In addition, videos that begin with or contain empty edits will be blocked; the file can contain an empty edit in its edit list only if it is the last edit.

Film Audio Source Profile

For every film that 5.1 Surround audio is available in any competing format or market, it must be provided to iTunes in addition to the stereo tracks.

Surround

- LPCM in either Big Endian or Little Endian, 16-bit or 24-bit, at least 48kHz

- Expected channels: L, R, C, LFE, Ls, Rs

Stereo

- LPCM in either Big Endian or Little Endian, 16-bit or 24-bit, at least 48kHz

- Expected Dolby Pro Logic channels: Lt, Rt or expected stereo channels: L, R

Film Audio/Video and Alt-Audio Container

- Deliver all content in a QuickTime .mov file container.

- The QuickTime `.mov` file extension is expected for all audio and video content.

- Each audio channel must have an assignment. The channel assignments must match one of the options in the table below. For Option 1 (one track with all six channels), the order of the channel assignments can vary as noted in Option 1a, 1b, 1c, and 1d. Note that "Lt" and "Rt" are only used for Dolby matrix audio mixdown.

5.1 Surround Tracks						Stereo Tracks	
L	R	C	LFE	Ls	Rs	Lt	Rt

Option 1a	Track 1 -- six channels						Track 2	Track 3
	L	R	C	LFE	Ls	Rs	Lt	Rt

Option 1b	Track 1 -- six channels						Track 2	Track 3
	L	R	Ls	Rs	C	LFE	Lt	Rt

Option 1c	Track 1 -- six channels						Track 2	Track 3
	L	C	R	Ls	Rs	LFE	Lt	Rt

Option 1d	Track 1 -- six channels						Track 2	Track 3
	C	L	R	Ls	Rs	LFE	Lt	Rt

One track containing all Surround channels; Matrix Stereo with Lt in one track and Rt channel in another track

Option 2	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7	Track 8
	L	R	C	LFE	Ls	Rs	Lt	Rt

One track for each channel

Option 3	Track 1 -- six channels						Track 2 -- two channels	
	L	R	C	LFE	Ls	Rs	Lt	Rt

One track containing all Surround channels; Matrix Stereo with Lt and Rt channels in one track

Option 4	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7 -- two channels	
	L	R	C	LFE	Ls	Rs	Lt	Rt

One track for each Surround channel; Matrix Stereo with Lt and Rt channels in one track

Option 5		Track 1	Track 2
		L	R

No Surround. Stereo with L in one track and R in another track

Option 6		Track 1 -- two channels	
		L	R

No Surround. Stereo with both L and R channels in one track

> **Important:** Refer to "Table 1: Audio Channel Assignment Labels" (page 41) for label descriptions and "How to Apply Audio Channel Assignments" (page 35) for instructions on applying audio channel assignments.

Note: For more information on alternate audio, see the "Assets and Data Files" section in the *iTunes Package Film Specification* .

Film Closed Captioning Profile

- English text in EIA 608 format.

- Delivered in the same package with the video it references.

- In a Scenarist SCC formatted file, using `.scc` file extension.

- The timecode framerate can only be 29.97. This framerate is independent from your video source framerate.

- The timecode format must match the framerate format of the source video, either drop frame (DF) or non-drop frame (NDF).

  - Drop frame format has colons for the first two time delimiters and a semi-colon for the last time delimiter (HH:MM:SS;FF)

  - Non-drop frame format has colons for all the time delimiters (HH:MM:SS:FF)

Source Video Framerate	Description	Timecode Format	Timecode Example
29.97, 59.94i	NTSC Video	DF	HH:MM:SS;FF
25, 50i	PAL Video	NDF	HH:MM:SS:FF
24	Film	NDF	HH:MM:SS:FF
23.976	NTSC Film	NDF	HH:MM:SS:FF

- SCC files must be validated for proper sync against the associated video file using QuickTime 7 Pro.

- Captions should display and synchronize to within one second of the initial, audible dialog to be represented in text.

The timecodes of the captions are relative to the start of the program, and not the QuickTime movie's timecode track.

Currently, the iTunes Store does not support EIA 708 (ATSC closed captioning) or Teletext.

MacCaption is a tool you can use to create .scc files: http://www.cpcweb.com/products/. (Note that this product is not endorsed by Apple. Apple cannot and does not provide support for third-party products.)

> **Note:** The closed caption file must be provided unless it does not exist.

Film iTunes Timed Text Profile

Below is a summary of delivery requirements for iTunes Timed Text. Refer to Chapter 5 in the *iTunes Package Film Specification* for complete details.

- Delivered in an iTunes Timed Text (iTT) formatted file, using `.itt` file extension.
- Delivered in the same package with the video it references as an asset in the `<assets>` block.
- Only one `div` element is allowed in an iTT document.
- `timeBase` must be set to `smpte`.
- `dropMode` must be set to `"dropNTSC"` or `"nonDrop"`; iTunes Timed Text does not support `dropPAL`.
- Only `sansSerif` may be specified as the typeface in `fontFamily`.

The iTT file format is a subset of the Timed Text Markup Language, Version 1.0 W3C Candidate Recommendation 23 February 2010 (TTML) (http://www.w3.org/TR/2010/CR-ttaf1-dfxp-20100223/) from the World Wide Web Consortium (W3C) (http://w3.org/). All iTT documents are TTML documents that use the restricted subset of TTML.

Film Dub Card Video Profile

The full feature-length video asset is comprised of a set of data files, which play specific roles for their asset. The following table describes the optional data file for dub card video.

Asset Type	Data File	Description
Full	**Role:** video.end.dub_credits An optional data file containing the credits associated with an audio track.	A video-only sequence containing one or more still credits specific to the locale-matched audio. iTunes products will include dub credit video sequences for the associated audio dubs following the main program. Locale: Required

Dub Card Video Profile

- Apple ProRes 422 (HQ)

- Movie correctly tagged with color parameter: ITU BT.709

- Video dimensions, pixel aspect ratio, and frame rate must match full program video

- Minimum of 4 seconds per dub card

- Crop dimensions from full program video will be applied to all dub card video — effective crop must match full program video

- Crop attributes must not be supplied for dub card video

- Sound tracks should not be supplied for dub card video — sound tracks will be ignored

- Dub card video will be deinterlaced if necessary so the field order does not need to match — progressive is preferred

- Dissolves and scrolling credits are not supported

- First and last frames do not need to be black frames

Film Chapter Image Profile

- JPEG with `.jpg` extension (quality unconstrained)

- RGB (screen standard)

- Must be same aspect ratio as video source

- 640 minimum horizontal dimension (larger for HD sourced)

- Variable size vertical dimension (based on aspect ratio of video source)

- Only active pixel area may be included (except where necessary to match the overall aspect ratio of the full program)

- Chapter images must be cropped (no letterbox, pillarbox, or windowbox, except as noted above)

- Chapter images must contain picture content

- Chapter image files must be unique with different checksums

---

Important: CMYK (print standard) images will **not** be accepted.

Film Poster Art Profile

- JPEG with `.jpg` extension (quality unconstrained) or PNG with `.png` extension

- RGB (screen standard)

- 1400 x 2100 pixels minimum size

- 2:3 aspect ratio

- Poster art (one-sheet) from film. Must contain key art and title. DVD cover, release date, website, or promotional tagging may not be included.

- Poster art must not display film ratings.

Do not increase the size of a smaller image to meet the minimum size standard. Excessively blurry or pixelated images will be rejected.

> **Important:** CMYK (print standard) images will **not** be accepted.

Film Content Considerations

- The full movie asset should not contain FBI, MPAA, or release date tagging.

- The trailer asset should not contain FBI, MPAA, or release date tagging.

- A minimum of 1 black frame at the beginning and end of each video is required.

- Trailer should be same aspect ratio as the full asset.

- Promotional bumpers, including URLs, are NOT accepted. For more details, please contact your iTunes Technical Representative.

- Trailers must contain content suitable for a general audience.

- Poster art should not contain DVD tagging, release date tagging, or website tagging.

XML

- All XML must be encoded in UTF-8.

- No byte order markers (BOM) can be used.

- There should be no null data or empty tags in the XML. If not used, elements should be removed.

- The XML must be formatted to use line breaks and indentations.

For further information, please refer to the appropriate media type metadata specification, or consult with your iTunes Technical Representative.

Revision History

Previous Spec Revisions

The following table lists the previously-released specifications and the revisions:

Date/Version	Summary
March 20, 2013 - Version 5.1 Revision 1	Corrected the ProRes SD Profiles for NTSC and PAL.
February 28, 2013 - Version 5.1	Added best practices for delivering Mastered for iTunes content. Changed requirements for ringtone album cover art. Clarified acceptable frame rates for closed captioning for TV and Film. Videos with empty edits other than the last edit will be blocked.
November 7, 2012 - Version 5.0 Revision 1	Clarified SD Source video for film. Added new video source validations. Added Apple ProRes to NTSC and PAL SD TV source profiles. Changed requirements for QuickTime audio channel assignments. Closed captions can be added to MPEG-2 sources
May 30, 2012 - Version 5.0	Album cover art and poster art requirements have changed. Removed TIFF from the list of recommended image formats and removed DPI requirements. Added delivery requirements for dub card video. 96Khz audio is now supported.
September 22, 2011 - Version 4.8	Added crop dimensions for TV. Clarified content considerations for TV. Clarified closed captioning for TV. Added delivery requirements for iTT files.
July 13, 2011 - Version 4.7 Revision 2	Clarified delivery requirements for 5.1 audio and closed captioning. Added the profile for closed captioning for TV. Film poster art requirements have changed.
April 15, 2011 - Version 4.7	Clarified HD cropping for TV. Added color space requirement for HD film source. Clarified closed captioning text for film.
February 9, 2011 - Version 4.6	Removed asset specifications for books (a new iBooks Store asset guide has been created). Renamed this asset guide to: *iTunes Video and Audio Asset Guide*.

Date/Version	Summary
November 5, 2010 - Version 4.5	Clarified surround sound for HD music video audio source profile. Clarified delivery of HD source for music videos. Added a chapter for book source profiles. Put back 25 fps in the HD TV source profile that was incorrectly removed. Added two new best practice items to the TV Content Considerations section.
August 5, 2010 - Version 4.4	Added source profile for HD music video and cropping information. Clarified album cover art. Added surround sound to HD music video audio source profile.
February 3, 2010 - Version 4.3	Clarified that ALAC in a CAF container is allowed. Added source profile for pre-cut ringtones. Clarified that film ratings should not appear on poster art.
December 18, 2009 - Version 4.2	Clarified quality standards. Clarified closed captioning.
November 10, 2009 - Version 4.1	Clarified audio requirements for music and film.
September 11, 2009 - Version 4.0	Added best practices content for Film. Clarified requirements for SCC files.
July 1, 2009 - Version 3.3.2	Clarified image and audio requirements. Clarified frame rate requirements for TV.
May 12, 2009 - Version 3.3.1	Added support for PNG format images for cover art, poster art, and video screen captures. PNG images are not currently supported for chapter thumbnail images.
March 17, 2009 - Version 3.3	Added updated PAL support for film. Added closed-captioning to Film Content Profile. Added 24-bit support for audio. Added best practices content for TV. Clarified how to send stereo sound for Film and TV.
October 1, 2008 - Version 3.2	Added audio source specification to Music Audio Content Profile, added HD format to Television Content Profile and Appendix I, which provides audio channel assignments instructions.
May 8, 2008 - Version 3.1.1	Complete reformatting of the Guide. Separation of content type profiles. Addition of Movie HD and SD specification. Addition of image specifications for TV and Film.
April 2, 2007 - Version 2.3	Introduction of Asset Specification Guide.

Audio Channel Assignments

How to Apply Audio Channel Assignments

Step 1: Open the Movie Properties window from the **Window** > **Show Movie Properties** menu.

> **Important:** You must use the Pro version of QuickTime.
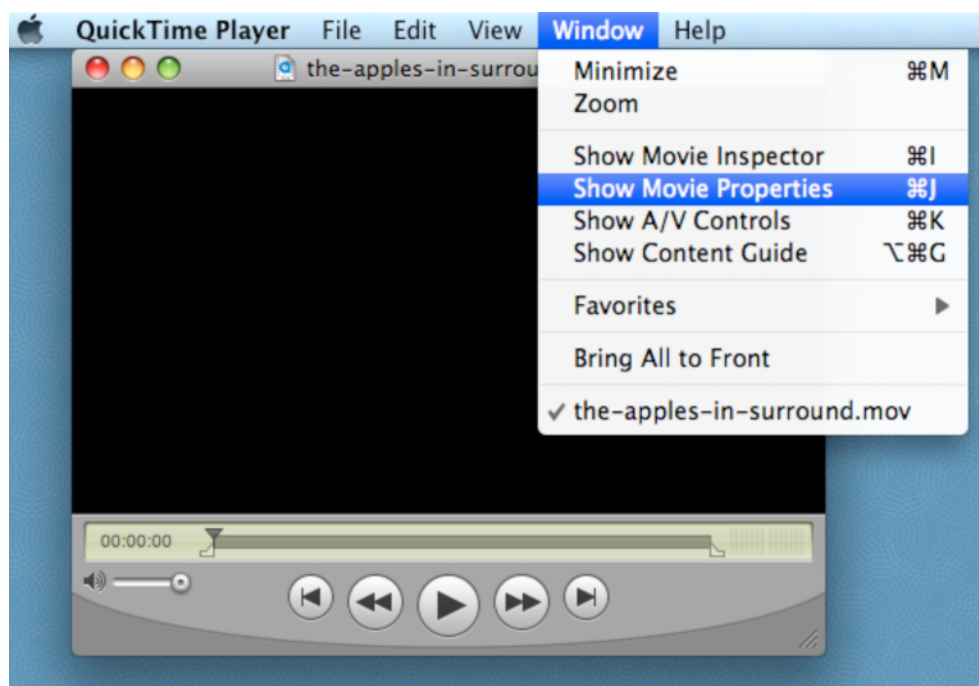
Standard Movie Properties window:
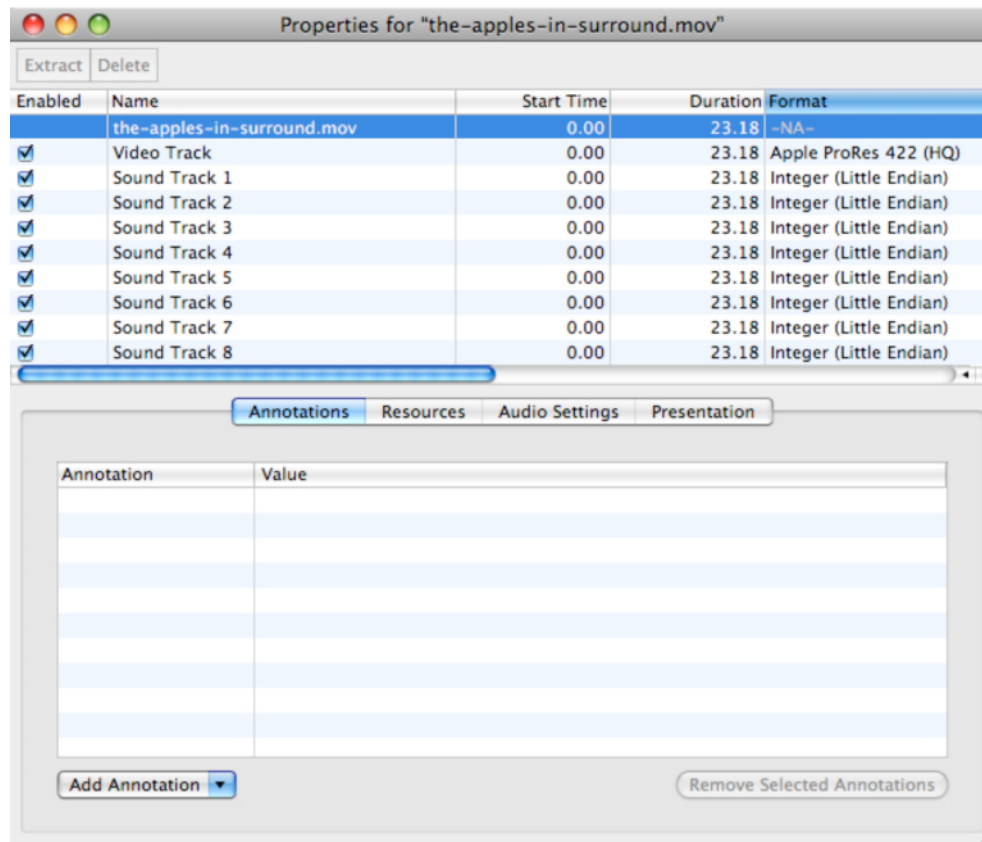
Step 2: Go to the **View** > **Columns** menu and choose **Channels**. You may add additional columns like ID, Data Rate, and so on.
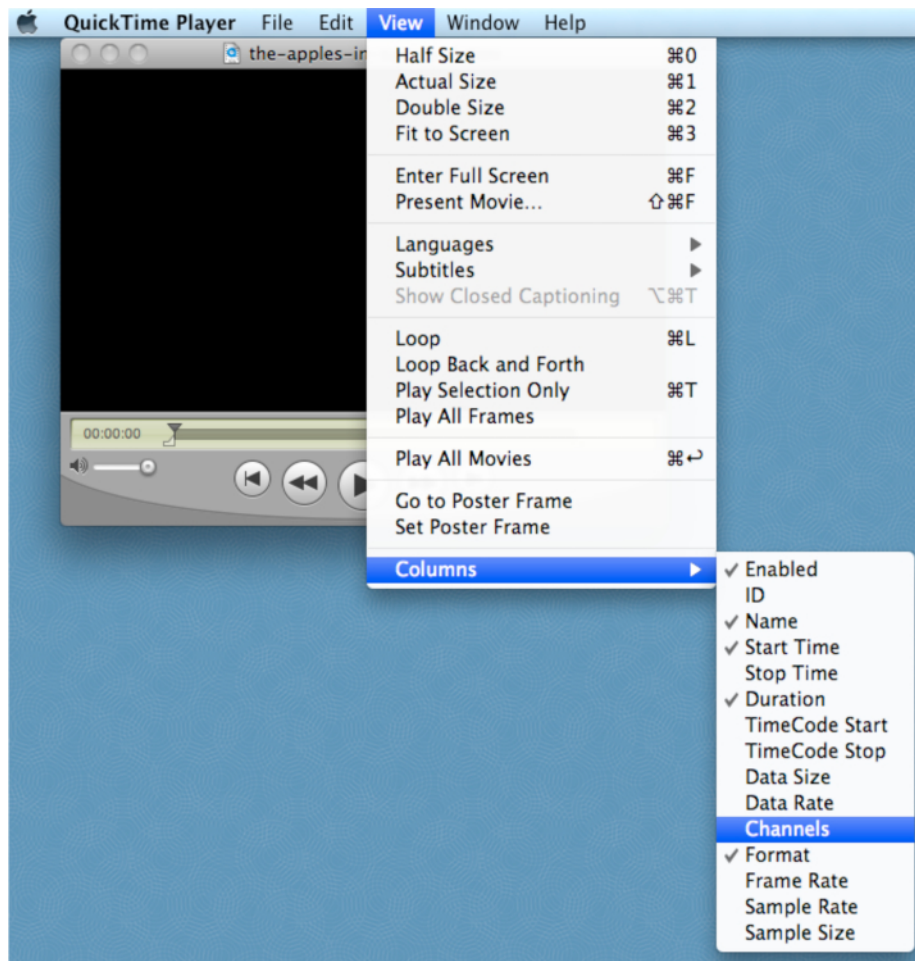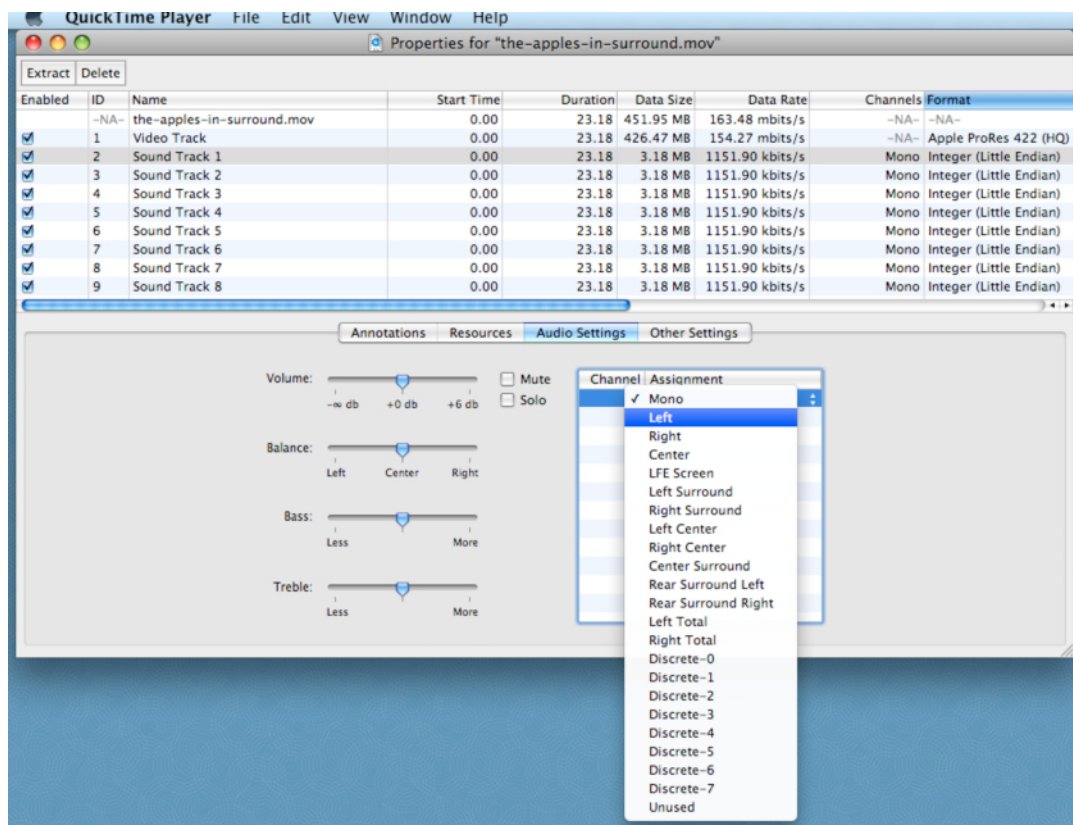
Notice the Channels column in the following screenshot indicates the audio tracks are Mono.
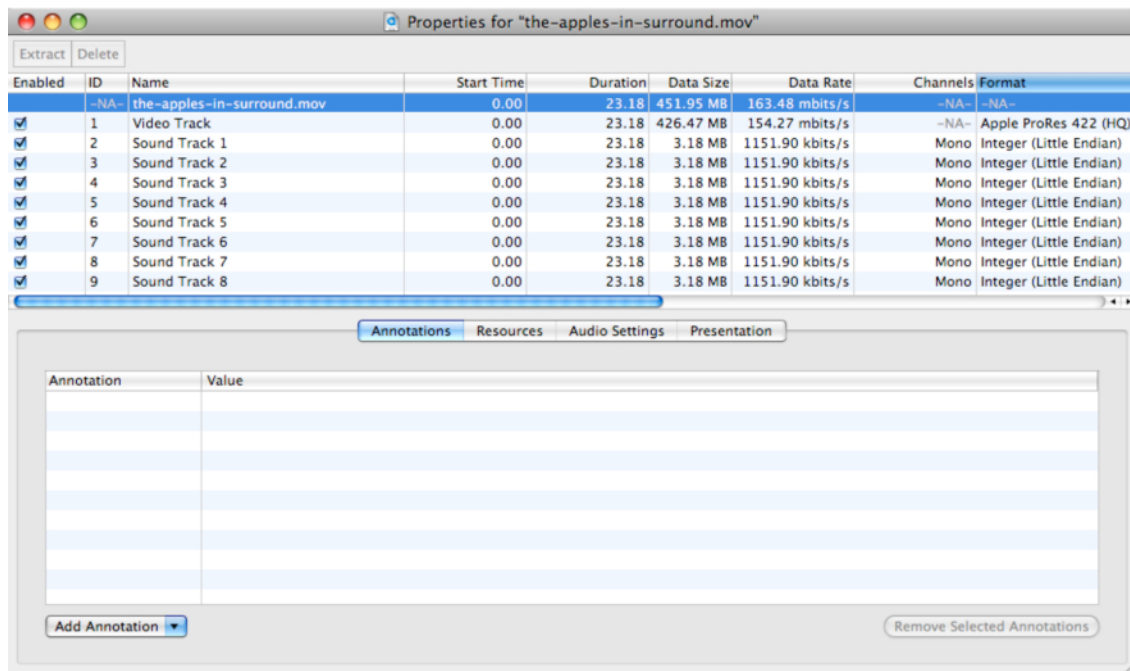
Step 3: Select the sound track to make the audio channel assignment and click the **Audio Settings** tab.
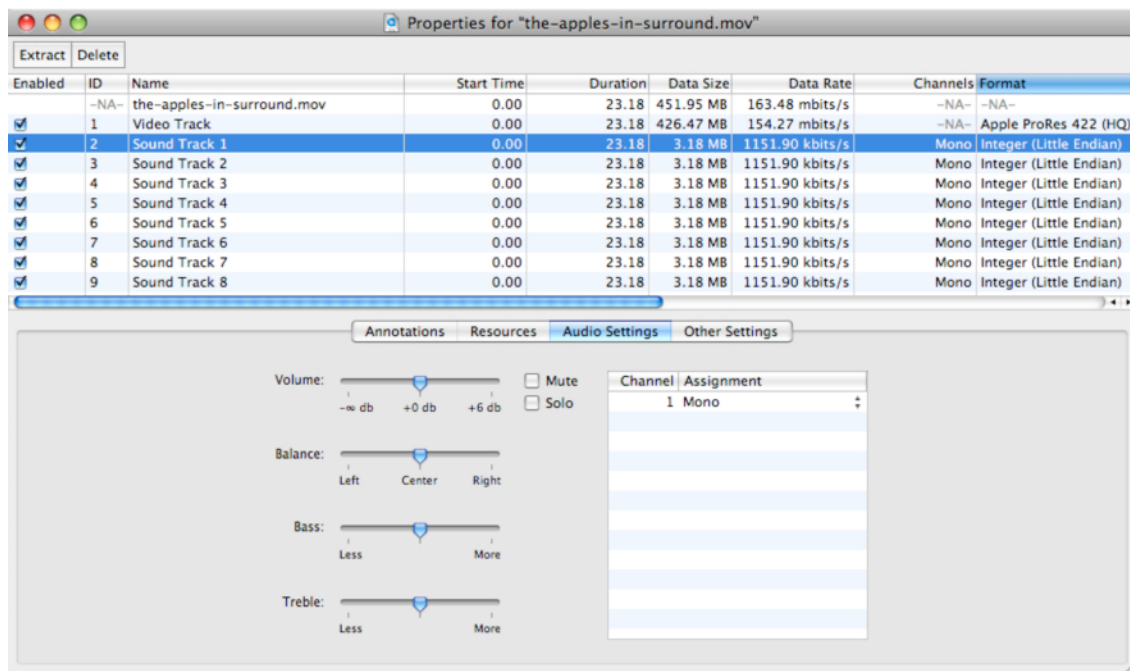
Step 4: To make the Channel Assignment, choose the appropriate setting from the pop-up menu. Repeat this process for each audio track.

In this example all the tracks have been properly assigned for the channel assignments of Option 2 as indicated in the Audio/Video Container section for "Television Content Profiles" (page 15) and "Film Content Profiles" (page 24).

Notice the Channels column in the following screenshot no longer indicates the audio tracks are Mono.
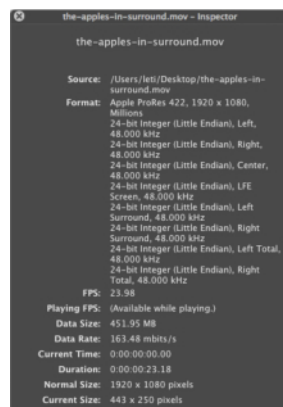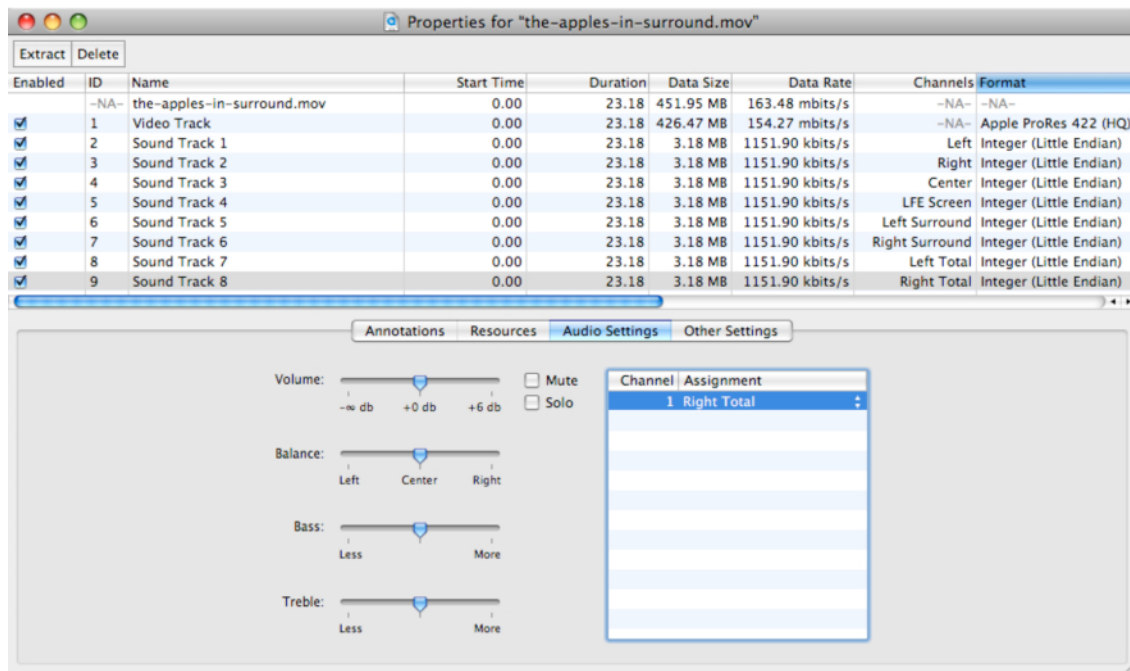
Notice this indicates the assignment. This should not indicate mono.

Step 5: Save the file.

Step 6: Open the Movie Inspector from the **Window** > **Show Movie Inspector** menu to verify assignments were applied correctly.

Table 1: Audio Channel Assignment Labels

Label	Description
L	Left
R	Right
C	Center
LFE	LFE Screen
Ls	Left Surround
Rs	Right Surround
Lt*	Left Total
Rt*	Right Total

* Lt and Rt are supported in the latest version of QuickTime.