

Unión Internacional de Telecomunicaciones

UIT-T

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

P.880

(05/2004)

SERIE P: CALIDAD DE TRANSMISIÓN TELEFÓNICA,
INSTALACIONES TELEFÓNICAS Y REDES LOCALES
Métodos de evaluación objetiva y subjetiva de la calidad

**Evaluación continua de la calidad vocal que
varía con el tiempo**

Recomendación UIT-T P.880

UIT-T



RECOMENDACIONES UIT-T DE LA SERIE P

CALIDAD DE TRANSMISIÓN TELEFÓNICA, INSTALACIONES TELEFÓNICAS Y REDES LOCALES

Vocabulario y efectos de los parámetros de transmisión sobre la opinión de los clientes	Series	P.10
Líneas y aparatos de abonado	Series	P.30
		P.300
Patrones de transmisión	Series	P.40
Aparatos para mediciones objetivas	Series	P.50
		P.500
Medidas electroacústicas objetivas	Series	P.60
Medidas relativas a la sonoridad vocal	Series	P.70
Métodos de evaluación objetiva y subjetiva de la calidad	Series	P.80
		P.800
Calidad audiovisual en servicios multimedios	Series	P.900
Aspectos de calidad de transmisión y de calidad de servicio en los puntos extremos de redes de protocolo Internet	Series	P.1000

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

Recomendación UIT-T P.880

Evaluación continua de la calidad vocal que varía con el tiempo

Resumen

En esta Recomendación se describe la metodología denominada evaluación continua de la calidad vocal que varía con el tiempo (CETVSQ) que sirve para evaluar los efectos de las fluctuaciones temporales de la calidad vocal en la calidad instantánea percibida. El método consta de dos partes: en primer lugar, la valoración instantánea que se regula mediante un cursor mientras que se escucha la secuencia vocal en una escala continua, y en segundo lugar, la valoración global en una escala de cinco categorías convencional al final de la secuencia vocal.

Orígenes

La Recomendación UIT-T P.880 fue aprobada el 14 de mayo de 2004 por la Comisión de Estudio 12 (2001-2004) del UIT-T por el procedimiento de la Recomendación UIT-T A.8.

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 Asamblea Mundial de Normalización de las Telecomunicaciones (AMNT), 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 1 de la AMNT.

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 "Administración" se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

La observancia de esta Recomendación es voluntaria. Ahora bien, la Recomendación puede contener ciertas disposiciones obligatorias (para asegurar, por ejemplo, la aplicabilidad o la interoperabilidad), por lo que la observancia se consigue con el cumplimiento exacto y puntual de todas las disposiciones obligatorias. La obligatoriedad de un elemento preceptivo o requisito se expresa mediante las frases "tener que, haber de, hay que + infinitivo" o el verbo principal en tiempo futuro simple de mandato, en modo afirmativo o negativo. El hecho de que se utilice esta formulación no entraña que la observancia se imponga a ninguna de las partes.

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 no 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 2005

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

ÍNDICE

	Página
1 Alcance	1
2 Referencias	1
3 Abreviaturas, siglas o acrónimos	1
4 Descripción del método	1
4.1 Origen y motivación	1
4.2 Preparación de la prueba	2
4.3 Sesión de escucha	2
4.4 Análisis estadístico	4
BIBLIOGRAFÍA	6

Recomendación UIT-T P.880

Evaluación continua de la calidad vocal que varía con el tiempo

1 Alcance

En la presente Recomendación se define un método de evaluación subjetiva de la calidad vocal transmitida de las secuencias de voz largas en las que la calidad varía con el tiempo. El método consiste en la valoración continua mientras se escucha la secuencia y en la valoración global al final de la secuencia. Por consiguiente, además de medir la calidad global percibida (como en la mayor parte de los métodos recomendados), se mide la calidad instantánea percibida, es decir, la calidad percibida en cada instante mientras se escucha la secuencia vocal. En esta versión, el método no es aplicable para seleccionar los códecs vocales. Sin embargo, puede servir de herramienta de diagnóstico de los efectos de la degradación en la calidad instantánea y global percibida, en particular en el caso de las degradaciones discontinuas que varían con el tiempo (por ejemplo, las debidas a la pérdida de paquetes IP, el traspaso a redes móviles, etc.). También puede servir para desarrollar y validar instrumentos de medición objetiva destinados a predecir la calidad vocal mediante la detección y análisis de diversos tipos de degradaciones en la señal vocal; de hecho, mediante el método de evaluación continua de la calidad vocal que varía con el tiempo (CETVSQ) se obtienen evaluaciones instantáneas y globales de la calidad subjetiva.

2 Referencias

- [1] Recomendación UIT-T P.800 (1996), *Métodos de determinación subjetiva de la calidad de transmisión*.
- [2] Recomendación UIT-R BT.500-11 (2002), *Metodología para la evaluación subjetiva de la calidad de las imágenes de televisión*.

3 Abreviaturas, siglas o acrónimos

En esta Recomendación se utilizan las siguientes abreviaturas, siglas o acrónimos.

ACR	Evaluación por categorías absolutas (<i>absolute category rating</i>)
ANOVA	Análisis de la varianza (<i>analysis of variance</i>)
CETVSQ	Evaluación continua de la calidad vocal que varía con el tiempo (<i>continuous evaluation of time varying speech quality</i>)
MOS	Nota media de opinión (<i>mean opinion score</i>)
QoS	Calidad del servicio (<i>quality of service</i>)
SSCQE	Evaluación continua de la calidad con un solo estímulo (<i>single stimulus continuous quality evaluation</i>)

4 Descripción del método

4.1 Origen y motivación

El método de evaluación continua de la calidad vocal que varía con el tiempo (CETVSQ) fue creado porque la calidad del servicio (QoS) de las nuevas redes varía, incluso durante una misma conversación, debido a que se producen ciertas degradaciones (como la pérdida de paquetes IP, el traspaso a redes móviles, etc.). Estas degradaciones se caracterizan por ser distorsiones de la calidad transitorias (a diferencia de las degradaciones "continuas" como la relación señal/ruido) de densidad relativamente alta. Las muestras de corta duración no cuentan en esta densidad. Además, dadas las

características técnicas de las redes móviles o IP, la calidad vocal puede variar sobremanera durante la misma comunicación. Los métodos de evaluación de la calidad vocal descritos en la Rec. UIT-T P.800 [1] utilizan pruebas de escucha subjetiva con estímulos de corta duración (8 s). Si bien estos métodos resultan convenientes para evaluar la calidad vocal constante con el tiempo, a no ser que se empleen muchas muestras no pueden tener en cuenta la producción y distribución realista de estas degradaciones. Tampoco pueden tener en cuenta las fluctuaciones de calidad temporales de larga duración para las cuales se producen procesos mnésicos (efectos de la memoria) que afectan a la calidad global percibida. La falta de una metodología para evaluar la calidad vocal que varía con el tiempo ha motivado la elaboración de la presente Recomendación. Así pues, la evaluación de cómo afectan a la calidad percibida estos tipos de degradaciones y de sus fluctuaciones temporales durante una determinada comunicación exige secuencias de voz más largas que las que generalmente se emplean en los métodos subjetivos convencionales. Además, el método de valoración continua permite estudiar el efecto de las degradaciones transitorias sobre la calidad percibida en cada instante de la secuencia de escucha, y sobre la calidad global percibida (al final de la secuencia). Este método se inspiró en el método SSCQE (evaluación continua de la calidad con un solo estímulo) empleado en el campo de vídeo (Rec. UIT-R BT.500-11 [2] y se ha comprobado su validez para la calidad vocal en diversos estudios previos ([B-1], [B-2], [B-3], [B-4]).

4.2 Preparación de la prueba

4.2.1 Estímulo

El material vocal que se utilice en las pruebas debe ser sencillo, con sentido y de fácil comprensión. No deben emplearse secuencias de corta duración y, por el momento, pueden utilizarse secuencias de duración entre 45 segundos y 3 minutos. Las grabaciones originales, en particular el entorno y procedimiento de grabación, los sistemas emisor y de grabación, los altavoces y los niveles de voz podrán ser iguales a los descritos en B.1/P.800 [1].

4.2.2 Fuentes

El número y la selección de las condiciones dependerá del objetivo de la prueba. La única restricción es la duración de la misma. En la medida de lo posible, se recomienda incluir en el conjunto de condiciones de pruebas algunas condiciones de control, es decir, condiciones sin variación alguna de los parámetros físicos.

4.3 Sesión de escucha

4.3.1 Oyentes

En la prueba participarán un número conveniente de oyentes legos en la materia (al menos 24). Todos los oyentes tendrán como lengua materna el idioma utilizado en la prueba y ninguno de ellos habrá participado en un experimento de este tipo en los últimos seis meses.

4.3.2 Presentación del audio y entorno de prueba

La presentación del audio será conforme con las directrices que figuran en la Rec. UIT-T P.800 [1]. Estas directrices tratan sobre los sistemas, los niveles y el entorno de escucha.

4.3.3 Dispositivo para registrar la valoración continua y su configuración

Para registrar la evaluación continua de la calidad debe utilizarse un cursor electrónico (por ejemplo, una resistencia variable) conectado a un computador. Este dispositivo debe tener las siguientes características:

- mecanismo del cursor sin posición de "reinicio" (es decir, sin que vuelva automáticamente a una posición predefinida);
- deslizamiento lineal de unos 10 cm;
- posición fija o de sobremesa;

- las muestras de la "posición del cursor" se han de tomar dos veces por segundo (velocidad suficiente para capturar con precisión las respuestas de los participantes);
- la "posición del cursor" podrá codificarse desde 0 (valor inferior de la escala) hasta un mínimo de 100 (valor superior de la escala), para lograr una resolución aceptable. La posición inicial del cursor debe ser la correspondiente a la mitad de la escala.

4.3.4 Procedimiento de evaluación

Para cada secuencia de voz, el participante debe realizar dos tareas: la valoración continua mientras escucha la secuencia y la valoración global al final de la secuencia. Para mayor información, así como para consultar los resultados de los estudios anteriores, véanse [B-1] y [B-2].

a) Valoración continua

En primer lugar, los participantes tienen que valorar continuamente la calidad vocal de la secuencia desplazando el cursor a lo largo de escala de modo que la posición de éste indique la calidad instantánea; podrán deslizar libremente el cursor a lo largo de toda la escala. Para ayudar al participante a ajustar la posición del cursor a la calidad vocal, se indicarán cinco posiciones en la escala, a saber, excelente, buena, regular, mediocre y mala.

Escala de calidad continua

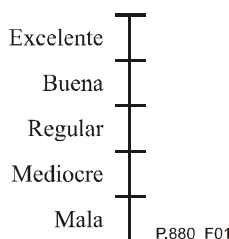


Figura 1/P.880 – Escala continua para la valoración instantánea

b) Evaluación global

En segundo lugar, los participantes tienen que valorar la calidad global de cada secuencia al final de la misma, utilizando para ello la siguiente escala de calidad de la escucha de 5 categorías (la misma MOS empleada en la ACR).

Escala de la calidad global (ACR)

Calidad de la voz	Nota
Excelente	5
Buena	4
Regular	3
Mediocre	2
Mala	1

4.3.5 Procedimiento de prueba

Antes de evaluar las secuencias vocales de prueba, los participantes tienen que entrenarse con unas secuencias escogidas. Estas secuencias deben tener niveles de calidad diferentes y presentar distintas fluctuaciones de calidad representativas de la gama de fluctuaciones temporales y los niveles de calidad que habrán de evaluar durante la prueba real.

Por lo general, la prueba constará de varias sesiones, separadas por pausas. El conjunto completo de estímulos (secuencias) se presenta en un orden aleatorio diferente a cada grupo de participantes.

4.3.6 Instrucciones para los participantes

En el cuadro 1 se muestra un ejemplo de instrucciones típicas (para secuencias de T segundos de duración). Las instrucciones deben darse por escrito (y oralmente, si cabe) antes de comenzar el experimento.

Cuadro 1/P.880 – Ejemplo de instrucciones que podrían darse a los oyentes

En esta prueba escuchará varias secuencias vocales de T segundos de duración por el terminal telefónico. La calidad de la secuencia puede variar de diferentes formas. Para cada secuencia deberá realizar lo siguiente:

- 1) Dar su opinión sobre la calidad vocal mientras escucha la secuencia, es decir, en todo instante, para lo cual utilizará el cursor situado en la mesa frente a usted. La escala de puntuación es continua, de modo que puede desplazar el cursor libremente de acuerdo con la calidad vocal que perciba.

Las posiciones indicadas a lo largo de la escala (excelente, buena, regular, mediocre y mala) le servirán de orientación para ajustar el cursor según su opinión. Por ejemplo, si considera que la calidad corresponde exactamente a la posición entre regular y buena, colocará el cursor en la mitad entre esas dos posiciones indicadas. Ahora bien, no tiene por qué colocar el cursor justo en una de las posiciones indicadas o entre dos de ellas a no ser que considere que dicha posición representa precisamente su opinión.

No olvide desplazar el cursor cuando la calidad varíe.

- 2) Al final de cada secuencia tendrá que puntuar la calidad global de toda la secuencia. Para ello deberá pulsar el botón correspondiente a su opinión de acuerdo con la siguiente escala de puntuación:

Opinión global sobre la calidad vocal de la secuencia que acaba de escuchar:

Excelente

Buena

Regular

Mediocre

Mala

Dispone de cinco segundos para responder pulsando el botón correspondiente a su opinión. Habrá una breve pausa antes de pasar a la siguiente secuencia.

Comenzaremos con una breve sesión de entrenamiento para que se familiarice con el procedimiento de prueba. La duración de las sesiones de prueba reales será de 10 a 15 minutos.

Gracias por su colaboración.

4.4 Análisis estadístico

Si T-s es la duración de cada secuencia (en segundos), para cada participante se tendrá un archivo de datos con $2 \times T$ -s valores (lo que corresponde a una nota instantánea cada 500 ms durante T segundos), más un valor escalar (o sea, la valoración de la calidad global). Los $2 \times T$ -s valores instantáneos (entre $t = 0$ y $t = 2T-1$) se transforman linealmente en valores de 1 a 5 utilizando la ecuación $S(t) = 1 + 4$ (posición del cursor/máximo), siendo S(t) la nota de opinión instantánea.

La valoración instantánea media (y su desviación típica) de cada secuencia se calcula promediando las valoraciones instantáneas de cada participante con respecto al número de participantes para cada instante de tiempo t (es decir, cada 500 ms). La MOS global media (y su desviación típica) de cada secuencia se calcula promediando las valoraciones globales de cada participante con respecto al número de participantes en la escala ACR.

Las diferencias considerables entre la nota continua y la nota global indicarán probablemente que se han producido degradaciones transitorias. Esto dependerá de lo reciente que sean, del número de veces que se hayan producido y de la duración de las degradaciones relativa a la nota global. Habrá

que estudiar más este asunto antes de poder utilizar esta información en la planificación de las transmisiones.

Podrán realizarse análisis estadísticos (por ejemplo, análisis de la varianza) a fin de determinar los efectos importantes que se han producido en las diferentes condiciones experimentales.

Observación: Las respuestas de algunos participantes pueden descartarse si presentan grandes variaciones y figuran fuera de dos desviaciones típicas entre sujetos (calculadas para todos los sujetos) más del 10% del tiempo.

BIBLIOGRAFÍA

- [B-1] Contribución UIT-T [COM 12-94] (1999), *Continuous assessment of time-varying subjective vocal quality and its relationship with overall subjective quality*, (Source FT R&D).
- [B-2] GROS (L.), CHATEAU (N.), Instantaneous and Overall Judgments for Time-Varying Speech Quality: Assessments and Relationships, *Acta Acustica*, 87, 367-377, 2001.
- [B-3] Contribución UIT-T [COM 12-D139] (2000), *Study of the relationship between instantaneous and overall subjective speech quality for time-varying quality speech sequences: influence of recency effect*, (Source FT R&D).
- [B-4] Contribución UIT-T [COM 12-22] (2000), *Assessment of time-varying speech quality: comparison between listening situation and conversational situation*, (Source FT R&D).

SERIES DE RECOMENDACIONES DEL UIT-T

Serie A	Organización del trabajo del UIT-T
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	Redes de cable y transmisión de programas radiofónicos y televisivos, 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	Gestión de las telecomunicaciones, incluida la RGT y el mantenimiento de redes
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, comunicaciones de sistemas abiertos y seguridad
Serie Y	Infraestructura mundial de la información, aspectos del protocolo Internet y Redes de la próxima generación
Serie Z	Lenguajes y aspectos generales de soporte lógico para sistemas de telecomunicación