

Closed Caption en ISDB-Tb

¿Qué es el Closed Caption?

El Closed Caption¹ (CC), o subtítulo oculto, consiste en enviar en una porción no visible de la señal de video una transcripción textual de lo que van hablando los interlocutores, para ayudar a comprender el contenido de dicha señal a personas sordas, hipoacúsicas o con dificultades para captar la señal de audio. A diferencia de lo que sería el Open Caption que es siempre visible, el Closed Caption sólo se hace visible sobre la señal de video cuando el usuario del TV o dispositivo receptor lo activa, generalmente por medio de una tecla del control remoto con las siglas CC o SUB.

La norma ISDB-Tb adopta la idea del Closed Caption, presente ya desde las transmisiones analógicas, y lo lleva al terreno del video digital comprimido a través del agregado de un componente de datos asociado al programa que se quiere subtítular. En la terminología MPEG, este componente de datos es transmitido en un formato conocido como "Packetized Elementary Stream" o "PES", que en el caso del CC será del tipo 0x06, que significa que contiene datos, a diferencia de otros PES que pueden contener video, audio, etc. Este componente de datos estará identificado a su vez con un "PID" (Packet Identifier) establecido en la tabla PMT (Program Map Table) del programa correspondiente².

El laberinto normativo

Para definir completamente el Closed Caption según la norma ISDB-Tb (también conocida como ISDB-T Internacional) es necesario recorrer un intrincado laberinto de normas, reportes técnicos y guías de operación, partiendo por las normas definidas en el estándar Brasileño, continuando con las referencias que hacen éstas a las normas Japonesas y terminando finalmente con los Reportes Técnicos y Guías de Operación de ambas normas.

Podemos resumir el conjunto de normativas involucradas en el Closed Caption ISDB-Tb a continuación:

Normas base:

- ABNT NBR 15606-1, Section 11.6: Subtitles and superimposed characters
- ARIB STD-B24, Volume 1, Part 3: Coding of Caption and Superimpose

Codificación de Caracteres:

- ABNT NBR 15606-1, Chapter 11.4: Character coding
- ARIB STD-B24, Volume 1, Part 2, Chapter 7: Character coding

¹ Generalmente identificado con el símbolo 

² En la terminología MPEG, el término "programa" corresponde a un video con su audio y datos asociados, mientras que en la norma ISDB-Tb este término se reemplaza por "servicio". Ambos términos tienen el mismo significado a lo largo de esta nota.

Guías de Operación y Reportes Técnicos:

- ABNT NBR 15608-3, Chapter 22: Multiple captions transmission
- ABNT NBR 15608-3, Annex A: Caption PES structure
- ARIB TR-B14, Volume 3, Section 2, Chapter 4: Operation of caption and superimpose encoding (Profile A)
- ARIB TR-B14, Volume 3, Section 4, Chapter 6: Operation of closed caption coding (Profile C)

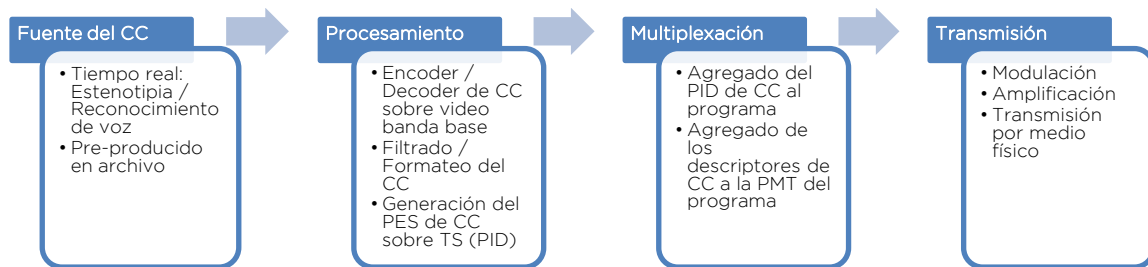
Transmisión de CC en video banda base:

- ANSI/CEA-608: Line 21 Data Services
- ANSI/CEA-708: Digital Television (DTV) Closed Captioning
- SPMTE 334-1: Vertical Ancillary Data Mapping of Caption Data and Other Related Data
- SPMTE 334-2: Caption Distribution Packet (CDP) Definition
- ARIB STD-B37: Structure and Operation of Closed Caption Data conveyed by Ancillary Data Packets

Como se desprende de lo anterior, el estudio a fondo del tema puede llevar bastante tiempo, y así y todo resultar un tanto complicado, debido a la cantidad de documentos a revisar para entender cada tema. Trataremos de resumir a continuación los aspectos más importantes de las mencionadas normas y su aplicación práctica en los sistemas de TV Digital ISDB-Tb.

La cadena de transmisión del CC

Partiremos definiendo los componentes en la cadena de transmisión del CC, desde la fuente que origina el CC, pasando por su procesamiento y su multiplexación, para ser finalmente transmitido a través de un medio físico, siendo lo habitual la transmisión a través del aire en los sistemas de TDT, aunque también podrá darse el caso de ser transmitido por otros tipos de redes³.



Fuente del CC

El CC podrá originarse básicamente de 2 maneras:

1. En tiempo real, generalmente para programas en vivo o que no han sido previamente subtítulos.
2. Pre-producido en archivo digital o cinta.

Para el primer caso, existen empresas dedicadas a brindar servicios de CC en tiempo real a los canales de TV; estas suelen emplear teclados de estenotipia para una rápida transcripción del texto hablado, que luego será enviado al canal de TV mediante un enlace de datos. Tradicionalmente se solían usar líneas telefónicas (para recibir el audio a transcribir y para el envío de los caracteres), hoy en día Internet es el medio más usado para esta tarea.

Otra opción para la transcripción automática del texto será usar un software de reconocimiento de voz, que permita transcribir en tiempo real el audio de la señal, el cual preferentemente se deberá

³ Redes satelitales, de cable, fibra óptica, etc.

tomar lo más limpio posible⁴ para un mejor reconocimiento, pudiendo también emplear patrones predefinidos para cada interlocutor del programa.

Para el segundo caso, se usará algún software de edición no lineal que permita agregar una pista de CC sincronizada con el video.

En cualquiera de los 2 casos, el CC deberá ser transportado en forma sincrónica con su señal de video asociada.

Procesamiento del CC

La forma de transportar el CC sobre señales analógicas (Video Compuesto o CVBS) está definida en la norma CEA-608 (anteriormente EIA-608). Originalmente esta norma hacía referencia a señales NTSC (525 líneas, 60 Hz), pero se ha adaptado su uso a señales PAL (625 líneas, 50 Hz), cambiando la línea 21 por otras⁵, y manteniendo el resto del sistema tal cual lo definido en dicha norma. Esta norma define la forma de transmitir caracteres de texto y de control en una línea de la porción de video no visible conocida como VBI (Vertical Blanking Interval) de forma tal que los datos transmitidos estarán intrínsecamente sincronizados con cada cuadro de video.

Existen dispositivos “Encoders de CC” que reciben la señal de Video Compuesto y los datos de CC (caracteres de texto y control) provenientes de alguna interfaz serial o IP, y generan a la salida una señal de video con el CC embebido en la línea configurada. A su vez, los dispositivos receptores (TV y STB) deberán poseer un “Decoder de CC” integrado para interpretar los datos presentes en la línea del CC, y mostrarlos a petición del televidente.

Para transportar el CC sobre señales digitales SDI-SD (Serial Digital Interface - Standard Definition) es posible hacerlo de 2 maneras: una es “digitalizando” la forma de onda analógica de la señal definida en la norma CEA-608. La otra, digital ya en forma nativa, es usando una porción específica de la señal SDI conocida como VANC (Vertical ANCillary data) la cual permite el envío de datos, nuevamente en forma sincrónica con los cuadros de video. Esta porción VANC sería el equivalente en video digital a la porción VBI en video analógico, con la ventaja de que permite transmitir una mayor cantidad de datos que lo que permitía el VBI.

Esta forma de transmitir el CC en el VANC de la señal SDI está definida en la norma SMPTE 334-1, y es aplicable tanto para señales SD como HD, siendo aun posible que los datos propiamente dichos (caracteres de texto y control) sigan estando en el formato definido por la norma CEA-608.

Pensando ya en sistemas puramente digitales, se definió la norma CEA-708 (anteriormente EIA-708), que extiende el alfabeto de caracteres y sus posibles formatos con respecto a su antecesora CEA-608. El transporte del CC en formato CEA-708 se debe previamente encapsular en CDP (Caption Distribution Packets) definidos en la norma SMPTE 334-2, para ser luego insertados en el VANC de la señal SDI según lo estipulado por la norma SMPTE 334-1⁶.

Cabe hacer una mención al conjunto de normas “paralelas” a las descritas anteriormente, que son usadas en Japón para la generación y el transporte del CC sobre video banda base, en particular la norma B37. Básicamente esta norma define una manera similar de transportar el CC sobre el VANC de señales SDI-SD/HD, pero con un formato distinto al definido en las normas CEA-608/708 y SMPTE 334-1/2, lo cual hace que ambos conjuntos de normas sean incompatibles entre sí. Por este motivo, de aquí en más sólo nos referiremos a las normas mencionadas en los párrafos anteriores para la transmisión del CC sobre el VANC de señales SDI, dado que su uso es el que predomina en nuestra región.

Una vez insertado el CC sobre la señal de video, el próximo paso de la cadena será que un equipo decodifique los caracteres de texto y control, y los “traduzca” lo más fielmente posible a lo

⁴ Proveniente de un submaster de la mezcladora de audio.

⁵ En Argentina (sistema PAL-N) se usó la línea 18 porque coincide en tiempo con la línea 21 de NTSC y se vio que los dispositivos preparados para NTSC funcionaban sin problemas en PAL-N cuando se enviaba el CC en la línea 18. En otros países con otros sistemas PAL se usó la línea 22.

⁶ Para señales SDI de 50Hz (25/50 cuadros por segundo), éste sería el único formato admitido por la norma SMPTE 334-1 para el envío de CC.

definido en las normas ISDB-Tb para generar el PES de CC. Este proceso no será sencillo ni directo, dada las diferencias inherentes en las definiciones y el formato de los caracteres entre las distintas normas.⁷ Es por ello que comúnmente se procederá a “filtrar” todos (o gran parte de) los caracteres de control provenientes de la señal de video, para luego generarlos e introducirlos en forma nativa con el dispositivo que genere el PES de CC, esta vez ya en forma nativa según lo definido en las normas ISDB-Tb.

Al dispositivo que recibe como entrada los caracteres de texto y control, y genera a su salida el PES de CC sobre un Transport Stream (en un PID determinado), lo llamaremos de aquí en más “Cliente Closed Caption”.

Es importante que el Cliente Closed Caption le permita al Broadcaster la mayor versatilidad posible en cuanto a las opciones de formateo del CC según las posibilidades que brinda la norma ISDB-Tb, puesto que esto le permitirá adaptar la funcionalidad, la apariencia y la estética del mismo a cada caso particular. Por ejemplo, se podrán modificar la posición en la pantalla del cuadro de CC, la cantidad de líneas a mostrar, el tamaño y formato de los caracteres, el espaciado entre caracteres y entre líneas, los colores del texto y del fondo (incluyendo transparencias), etc.

Una variante que se puede utilizar en sistemas ISDB-Tb puros, en los cuales no exista la infraestructura del Encoder / Decoder de CC, y/o que no se requiera la presencia del CC sobre el video banda base en ningún punto de la cadena de transmisión, consistirá en enviar directamente los caracteres de CC al Cliente Closed Caption, ya sea operando directamente sobre el mismo Cliente CC (mediante un teclado y una interfaz de usuario), o bien, enviando los datos mediante algún vínculo serial o IP, incluso a través de Internet. De esta manera, el Cliente CC generará el stream de CC en norma ISDB-Tb, que contendrá los paquetes PES que alimentarán al Multiplexor, en un solo paso, sin tener que embeber y desembeber el CC sobre la señal de video.

En este caso podría presentarse una limitación al querer sincronizar el CC con el video a nivel de cuadro, dado que ambos streams estarán llegando en paralelo al Multiplexor, sin una base de tiempo en común. En estos casos, se podrá dotar al Cliente Closed Caption de una entrada de referencia para tomar de allí la base de tiempo del video comprimido, transmitida en los PCRs (Program Clock Reference) del programa de referencia, para el estampado de los campos PTS (Presentation Time Stamp) de los paquetes PES del stream de CC⁸. Adicionalmente se le podrá incorporar un retardo ajustable para compensar cualquier demora que pueda existir en el camino del video que no esté presente en el camino del CC, por ejemplo, el retardo causado por el Encoder de video.

De cualquier forma, este ajuste en la mayoría de los casos no requiere precisión a nivel de cuadro, sobre todo en programas en vivo, dado que por la naturaleza del CC en tiempo real, siempre existirá una demora de la persona que tipea el texto o del software que reconoce el audio, lo cual a veces puede incluso compensar el retardo del Encoder de video, de forma tal que ambos streams lleguen al Multiplexor en forma aproximadamente sincronizada sin mayores ajustes.

El camino de la señal de video (y su audio asociado) deberá pasar siempre por un Encoder de video y audio, que tomará como entrada la señal de video y audio en banda base, ya sea analógica o digital SD/HD, y generará a su salida un Transport Stream con los PES de video y audio en formato digital comprimido, cada uno con su correspondiente PID. Este Transport Stream llegará a una de las entradas del Multiplexor.

El Cliente Closed Caption podrá ser una entidad independiente al Encoder de video, por ejemplo, un software corriendo sobre un servidor, o bien podrá estar integrado al propio Encoder de video, en cuyo caso el sincronismo con el video estará asegurado.

En cualquier caso, el Cliente Closed Caption generará el PES de CC en un PID asignado sobre su Transport Stream de salida, que también llegará a otra entrada del Multiplexor.

⁷ Cabe hacer mención especial en este punto a la importancia que tendrá trabajar con equipos y sistemas preparados para la norma ISDB-T Internacional (ISDB-Tb) y no los diseñados específicamente para la norma ISDB-T japonesa, dado que los alfabetos y las tablas de codificación de los caracteres son distintos entre ambas normas.

⁸ En la práctica se pudo observar que los dispositivos receptores no requieren que el PTS del CC esté sincronizado con el PCR del programa, sino que lo muestran inmediatamente luego de recibir cada paquete de datos de CC.

Multiplexación del CC

Una de las tareas del Multiplexor será incorporar el PID de CC al programa que se quiere subtítular. Para esto, el Multiplexor debe permitir agregar componentes externos a un programa, o bien, armar el programa manualmente desde cero, agregando cada uno de sus componentes (video, audio, datos).

La otra tarea del Multiplexor será generar los descriptores necesarios en la tabla PMT del programa en cuestión para señalar la presencia del stream de CC asociado a dicho programa. Para esto, será necesario agregar en la mencionada tabla PMT dos descriptores que dan información sobre el stream de CC:

- o Stream Identifier descriptor
- o Data Component descriptor

Estos descriptores deberán estar definidos a nivel del Elementary Stream de CC dentro de la tabla PMT, es decir, dentro del loop de descriptores del componente de CC.

El primer descriptor sirve para identificar fácilmente el tipo de componente dentro de un programa, y diferenciarlo del resto de los componentes del programa (video, audio, u otros tipos de datos). Tiene un único campo llamado "Component Tag", y según la norma ISDB-Tb, el rango de valores para un stream de CC de un servicio de tipo Full-Seg (SD/HD) van del 0x30 al 0x37, siendo 0x30 el valor usado para el subtítulo principal del programa y el resto de los valores para subtítulos secundarios, por ejemplo, en otros idiomas. Para servicios de tipo One-Seg, el valor que deberá tener este campo será 0x87 (un único subtítulo).

El segundo descriptor detalla las características del componente de datos que corresponden al Elementary Stream de CC.

El primer campo de este descriptor, llamado "Data Component ID", identifica el método de codificación de datos usado. Su valor para los streams de CC está fijado por norma en 0x0008 para servicios Full-Seg (SD/HD) y en 0x0012 para servicios One-Seg. Al tomar cualquiera de estos valores, los siguientes 2 parámetros de este descriptor, agrupados bajo el nombre "additional_data_component_info" darán información extendida sobre este componente de datos.

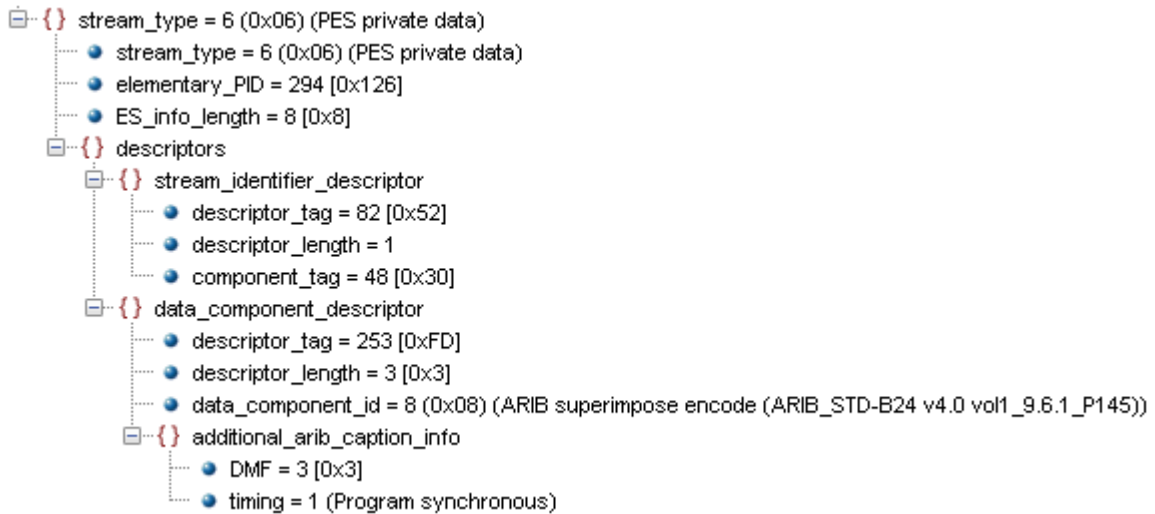
El primer parámetro, llamado "DMF" (Display Mode Flag), indica la forma de presentación del CC cuando se recibe el stream (primeros 2 bits), y cuando se graba y reproduce el mismo (tercer y cuarto bit). El valor predeterminado de este parámetro para el stream correspondiente al subtítulo principal de un servicio de tipo Full-Seg (SD/HD) está fijado por norma en 0x3 = 0011_b, e indica en este caso que éste será el idioma de subtítulos que deberá mostrarse en forma automática para ese programa⁹. Para servicios de tipo One-Seg, este parámetro está fijado por norma en 0xA = 1010_b lo cual significa que en cualquier caso la presentación del CC será seleccionable por el usuario.

El segundo parámetro, llamado "Timing", indica el tipo de sincronización que tendrá el stream de CC para mostrarse en la pantalla. El valor de este parámetro deberá ser 01_b lo cual significa que este stream estará sincronizado con su programa asociado.

La definición del Elementary Stream de CC y sus descriptores para el subtítulo principal de cualquier servicio de tipo Full-Seg (SD/HD), deberá ser como se muestra en la siguiente figura¹⁰:

⁹ Esta definición que hace la norma para este caso particular, rompe un poco con el esquema general de uso de este parámetro.

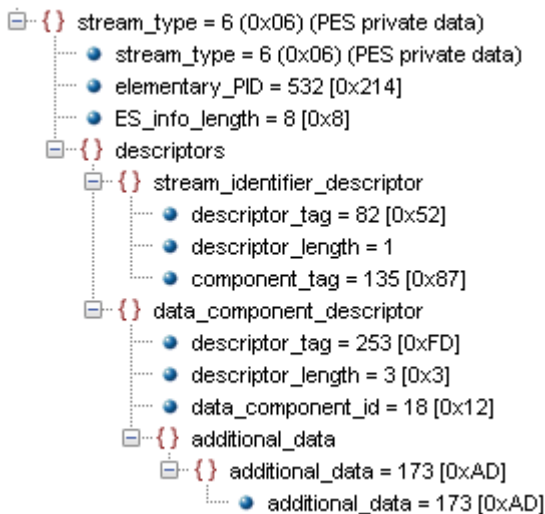
¹⁰ En este ejemplo, el PID del stream de CC es el 294 = 0x0126



Los valores que deberán tener estos descriptores desde el punto de vista de la configuración de un Multiplexor, para cualquier servicio Full-Seg, se resumen en la siguiente tabla:

| Stream_Type | 0x06 | ISO/IEC 13818-1 PES packets containing private data | |
|-------------------------------------|------|---|-------------|
| Stream identifier descriptor | | | |
| | | Desc. Tag | 0x52 |
| | | Desc. Length | 1 |
| | | Hex Data | 0x30 |
| Data component descriptor | | | |
| | | Desc. Tag | 0xFD |
| | | Desc. Length | 3 |
| | | Hex Data | 0x 00 08 3D |

La definición del Elementary Stream de CC y sus descriptores para el subtítulo de cualquier servicio de tipo One-Seg, deberá ser como se muestra en la siguiente figura¹¹:



¹¹ En este ejemplo, el PID del stream de CC es el 532 = 0x0214

Los valores que deberán tener estos descriptores desde el punto de vista de la configuración de un Multiplexor, para cualquier servicio One-Seg, se resumen en la siguiente tabla:

| | | | |
|-------------------------------------|-------------|--|-------------|
| Stream_Type | 0x06 | ISO/IEC 13818-1 PES packets containing private data | |
| Stream identifier descriptor | | | |
| | | Desc. Tag | 0x52 |
| | | Desc. Length | 1 |
| | | Hex Data | 0x87 |
| Data component descriptor | | | |
| | | Desc. Tag | 0xFD |
| | | Desc. Length | 3 |
| | | Hex Data | 0x 00 12 AD |

Dependiendo de las facilidades que brinde cada Multiplexor, es probable que algunos de los datos mostrados en las tablas de arriba no sea necesario ingresarlos a mano. Por ejemplo, los multiplexores Videoswitch calculan automáticamente el campo “Descriptor Length” de cada descriptor, de acuerdo a la cantidad de datos ingresados en el campo “Hex Data”.

Una vez que el programa esté definido en forma completa, el Multiplexor generará su BTS (Broadcast Transport Stream) de salida para alimentar al Modulador/Excitador, el cual se encargará de modular la señal según lo especificado por la norma ISDB-Tb para adaptarla al medio físico de transmisión, sin realizar modificaciones en los datos del programa. Luego esta señal será amplificada y transmitida por el medio físico.

Cuando la señal llegue a un dispositivo receptor ISDB-Tb, luego de demodular y decodificar la señal, éste podrá verificar en la PMT del programa subtulado la presencia de un componente de datos del tipo CC, y como consecuencia de esto, permitirá mostrar u ocultar el CC para ese servicio cuando el televidente así se lo ordene desde su control remoto.

Conclusiones

Para cumplir con todas las normativas del estándar ISDB-Tb que permiten la correcta generación, procesamiento y transmisión del Closed Caption, es necesario conocer y estudiar a fondo las distintas normativas involucradas. Para lograr una implementación exitosa del Closed Caption en señales ISDB-Tb, será preciso emplear equipos y sistemas que estén diseñados para el manejo del Closed Caption en cada una de sus etapas, de acuerdo a los requerimientos impuestos por las normas mencionadas.



Videoswitch S.R.L.
 Uspallata 3579
 Ciudad de Buenos Aires
 C1437JCS - Argentina
 Tel: +54-11 4911-5551
<http://www.videoswitch.tv>
info@videoswitch.tv