SMPTE RP 157:2012
Revision of RP 157-1995

SMPTE RECOMMENDED PRACTICE

Key and Alpha Signals

Table of Contents

Approved November 1, 2012

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 Part XIII of its Operations Manual.

SMPTE RP 157 was prepared by Technology Committee 10E.

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.

A key signal, also called an alpha signal or a traveling matte, is a video signal (digital or analog) used to control the contribution of an associated video signal, generally called the fill, into a composite with another signal or signals, generally called the background, by modulating the fill and the background before they are summed. The different names originated in different segments of the motion imaging industry; "key" comes from television, "alpha" from computer graphics and "matte" from film. The terms are often used interchangeably, so one will see references to key signals carried on alpha channels, as an example.

This Recommended Practice was developed in the mid-1980s when composite analog image formats and interfaces were predominant. Many (but not all) of the issues this document was originally written to resolve do not exist in digitally-interfaced facilities. In order to make the document more useful in current practice, the legacy analog and digital composite specifications have been moved to a separate section. The Scope has been expanded to take cognizance of the convergence of TV, film and computer graphics technology and terminology.

1 Scope

This document describes the requirements for the key, alpha or matte signal with respect to its associated fill. In the remainder of the document the term “key” is used, but the requirements are equally applicable to all three. This description is given for digital and analog signals and is valid for all image formats, including without limitation SDTV, HDTV and UHDTV.

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 documents contain provisions that, through reference in this text, constitute provisions of this Recommended Practice. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this recommended practice are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

None.

4 General Specifications

These specifications are applicable to all image formats. For digital systems, the terms “video” and “video signal” refer to the digital representation of the image as specified in the relevant image format standard. For a listing of some of these image format standards, see the Bibliography.

Note: The Bibliography also lists a useful reference for a definition of terms used in this document (SMPTE EG 28).

A key video signal represents the opacity or transparency of its associated fill video signal. When the value of the key signal indicates that the fill is opaque, the fill will obscure all video signals of lower priority (collectively, the background) in the composited image and the fill will be visible unless it is itself obscured by video signals of higher priority. When the value of the key signal indicates that the fill video is transparent, the fill will not be visible. When the value of the key signal indicates that the fill video signal is partially transparent, the fill will be mixed with the background in proportion to the degree of transparency represented by the value of the key:

$$\text{Output} = \text{Fill} * \text{Key} + \text{Background} * (1 - \text{Key})$$

Where the value of Key ranges from 0.0 (fully transparent) to 1.0 (fully opaque), inclusive. This is a generic equation; in practice the quantization levels representing full transparency and full opacity shall be defined by the levels representing black and white in the video format being used, as specified below.

Full specification of the implementation of this equation in digital systems is outside the scope of this Recommended Practice.

The key signal and the interface channel in which it is conveyed shall have the same video signal format, including sampling, quantization and filtering as the luminance (Y') or primary (R', G' and B') video signal(s) with which it is associated. There is one exception, which is the electro-optical transfer function. This is explained below.

The key information shall be treated as a video signal, with black representing complete transparency and white representing complete opacity.

Values between black and white indicate partial transparency. Black and white levels shall conform to the specifications for the image format of the fill Specification of the transfer function for values between black and white is outside the scope of this Recommended Practice. However, for most applications, the relation between opacity and key value may be assumed to be linear.

In component digital systems, if the key signal is carried on a separate interface from the fill signal, the key shall be carried in the Y or Luma channel. It is possible to use the color-difference channels of the key signal interface for other purposes. Any such information does not form part of the key signal.

Note 1: In facilities employing multiple video formats, it is possible that the key video signal and fill video signal will be of different formats. The intent of this practice is that the key signal be constituted according to the equipment through which it will pass. For example, a key signal which is to be handled by digital or analog equipment would conform to specifications for a specific digital or analog video signal.

Note 2: Signals such as chroma keys (blue screen, green screen, etc., in film parlance) that are derived from limited-bandwidth color-difference signals will of necessity have lower bandwidth than the luminance channels of the signals from which they were derived. However, to maintain sample congruence with the luminance channel, chroma keys should be up-sampled to the sample rate of the luminance channel before distribution.

Note 3: Chroma keyers have traditionally used the opposite polarity to all other key signals. The method of implementation within equipment is at the discretion of the designer, but any key signal outputs from chroma keyers for external connection need to conform to this practice.

5 Synchronization and Timing

The specifications in this section are applicable to all video formats, analog and digital and component and composite. Specifications applicable only to legacy analog and composite digital formats are given in Annex A.

5.1 Synchronization

The key signal shall incorporate the same synchronizing elements as a video signal of the same format. The synchronizing element for digital component signals are the synchronizing elements End of Active Video (EAV) and Start of Active Video (SAV).

The synchronizing element for analog signals is composite sync (CVBS), while for digital composite signals it is the Timing and Reference signal (TRS).

5.2 Key to Fill Pixel Offset

In digital component systems, key samples shall be co-sited with the corresponding fill luminance or RGB samples with zero pixel offset relative to each other in either direction. In systems capable of sub-pixel changes in position, the differential positioning between key and fill should be maintained to a spatial shift of less than 1/25 of the width of a pixel either vertically or horizontally.

5.3 Timing

The key signal shall be timed coincident with its associated fill video signal.

6 Pre-Multiplied Alpha

In computer graphics (CG) applications, fills are typically modulated at their edges by the alpha or key signal in order to reduce or eliminate aliasing artifacts. This is referred to as pre-multiplied alpha or sometimes anti-aliased key. If these fills are then composited with the background using the equation of Section 4, dark fringes will appear at the edges of the CGI fill. Designers and users are advised that this can occur when CG images are composited in a video environment.

Annex A Legacy Formats (Normative)

A.1 Legacy Formats

Specifications and considerations relating to analog and digital composite formats have been moved to this Annex.

Note: Much composite analog equipment does not meet these specifications. Designers and users of systems using such equipment need to take this into account. The tolerances specified are those required to ensure no discernible error in the composited picture.

A.2 Synchronization

In analog composite formats, key signals shall include sync and burst.

Note: For composite analog systems, burst is an essential part of the synchronization information. Many devices that operate in the composite analog domain do not operate properly when burst is not present, even if the signal has no color-difference information.

A.3 Picture Phase

In analog systems, the horizontal picture phase requirement is typically achieved by adjustment of the latency of the key signal path relative to the fill signal path. In composite digital systems, key samples shall be co-sited with the corresponding fill samples.

A.4 Timing

In analog systems, the timing tolerance shall be ± 0 lines vertically and ± T/25 horizontally, where T is the inverse of the highest frequency of interest (the Nyquist interval). For standard definition systems, T/25 = 5 ns. In composite systems, both analog and digital, SC/H timing shall match that of the associated fill video signal.

A.5 Blanking

When the fill video is required to be opaque at the horizontal edges of the picture, use of transmission blanking width on the key signal may give rise to undesirable edge effects. Facilities using narrow blanking for video should use the same blanking width for key signals. Edge artifacts will be removed when transmission blanking is applied. Facilities using transmission blanking width may wish to use narrower blanking for key signals to avoid edge artifacts.

Annex B Bibliography (Informative)

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 170:2004) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 170M-2004). Documents with the same root number (e.g. 170) and publication year (e.g. 2004) are functionally identical.

SMPTE EG 28:1993 (Archived 2004), Annotated Glossary of Essential Terms for Electronic Production

SMPTE ST 170:2004 (Archived 2010), Television — Composite Analog Video Signal — NTSC for Studio Applications

SMPTE ST 274:2008, Television — 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

SMPTE ST 296:2012, 1280 x 720 Progressive Image 4:2:2 and 4:4:4 Sample Structure — Analog and Digital Representation and Analog Interface

SMPTE ST 2036-1:2009, Ultra High Definition Television — Image Parameter Values for Program Production

Recommendation ITU-R BT.601-7 (03/11), Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios

Recommendation ITU-R BT.656-6 (12/07), Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT.601

Recommendation ITU-R BT.1700 (02/05), Characteristics of Composite Video Signals for Conventional Analogue Television Systems