

## RECOMENDACIÓN UIT-R BS.1548

**Condiciones de usuario para sistemas de codificación  
audio en radiodifusión digital**

(Cuestión UIT-R 208/10)

(2001)

La Asamblea de Radiocomunicaciones de la UIT,

*considerando*

- a) que el sistema de sonido multicanal, solo o con imagen asociada, es el objeto de la Recomendación UIT-R BS.775;
- b) que los sistemas de codificación audio para la radiodifusión de televisión terrenal digital son el objeto de la Recomendación UIT-R BS.1196;
- c) que la codificación audio de baja velocidad binaria es el objeto de la Recomendación UIT-R BS.1115;
- d) que el sistema de codificación considerado en la Recomendación UIT-R BS.1115 permite modos de codificación monofónico y estereofónico de dos canales;
- e) que es preciso ofrecer la mayor calidad básica posible de audio y de imagen estereofónica en sistemas de sonido para radiodifusión sonora y de televisión, lo que supone, en general, evitar la deformación del material de origen;
- f) que es necesario considerar minuciosamente las condiciones de interfuncionamiento y explotación de la red con conexiones de programas, como los enlaces de contribución y distribución;
- g) que es necesario considerar minuciosamente el interfuncionamiento con los equipos de audio multicanal de uso público, como los decodificadores matriciales de sonido ambiental o los decodificadores multicanal discretos;
- h) que la integración de un sistema de sonido multicanal en un servicio de radiodifusión existente con las condiciones de la Recomendación UIT-R BS.1115 obliga a considerar la compatibilidad con los receptores existentes, para mantener el servicio;
- j) que, en un plano más general y teniendo en cuenta las numerosas aplicaciones de estos sistemas, es necesario especificar todas las condiciones técnicas, de calidad y operacionales;
- k) que las prestaciones de los sistemas de codificación audio multicanal según la Recomendación UIT-R BS.1196 dependen en buena parte de la configuración de explotación del sistema (velocidad binaria, utilización de sistemas de pre-matrización, utilización de la codificación compuesta, etc.);
- l) que algunos servicios de radiodifusión ya utilizan o han especificado la utilización de los sistemas descritos en la Recomendación UIT-R BS.1196;
- m) que, visto lo anterior, los organismos de radiodifusión necesitan urgentemente toda la información necesaria para definir todos los parámetros de codificación disponibles de los sistemas que figuran en la Recomendación UIT-R BS.1196;
- n) que, en lo posible, debe evitarse la introducción de sistemas incompatibles con características de prestación similares;

o) que es importante que los organismos de radiodifusión que todavía no están ofreciendo el servicio puedan elegir el sistema que sea más adecuado para su aplicación y que tenga la mejor eficacia en relación con los costos,

*recomienda*

1 que los sistemas de codificación audio de radiodifusión sonora y de televisión digital para aplicaciones de contribución y distribución sean definidos según las condiciones del Anexo 1;

2 que los sistemas de codificación audio para la radiodifusión sonora y de televisión digital para aplicaciones de emisión sean definidos según las condiciones del Anexo 2.

## ANEXO 1

### Condiciones de contribución y distribución

Los sistemas de codificación audio para radiodifusión sonora y de televisión digital en aplicaciones de contribución y distribución deben cumplir las siguientes condiciones.

#### 1 Condiciones de servicio

##### 1.1 Configuraciones de canales

Los sistemas deben soportar las siguientes configuraciones de canales para servicios de audio, según las necesidades de las aplicaciones (véase la Recomendación UIT-R BS.775 – Sistema de sonido estereofónico multicanal con y sin acompañamiento de imagen):

1 canal	1/0	Monofónico
2 canales	2/0	Izquierda, derecha
3 canales	3/0 2/1	Izquierda, derecha, centro Izquierda, derecha/ambiente
4 canales	3/1 2/2	Izquierda, derecha, centro/ambiente Izquierda, derecha/ambiente izquierda, ambiente derecha
5 canales	3/2	Izquierda, derecha, centro/ambiente Izquierda, ambiente derecha

También deben soportar un canal facultativo de intensificación de bajas frecuencias (LFE, *low frequency enhancement*).

Por otra parte, para aplicaciones de contribución, puede ser necesario en algunos casos transmitir programas creados en otros formatos, por ejemplo 3/4, lo que supone que el sistema de codificación ha de permitir la integración de otros canales de alta calidad.

## 1.2 Atribución flexible de canales

El sistema dispondrá de una cadena de bits que suministre la información de identificación necesaria para la señalización y el control de las configuraciones de sonido. Es necesario que el sistema de transmisión pueda hacer una conmutación dinámica entre las configuraciones de canales mencionadas en el § 1.1.

## 1.3 Datos auxiliares

El sistema de codificación audio ha de permitir la transmisión de datos auxiliares. Las distintas clases de información anexa son: control de gama dinámica, control de sonoridad, datos de usuario, y otros metadatos necesarios para el codificador de emisión que codificará la señal de audio entregada finalmente al consumidor.

# 2 Condiciones de prestaciones

## 2.1 Calidad de audio

### 2.1.1 Calidad de audio básica

Es importante que la calidad del sonido reproducido después de la cascada de contribución/distribución de referencia (cinco códecs de contribución y tres códecs de distribución en tándem) no pueda distinguirse, subjetivamente, de la fuente en la mayor parte del material de programas de audio. Si se utiliza el método de triple estímulo doblemente ciego con referencia oculta, descrito en la Recomendación UIT-R BS.1116 – Métodos para la evaluación subjetiva de pequeñas degradaciones en los sistemas de audio incluyendo los sistemas de sonido multicanal, la nota media ha de ser superior a 4,5 en una escala de degradación de 5 notas, para los oyentes en general, en la posición de escucha de referencia. Ninguna nota podrá ser inferior a 4 (véase la Recomendación UIT-R BS.775).

NOTA 1 – El intervalo de confianza (barra de error) para la nota de apreciación media de un códec y de un elemento es la gama por encima y por debajo de la nota media general, donde hay una determinada probabilidad (generalmente un 95%) de que se encuentre la nota efectiva. Un códec y un elemento pueden tener una nota efectiva tan baja como el límite inferior del intervalo de confianza con respecto a la nota general. Para poder hacer una evaluación significativa de las prestaciones esperadas de códecs en cascada, es necesario que el intervalo de confianza asociado con las notas medias de cada uno de los códecs sea aproximadamente igual o menor a la diferencia entre las notas que se comparan.

NOTA 2 – La cascada de contribución/distribución, si está en tándem con el códec de emisión, no debe afectar de forma significativa la calidad de funcionamiento con respecto a la calidad de audio básica del códec de emisión. Es necesario efectuar ulteriores estudios para elaborar una especificación precisa.

NOTA 3 – Los parámetros de calidad objetiva de audio para las aplicaciones de contribución/distribución pueden incorporarse ulteriormente conforme a la Recomendación UIT-R BS.1387.

### 2.1.2 Definición de cuantización

Es necesaria una definición mínima de 18 bits para distribución y, preferentemente, 20 bits o más para contribución.

### 2.1.3 Frecuencia de muestreo

Con arreglo a la Recomendación UIT-R BS.646 – Codificación en la fuente de las señales de sonido digitales en los estudios de producción de radiodifusión, la frecuencia de muestreo será de 48 kHz.

### 2.1.4 Anchura de banda

Canales audio principales: 20-20 000 Hz.

Canal LFE: 15-120 Hz.

### 2.1.5 Acentuación

El sistema de codificación audio no puede tener acentuación.

### 2.1.6 Posibilidad de operación en tándem

La posibilidad de operación en tándem es diferente según la aplicación, de acuerdo con el siguiente Cuadro:

Distribución	3 códecs en cascada
Contribución	5 códecs en cascada

Estas cifras reflejan los resultados de experiencias anteriores de evaluación de sistemas de radiodifusión de sonido de dos canales (véase la Recomendación UIT-R BS.1115 – Codificación del sonido a baja velocidad binaria) y no son necesariamente representativas de la situación real de la radiodifusión sonora y de televisión. Es necesario disponer de mayor información para elaborar especificaciones concretas sobre este particular.

### 2.1.7 Posibilidad de segundo tratamiento

La aplicación determina, en buena parte, las posibilidades de segundo tratamiento que son necesarias. Puede hacerse un tratamiento de interferencia, desvanecimiento y control de la gama dinámica en distribución. Es más difícil determinar las posibilidades de tratamiento de las señales entre enlaces de contribución.

## 2.2 Retardo de codificación

El retardo de codificación debe ser el mismo para todos los canales de un programa. Ha de ser tan bajo como sea posible, teniendo en cuenta las prestaciones de codificación (importancia de la reducción de velocidad binaria) que son necesarias. Cuando se trata de sonido de televisión, el retardo de audio y el retardo de vídeo deben coincidir. Preferentemente, el codificador de audio debe producir tramas de audio codificadas (unidades de acceso) que coincidan exactamente con el intervalo de tiempo de la trama de vídeo correspondiente.

## 2.3 Compensación de errores

En la cadena de bits de audio debe incorporarse un mecanismo que permita al decodificador identificar errores de canal residuales y tomar las medidas apropiadas para su ocultación.

## 2.4 Tiempo de recuperación

El tiempo de recuperación ha de ser lo más reducido posible. Si es estimado por unidades de acceso de audio (AAU, *audio access unit*), el tiempo de recuperación será, como máximo, de unas cuantas AAU, y la señal de audio debe ser recuperada, preferentemente, al recibir la primera unidad AAU sin errores.

# 3 Condiciones de funcionamiento y operacionales

## 3.1 Velocidad binaria y principio de codificación

Para los enlaces de distribución y contribución, la Recomendación UIT-R BS.1115 propone la utilización del sistema de Capa II Organización internacional de normalización (ISO)/Comisión electrotécnica internacional (CEI) IS 11172-3 con una velocidad binaria de 180 kbit/s por canal o superior. En distintas circunstancias, es posible que el sistema sea aplicado con una velocidad binaria superior o que se utilicen otros sistemas.

Puede ocurrir, por ejemplo, en las siguientes circunstancias:

- un margen de codificación superior que permita el tratamiento de las señales eventualmente introducidas en el intervalo de generación de codificación (no se hicieron pruebas ni verificaciones sobre este particular en el proceso de definición de la Recomendación UIT-R BS.1115);
- para reducir la velocidad binaria en los enlaces de contribución y distribución;
- para elevar el nivel de calidad;
- una sincronización apropiada y la conmutación con las señales de vídeo asociadas.

### 3.2 Codificación compuesta

En el material de programa de dos canales o multicanal muchas veces hay cierta correlación estadística entre canales. La función de codificación compuesta es tal vez una solución eficaz para reducir la irrelevancia o la redundancia entre canales, y elevar así la eficiencia de codificación. En algunos sistemas de codificación se aplican criterios de apreciación para eliminar parcialmente la irrelevancia entre canales, combinando dos o más canales en regiones de frecuencias donde el oído humano tiene muy poca capacidad para distinguir la dirección de la fuente. El inconveniente de esta técnica es que, generalmente, no se puede reintegrar la información de sonido en los canales originales en una fase posterior. Estos principios de codificación compuesta no deben ser utilizados para contribución ni para muchas aplicaciones de distribución.

## ANEXO 2

### Condiciones de emisión

Los sistemas de codificación audio para la radiodifusión sonora y de televisión digital en aplicaciones de emisión deben cumplir las siguientes condiciones.

## 1 Condiciones de servicio

### 1.1 Configuraciones de canales

El sistema debe soportar las siguientes configuraciones de canales para servicios de audio, según las necesidades de las aplicaciones (véase la Recomendación UIT-R BS.775):

1 canal	1/0	Monofónico
2 canales	2/0	Izquierda, derecha
3 canales	3/0 2/1	Izquierda, derecha, centro Izquierda, derecha/ambiente
4 canales	3/1 2/2	Izquierda, derecha, centro/ambiente Izquierda, derecha/ambiente izquierda, ambiente derecha
5 canales	3/2	Izquierda, derecha, centro/ambiente izquierda, ambiente derecha

También debe soportar un canal LFE facultativo.

## 1.2 Servicios de audio

Además del servicio de audio principal, el sistema puede ofrecer los siguientes servicios de audio asociados, según las necesidades de las aplicaciones: servicio de varios idiomas, servicios de audio para personas con discapacidades de visión o auditivas, y alguna capacidad para datos auxiliares (véase la Recomendación UIT-R BS.775) [ATSC, 1992]. El servicio de varios idiomas, que se ofrece con uno o varios canales independientes, permite distribuir un programa con los comentarios en varios idiomas. El servicio para personas con discapacidad de visión consiste generalmente en una descripción sonora del contenido de las imágenes, y el servicio para personas con discapacidad auditiva distribuye los diálogos limpios, sin música ni efectos especiales, o reduciendo la importancia de éstos, para facilitar la comprensión del diálogo. En los datos auxiliares puede incluirse el control de gama dinámica, el control de sonoridad o los datos de usuario.

Repartición de los distintos servicios:

- *Servicio principal* (todos los canales de un servicio principal son asignados al mismo programa, incluido el canal LFE facultativo).
- *Servicio(s) ampliado(s)*, que pueden ser:
  - *Servicios independientes* (para programas adicionales que son independientes del programa de servicio principal, como el comentario y otros servicios en dos o más canales; el Cuadro del § 1.1 indica las posibilidades de configuración de canales).
  - *Servicios alternativos* (para programas que reemplazan a uno o varios de los canales del servicio principal, por ejemplo un comentario en varios idiomas o adaptado a discapacidades auditivas).
  - *Servicios adicionales* (con canales que se añaden a los canales del servicio principal, por ejemplo un comentario, o canales adicionales para sistemas de sonido especial como el 3D TV).

En todos los sistemas de transmisión es necesaria una capa de sistema que pueda realizar operaciones de multiplexión. Por tanto, no es necesario que todos los servicios de audio mencionados anteriormente sean transportados por una sola cadena de bits.

## 1.3 Atribución flexible de los canales

Es necesaria una cadena de bits con información de identificación para la señalización y el control de las configuraciones de sonido. También es necesario que el sistema de transmisión pueda hacer una conmutación dinámica entre cualquier par de las configuraciones mencionadas en el § 1.1.

## 1.4 Datos auxiliares

La posibilidad de transmisión de datos auxiliares es una necesidad en los sistemas de codificación audio. Hay varios tipos de información anexa: control de la gama dinámica, control de sonoridad y datos de usuario.

# 2 Condiciones de prestaciones

## 2.1 Calidad de audio

### 2.1.1 Calidad de audio básica

Habitualmente, la organización de radiodifusión puede modificar la calidad de audio adoptando una velocidad binaria más o menos superior. El objetivo es reproducir el sonido de forma que la calidad después de la decodificación sea, subjetivamente, similar a la señal original para la mayor parte del

material de programa de audio. Cuando se utiliza el método de prueba de triple estímulo doblemente ciego con referencia oculta, descrito en la Recomendación UIT-R BS.1116, es preciso que las notas medias siempre sean superiores a 4, de conformidad con la Recomendación UIT-R BS.1116 y la escala de degradación de 5 notas, desde la posición de escucha de referencia. En la práctica, las condiciones comerciales algunas veces obligan a funcionar con una velocidad binaria inferior a la que sería necesaria para lograr este nivel de calidad. De todas maneras, el sistema ha de permitir que la organización de radiodifusión obtenga este nivel de calidad.

NOTA 1 – Los parámetros de calidad objetiva de audio para las aplicaciones de contribución/distribución pueden incorporarse ulteriormente conforme a la Recomendación UIT-R BS.1387.

### **2.1.2 Calidad de la imagen estereofónica**

En las configuraciones estereofónicas (dos canales o multicanal), debe conservarse la calidad de la imagen sonora del material de origen. En las configuraciones que tienen un canal central (3/0, 3/1, 3/2) debe mantenerse la estabilidad direccional de la imagen sonora frontal dentro de límites razonables, en un espacio de escucha que es más grande que el espacio de los sistemas habituales estereofónicos de dos canales. En las configuraciones con sonido ambiente (2/1, 2/2, 3/1, 3/2) la sensación de realidad espacial (ambiente) debe ser notablemente superior a la que permite el sistema estereofónico de dos canales (véase la Recomendación UIT-R BS.775).

### **2.1.3 Definición de cuantización**

La definición mínima necesaria debe ser de 16 bits.

### **2.1.4 Frecuencia de muestreo**

De conformidad con la Recomendación UIT-R BS.646, la frecuencia de muestreo será de 48 kHz.

### **2.1.5 Anchura de banda**

Canales de audio principales: 20-20 000 Hz.

Canal LFE: 15-120 Hz.

### **2.1.6 Acentuación**

En los sistemas de codificación audio no debe utilizarse acentuación.

### **2.1.7 Posibilidades de segundo tratamiento**

Las posibilidades de segundo tratamiento necesarias dependen, principalmente, de la aplicación. Para enlaces de emisión, puede limitarse a la ecualización y al control de la gama dinámica (por ejemplo, adaptar la gama dinámica del material de programa a los valores del entorno de escucha).

## **2.2 Retardo de codificación**

Todos los canales de un programa deben tener el mismo retardo de codificación. Para el sonido de televisión, el retardo de audio debe coincidir con el retardo de vídeo.

## **2.3 Compensación de errores**

En la cadena de bits de audio debe incorporarse un mecanismo que permita al decodificador identificar errores de canal residuales y tomar las medidas apropiadas para su ocultación.

## 2.4 Tiempo de recuperación

El tiempo de recuperación ha de ser lo más reducido posible. Si es estimado por AAU, el tiempo de recuperación será, como máximo, de unas cuantas AAU, y la señal de audio debe ser recuperada, preferentemente, al recibir la primera unidad AAU sin errores.

## 3 Condiciones de funcionamiento y operacionales

### 3.1 Compatibilidad

#### 3.1.1 Compatibilidad descendente (Recomendación UIT-R BS.775)

Es indispensable que un formato de cadena de bits multicanal pueda ser decodificado por distintos tipos de decodificadores más o menos complejos. También es necesario que el decodificador permita organizar una presentación con un número de canales inferior al número de canales transmitidos, según las posibilidades de reproducción del usuario, perdiendo únicamente el efecto de localización estéreo o multicanal.

Se han definido dos métodos que garantizan la compatibilidad descendente con sistemas receptores poco complejos. El primero supone la utilización del proceso matricial. En este caso, un receptor económico necesita únicamente los canales A y B, como en el sistema 2/0 que no utiliza la matriz de compatibilidad con sistemas anteriores. El segundo método puede aplicarse al sistema de entrega discreto 3/2. Una combinación digital de las señales entregadas, mediante ecuaciones, permite disponer del número necesario de señales. Cuando se trata de señales codificadas en origen de baja velocidad binaria, es posible hacer la adaptación descendente de las señales 3/2 antes del tramo de síntesis del proceso de decodificación (la sección de mayor complejidad).

#### 3.1.2 Compatibilidad con sistemas anteriores, conforme a la Recomendación UIT-R BS.1115

Esta condición se aplica a los casos en los que una aplicación existente monofónica/estereofónica basada en la Recomendación UIT-R BS.1115 es convertida para sonido multicanal, y es necesario mantener los servicios a los actuales destinatarios. La Recomendación UIT-R BS.1115 recomienda un sistema de Capa II ISO/CEI IS 11172-3 (para aplicaciones que sólo necesitan un sistema de codificación audio monofónico y estereofónico de baja velocidad binaria). Cuando se trata de sistemas que ya utilizan este tipo de codificación audio monofónico o estereofónico, la compatibilidad con sistemas anteriores para la codificación multicanal de baja velocidad binaria significa que un decodificador ISO/CEI IS 11172-3 debe estar en condiciones de decodificar correctamente la información estereofónica básica, que es la adaptación descendente de toda la información de audio de los canales de la fuente. Esta condición podrá cumplirse aplicando el método de difusión simultánea o el método de matrización.

##### *El método de difusión simultánea*

Una solución es seguir ofreciendo el servicio de canal conforme con la Recomendación UIT-R BS.1115, y completar con el nuevo servicio de canal 3/2. Es lo que acostumbra a llamarse difusión simultánea. La ventaja es que permite suspender en un futuro el primer servicio (véase la Recomendación UIT-R BS.1115), y que permite hacer una mezcla óptima de programas 2/0 y 3/2 independientemente.

##### *El método de matrización*

Otra solución es utilizar matrices de compatibilidad. Pueden introducirse ecuaciones matriciales para garantizar la compatibilidad con los receptores existentes. En este caso, los canales de emisión de izquierda y derecha existentes serán utilizados para transmitir las señales de matrices

compatibles A y B. Otros canales de emisión se utilizan para transmitir las señales de matrices T, Q<sub>1</sub> y Q<sub>2</sub>. Una posible ventaja de este método es que requiere menos capacidad adicional de datos para completarlo con el nuevo servicio.

### **3.1.3 Compatibilidad con sistemas posteriores, conforme a la Recomendación UIT-R BS.1115**

Para las aplicaciones con el nuevo sistema multicanal y también con el sistema monofónico/estereofónico definido en la Recomendación UIT-R BS.1115, podría ser necesario que los decodificadores estén en condiciones de decodificar una cadena de bits de audio ISO/CEI IS 11172-3.

### **3.2 Velocidad binaria**

Se recomienda el sistema de Capa II ISO/CEI IS 11172-3 [ISO/CEI, 1993], con una velocidad binaria de 128 kbit/s por canal (véase la Recomendación UIT-R BS.1115). Un valor de  $2 \times 128$  kbit/s puede ser considerado como el límite superior para el servicio de dos canales. Un valor de  $5 \times 128$  kbit/s podría ser el límite superior del servicio multicanal cuando no sea necesario garantizar la compatibilidad con sistemas anteriores (véase el § 3.1.2). Dado que las técnicas de codificación compuesta permiten, previsiblemente, una mayor ganancia de codificación, el nuevo sistema de codificación multicanal podría limitarse a una velocidad binaria de 512 kbit/s, para el servicio principal de cinco canales.

El sistema de codificación debe garantizar la calidad de audio definida en el § 2.1 con la misma velocidad binaria que sería necesaria en el sistema recomendado en la Recomendación UIT-R BS.1115, o inferior a este valor.

### **3.3 Complejidad del decodificador**

Es importante limitar a valores razonables la complejidad del decodificador para el programa de audio, para poder utilizar un sistema decodificador más económico. Cuando se trata de reproducir un número inferior de canales  $M$  de un programa de audio con canales  $N$ , se utilizará un decodificador de menor complejidad que el decodificador completo de canal  $N$ .

## **Referencias Bibliográficas**

- ATSC [1992] Servicios de audio digital y de información anexa para un servicio de televisión avanzada. Documento ATSC T3/186. Comité para sistemas de televisión avanzada.
- ISO/CEI [1993] Codificación de las imágenes móviles y el audio correspondiente para el soporte de almacenamiento digital a velocidades binarias inferiores a 1,5 Mbit/s. Norma ISO/CEI IS 11172-3. Organización Internacional de Normalización/Comisión Electrotécnica Internacional.
-