

G10 INSTRUMENTOS DE MUSICA; ACUSTICA

G10L ANALISIS O SINTESIS DE LA VOZ; RECONOCIMIENTO DE LA VOZ; PROCESAMIENTO DE LA VOZ O EL HABLA; CODIFICACIÓN O DESCODIFICACIÓN DEL AUDIO O LA VOZ [4]

Nota

Esta subclase no cubre :

- los dispositivos para el almacenamiento de señales de voz o de audio, que están cubiertos por las subclases G11B y G11C; [2010.01]
- la codificación de señales de voz comprimidas para la transmisión o el almacenamiento, que está cubierta por el grupo H03M 7/30. [2010.01]

13/00	Síntesis de la voz; Sistemas de síntesis de la voz a partir de texto [7,8]	15/14 utilizando técnicas de programación dinámica, p. ej. normalización temporal por comparación dinámica [DTW] (G10L 15/18 tiene prioridad) [7,8]
13/02 Métodos de producción de voz sintética; Sintetizadores de voz [7,8,2013.01]	15/16 utilizando redes neuronales artificiales [7,8]
13/027 Conceptos para sintetizadores de voz; Generación de frases en lenguaje natural a partir de conceptos basados en lenguaje máquina (generación de parámetros para síntesis de voz a partir de texto G10L 13/08) [2013.01]	15/18 utilizando una modelización del lenguaje natural [7,8,2013.01]
13/033 Edición de voz, p. ej. manipulando la voz del sintetizador [2013.01]	15/183 utilizando dependencias del contexto, p. ej. modelos del lenguaje [2013.01]
13/04 Detalles de sistemas de síntesis de voz, p. ej. estructura del sintetizador o gestión de memoria [7,8,2013.01]	15/187 Contextos fonémicos, p. ej. reglas de pronunciación, impedimentos fonotácticos, n-grama de fonemas [2013.01]
13/047 Arquitectura de los sintetizadores de voz [2013.01]	15/19 Contexto gramatical, p. ej. desambiguación de las hipótesis de reconocimiento basadas en reglas de secuencias de palabras [2013.01]
13/06 Unidades de voz elementales utilizadas en sintetizadores de voz; reglas de concatenación [7,8,2013.01]	15/193 Gramática formal, p. ej. autómatas finitos, gramática libre de contexto o cadenas de palabras [2013.01]
13/07 Reglas de concatenación [2013.01]	15/197 Gramática probabilística, p. ej. n-gramas de palabras [2013.01]
13/08 Análisis de texto o generación de parámetros para la síntesis de la voz a partir de texto, p. ej. conversión grafema-fonema, generación de métrica o determinación de la entonación o de la acentuación [7,8,2013.01]	15/20 Técnicas de reconocimiento de la voz especialmente adaptadas para trabajar en ambientes adversos, p. ej. en presencia de ruido o para voz emitida en situaciones de estrés (G10L 21/02 tiene prioridad) [7,8]
13/10 Reglas de prosodia derivadas del texto; Acento o entonación [2013.01]	15/22 Procedimientos utilizados durante el proceso de reconocimiento de la voz, p. ej. diálogo hombre-máquina [7,8]
15/00	Reconocimiento de la voz (G10L 17/00 tiene prioridad) [7,8,2013.01]	15/24 Reconocimiento de la voz mediante la utilización de medios no acústicos [7,8,2013.01]
15/01 Valoración o evaluación de los sistemas de reconocimiento de voz [2013.01]	15/25 utilizando la posición de los labios, el movimiento de los labios o el análisis del rostro [2013.01]
15/02 Extracción de características para el reconocimiento de la voz; Selección de la unidad de reconocimiento [7,8]	15/26 Sistemas de síntesis de texto a partir de la voz (G10L 15/08 tiene prioridad) [7,8]
15/04 Segmentación o detección de los límites de las palabras [7,8,2013.01]	15/28 Detalles estructurales de sistemas de reconocimiento de la voz [7,8,2013.01]
15/05 Detección de los límites de las palabras [2013.01]	15/30 Reconocimiento distribuido, p. ej. en sistemas cliente-servidor, para teléfonos móviles o aplicaciones en red [2013.01]
15/06 Creación de plantillas de referencia; Entrenamiento de sistemas de reconocimiento de la voz, p. ej. adaptación a las características de la voz de la persona que habla (G10L 15/14 tiene prioridad) [7,8,2013.01]	15/32 Reconocedores múltiples usados en serie o en paralelo; Basados en sistemas de puntuaciones combinadas, p. ej. sistemas de voto. [2013.01]
15/065 Adaptación [2013.01]	15/34 Adaptación de un reconocedor individual a un procesamiento en paralelo, p. ej. usando procesadores múltiples o computación en la nube (cloud computing). [2013.01]
15/07 a la persona que habla [2013.01]		
15/08 Búsqueda o clasificación de la voz [7,8]		
15/10 utilizando medidas de distorsión o distancia entre la voz desconocida y las plantillas de referencia [7,8]		
15/12 utilizando técnicas de programación dinámica, p. ej. normalización temporal por comparación dinámica [DTW] [7,8]		

- 17/00** **Identificación o verificación de la persona que habla [7,8,2013.01]**
- 17/02** . Operaciones anteriores al procesamiento, p. ej. elección del segmento; Representación o modelado de patrones, p. ej. basados en Análisis Discriminante Linear (LDA) o componentes principales; Elección o extracción de características [2013.01]
- 17/04** . Entrenamiento, registro o construcción de modelos [2013.01]
- 17/06** . Técnicas de toma de decisiones; Estrategia de ajuste de patrones [2013.01]
- 17/08** . . Utilización de métricas de distorsión o una distancia concreta entre el patrón sonda y las plantillas de referencia [2013.01]
- 17/10** . . Sistemas multimodales, p. ej. basados en la integración de múltiples motores de reconocimiento o la fusión de sistemas expertos [2013.01]
- 17/12** . . Normalización de la puntuación [2013.01]
- 17/14** . . Utilización de la categorización fonémica o el reconocimiento de la voz antes del reconocimiento de la persona que habla o su verificación [2013.01]
- 17/16** . Modelos ocultos de Márkov [HMMs] [2013.01]
- 17/18** . Redes neuronales artificiales; Enfoque conexionista [2013.01]
- 17/20** . Transformación de patrones operaciones dirigidas al incremento de la robustez del sistema, p. ej. contra el ruido del canal o las diferentes condiciones de trabajo [2013.01]
- 17/22** . Procedimientos interactivos; Interfaces hombre-máquina [2013.01]
- 17/24** . . haciendo que el usuario pronuncie una contraseña o una frase predefinida [2013.01]
- 17/26** . Reconocimiento de características especiales de la voz, p. ej. para utilización en detectores de mentiras; Reconocimiento de voces de animales [2013.01]
- 19/00** **Técnicas de análisis-síntesis de la voz o de señales de audio para la reducción de la redundancia, p. ej. en codificadores vocales; Codificación o decodificación de la voz o de señales de audio, utilizando modelos filtro-fuente o el análisis psicoacústico (en instrumentos musicales G10H) [7,8,2013.01]**
- 19/002** . Distribución dinámica de bits (para codificadores de audio perceptuales G10L 19/032) [2013.01]
- 19/005** . Corrección de errores producidos por el canal de transmisión, si se refieren al algoritmo de codificación [2013.01]
- 19/008** . Codificación de señales de audio multicanal de o de decodificación mediante la correlación entre canales para reducir la redundancia, p. ej. estéreo conjunto, codificación de la intensidad o matizado [2013.01]
- 19/012** . Codificación del ruido de confort o el silencio [2013.01]
- 19/018** . Marcado digital del audio, p. ej. integrando datos inaudibles en la señal de audio [2013.01]
- 19/02** . utilizando análisis espectrales, p. ej. codificadores vocales de transformación o codificadores vocales subbanda [7,8,2013.01]
- 19/022** . . Bloqueo, p. ej. agrupación de muestras en el tiempo; Elección de las ventanas de análisis; Factorización de interferencias [2013.01]
- 19/025** . . . Detección de transitorios o accesos para la conmutación de la resolución de tiempo y frecuencia [2013.01]
- 19/028** . . Sustitución del ruido, p. ej. Sustituyendo componentes espectrales no tonales por fuentes de ruido (ruido de confort para transmisiones de voz discontinua G10L 19/012) [2013.01]
- 19/03** . . Predicción espectral para prevenir el preeco; Conformación temporal del ruido (temporary noise shaping)[TNS], p. ej. en MPEG2 o MPEG4 [2013.01]
- 19/032** . . Cuantificación o descuantificación de componentes espectrales [2013.01]
- 19/035** . . . Cuantificación logarítmica o escalar [2013.01]
- 19/038** . . . Cuantificación vectorial, p. ej. audio TwinVQ [2013.01]
- 19/04** . utilizando técnicas de predicción [7,8,2013.01]
- 19/06** . . Determinación o codificación de las características del espectro, p. ej. de los coeficientes de predicción a corto plazo [7,8,2013.01]
- 19/07** . . . Codificadores vocales de par de espectro de líneas (line spectrum pair) [LSP] [2013.01]
- 19/08** . . Determinación o codificación de la función de excitación; Determinación de los parámetros de predicción a largo plazo [7,8,2013.01]
- 19/083** . . . siendo la función de excitación un incremento de excitación (G10L 25/90 tiene prioridad) [2013.01]
- 19/087** . . . Utilizando modelos de excitación mixtos, p. ej. MELP, MBE, split band LPC o HVXC [2013.01]
- 19/09** . . . Predicción a largo plazo, p. ej. Eliminando redundancias periódicas, p. ej. utilizando un libro de códigos adaptable o predictor del tono [2013.01]
- 19/093** . . . usando modelos de excitación sinusoidales [2013.01]
- 19/097** . . . utilizando la descomposición en forma de ondas prototipo o codificadores interpoladores de formas de onda prototipos [PWI] [2013.01]
- 19/10** . . . siendo la función de excitación una excitación de impulsos múltiples [7,8,2013.01]
- 19/107** Excitación de pulsaciones débiles, p. ej. utilizando libros de códigos algebraico [2013.01]
- 19/113** Excitación de pulsaciones normales [2013.01]
- 19/12** . . . Determinación o codificación de una excitación de código, p. ej. en codificadores vocales de predicción lineal excitados por código [CELP] [7,8,2013.01]
- 19/125** Excitación del tono, p. ej. innovación sincronizada del tono CELP [PSI-CELP] [2013.01]
- 19/13** Predicción lineal de excitación residual [RELP] [2013.01]
- 19/135** Predicción lineal de excitación de suma vectorial [VSELP] [2013.01]
- 19/16** . Estructura de codificadores de voz [2013.01]
- 19/18** . . . Codificadores de voz que utilizan modos múltiples [2013.01]
- 19/20** utilizando codificación específica de clase de sonido, codificadores híbridos o codificación basada en objeto [2013.01]
- 19/22** Decisión del modo, p. ej. basado en el contenido de la señal de audio frente a parámetros externos [2013.01]

19/24	Codecs de frecuencia variable, p. ej. para generar diferentes calidades utilizando una representación escalable como una codificación jerárquica o en capas [2013.01]	21/14	mostrando la información en el dominio de la frecuencia [2013.01]
19/26	prefiltrado o postfiltrado [2013.01]	21/16	Transformando en una representación no visible (dispositivos o métodos para habilitar los oídos de los pacientes para sustituir una percepción auditiva directa por otro tipo de percepción A61F 11/04) [2013.01]
21/00	Tratamiento de la señal de la voz para producir otra señal audible o no audible, p. ej. visual o táctil, con el fin de modificar su calidad o su inteligibilidad (G10L 19/00 tiene prioridad) [7,8,2013.01]		21/18	Detalles del proceso de transformación [2013.01]
21/003	Cambio de la calidad de la voz, p. ej. tono o formantos [2013.01]	25/00	Técnicas de análisis del habla o voz no restringidos a un solo de los grupos G10L 15/00 Hasta G10L 21/00 (silenciar los amplificadores basados en semiconductores, cuando algunas de las características especiales de una señal son detectadas por un detector de voz, p. ej. detectar cuando no hay ninguna señal, H03G 3/34) [2013.01]	
21/007	caracterizada por el proceso utilizado [2013.01]	25/03	caracterizada por el tipo de parámetros extraídos [2013.01]
21/01	Corrección del eje de tiempos [2013.01]	25/06	siendo los parámetros extraídos coeficientes de correlación [2013.01]
21/013	Adaptación al tono objetivo [2013.01]	25/09	siendo los parámetros extraídos los ratios de cruce por cero [2013.01]
21/02	Mejora de la inteligibilidad de la voz, p. ej. reducción de ruido o eliminación de ecos (reducción de efectos de eco en los sistemas de transmisión en línea H04B 3/20; supresión de eco en teléfonos de manos libres H04M 9/08) [7,8,2013.01]	25/12	siendo los parámetros extraídos coeficientes de predicción [2013.01]
21/0208	Filtración del ruido [2013.01]	25/15	siendo los parámetros extraídos información de la estructura de la formación del habla [2013.01]
21/0216	caracterizado por el método utilizado para la estimación del ruido [2013.01]	25/18	siendo los parámetros extraídos información espectral de cada una de las subbandas [2013.01]
21/0224	Procesado en el dominio del tiempo [2013.01]	25/21	siendo los parámetros extraídos información de potencia [2013.01]
21/0232	Procesado en el dominio de la frecuencia [2013.01]	25/24	siendo los parámetros extraídos el cepstrum [2013.01]
21/0264	caracterizado por el tipo de medición de parámetros, p. ej. técnicas de correlación, técnicas zero crossing o técnicas predictivas [2013.01]	25/27	caracterizadas por la técnica de análisis [2013.01]
21/0272	Separación de la señal de voz [2013.01]	25/30	utilizando redes neuronales [2013.01]
21/028	utilizando las propiedades de la fuente de sonido [2013.01]	25/33	utilizando lógica difusa [2013.01]
21/0308	caracterizado por el tipo de medición de parámetros, p. ej. técnicas de correlación, técnicas zero crossing o técnicas predictivas [2013.01]	25/36	utilizando teoría del caos [2013.01]
21/0316	cambiando la amplitud [2013.01]	25/39	utilizando algoritmos genéticos [2013.01]
21/0324	Detalles de su proceso [2013.01]	25/45	caracterizadas por el tipo de ventana de análisis [2013.01]
21/0332	implicando la modificación de las formas de onda [2013.01]	25/48	especialmente adaptadas para uso particular [2013.01]
21/034	Ajuste automático [2013.01]	25/51	para comparación o discriminación [2013.01]
21/0356	para sincronización con otras señales, p. ej. señales de vídeo [2013.01]	25/54	para recuperación [2013.01]
21/0364	para mejorar la inteligibilidad [2013.01]	25/57	para procesar las señales de vídeo [2013.01]
21/038	utilizando técnicas de extensión de banda [2013.01]	25/60	para medir la calidad de las señales de voz [2013.01]
21/0388	Detalles de su proceso [2013.01]	25/63	para estimar un estado emocional [2013.01]
21/04	Compresión o expansión temporales [7,8,2013.01]	25/66	para extraer parámetros relacionados con cuestiones de salud (detección y medidas para diagnóstico A61B 5/00) [2013.01]
21/043	cambiando la velocidad [2013.01]	25/69	para evaluar señales de voz sintéticas o decodificadas [2013.01]
21/045	utilizando dispersión o inserción de una forma de onda [2013.01]	25/72	para transmitir los resultados de los análisis [2013.01]
21/047	caracterizado por el tipo de forma de onda a ser dispersada o insertada [2013.01]	25/75	para modelar parámetros de la cavidad vocal [2013.01]
21/049	caracterizado por la interconexión de las formas de onda [2013.01]	25/78	Detección de presencia o ausencia de señales de voz (con conmutación de dirección de transmisión por frecuencia vocal en sistemas telefónicas bidireccionales de altavoz H04M 9/10) [2013.01]
21/055	por sincronizar con otras señales, p. ej. señales de vídeo [2013.01]	25/81	para discriminar voz de música [2013.01]
21/057	para mejorar la inteligibilidad [2013.01]	25/84	para discriminar voz de ruido [2013.01]
21/06	Transformación de la voz en una representación no audible, p. ej. visualización de la voz o tratamiento de la voz para ayudas táctiles (G10L 15/26 tiene prioridad) [7,8,2013.01]	25/87	Detección de puntos discretos dentro de una señal de voz [2013.01]
21/10	Transformando en información visible [2013.01]			
21/12	mostrando la información en el dominio del tiempo [2013.01]			

G10L

- 25/90
- 25/93
- . Detección del tono de una señal de habla [2013.01]
 - . Discriminación entre las partes con voz y sin voz de una señal de habla (G10L 25/90 tiene prioridad) [2013.01]

99/00

Materia no prevista en otros grupos de esta subclase [2013.01]