

## RECOMENDACIÓN UIT-R BT.710-3

**EVALUACIÓN SUBJETIVA DE LA CALIDAD DE LAS IMÁGENES  
EN SISTEMAS DE TELEVISIÓN DE ALTA DEFINICIÓN**

(Cuestión UIT-R 211/11)

(1990-1992-1994-1997)

La Asamblea de Radiocomunicaciones de la UIT,

*considerando*

- a) que varias administraciones y organizaciones en el mundo están evaluando actualmente sistemas de televisión de alta definición (TVAD), y que en muchos lugares del mundo la radiodifusión de TVAD está llamada a ser el medio principal de difusión del próximo siglo;
- b) que las evaluaciones subjetivas son un elemento vital en el diseño y la selección de sistemas de TVAD;
- c) que la Recomendación UIT-R BT.500 presenta métodos de evaluación subjetiva generales, y que muchos de los detalles metodológicos allí expuestos son también apropiados en el contexto de la TVAD;
- d) que, no obstante, puede servir de ayuda aclarar los métodos de evaluación y las condiciones de observación apropiados para la TVAD, en las áreas clave que actualmente están en estudio, por medio de una Recomendación separada,

*recomienda*

- 1** que las evaluaciones subjetivas de la calidad de las imágenes en sistemas de TVAD, deberían realizarse en las condiciones de visualización dadas en el Anexo 1;
- 2** que la evaluación del comportamiento respecto a fallos del sistema de emisión de TVAD analógica se realice utilizando un método de degradación con doble estímulo (véase la Recomendación UIT-R BT.500), tomando como referencia o bien la imagen de TVAD de estudio o bien una emisión sin degradación;
- 3** que la evaluación del comportamiento respecto a fallos del sistema de emisión de TVAD se realice utilizando un método de degradación con doble estímulo (véase la Recomendación UIT-R BT.500), tomando como referencia o bien la imagen de TVAD de estudio o la de emisión sin degradación;
- 4** que el comportamiento respecto a fallos del contenido de imagen del sistema de TVAD digital se determine por medio del procedimiento definido en la Recomendación UIT-R BT.1129;
- 5** que, en ausencia de una referencia de alta calidad, se tenga en cuenta el método de escala gráfica o el método de valoración cuantitativa para evaluar la calidad global de la imagen (previa o posterior al tratamiento) suministrada por un sistema de TVAD de estudio;
- 6** que, en el caso de estar disponible una referencia de alta calidad, el método de doble estímulo con escala de calidad continua (véase la Recomendación UIT-R BT.500) se considere para la evaluación de la calidad global (previa o posterior al tratamiento) de un sistema de TVAD de estudio;
- 7** que, al abordar la calidad de funcionamiento para todo tipo de contenidos de programa y condiciones de transmisión que puedan darse en la práctica, se tenga en cuenta la descripción de las características de fallo compuestas que aparece en el Apéndice 2 al Anexo 1 de la Recomendación UIT-R BT.500;
- 8** que, en la interpretación de los resultados de estudios particulares, se tome nota debidamente de cualquier limitación que la tecnología actual pueda imponer sobre los resultados del estudio (por ejemplo, efectos debidos a limitaciones en los dispositivos de toma o presentación);
- 9** que se ponga cuidado en distinguir la influencia del formato de presentación cuando éste es distinto del formato del sistema básico (por ejemplo, cualquier conversión ascendente). Si es adecuado y aplicable, se pueden organizar evaluaciones para tener en cuenta los diferentes formatos.

NOTA 1 – En el Anexo 2 figura información relativa a los métodos de evaluación subjetiva para determinar la calidad de las imágenes en sistemas TVAD.

NOTA 2 – La información sobre los factores de evaluación para la evaluación global de la TVAD figura en el Anexo 3.

## ANEXO 1

## CUADRO 1

**Condiciones de observación para la evaluación subjetiva de la calidad de las imágenes TVAD**

Condición	Descripción	Valores <sup>(1)</sup>
a	Relación entre la distancia de observación y la altura de la imagen	3
b	Luminancia de cresta en la pantalla (cd/m <sup>2</sup> ) <sup>(2)</sup>	150-250
c	Relación entre el valor de luminancia de la pantalla con tubo inactivo (corte del haz) y el valor de cresta <sup>(3)</sup>	≤ 0,02
d	Relación entre el valor de luminancia cuando se representa en la pantalla sólo el nivel de negro en una habitación completamente a oscuras y el valor correspondiente al blanco más intenso <sup>(4)</sup>	Aproximadamente 0,01
e	Relación entre el valor de luminancia del fondo situado detrás del monitor de imágenes y el valor de luminancia de cresta de la imagen	Aproximadamente 0,15
f	Iluminación procedente de otras fuentes <sup>(5)</sup>	Baja
g	Cromaticidad del fondo	$D_{65}$
h	Ángulo subtendido por aquella parte del fondo que satisface las especificaciones anteriores <sup>(6)</sup> . Este valor debería respetarse para todos los observadores	53° de altura × 83° de anchura
i	Colocación de los observadores	Dentro de ± 30° horizontalmente desde el centro de la pantalla. El límite vertical está en estudio
j	Tamaño de la pantalla <sup>(7)</sup>	1,4 m (55 pulgadas)

- (1) Dado que actualmente puede ser imposible conseguir estas condiciones en su totalidad para las pruebas, se facilitan valores alternativos sobre una base provisional. Debería reconocerse, no obstante, que los resultados de las pruebas realizadas bajo las condiciones provisionales pueden en general no ser comparables con los obtenidos en las situaciones en que se aplican objetivos de presentación más bajos.
- (2) Valor de luminancia de cresta en la pantalla, correspondiente a la señal de vídeo con una amplitud del 100%. Hasta que el nivel especificado sea técnicamente realizable deberán utilizarse valores  $\geq 70$  cd/m<sup>2</sup>.
- (3) En este punto podría influir la iluminación de la habitación y la gama de contraste de la pantalla.
- (4) El nivel de negro corresponde a la señal de vídeo con amplitud del 0%.
- (5) La iluminación de la habitación será tal que satisfaga las condiciones c y e.
- (6) Se recomienda un mínimo de 28° de altura × 48° de anchura.
- (7) Deberán utilizarse valores mayores o iguales que  $\geq 76,2$  cm (30 pulgadas), si no se dispone de pantallas del tamaño especificado.

## ANEXO 2

**Métodos de evaluación subjetiva de la calidad de las imágenes en un entorno TVAD****1 Introducción**

En la Recomendación UIT-R BT.500 se describen los métodos utilizados en las pruebas subjetivas de los sistemas de televisión convencionales. Los conceptos principales de la metodología de evaluación se aplican por igual a todas las formas de televisión, pero la utilización de las especificaciones detalladas de los métodos de la televisión convencional en la TVAD requiere un estudio cuidadoso.

El UIT-R está examinando los adelantos en materia de TVAD para determinar, si fuera necesario, cuáles son los cambios requeridos en los métodos de evaluación subjetiva para incorporar estos adelantos. No se han completado aún los estudios a este respecto.

Antes de seguir adelante, es importante hacer hincapié en los puntos siguientes:

- La calidad de la imagen no es el único factor que ha de considerarse en la selección de normas. Otros factores tales como la complejidad del sistema, su disponibilidad, sus posibilidades futuras, etc., deben formar parte de los criterios de evaluación global.
- Los resultados de los experimentos de evaluación subjetiva no constituyen por sí mismos leyes físicas. Ofrecen pautas para un conjunto determinado de condiciones de ensayo, sin que constituyan hechos absolutos acerca de un sistema.
- Las diferencias conceptuales entre los términos de escalas de notas de calidad y de degradación utilizados actualmente no son uniformes, aunque tradicionalmente, en el proceso de los resultados se utiliza la aproximación de que lo son. Se están realizando estudios sobre otros métodos de evaluación con menos inconvenientes, pero la interpretación de los resultados de los métodos actuales debe tener en cuenta dichas desventajas.
- El elemento clave de las evaluaciones subjetivas suele ser la selección del material de prueba. Las directrices apuntan hacia un material que sea crítico pero no en exceso. La decisión sobre lo que puede ser crítico requiere una comprensión total de la forma en que trabajan los sistemas de TVAD.

## **2 Evaluación de la calidad de la imagen en un entorno TVAD**

### **2.1 Áreas de evaluación de la calidad de la imagen**

#### **2.1.1 Evaluación de los formatos de estudio de TVAD**

Es necesario evaluar:

- la calidad de la imagen básica,
- la calidad de la imagen tras un procesamiento post-producción, por ejemplo la incrustación cromática, la cámara lenta y la manipulación de la imagen, y una posible conversión a otros formatos, incluido el de película.

#### **2.1.2 Evaluación de los formatos de estudio convencionales (y de película) obtenidos a partir de las fuentes de estudio de TVAD**

Es necesario evaluar la adecuación, en términos de calidad de la imagen, de los formatos de estudio convencionales y de los de película obtenidos a partir de las fuentes de estudio de TVAD.

#### **2.1.3 Evaluación de los formatos de emisión de TVAD**

Es necesario evaluar:

- la calidad de la imagen básica,
- las características de los fallos,
- el comportamiento en cuanto al eco, y
- la susceptibilidad a la interferencia.

#### **2.1.4 Evaluación de las imágenes de televisión convencional incorporadas en las emisiones de TVAD**

Algunos de los formatos de emisión de TVAD que se consideran actualmente incluyen la incorporación de un formato de televisión convencional («compatibilidad hacia atrás»). Así pues, es necesario evaluar, en términos de calidad de la imagen, la adecuación de las imágenes de televisión convencional incorporadas a las emisiones de TVAD.

## **2.2 Aspectos de la evaluación de la calidad de la imagen**

### **2.2.1 Métodos de evaluación**

#### **2.2.1.1 Evaluación de la calidad de la imagen**

La escala de calidad de cinco notas que se utiliza actualmente en las evaluaciones subjetivas no tiene una distinción uniforme desde el punto de vista conceptual y se han señalado dificultades para comparar los resultados obtenidos en distintos laboratorios, especialmente cuando se requiere una traducción idiomática de los términos. Además, dada la sensibilidad de las evaluaciones de la calidad que utilizan términos conceptuales ante la gama de condiciones que se dan en los ensayos, no es aconsejable interpretar los términos de una manera absoluta o comparar los resultados de ensayos realizados con gamas de calidad distintas (por ejemplo, TVAD y televisión convencional).

Se ha utilizado con éxito una escala de calidad de siete notas para establecer el significado de la calidad de la TVAD y esa técnica puede ser útil en el futuro. Por otra parte, en el § 6 de la Recomendación UIT-R BT.500 se presentan alternativas a los métodos basados en la escala de cinco notas. Teniendo esto en cuenta, se propone que para las evaluaciones de entorno TVAD se utilice, en general, el método de doble estímulo con escala de calidad continua que figura en la Recomendación UIT-R BT.500.

#### **2.2.1.2 Evaluación de las degradaciones de la imagen**

En cierta medida, los mismos problemas que plantea la escala de calidad de cinco notas se han señalado para la escala de degradación de cinco notas. Teniendo esto en cuenta, se recomienda que, cuando haya que evaluar las degradaciones de la calidad de la imagen, se utilice en general el método de degradación con doble estímulo que figura en la Recomendación UIT-R BT.500.

### **2.2.2 Condiciones de observación para las evaluaciones subjetivas en un entorno TVAD**

#### **2.2.2.1 Evaluación de los formatos de estudio de TVAD**

El Informe UIT-R BT.801 contiene los objetivos de presentación de la imagen para los formatos de estudio de TVAD.

#### **2.2.2.2 Evaluación de los formatos de estudio convencionales obtenidos a partir de fuentes de estudio de TVAD**

Puesto que estas evaluaciones se refieren a sistemas de TV ya considerados en los textos del UIT-R, las evaluaciones de los formatos de estudio convencionales deben utilizar las condiciones de observación ya acordadas y que figuran en la Recomendación UIT-R BT.500.

#### **2.2.2.3 Evaluación de los formatos de emisión de TVAD**

No está claro cómo pueden aplicarse a las condiciones probables de la observación en los hogares los objetivos de presentación para las imágenes de estudio de TVAD. No obstante, en las evaluaciones subjetivas de los formatos de emisión de TVAD se debería tener en cuenta de alguna manera los objetivos de calidad superior fijados para la TVAD de estudio.

Es probable que, dadas las limitaciones de la emisión, los formatos de emisión de TVAD no puedan reproducir totalmente el nivel de calidad de imagen posible en el estudio de TVAD. No obstante, reconociendo la finalidad de los formatos de emisión de reproducir en la mayor medida posible la imagen original del estudio y, a fin de mantener la coherencia de las pruebas subjetivas a lo largo de la cadena estudio-emisión de TVAD, se propone que para los ensayos de los formatos de emisión de TVAD así como para las pruebas de los formatos de estudio de TVAD se utilicen las condiciones de observación que se señalan en el Anexo 1.

#### **2.2.2.4 Evaluación de las imágenes de televisión convencional incorporadas en las emisiones TVAD**

Como se trata de imágenes de televisión convencional, las condiciones de observación de las Recomendaciones UIT-R BT.500 y UIT-R BT.1129 son aplicables.

### **3 Evaluación de la calidad de la imagen de los formatos de estudio de TVAD**

#### **3.1 Evaluación de la calidad de la imagen básica**

De lo que se trata aquí es de la calidad de la imagen del formato de estudio TVAD antes del procesamiento post-producción. Los factores que afectarán probablemente a la calidad de la imagen básica son, entre otros, la resolución espacial, la resolución temporal, la gama de colores y las características de linealidad. El Anexo 3 a esta Recomendación resume los trabajos sobre factores que intervienen en la evaluación de la calidad de la imagen de TVAD.

Existe el acuerdo general de que un aumento de la gama de colores y la inclusión de la codificación de luminancia constante es un objetivo deseable para el sistema de estudio de TVAD. Sin embargo, estas características conllevan la necesidad de un procesamiento más complejo de las señales en la cámara y en la pantalla, y puede ser necesario evaluar los compromisos entre las ventajas de estos objetivos y las posibles desventajas debidas al procesamiento complejo de las señales.

La evaluación del valor de una mayor gama de colores y la repercusión del procesamiento adicional requiere disponer de un dispositivo de visualización con una gama considerablemente mayor que las actuales pantallas de tubos de rayos catódicos (CRT) y una señal de origen debidamente procesada para la visualización de esta gran gama. Además, el procesamiento no lineal exige una pantalla CRT corriente para transformar la señal de origen de gama grande en una señal apropiada para esta pantalla de gama más pequeña. Las imágenes fijas, que contienen una gama de colores normales y algunos colores que están fuera de la gama más pequeña, deben evaluarse comparando las dos visualizaciones. Una fuente para colores de alta pureza son pelotas de hilos con colores saturados. Los materiales de este tipo pueden proporcionar colores muy saturados que están fuera de la gama de las actuales pantallas CRT pero se prestan a una composición de escenas razonable. La evaluación subjetiva de esta escena en pantallas CRT normales y en pantallas de alta pureza debe proporcionar una estimación de la ganancia de calidad.

La evaluación de los métodos de codificación de luminancia constante en comparación con los métodos de codificación de luminancia no constante debe realizarse comparando una pantalla *RGB* de anchura de banda completa con una pantalla en la cual puedan presentarse señales de luminancia constante o de luminancia no constante. La escena debe incluir detalles en colores saturados junto con los elementos normales de la escena. Las sombras formadas en pelotas de hilos rojos, verdes y azules saturados son un medio para proporcionar el detalle en colores saturados.

Los métodos normalmente utilizados para evaluar la calidad de la imagen (es decir, métodos de doble estímulo) requieren típicamente una condición de referencia, que proporcione calidad superior a la del sistema probado. Sin embargo, la alta calidad de un sistema de estudio de TVAD dificulta hallar condiciones de referencia apropiadas. Por este motivo, puede ser conveniente utilizar escenas observadas directamente (imágenes fijas y en movimiento) con el fin de proporcionar la condición de referencia para las evaluaciones de sistemas de estudio de TVAD.

##### **3.1.1 Metodología**

Podrá utilizarse el método de doble estímulo con escala de calidad continua. La referencia para la evaluación de la calidad de las imágenes podrá ser la escena observada directamente (sujeta al encuadre apropiado). La prueba podría consistir en la misma escena observada por el sistema que se prueba.

##### **3.1.2 Condiciones de observación**

Véase el Anexo 1.

##### **3.1.3 Material de evaluación**

El material de prueba podrá comprender varias imágenes fijas y secuencias en movimiento. Las fuentes de las imágenes fijas podrán ser transparencias (iluminadas por detrás) o fotogramas (directamente iluminados). Las fuentes para las secuencias de movimiento podrán ser dioramas en movimiento. La condición de referencia se proporciona cuando se observa directamente una fuente, mientras que la condición de pruebas se proporciona cuando la misma fuente es observada a través de una cámara y un monitor. Se mantendrá un encuadre idéntico para las dos condiciones reflejando los materiales de prueba para ambos en el mismo espejo de observación de 16:9. El cambio de condiciones se efectuará mediante obturadores en los trayectos ópticos. Dicho cambio debe hacerse bajo control del experimentador.

Las pruebas consisten en efectuar comparaciones del material de prueba obtenido con una cámara vídeo con el mismo material observado directamente. Para minimizar la posible contaminación de los resultados por diferencias inherentes a la televisión con respecto al «mundo real», será necesario controlar varios factores, entre los que cabe citar:

- *diferencias de paralaje*: durante el tiempo de observación, el observador no deberá poder moverse apreciablemente pues esto resultará en cierto paralaje de movimiento en la escena observada directamente pero no en la escena mostrada en el monitor;
- *profundidad visible*: el espejo de observación presentará alternadamente la imagen de televisión y la escena de origen. La composición y la iluminación de las escenas de origen deben fijarse para garantizar que se minimizan las diferencias de profundidad entre la imagen de televisión y la escena observada directamente;
- *iluminación de la escena*: el espejo de observación presentará alternadamente la imagen de televisión y la escena de origen. Habrá que ajustar la iluminación de la escena de origen cuando se varía el trayecto de visualización para mantener la intensidad y la temperatura de color ( $D_{65}$ ) constantes en ambas imágenes. Es posible que la temperatura de color deba fijarse escena por escena.

Se han propuesto varios criterios para la composición de escenas de origen, entre los que figuran:

- la resolución espacial estática,
- la resolución espacial dinámica,
- la reproducción de la luminancia,
- la reproducción de colores,
- la reproducción del movimiento.

Además, podría ser útil complementar estos criterios con otros, para escenas con fines especiales, que permitieran evaluar:

- efectos de profundidad aparente (por ejemplo, en escenas panorámicas),
- reproducción de tonos familiares (por ejemplo, tonos de piel),
- sensación de presencia (por ejemplo, en un giro rápido de la cámara para tomar una panorámica), y
- características de parpadeo (por ejemplo, con subcampos blancos, grandes).

Es importante establecer el conjunto de imágenes de prueba normalizadas que ha de utilizarse para diversas evaluaciones subjetivas de la calidad de imagen de TVAD, como se ha hecho para material de programas con 4:2:2 (véase también la Recomendación UIT-R BT.1210).

### 3.1.4 Interpretación de los resultados

El sistema probado debe aproximarse lo más posible al nivel de calidad proporcionado por la referencia observada directamente. Al considerar los resultados, deben tenerse en cuenta dos aspectos:

- es probable que un sistema de estudio de TVAD efectúe compromisos entre las diversas características que se relacionan con la calidad. Además, para considerar la calidad promediada en las diversas partes de un material de prueba, sería prudente examinar las reacciones ante cada escena de origen para identificar las características que podrán mejorarse;
- al interpretar los resultados, es necesario identificar y, en la medida posible, realizar los ajustes necesarios para tener en cuenta la posible contaminación de los resultados debida al grado de desarrollo técnico de los equipos (estado de implantación).

## 3.2 Evaluación de la calidad de la imagen de TVAD tras el procesamiento post-producción

Se consideran dos casos: el procesamiento post-producción y la conversión de normas.

### 3.2.1 Procesamiento post-producción

Los casos principales de procesamiento post-producción son la incrustación cromática, el movimiento en cámara lenta y la manipulación de la imagen. Las evaluaciones efectuadas cuando se elaboraron las normas de relación 4:2:2 de la Recomendación UIT-R BT.601 sugirieron que la incrustación cromática es la operación post-producción más exigente. Con una frecuencia de trama y un sistema de exploración determinados es probable que suceda lo mismo en la TVAD.

### **3.2.1.1 Evaluación de la incrustación cromática**

#### **3.2.1.1.1 Metodología**

Debe utilizarse el método de la escala de degradaciones con doble estímulo siempre que se disponga de toda la gama de calidades de imagen. La referencia para las evaluaciones de la incrustación cromática puede ser una incrustación que utilice una señal *RGB* en toda la anchura de banda como primer plano. La prueba puede consistir en una incrustación que utilice como primer plano la señal de anchura de banda reducida de diferencia de color. Debe optimizarse la calidad de las imágenes de prueba y de referencia de una en una, como se hace en la práctica. Se está considerando aún la metodología apropiada para el caso de no poder tener una gama completa de calidades de imagen.

#### **3.2.1.1.2 Condiciones de observación**

Véase el Anexo 1.

#### **3.2.1.1.3 Material de evaluación**

El material de prueba debe ser crítico en cuanto a los tipos de degradación probables para el procesamiento de la incrustación cromática. El material probablemente más exigente contendrá detalles finos de movimiento. No se conocen secuencias de prueba específicas para la incrustación cromática en la TVAD, si bien las escenas con crestas de olas y cintas trenzadas (en movimiento) y vasos (transparentes) pueden ser adecuados para la evaluación de la incrustación cromática; no obstante, se dispone de una imagen fija. La calidad de la incrustación depende en gran medida de la iluminación de la escena y ha de asegurarse que ésta se optimiza y se mantiene un nivel coherente (véase también la Recomendación UIT-R BT.1210).

#### **3.2.1.1.4 Interpretación de los resultados**

El material de prueba no debe degradarse apreciablemente en relación con el material de referencia.

### **3.2.1.2 Evaluación del movimiento lento y de la manipulación de la imagen**

#### **3.2.1.2.1 Metodología, condiciones de observación, material de evaluación, interpretación de resultados**

La evaluación de esta categoría plantea problemas ya que no es probable que se disponga de una señal de referencia de gran calidad. La inclusión de una señal de referencia es lo que da a los métodos de doble estímulo sus propiedades. Se está estudiando un método de valoración cuantitativa que puede ser adecuadamente estable y reproducible sin una referencia. Como alternativa, en algunos casos puede haber un medio de generar secuencias de referencia de gran calidad. Por ejemplo, puede conseguirse un movimiento lento de gran calidad mediante una toma por separado de la secuencia de origen a una velocidad de imágenes superior.

### **3.2.2 Calidad de la imagen tras la conversión de normas TVAD-TVAD**

#### **3.2.2.1 Metodología**

El objetivo declarado en todas las administraciones es lograr una norma única de TVAD de estudio en todo el mundo y uno de los motivos para ello es permitir el intercambio internacional de programas sin conversión de normas. No obstante, surgirán sin duda situaciones en que se necesite la conversión a partir de otros formatos de TVAD o de película. Además, pueden ser necesarias conversiones similares antes de la generación de un formato de emisión con una frecuencia de campo distinta a la de la fuente. Este tipo de conversiones deberían considerarse al estudiar un formato de emisión.

La conversión de las normas de frecuencia de campo puede dar lugar a la aparición de efectos secundarios transitorios y para lograr una mejor evaluación de todo el sistema, se propone un método de evaluación en dos pasos.

#### **3.2.2.1.1 Evaluaciones primarias**

Se consideran que son las evaluaciones principales y de mayor utilidad. Debe utilizarse un método de doble estímulo con escala de calidad continua. La señal de referencia debe ser, en el caso ideal, la misma imagen o secuencia utilizada como entrada al convertidor de normas, pero en este caso utilizando los parámetros de exploración de la señal de salida del convertidor. Si ello no es posible, o si dichos ensayos ofreciesen interés por otros motivos, la señal de referencia debe ser la señal de entrada al convertidor.

### **3.2.2.1.2 Evaluaciones auxiliares**

Se pedirá a una serie de observadores expertos que, utilizando el método de un solo estímulo (véase la Recomendación UIT-R BT.500), asignen una nota de calidad global a varios programas representativos convertidos. También puede ser posible evaluar las frecuencias de detección de los efectos secundarios, si bien esto requiere nuevos estudios.

### **3.2.2.2 Condiciones de observación**

Las que se indican en el Anexo 1.

### **3.2.2.3 Material de prueba**

#### **3.2.2.3.1 Evaluaciones primarias**

Deben utilizarse una serie relativamente grande de imágenes fijas y en movimiento. También puede utilizarse para las evaluaciones TVAD-TVAD el material enunciado en el § 3.1.3 (véase también la Recomendación UIT-R BT.1210).

#### **3.2.2.3.2 Evaluaciones auxiliares**

Puede pedirse a una serie de observadores expertos que den notas de calidad global a varios programas de 5 a 20 min de duración en los que se incluyan ejemplos de los distintos tipos de movimiento y escenas con gran detalle.

#### **3.2.2.3.3 Características del material de prueba**

Es probable que el material de prueba crítico para la conversión de normas incluya áreas de gran detalle que tienen velocidades y direcciones de movimiento distintas.

### **3.2.2.4 Interpretación de los resultados**

Los resultados deben interpretarse con cautela para que ninguna de las diferencias de calidad inherente a las dos normas de estudio de TVAD se atribuya al proceso de conversión. Contribuirá a ello la utilización de secuencias de referencia inyectadas directamente a la salida según la norma.

La calidad subjetiva de la imagen convertida debe ser virtualmente equivalente a la imagen de entrada, a menos que quede limitada por los parámetros de cualquiera de las normas.

## **4 Evaluación de la calidad de las imágenes de estudio convencionales obtenidas a partir de imágenes de TVAD en el estudio**

### **4.1 Áreas de evaluación**

La interfaz entre la TVAD y la televisión convencional puede implicar conversiones del número de líneas, de la frecuencia de trama y del formato de imagen, aunque también son posibles los casos en que no haya conversión de la frecuencia de trama. La calidad de la imagen convencional debe ser la misma que en el caso de producción directa según la norma convencional.

### **4.2 Degradaciones por la conversión de normas**

#### **4.2.1 Degradaciones debidas a la conversión del número de líneas**

Las conversiones que implican cambios en el número de líneas activas pueden resultar en perturbaciones perceptibles en forma de movimientos verticales de los extremos. Estas perturbaciones pueden ser más pronunciadas en las conversiones que aumentan el número de líneas que en las que disminuyen dicho número.

#### **4.2.2 Degradaciones debidas a la conversión de la frecuencia de trama**

Las conversiones que implican cambios de la frecuencia de trama introducirán efectos secundarios, tales como el baile vertical de imagen, limitados a las zonas de movimiento de la imagen. El nivel de estas degradaciones depende de la relación de las frecuencias de trama en cuestión y de la complejidad del algoritmo de conversión. Algunas técnicas, tales como la compensación con adaptación del movimiento, pueden reducir estos efectos secundarios a niveles muy reducidos.

### 4.2.3 Degradaciones debidas a la conversión del formato de imagen

Las conversiones del formato amplio de la TVAD al formato 4:3 convencional pueden dar lugar a pérdidas de resolución o de contenido significativo de la imagen. No obstante, no es éste probablemente un caso en el que la evaluación subjetiva ofrezca directrices útiles.

## 4.3 Evaluación de la calidad de la televisión convencional obtenida a partir de una señal de TVAD

### 4.3.1 Metodología

Es evidente que la calidad de los convertidores de TVAD a televisión convencional sólo puede evaluarse completamente con secuencias de imágenes en movimiento. Se estima que para la evaluación de las pequeñas degradaciones en la gama limitada prevista, el método de doble estímulo con escala de calidad continua es el más útil. Se pide a los evaluadores que observen un par de secuencias, una de las cuales es un formato directo de estudio convencional y la otra, un formato convencional adecuado obtenido a partir de la TVAD.

### 4.3.2 Condiciones de observación

Las condiciones de observación serán las indicadas en las Recomendaciones UIT-R BT.500 y UIT-R BT.1129.

### 4.3.3 Material de evaluación

Puede utilizarse un amplio abanico de materiales de programa relativamente críticos como material de evaluación (véase también la Recomendación UIT-R BT.1210). También puede ser apropiado el empleo de imágenes fijas.

Pueden incluirse asimismo otros dos tipos de secuencias, que probablemente serán más críticas:

- escenas con movimiento de zoom;
- escenas con movimientos en sentidos contrarios, como las de un mercado.

Las pruebas deben también realizarse con material que tenga un proceso post-producción.

### 4.3.4 Interpretación

En el caso ideal, una interfaz de TVAD a televisión convencional debe dar la misma calidad que la de la televisión convencional directa. Como esta exigencia no se cumplirá probablemente para las imágenes en movimiento, debe investigarse la frecuencia de las degradaciones de la imagen evaluada en los programas de televisión. Esto puede implicar un método de dos etapas, como en el caso de la conversión TVAD-TVAD.

## 5 Evaluación de la calidad de los sistemas de emisión de TVAD a partir de la norma de estudio de TVAD

### 5.1 Áreas de evaluación

Las características del sistema que ofrecen interés son las siguientes:

#### 5.1.1 Calidad básica

Se trata de la calidad de la imagen en condiciones de recepción perfecta, es decir, cuando la relación señal/ruido,  $S/N$ , o portadora/ruido,  $C/N$ , es elevada.

#### 5.1.2 Características de fallo

Relación entre la calidad de la imagen y el ruido (que tiene una característica adecuada para el sistema de modulación utilizado). Ha de revisarse la gama en que se efectuarán las evaluaciones tras una primera pasada disponiendo que con 5 a 8 puntos se cubra toda la gama. Para los sistemas MA la gama de interés es generalmente de una  $S/N$  comprendida entre 25 y 55 dB, y para los sistemas MF, de una  $C/N$  entre 0 y 30 dB.

#### 5.1.3 Comportamiento en cuanto al eco

Relación entre la calidad de la imagen y la amplitud y el retardo del eco. Normalmente ofrece más interés en los sistemas MA. Ha de revisarse la gama en que se efectuarán las evaluaciones tras las pasadas preliminares, si bien un enfoque adecuado podría ser obtener información en tres curvas, añadiendo una señal retardada a una señal no retardada con demoras de 150 ns, 1  $\mu$ s, y 5  $\mu$ s, respectivamente, y con amplitudes de eco comprendidas entre -5 y -25 dB, en relación con la señal deseada.

#### 5.1.4 Comportamiento en cuanto a interferencias

Es necesario evaluar las características de interferencia cocanal y de canal adyacente.

Puede ser adecuado evaluar los § 5.1.2 a 5.1.4 anteriores con aleatorización y sin ella.

### 5.2 Metodología

#### 5.2.1 Calidad básica

El problema básico de diseño de la emisión de TVAD es cumplir, en la mayor medida posible, los requisitos visuales de la TVAD con la anchura de banda disponible. Para ello, puede utilizarse un submuestreo espacial o temporal, o ambos.

Estas técnicas pueden introducir degradaciones detectables o pérdidas de la calidad, superiores a las atribuibles al formato de estudio. El submuestreo espacial puede dar lugar a pérdidas detectables en la resolución horizontal, vertical o diagonal. El submuestreo temporal puede traducirse en reducciones detectables de la calidad de las imágenes en movimiento. El submuestreo espacial-temporal puede dar lugar a pérdidas detectables de la resolución espacial en las secuencias de imágenes en movimiento.

Evidentemente, para la evaluación de los formatos de emisión de TVAD se requieren imágenes de alta resolución y secuencias de imágenes en movimiento. Pero en esta Recomendación se propone un método de evaluación de la calidad básica en dos etapas para conseguir una evaluación global adecuada y representativa.

##### 5.2.1.1 Evaluaciones primarias

Se considera que éstas son las evaluaciones principales y las más útiles. Debe utilizarse un método de doble estímulo con escala de calidad continua. La referencia debe ser la señal de origen de estudio y la señal de prueba debe ser la señal de emisión.

##### 5.2.1.2 Evaluaciones auxiliares

Puede pedirse a una serie de observadores expertos que establezcan una escala de calidad global para varios programas representativos en el formato de emisión. También puede ser posible evaluar la frecuencia de detección de los efectos secundarios, si bien esto requiere nuevos estudios.

#### 5.2.2 Características de fallo, comportamiento en cuanto a eco y comportamiento en cuanto a interferencia

Debe utilizarse un método de degradación con doble estímulo conforme a la Recomendación UIT-R BT.500.

Pueden adoptarse dos enfoques:

- *características de fallo acumulativas*: en este caso, se consideran los puntos en los cuales se producen pérdidas objetables con respecto a la referencia de alta calidad no degradada; y
- *características de fallo no acumulativas*: en este caso, se consideran los puntos en los cuales se producen pérdidas objetables con respecto al formato de emisión no degradado.

### 5.3 Condiciones de observación

Véase el Anexo 1.

### 5.4 Material de prueba

#### 5.4.1 Calidad básica

El material de prueba debe elegirse entre una gama de imágenes fijas de alta resolución y de secuencias de imágenes en movimiento que sean críticas, pero no en exceso.

El material que pueda ser crítico exigirá grandes detalles con movimientos simultáneos a diferentes velocidades y en distintas direcciones (véase también la Recomendación UIT-R BT.1210).

#### 5.4.2 Características de fallo, comportamiento en cuanto a eco y comportamiento en cuanto a interferencia

Pueden lograrse resultados adecuados utilizando únicamente una pequeña gama de imágenes fijas y en movimiento. Generalmente, puede calcularse y utilizarse de forma bastante precisa la nota de calidad media global (véase también la Recomendación UIT-R BT.1210).

## **5.5 Interpretación de los resultados**

### **5.5.1 Calidad básica**

Parece razonable argüir que, por motivos de eficacia, la calidad de la señal de emisión de TVAD debe estar más próxima a la calidad de la señal de estudio de TVAD que a la de la señal *RGB* de la televisión convencional.

Por regla general, para la mayor parte del material de prueba debe haber un aumento suficiente de calidad en comparación con la de la televisión convencional. Además, cualquier efecto secundario temporal no llamará demasiado la atención para no disminuir la calidad de la TVAD.

### **5.5.2 Características de fallo, comportamiento en cuanto a eco y comportamiento en cuanto a interferencia**

Pendiente de otros estudios.

## **6 Evaluación de la calidad de las imágenes compatibles incorporadas en los formatos de emisión de TVAD**

### **6.1 Áreas de evaluación**

Algunos sistemas de emisión de TVAD están pensados para permitir la recepción simultánea en receptores de TVAD y convencionales. El § 5 se refiere a la evaluación de la calidad de la emisión de TVAD en sí. Este punto se refiere a la calidad de la señal convencional recibida simultáneamente.

En general, debe establecerse un compromiso de diseño entre la calidad lograda en la presentación TVAD y la calidad de la presentación convencional, que lleve a un grado de compatibilidad basado en el nivel de degradación introducido. Ello puede implicar un estudio de los mismos factores enumerados en el § 5, pero en este caso para la imagen compatible.

Los sistemas propuestos de emisión TVAD implican un procesamiento temporal y otros mecanismos, que pueden dar lugar a degradaciones en las imágenes compatibles.

### **6.2 Metodología**

Para la calidad básica, debe utilizarse el método de doble estímulo con escala de calidad continua, y la referencia será el material elaborado directamente en el formato de emisión convencional y/o el material convertido directamente del formato del estudio de TVAD. Para las características de fallo, comportamiento del eco y de la interferencia debe utilizarse el método de degradación con doble estímulo, con el material elaborado directamente en el formato de emisión convencional (pero no degradado) y el material convertido directamente del formato de estudio de TVAD (pero no degradado), como referencia. En todos los casos, la señal de prueba debe ser la imagen compatible recibida.

### **6.3 Condiciones de observación**

Las que se indican en las Recomendaciones UIT-R BT.500 y UIT-R BT.1129 para la televisión convencional.

### **6.4 Material de prueba**

Ha de utilizarse una gama de imágenes fijas y en movimiento.

Las características del material de prueba deben ser, por lo general, las indicadas en el § 5 (es decir, críticas, pero no en exceso) (véase también la Recomendación UIT-R BT.1210).

### **6.5 Interpretación de los resultados**

La interpretación de lo que debe ser la calidad de las imágenes «compatibles» en términos cuantitativos plantea problemas, no sólo porque las escalas carezcan de intervalos.

Los resultados para cada imagen o secuencia de prueba deben presentarse por separado.

La calidad de la imagen incorporada debe, en principio, ser «equivalente» a la de la señal de referencia. En la práctica, debe lograrse un «grado de compatibilidad» convenido.

## **7 Evaluación de la calidad de la película cinematográfica obtenida a partir de material de origen TVAD**

Pendiente de otros estudios.

## **8 Comparaciones de posibles formatos de TVAD**

A veces puede ser necesario comparar posibles formatos de TVAD a efectos de selección. Se considera que estas comparaciones pueden ser de máxima utilidad para determinar las mejores características de los diversos formatos probados.

### **8.1 Comparaciones de sistemas de estudio de TVAD**

Se han encontrado tres maneras de comparar los posibles formatos de estudio:

- directamente, por comparación uno al lado del otro;
- indirectamente, por comparaciones con una condición de referencia común en un solo experimento; y
- teóricamente, estableciendo colocaciones relativas en términos de criterios determinados psicofísicamente.

### **8.2 Comparaciones directas**

Pendiente de otros estudios.

### **8.3 Comparaciones indirectas**

Una comparación indirecta requiere una condición de referencia común con la cual se evalúa cada sistema probado. Los métodos subjetivos normalmente utilizados para comparaciones indirectas (es decir, los métodos de doble estímulo) utilizan condiciones de referencia que, típicamente, proporcionan una calidad superior a la de cualquiera de las condiciones sometidas a prueba.

Sin embargo, la alta calidad de los posibles sistemas de estudio de TVAD hace difícil encontrar estas condiciones de referencia. Por este motivo, puede ser apropiado utilizar escenas observadas directamente para proporcionar la condición de referencia.

Para una prueba indirecta válida, la referencia observada directamente debe mantenerse constante a través de todos los sistemas probados. Naturalmente, para imágenes fijas esto puede realizarse por medio de transparencias o fotografías. Sin embargo, para imágenes en movimiento, es necesario utilizar secuencias de movimiento totalmente reproducibles para la referencia, lo que puede hacerse mediante escenas controladas mecánicamente (por ejemplo, dioramas).

Es igualmente importante asegurar que, salvo para las diferencias inherentes a los propios formatos, los materiales de prueba se mantienen constantes a través de todos los sistemas probados. Esto se conseguirá si la cámara vídeo para el sistema probado se utiliza para tomar la imagen fija o secuencia de referencia, en la medida en que se mantenga constante la referencia.

Cabe señalar que todos los sistemas considerados deben probarse en un solo contexto experimental (es decir, que los observadores deben ver, en el curso del experimento, una secuencia aleatoria de los sistemas probados). Esto puede hacerse alternando las cámaras utilizadas para reflejar los sistemas probados. El monitor que debe mantenerse constante, debe seleccionarse de modo que sea adecuado para todos los sistemas probados. Quizás no sea siempre posible extrapolar los datos generales aplicables a las condiciones fijadas por el Anexo 1, a partir de los resultados obtenidos con las condiciones de observación permitidas por los equipos actuales. Se debe prestar atención al interpretar los resultados de las pruebas para distinguir los valores relacionados con la norma de sistema con respecto a los pertinentes a la realización práctica.

Las escenas directamente observadas pueden proporcionar una referencia de calidad considerablemente superior a la de los sistemas probados. Ello puede plantear dos cuestiones:

- las diferencias de las reacciones subjetivas a los sistemas probados pueden minimizarse artificialmente. Cuando los observadores juzgan, tienden a estar influidos por la gama y distribución de la calidad observada. Cuando la calidad (incluida la mostrada por la referencia) abarca una gran gama, se tiende a juzgar los casos con una calidad algo similar como más similares de lo que serían si se evaluaran en un contexto más restringido o se compararan directamente;

- el método de prueba preferido puede variar. Si las condiciones (incluida la referencia y la prueba) abarcan una amplia gama de calidad, puede utilizarse el método de degradación con doble estímulo para las comparaciones directas. Cuando las condiciones abarcan una gama de calidad más pequeña, se prefiere el método de doble estímulo con escala de calidad continua.

De este modo, según la finalidad de la prueba, hay dos opciones. Si las pruebas están destinadas a clasificar sistemas en relación con una norma «perfecta», puede utilizarse una referencia superior y el método de degradación con doble estímulo. En este caso, no obstante, es posible que no se detecten diferencias finas entre sistemas. Por otra parte, si las pruebas se efectúan para hacer discriminaciones finas entre sistemas, debe evitarse una referencia superior y utilizarse el método de doble estímulo con escala de calidad continua. En este último caso, puede ser necesario limitar la calidad a la escena observada directamente por medio de la composición, iluminación, filtrado óptico, etc.

### **8.3.1 Metodología**

De acuerdo con la gama de calidad de la prueba, podrá utilizarse el método de doble estímulo con escala de calidad continua o el método de degradación con doble estímulo.

Si se utiliza el método de doble estímulo con escala de calidad continua, puede ser conveniente considerar una variante de este método. En esta variante, se utilizan exposiciones relativamente largas para estimular la detección de efectos sutiles, particularmente en las secuencias en movimiento.

Si se utiliza el método de doble estímulo con escala de calidad continua, cada prueba entrañará múltiples presentaciones alternadas de las condiciones de referencia y de prueba. Para la mitad de las pruebas (determinadas aleatoriamente), la condición de referencia ha de presentarse primero; para las pruebas restantes, debe presentarse primero la condición de prueba. Si se utiliza el método de degradación con doble estímulo, cada prueba consistirá en una sola alternación entre referencia y prueba, presentándose primero la referencia.

En el § 3.1.3 figura información sobre el material de prueba y las precauciones que deberán tomarse a fin de minimizar la posibilidad de contaminación de los resultados.

Cada observador deberá ver el espejo de observación a través de una apertura que permita la observación binocular, pero con muy poco o ningún movimiento de la cabeza. La observación puede realizarse individualmente o en pequeños grupos. Sin embargo, si más de un observador trabaja al mismo tiempo, el ángulo de visión (hacia la escena) debe mantenerse constante para todos los observadores.

Como se sabe que diferentes grupos lingüísticos utilizan diferentemente los términos de escala de calidad y de degradaciones, todas las pruebas deben efectuarse en un solo idioma con observadores que hablen bien ese idioma.

### **8.3.2 Condiciones de observación**

Véase el Anexo 1.

### **8.3.3 Material de la evaluación**

Se están estudiando los materiales disponibles.

### **8.3.4 Interpretación de los resultados**

La interpretación de los resultados se efectúa sobre la base de las clasificaciones relativas de los posibles sistemas con respecto a la referencia común observada directamente. Deben tenerse en cuenta los aspectos indicados en el § 3.1.4.

## **8.4 Comparaciones teóricas**

La base de este método es considerar, parámetro por parámetro, la clasificación de los posibles sistemas desde el punto de vista de los criterios psicofísicos pertinentes. Este método es el que se propone.

## **9 Comparaciones de posibles formatos de emisión de TVAD**

Al igual que los sistemas de estudio de TVAD, las comparaciones pueden ser directas, indirectas o teóricas. En este caso, sólo se consideran las comparaciones indirectas.

## 9.1 Calidad básica

Esta es generalmente como para los formatos de estudio de TVAD (véase el § 8.3). En este caso puede utilizarse el método de doble estímulo con escala de calidad continua, con una sola referencia de alta calidad.

## 9.2 Características de fallo

Estas pruebas son, por lo general, las indicadas en el § 5. Sin embargo, el propósito es comparar las características de fallo para todos los posibles sistemas.

## 10 Otros aspectos

Se estima que deben considerarse también los siguientes aspectos conexos: los métodos de evaluación para la conversión de TVAD a película cinematográfica de 35 mm, la utilización de los descriptores de escala de calidad del UIT-R, la interpretación de objetivos de calidad en términos de resultados numéricos de evaluación, la calidad subjetiva y las características de la relación señal/ruido de la señal de TVAD, y la relación entre los aspectos de imagen y sonido de la TVAD.

La campaña de pruebas subjetivas llevada a cabo en Italia y destinada a evaluar la distancia de observación preferida para los programas de TVAD con distintos tipos de material y tamaños de pantalla reveló que, por término medio, la distancia de observación preferida era de unas 5 ó 6 veces la altura de la imagen (5-6 *H*).

Según estudios anteriores, la distancia de observación preferida en este caso para la televisión de resolución convencional era de 8 a 9 *H*.

Hay que seguir estudiando la relación entre la distancia de observación preferida media y la distancia de observación para evaluaciones subjetivas.

### ANEXO 3

#### **Factores de evaluación adecuados para la evaluación global de la TVAD**

En un reciente estudio a gran escala efectuado en Canadá y Estados Unidos de América, los observadores evaluaron la TVAD (MUSE-E, a través de satélite) en términos absolutos y comparándola con la NTSC de calidad de estudio. En las evaluaciones se consideraron la calidad de la imagen global y factores de evaluación específicos, incluyendo la nitidez de la imagen, la calidad del color, la presentación del movimiento y de la profundidad, el brillo de la imagen, el tamaño de la pantalla y la forma de ésta (formato de imagen). Los resultados mostraron que:

- las opiniones absolutas sobre la TVAD únicamente, se concentran al final de la escala de calidad, lo que apunta hacia posibles problemas si se comparan los resultados de pruebas por separado con distintos sistemas de TVAD sobre la base de opiniones de calidad absolutas;
- los observadores pudieron responder diferencialmente a los distintos factores de evaluación, lo que indica que el método de factores específicos puede ser útil en evaluaciones futuras;
- las opiniones sobre calidad global de la imagen estaban íntimamente relacionadas con la mayoría, si bien no con la totalidad, de los factores de evaluación en los que se percibía que la TVAD difería de la NTSC, lo que sugiere que la calidad global de la imagen puede no llegar a reflejar totalmente las reacciones del observador;
- los juicios sobre los factores específicos estaban, en cierta medida, relacionados, lo que apunta hacia posibles jerarquías entre los factores utilizados en las evaluaciones y hacia la posible existencia de factores de calidad básica de orden inferior;
- las opiniones sobre la calidad global de la imagen se veían, en cierta medida, afectadas por factores distintos de evaluación en función de la distancia de observación, lo que indica la necesidad de considerar cuidadosamente las distancias de observación que han de utilizarse en las evaluaciones.