

Unión Internacional de Telecomunicaciones

UIT-R

Sector de Radiocomunicaciones de la UIT

Recomendación UIT-R BS.2051-0
(02/2014)

**Sistemas de sonido avanzados para
la producción de programas**

Serie BS
Servicio de radiodifusión
(sonora)



Unión
Internacional de
Telecomunicaciones

Prólogo

El Sector de Radiocomunicaciones tiene como cometido garantizar la utilización racional, equitativa, eficaz y económica del espectro de frecuencias radioeléctricas por todos los servicios de radiocomunicaciones, incluidos los servicios por satélite, y realizar, sin limitación de gamas de frecuencias, estudios que sirvan de base para la adopción de las Recomendaciones UIT-R.

Las Conferencias Mundiales y Regionales de Radiocomunicaciones y las Asambleas de Radiocomunicaciones, con la colaboración de las Comisiones de Estudio, cumplen las funciones reglamentarias y políticas del Sector de Radiocomunicaciones.

Política sobre Derechos de Propiedad Intelectual (IPR)

La política del UIT-R sobre Derechos de Propiedad Intelectual se describe en la Política Común de Patentes UIT-T/UIT-R/ISO/CEI a la que se hace referencia en el Anexo 1 a la Resolución UIT-R 1. Los formularios que deben utilizarse en la declaración sobre patentes y utilización de patentes por los titulares de las mismas figuran en la dirección web <http://www.itu.int/ITU-R/go/patents/es>, donde también aparecen las Directrices para la implementación de la Política Común de Patentes UIT-T/UIT-R/ISO/CEI y la base de datos sobre información de patentes del UIT-R sobre este asunto.

Series de las Recomendaciones UIT-R

(También disponible en línea en <http://www.itu.int/publ/R-REC/es>)

Series	Título
BO	Distribución por satélite
BR	Registro para producción, archivo y reproducción; películas en televisión
BS	Servicio de radiodifusión (sonora)
BT	Servicio de radiodifusión (televisión)
F	Servicio fijo
M	Servicios móviles, de radiodeterminación, de aficionados y otros servicios por satélite conexos
P	Propagación de las ondas radioeléctricas
RA	Radio astronomía
RS	Sistemas de detección a distancia
S	Servicio fijo por satélite
SA	Aplicaciones espaciales y meteorología
SF	Compartición de frecuencias y coordinación entre los sistemas del servicio fijo por satélite y del servicio fijo
SM	Gestión del espectro
SNG	Periodismo electrónico por satélite
TF	Emisiones de frecuencias patrón y señales horarias
V	Vocabulario y cuestiones afines

Nota: Esta Recomendación UIT-R fue aprobada en inglés conforme al procedimiento detallado en la Resolución UIT-R 1.

Publicación electrónica
Ginebra, 2016

© UIT 2016

Reservados todos los derechos. Ninguna parte de esta publicación puede reproducirse por ningún procedimiento sin previa autorización escrita por parte de la UIT.

RECOMENDACIÓN UIT-R BS.2051*

**Sistemas de sonido avanzados para
la producción de programas**

(Cuestión UIT-R 135/6)

(2014)

Cometido

En la presente Recomendación se especifica un sistema de sonido avanzado con y sin acompañamiento de imagen, superior a los sistemas que se exponen en la Recomendación UIT-R BS.775. Los sistemas de sonido avanzados utilizan datos de audio junto con un conjunto adecuado de metadatos para determinar las escenas de sonido que habrán de entregarse/radiodifundirse. Las especificaciones comprenden requisitos para señalar las propiedades del contenido de sonido avanzado y la disposición de los altavoces que se utilizará en la producción de contenidos para los sistemas de sonido avanzados. El sistema de sonido avanzado puede emplearse con el componente de sonido de programas de televisión y de imágenes digitales en pantalla grande (LSDI), así como con programas de audio únicamente.

Palabras clave

Audio inmersivo, audio multicanal, sistemas de sonido avanzados, sistema de sonido basado en canales, sistema de sonido basado en escenarios, sistema de sonido basado en objetos, posiciones de los altavoces

La Asamblea de Radiocomunicaciones de la UIT,

considerando

- a) que en la Recomendación UIT-R BS.775 – Sistema de sonido estereofónico multicanal con y sin acompañamiento de imagen se describe un sistema de sonido estereofónico multicanal con tres canales frontales y dos canales posteriores/laterales, así como con un canal facultativo de efectos de baja frecuencia (LFE), como el nivel más elevado en la jerarquía de sistemas de sonido multicanal que oscila entre 1/0 (monofónico) y 3/2;
- b) que en la Recomendación UIT-R BT.1769 – Valores de los parámetros de una jerarquía ampliada de formatos de imágenes LSDI para la producción y el intercambio de programas internacionales se especifica una jerarquía ampliada de formatos de imágenes LSDI para la producción y el intercambio internacional de programas;
- c) que en la Recomendación UIT-R BT.709 – Valores de los parámetros de la norma de TVAD para la producción y el intercambio internacional de programas se definen los parámetros del formato de imagen para la TVAD;

* La Comisión de Estudio 6 de Radiocomunicaciones introdujo cambios de edición en la presente Recomendación en el año 2016 con arreglo a la Resolución UIT-R 1.

- d) que en la Recomendación UIT-R BT.2020 – Valores de los parámetros de los sistemas de TVEAD para la producción y el intercambio internacional de programas se especifican los parámetros de los sistemas de imagen TVEAD;
- e) que el amplio campo de visión de la imagen de los programas de LSDI ampliadas, TVAD y televisión de extremadamente alta definición (TVEAD) goza de un sonido mejorado espacialmente superior a los canales de sonido 5.1;
- f) que en la Recomendación UIT-R BS.1909 – Requisitos de calidad de funcionamiento para un sistema de sonido estereofónico multicanal avanzado para uso con o sin acompañamiento de imagen se especifican los requisitos para un sistema de sonido avanzado para uso con o sin acompañamiento de imagen;
- g) que en el Informe UIT-R BS.2159 – Tecnología de sonido multicanal en aplicaciones domésticas y de radiodifusión figuran los resultados de los experimentos de evaluación subjetiva sobre la disposición de los altavoces para cumplir con los requisitos descritos en la Recomendación UIT-R BS.1909,

recomienda

- 1 que los descriptores/metadatos estáticos o dinámicos que cumplan con los criterios descritos en el Anexo 1 se empleen para señalar las propiedades de todas las señales sonoras que se utilicen en un sistema de sonido avanzado a fin de representar completamente el contenido de audio deseado;
- 2 que se tengan en cuenta las posiciones y configuraciones de los altavoces descritas en el Anexo 1 para producir programas de sonido avanzados;
- 3 que las configuraciones del número adecuado de elementos de audio¹ y la posición de los altavoces en la producción de programas se elijan de mutuo acuerdo entre el productor y el receptor del intercambio de programas;
- 4 que las interacciones de los consumidores con el audio recibido se faciliten en el sistema de sonido de producción y radiodifusión, y que sus características se elijan de mutuo acuerdo entre el productor y el receptor del intercambio de programas,

recomienda además

que se prosiga trabajando para proporcionar información adicional sobre las características de cualquier sistema de sonido avanzado conforme a la presente Recomendación para satisfacer los requisitos de calidad de la Recomendación UIT-R BS.1909.

¹ Un elemento de audio se considera una señal con metadatos estática o dinámica a lo largo de la duración del programa, lo que permite la entrega de contenido por canal, objeto o escena.

Anexo 1 (normativo)

Sistema de sonido avanzado para la producción de programas

1 Introducción

El sistema de sonido que se describe en la presente Recomendación es un sistema de sonido avanzado que permite que los metadatos asociados a cada uno de los trenes de audio sean estáticos o dinámicos durante el transcurso de un programa. Esto, por ejemplo, permite que un programa esté representado por elementos constituidos por una combinación de señales de objeto y señales de canal.

Por señales de canal se entienden las señales de audio que se mezclan para un número predeterminado de canales y están acompañadas por metadatos estáticos durante el transcurso de un programa. Cada uno de esos canales está asociado con una posición del altavoz concreta. Los flujos de trabajo de producción, las redes de radiodifusión y los sistemas de reproducción están definidos por una serie de posiciones del altavoz. Algunos ejemplos son los sistemas conformes a la Recomendación UIT-R BS.775.

Las señales de objeto son señales de audio que, al acompañarse de metadatos dinámicos durante el transcurso de un programa, representan elementos de audio que permiten que un reproductor reproduzca los objetos de audio de manera óptima en función del sistema de reproducción y el entorno auditivo. La adopción de un enfoque basado en los objetos podría asimismo permitir a los usuarios interactuar plenamente con el contenido de audio.

Los elementos basados en objetos y los elementos basados en canales pueden asociarse entre sí o existir independientemente. Para permitir cualquier combinación los elementos basados en objetos o los elementos basados en canales, todas las señales deberían estar acompañadas de los metadatos o descriptores necesarios, incluida la ubicación espacial independiente del tiempo (estática) y/o dependiente del tiempo (dinámica) de la señal. Esas señales pueden reproducirse por medio de una configuración de altavoces con diversas técnicas de reproducción y/o correspondencia.

Los programas de sonido avanzados consisten en señales de audio y sus metadatos correspondientes.

En el § 2 se exponen los requisitos de los metadatos para contenido de audio del sistema de sonido avanzado.

En el § 3 se describen las posiciones de los altavoces para los sistemas en entornos de producción. Habida cuenta de que para reproducir señales de audio es necesario un proceso de reproducción o de correspondencia, los números y la ubicación de los altavoces han de estar bien definidos. Esa información posibilita la reproducción de las señales de audio en función de una configuración predeterminada del altavoz en unas condiciones de reproducción.

2 Requisitos de los metadatos/descriptores para el contenido de audio de los sistemas de sonido avanzados

Todos los ficheros y flujos de audio utilizados en un sistema de sonido avanzado precisarán metadatos adecuados concomitantes. A diferencia de los sistemas fijos simples basados en canales, donde la ordenación de los canales, suele bastar para determinarlos, el sistema avanzado requerirá descripciones completas de todos los elementos de audio que se empleen para garantizar que se

tratan debidamente. Por consiguiente, los modelos de metadatos que hayan sido normalizados por una organización de normalización han de proporcionar definiciones coherentes para el audio. El modelo habrá de cumplir con los requisitos que figuran a continuación:

- Incluir toda la información necesaria para reproducir un programa en todas las hipotéticas configuraciones de reproducción enunciadas en la Recomendación UIT-R BS.1909 en función de una única representación.
- Poder describir el formato de cualquier elemento de audio basado en canales, objetos o escenas.
- Disponer de flexibilidad suficiente para describir cualquier combinación de elementos.
- La descripción de los elementos de los metadatos ha de ser perfectamente detallada para que pueda utilizarlos cualquier reproductor.
- Quedar especificado en un esquema XML abierto que permita que los metadatos se representen en XML (como método principal, por supuesto podrían convertirse a otros formatos como JSON).
- Poder ser incorporado al formato de un fichero de audio existente.
- Permitir que las definiciones generalizadas (en particular las configuraciones existentes basadas en canales) sean abiertas y accesibles sin restricciones a partir de un conjunto de definiciones de referencia.

3 Configuraciones de los altavoces en los sistemas de sonido avanzados

Las señales basadas en canales (incluidas las de los sistemas de sonido avanzados) precisan una configuración de los altavoces en la que el número y la posición de los altavoces estén perfectamente definidos. Las señales basadas en objetos pueden reproducirse a través de altavoces configurados para señales basadas en canales o altavoces adicionales en el caso de otros sistemas de reproducción avanzados.

Para garantizar una definición coherente de la configuración de los altavoces, en los sistemas de sonido multicanal avanzado se ha definido un conjunto de parámetros que especifican la etiqueta, la designación y la posición de cada altavoz, según figura en el Cuadro 1:

- Etiqueta SP: abreviatura de la designación del canal.
- Acimut: el ángulo de acimut expresado en grados, los valores positivos rotan a la izquierda si se mira al frente.
- Elevación: el ángulo de elevación expresado en grados, los valores positivos aumentan respecto al plano horizontal.

CUADRO 1 (Fin)

Etiqueta SP	Acimut	Elevación	A	B	C	D	E	F	G	H
			0+2+0	0+5+0	2+5+0	4+5+0	4+5+1	3+7+0	4+9+0	9+10+3
B+000	0	-30					X			X
B+022	+22,5	-30								
B-022	-22,5	-30								
B+030	+30	-30								
B-030	-30	-30								
B+045	+45	-30								X
B-045	-45	-30								X
B+060	+60	-30								
B-060	-60	-30								
B+090	+90	-30								
B-090	-90	-30								
B+110	+110	-30								
B-110	-110	-30								
B+135	+135	-30								
B-135	-135	-30								
B+180	+180	-30								
LFE1	+45	-30		X	X	X	X	X	X	X
LFE2	-45	-30						X		X

NOTA – Se parte de la premisa de que los altavoces se sitúan en una esfera. De no ser el caso, sería preferible que estuvieran alineados temporalmente (respecto al oyente) con una precisión de 100 μ s, incluida la trayectoria acústica. Resulta útil determinar, a través de los metadatos, si el entorno de producción estaba alineado temporalmente.

La disposición de los altavoces que figura en el Cuadro 1 está ilustrada en el Cuadro 2.

CUADRO 2

Disposición de los altavoces en los sistemas de sonido avanzados

<p>Capa inferior Comprende altavoces de efectos de bajas frecuencias (LFE)</p>	
<p>Capa intermedia Mismo nivel que el oyente</p>	
<p>Capa superior Comprende el altavoz superior (T+000)</p>	

3.1 Posición de los altavoces en los entornos de producción

Habida cuenta de que para el diseño práctico de las configuraciones de los altavoces en los estudios de producción podría ser necesario tener en cuenta varias características, como por ejemplo el tamaño de la sala, la imagen asociada u otras limitaciones, las posiciones de los altavoces podrían registrar variaciones en los ángulos de acimut y de elevación. Para garantizar la adaptación y reproducción adecuadas de los elementos basados en el canal, las posiciones de los altavoces utilizadas en un entorno de producción concreto deberían almacenarse como parte de los metadatos conforme al § 2. Cuando se transmite el contenido a una ubicación o sistema de reproducción diferente, es preciso cerciorarse de que se realicen los ajustes necesarios al programa para satisfacer todos los requisitos de calidad del sistema de sonido avanzado. Los requisitos de calidad están especificados en la Recomendación UIT-R BS.1909.

No obstante, los altavoces deben situarse dentro de los sectores comprendidos entre los valores mínimos de acimut/elevación y los valores máximos de acimut/elevación que figuran en los Cuadros 3 a 6 para atenuar las variaciones en la calidad del sonido en función de las diferentes posiciones de los altavoces. En los cuadros debe entenderse que «a .. b» denota el menor de los dos sectores que puede obtenerse por rotación dextrógira o levógira entre el ángulo «a» y el ángulo «b».

CUADRO 3

Configuración del altavoz para el sistema de sonido H (9+10+3)

Etiqueta SP	Canal		Acimut		Elevación	
	Etiqueta	Designación	Rango		Rango	
M+060	FL	Frontal izquierdo	+45	+60	0	+5
M-060	FR	Frontal derecho	-45	-60	0	+5
M+000	FC	Frontal central	0		0	+5
LFE1	LFE1	Efectos de bajas frecuencias-1	+30	+90	-15	-30
M+135	BL	Posterior izquierdo	+110	+135	0	+15
M-135	BR	Posterior derecho	-110	-135	0	+15
M+030	FLc	Frontal izquierdo central	+22,5	+30	0	+5
M-030	FRc	Frontal derecho central	-22,5	-30	0	+5
M+180	BC	Central posterior	+180		0	+15
LFE2	LFE2	Efectos de bajas frecuencias-2	-30	-90	-15	-30
M+090	SiL	Lateral izquierdo	+90		0	+15
M-090	SiR	Lateral derecho	-90		0	+15
U+045	TpFL	Frontal superior izquierdo	+45	+60	+30	+45
U-045	TpFR	Frontal superior derecho	-45	-60	+30	+45
U+000	TpFC	Frontal superior centro	0		+30	+45
T+000	TpC	Central superior	-		+90	
U+135	TpBL	Posterior superior izquierdo	+110	+135	+30	+45
U-135	TpBR	Posterior superior derecho	-110	-135	+30	+45
U+090	TpSiL	Lateral superior izquierdo	+90		+30	+45
U-090	TpSiR	Lateral superior derecho	-90		+30	+45
U+180	TpBC	Posterior superior central	+180		+30	+45

CUADRO 3 (Fin)

Configuración del altavoz para el sistema de sonido H (9+10+3)

Etiqueta SP	Canal		Acimut		Elevación	
	Etiqueta	Designación	Rango		Rango	
B+000	BtFC	Frontal central inferior	0		-15	-30
B+045	BtFL	Frontal inferior izquierdo	+45	+60	-15	-30
B-045	BtFR	Frontal inferior derecho	-45	-60	-15	-30

CUADRO 4

Configuración del altavoz para el sistema de sonido G (4+9+0)

Etiqueta SP	Canal		Ubicación del altavoz, coordenadas polares		Ubicación del altavoz, coordenadas cartesianas				
			Acimut	Elevación	X (izquierda/ derecha)	Y (posterior/ frontal)	Z (inferior/ superior)		
	Etiqueta	Designación	Rango					Rango	
M+000	C	Central	0		0	0	+1	0	
M+SC	Lsc	Pantalla izquierda	Borde izquierdo de la pantalla		0	-1*	+1	0	
M-SC	Rsc	Pantalla derecha	Borde derecho de la pantalla		0	+1*	+1	0	
M+030	L	Izquierdo	+30	+45	0	-1	+1	0	
M-030	R	Derecho	-30	-45	0	+1	+1	0	
M+090	Ls	Ambiente izquierdo	+90	+110	0	-1	0	0	
M-090	Rs	Ambiente derecho	-90	-110	0	+1	0	0	
M+135	Lrs	Ambiente posterior izquierdo	+135	+150	0	-1	-1	0	
M-135	Rrs	Ambiente posterior derecho	-135	-150	0	+1	-1	0	
U+045	Ltf	Frontal superior izquierdo	+30	+45	+30	+45	-0,5	+0,5	+1
U-045	Rtf	Frontal superior derecho	-30	-45	+30	+45	+0,5	+0,5	+1
U+110	Ltr	Posterior superior izquierdo	+110	+135	+30	+45	-0,5	-0,5	+1
U-110	Rtr	Posterior superior derecho	-110	-135	+30	+45	+0,5	-0,5	+1
LFE1	LFE	Efectos de bajas frecuencias							

* Tomando como referencia la pantalla, -1 = borde izquierdo de la pantalla, +1 = borde derecho de la pantalla.

CUADRO 5

Configuración del altavoz para el sistema de sonido F (3+7+0)

Etiqueta SP	Canal		Acimut		Elevación	
	Etiqueta	Designación	Rango		Rango	
M+000	C	Central	0		0	
M+030	L	Izquierdo	+30		0	
M-030	R	Derecho	-30		0	
U+045	LH	Altura izquierda	+30	+45	+30	+45
U-045	RH	Altura derecha	-30	-45	+30	+45
M+090	LS	Lateral izquierdo	+60	+150	0	
M-090	RS	Lateral derecho	-60	-150	0	
M+135	LB	Posterior izquierdo	+60	+150	0	
M-135	RB	Posterior derecho	-60	-150	0	
UH+180	CH	Altura central	180		+45	+90
LFE1	LFE1	Efectos de bajas frecuencias izquierdo	+30	+90	-15	-30
LFE2	LFE2	Efectos de bajas frecuencias derecho	-30	-90	-15	-30

CUADRO 6

Configuración del altavoz para el sistema de sonido B (0+5+0)
(de la Recomendación UIT-R BS.775)

Etiqueta SP	Canal		Acimut		Elevación	
	Etiqueta	Designación	Rango		Rango	
M+030	L	Izquierda	+30		0	
M-030	R	Derecha	-30		0	
M+000	C	Central	0		0	
LFE1	LFE	Efectos de bajas frecuencias de la izquierda	-		-	
M+110	Ls	Ambiente izquierda	+100	+120	0	+15
M-110	Rs	Ambiente derecha	-100	-120	0	+15

La disposición de los altavoces que figuran en los Cuadros 3 a 6 aparece ilustrada en el Cuadro 7.

CUADRO 7

Disposición de los altavoces en los sistemas de sonido avanzados

Sistema de sonido	Capa superior	Capa intermedia	Capa inferior
H (9+10+3)			
Capa superior 3/3/3 (Nota)			
Capa intermedia 5/2/3			
Capa inferior 3/0/0.2			
G (4+9+0)			
Capa superior 2/0/2			
Capa intermedia 5/2/2			
Capa inferior 0.0.0.1			
F (3+7+0)			
Capa superior 2/0/1			
Capa intermedia 3/2/2			
Capa inferior 0/0/0.2			
B (0+5+0)			
Capa superior 0/0/0			
Capa intermedia 3/0/2			
Capa inferior 0/0/0.1			

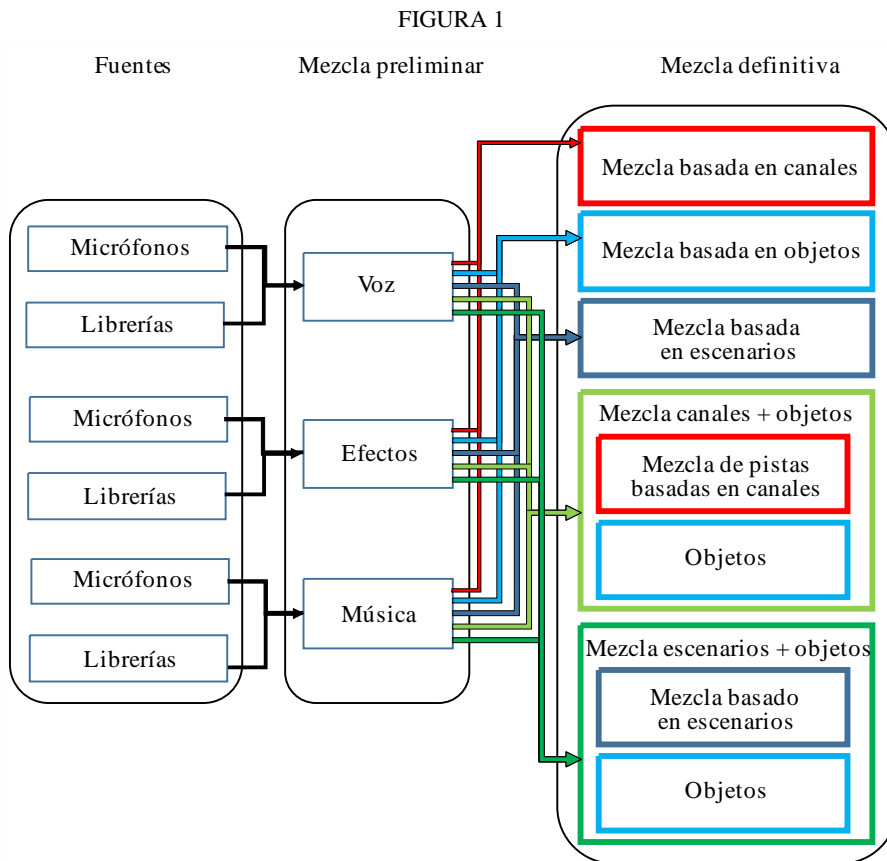
NOTA – X/Y/Z.LFE identifica el número de altavoces frontales/laterales/posteriores (más los LFE) de cada capa.

Apéndice 1

Producción del programa de sonido y características básicas del sistema de sonido avanzado

En el proceso de mezcla, una serie de elementos que dependen de la naturaleza de la producción y del proceso de adquisición de la señal, determinados habitualmente por el ingeniero de sonido, se ven reducidos a una representación de la intención del autor, manteniendo separado sólo un número reducido de elementos. La diferencia en la producción entre enfoques basados en canales, en objetos, en escenarios y enfoques que los combinen reside en los métodos de combinación de las mezclas preliminares y definitivas. El enfoque basado en los canales supone combinar todos los elementos en un conjunto predeterminado de canales, mientras que el enfoque híbrido basado en canales + basado en objetos da cabida a que los objetos se combinen en canales o se mantengan como objetos separados. De manera en similar, en el enfoque híbrido basado en escenarios + basado en objetos, los elementos o bien están almacenados en un formato basado en escenarios (por ejemplo, HOA), o bien se mantienen como objetos separados. En un enfoque únicamente basado en los objetos, se mantienen separados todos los elementos que recreen una determinada experiencia.

El flujo de trabajo por producción, mezcla en directo y posproducción de los sistemas basados en canales es el mismo que en el caso del modelo híbrido de combinación de pistas y objetos. El formato de transmisión de la mezcla definitiva difiere entre el modelo basado en canales y el modelo híbrido que transmite señales de combinación de pistas basadas en canales, así como señales de objetos.



La Figura muestra el aspecto que podría tener una cadena de radiodifusión característica en un sistema de sonido avanzado. La producción puede utilizar cualquier tipo de fuente de audio y contenido, pero ha de estar perfectamente descrita con los metadatos correctos para describir las señales de audio. Esta información queda almacenada en un formato de fichero compatible con esos metadatos. Esta etapa de la distribución adaptará las representaciones procedentes de la producción a una representación más compacta que conservará o generará metadatos nuevos para posibilitar una reproducción ulterior. El fichero o flujo de distribución pasa después a la etapa de radiodifusión, donde se reproducirá en un formato de radiodifusión particular. La radiodifusión en un mayor ancho de banda permitirá transmitir y reproducir muchos objetos y canales, mientras que la radiodifusión de bajo ancho de banda podría tener que reducirse a formatos estereofónicos más tradicionales. El formato de radiodifusión debería conservar todos los metadatos que sean necesarios en el extremo receptor.

Cada dispositivo receptor tendrá su propio reproductor, que habrá sido diseñado conforme a las posibles disposiciones de los altavoces en ese dispositivo. Por ejemplo, una cadena de música tendrá que tener mucha flexibilidad para permitir la disposición de varios altavoces en sitios distintos, mientras que en una televisión los altavoces internos estarán fijos en posiciones conocidas. La transmisión de contenido futuro podría recibir una representación, que mantendrá total flexibilidad para la interacción o personalización de los usuarios.

FIGURA 2

