CD STANDARD

MXF Multichannel Audio Controlled Vocabulary

Table of Contents

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards,

Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in its Standards Operations Manual. This SMPTE Engineering Document was prepared by Technology Committee 31FS.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

Developments in the moving image production industries have led to increasing definition and use of many multichannel audio (herein abbreviated "MCA") schemes, with channel counts in excess of 20. Since the MXF file format was published in 2004 the requirement to identify such audio for efficient production, storage and distribution within MXF has led to standardization for audio labeling found in SMPTE ST 377-4:20xx, *MXF Multichannel Audio Labeling Framework*. Metadata associated with content adds value to that content in both automation – with the ability to save on costly and error-prone human input; and in manual handling – where the efficient searching and description of these assets enhances their re-use and thus the realizable value.

Since SMPTE ST 377-4 was published in 2012, additional granularity and a controlled vocabulary describing multichannel audio has been desired in order to give the uniqueness necessary for a device to perform logical operations on the audio and facilitate automated processes.

This engineering document defines a controlled vocabulary for selected subdescriptors in SMPTE ST 377-4:20xx to allow for automation and enable multiple applications to utilize these properties. These values are intended for easy human and machine recognition.

This engineering document defines values for each MCALabelSubDescriptor subclass in SMPTE 377-4:20xx that are known to be in wide use at the time of publication. Applications using the MCA framework may later define additional values to suit the needs of the application. While it is possible to create application-specific documents with these values, it would ultimately be best to have all of the values in one document that can serve as the controlled vocabulary for all multichannel audio.

It is expected that a number of applications will use this controlled vocabulary to increase interoperability for SMPTE ST 377-4:20xx.

1 Scope

This standard defines a controlled vocabulary for the MCA labeling framework found in SMPTE ST 377-4:20xx.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; then formal languages; then figures; and then any other language forms.

3 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this engineering document. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this engineering document are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE ST 377-4:xxxx, MXF Multichannel Audio Labeling Framework

SMPTE ST 2067-8:2013, Interoperable Master Format — Common Audio Labels

4 Terms and Definitions

In addition to the terms and definitions in SMPTE ST 377-4:20xx, the below terms and definitions apply for this document.

4.1 Descriptive Video

descriptive video, defined as “Descriptive Video Service” in SMPTE ST 2067-8:2013 section 7 and further defined in Table 2 below.

Note: Descriptive Video is also called “Audio Description (AD)”.

4.2 Lt-Rt

Left Total and Right Total audio, defined as “Lt-Rt” in SMPTE ST 2067-8:2013 section 6.

Note: Lt-Rt audio contains surround information encoded into the two channels.

Note: “Lt-Rt” is sometimes written “LtRt”.

4.3 Mono

monaural audio, defined as “Mono One” in SMPTE ST 2067-8:2013 section 6.

4.4 Stereo

“Stereo”, defined as “Standard Stereo” in SMPTE ST 2067-8:2013 section 6.

Note: Standard Stereo audio does not contain any surround information.

Note: Standard Stereo is also referred to as LoRo (Left-only, Right-only).

5 Controlled Vocabulary

5.1 Controlled Vocabulary Usage

The properties (e.g. MCA Content) in this section are defined in ST 377-4:20xx. The rules and tables below define the controlled vocabulary values for each of these properties using Name, Symbol, and Definition.

Note: Except where noted, requirements are written to define the complete set of valid values. That is, an application can only supplement these values by revising or amending this engineering document.

5.2 MCA Partition Kind

When MCA Partition Kind is encoding one of the partition kinds defined in Table 1, the value in the “Symbol” column shall be used.

Table 1: Valid MCA Partition Kind Vocabulary

Name	Symbol	Definition
Full Length	FL	The entire program with no partitioning
Reel	REEL	A segment of the program that traditionally corresponded to a reel of film but today can be any segment of the program that is determined to be convenient for post-production purposes
Act	ACT	A segment of the program corresponding to a defined part of the story.
Part	PART	A presentation segment of the program, such as of a program presented over more than one continuous showing, for example, a mini-series.

5.3 MCA Content

The `MCA Content` item shall be equal to one of the values in the "Symbol" column specified in Table 2 below:

Table 2: Valid MCA Content Vocabulary

Name	Symbol	Definition
Primary	PRM	Primary program content that combines dialog, music, effects and narration (if applicable).
Secondary Audio Program	SAP	Program content that combines a Primary (PRM) with a Voice Over (VO). SAP is generally used as an way of providing program in another language without having to do a full translation and mix.
Hearing Impaired	HI	Audio content designed for hearing-impaired audiences, generally consisting of primary program content with its dialog content mixed considerably louder than its music and effects content, optionally employing audio compression or limiting.
Descriptive Video	DV	Audio content designed for visually-impaired audiences that combines primary program content with narration content describing selected action in the picture.
Dialog	DX	If MCA Language Attributes=Original: The composite of all spoken original language dialog content, including crowds, walla and any other dialog-related content of the primary program, in their final balances, in a single Soundfield Group. This can contain production effects that were recorded on the set or location. If MCA Language Attributes=Dubbed: The composite of all dubbed spoken dialog content for a given language translation, including any optional material that might have been used from the primary program, in their final balances, in a single Soundfield Group. Note that a dubbed dialog has no effects fill, and therefore is intended to be mixed with an ME to create a dubbed dialog program.
Music	MX	The composite of all music content of the primary program, in its final balances, in a single Soundfield Group.
Effects	FX	The composite of all sound effects content, including backgrounds, Foley, hard effects and any other effects-related content of the primary program, in their final balances, in a single Soundfield Group. FX does not contain any production effects or filled effects.
Filled Effects	FFX	The composite of all sound effects content, including backgrounds, Foley, hard effects and any other effects-related content of the primary program, in their final balances, in a single Soundfield Group. FFX also contains production effects or filled effects that are added to replace production effects in areas where dialog was removed. This cannot be mixed with original dialog to create a PRM as the effects will double up. It is generally created in the process of creating an ME but is delivered separately so that different music can be used for different territories if required.
Music and Effects	ME	The composite of all music and effects content of the primary program, without dialog content but with additional material added to fill production effects in areas where dialog was removed, in their final balances, in a single Soundfield Group. A ME does not contain any optional material. An ME can be created by mixing MX with FFX.
Optional Music and Effects	OP	Optional music and effects content that can either be dubbed or used as-is for localization, in a single Soundfield Group. More

		than one may be present, each containing different optional material, in which case the MCA Content Differentiator is used to differentiate. This cannot be used to create a primary program.
Music and Effects with Optional	MESP	A Group of Soundfield Groups containing a ME Soundfield Group and one or more OP Soundfield Groups.
DME	DME	Dialog, Music, and Effects. A Group of Soundfield Groups containing an original language DX Soundfield Group, a MX Soundfield Group and a FX Soundfield Group.
NDME	NDME	Narration, Dialog, Music, and Effects. A Group of Soundfield Groups containing program narration, an original language DX Soundfield Group, a MX Soundfield Group and a FX Soundfield Group.
Program Narration	PNAR	Narration content inherent to the primary program, for example, in a documentary.
Optional Narration	ONAR	Optional Narration content that is not inherent to the primary program but relates to it. This can be recorded and supplied with the primary program.
Voice Over	VO	Localized dialog intended to be heard over the primary program, i.e. as an alternative to replacing the original dialog with translated dialog.
Visually Impaired	VI	Narration content designed for visually-impaired audiences that describes selected action in the picture. The content is generally a narrative only and is designed to be heard while listening to the primary program.
Recorded Commentary	CM	Commentary on selected material in the picture that has been previously recorded to go along with the primary program. When using this MCA Content, an associated MCA Content Subtype shall be used to describe the type of commentary
Live Commentary	LCM	Live Commentary content, such as would be created by an event announcer, for example, a sports announcer. When using this MCA Content, an associated MCA Content Subtype shall be used to describe the type of commentary
Custom	x-<private>	Any Audio Channel, Soundfield Group or Group of Soundfield Groups not covered by another category. Symbol begins with "x-" and is followed by 1-4 character private tag that uniquely describes the content.

5.4 MCA Use Class

The MCA Use Class item shall be equal to the value in the "Symbol" column in Table 3 below:

Table 3: Valid MCA Use Class Vocabulary

Name	Symbol	Definition
Finished Composite	FCMP	The associated MCA Content is a composite, complete work and need not be mixed prior to presentation. It can be mixed with other Soundfield Groups to create other particular desired content per the allowed combinations in Table 4.
Intermediary Composite	ICMP	The associated MCA Content is a composite but requires mixing with other elements prior to presentation. It can be mixed with other Soundfield Groups to create desired FCMP content per the allowed combinations in Table 4.
Simplified	SMPL	The associated MCA Content contains a simplified mix, typically mono, stereo or Lt-Rt, intended for special uses. It can be mixed with other SMPL Soundfield Groups to create desired SMPL content per the allowed combinations in Table 4.
Singular	SING	Element contains a single content (such as narration voice) that can be output on its own or be mixed with other Soundfield Groups.

5.4.1 MCA Use Class Constraints

Only certain MCA Use Class values are compatible with given MCA Content values.

For a given value of MCA Content in Table 4, valid MCA Use Class values are those shown in a column with a corresponding "X" and a superscript that references an entry in Table 5. MCA Use Class shall contain only values in accordance with Table 4.

Table 4: Valid values of MCA Use Class for a given MCA Content Vocabulary

		MCA Use Class			
		Finished Composite (FCMP)	Intermediary Composite (ICMP)	Simplified (SMPL)	Singular (SING)
MCA Content	Program (PRM)	X[1]		X[2]	
	Second Audio Program (SAP)	X[3]		X[4]	
	Hearing Impaired (HI)	X			
	Descriptive Video (DV)	X			
	Recorded Commentary (CM)	X[5]			X[6]
	Dialog (DX)		X[7]	X[8]	
	Music (MX)		X[9]	X[10]	
	Effects (FX)		X[11]	X1[12]	
	Music and Effects (ME)		X[13]	X[14]	
	Optional Music and Effects (OP)		X[15]	X[16]	
	Music and Effects with Optional (MESP)		X[17]	X[18]	
	DME		X[19]	X[20]	
	NDME		X[21]	X[22]	
	Filled Music and Effects (FFX)		X[23]	X[24]	
	Program Narration (PNAR)				X
	Optional Narration (ONAR)				X
	Live Commentary (LCM)				X
	Voice-Over (VO)				X
	Visually Impaired (VI)				X

As shown in Table 4, there are several instances where a given MCA Content value corresponds with more than one MCA Use Class. Each combination has unique meaning described in Table 5.

Table 5: MCA Content and MCA Use Class Combination Meanings

Note	Combination (MCA Content & MCA Use Class)	Description
1	PRM & FCMP	Primary program content that is a composite, complete work in its native Soundfield Group and needs no further processing prior to presentation. It can be mixed with other Soundfield Groups to create other particular desired content per the allowed combinations in Table 4.
2	PRM & SMPL	Simplified primary program content that is simplified by means of downmixing to a smaller Soundfield Group, or can be comprised of DX&SMPL, MX&SMPL, FX&SMPL and PNAR (if applicable). Simplified

		primary program is typically a mono, stereo or Lt-Rt Soundfield Group and is used only for deliverables to specific markets, for example, airline or advertising.
3	SAP & FCMP	SAP content that is a composite, complete work in its native Soundfield Group and needs no further processing prior to presentation.
4	SAP & SMPL	Simplified SAP content that is typically a mono, stereo or Lt-Rt Soundfield Group and is used only for deliverables to specific markets, for example, airline or advertising.
5	CM & FCMP	Recorded commentary that is mixed with PRM to make a PRM & FCMP
6	CM & SING	Recorded commentary that is only the narration. It would typically be mixed with PRM to make a final output or could be output separately for downstream use.
7	DX & ICMP	Dialog (whether original or dubbed depending on the value in MCA Language Attributes) that is a composite within its native Soundfield Group but must be mixed with other Soundfield Groups to make a PRM&FCMP.
8	DX & SMPL	Dialog (whether original or dubbed depending on the value in MCA Language Attributes) that is a simplified representation of DX&ICMP. Typically a mono, stereo or Lt-Rt Soundfield Group. It can be mixed with other SMPL Soundfield Groups to make a PRM&SMPL.
9	MX & ICMP	Music that is a composite within its native Soundfield Group but must be mixed with other Soundfield Groups to make a PRM&FCMP.
10	MX & SMPL	Music that is a simplified representation of MX&ICMP. Typically a mono, stereo or Lt-Rt Soundfield Group.
11	FX & ICMP	Effects that is a composite within its native Soundfield Group but must be mixed with other Soundfield Groups to make a PRM&FCMP.
12	FX & SMPL	Effects that is a simplified representation of FX&ICMP. Typically a mono, stereo or Lt-Rt Soundfield Group.
13	ME & ICMP	Music and Effects that is a composite within its native Soundfield Group and can be mixed with dubbed DX&ICMP to create a final dubbed PRM&FCMP
14	ME & SMPL	Music and Effects as a simplified representation, typically a mono, stereo or Lt-Rt Soundfield Group and can be mixed with dubbed DX&SMPL to make a PRM&SMPL
15	OP & ICMP	Optional Music and Effects in its native Soundfield Group, which can be output separately and used to create a dubbed DX&ICMP at a mixing facility. This cannot be used to create a PRM.
16	OP & SMPL	Optional Music and Effects as a simplified representation, typically a mono, stereo or Lt-Rt Soundfield Group, which can be output separately and used to create a dubbed DX&SMPL at a mixing facility. This cannot be used to create a PRM.
17	MESP & ICMP	Music and Effects with Optional Group of Soundfield Groups that can be used in creating a dubbed DX&ICMP and a final dubbed PRM&FCMP
18	MESP & SMPL	Music and Effects with Optional Group of Soundfield Groups as a simplified representation, for example, an Lt-Rt ME and one or more mono OP. This can be used in creating a dubbed DX&SMPL and a PRM&SMPL
19	DME & ICMP	DME Group of Soundfield Groups that can be mixed together to create a PRM&FCMP. This GSG is commonly called “Stems”.
20	DME & SMPL	DME Group of Soundfield Groups where each Soundfield Group is a simplified mix, generally either mono, stereo or Lt-Rt. These can be mixed together to create PRM&SMPL, but not PRM&FCMP. This GSG is commonly called a “Split Track”
21	NDME & ICMP	NDME group of Soundfield Groups that can be mixed together to create a PRM&FCMP. This would apply only to content that contained program narration, otherwise “DME” is used.
22	NDME & SMPL	NDME Group of Soundfield Groups where each Soundfield Group is a simplified mix, generally either mono, stereo or Lt-Rt. These can be

		mixed together to create PRM&SMPL, but not PRM&FCMP. This would apply only to content that contained program narration, otherwise "DME" is used.
23	FFX & ICMP	Filled Effects that is a composite within its native Soundfield Group but must be mixed with a ME Soundfield Group to make a ME or MESP to be used in dubbing.
24	FFX & SMPL	Filled Effects that is a simplified representation of FFX&ICMP. Typically a mono, stereo or Lt-Rt Soundfield Group.

5.5 MCA Content Subtype

When present, the MCA Content Subtype item shall be equal to one of the values in the "Symbol" column in Table 6.

Table 6: Valid MCA Content Subtype Vocabulary

Name	Symbol	Definition
Director	DIR	Recorded Commentary with the Director as the primary speaker. Typically used with the MCA Content "CM"
Technical	TECH	Recorded Commentary with a technical person as the primary speaker, such as one that designed the visual effects. Typically used with the MCA Content "CM"
Writer	WRT	Recorded Commentary with the screenplay writer as the primary speaker. Typically used with the MCA Content "CM"
Cast	CAST	Recorded Commentary with one or more of the cast members. Typically used with the MCA Content "CM"
Announcer	ANN	Live Commentary with a person announcing the event. Typically used with the MCA Content "LCM"
Commentator	CTR	Live Commentary with a person or people who are commenting on an event or other piece of content. Typically used with the MCA Content "LCM"
Other	OTHER	Placeholder for other types of Live Commentary. Typically used with the MCA Content "LCM" If "OTHER" is used for SubType, an MCA Additional Attributes value should be present, including identification of the other SubType.

5.6 MCA Spoken Language Attribute

If present, MCA Spoken Language Attribute shall be equal to one of the values in the "Symbol" column in Table 7 below:

Table 7: Valid MCA Spoken Language Attribute Vocabulary

Name	Symbol	Definition
Original	ORIGINAL	Indicates that the MCA Content contains the original spoken dialog
Dubbed	DUBBED	Indicates that the MCA Content contains dubbed dialog which replaces some or all of the original spoken dialog with another language

5.7 MCA Additional Spoken Language Attributes

If present, MCA Spoken Language Attribute shall contain an entry for each RFC 5646 Additional Spoken Languages with its corresponding attribute represented as a value in the "Symbol" column in Table 7.