

RECOMENDACIÓN UIT-R BS.1688

Sistema de sonido en banda base y codificación de la fuente de audio en las interfaces de distribución de aplicaciones de generación digital de imágenes en pantalla gigante

(Cuestión UIT-R 15/6)

(2004)

La Asamblea de Radiocomunicaciones de la UIT,

considerando

- a) que la Recomendación UIT-R BT.1662 – Cadena de referencia general y gestión del margen de tratamiento ulterior de la esencia del programa en aplicaciones de imágenes digitales en pantalla grande, especifica una cadena de referencia general para aplicaciones de imágenes digitales en pantalla grande (LSDI) típicas y que sus principios se aplican igualmente a la esencia de la imagen y a la esencia del sonido;
- b) que una característica importante de la cadena de referencia general es la identificación de los principales bloques funcionales en una cadena genérica y de las interfaces entre ellos;
- c) que para la distribución de aplicaciones LSDI, es necesario identificar el sistema de sonido en banda base de referencia así como el método de codificación de la fuente¹ utilizado para la distribución de programas de audio:
 - en la interfaz entre la copia maestra de distribución² del programa LSDI y el canal de distribución;
 - en la interfaz entre el canal de distribución y el sistema de audio en el lugar de presentación;
- d) que la Recomendación UIT-R BT.1666 – Requisitos de usuario para aplicaciones de imágenes digitales en pantalla grande destinadas a su presentación en salas de cine, identifica la calidad global subjetiva de imagen y sonido que debe proporcionarse a nivel de calidad de funcionamiento más elevado de la familia de aplicaciones destinadas para su presentación LSDI en salas de cine, o similares;
- e) que la Recomendación UIT-R BS.775 – Sistema de sonido estereofónico multicanal con y sin acompañamiento de imagen, especifica un sistema de sonido de referencia 5-1 con cinco canales³ más un canal de efectos de baja frecuencia opcional, como el nivel más elevado en una jerarquía de sistemas de sonido multicanal que va de 1/0 (monofónico) hasta 3/2;

¹ La UIT define «codificación de la fuente» como la codificación de la señal digital original (respectivamente vídeo, audio o datos) en representación con velocidad binaria reducida antes de aplicar la protección contra los bit erróneos en el canal de distribución.

² El término «copia maestra de distribución» se utiliza aquí para denominar la copia maestra del programa obtenida a partir de la copia maestra terminada original del programa una vez adaptada (en contenido y calidad) al medio de distribución previsto.

³ Los cinco canales principales de sonido son los canales frontal izquierdo, frontal central y frontal derecho más los canales de sonido panorámico izquierdo y derecho.

- f) que el sistema de sonido de referencia 5-1 especificado en la Recomendación UIT-R BS.775 puede proporcionar la calidad de sonido identificada en la Recomendación UIT-R BT.1666;
- g) que en algunos casos es importante optimizar el programa de sonido (por ejemplo, estereofónico de 2 canales, matriz panorámica de 2 canales, o 5-1 canales) para la presentación en un lugar específico (por ejemplo, una sala pequeña o una sala grande);
- h) que para la distribución del contenido LSDI con canales limitados en anchura de banda (por ejemplo, los de satélite), una disminución de la velocidad binaria de la señal de audio reducirá los costes de transmisión;
- j) que el UIT-R ha llevado a cabo pruebas de audio subjetivas, ha evaluado los resultados de dichas pruebas realizadas por otros organismos y ha documentado, en la Recomendación UIT-R BS.1548 – Requisitos de usuario para sistemas de codificación audio en radiodifusión digital, la calidad que pueden proporcionar diversos algoritmos de codificación de la fuente a ciertas velocidades binarias;
- k) que las aplicaciones LSDI pueden beneficiarse de la utilización de los equipos y sistemas desarrollados para dar soporte a la televisión digital;
- l) que el audio digital en la radiodifusión y en otras áreas profesionales utiliza una velocidad de muestreo de 48 kHz sin aplicar acentuación dependiente de la frecuencia,

recomienda

- 1** que para aplicaciones LSDI el sistema de sonido en banda base digital de referencia en la interfaz entre la copia maestra de distribución de programas y el canal de distribución y en la interfaz entre el canal de distribución y el sistema de audio en la sala de presentación se base en el sistema de sonido de referencia jerárquico indicado en la Recomendación UIT-R BS.775⁴ que especifica una jerarquía que va desde la monofonía hasta 5-1 canales de sonido, pasando por una estereofonía de 2 canales;
- 2** que cada canal del sistema de sonido digital de referencia utilice una representación de señal MIC con una velocidad de muestreo de 48 kHz y un mínimo de 16 bits/muestra sin acentuación, como se especifica en la Recomendación UIT-R BS.1548, a menos que el suministrador del programa y el destinatario hayan llegado a un previo acuerdo en el sentido de utilizar una velocidad de muestreo o una representación de señal diferentes;
- 3** que en los casos en que el lugar de presentación LSDI recibe un programa con un número de canales de sonido superior al número de canales que sus equipos pueden presentar, se realice un mezclado descendente basándose en las especificaciones indicadas en el Anexo 4 a la Recomendación UIT-R BS.775;

⁴ Si bien en algunos casos la «interfaz» puede ser física, en muchos casos esta interfaz puede ser la idea de un ente conceptual incorporado al equipo sin que exista de hecho ninguna interfaz física.

4 que en los casos en que el lugar de presentación LSDI recibe un programa con más de dos canales y el equipo de presentación sólo puede aceptar dos canales en la interfaz física en un sistema de audio que contiene un decodificador de sonido panorámico de matriz⁵, se efectúe un mezclado descendente del audio basándose en las ecuaciones del Anexo 1,

5 que en los casos en que el lugar de presentación LSDI recibe un programa con un número de canales de sonido más pequeño que el número de canales que sus equipos pueden presentar, se realice una conversión ascendente basándose en las especificaciones que figuran en el Anexo 5 a la Recomendación UIT-R BS.775;

6 que en los casos en que el lugar de presentación LSDI recibe un programa con dos canales, se realice una conversión ascendente del audio con un decodificador de sonido panorámico de matriz y que los suministradores del programa sean conscientes de que es probable que los programas de dos canales se reproduzcan de esta forma;

7 que los suministradores de programas se aseguren de que el programa de audio distribuido es adecuado para el lugar previsto⁶, y consideren seriamente la posible necesidad de distribuir otras versiones del programa que quizá contengan distintos números de canales, optimizados para otros lugares de presentación previstos;

8 que los suministradores de programas sean conscientes de que las interfaces físicas disponibles más comunes en las salas de cine están adaptadas a una señal codificada panorámica de matriz LtRt de dos canales y que la provisión de esta señal en paralelo con cualquier distribución 5-1 debe ser considerada seriamente a menos que se sepa con certeza que todas las salas pueden presentar satisfactoriamente un programa 5-1 elaborado para obtener resultados satisfactorios cuando se reproduce en los lugares previstos,

recomienda también

1 que para la distribución de programas en sala de cine o similares, se prefiera el formato de audio 5-1, que es con el que se obtiene el mejor resultado dentro de la jerarquía especificada en la Recomendación UIT-R BS.775;

2 que el audio se distribuya en la representación MIC en banda base original, a menos que lo prohíban las especificaciones relativas al medio de distribución concreto;

⁵ Muchas salas cinematográficas poseen hoy en día un «decodificador» de sonido panorámico de matriz» que puede efectuar un mezclado ascendente o «decodificar un programa de audio de dos canales presentado mediante un altavoz alimentado por los canales izquierdo, central, derecho y panorámico.

⁶ Si el lugar previsto es una sala cinematográfica o similar, es importante que el suministrador del programa compruebe el mezclado en una sala grande; los mezclados creados y utilizados en salas pequeñas con altavoces de campo cercano puede que no sean adecuados para la presentación en salas grandes.

3 que en los casos en que se utilice codificación de la fuente de audio en el canal de distribución a una sala de cine o similar, el sistema utilizado para la codificación de la fuente de audio sea el AC-3, como indica el Anexo 2 a la Recomendación UIT-R BS.1196⁷, salvo que se haya llegado a un acuerdo previo entre el suministrador del programa y el destinatario en el sentido de emplear otro sistema de codificación de la fuente o que se utilice otro sistema de codificación de la fuente en paralelo con el sistema antes especificado;

4 que en los casos en que se utilice codificación de la fuente de audio en el canal de distribución de una aplicación LSDI donde la eficacia de la velocidad binaria reviste una importancia fundamental, el sistema utilizado para la codificación de fuente de audio sea el ISO/CEI IS 13818-7 (AAC);

5 que si futuras contribuciones al UIT-R muestran que los sistemas de sonido que utilizan canales adicionales o tecnología más avanzada pueden proporcionar una ventaja significativa con respecto al sistema de sonido básico de 5-1 canales o con respecto a los sistemas de codificación de la fuente especificados en esta Recomendación, se considere la introducción de modificaciones en la presente Recomendación.

NOTA 1 – La Norma ISO/CEI IS 13818-7 está disponible en versión electrónica en la siguiente dirección: <http://www.iso.org/itu>.

Anexo 1

Mezclado descendente de una señal LtRt codificada de sonido panorámico de matriz 3/1

Las siguientes ecuaciones especifican la forma de codificar un formato de sonido 3/1 a una señal panorámica codificada de matriz compatible estereofónica. Estas ecuaciones se han utilizado en las salas cinematográficas y en los vídeos domésticos desde 1980.

$$L_t = \text{Izquierdo} + 0,707 \text{ Central} - 0,707 J (\text{Panorámico izquierdo} + \text{Panorámico derecho})$$

$$R_t = \text{Derecho} + 0,707 \text{ Central} + 0,707 J (\text{Panorámico izquierdo} + \text{Panorámico derecho})$$

donde J representa un desplazamiento de fase de 90°.

NOTA 1 – Para evitar la sobrecarga si este mezclado descendente se lleva a cabo en el dominio digital debe realizarse o un ajuste escalonado o deben aplicarse otros medios.

⁷ De acuerdo con la Recomendación UIT-R BS.1548, el sistema AC-3 proporciona una calidad de radiodifusión con una velocidad binaria de 96 kbit/s por canal y el sistema ISO/CEI IS 13818-7 proporciona una calidad de radiodifusión con una velocidad binaria de 72 kbit/s por canal.