



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

UIT-T

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

P.931

(12/98)

SERIE P: CALIDAD DE TRANSMISIÓN TELEFÓNICA,
INSTALACIONES TELEFÓNICAS Y REDES LOCALES

Calidad audiovisual en servicios multimedios

**Retardo de las comunicaciones multimedios,
sincronización y medición de la velocidad de
tramas**

Recomendación UIT-T P.931

(Anteriormente Recomendación del CCITT)

RECOMENDACIONES DE LA SERIE P DEL UIT-T

CALIDAD DE TRANSMISIÓN TELEFÓNICA, INSTALACIONES TELEFÓNICAS Y REDES LOCALES

| | | |
|---|--------------|---------------|
| Vocabulario y efectos de los parámetros de transmisión sobre la opinión de los clientes | Serie | P.10 |
| Líneas y aparatos de abonado | Serie | P.30 P.300 |
| Patrones de transmisión | Serie | P.40 |
| Aparatos para mediciones objetivas | Serie | P.50 P.500 |
| Medidas electroacústicas objetivas | Serie | P.60 |
| Medidas relativas a la sonoridad vocal | Serie | P.70 |
| Métodos de evaluación objetiva y subjetiva de la calidad | Serie | P.80 P.800 |
| Calidad audiovisual en servicios multimedios | Serie | P.900 |

Para más información, véase la Lista de Recomendaciones del UIT-T.

RECOMENDACIÓN UIT-T P.931

RETARDO DE LAS COMUNICACIONES MULTIMEDIOS, SINCRONIZACIÓN Y MEDICIÓN DE LA VELOCIDAD DE TRAMAS

Resumen

Un aspecto de los verdaderos sistemas de comunicaciones multimedios, que los diferencia de un simple conjunto de canales no relacionados entre sí, es su capacidad de mantener una relación temporal entre los distintos medios.

La presente Recomendación especifica los parámetros y métodos de medición que permiten evaluar la sincronización relativa entre los canales de los medios, junto con otros dos aspectos fundamentales de la calidad temporal. El tiempo de transmisión o retardo a través de un canal es decisivo al evaluar la conveniencia de un sistema para la conversación y otros usos interactivos. Los intervalos de tiempo entre llegadas de tramas y su inversa, la velocidad de tramas, caracterizan la capacidad de un sistema para suministrar información de manera continua y consistente.

Los sistemas multimedios combinan canales de vídeo, audio y datos perfeccionando así las comunicaciones. Todos esos medios están contemplados en la presente Recomendación. El retardo de vídeo puede variar en gran medida en el curso de secuencias breves, las secuencias de audio y vídeo pueden sufrir distorsiones durante la transmisión, mientras que los trenes de datos pueden tener una estructura reducida o no tener ninguna y pueden contener bits erróneos. Si bien cada medio presenta desafíos específicos para su medición, los métodos aquí especificados satisfacen y responden a todos ellos. El método basado en el error cuadrático medio prevé y mide las variaciones instantáneas del retardo de vídeo que puedan presentarse. El método de retardo de audio es útil para los canales que no conservan la forma de onda vocal original. Existen también métodos para canales de datos que aprovechan las estructuras nativas y toleran bits erróneos. Todos estos métodos admiten señales de prueba representativas de las aplicaciones previstas del sistema.

Los métodos abarcan la captación de secuencias de tramas de entrada y de salida de los medios con una escala de tiempo común, efectúan comparaciones entre las tramas para determinar las tramas de salida activas (no repetidas) y establecen la concordancia entre las tramas de salida activas y las tramas de entrada únicas, para determinar el tiempo de transmisión y lograr la sincronización. Los métodos permiten acumular las distribuciones de retardo, desfases temporales e intervalos de tiempo entre llegadas de tramas, que representan los parámetros deseados en sus formas elementales.

Orígenes

La Recomendación UIT-T P.931 ha sido preparada por la Comisión de Estudio 12 (1997-2000) del UIT-T y fue aprobada por el procedimiento de la Resolución N.º 1 de la CMNT el 3 de diciembre de 1998.

PREFACIO

La UIT (Unión Internacional de Telecomunicaciones) es el organismo especializado de las Naciones Unidas en el campo de las telecomunicaciones. El UIT-T (Sector de Normalización de las Telecomunicaciones de la UIT) es un órgano permanente de la UIT. Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Conferencia Mundial de Normalización de las Telecomunicaciones (CMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución N.º 1 de la CMNT.

En ciertos sectores de la tecnología de la información que corresponden a la esfera de competencia del UIT-T, se preparan las normas necesarias en colaboración con la ISO y la CEI.

NOTA

En esta Recomendación, la expresión *empresa de explotación reconocida (EER)* designa a toda persona, compañía, empresa u organización gubernamental que explote un servicio de correspondencia pública. Los términos *Administración*, *EER* y *correspondencia pública* están definidos en la *Constitución de la UIT (Ginebra, 1992)*.

PROPIEDAD INTELECTUAL

La UIT señala a la atención la posibilidad de que la utilización o aplicación de la presente Recomendación suponga el empleo de un derecho de propiedad intelectual reivindicado. La UIT no adopta ninguna posición en cuanto a la demostración, validez o aplicabilidad de los derechos de propiedad intelectual reivindicados, ya sea por los miembros de la UIT o por terceros ajenos al proceso de elaboración de Recomendaciones.

En la fecha de aprobación de la presente Recomendación, la UIT ha recibido notificación de propiedad intelectual, protegida por patente, que puede ser necesaria para aplicar esta Recomendación. Sin embargo, debe señalarse a los usuarios que puede que esta información no se encuentre totalmente actualizada al respecto, por lo que se les insta encarecidamente a consultar la base de datos sobre patentes de la TSB.

© UIT 1999

Es propiedad. Ninguna parte de esta publicación puede reproducirse o utilizarse, de ninguna forma o por ningún medio, sea éste electrónico o mecánico, de fotocopia o de microfilm, sin previa autorización escrita por parte de la UIT.

ÍNDICE

| | Página |
|---|---------------|
| 1 Alcance y aplicación..... | 1 |
| 1.1 Alcance..... | 1 |
| 1.2 Aplicación | 2 |
| 1.2.1 Canales de usuario a usuario | 2 |
| 1.2.2 Configuraciones aplicables | 2 |
| 1.2.3 Interfaces aplicables – Vídeo | 4 |
| 1.2.4 Interfaces aplicables – Audio | 4 |
| 1.2.5 Interfaces aplicables – Datos..... | 4 |
| 2 Referencias | 5 |
| 3 Términos y definiciones | 5 |
| 3.1 Términos generales..... | 5 |
| 4 Abreviaturas | 8 |
| 5 Cálculos de tiempo para medios generales de comunicación | 9 |
| 5.1 Cálculos para canal único..... | 10 |
| 5.2 Tramas de medios..... | 10 |
| 5.3 Cálculos de sincronización..... | 11 |
| 6 Mediciones de vídeo..... | 12 |
| 6.1 Recopilación de tramas de vídeo para mediciones..... | 12 |
| 6.1.1 Descripción de las tramas de vídeo..... | 12 |
| 6.1.2 Tramas de vídeo en interfaces compuestas | 12 |
| 6.1.3 Tramas de vídeo en interfaces de componentes digitales | 15 |
| 6.2 Método del error cuadrático medio para mediciones de vídeo..... | 16 |
| 6.2.1 Generalidades..... | 16 |
| 6.2.2 Calibrado de la diferencia distinguible mínima entre tramas..... | 18 |
| 6.2.3 Pruebas de una secuencia para encontrar diferencias distinguibles | 19 |
| 6.2.4 Categorización de tramas activas y tramas repetidas..... | 20 |
| 6.2.5 Pruebas de correspondencia entre tramas (tramas concordantes)..... | 20 |
| 6.2.6 Clasificación de secuencias fuente para los métodos del error cuadrático medio..... | 21 |
| 6.2.7 Consideraciones sobre el empleo de secuencias fuente con reducción 3:2... | 22 |
| 6.2.8 Factores que influyen en la exactitud y estabilidad de las mediciones..... | 23 |
| 6.3 Métodos de medición para vídeo basados en códigos horarios dentro de la trama.... | 23 |
| 7 Mediciones de audio..... | 24 |
| 7.1 Recopilación de tramas de audio para mediciones..... | 24 |

| | Página |
|---|---------------|
| 7.1.1 Descripción de las tramas de audio..... | 24 |
| 7.1.2 Conversión de señales analógicas en digitales..... | 24 |
| 7.1.3 Asignación de indicaciones de tiempo..... | 24 |
| 7.2 Medición del retardo en audio..... | 25 |
| 7.2.1 Generalidades..... | 26 |
| 7.2.2 Preparación de la señal..... | 27 |
| 7.2.3 Fase aproximada..... | 28 |
| 7.2.4 Fase de precisión..... | 30 |
| 7.2.5 Combinación de los resultados de las fases aproximada y de precisión..... | 33 |
| 8 Consideraciones sobre las mediciones combinadas de audio/vídeo..... | 34 |
| 8.1 Mediciones de la actividad de canal audio/vídeo y de la sincronización..... | 34 |
| 8.2 Asociación de las distintas mediciones entre sí..... | 34 |
| 9 Mediciones de datos..... | 35 |
| 9.1 Recopilación de tramas de datos para medición..... | 35 |
| 9.1.1 Consideraciones sobre la definición de tramas de datos..... | 35 |
| 9.1.2 Asignación de indicaciones de tiempo..... | 36 |
| 9.2 Medición de retardos para datos..... | 36 |
| 9.2.1 Concordancia de datos estructurados..... | 36 |
| 9.2.2 Concordancia de trenes de bits sin estructurar..... | 36 |
| 10 Requisitos de estabilidad de temporizador y sincronización..... | 37 |
| 10.1 Resolución..... | 37 |
| 10.2 Precisión y estabilidad (error de intervalo de tiempo admisible)..... | 37 |
| 10.3 Error de fijación del tiempo..... | 40 |
| Apéndice I – Bibliografía..... | 40 |
| Apéndice II – Símbolos matemáticos y códigos convencionales..... | 41 |

Introducción

Un aspecto de los verdaderos sistemas de comunicaciones multimedios, que los diferencia de un simple conjunto de canales no relacionados entre sí, es su capacidad de mantener una relación temporal entre los distintos medios.

La presente Recomendación especifica los parámetros y métodos de medición que permiten evaluar la sincronización relativa entre los canales de los medios, junto con otros dos aspectos fundamentales de la calidad temporal. El tiempo de transmisión o retardo a través de un canal es decisivo al evaluar la conveniencia de un sistema para la conversación y otros usos interactivos. Los intervalos de tiempo entre llegadas de tramas y su inversa, la velocidad de tramas, caracterizan la capacidad de un sistema para suministrar información de manera continua y consistente.

Los sistemas multimedios actuales combinan canales de vídeo, audio y datos perfeccionando así las comunicaciones. Todos esos medios están contemplados en la presente Recomendación. El retardo de vídeo puede variar en gran medida en el curso de secuencias breves, las secuencias de audio y vídeo pueden sufrir distorsiones durante la transmisión, mientras que los trenes de datos pueden tener una estructura reducida o no tener ninguna y pueden contener bits erróneos. Si bien cada medio presenta desafíos específicos para su medición, los métodos aquí especificados satisfacen y responden a todos ellos. El método basado en el error cuadrático medio prevé y mide las variaciones instantáneas de retardo de vídeo que puedan presentarse. El método de retardo de audio es útil para los canales que no conservan la forma de onda vocal original. Existen también métodos para canales de datos que aprovechan las estructuras nativas y toleran bits erróneos. Todos estos métodos admiten señales de prueba representativas de las aplicaciones previstas del sistema.

Recomendación P.931

RETARDO DE LAS COMUNICACIONES MULTIMEDIOS, SINCRONIZACIÓN Y MEDICIÓN DE LA VELOCIDAD DE TRAMAS

(Ginebra, 1998)

1 Alcance y aplicación

1.1 Alcance

La presente Recomendación abarca métodos de prueba para sistemas de transmisión multimedios que utilizan instalaciones de transporte digital. Ofrece un conjunto de parámetros de medición para caracterizar los siguientes aspectos de la calidad de funcionamiento de un sistema:

- 1) intervalos de tiempo entre llegadas de tramas de vídeo activas, que es la inversa de la velocidad de tramas elemental;
- 2) tiempo de transmisión del canal visual, también llamado retardo de vídeo;
- 3) tiempo de transmisión del canal audio (o retardo de audio);
- 4) tiempo de transmisión del canal de datos o retardo (e intervalos de tiempo entre llegadas de tramas);
- 5) sincronización temporal entre canales.

El alcance de la medición se limita a los casos con presencia de interfaces apropiadas de entrada y salida de los medios o cuando estas interfaces pueden obtenerse mediante accesorios de prueba facultativos.

Las siguientes aplicaciones quedan fuera del alcance de la presente Recomendación:

- 1) Aspectos relativos a la medición de la calidad de funcionamiento del sistema distintos del retardo, la sincronización y la velocidad de tramas. Las mediciones temporales no son suficientes para caracterizar la calidad de un sistema de transmisión multimedios. Por ejemplo, la calidad de reproducción de las tramas de vídeo desde la entrada hasta la salida es también importante, evidentemente, para los usuarios. La optimización de esos aspectos subjetivos de la calidad de funcionamiento para todos los parámetros de calidad puede prevalecer sobre la optimización de los resultados de las mediciones paramétricas realizadas conforme a esta Recomendación.
- 2) La selección ilimitada de un contenido fuente útil y representativo. Los métodos de medición aquí especificados requieren imponer ciertas restricciones a las señales de origen para prueba. Las secuencias fuente en vídeo con imágenes de mucho movimiento suelen producir en algunas aplicaciones multimedios un retardo mayor, una disminución de la velocidad de trama y un desfase de la sincronización audiovisual. Por consiguiente, las mediciones deberán basarse en cuadros de prueba realistas, que correspondan a la aplicación del sistema multimedios sometido a evaluación. En las cláusulas correspondientes a cada método de medición se indican otras limitaciones.
- 3) La medición de los aspectos relativos a la calidad de funcionamiento de los sistemas en los que no es posible acceder a las interfaces de entrada y de salida.
- 4) Los límites de los parámetros, no contemplados por el momento. La presente Recomendación sólo aporta métodos de medición de dichos parámetros, sin cuantificar la evaluación.

1.2 Aplicación

1.2.1 Canales de usuario a usuario

En condiciones ideales, la medición del retardo tendría lugar en las interfaces de usuario, para caracterizar todo el retardo de usuario a usuario. El canal de usuario a usuario completo comienza y termina en los dispositivos de interfaz. Consideremos, por ejemplo, el canal visual, con su cámara y componentes de visualización, que se muestra en la figura 1.

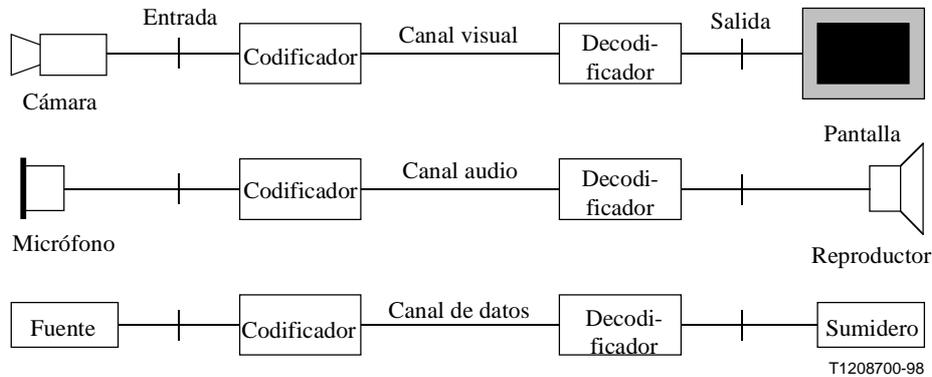


Figura 1/P.931 – Canales de usuario a usuario en un sistema multimedia

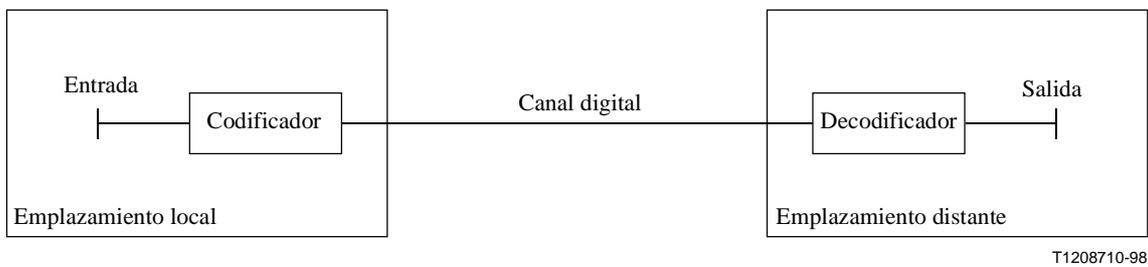
Desgraciadamente, las señales entran y salen de este canal bajo la forma de señales luminosas, por lo que la generación y recopilación de señales convenientes para medir constituye una tarea difícil. Para simplificar la interconexión entre el equipo de medición y el canal, el canal de prueba se especifica entre los conectores de la interfaz eléctrica a la salida de la cámara y a la entrada de la pantalla. La ventaja de este procedimiento consiste en que proporciona una estructura física y lógica mayor a la interfaz de prueba. Los retardos adicionales que aportan la cámara y la pantalla pueden evaluarse separadamente (se espera que estos retardos se circunscriban al intervalo muestra/pantalla, retardos que pueden ser constantes para las pantallas y suelen ser independientes de la señal de prueba), añadiéndose a las mediciones de retardo variable efectuadas de conformidad con la presente Recomendación.

Es posible identificar de manera similar las interfaces de entrada y salida en canales audio de usuario a usuario y en canales de datos de usuario a usuario.

1.2.2 Configuraciones aplicables

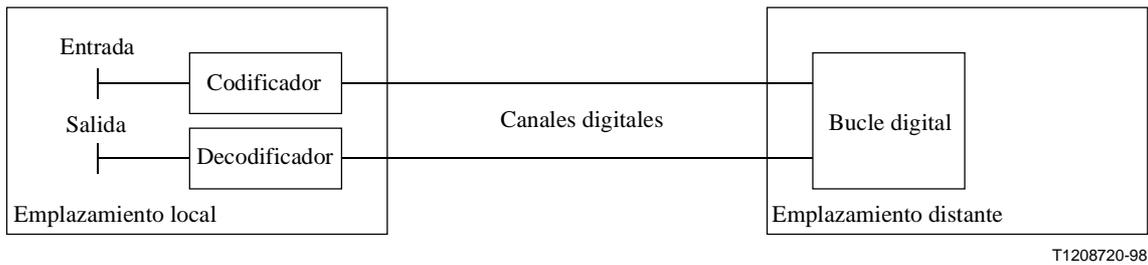
Las siguientes configuraciones de canal constituyen aplicaciones apropiadas de esta Recomendación. Cada figura muestra las interfaces de entrada y de salida necesarias.

Las figuras 2, 3 y 4 sólo contienen codificador, decodificador y canal digital para simplificar las cosas. Los componentes que puede comprender el canal de medios en dichas pruebas no están estrictamente limitados.



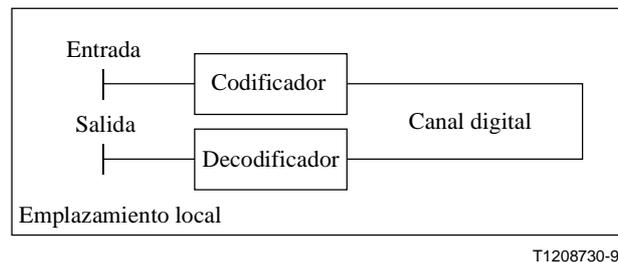
T1208710-98

Figura 2/P.931 – Medición de extremo a extremo¹



T1208720-98

Figura 3/P.931 – Medición digital distante en bucle

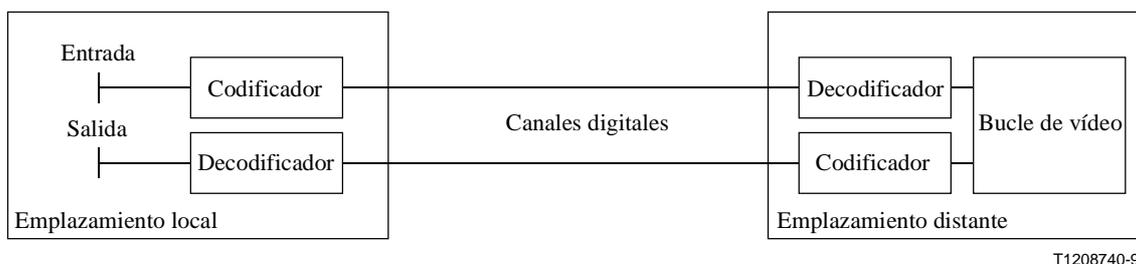


T1208730-98

Figura 4/P.931 – Medición de un sistema local

La figura 5 presenta la configuración de medición de un canal vídeo de aplicación limitada. Se trata de la medición del retardo en ambos sentidos de un sistema unidireccional sin retorno, para la que no haría falta más que un único dispositivo de medición.

¹ Obsérvese que, si el canal digital contiene componentes de tratamiento, los resultados se han de interpretar con cautela, para tener en cuenta los efectos de interacción.



T1208740-98

Figura 5/P.931 – Medición en bucle de vídeo distante

El bucle de vídeo no es adecuado en los casos en que los canales digitales presentan un retardo asimétrico. Para sistemas de teleconferencia en vídeo que codifican menos de 30 tramas por segundo, la codificación del trayecto de ida puede influir en el retardo de transmisión del trayecto de retorno, por lo que en este caso la prueba en bucle podría no ser apropiada. En tales circunstancias, la medición en ambos sentidos no permitiría la evaluación deseada en un solo sentido.

NOTA – La evaluación en un solo sentido de sistemas simétricos equivale simplemente a la mitad del retardo en ambos sentidos.

1.2.3 Interfaces aplicables – Vídeo

El trabajo que llevó a la elaboración de esta Recomendación se efectuó al principio mediante interfaces de señales de vídeo analógicas compuestas. Sin embargo, el diseño de sistemas de videoconferencia está evolucionando rápidamente, pasando de una simple adición de componentes (mediante la interfaz compuesta) a sistemas más integrados. Por añadidura, la demanda de productos vídeo de alta calidad supera las capacidades de la señal analógica compuesta. Para seguir el ritmo de las tecnologías más recientes, será necesario aplicar esta Recomendación a las nuevas interfaces. Es probable que esa aplicación se haga a las interfaces de componentes digitales, interfaces RGB de unidades de visualización de ordenadores e interfaces de cámara digital.

Los parámetros de medición aquí definidos sólo requieren la capacidad de suministrar tramas de vídeo a la entrada y reunir y comparar tramas de vídeo a la salida de un canal visual. Ninguna técnica precisa por sí misma de una interfaz compuesta.

Para facilitar las mediciones descritas en esta Recomendación respecto de sistemas totalmente integrados será necesario, en soporte de las pruebas, contar con características optativas de acceso a las interfaces. La complejidad de los distintos sistemas tal vez no aumente de manera notable si casi todas las funciones adicionales requeridas para la implementación de dichas interfaces están contenidas en el subsistema optativo. Estas características de interfaz pueden ser útiles en otras actividades, como la localización de averías y el aseguramiento de la calidad de fabricación, por lo que los fabricantes deberían recibirlas con agrado.

1.2.4 Interfaces aplicables – Audio

Esta especificación se aplica a todas las interfaces de audio normalizadas.

1.2.5 Interfaces aplicables – Datos

Esta especificación se aplica a todas las interfaces de datos normalizadas.

2 Referencias

Las siguientes Recomendaciones del UIT-T y otras referencias contienen disposiciones que, mediante su referencia en este texto, constituyen disposiciones de la presente Recomendación. Al efectuar esta publicación, estaban en vigor las ediciones indicadas. Todas las Recomendaciones y otras referencias son objeto de revisiones por lo que se preconiza que los usuarios de esta Recomendación investiguen la posibilidad de aplicar las ediciones más recientes de las Recomendaciones y otras referencias citadas a continuación. Se publica periódicamente una lista de las Recomendaciones UIT-T actualmente vigentes.

- Recomendación UIT-T P.84 (1993), *Método de pruebas de escucha subjetiva para la evaluación de equipos de multiplicación de circuitos digitales y de sistemas de voz por paquetes.*
- Recomendación UIT-T P.800 (1996), *Métodos de determinación subjetiva de la calidad de transmisión.*
- Recomendación UIT-T P.861 (1998), *Medición objetiva de la calidad de los códecs vocales de banda telefónica (300-340 Hz).*
- Recomendación UIT-T P.910 (1996), *Métodos de evaluación subjetiva de la calidad vídeo para aplicaciones multimedios.*
- Recomendación UIT-T P.920 (1996), *Métodos de prueba interactivos para comunicaciones audiovisuales.*
- Recomendación UIT-T P.930 (1996), *Principios de un sistema de degradaciones de referencia para vídeo.*
- Recomendación UIT-T X.25 (1996), *Interfaz entre el equipo terminal de datos y el equipo de terminación del circuito de datos para equipos terminales que funcionan en el modo paquete y están conectados a redes públicas de datos por circuitos especializados.*
- Recomendación UIT-R BT.470-6 (1998), *Sistemas de televisión convencionales.*
- Recomendación UIT-R BT.500-9 (1998), *Metodología para la evaluación subjetiva de la calidad de las imágenes de televisión.*
- Recomendación UIT-R BT.601-5 (1995), *Parámetros de codificación de televisión digital para estudios con formatos de imagen normal 4:3 y de pantalla ancha 16:9.*

3 Términos y definiciones

En esta Recomendación se definen los términos siguientes.

3.1 Términos generales

3.1.1 sistema de comunicación multimedios: Sistema que maneja más de un tren de medios de manera sincronizada desde el punto de vista del usuario. El sistema puede admitir la interconexión de múltiples partes y conexiones, así como la inserción o supresión de recursos y usuarios en el curso de una sola sesión de comunicación.

3.1.2 tren de medios: Secuencia de unidades de presentación, destinada a hacer llegar un contenido específico.

3.1.3 niveles de jerarquía de codificación: Unidades anidadas de representación de señales en las que puede descomponerse el tren de medios.

- 3.1.4 niveles de jerarquía de contenido:** Unidades anidadas de información en las que puede descomponerse el tren de medios.
- 3.1.5 unidad de presentación:** La menor división adecuada de un tren de medios (definida por el sistema de medición) que transporta una unidad de contenido independiente y autónoma a partir de los niveles de jerarquía de contenido presentes en el tren.
- 3.1.6 trama de vídeo:** Unidad de presentación del canal visual. Es el nivel más bajo de la jerarquía de contenido en un tren de medios vídeo en que las diferencias entre las unidades secuenciales aparecen en toda la extensión de la unidad de presentación. De ser necesario, la jerarquía de contenido de esta norma puede reutilizar los términos en ciertas jerarquías de codificación. En los formatos de 525 y 625 líneas, una trama de vídeo se define como un campo, cuyas especificaciones figuran en la Recomendación BT.470-6.
- 3.1.7 trama de audio:** Unidad de presentación del canal audio. Está formado por un grupo de muestras sucesivas de audio. El número preferido de muestras en una trama de audio depende de la velocidad de muestreo de audio y se indica en la cláusula 5. Dichas tramas de audio no guardan relación con las tramas definidas por determinados códecs audio/vocales.
- 3.1.8 trama de datos:** Unidad de presentación del canal de datos. Está formado por un grupo de bits sucesivos de datos. El número preferido de bits en una trama de datos depende de la aplicación del canal de datos.
- 3.1.9 canal digital:** Medio para transportar información de un punto a otro en forma digital. Un canal digital puede instalarse en una red formada por componentes de comunicación digitales.
- 3.1.10 canal visual:** Medio para llevar tramas de vídeo de un punto a otro. Una secuencia de tramas entregada a la entrada del canal da por resultado una secuencia de tramas similar (no necesariamente idéntica) a la salida del canal. El canal visual puede comprender los siguientes componentes: dispositivos de conversión del formato vídeo, codificadores (compresores) y decodificadores (descompresores), memorias intermedias de reconfiguración de la velocidad, multiplexores y demultiplexores, moduladores y demoduladores, facilidades de transmisión, conmutadores, unidades de conferencia multipunto y otros componentes necesarios para obtener las características de canal deseadas.
- 3.1.11 canal audio:** Medio para llevar señales de audio de un punto a otro. Una forma de onda de audio entregada a la entrada del canal da por resultado una forma de onda similar (no necesariamente idéntica) a la salida del canal. El canal audio puede comprender los siguientes componentes: codificadores (compresores) y decodificadores (descompresores), memorias intermedias, multiplexores y demultiplexores, moduladores y demoduladores, facilidades de transmisión, conmutadores, unidades de conferencia multipunto y otros componentes necesarios para obtener las características de canal deseadas.
- 3.1.12 canal de datos:** Medio para llevar datos de un punto a otro. Una secuencia de bits de datos entregada a la entrada del canal da por resultado una secuencia de bits similar (no necesariamente idéntica) a la salida del canal. El canal de datos puede comprender los siguientes componentes: dispositivos de conversión de formato, codificadores (compresores) y decodificadores (descompresores), memorias intermedias, dispositivos de segmentación y reensamblado del tren, multiplexores y demultiplexores, moduladores y demoduladores, facilidades de transmisión, conmutadores, unidades de conferencia multipunto y otros componentes necesarios para obtener las características de canal deseada.
- 3.1.13 integridad de campo:** Atributo de un canal visual presente cuando el contenido de campos impares (pares) en la secuencia fuente de formato de 525 ó 625 líneas se transporta en los campos impares (pares) a la salida.

3.1.14 trama de vídeo repetida: Trama de vídeo de salida que no puede distinguirse de la o las tramas precedentes de la misma secuencia (mientras que las tramas de la secuencia de entrada correspondiente presentan diferencias distinguibles). Se supone que una trama repetida se genera en algún punto intermedio del canal visual. Dado que las tramas repetidas no han recorrido el canal desde la entrada hasta la salida, no se utilizan para calcular la distribución de retardos del canal visual. Las tramas repetidas tampoco transportan nuevos estímulos visuales y se excluyen del cálculo de los intervalos de tiempo entre llegadas de tramas (y, por consiguiente, la velocidad de tramas elemental).

3.1.15 trama de vídeo no repetida (trama activa): Trama de vídeo de salida que puede distinguirse de la o las tramas precedentes de la misma secuencia (cuando las tramas de la secuencia de entrada correspondiente presentan diferencias distinguibles). Se supone que la trama activa ha recorrido el canal desde la entrada hasta la salida y su retardo puede incluirse en la distribución de retardos del canal visual. Puesto que las tramas activas transportan nuevos estímulos visuales, constituyen la base de cálculo de los intervalos de tiempo entre llegadas de tramas (y, por consiguiente, la velocidad de tramas elemental).

NOTA – En este proceso se interpretará como trama activa cualquier trama interpolada que se haya generado en un decodificador.

3.1.16 concordancia de tramas: Proceso consistente en comparar una secuencia de tramas con otra secuencia de tramas, para determinar la correspondencia entre las tramas de cada secuencia y la correspondencia de cada trama.

NOTA – Una manera de probar la correspondencia entre dos tramas de vídeo consiste en comparar sus representaciones digitales píxel por píxel, y calcular la sumatoria de todos los píxeles como la media cuadrática de las diferencias (denominada generalmente error cuadrático medio).

3.1.17 identificación de tramas de vídeo repetidas: Proceso consistente en comparar cada trama de vídeo de salida con su o sus tramas precedentes en la secuencia y cuantificar el grado de correspondencia entre cada par. Cuando la correspondencia entre un par de tramas es alta (las únicas diferencias son imputables al ruido durante la medición), el par es indistinguible; cuando la secuencia de entrada correspondiente de tramas presenta diferencias distinguibles, la trama en cuestión se clasifica como trama repetida. Véase al respecto 3.1.16, nota.

3.1.18 identificación de tramas de vídeo activas: Proceso consistente en comparar cada trama de vídeo de salida con su o sus tramas precedentes en la secuencia y cuantificar el grado de correspondencia entre cada par. Cuando la correspondencia entre un par de tramas es limitada (caso de que las diferencias medidas puedan distinguirse del ruido de medición), y la secuencia de entrada de tramas correspondiente presenta diferencias distinguibles, la trama en cuestión se clasifica como trama activa. Véase al respecto 3.1.16, nota.

3.2 Marco de los parámetros medibles

En esta subcláusula se presentan definiciones de alto nivel de los principales parámetros de medición. En las subcláusulas que le siguen figuran también definiciones de métodos específicos.

3.2.1 retardo de transmisión: Tiempo que requiere una trama dada para recorrer el canal de transmisión. Para calcular dicho tiempo, se registran en primer lugar los momentos en que las tramas se sitúan en el canal, para buscar a continuación una trama de salida que haya recorrido el canal, anotando su tiempo de llegada a la salida. Seguidamente, la trama de salida deberá compararse de manera exclusiva con una trama de entrada. El retardo de transmisión viene dado entonces por la diferencia entre el tiempo de llegada y el tiempo de entrada.

NOTA – Cuando la actividad en el canal es escasa o nula, los métodos aquí descritos encontrarán dificultades para efectuar mediciones válidas. Sin embargo, este parámetro se vuelve superfluo después de visualizar la primera trama cuando todas las tramas que le siguen son idénticas (sin actividad o vídeo inmóvil).

3.2.2 distribución de retardos del tren de medios: Conjunto de retardos calculados para una secuencia de tramas de salida y expresado de manera que ponga en evidencia toda variación producida entre las distintas mediciones. Si así procede, pueden suministrarse también estadísticas sumatorias clásicas.

3.2.3 intervalos de tiempo entre llegadas de tramas activas: Tiempo transcurrido entre las sucesivas tramas activas a la salida del canal. Se calcula seleccionando una trama activa (o una trama no repetida que haya recorrido el canal) y anotando su tiempo de llegada a la salida. A continuación deberá hallarse la trama activa más reciente (previa), anotando su tiempo de llegada. El intervalo de tiempo entre llegadas de las tramas activas en el canal es entonces igual al tiempo de llegada de la trama presente menos el tiempo de llegada de la trama anterior. Véase 3.2.1, nota.

3.2.4 distribución de los intervalos de tiempo entre llegadas de tramas activas: Conjunto de intervalos de tiempo entre llegadas calculados para una secuencia de tramas de salida activas y expresados de manera que se manifieste claramente cualquier variación producida entre las distintas mediciones. Si procede, pueden también suministrarse estadísticas sumatorias clásicas.

3.2.5 velocidad de tramas elemental: Valor inverso del intervalo de tiempo entre llegadas de tramas activas correspondiente a la trama activa presente. La velocidad de tramas elemental es igual a 1 dividido por la diferencia entre los tiempos de llegada de la trama presente y de la trama anterior.

3.2.6 estadística de velocidades de tramas: Conjunto de datos estadísticos calculados para una secuencia de tramas activas de salida, expresado de manera que se manifieste claramente cualquier variabilidad. Al comunicar estadísticas sumatorias de velocidad de tramas, su cálculo deberá efectuarse utilizando la distribución de intervalos de tiempo entre llegadas y tomando los valores inversos.

3.2.7 relación de salto de trama elemental: Relación entre los intervalos de tiempo entre llegadas a la entrada y a la salida (o velocidades de tramas elementales). Es la medida de las modificaciones que aparecen entre las velocidades de tramas activas a la entrada y la salida.

4 Abreviaturas

En esta Recomendación se utilizan las siguientes siglas.

| | |
|------|---|
| ANSI | Instituto nacional de normas de los Estados Unidos (<i>American National Standards Institute</i>) |
| ATM | Modo de transferencia asíncrono (<i>asynchronous transfer mode</i>) |
| DC | Corriente continua (<i>direct current</i>) |
| DFT | Transformada discreta de Fourier (<i>discrete Fourier transform</i>) |
| EAV | Fin de vídeo activo (<i>end of active video</i>) |
| FFT | Transformada rápida de Fourier (<i>fast Fourier transform</i>) |
| fps | Tramas por segundo (<i>frames per second</i>) |
| GPS | Sistema mundial de determinación de posición (<i>global positioning system</i>) |
| IIR | Respuesta a impulsos infinita (<i>infinite impulse response</i>) |
| MSE | Error cuadrático medio (<i>mean square error</i>) |
| MTIE | Error de intervalo de tiempo máximo (<i>maximum time interval error</i>) |
| PSD | Densidad espectral de potencia (<i>power spectral density</i>) |
| PSNR | Relación señal/ruido de cresta (<i>peak signal-to-noise ratio</i>) |

| | |
|-------|---|
| RGB | Rojo, verde, azul (<i>red, green, blue</i>) |
| RMS | Media cuadrática (<i>root mean square</i>) |
| SAV | Inicio de vídeo activo (<i>start of active video</i>) |
| SMPTE | Sociedad de ingenieros de imágenes en movimiento y televisión (<i>Society of Motion Picture and Television Engineers</i>) |
| TIE | Error de intervalo de tiempo (<i>time interval error</i>) |

5 Cálculos de tiempo para medios generales de comunicación

En esta cláusula se presenta un modelo general de cálculo de los parámetros de esta especificación, y se expone la aplicación del modelo a cada uno de los medios que aquí se contemplan. La figura 6 ilustra el modelo general. Dos secuencias de **unidades de presentación**, P y P', se introducen en la interfaz de entrada y abandonan la interfaz de salida del canal. Al pasar por la interfaz la última porción de cada unidad de presentación, el sistema de medición procede a la lectura de un temporizador, T, asociando el valor T(n) a la unidad de presentación de entrada n, P(n). En la interfaz de salida, el sistema de medición procede a la lectura de un temporizador, T', asociando el valor T'(m) a la unidad de presentación de salida m, P'(m).

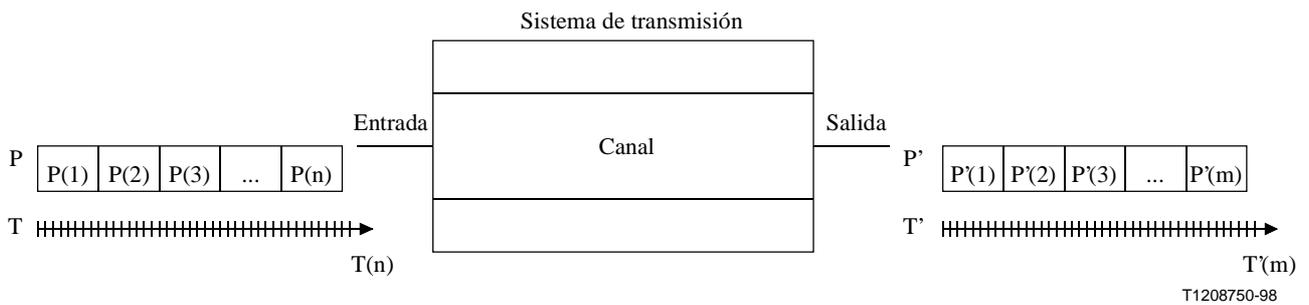


Figura 6/P.931 – Modelo general para unidades de presentación

Obsérvese que los temporizadores T y T' pueden ser un mismo temporizador en algunas configuraciones de medición (por ejemplo, las configuraciones locales y en bucle que se muestran en las figuras 3, 4 y 5). En otros casos (por ejemplo, las configuraciones de extremo a extremo, véase la figura 2), los temporizadores deberán estar sincronizados. Además, un mismo temporizador da indicaciones de tiempo casi simultáneas para todos los canales en una interfaz determinada, permitiendo así efectuar cálculos directos entre los canales y dentro de cada canal.

El inicio y el fin de las unidades de presentación pueden no coincidir, como se ve en la figura 6. Las indicaciones de tiempo deberán vincularse al fin de la unidad de presentación para contemplar tanto los aspectos prácticos como perceptivos. Las indicaciones de tiempo de comienzo pueden almacenarse separadamente, a los efectos de otras mediciones.

Obsérvese que P y P' se ilustran arriba como trenes continuos, pero también son posibles en muchos medios las llegadas no periódicas. De hecho, las unidades de presentación completas pueden necesitar más tiempo para salir del sistema que para entrar en él. Es éste uno de los motivos para indicar el tiempo de fin de las unidades de presentación. Otro motivo es de índole práctica, ya que el fin de una unidad de presentación suele preverse con mayor facilidad que su inicio.

5.1 Cálculos para canal único

El retardo de canal correspondiente a la unidad de presentación $P'(m)$, una vez comprobada la concordancia entre $P'(m)$ y $P(n)$, es:

$$t_p(m) = T'(m) - T(n)$$

La concordancia suele ser irrelevante en un canal sin pérdidas, pero los métodos aquí definidos se ocuparán tanto de la distorsión como de la pérdida completa de unidades de presentación. Deberá haber diferencias entre las sucesivas unidades de presentación a la entrada, ya que de no ser así existirá una cierta ambigüedad en los resultados de la concordancia.

El intervalo de tiempo entre llegadas para la unidad de presentación $P'(m)$ a la salida del canal es:

$$b'_p(m) = T'(m-1)$$

donde $m-1$ representa el índice de la unidad de presentación precedente.

La velocidad de tramas elemental para $P'(m)$ a la salida del canal es:

$$f'_p(m) = \frac{1}{T'(m) - T'(m-1)}$$

Cuando los sistemas descartan automáticamente las unidades de presentación (como ocurre en el caso de vídeo en canales digitales de baja velocidad binaria), las velocidades de tramas respectivas a la entrada y a la salida serán diferentes. La relación de salto de trama elemental es:

$$\frac{b'_p(m)}{b_p(n)} = \frac{f_p(n)}{f'_p(m)}$$

donde $b_p(n) = T(n) - T(n-1)$

y $P'(m)$ concuerda con $P(n)$

Con fines estadísticos y de análisis gráfico, pueden recopilarse conjuntos completos de mediciones separadas de retardos e intervalos de tiempo entre llegadas, representados por:

$$t_p = \{t_p(1), t_p(2), t_p(3), t_p(4), \dots, t_p(M)\}$$

y

$$b'_p = \{b'_p(2), b'_p(3), b'_p(4), \dots, b'_p(M)\}$$

5.2 Tramas de medios

En el canal visual, las unidades de presentación son tramas de vídeo. En la cláusula 6 se definen estas unidades para cada interfaz de vídeo como un único campo del formato de 525 líneas, por ejemplo. V y V' son las variables de las unidades de los trenes de vídeo de entrada y de salida.

En el canal audio, las unidades de presentación son tramas de audio. Las tramas de audio son un grupo de muestras digitalizadas que representan el tren de audio (véase el cuadro 1). En la cláusula 7 se definen estas tramas para las interfaces de audio. A y A' son las variables de las unidades de los trenes de audio de entrada y de salida. La longitud recomendada de la trama de audio es de aproximadamente la misma duración que la **trama de vídeo** correspondiente, si existe (por ejemplo, 16,66 ... ms para el formato de 525 líneas).

En el canal de datos, las unidades de presentación son tramas de datos. Las unidades de datos conforman una sola palabra de prueba única o un grupo de bits que representa cierta unidad de

presentación para la aplicación de usuario del tren de datos. En la cláusula 9 se definen dichas unidades para las interfaces de datos. D y D' son las variables de las unidades de los trenes de datos de entrada y de salida.

Cuadro 1/P.931 – Longitud recomendada de la trama de audio

| Muestras/segundo | Tamaño recomendado de la muestra, duración |
|------------------|--|
| 8 000 | 128, 16 ms |
| 16 000 | 256, 16 ms |
| 32 000 | 512, 16 ms |
| 44,1 k | 512, 11,61 ms |
| 48 k | 512, 10,66 ms |

5.3 Cálculos de sincronización

Como ejemplo de cálculo temporal de parámetros de sincronización, considérense unos trenes de medios de audio y vídeo en donde:

- A(m) y V(n) están asociados entre sí, bien por definición del dispositivo de medición o porque la asociación se conoce por anticipado.
- V'(q) y V(n) concuerdan entre sí, como resultado de los métodos descritos en esta especificación.
- A'(p) y A(m) concuerdan entre sí, como resultado de los métodos descritos en esta especificación.

El desplazamiento temporal entre las tramas de audio y vídeo asociadas a la entrada es:

$$O_{AV}(m,n) = T_A(m) - T_V(n)$$

Este parámetro indica la posición de la trama de audio respecto de la trama de vídeo en el tiempo.

El desplazamiento temporal entre las tramas de audio y vídeo asociadas a la salida es:

$$O'_{AV}(p,q) = T'_A(p) - T'_V(q)$$

Obsérvese que $O'_{AV}(p,q)$ representa la relación de sincronización percibida por un usuario a la salida del canal.

El desfase temporal entre las tramas de audio y vídeo asociadas a la salida, introducido por el canal del sistema de transmisión, es:

$$S_{AV}(p,q) = O'_{AV}(p,q) - O_{AV}(m,n)$$

Aplicando la convención sobre retardo de tiempo positivo, el retardo de tiempo de sincronización es un valor positivo. Cuando el que dirige es el segundo canal, el valor de este parámetro tiene signo negativo.

Para sistemas que funcionan a bajas velocidades de tramas, un desfase temporal distinto de cero puede producir subjetivamente una sincronización más precisa con el movimiento de los labios. Para sistemas con retardo de vídeo variable, la variación del retardo de audio para mantener el desfase en cero no suele ser conveniente, ya que las señales sonoras se volverían ininteligibles durante esa manipulación.

Si un sistema de transmisión es capaz de aceptar audio multicanal, estos cálculos son los apropiados para evaluar la sincronización entre los canales audio.

6 Mediciones de vídeo

6.1 Recopilación de tramas de vídeo para mediciones

En esta cláusula se describe la jerarquía de elementos que pueden existir en una secuencia de vídeo, y se especifica el nivel de cada jerarquía de codificación de vídeo que constituye la unidad fundamental para efectuar comparaciones. La presente Recomendación define dichas unidades fundamentales como **tramas de vídeo**, redefiniendo los términos de la jerarquía de codificación cuando es necesario. Primero examinaremos el tema en su nivel conceptual, para propiciar la extensión de los métodos más allá de las interfaces eléctricas específicas expuestas más adelante.

6.1.1 Descripción de las tramas de vídeo

En una secuencia de vídeo pueden presentarse muchos niveles de granularidad de la información. La información en las secuencias de vídeo puede dividirse como se ve en la figura 7.

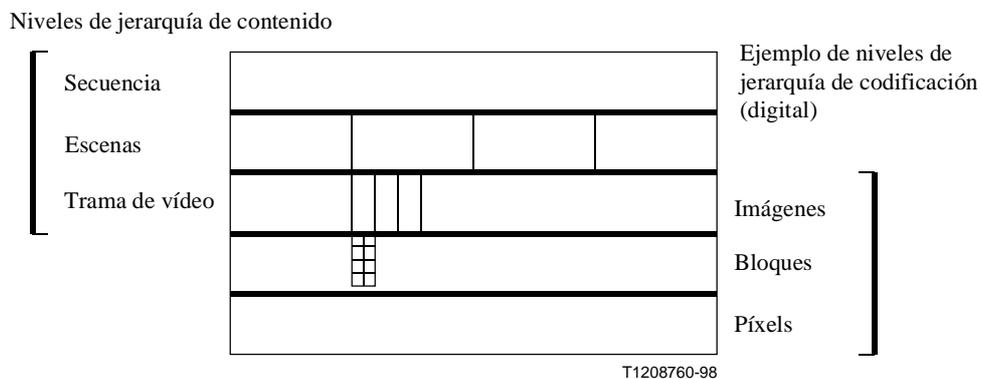


Figura 7/P.931 – Ejemplo de jerarquía de una secuencia de vídeo

La figura 7 muestra una secuencia de vídeo completa compuesta de cuatro escenas de contenido variado. Cada escena consta de muchas imágenes o tramas que difieren entre sí al transmitir el movimiento o cambio. Las imágenes constituyen el nivel autónomo más bajo de este ejemplo de jerarquía de contenido en el dominio del tiempo, ya que los cambios de contenido más pequeños tienen lugar en este nivel a lo largo del tiempo. Este nivel es la unidad de presentación fundamental para el usuario. Las **tramas de vídeo** las definimos a este nivel, en el que aparecen diferencias entre las unidades secuenciales a lo largo de la unidad de presentación.

Las tramas de vídeo también pueden encontrarse en el nivel más alto de la jerarquía de codificación de vídeo. La figura 7 muestra un ejemplo de jerarquía de codificación digital en donde esto ocurre realmente. No obstante, si se admite una codificación dependiente entre conjuntos de imágenes, las jerarquías de contenido y de codificación podrían superponerse en más de un nivel. Las tramas de vídeo seguirían definiéndose en el nivel de contenido más bajo.

6.1.2 Tramas de vídeo en interfaces compuestas

La figura 8 muestra la jerarquía de secuencias de vídeo para señales en los formatos de 525 líneas y 625 líneas.

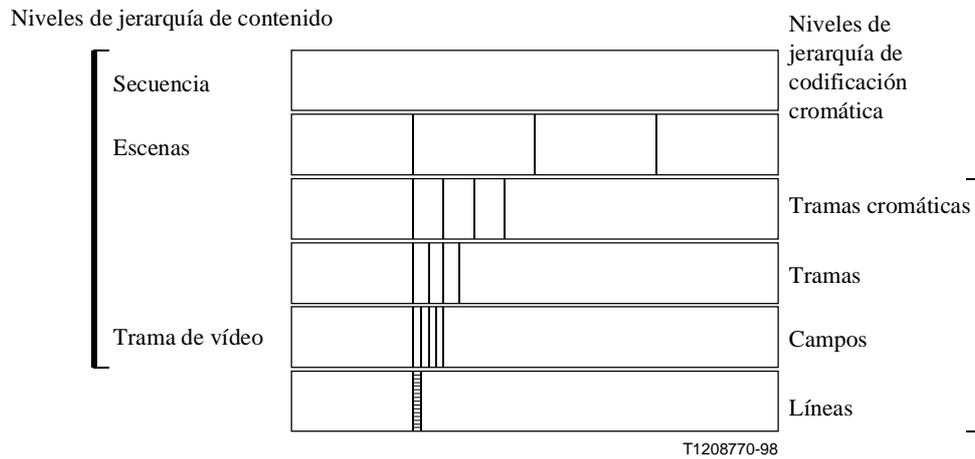


Figura 8/P.931 – Jerarquía de secuencias de vídeo de 525 y 625 líneas

En el caso de formatos entrelazados de 525 y 625 líneas, una **trama de vídeo** se define como un campo, tal como se especifica en la Recomendación BT.470-6. Los campos sucesivos pueden transportar nuevas informaciones al usuario a través de la unidad de presentación, aunque no en cada línea de la unidad de presentación. Además, puesto que la velocidad de campos es superior a la velocidad de tramas, esta definición permite un muestreo más fino de la secuencia de vídeo.

Esta definición también es necesaria en el caso de sistemas de transmisión que, (debido a las restricciones del sistema, tales como la velocidad binaria de transmisión), transmiten tramas de vídeo activas a velocidades inferiores a la velocidad de tramas de interfaz de vídeo y visualizan lo antes posible cada nueva trama activa. Este método de actualización de la visualización puede dar lugar a diferencias considerables entre los campos sucesivos.

6.1.2.1 Conversión de analógico a digital

Los métodos de medición que se describen en las subcláusulas que siguen requieren la digitalización de la señal analógica compuesta. La Recomendación UIT-R BT.601-5 proporciona un método de alta fidelidad de muestreo de la zona de líneas activas analógicas con señales de luminancia de 525 ó 625 líneas. La experiencia indica que la señal de luminancia suministra información suficiente para comparar y establecer la concordancia de las tramas de vídeo a los efectos de las mediciones en esta especificación.

La figura 9 indica cómo pueden organizarse las muestras digitalizadas correspondientes a tramas de vídeo.

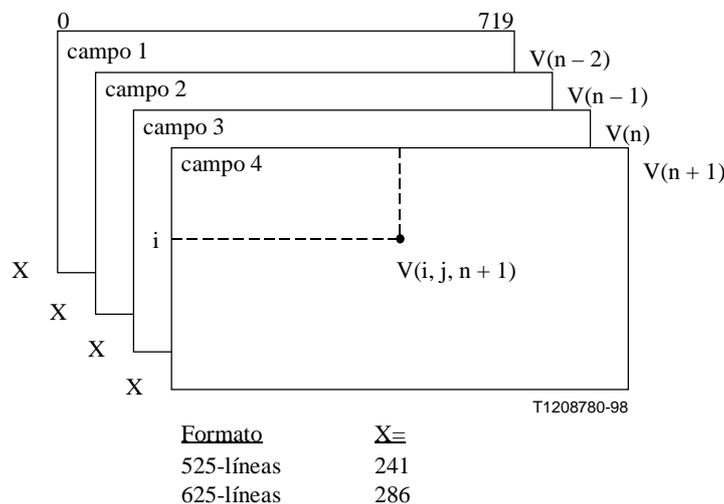


Figura 9/P.931 – Coordenadas recomendadas para la digitalización de tramas de vídeo

$V(n)$ es la trama de vídeo n en el momento $T(n)$, $V(n - 1)$ es la trama de vídeo precedente y $V(n + 1)$ es la trama de vídeo siguiente. $V(i,j,n)$ es el píxel de luminancia (i,j) de la trama de vídeo n en el momento $T(n)$.

Son posibles otros sistemas de coordenadas, pero deberá emplearse la numeración a partir del ángulo superior izquierdo hacia el ángulo inferior derecho, a fin de facilitar la comparación entre diferentes implementaciones de medición. Las comparaciones también se facilitarán aplicando la interfaz de componente digital que se describe en 6.1.3. Un ejemplo de otros sistemas de coordenadas puede ser la asignación de números de línea específicos a partir de las líneas de sincronización verticales [10].

6.1.2.2 Asignación de indicaciones de tiempo

En los sistemas de 525 líneas, el tiempo $T(n)$ asociado a la trama de vídeo n deberá obtenerse inmediatamente después de la digitalización del último píxel de la trama, coordenadas $(242, 719)$, y antes de que empiece la línea siguiente $(10,222 \mu s)$.

En los sistemas de 625 líneas, el tiempo $T(n)$ asociado a la trama de vídeo n deberá obtenerse inmediatamente después de la digitalización del último píxel de la trama, coordenadas $(286, 719)$, y antes de que empiece la línea siguiente $(10,666 \mu s)$.

6.1.2.3 Ganancia, zona activa y alineación espacial

Esta especificación requiere un factor de corrección para la ganancia (g), el desplazamiento de nivel (l), la deriva horizontal (h) y la deriva vertical (v) antes de proceder a las mediciones, en aras de la exactitud.

En la bibliografía [3] se define un método para medir la ganancia media y el desplazamiento de nivel, así como métodos manuales y automatizados para medir la zona de vídeo activo y la deriva de vídeo activo. Dichos métodos se pueden utilizar para asegurar la calidad de la comparación entre trama y trama y los métodos de concordancia de tramas. Una vez medidos y utilizados estos parámetros, deberán especificarse las coordenadas de la zona de vídeo activo y de la deriva de vídeo activo.

Si el sistema de transmisión redimensiona las tramas de entrada (o sea, expande o contrae la escala horizontalmente, verticalmente o en ambas direcciones), estas diferencias adicionales aparecen al comparar las secuencias de entrada y de salida, por lo que puede ser conveniente el empleo de un factor de corrección por las modificaciones de tamaño (z), para reducir al mínimo las diferencias

antes de analizar la concordancia de las tramas. Cuando el redimensionamiento es deliberado y a gran escala, es muy importante compensarlo. No se sabe si prevalece en los sistemas de comunicación multimedios un redimensionamiento de grandes proporciones ni en qué medida daría lugar el redimensionamiento a valores importantes de ruido, pero hay una oportunidad de compensar ese redimensionamiento y proseguir el proceso que aquí se especifica.

De no aplicarse los métodos mencionados, podrían producirse errores de medición. Por ejemplo, si el canal presenta una deriva de vídeo de 4 píxels, un barrido panorámico de la cámara en la misma dirección podría dar lugar a tramas de concordancia incorrecta, con retardos inadecuados. Si no se compensan, las derivas y ganancias o los desplazamientos de nivel aumentarían, por su parte, el nivel de ruido reduciendo la eficacia de la concordancia de las tramas basada en el error cuadrático medio.

En algunas circunstancias, el empleo de métodos simplificados asegura la exactitud de las mediciones. A continuación se indica en qué casos es posible una simplificación de las mediciones.

6.1.2.4 Opciones reconocidas

Una opción reconocida para esta especificación consiste en definir una zona de subtrama por sus coordenadas x-y, para efectuar todas las comparaciones y operaciones de establecimiento de la concordancia en dicha subtrama. Otra solución sería definir la subtrama mediante la zona de acción segura y/o zona de título seguro [7]. Esto permite en la medición excluir las fronteras de las tramas sin determinar exactamente la zona de vídeo activo y mejorar el umbral de ruido de captación evitando errores en las fronteras de las tramas. Si se adopta cualquiera de estas opciones, deberán especificarse las coordenadas de la subtrama.

Al definir la zona de subtrama, deberían considerarse los siguientes aspectos:

- 1) Evitar los límites que no contienen imagen alguna.
- 2) Incluir las zonas de imagen fija que contienen bordes tanto horizontales como verticales y una gama completa de valores de píxel siempre que sea posible, para ayudar a determinar los factores de corrección espacial.
- 3) Incluir las zonas que contienen movimiento, en cuyo defecto la medida del retardo produciría resultados ambiguos.
- 4) El tamaño de la zona de subtrama en píxels es proporcional al tiempo de cálculo. Además, zonas reducidas pueden mejorar la operación de establecimiento de la concordancia de las tramas al localizarse el movimiento, pero deben tener un tamaño suficiente para evitar ambigüedades.

Otra opción es el empleo de un muestreo horizontal inferior durante la digitalización para el almacenamiento de datos. La experiencia indica que una resolución horizontal de 320 píxels en el periodo de línea activa digital puede ser suficiente para efectuar el cálculo de la calidad en estas mediciones y que tal vez no sea necesario introducir correcciones de alineación espacial. Si se adopta esta opción, deberán especificarse los píxels horizontales en el periodo de línea activa digital.

Al medir sistemas cuya sucesión de imágenes es inferior a 30 por segundo, es admisible registrar campos alternos para reducir los requisitos de almacenamiento de la captación, anotando el campo seleccionado (par o impar). Un ejemplo de sistema son los terminales compatibles con la Recomendación H.320, con codificación de vídeo conforme a la Recomendación H.261. Si se adopta esta opción, deberá especificarse el campo seleccionado.

6.1.3 Tramas de vídeo en interfaces de componentes digitales

Cuando se trabaja con vídeo de componentes, se define de nuevo una trama de vídeo como un campo con formato de 525 ó 625 líneas.

6.1.3.1 Organización de las señales

Las señales de vídeo en las interfaces digitales ([8] para conexiones en paralelo y [9] para conexiones en serie) ya han sido convertidas desde su forma analógica, por lo que no se requiere aquí ninguna digitalización adicional. Estas interfaces multiplexan muestras de 8 ó 10 bits de los componentes de vídeo en el orden (C_B, Y, C_R, Y, C_B, Y,...). Los sistemas de medición sólo pueden emplear muestras Y.

Las palabras de información de la secuencia de temporización se insertan en el tren de bits para identificar la línea activa digital. Las líneas activas aparecen después de la palabra de inicio de vídeo activo (SAV, *start of active video*) y terminan antes de la palabra de fin de vídeo activo (EAV, *end of active video*), y contienen 720 muestras de luminancia. Las coordenadas recomendadas para las muestras son las mismas que en el caso de interfaces compuestas.

6.1.3.2 Asignación de indicaciones de tiempo

En este caso también son aplicables los mismos requisitos que para las interfaces compuestas (véase 6.1.2.2).

6.1.3.3 Ganancia, zona activa y alineación espacial

Son también aplicables en este caso los mismos requisitos (véase 6.1.2.3) y opciones (véase 6.1.2.4) de las interfaces compuestas.

6.2 Método del error cuadrático medio para mediciones de vídeo

En esta subcláusula se presenta el método de medición de sistemas aplicando el criterio del error cuadrático medio. Las implementaciones de este método deberán poder suministrar secuencias de tramas de vídeo apropiadas a la entrada de canal. El método requiere asimismo la captación y, de ser necesario, la digitalización del componente luminancia de las secuencias de tramas de vídeo en las interfaces de canal. El suministro y la captación deberán efectuarse de conformidad con lo dispuesto en 6.1 para las interfaces en uso. Una vez halladas las tramas activas y sus correspondientes tramas concordantes en la secuencia de entrada, deberán efectuarse los cálculos temporales deseados según se especifica en la cláusula 5.

El método satisface varias circunstancias especiales, entre ellas los sistemas de alta calidad de transmisión que conservan la integridad de campo entrelazado y el empleo de secuencias de vídeo fuente resultantes del proceso de reducción 3:2 de una película de 24 cuadros por segundo.

6.2.1 Generalidades

La detección de tramas activas dentro una secuencia de tramas de vídeo y la localización de las tramas concordantes entre las secuencias requiere un método normalizado de comparación. Dicho método compara las tramas de vídeo píxel por píxel y hace equivaler la diferencia entre un par de tramas al error cuadrático medio calculado con todos los píxels de interés. En consecuencia, para un par de tramas (una de la secuencia de entrada y una de la secuencia de salida) el error cuadrático medio (MSE, *mean square error*) es:

$$M[V'(m), V(n)] = \frac{1}{K_s} \sum_{j=J_{\min}}^{J_{\max}} \sum_{i=I_{\min}}^{I_{\max}} [V'(i, j, m) - V(i, j, n)]^2$$

donde $V'(i, j, m)$ es el valor del píxel i, j en la trama de salida en el momento $T'(m)$, y $V(i, j, n)$ es el valor del píxel i, j en la trama de entrada en el momento $T(n)$. K_s es el número total de píxels en la subtrama rectangular de interés, dado por:

$$K_s = (I_{\max} - I_{\min} + 1) \times (J_{\max} - J_{\min} + 1)$$

Obsérvese que $V'(i,j,m)$ se ha corregido para tener en cuenta cualquier ganancia, desplazamiento de nivel, deriva horizontal, deriva vertical y escalamiento espacial (si es necesario) entre la entrada y la salida (con los consiguientes factores de corrección g, l, h, v, z):

$$V'(i, j, m) = \frac{V^*(x + v, y + h, m) - 1}{g}$$

donde $V^*(x,y,m)$ es el píxel de salida antes de aplicarse los factores de corrección. Si se precisa el redimensionamiento del vídeo de salida para que concuerde con la entrada:

$$V'(i, j, m) = \frac{V^{**}(\hat{i} + v, \hat{j} + h, m) - 1}{g}$$

en que $V^{**}(m) = f(V^*(m), z)$

y donde $f(V^*(m), z)$ representa una función de redimensionamiento.

Para comparaciones entre tramas adyacentes dentro de una secuencia (por ejemplo, para detectar las tramas activas en la interfaz de salida), $V(i,j,n)$ pasa a ser $V'(i,j,m - 1)$ en la ecuación de MSE expuesta más arriba.

El MSE es un importante factor en el cálculo de la relación señal/ruido de cresta (PSNR, *peak signal-to-noise ratio*), similar a [3]:

$$PSNR = 20 \log_{10} \left[\frac{V_{cresta}}{\sqrt{M[V'(m), V(n)]}} \right] \text{dB}$$

La figura 10 ilustra el proceso de medición de alto nivel conforme al método del MSE.

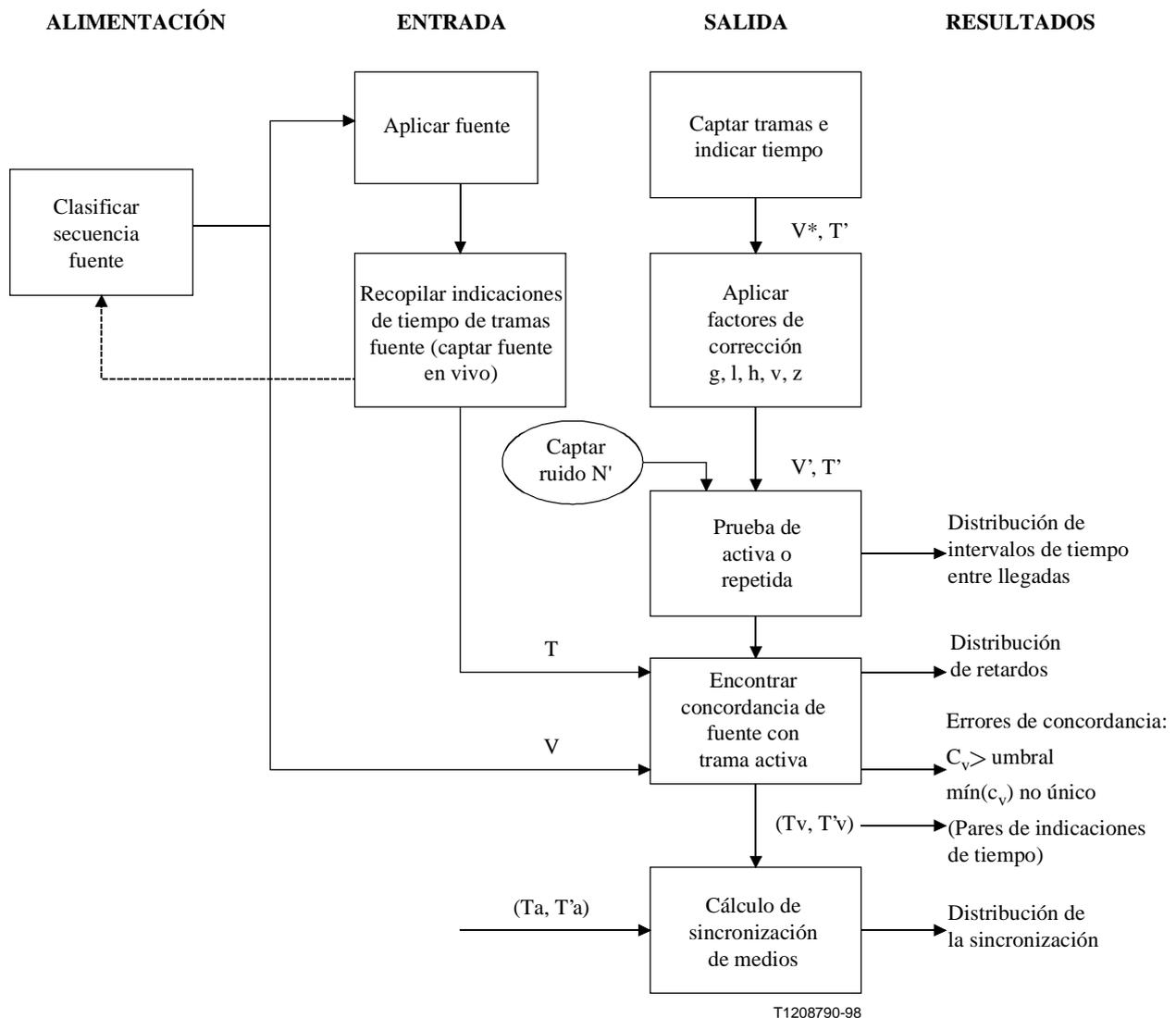


Figura 10/P.931 – Diagrama de flujo de las mediciones de vídeo basadas en el MSE

6.2.2 Calibrado de la diferencia distinguible mínima entre tramas

En esta subcláusula se especifica el método para determinar el ruido (o variaciones no deseadas) en los procesos de digitalización y almacenamiento que recopilan secuencias de tramas de vídeo para su comparación. Este nivel de ruido depende de las opciones específicas elegidas (por ejemplo, el formato de digitalización) y su conocimiento es necesario para efectuar mediciones válidas.

Las condiciones de prueba para calibrar el ruido de captación son las siguientes:

- 1) Aplicar a la entrada del canal una escena de **vídeo inmóvil**. En [2] se define el vídeo inmóvil como "una composición de imágenes de vídeo que no transmite movimiento ni cambios". Es importante mantener la misma relación señal/ruido del vídeo de entrada durante el calibrado y la medición. Una de las técnicas utiliza una secuencia fuente compuesta por una sola trama repetida procedente de una o más de las secuencias de vídeo en movimiento que se tiene intención de seguir probando. El procedimiento no reproduce el ruido en la secuencia fuente y sólo sirve en los casos en que el ruido de captación es considerablemente mayor que el ruido fuente. Para algunas secuencias de prueba y vídeo en vivo quizás sea posible dividir la trama de vídeo en el espacio y definir una subtrama inmóvil (por ejemplo, el fondo) para este calibrado y una subtrama en movimiento para otras mediciones. Las señales de prueba inmóviles (barras cromáticas SMPTE) también se han utilizado con éxito (de nuevo, el ruido

fuente presente en la señal de prueba inmóvil deberá ser el mismo que en la secuencia de prueba).

- 2) Captar (digitalizar y almacenar) la secuencia de tramas correspondiente a la salida del canal. Deberían bastar 30 a 60 tramas. Si el canal emplea compresión digital, se admitirá alcanzar un nivel de calidad sostenido en situación inmóvil, evitando así toda **respuesta de corte de escena** que pudiera degradar la medición del ruido.

En general, el ruido de captación a la entrada será diferente del ruido a la salida. Existen códecs que filtran el ruido fuente para mejorar la señal que hay que codificar.

Para una secuencia de 30 tramas, se calcula el conjunto de valores de MSE de $30 - 1 = 29$ tramas adyacentes, $M[V'(m), V'(m - 1)]$.

El nivel de ruido de captación a la salida es el valor máximo de MSE del conjunto.

$$v'(m) = M[V'(m), V'(m - 1)] \text{ para } m = 2, 3, \dots, 30$$

donde $v'(m)$ es el valor de MSE correspondiente a la trama $V'(m)$, y el ruido de captación N' , es:

$$N' = \text{máx}(v')$$

siendo v' el conjunto de valores de MSE para la secuencia V' . La variación dentro del conjunto de valores de MSE deberá ser pequeña ($< 20\%$), ya que se calcula la media de muchos píxels para cada valor del conjunto. Para la secuencia de entrada se tiene $N = \text{máx}(v)$.

Para contar con un cierto margen entre el nivel de ruido de captación y un umbral de detección de tramas activas, se definen tramas de salida cuyo $v'(m) = M[V'(m), V'(m - 1)] \leq 1,5 \times N'$ corresponde a tramas repetidas. Para una secuencia de origen, se definen tramas cuyo $v(n) = M[V(n), V(n - 1)] \leq 1,5 \times N$ corresponde a **tramas indistinguibles**. Puede haber pequeñas diferencias entre tramas repetidas y tramas indistinguibles, pero el sistema de medición no es capaz de detectarlas de manera fiable. La selección y clasificación de secuencias fuente para la prueba deberá tener en cuenta este umbral. No cabe esperar que se detecten tramas activas cuando las tramas de la secuencia fuente son indistinguibles para el dispositivo de medición. El margen propuesto permite la detección de tramas activas con un mayor grado de confianza.

6.2.3 Pruebas de una secuencia para encontrar diferencias distinguibles

Dada una secuencia de vídeo V , se calcula el conjunto de valores v de MSE y se compara cada miembro del conjunto con el umbral de tramas indistinguibles ($1,5 \times N$). Todas las tramas $V(n)$ cuyos $M[V(n), V(n - 1)] > 1,5 \times N$ presentan diferencias distinguibles respecto de sus tramas precedentes. A todo canal sometido a prueba se le harán llegar tramas de entrada con diferencias distinguibles para la prueba de tramas activas y tramas repetidas.

Cuando se consideren secuencias fuente a utilizar con sistemas de transmisión de alta calidad que conservan la integridad de campo, es más apropiado comparar la **trama de vídeo** actual $V(n)$ con $V(n - 2)$ respecto a campos equivalentes por pares y evitar el error de comparación imputable de desplazamiento espacial entre campos.

El procedimiento siguiente permite efectuar pruebas condicionales para asegurarse de que una trama de vídeo posee diferencias distinguibles (para secuencias fuente con campos entrelazados).

- 1) Calcular $M[V(n), V(n - 1)]$.
- 2) Si el resultado es $\leq 1,5N$, declarar las tramas indistinguibles. En caso contrario, continuar.
- 3) Calcular $M[V(n), V(n - 2)]$.
- 4) Si el resultado es $\leq 1,5N$, declarar las tramas indistinguibles. En caso contrario, continuar.
- 5) La trama $V(n)$ posee diferencias distinguibles.

En 6.2.6 se hace un examen más pormenorizado del tema de la caracterización de escenas fuente o de entrada. Se señala que las secuencias creadas empleando el proceso de reducción 3:2 quedarán fuera de esta clasificación, aun cuando puedan seguir siendo utilizadas en virtud de lo dispuesto en 6.2.7.

6.2.4 Categorización de tramas activas y tramas repetidas

Dada una secuencia de vídeo de salida V' , se calcula el conjunto de valores de MSE $M[V'(m), V'(m-1)]$ y se compara cada valor del conjunto con el umbral, con el objeto de encontrar tramas indistinguibles ($1,5 \times N'$).

Téngase en cuenta que muchos sistemas de transmisión de alta calidad conservan la integridad de campo, aunque introduciendo una distorsión mínima. Para estos sistemas también es apropiado comparar la **trama de vídeo** actual $V'(m)$ con $V'(m-2)$ respecto a campos equivalentes por pares y evitar el error de comparación imputable al desplazamiento espacial entre campos. Cuando las pruebas se efectúan entre interfaces no entrelazadas o se emplean las opciones reconocidas de índice y resolución de captación reducidos, probablemente no sea necesaria la comparación con $V'(m-2)$.

Una trama $V'(m)$ cuyo MSE resulte ser $M[V'(m), V'(m-1)]$ y $M[V'(m), V'(m-2)] > 1,5 \times N'$, en respuesta a una secuencia de entrada que posea diferencias distinguibles, tiene una correspondencia limitada respecto a cualquier trama y deberá clasificarse como **trama activa**.

Una trama $V'(m)$ cuyo MSE resulte ser $M[V'(m), V'(m-1)]$ o $M[V'(m), V'(m-2)] \leq 1,5 \times N'$, en respuesta a una secuencia de entrada que posea diferencias distinguibles, tiene una alta correspondencia con $V'(m-1)$ o $V'(m-2)$ y deberá calificarse de **trama repetida**.

6.2.5 Pruebas de correspondencia entre tramas (tramas concordantes)

Dadas la trama activa m y la secuencia de entrada de tramas X , se calcula el conjunto de valores X de MSE, $M[V'(m), V]$. La trama de entrada que presenta la mejor correspondencia es la que produce el valor mínimo de MSE dentro del conjunto de:

$$c_v(x) = M[V'(m), V(x)] \text{ para } 1 \leq X \leq X$$

c_v es el conjunto de valores MSE correspondientes a la trama $V'(m)$ en *comparación* con cada trama de la secuencia V , y la trama de entrada que concuerda mejor con $V'(m)$ viene definida por:

$$C_v = \text{mín}(c_v)$$

[el error mínimo (MSE) representa la máxima correspondencia o mejor concordancia entre tramas].

Hay reglas que permiten mejorar el proceso de concordancia y reducir las ambigüedades. Puede suceder que una sola trama activa guarde una estrecha correspondencia con más de una trama de entrada. Estos casos deberían reducirse al mínimo utilizando criterios de concordancia basados en métodos de comparación de píxels (MSE), pero determinadas circunstancias aumentan la probabilidad de que haya ambigüedades. Son las siguientes:

- Una distorsión espacial extrema como consecuencia de una velocidad binaria de transmisión baja, el empleo de un formato de trama digital de baja resolución, etc.
- Contenido fuente – movimiento intenso (causa de arrastre corto u otras distorsiones), movimiento repetitivo, intervalos inmóviles dentro de una secuencia.
- Una velocidad baja de tramas de salida activas produce muchas tramas fuente de posible concordancia.
- El empleo de la interpolación de tramas puede hacer aún más difícil la concordancia.

La experiencia ha demostrado que la aplicación de este método permite hacer frente a casos en los que se dan esas circunstancias difíciles.

Las siguientes reglas pueden ser útiles en la solución de concordancias ambiguas, con lo que se facilita la automatización del método:

- 1) Requerir una concordancia uno a uno: No más de una trama activa puede concordar con una trama de entrada dada. Uno de los motivos posibles de la existencia de una doble concordancia es la detección errónea de una trama activa. Si los resultados de la comparación ponen de manifiesto una concordancia doble, hay que informar de esta aparente irregularidad.
- 2) Aplicar la secuencia: Por ejemplo, (con tramas no entrelazadas) si $V'(m)$ concuerda con $V(n)$, la siguiente trama activa $V'(m+2)$ habrá de concordar con $V(n+1)$ o $V(n+2)$ o $V(n+3)$, etc. No se admite que una $V'(m+2)$ concuerde con $V(n-1)$ o $V(n)$, pero semejante suceso deberá señalarse como un posible error.
- 3) Reconocer un retardo mínimo: No se admiten concordancias que den lugar a un retardo menor que t_{\min} . Se excluye el retardo negativo estableciendo $t_{\min} \geq 0$.
- 4) Aceptar el estado de no concordancia: Algunas tramas activas pueden contener demasiada distorsión como para concordar con la secuencia transmitida. Tales tramas deberán ser contadas y notificadas junto con el umbral de no concordancia utilizado. Los usuarios del sistema de medición deberán determinar la gama habitual de valores MSE de concordancia para el sistema de transmisión sometido a prueba y establecer dicho umbral por encima de esa gama.
- 5) Diagnóstico: El proceso de concordancia podría repetirse a partir del extremo opuesto de la secuencia, para comprobar la presencia de un menor número de concordancias ambiguas o de no concordancias. Además hay que invertir los sentidos de las reglas.
- 6) Probar la trama siguiente: Si la trama activa siguiente de la secuencia de salida tiene una sola concordancia en la secuencia transmitida, utilícese su concordancia y aplíquense las reglas antes mencionadas a la trama activa precedente.
- 7) Seleccionar el mejor caso aleatoriamente: Si se mantiene la ambigüedad después de aplicar las reglas mencionadas, puede emplearse una selección aleatoria. Sin embargo, se recomienda el cálculo de MSE con resolución suficiente para reducir al mínimo tales casos. Los errores introducidos en la distribución por un proceso aleatorio deberían anularse al cabo de una secuencia y algunas estadísticas sumatorias no se verían afectadas. Las selecciones así efectuadas deberán ser contadas y notificadas.
- 8) Si los resultados utilizando una escena específica tienden a requerir intervenciones masivas y soluciones en base a estas reglas, deberá intentarse la medición utilizando una escena diferente.

6.2.6 Clasificación de secuencias fuente para los métodos del error cuadrático medio

El éxito de los métodos basados en el MSE depende de la utilización de frecuencias fuente apropiadas. Como se ha dicho, la secuencia fuente deberá presentar diferencias de trama a trama que sean distinguibles para el dispositivo de medición y han de evitarse el movimiento repetitivo y los intervalos de vídeo inmóvil, que sin duda provocarán ambigüedades de concordancia. Cuando se utilice una zona de subtrama, el proceso deberá adoptar una zona similar como base de clasificación. El siguiente procedimiento permite determinar los fragmentos convenientes para la medición en vídeo:

- 1) Elegir la primera trama de vídeo de la secuencia fuente y compararla con las demás tramas de la secuencia.
- 2) Registrar el número y posición de todas las tramas indistinguibles (como se describe en 6.2.2).

- 3) Análisis: El intervalo existente entre todas las tramas indistinguibles deberá ser suficiente para resolver las ambigüedades de alineación de entrada y salida utilizando ciertas estimaciones previas de intervalos de tiempo entre llegadas de tramas y otras informaciones. Por ejemplo, si se sabe que el retardo de transmisión es < 2 s, las tramas indistinguibles pueden aparecer con una separación ≥ 2 s.
- 4) Repetir los pasos precedentes al menos con las primeras X tramas (por ejemplo, $X = 60$). Si se da prioridad a las primeras tramas de la secuencia es porque ciertos resultados de las reglas de comparaciones pueden depender de los resultados previos.

Obsérvese además que:

Las secuencias fuente con tramas adyacentes similares (distinguibles, pero en escaso grado) tendrán mayores oportunidades de producir concordancias ambiguas con las tramas activas. Una prueba de diferencias de trama a trama revelará la amplitud de estas similitudes, según se describe en 6.2.3.

En la bibliografía [1] se especifica un conjunto de secuencias fuente apropiado para la evaluación de sistemas de teleconferencia/videotelefonía. Sin embargo, algunas secuencias contienen secciones de vídeo inmóvil que deberán evitarse en las mediciones del retardo (las secciones inmóviles son deseables para el calibrado). Se dispone de otros conjuntos de escenas, que pueden servir para otras aplicaciones. Incluso el vídeo en vivo está permitido, a condición de que el sistema de medición pueda captar y almacenar durante una medición secuencias tanto de entrada como de salida y de que se evalúe seguidamente la conveniencia de la secuencia de entrada. Deberán evaluarse todas las secuencias de entrada mediante una implementación de medición específica conforme a los procedimientos descritos aquí y en 6.2.3.

6.2.7 Consideraciones sobre el empleo de secuencias fuente con reducción 3:2

Varias situaciones pueden dar lugar a tramas duplicadas (y, por consiguiente, indistinguibles) en una secuencia de origen:

- 1) El contenido de la secuencia permanece inmóvil, sin movimiento ni cambios.
- 2) Se ha creado la secuencia de vídeo a partir de una película mediante el proceso de reducción 3:2. Cada quinto campo (en el formato de 525 líneas) es una repetición del campo correspondiente anterior (par o impar).
- 3) Luego de un proceso de reducción 3:2, la secuencia de vídeo es objeto de una elaboración que produce repeticiones de campo no periódicas.

Estrictamente hablando, la presencia de cualquiera de estas tres situaciones haría que la secuencia fuente no pasara con éxito las pruebas de calificación. Es imposible efectuar mediciones de retardo o de velocidad de tramas con una secuencia inmóvil. Sin embargo son posibles ciertas mediciones en los casos de reducción 3:2.

Dos problemas se presentan al aplicar el método de MSE de medición en los casos originados en la reducción 3:2:

- El proceso de concordancia de las tramas activas con sus secuencias fuente correspondientes es más complejo cuando algunas tramas fuente se repiten. De producirse concordancias ambiguas, los cálculos de retardo para esas tramas son dudosos. Puede resolverse la ambigüedad examinando la secuencia fuente de las tramas repetidas.
- La velocidad de tramas real puede ser mayor que la que se calcula utilizando las tramas activas únicamente. Algunas tramas de la secuencia de salida clasificadas como repetidas pueden haber recorrido el canal. Es posible hallar esas tramas mediante un proceso de concordancia complementario.

Para efectuar mediciones con una secuencia fuente de reducción 3:2 es necesario crear un registro de las tramas repetidas en la secuencia fuente, tal como se describe en 6.2.4 (pero aplicado a la secuencia fuente). Tras la clasificación de tramas activas y tramas repetidas en la secuencia de salida cabe esperar que el proceso de concordancia establezca pares de tramas activas con tramas fuente no repetidas y, en algunos casos, con tramas fuente repetidas. La regla 1 de 6.2.5 requiere una concordancia de tramas uno a uno, por lo que se determinará la mejor concordancia remitiéndose al registro de tramas fuente repetidas para elegir la trama fuente no repetida. Puesto que las tramas originales que se convierten en tramas activas siempre llegan antes que sus repeticiones, el proceso de concordancia siempre seguirá la secuencia normal, del inicio al fin (nunca se intentará la regla 5).

Una vez efectuado el proceso de concordancia en todas las tramas activas, puede comenzar un segundo proceso de concordancia con las tramas repetidas. En este paso se examinan tanto las tramas fuente repetidas como cualquier trama no concordante del primer paso, eludiendo así la mayoría de las tramas fuente no repetidas. De existir una sola concordancia entre tramas repetidas en las secuencias fuente y de salida (tras establecer la concordancia de las tramas de salida activas adyacentes), el tiempo de llegada de la trama repetida deberá incluirse en la distribución de intervalos de tiempo entre llegadas a la salida. Ha de aplicarse nuevamente la concordancia de tramas uno a uno, en la eventualidad de que una trama fuente repetida sea duplicada por el sistema de transmisión. Ningún cálculo de tiempo debe utilizar las tramas de salida repetidas sin considerar las tramas fuente de concordancia única.

6.2.8 Factores que influyen en la exactitud y estabilidad de las mediciones

En muchos sistemas de transmisión de vídeo, el decodificador debe suministrar tramas de vídeo a la salida conforme a un régimen de presentación visual periódica (como una interfaz analógica compuesta). Si los relojes de presentación visual de entrada y salida no están sincronizados, debe añadirse una memoria intermedia al decodificador. Cuando el decodificador dispone de una trama de vídeo lista para visualizar, debe esperar la siguiente oportunidad de salida, aumentando así el retardo global del sistema. A este intervalo de espera del decodificador se le llama *retardo de salida*.

El retardo de salida está limitado por el intervalo entre actualizaciones de visualización. En el caso de sistemas de transmisión con interfaces compuestas capaces de actualizar en las fronteras de los campos, el retardo de salida máximo es de 16,7 ms. Con actualizaciones en las fronteras de las tramas de 525 líneas, el máximo es de 33 ms. El retardo de salida real será un valor aleatorio situado entre 0 y el valor máximo.

Cuando los relojes de forma de onda de entrada y de salida tienen un pequeño desplazamiento de frecuencia, el retardo de salida variará con el tiempo. Si el desplazamiento de frecuencia es constante, el retardo de salida oscilará entre sus valores extremos a intervalos regulares. Si los relojes están sincronizados con osciladores independientes controlados por un haz de cesio, el retardo de salida sufrirá modificaciones de < 1 ms cada 13 900 horas. Cuando la precisión del reloj proviene de osciladores de cuarzo independientes (con un desplazamiento de 2 ppm), el retardo de salida variará de manera cíclica a lo largo de la gama de 33 ms cada 4,58 horas.

Las indicaciones de tiempo con resolución de subcampo permiten caracterizar el retardo de salida como un componente inextricable de todo el retardo del sistema de transmisión.

6.3 Métodos de medición para vídeo basados en códigos horarios dentro de la trama

En algunas situaciones es posible insertar en el vídeo de entrada símbolos visibles que pueden utilizarse para identificar de manera exclusiva cada trama de entrada. Dichos símbolos se llevarán a la salida del sistema y podrían utilizarse como un modo sencillo de medición de la velocidad de tramas y el retardo. Los métodos correspondientes y las limitaciones a su empleo quedan en estudio.

7 Mediciones de audio

7.1 Recopilación de tramas de audio para mediciones

7.1.1 Descripción de las tramas de audio

Una trama de audio es un grupo de muestras de audio consecutivas. El número preferido de muestras en una trama de audio depende de la velocidad de muestreo de audio y se indica en la cláusula 5. La siguiente figura 11 muestra la posición de las tramas de audio en el contexto de jerarquías de codificación y contenido de audio (en este ejemplo, las tramas de codificación son de menor duración que las tramas de audio).

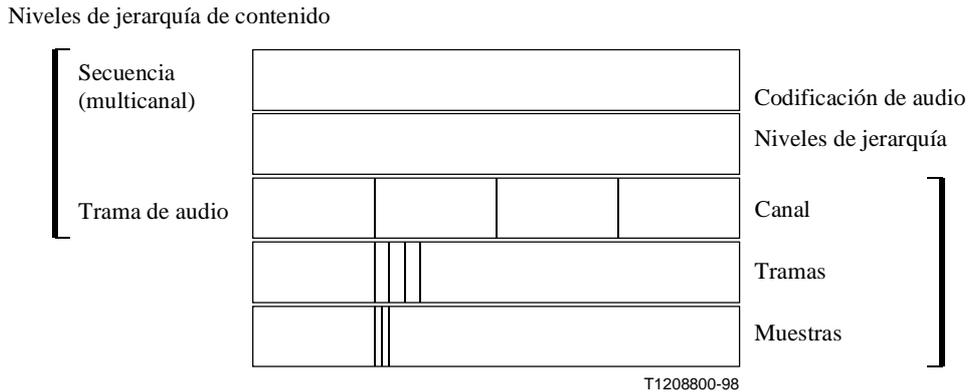


Figura 11/P.931 – Jerarquías de contenido y codificación de audio

7.1.2 Conversión de señales analógicas en digitales

Los métodos de medición descritos a continuación requieren la digitalización de las señales analógicas de audio. El proceso de digitalización da por resultado muestras de audio que pueden agruparse en tramas de audio. La velocidad de muestras de audio está determinada por la anchura de banda necesaria para las mediciones subsiguientes. Empleando el teorema de Nyquist, la velocidad de muestreo deberá equivaler por lo menos al doble de la anchura de banda de la medición. Para señales de audio exclusivamente vocales, es suficiente una velocidad de muestreo de 8000 muestras por segundo. Podrían requerirse velocidades de muestreo superiores para otras señales de audio. El proceso de digitalización deberá dar como resultado una precisión de 8 bits como mínimo. En función de la relación señal/ruido de la señal de audio, una precisión adicional de hasta 16 bits favorecerá generalmente los métodos de medición que siguen. El proceso de digitalización deberá incluir un filtrado de paso bajo apropiado antialiasing, y deberá ser concordante con la impedancia y el equilibrado de las señales de audio.

7.1.3 Asignación de indicaciones de tiempo

El tiempo $T(n)$ asociado a la trama de audio n ha de captarse inmediatamente después de digitalizar la última muestra de la trama de audio n y antes de digitalizar la muestra siguiente. Las indicaciones de tiempo para cada muestra dentro de una trama pueden deducirse de la indicación de tiempo de la trama de audio, puesto que se conoce la velocidad de muestreo.

7.2 Medición del retardo en audio

Muchos canales que podrían interesar están en condiciones de entregar señales de audio utilizables sin conservar las formas de onda de audio desde la entrada hasta la salida. Hay ejemplos de especificaciones de codificación de audio en varias Recomendaciones, G.728 (16 kbit/s), G.729 (8 kbit/s), CELP y G.723 (6,4 kbit/s y 5,3 kbit/s). Lo anterior significa que una medición de retardo sólida no debe basarse únicamente en las formas de onda de audio. Las mediciones que aquí se describen corresponden a una fase aproximada que emplea envolventes de audio y a una fase de precisión que emplea densidades espectrales de potencia de audio (PSD, *power spectral densities*). Las envolventes de audio y las PSD de audio suelen conservarse con relativa precisión en la mayoría de los canales. La figura 12 presenta un diagrama de flujo del proceso de medición.

Este proceso en dos etapas resulta eficaz, ya que en la fase aproximada se explora una amplia gama de posibles valores de retardo, pero de baja resolución. Si la misma gama se explorase con una alta resolución, los cálculos requeridos serían muchos más. Una vez concluido el trabajo de la fase aproximada, sus medidas de baja resolución pueden afinarse frecuentemente a medidas de alta resolución gracias a la fase de precisión que sigue. En la fase de precisión, sólo se necesita explorar una gama reducida de posibles valores de retardo, correspondiente a las ambigüedades resultantes de la medición aproximada. En algunos canales, las PSD de audio no se conservan suficientemente y la medición precisa no es posible. En otros casos, múltiples mediciones precisas darán resultados diversos. En tales situaciones, la medida aproximada del retardo, con toda su ambigüedad, se aceptará como medida de retardo definitiva.

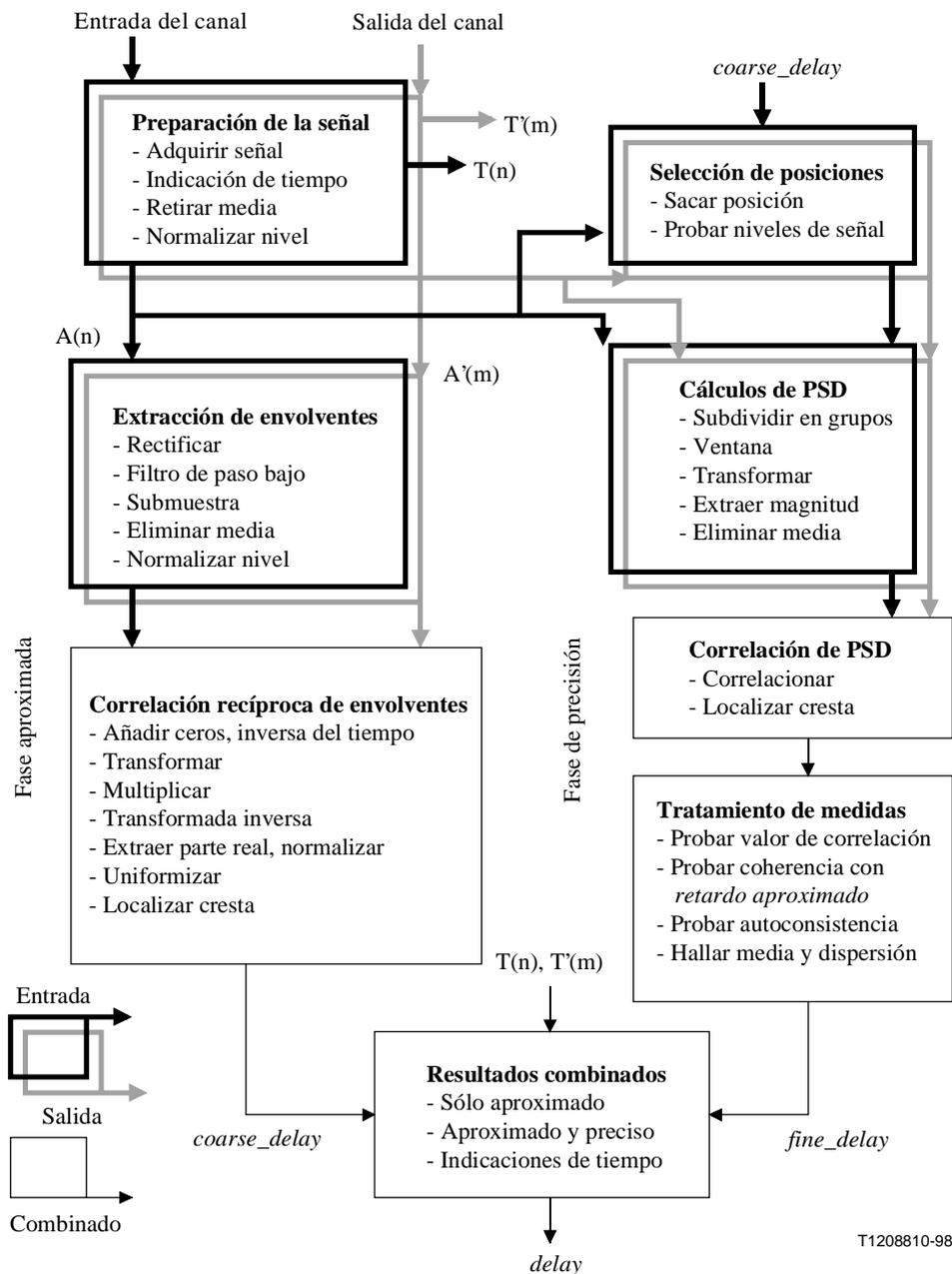


Figura 12/P.931 – Diagrama de flujo de mediciones de retardo de audio

7.2.1 Generalidades

Antes de proceder a las mediciones, ha de reunirse una serie de tramas de audio de la entrada y la salida del canal. Cuantas más tramas de audio se utilicen, mayor será la fiabilidad y complejidad de las mediciones. Si un grupo de tramas de audio contiene únicamente las pausas entre las palabras o frases de una conversación hablada o bien tonos continuos, no será posible efectuar una medición fiable y deberán obtenerse tramas de audio adicionales antes de proceder a la medición. Para detectar una condición de nivel de audio insuficiente, deberá compararse el nivel de RMS de muestras de audio del grupo de tramas de audio obtenidas de la entrada del canal con el nivel de RMS nominal de la entrada del canal. Dicho nivel nominal puede tomarse de las especificaciones de la entrada del canal o, en su defecto, medirse. Si el nivel de RMS de las muestras del grupo de tramas de audio obtenidas de la entrada del canal está más de 30 dB por debajo del valor nominal de entrada del canal, será necesario obtener tramas de audio adicionales antes de proceder a la medición. Similarmente, el nivel de RMS de muestras de audio del grupo de tramas de audio obtenidas de la

salida del canal deberá compararse con el nivel de RMS nominal de la salida del canal. Si el nivel de RMS de las muestras del grupo de tramas de audio obtenidas de la salida del canal está más de 30 dB por debajo del nivel nominal de salida del canal, deberán obtenerse tramas de audio adicionales antes de proceder a la medición.

Para señales vocales típicas, el uso de grupos mayores de tramas de audio reduce la posibilidad de que el grupo contenga sólo silencios. Aparte de esta consideración, un número mayor de tramas de audio aporta más datos para la medición que, por lo tanto, será más fiable. Para señales de audio exclusivamente vocales es preferible tomar 256 tramas de audio de la entrada y de la salida del canal. La medición también será posible con 128 ó 64 tramas. Con una velocidad de muestreo de 8000 tramas por segundo, conteniendo cada trama 128 muestras, las opciones mencionadas corresponden aproximadamente a 4 segundos, 2 segundos o 1 segundo respectivamente de señales vocales. El retardo de audio previsto no deberá representar más del 25% de la duración de la señal vocal utilizada en la medición. Utilizando 256 tramas de señal vocal (4 segundos), son medibles retardos de hasta 1 segundo. Utilizando 64 tramas (1 segundo), sólo se medirán retardos de 250 ms. Las mediciones se calculan más eficazmente cuando el número de tramas obtenidas sea una potencia de 2.

7.2.2 Preparación de la señal

Cuando se precisa medir el retardo de audio, la trama de audio de adquisición más reciente a partir de la entrada del canal, $A(n)$, se concatena con cierto número de tramas de audio obtenidas anteriormente [por ejemplo, $A(n)$, $A(n-1)$, ... $A(n-255)$], para formar el historial más reciente de muestras de la entrada del canal. De manera similar, las tramas de audio obtenidas más recientemente de la salida del canal, $A'(m)$, se concatena con el mismo número de tramas de audio obtenidas anteriormente [por ejemplo, $A'(m)$, $A'(m-1)$, ... $A'(m-255)$], para formar el historial más reciente de muestras de la salida del canal. Como se indica en la cláusula 5, la diferencia de tiempo entre ambos procesos de adquisición es $T'(m) - T(n)$. Los valores positivos indican que la adquisición a la salida del canal se produce después de la adquisición a la entrada del canal.

Las muestras de la entrada se ordenan en una matriz denominada *ref*, cuyos elementos son $ref(1)$, $ref(2)$, ... $ref(L1)$. Las muestras de la salida se ordenan en una matriz de las mismas dimensiones denominada *test*, cuyos elementos son $test(1)$, $test(2)$, ... $test(L1)$. Seguidamente se extrae el valor medio de cada matriz para eliminar todo componente DC de las señales de audio digitalizadas:

$$\begin{cases} ref(i) = ref(i) - \frac{1}{L1} \cdot \sum_{j=i}^{L1} ref(j) \\ test(i) = test(i) - \frac{1}{L1} \cdot \sum_{j=1}^{L1} test(j) \end{cases} \quad \text{para } 1 \leq i \leq L1$$

A continuación, se normaliza cada matriz a un nivel de RMS común:

$$\begin{cases} ref(i) = ref(i) \cdot \left[\frac{1}{L1-1} \sum_{j=1}^{L1} ref(j)^2 \right]^{-\frac{1}{2}} \\ test(i) = test(i) \cdot \left[\frac{1}{L1-1} \sum_{j=1}^{L1} test(j)^2 \right]^{-\frac{1}{2}} \end{cases} \quad \text{para } 1 \leq i \leq L1$$

7.2.3 Fase aproximada

La metodología de medición comienza con una fase aproximada en la que se extraen las envolventes de audio correlacionadas entre sí. Las envolventes de audio se conservan sin grandes modificaciones en la mayoría de los canales.

7.2.3.1 Extracción de envolventes

El cálculo de envolventes de audio es como sigue. Las señales de audio digitalizadas en *ref* y *test* se rectifican tomando el valor absoluto de cada muestra. Puesto que la fase de precisión requerirá las señales de audio originales digitalizadas en *ref* y *test*, las señales rectificadas y otros resultados intermedios subsiguientes se almacenan en las matrices temporales *ref_temp* y *test_temp*:

$$ref_temp(i) = |ref(i)|, test_temp(i) = |test(i)|, 1 \leq i \leq L1$$

A continuación, las señales rectificadas se someten a un filtrado de paso bajo, para crear envolventes de audio con una anchura de banda de aproximadamente 125 Hz. Es este filtrado de paso bajo y el submuestreo subsiguiente lo que determina la resolución reducida y la carga de cálculo reducida de la fase aproximada. El factor de reducción de anchura de banda y el factor de submuestreo se especifican mediante la variable *B*. En el cuadro 2 se presentan valores apropiados de *B* correspondientes a algunas velocidades de muestreo de audio corrientes.

Cuadro 2/P.931 – Valores del factor de reducción de anchura de banda B

| Velocidad de muestreo en audio (muestras/segundo) | B |
|--|-----|
| 8 000 | 32 |
| 16 000 | 64 |
| 32 000 | 128 |
| 44 100 | 176 |
| 48 000 | 192 |

Cuando la velocidad de muestreo en audio se eleva a 8000 muestras por segundo, será necesario reducir la anchura de banda por un factor de $B = 32$, pasando de una anchura de banda nominal de 4000 Hz a una anchura de banda nominal de 125 Hz. La reducción de anchura de banda requerida puede calcularse con suficiente aproximación empleando un filtro Butterworth de paso bajo de séptimo orden de respuesta a impulso infinita (IIR, *infinite impulse response*), con un punto de -3 dB en 125 Hz. La aplicación directa es la siguiente:

$$out(i) = \sum_{j=0}^7 b_j \cdot in(i-j) - \sum_{j=1}^7 a_j \cdot out(i-j), 1 \leq i \leq L1$$

donde

$$out(i) = in(i) = 0, i \leq 0$$

Se ha de tener la precaución de eliminar cualquier muestra de la salida del filtro que pueda contener un transitorio de puesta en funcionamiento del filtro. En la implementación directa que aquí se describe, el transitorio de puesta en funcionamiento se limita a unas 400 muestras. En el cuadro 3 se presentan los coeficientes de filtro.

Cuadro 3/P.931 – Valores de coeficiente para filtro de Butterworth de paso bajo de séptimo orden IIR

| j | a_j | b_j |
|----------|----------------------|-------------------------------|
| 0 | 1,00000000 | 0,00553833 x 10 ⁻⁷ |
| 1 | -6,55883158 | 0,03876830 x 10 ⁻⁷ |
| 2 | 18,44954612 | 0,11630512 x 10 ⁻⁷ |
| 3 | -28,85178274 | 0,19384125 x 10 ⁻⁷ |
| 4 | 27,08958968 | 0,19384206 x 10 ⁻⁷ |
| 5 | -15,27097592 | 0,11630465 x 10 ⁻⁷ |
| 6 | 4,78557610 | 0,03876843 x 10 ⁻⁷ |
| 7 | -0,64312159 | 0,00553831 x 10 ⁻⁷ |

Ambas matrices, *ref_temp* y *test_temp*, se someten a un filtrado de paso bajo mediante dicho filtro. Las siguientes *ref_temp* y *test_temp* se submuestran reteniendo únicamente cada *B*-ésima muestra, con lo que se obtiene un total de *L2* muestras. Por ejemplo, para *B* = 32, se retendrían las muestras 1, 33, 65, etc. Cuando son 256 tramas de audio, cada una de ellas con 128 muestras, se utiliza como dato de cálculo en la fase aproximada *L1* = 32 768 y *L2* = *L1*/*B* = 1024 muestras, como resultado del proceso de submuestreo. Tanto *ref_temp* como *test_temp* contienen ahora envolventes de audio. Por último, se normalizan las envolventes de audio en *ref_temp* y *test_temp*. Se elimina el valor medio de cada matriz y cada matriz se divide por su desviación típica, para normalizar cada una de ellas a un nivel RMS común.

$$ref_temp(i) = ref_temp(i) - \frac{1}{L2} \cdot \sum_{j=1}^{L2} ref_temp(j)$$

$$test_temp(i) = test_temp(i) - \frac{1}{L2} \cdot \sum_{j=1}^{L2} test_temp(j)$$

$$ref_temp(i) = ref_temp(i) \cdot \left[\frac{1}{L2-1} \sum_{j=1}^{L2} ref_temp(j)^2 \right]^{-\frac{1}{2}}$$

$$test_temp(i) = test_temp(i) \cdot \left[\frac{1}{L2-1} \sum_{j=1}^{L2} test_temp(j)^2 \right]^{-\frac{1}{2}}, 1 \leq i \leq L2$$

7.2.3.2 Correlación recíproca de envolventes

La correlación recíproca entre envolventes de audio en *ref_temp* y *test_temp* se calcula mediante una convolución circular, a su vez calculada mediante transformadas discretas de Fourier (DFT, *discrete Fourier transforms*) o transformadas rápidas de Fourier (FFT, *fast Fourier transforms*). En primer lugar, se ensancha la matriz *ref_temp* de la longitud *L2* a la longitud *2 · L2* añadiendo *L2* ceros. En el ejemplo mencionado, se añadirán *L2* = 1024 ceros, para llegar a un tamaño final de matriz de 2048. A continuación, la matriz *test_temp* se invierte en el tiempo. Para efectuarlo por sustitución, se intercambian las muestras 1 y *L2* de *test_temp* así como las muestras 2 y *L2* - 1, 3 y *L2* - 2, etc. Si *L2* es par, el último intercambio es entre las muestras *L2*/2 y *L2*/2 + 1. Si *L2* es impar, el último intercambio es entre las muestras *L2*/2 - 1/2, y *L2*/2 + 3/2. Efectuada esta inversión temporal, *test_temp* se ensancha de la longitud *L2* a la longitud *2 · L2* añadiendo *L2* ceros.

Seguidamente se transforman *ref_temp* y *test_temp* mediante DFT o FFT. Si la longitud de la matriz $2 \cdot L2$ es una potencia de dos, puede utilizarse FFT. Si $2 \cdot L2$ no es una potencia de dos, puede utilizarse DFT. Como alternativa, el número de ceros añadido en el paso anterior puede aumentarse hasta que la longitud de la matriz sea una potencia de dos, tras lo cual puede utilizarse FFT. En todo caso puede emplearse un algoritmo de transformación por sustitución, que da como resultado versiones transformadas de *ref_temp* y *test_temp* que sustituyen a las versiones anteriores. Las transformaciones dan como resultado números complejos.

A continuación, las muestras complejas almacenadas en *ref_temp* y *test_temp* se multiplican muestra por muestra y los resultados complejos pasan a una nueva matriz denominada *cross_corr*, cuya longitud es igual a la de *ref_temp* y *test_temp*:

$$cross_corr(i) = ref_temp(i) \cdot test_temp(i), \text{ para } i = 1 \text{ to } 2 \cdot L2$$

La matriz *cross_corr* se somete ahora a la transformada rápida de Fourier inversa o la transformada discreta de Fourier inversa, en función de su longitud. Puede emplearse una transformación por sustitución. Teóricamente, los contenidos resultantes de *cross_corr* deberían ser números reales. En la práctica, los cálculos con precisión finita producen un pequeño componente imaginario. Al llegar aquí, se retiene la parte real de *cross_corr* y se descarta la parte imaginaria. Luego se normaliza cada resultado en *cross_corr*:

$$cross_corr(i) = cross_corr(i)/(L2 - 1), 1 \leq i \leq 2 \cdot L2$$

Obsérvese que esta normalización se requiere para obtener valores de correlación recíproca verdaderos entre -1 y $+1$, pero no afecta a los pasos siguientes de uniformización o localización de crestas.

La matriz *cross_corr* conserva los valores de las correlaciones recíprocas entre los envolventes vocales de *ref_temp* y *test_temp* en cada posible deriva de dichas envolventes. Los resultados se uniformizan a través de un filtro FIR simétrico de paso bajo de segundo orden y se almacenan en una matriz de correlación recíproca uniformizada:

$$cross_corr_s(i) = .25 \cdot cross_corr(i - 1) + .5 \cdot cross_corr(i) + .25 \cdot cross_corr(i + 1), \\ 2 \leq i \leq 2 \cdot L2 - 1, cross_corr_s(i) = cross_corr(i), i = 1, 2 \cdot L2$$

Después de esta uniformización, se selecciona el valor mayor de *cross_corr_s* como indicación del retardo aproximado:

$$coarse_delay = (L2 - j) \cdot B \text{ muestras}$$

donde:

$$cross_corr_s(j) > cross_corr_s(i), 1 \leq i \leq 2 \cdot L2, i \neq j$$

El margen de incertidumbre que en este punto tiene el valor de *coarse_delay* se cifra en $\pm B$ muestras. Si *cross_corr_s* no contiene ningún valor máximo único, será necesario proceder a una nueva medición empleando nuevas muestras de audio.

7.2.4 Fase de precisión

En muchos casos, el margen de incertidumbre de $\pm B$ muestras inherente a la medición aproximada del retardo de audio puede reducirse mediante la fase de precisión de la medición del retardo.

7.2.4.1 Selección de posiciones

La fase de precisión tiene lugar en *n1* posiciones respecto de la señal de audio adquirida. Cuando las señales de audio son exclusivamente vocales y se emplean 256 tramas de audio en la medición del retardo de audio, el valor de *n1* es de 6. Otros valores de *n1* podrían ser más apropiados para otras

señales de audio. En cada posición, se explora una gama de valores de retardo potencial entre $-3 \cdot B$ a $3 \cdot B$ muestras.

Las posiciones en las que se lleva a cabo la fase de precisión se seleccionan aleatoriamente. En cada posición se obtienen $8 \cdot B$ muestras, de la matriz *ref*, que se almacenan en *ref_temp*, mientras que de la matriz *test* se obtienen $2 \cdot B$ muestras, que se almacenan en *test_temp*. Las muestras obtenidas de *test* están desplazadas por el retardo de la medición aproximada:

$$ref_temp(i) = ref(location - 4 \cdot B - 1 + i), 1 \leq i \leq 8 \cdot B$$

$$test_temp(i) = test(location + coarse_delay - B - 1 + i), 1 \leq i \leq 2 \cdot B$$

donde *location* es una variable de intervalo pseudo-aleatoria distribuida uniformemente:

$$[\text{máx}(4 \cdot B + 1, 1 - coarse_delay + B), \text{mín}(Ll - 4 \cdot B + 1, Ll - coarse_delay - B + 1)]$$

La medición de retardo con precisión no será efectiva en las regiones de silencio o de tono sostenido. En cada posición se efectúan dos pruebas de nivel de las regiones de silencio, para asegurarse de que la señal de audio se encuentra allí dentro de 30 dB a partir del nivel de señal de audio medio:

$$-30 \leq 10 \cdot \log_{10} \left[\frac{1}{8 \cdot B - 1} \sum_{i=1}^{8 \cdot B} ref_temp(i)^2 \right], -30 \leq 10 \cdot \log_{10} \left[\frac{1}{2 \cdot B - 1} \sum_{i=1}^{2 \cdot B} test_temp(i)^2 \right]$$

De no dar resultado ninguna de las pruebas de nivel, deberá seleccionarse una nueva posición.

7.2.4.2 Cálculos de la densidad espectral de potencia

La fase de precisión trabaja mediante densidades espectrales de potencia (PSD) de audio correlacionadas entre sí, en cada una de las posiciones seleccionadas. El cálculo de PSD se realiza como sigue. Las $8 \cdot B$ muestras de *ref_temp* se dividen en grupos de $2 \cdot B$ muestras por grupo. Existen $6 \cdot B + 1$ grupos de este tipo. Cada grupo de muestras se almacena en una matriz denominada *ref_temp_i*:

$$ref_temp_i(j) = ref_temp(i + j - 1), 1 \leq i \leq 6 \cdot B + 1, 1 \leq j \leq 2 \cdot B$$

Cada matriz *ref_temp_i* y la matriz *test_temp* se multiplican por una ventana de Hamming y a continuación se transforman al dominio de las frecuencias mediante DFT o FFT de $2 \cdot B$ de longitud. Estos pasos pueden realizarse por sustitución:

$$ref_temp_i(j) = ref_temp_i(j) \cdot \{.54 - .46 \cdot \cos(2\pi(j - 1)/(2 \cdot B - 1))\}, 1 \leq i \leq 6 \cdot B + 1, 1 \leq j \leq 2 \cdot B$$

$$test_temp(j) = test_temp(j) \cdot \{.54 - .46 \cdot \cos(2\pi(j - 1)/(2 \cdot B - 1))\}, 1 \leq j \leq 2 \cdot B$$

$$ref_temp_i = \text{FFT}(ref_temp_i), 1 \leq i \leq 6 \cdot B + 1$$

$$test_temp = \text{FFT}(test_temp)$$

En el dominio de frecuencias, sólo las primeras $B + 1$ muestras complejas de cada matriz son únicas, por lo que sólo se guardan dichas muestras. Se extrae la magnitud de cada muestra retenida, obteniéndose la raíz cuadrada de la densidad espectral de potencia de cada trama. En aras de la simplicidad, estos resultados se indican como PSD.

$$ref_temp_i(j) = |ref_temp_i(j)|, 1 \leq i \leq 6 \cdot B + 1, 1 \leq j \leq B + 1$$

$$test_temp(j) = |test_temp(j)|, 1 \leq j \leq B + 1$$

Se eliminan seguidamente los valores medios de cada PSD:

$$ref_temp_i(j) = ref_temp_i(j) - \frac{1}{B+1} \cdot \sum_{j=1}^{B+1} ref_temp_i(j), 1 \leq i \leq 6 \cdot B + 1, 1 \leq j \leq B + 1$$

$$test_temp(j) = test_temp(j) - \frac{1}{B+1} \cdot \sum_{j=1}^{B+1} test_temp(j), 1 \leq j \leq B + 1$$

7.2.4.3 Correlación recíproca de densidades espectrales de potencia

Se calcula un valor de correlación recíproca entre la PSD almacenada en la matriz *test_temp* y cada una de las $6 \cdot B + 1$ PSD almacenadas en las matrices *ref_temp_i*.

$$cross_corr(i) = \frac{\left(\sum_{j=1}^{B+1} ref_temp_i(j) \cdot test_temp(j) \right)}{\left(\sum_{j=1}^{B+1} ref_temp_i(j)^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{B+1} test_temp(j)^2 \right)^{\frac{1}{2}}}, 1 \leq i \leq 6 \cdot B + 1$$

La matriz *cross_corr* contiene ahora los valores de las correlaciones recíprocas entre las PSD de referencia y de prueba en cada deriva del dominio de tiempo. Obsérvese que el segundo factor del denominador de la ecuación de *cross_corr* es una constante normalizadora, necesaria para obtener valores verdaderos de correlación recíproca entre -1 y $+1$. No tiene repercusiones en la localización de crestas que sigue, pero influye en el tratamiento ulterior de las mediciones del retardo con precisión. El mayor valor contenido en *cross_corr* se adopta como una indicación precisa del retardo:

$$fine_delay_k = (3 \cdot B + 1) - j \text{ muestras}$$

$$corr_k = cross_corr(j), 1 \leq k \leq n1$$

donde

$$cross_corr(j) > cross_corr(i), 1 \leq i \leq 6 \cdot B + 1, i \neq j$$

Si *cross_corr* no contiene un valor máximo único, será necesario repetir el procedimiento de la fase de precisión en una nueva posición. Toda la fase de precisión, empezando con la selección de una posición, se repite *n1* veces, produciendo *n1* mediciones de retardo de precisión que se almacenan en *fine_delay_1*, *fine_delay_2*, ... *fine_delay_n1*, y *n1* valores de correlación correspondientes, que se almacenan en *corr_1*, *corr_2*, ... *corr_n1*, respectivamente. Obsérvese que cada estimación precisa del retardo se encontrará entre $-3 \cdot B$ y $3 \cdot B$ inclusive.

7.2.4.4 Tratamiento de la medición de precisión del retardo

Efectuadas las mediciones de precisión de retardo *n1* y calculados los valores correspondientes de correlación recíproca, se elaboran los resultados para determinar su modo de empleo.

En primer lugar, cada uno de los valores de correlación *n1* se contrasta con un valor umbral:

$$\sqrt{\frac{1}{2}} \leq corr_k \Rightarrow fine_delay_k \text{ se retiene}, 1 \leq k \leq n1$$

Con este proceso, sólo se retienen las mediciones de precisión del retardo que justifiquen por lo menos la mitad de la varianza de PSD. El número de mediciones de precisión de retardo que pasa esta prueba es de n_2 , procediéndose a reenumerar las mediciones como $fine_delay_1, fine_delay_2, \dots, fine_delay_n_2$. Cuando $n_2 < n_1/2$, la fase de precisión no producirá un resultado útil. En tales circunstancias, el valor de $fine_delay$ se establece como "invalid", dándose por terminada la fase de precisión.

Cuando $n_2 \geq n_1/2$, la fase de precisión continúa probando la coherencia de las n_2 mediciones de precisión de retardo restantes con las mediciones de retardo aproximadas. Puesto que el margen de incertidumbre en la medición de retardo aproximada es de $\pm B$ muestras y se ha eliminado el retardo aproximado, sólo se retienen las mediciones de precisión de retardo entre $-B$ y B muestras:

$$|fine_delay_k| \leq B \Rightarrow fine_delay_k \text{ se retiene, } 1 \leq k \leq n_2$$

El número de mediciones de precisión de retardo que pasa esta prueba es de n_3 , procediéndose a reenumerar las mediciones como $fine_delay_1, fine_delay_2, \dots, fine_delay_n_3$. Cuando $n_3 < n_1/2$, la fase de precisión no producirá un resultado útil. En tales circunstancias, el valor de $fine_delay$ se establece como "invalid", dándose por terminada la fase de precisión.

Cuando $n_3 \geq n_1/2$, la fase de precisión continúa probando la coherencia entre las restantes n_3 mediciones de precisión de retardo. Esta prueba requiere explorar todos los posibles subconjuntos de tamaño $n_3, n_3 - 1$, hasta $n_1/2$. Sólo existe un posible subconjunto de tamaño $n_3, n_3 - 1$ posibles subconjuntos de tamaño $n_3 - 1, n_3 \cdot (n_3 - 1)/2$ posibles subconjuntos de tamaño $n_3 - 2$, y así sucesivamente. Para cada subconjunto, se prueba la dispersión de las mediciones de precisión de retardo:

$$\max_i \{fine_delay_i\} - \min_i \{fine_delay_i\} \leq \frac{B}{2}, fine_delay_i \in \text{subconjunto actual}$$

Al mayor subconjunto que pase esta prueba se le denomina subconjunto final. La fase de precisión no podrá producir un resultado útil, cuando:

- ningún subconjunto pasa la prueba, o
- no hay un único subconjunto de tamaño máximo que pase la prueba.

En cualquiera de estas situaciones, el valor de $fine_delay$ se establece como "invalid", dándose por terminada la frase de precisión.

El número de mediciones de precisión del retardo en el subconjunto final es n_4 . Estas n_4 mediciones de precisión del retardo se reenumeran ahora como $fine_delay_1, fine_delay_2, \dots, fine_delay_n_4$. El valor medio de estas n_4 mediciones de precisión del retardo se adopta como la medición de precisión del retardo final:

$$fine_delay = \frac{1}{n_4} \cdot \sum_{i=1}^{n_4} fine_delay_i$$

La dispersión de las n_4 mediciones en el subconjunto final se retiene como una medida del margen de incertidumbre en la medición de precisión del retardo final:

$$spread = \max_i \{fine_delay_i\} - \min_i \{fine_delay_i\}, fine_delay_i \in \text{subconjunto final}$$

7.2.5 Combinación de los resultados de las fases aproximada y de precisión

Si la fase de precisión no puede producir una medición de precisión de retardo útil, el valor de $fine_delay$ se establece como "invalid". En tal caso, sólo la medición aproximada expresará la medida del retardo. Si la fase de precisión puede producir una medida de retardo de precisión útil, se

procede a aumentar la medición aproximada teniendo en cuenta la medición de precisión, y el margen de incertidumbre se reduce a partir del de la medición aproximada únicamente:

$$fine_delay = "invalid" \Rightarrow delay = course_delay \pm B \text{ muestras}$$

$$fine_delay \neq "invalid" \Rightarrow delay = coarse_delay + fine_delay \pm spread \text{ muestras}$$

Estos valores de *delay* sólo son correctos cuando las adquisiciones de muestras de audio de la entrada y de la salida del canal son simultáneas. Tras convertir el retardo en segundos, pueden emplearse indicaciones de tiempo para corregir la medición del retardo en caso de adquisiciones no simultáneas:

$$delay = delay/sample_rate \text{ segundos}$$

$$delay = delay + T'(m) - T(n) \text{ segundos}$$

8 Consideraciones sobre las mediciones combinadas de audio/vídeo

En base a los conceptos y métodos definidos en las cláusulas 6 y 7, se pueden examinar algunas cuestiones relativas a mediciones para canales múltiples.

8.1 Mediciones de la actividad de canal audio/vídeo y de la sincronización

Ambos métodos de medición, para vídeo y audio, requieren niveles de señal mínimos en los canales sometidos a prueba, para producir resultados válidos. Cada método tiene sus requisitos específicos. Los métodos para vídeo requieren diferencias distinguibles entre las tramas de vídeo actual y precedente, mientras que los métodos para audio requieren el nivel de RMS de un grupo de tramas de audio, para contrarrestar favorablemente con los niveles de interfaz nominales.

Para poder comparar las mediciones de retardo de audio y vídeo, deben darse simultáneamente en cada entrada las condiciones de actividad necesarias y deberán efectuarse mediciones válidas en ambos canales (lo que depende de la actividad de salida). De no ser así, los cálculos de sincronización no son posibles y será necesario buscar otra oportunidad.

8.2 Asociación de las distintas mediciones entre sí

Los cálculos de sincronización que se describen en la cláusula 5 exigen una asociación entre las tramas de entrada y un proceso de concordancia de tramas, a fin de generar pares de indicaciones de tiempo para cada canal sometido a prueba. El cálculo del desplazamiento de tiempo entre canales y el error de tiempo que introduce el canal de transmisión no tienen en cuenta el retardo como un paso intermedio. En sentido estricto, este hecho impide comparar retardos medidos en dos momentos diferentes para evaluar la sincronización de dos canales.

Sin embargo, los canales audio son una excepción razonable, ya que transportan medios isócronos. Cuando la variación del retardo de audio medida se reduce al error experimental esperado, el retardo de audio medio puede considerarse un valor representativo constante. Este retardo medio puede compararse con una distribución de retardos de canal vídeo o de datos, para obtener una distribución del desfase temporal entre los canales, lo que permite comparar las mediciones de audio y vídeo efectuadas separadamente pero bajo las mismas condiciones de fuente u otras. De este modo, los dispositivos de prueba incapaces de efectuar mediciones simultáneas de canales múltiples pueden no obstante suministrar informaciones útiles cuando las condiciones lo permitan.

Si la variación del retardo de audio medida sobrepasa el error experimental esperado, no es posible utilizar mediciones por canales separados.

9 Mediciones de datos

En esta cláusula se especifican los métodos de medición del retardo en canales de datos integrados en sistemas de comunicación multimedios. Muchas normas de calidad de funcionamiento de datos definen el retardo o tiempo de transmisión para protocolos de comunicación específicos (por ejemplo, la Recomendación X.25 y las diversas Recomendaciones que definen el ATM).

9.1 Recopilación de tramas de datos para medición

Los canales de datos empleados en los sistemas de comunicación multimedios pueden ser muy variados en cuanto a sus objetivos y atributos específicos. Más que intentar abarcar todas las posibilidades, se presentan aquí los métodos generales aplicables a los canales de datos en dos categorías principales (definidas a continuación).

9.1.1 Consideraciones sobre la definición de tramas de datos

La presente Recomendación trata del canal de datos a nivel lógico, por encima de la capa física y de sus interfaces eléctricas. Sin embargo, los sistemas que admiten las mediciones descritas en esta Recomendación contarán con facilidades de prueba que satisfagan una o más normas de interfaces eléctricas. Las observaciones efectuadas en dichas interfaces eléctricas constituirán la base de las mediciones.

A los efectos de esta Recomendación, se define una secuencia de Z bits sucesivos como una trama de datos. La trama de datos n está representada por $D(n)$, y el primer bit de $D(n)$ es $D(1,n)$. La longitud de la trama de datos puede ser determinada por la aplicación del canal de datos.

Se consideran dos posibles configuraciones de los datos de usuario:

- 1) Los usuarios presentan bits de información incorporados dentro de una *estructura* normalizada. Tales estructuras pueden denominarse paquetes, células o tramas.
- 2) Los usuarios presentan trenes de bits *no estructurados*. El sistema de comunicación multimedios puede efectuar su propia segmentación de esos trenes de bits.

Cuando el usuario presenta un tren de bits estructurado y esa estructura permite reconocer las distintas tramas en cada interfaz de canal, la estructura original es considerada la trama de datos para las mediciones de los sistemas multimedios. Una posible excepción es cuando la estructura original contiene un gran número de bits y el tiempo de inserción de la estructura es relativamente grande en comparación con las tramas de audio y las tramas de vídeo. En tal situación quizás sea más eficaz tratar el tren de bits como si no estuviera estructurado. El caso ideal es un tiempo de inserción igual para las tramas en todos los medios, lo que permite una correspondencia uno a uno.

Cuando el canal de datos permite bits de entrada sin estructurar y el sistema de medición tiene la posibilidad de insertar los bits, puede utilizarse un generador de secuencias pseudoaleatorias. Esto ofrece varias ventajas:

- Las secuencias pueden generarse fácilmente, tanto en emplazamientos locales como distantes.
- Puede elegirse la longitud de repetición de las secuencias, para evitar concordancias ambiguas.
- La longitud de la trama de datos puede ser tan pequeña como un solo múltiplo de la longitud del registro de deriva de retroalimentación lineal, pudiéndose seleccionar de manera que concuerde lo más posible con la longitud de otras tramas de medios.

Cuando el canal de datos requiere una gran estructura, el generador de secuencias pseudoaleatorias puede suministrar los bits de información transportados por la estructura.

También es posible que el sistema de medición recoja el tren de bits de entrada a partir de la fuente usual del canal de datos. En este caso, la longitud de la trama de datos coincidirá por lo general con la estructura original. Cuando la fuente de datos produce un modelo de reposo, es poco probable que las mediciones tengan éxito.

9.1.2 Asignación de indicaciones de tiempo

El tiempo $T_D(n)$ asociado a una trama de datos n ha de leerse inmediatamente después de comunicarse el último bit de la trama a través de la interfaz y antes de que se comunique el bit siguiente a través de la interfaz.

Puesto que el tiempo de inserción de la trama es también una medida útil de la transmisión de datos, pueden asociarse al primer bit de una trama de datos indicaciones de tiempo adicionales, que deberán leerse antes de que el bit siguiente sea comunicado a través de la interfaz. En algunos sistemas, el tiempo de inserción a la entrada puede ser diferente del tiempo de inserción a la salida. Además, el tiempo de inserción puede no ser constante.

9.2 Medición de retardos para datos

A continuación se presentan dos maneras de establecer la concordancia entre tramas de datos de entrada y tramas de datos de salida.

9.2.1 Concordancia de datos estructurados

Se examinan métodos para determinar la correspondencia entre paquetes X.25 que son aplicables a muchas formas de datos estructurados [4]. Por lo general, los bits de encabezamiento comunican la suficiente información, por ejemplo, un número de secuencia, como para que los paquetes puedan ser diferenciados unos de otros. Estos identificadores incorporados constituyen una base apropiada para el establecimiento de la concordancia entre tramas de datos.

Caso de ser insuficiente la información del encabezamiento de un protocolo específico, quizás se pueda tratar de concordar identificación adicional dentro del campo de datos de usuario.

9.2.2 Concordancia de trenes de bits sin estructurar

En [5] se define la correspondencia entre bits por comparación de secuencias de bits. Los bits $D'(m)$ corresponden a una secuencia de entrada de longitud igual si existen enteros n y d tales que:

$$D'(i,m) = D(i + d,n) \text{ para casi todos los enteros } 1 \leq i \leq Z - d$$

y

$$D'(i,m) = D(i - Z + d,n + 1) \text{ para casi todos los enteros } Z - d + 1 \leq i \leq Z$$

donde d es el desplazamiento entero positivo ($d < Z$) que puede existir entre las asignaciones de tramas de entrada $D(n)$ y de salida $D'(m)$.

La figura 13 ilustra la correspondencia a través de las fronteras de las tramas de entrada.

Si las tramas de datos de entrada o de salida son realineadas de tal suerte que su desplazamiento de bits se fija en $d = 0$, la prueba de correspondencia se reduce a:

$$D'(i,m) = D(i,n) \text{ para casi todos los enteros } 1 \leq i \leq Z$$

Si los bits se comunican a través de las interfaces de manera periódica, es posible calcular la indicación de tiempo de cualquier bit en una trama de datos, ofreciéndose la posibilidad de establecer indicaciones de tiempo para el primero y último bit en toda nueva definición de $D(n)$ que cumpla $d = 0$.

Al admitirse la correspondencia para *casi todos* los bits en una trama de datos, se hace posible la búsqueda con éxito de la concordancia en presencia de errores binarios limitados. Cuando no hay errores binarios, deberá existir igualdad para todos los enteros i del intervalo $1 \leq i \leq Z$.

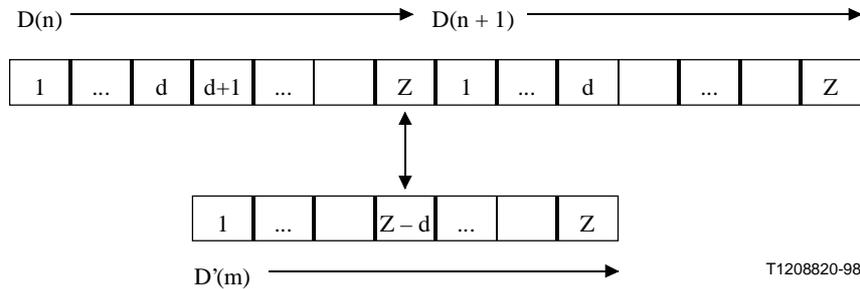


Figura 13/P.931 – Correspondencia entre tramas de datos de entrada y de salida

10 Requisitos de estabilidad de temporizador y sincronización

En esta cláusula se presentan las especificaciones mínimas de los temporizadores internos o relojes que suministran las indicaciones de tiempo a las tramas. Se han de considerar dos configuraciones de reloj, a saber:

- 1) Un único reloj, que suministra indicaciones de tiempo de entrada y de salida (generalmente presentes en aplicaciones de medición en bucle locales y a distancia). En este caso, sólo se aplican las especificaciones de exactitud, estabilidad y resolución, ya que con ellas se caracteriza plenamente la calidad de funcionamiento de un reloj.
- 2) Dos relojes, situados quizás en lugares diferentes (distantes) a la entrada y la salida del sistema de transmisión, que suministran las indicaciones de tiempo. Esta configuración se utiliza en la aplicación de medición de extremo a extremo. Todas las especificaciones de esta cláusula son aplicables a esta configuración.

10.1 Resolución

La resolución mínima de la escala de tiempo disponible a incluir en las indicaciones de tiempo es de $0,1 \mu\text{s}$ (10^{-7} segundos). Este valor representa la resolución pretendida del almacenamiento interno para mediciones. Aunque esta resolución total no será notificada cuando la precisión del reloj interno no la admita, permite a los diseñadores contar con una base precisa para la evaluación de la estabilidad y precisión del reloj.

10.2 Precisión y estabilidad (error de intervalo de tiempo admisible)

La precisión y la estabilidad de la escala de tiempo del reloj interno quedan plenamente establecidas con una especificación relativa al error de intervalo de tiempo máximo. El error de intervalo de tiempo (TIE, *time interval error*) se define como la variación de tiempo que se produce en las lecturas de un reloj dado en comparación con una escala horaria ideal durante un determinado periodo de observación S . El error de intervalo de tiempo máximo (MTIE, *maximum time interval error*) es el mayor TIE para todos los intervalos de medición posibles dentro del periodo de observación.

En la práctica, las mediciones de la transmisión efectuadas de conformidad con esta Recomendación durarán 1 segundo o más. Por ello, la especificación MTIE comenzarán con un intervalo de observación de 0,01 segundo (no se especifican intervalos menores).

En muchas aplicaciones de esta norma, el o los relojes se sincronizarán con una señal de referencia temporal, por ejemplo el sistema mundial de determinación de posición (GPS, *global positioning system*). En tal caso, el MTIE viene dado por la siguiente relación:

$$\text{MTIE, ns} \leq 10^{-2} S + 150$$

y se ilustra en la figura 14.

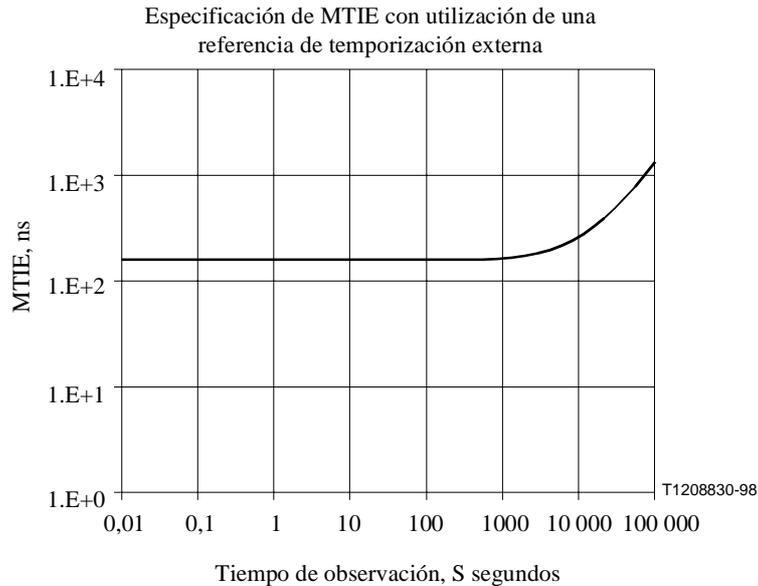


Figura 14/P.931 – MTIE con referencia de temporización

En otras aplicaciones se empleará un reloj interno único, o un reloj distante, que puede sincronizarse antes de las mediciones pero que seguidamente se basa en su precisión interna (conocida también como de funcionamiento en régimen libre) para mantener el tiempo. Esto permite efectuar mediciones durante un periodo de tiempo limitado cuando no se dispone de una fuente de temporización primaria (por ejemplo, el GPS). Para los relojes internos de tipo A, el MTIE viene definido por:

$$\text{MTIE, ns} \leq 10S + 150$$

y se ilustra en la figura 15.

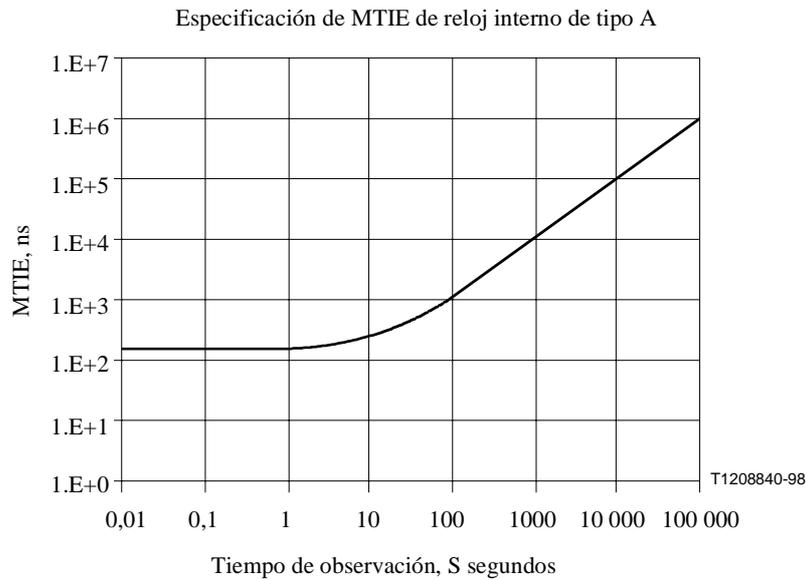


Figura 15/P.931 – Especificación de MTIE de reloj interno de tipo A (durante las mediciones)

Para los relojes internos de tipo B, el MTIE viene definido por:

$$\text{MTIE, ns} \leq 138,9 \times S + 150$$

y se ilustra en la figura 16.

En todos los casos, los informes de las mediciones deberán incluir el error máximo (determinado por la presencia de una fuente de referencia), el tiempo transcurrido desde el momento que se dispone de la referencia y el intervalo de medición efectivo.

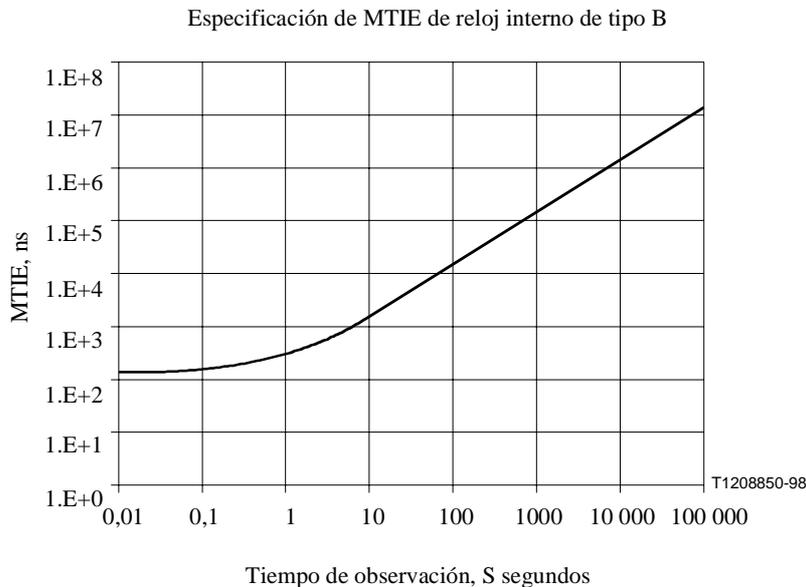


Figura 16/P.931 – Especificación de MTIE de reloj interno de tipo B (durante las mediciones)

10.3 Error de fijación del tiempo

Si se utilizan dos o más relojes en una medición, deberán estar sincronizados. Cuando los relojes se sincronizan directamente o con un tercer reloj de referencia, el error de fijación máximo deberá ser de $\pm 0,075 \mu\text{s}$ ($7,5 \times 10^{-8}$ segundos).

APÉNDICE I

Bibliografía

- [1] ANSI T1.801.01-1995, *American National Standard for Telecommunications – Digital Transport of Video Teleconferencing/Video Telephony Signals – Video Test Scenes for Subjective and Objective Performance Assessment.*
- [2] ANSI T1.801.02-1996, *American National Standard for Telecommunications – Digital Transport of Video Teleconferencing/Video Telephony Signals – Performance Terms, Definitions, and Examples.*
- [3] ANSI T1.801.03-1996, *American National Standard for Telecommunications – Digital Transport of One-Way Video Signals – Parameters for Objective Performance Assessment.*
- [4] ANSI T1.504a-1991, *American National Standard for Telecommunications – Packet Switched Data Communication Service – Performance Measurement Methods.*
- [5] ANSI T1.517-1995, *American National Standard for Telecommunications – Performance Parameters and Objectives for Integrated Services Digital Networks.*
- [6] ANSI T1.314-1991, *Video Codec for Audiovisual Services at $p \times 64$ kbits.*
- [7] SMPTE RP 27.3-1989², *Recommended Practice, Specifications for Safe Title Areas, Test Pattern for Television Systems.*
- [8] SMPTE 125M-1992, *SMPTE Standard for Television – Component Video Signal 4:2:2 – Bit-Parallel Digital Interface².*
- [9] SMPTE 259M-1993, *SMPTE Standard for Television – 10-Bit 4:2:2 Component and 4f_{sc} NTSC Composite Signals – Bit-Parallel Digital Interface².*
- [10] SMPTE 170M-1994, *SMPTE Standard for Television – Composite Analog Video Signal – NTSC for Studio Applications².*
- [11] Contribución UIT-T COM 12-75-E, *Visual Channel Delay and Frame Rate Measurement – Initial Results with a Prototype System, Unión internacional de Telecomunicaciones – Sector de Normalización de las Telecomunicaciones, Periodo de Estudios 1993-1996, Contribución a la Comisión de Estudio 12, origen AT&T, marzo de 1996.*
- [12] Netravali (A.N.) and Haskell (B.G.), *Digital Pictures: Representation and Compression, Plenum Publishing Corporation, Nueva York, NY, 1988.*

² Puede obtenerse en la Sociedad de Ingenieros de Imágenes Móviles y Televisión (SMPTE, *society of motion picture and television engineers*), 595 West Hartsdale Ave. White Plains, NY, 10607.

APÉNDICE II

Símbolos matemáticos y códigos convencionales

| | |
|--|--|
| V | Secuencia de tramas de vídeo adyacentes a la entrada del canal |
| V' | Secuencia de tramas de vídeo adyacentes a la salida del canal |
| V'(m) | Trama de vídeo de salida en el tiempo T'(m) |
| V'(i,j,m) | Píxel de luminancia (i,j) en la trama de vídeo de salida m en el tiempo T'(m) |
| V*(i,j,m) | Píxel de luminancia de salida antes de aplicarse los factores de corrección |
| g | Factor de corrección de ganancia |
| l | Factor de corrección de desplazamiento de nivel |
| h | Factor de corrección de deriva horizontal |
| v | Factor de corrección de deriva vertical |
| z | Factor de corrección de tamaño de trama |
| A | Secuencia de unidades de audio adyacentes a la entrada del canal |
| A' | Secuencia de unidades de audio adyacentes a la salida del canal |
| D | Secuencia de unidades de datos adyacentes a la entrada del canal |
| D' | Secuencia de unidades de datos adyacentes a la salida del canal |
| T | Tren de "medios de temporizador" a la entrada |
| T' | Tren de "medios de temporizador" a la salida |
| T _P | Conjunto de indicaciones de tiempo asociado a unidades de presentación general a la entrada |
| T' _P | Conjunto de indicaciones de tiempo asociado a unidades de presentación de salida |
| T' _P (m) | Valor de temporizador (indicación de tiempo) asociado a la unidad de presentación m de un tren de medios general (a la salida) |
| T' _P (m - 1) | Valor de temporizador de la unidad de presentación que precede a la unidad m |
| T _A | Conjunto de indicaciones de tiempo asociado a tramas de audio de entrada |
| T' _A | Conjunto de indicaciones de tiempo asociado a tramas de audio de salida |
| T' _A (m) | Valor de temporizador (indicación de tiempo) asociado a la trama de audio de salida m |
| t _P (m) | Retardo de canal para la unidad de presentación m |
| t _P | Conjunto de retardos de canal medidos de un tren de medios |
| b _P (m) | Intervalos de tiempo de llegada respecto de la unidad de presentación m |
| b _P | Conjunto de intervalos de tiempo de llegada medidos de un tren de medios |
| f _P (m) | Velocidad de trama elemental para la unidad de presentación m |
| M[V'(m), V'(m - 1)] | Error cuadrático medio (MSE) entre dos tramas adyacentes |
| M[V'(m), V(n)] | MSE entre una trama de salida y una trama de entrada |
| K _s = (I _{máx} - I _{mín} + 1) × (J _{máx} - J _{mín} + 1) | Total de píxeles en la subregión espacial para MSE |
| PSNR | Relación señal/ruido de cresta calculada para un nivel de cresta de vídeo, V _{peak} |

| | |
|----------------|---|
| v' | Conjunto de valores MSE correspondientes a tramas de salida adyacentes dentro de una secuencia de vídeo |
| c | Conjunto de valores MSE con los que se compara una trama de salida con la secuencia de entrada |
| c_v | Conjunto de valores MSE con los que se compara una trama de vídeo de salida con la secuencia de entrada |
| a' | Conjunto de valores de comparación para tramas de salida adyacentes dentro de un tren de audio |
| d' | Conjunto de valores de comparación para tramas de salida adyacentes dentro de un tren de datos |
| N' | Ruido de captación de salida calibrado |
| C_P | Valores de comparación para la unidad de presentación que mejor concuerda con una unidad activa específica |
| $O'_{AV(m,n)}$ | Desplazamiento de tiempo entre tramas asociadas de audio y vídeo a la entrada |
| $O'_{AV(p,q)}$ | Desplazamiento de tiempo entre tramas asociadas de audio y vídeo a la salida |
| $S'_{AV(m,n)}$ | Desfase de tiempo entre tramas asociadas de audio y vídeo a la salida, como consecuencia del sistema/canal de transmisión |

Variables utilizadas en la cláusula 7

| | |
|---------------------|---|
| B | Factor de reducción de anchura de banda y factor de submuestreo |
| $coarse_delay$ | Retardo medido en la fase aproximada |
| $cross_corr$ | Matriz temporal que finalmente contiene los valores de correlación recíproca |
| $cross_corr(i)$ | i -ésimo elemento de la matriz $cross_corr$ |
| $cross_corr_s$ | Versión uniformizada de $cross_corr$ en la fase aproximada |
| $cross_corr_s(i)$ | i -ésimo elemento de la matriz $cross_corr_s$ |
| $delay$ | Resultado final de la medición del retardo en dos etapas |
| $fine_delay$ | Retardo medido en la fase de precisión |
| $fine_delay_k$ | k -ésima medición de precisión del retardo |
| $L1$ | Número de muestras de audio sometidas a medición |
| $L2$ | Número de muestras de audio tras el submuestreo |
| $location$ | Posiciones en las que se efectúa una medición conforme a la fase de precisión |
| $n1$ | Número de mediciones efectuadas en la fase de precisión |
| $n2$ | Número de mediciones de la fase de precisión retenidas después de la primera prueba |
| $n3$ | Número de mediciones de la fase de precisión retenidas después de la segunda prueba |
| $n4$ | Número de mediciones de la fase de precisión retenidas después de la tercera prueba |
| ref | Matriz de muestras de audio provenientes de la entrada del canal |
| $ref(i)$ | i -ésimo elemento de la matriz de referencia |

| | |
|---------------|---|
| ref_temp | Matriz de almacenamiento temporal de muestras de audio de la entrada del canal durante su tratamiento |
| ref_temp(i) | i-ésimo elemento de la matriz ref_temp |
| ref_temp_i | Matriz de almacenamiento temporal de muestras de audio de la entrada del canal durante su tratamiento |
| ref_temp_i(j) | j-ésimo elemento de la matriz ref_temp_i |
| sample_rate | Velocidad de digitalización de la entrada y la salida del canal |
| spread | Dispersión en el subconjunto final de mediciones del retardo con precisión |
| test: | Matriz de muestras de audio provenientes de la salida del canal |
| test(i) | i-ésimo elemento de la matriz de prueba |
| test_temp | Matriz de almacenamiento temporal de muestras de audio de la salida del canal durante su tratamiento |
| test_temp(i) | i-ésimo elemento de la matriz test_temp |

Variables utilizadas en la cláusula 9

| | |
|--------|---|
| Z | Longitud de una trama de datos en bits |
| D(n) | Trama de datos de entrada n |
| D(i,n) | Bit i en la trama de datos de entrada n |
| d | Desplazamiento en bits entre tramas de datos de entrada y salida al determinar la correspondencia |

Variables utilizadas en la cláusula 10

| | |
|----------|---|
| MTIE, ns | Error de intervalo de tiempo máximo (MTIE), expresado en nanosegundos |
| S | Intervalo de observación para mediciones de MTIE |

SERIES DE RECOMENDACIONES DEL UIT-T

| | |
|----------------|---|
| Serie A | Organización del trabajo del UIT-T |
| Serie B | Medios de expresión: definiciones, símbolos, clasificación |
| Serie C | Estadísticas generales de telecomunicaciones |
| Serie D | Principios generales de tarificación |
| Serie E | Explotación general de la red, servicio telefónico, explotación del servicio y factores humanos |
| Serie F | Servicios de telecomunicación no telefónicos |
| Serie G | Sistemas y medios de transmisión, sistemas y redes digitales |
| Serie H | Sistemas audiovisuales y multimedios |
| Serie I | Red digital de servicios integrados |
| Serie J | Transmisiones de señales radiofónicas, de televisión y de otras señales multimedios |
| Serie K | Protección contra las interferencias |
| Serie L | Construcción, instalación y protección de los cables y otros elementos de planta exterior |
| Serie M | RGT y mantenimiento de redes: sistemas de transmisión, circuitos telefónicos, telegrafía, facsímil y circuitos arrendados internacionales |
| Serie N | Mantenimiento: circuitos internacionales para transmisiones radiofónicas y de televisión |
| Serie O | Especificaciones de los aparatos de medida |
| Serie P | Calidad de transmisión telefónica, instalaciones telefónicas y redes locales |
| Serie Q | Conmutación y señalización |
| Serie R | Transmisión telegráfica |
| Serie S | Equipos terminales para servicios de telegrafía |
| Serie T | Terminales para servicios de telemática |
| Serie U | Conmutación telegráfica |
| Serie V | Comunicación de datos por la red telefónica |
| Serie X | Redes de datos y comunicación entre sistemas abiertos |
| Serie Y | Infraestructura mundial de la información |
| Serie Z | Lenguajes de programación |