

SMPTE STANDARD

D-Cinema Packaging —
Timed Text Track File

Page 1 of 14 pages

Table of Contents	Page
Foreword	2
Intellectual Property	2
Introduction	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	3
4 Glossary of Acronyms and Terms	4
5 Overview (Informative)	5
6 Ancillary Resource Wrapping	5
6.1 Generic Stream Wrapper	5
6.2 Generic Stream Repetition	6
6.3 Indexing Generic Stream Data	6
7 Mapping XML Timed Text to the MXF Generic Container	6
7.1 Essence Encoding	6
7.2 Index Table	7
7.3 Timed Text Resource Constraints	7
7.4 Header Metadata Construction	7
7.5 Essence Descriptors	7
8 Essence Encryption	9
9 Synchronization	10
10 Random Access to Ancillary Resources (Informative)	10
11 Timed Text Track File Structure	10
11.1 Timed Text Track File Definition	10
11.2 Generic Container	10
11.3 Timed Text Resource	11
11.4 Header Metadata Constraints	11
11.5 Index Tables	11
Annex A Labels and Descriptor Sets (Normative)	12
A.1 Key UL Values	12
A.2 TimedText Descriptor Set	12
A.3 TimedText Resource SubDescriptor Set	13
Annex B Bibliography (Informative)	15
Annex C Summary of Changes from Version 2009 (Informative)	16

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.

SMPTE ST 429-5 was prepared by Technology Committee 21DC.

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.

Many applications, including D-Cinema, rely on XML expressions of timed text material for caption and subtitle essence. An MXF Generic Container mapping for XML text data allows convenient carriage of text resources with related ancillary resources, such as fonts and sub-pictures, in an MXF-oriented media workflow. The mapping optionally allows encryption of these contents for confidentiality.

This work was originally developed to support D-Cinema, but since that time other applications have been developed that have also identified a requirement for carrying XML timed text in MXF. To support these applications, the definition of the timed text GC mapping and the D-Cinema Track File are defined separately so that the GC mapping may be applied to applications other than D-Cinema.

1 Scope

This standard specifies the format of a Generic Container (GC) for XML timed text and Timed Text Track File for the distribution of timed text content using the Material Exchange Format (MXF). The GC and Track File provide for carriage of an XML document and optional supporting resources such as images or fonts. Encryption is optionally available for protecting against unauthorized disclosure of the file contents.

The standard defines data structures for interchange at the signal interfaces of networks or storage media, but does not define internal storage formats for compliant devices.

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, followed by 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 standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE ST 377:2004 Television — Material Exchange Format (MXF) — File Format Specification

SMPTE ST 377-1:2012 Television —Material Exchange Format (MXF)

SMPTE ST 379:2004 Television — Material Exchange Format (MXF) — MXF Generic Container

SMPTE ST 379-1:2009 —Material Exchange Format (MXF) — MXF Generic Container

SMPTE ST 379-2:2010 —Material Exchange Format (MXF) — MXF Constrained Generic Container

SMPTE ST 410:2008 Television —Material Exchange Format — Generic Stream Partition

SMPTE ST 429-3:2007 —D-Cinema Packaging — Sound and Picture Track File

SMPTE ST 429-6:2006 —D-Cinema Packaging — MXF Track File Essence Encryption

Internet Engineering Task Force (IETF) (1996, November) — [RFC 2045] Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies

Internet Engineering Task Force (IETF) (1996, November) — [RFC 2396] Uniform Resource Identifiers (URI): Generic Syntax

Internet Engineering Task Force (IETF) (2005, July) — [RFC 4122] A Universally Unique Identifier (UUID) URN Namespace

ISO/IEC —10646:2014 Information Technology — Universal Multiple-Octet Coded Character Set (UCS)

4 Glossary of Acronyms and Terms

Note: See also the list of acronyms and terms in ST 377M:2004 and ST 377-1.

Closed Caption: Timed-text intended for display on a device other than the theater screen.

Composition: A complete artistic or informational motion picture work, such as a feature, a trailer, or an advertisement, etc.

MXF: Material Exchange Format.

Open Caption: Timed-text intended for display on the theater screen, usually written in the same language as the sound essence.

Reel: A unit of essence comprising all or part of a composition.

Resource: An integral unit of data, such as an XML document, a font, or a sub-picture image.

Sub-picture: An ancillary image intended for display over a larger main image.

Subtitle: Timed-text intended for display on the theater screen, usually written in a language other than the language of the sound essence.

Timed-text: Text intended for display over a timeline, in synchronization with image and sound essence.

URI: A Uniform Resource Identifier, defined by [RFC 2396], is a text string which uniquely identifies a resource, such as an XML Schema. URI values often identify objects not accessible via a computer network.

UUID: Universally Unique Identifier, defined by [RFC 4122].

XML: eXtensible Markup Language, an abstract syntax for structured text with metadata.

5 Overview (Informative)

Subtitles, closed-captions or other forms of textual information often accompany sound and picture essence. The display of textual information varies according to purpose, but the essence encoding can be generalized as a timed-text resource (an XML document) that provides content, position and timing information (*i.e.*, essence plus metadata), and optional ancillary resources such as fonts and sub-pictures.

In some cases, the essence is entirely contained in one or more sub-pictures. In these cases, the timed-text resource contains only timing and position information.

Timed-text essence, therefore, consists of one XML timed-text resource plus optional supporting resources. To simplify the delivery of what may potentially be many dozens of files, this specification allows all resources to be wrapped in a single MXF file. The timed-text resource is contained in an MXF Generic Container (GC) in the Body partition, and any ancillary resources are individually contained in their own Generic Stream partitions.

Figure 1 below illustrates a simple Timed Text Track File for cinema containing text-based subtitle essence. The timed-text resource is contained in the Track File along with a font resource used to render the characters when the text is reproduced on the theater screen.

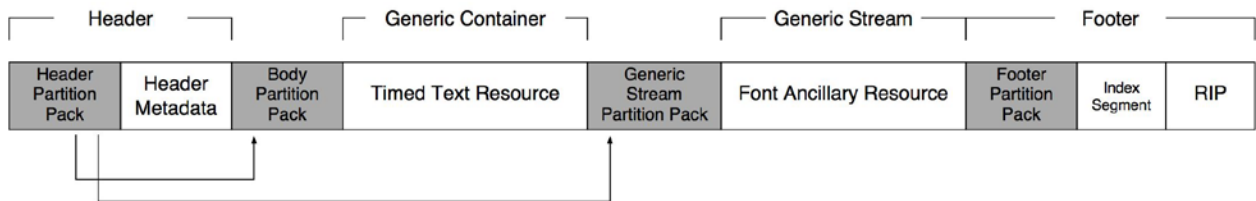


Figure 1 – Example Timed Text Track File Structure using a Font

Figure 2 illustrates a more complex Timed Text Track File containing sub-picture-based subtitles. The timed-text resource is contained in the Track File along with a number of sub-picture resources.

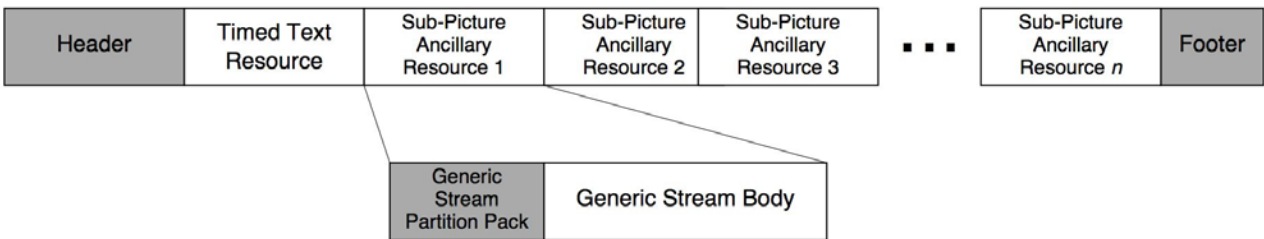


Figure 2 – Example Timed Text Track File Structure using Sub-Pictures

6 Ancillary Resource Wrapping

6.1 Generic Stream Wrapper

Each Ancillary Resource referenced by a Timed Text Generic Container shall be entirely contained within an MXF Generic Stream Partition constructed per ST 410 and located in the same MXF file. Each Generic Stream Partition in the file shall contain exactly one Ancillary Resource. Each Generic Stream Partition shall have a distinct BodySID per ST 410. The Generic Stream Partition shall consist of a Generic Stream Partition Pack