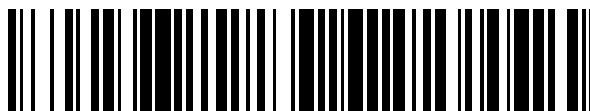


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 761 915**

51 Int. Cl.:

G09B 21/00 (2006.01)

H04N 7/088 (2006.01)

G10L 21/06 (2013.01)

H04N 9/82 (2006.01)

H04N 5/85 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.11.2009 PCT/US2009/065311**

87 Fecha y número de publicación internacional: **17.06.2010 WO10068388**

96 Fecha de presentación y número de la solicitud europea: **20.11.2009 E 09832312 (4)**

97 Fecha y número de publicación de la concesión europea: **13.11.2019 EP 2356654**

54 Título: **Método y procedimiento para descripciones auxiliares de programa basadas en texto para televisión**

30 Prioridad:

08.12.2008 US 330242

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

21.05.2020

73 Titular/es:

**HOME BOX OFFICE INC. (100.0%)
1100 Avenue of the Americas
New York, NY 10036, US**

72 Inventor/es:

CUTTNER, CRAIG DAVIS

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 761 915 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método y procedimiento para descripciones auxiliares de programa basadas en texto para televisión

Campo de la invención

5 La presente invención se refiere a un método y procedimiento para producir y/o distribuir servicios auxiliares basados en texto a telespectadores.

Información de antecedentes

10 Convencionalmente, para ofrecer contenido multimedia, tal como contenido que incluye vídeo y/o audio, a personas con discapacidad auditiva o visual se han utilizado diversas técnicas. A este respecto, el contenido de vídeo puede incluir imágenes, imágenes en movimiento, animación y otros contenidos, y el contenido de audio puede incluir diálogo, música, efectos sonoros y otros contenidos. Puede ser difícil ofrecer dichos contenidos a las personas con discapacidad auditiva o visual, de una manera con la cual el usuario pueda percibir en la práctica la información del contenido ofrecida.

15 Para las personas con discapacidad auditiva, al contenido de vídeo se le complementa convencionalmente con subtítulos cerrados (también denominados ocultos o no incrustados, en inglés, "closed caption"), al objeto de ofrecer a dichos usuarios un contenido multimedia relativamente completo. La subtitulación cerrada proporciona un texto en la pantalla que representa el contenido de audio, con nivel de detalle variable. Por ejemplo, los subtítulos cerrados pueden representar uno o varios de diálogo, música, efectos sonoros y otros componentes del contenido de audio. Según se muestra en la Figura 1, un sistema convencional de subtitulación cerrada puede incluir un taquígrafo 3, que escucha una producción 1 en directo o pregrabada. Mientras escucha, el taquígrafo 3 crea un flujo 5 de subtítulos cerrados basado en el contenido de audio.

25 Los subtítulos cerrados "dinámicos" (que van apareciendo a medida que se crean, en inglés "paint-on") y los subtítulos cerrados "estáticos" (que emergen de golpe como un conjunto, en inglés "pop-on") son dos tipos de subtítulos cerrados. Los subtítulos cerrados dinámicos son el tipo de subtítulos cerrados utilizado más frecuentemente. En la subtitulación cerrada dinámica, un taquígrafo 3 escucha una producción y escribe, en tiempo real, el flujo 5 de subtítulos cerrados, p. ej. durante una transmisión en directo de un programa de televisión o un evento deportivo. El flujo aparece "dinámicamente" en el dispositivo de visualización del usuario a medida que el taquígrafo 3 escribe el flujo 5 de subtítulos cerrados. Debido a esta creación en tiempo real de los subtítulos cerrados dinámicos, los usuarios advierten generalmente un retraso entre la recepción del contenido de vídeo y la recepción del flujo de subtítulos cerrados dinámicos correspondiente al contenido de audio que va asociado con el vídeo recibido.

35 Por el contrario, los subtítulos cerrados estáticos comprenden típicamente el flujo 5 de subtítulos cerrados y una instrucción asociada que da al dispositivo de visualización la orden de "hacer emerger" en el momento adecuado el flujo 5 de subtítulos cerrados. Por este motivo, la subtitulación cerrada estática requiere, antes de ser distribuida a los usuarios, cierta programación necesaria para escribir y temporizar con precisión las órdenes de "emersión". Gracias a esta programación adicional, los subtítulos cerrados estáticos se pueden visualizar de manera sincrónica con el contenido de vídeo correspondiente.

40 En cualquiera de estos tipos de subtitulación cerrada, se puede distribuir 7 el flujo 5 de subtítulos cerrados a los usuarios utilizando sistemas de subtitulación cerrada, tales como, por ejemplo, CEA-608 o CEA-708. Cuando se visiona un programa en particular en una ubicación 8 de visionado, se puede mostrar 9 el programa conjuntamente con los subtítulos cerrados en la pantalla. Así, los usuarios con discapacidad auditiva pueden percibir la información que está contenida en la parte de audio de un contenido multimedia.

45 Para las personas con discapacidad visual, se puede complementar el contenido de audio con una descripción adicional del contenido de vídeo, basada en el audio, al objeto de ofrecer contenido multimedia relativamente completo a dichos usuarios. La descripción adicional basada en audio puede proporcionar, con nivel de detalle variable, una descripción en audio del contenido de vídeo que incluya, por ejemplo, acciones, movimientos, ubicaciones, decorado o paisajes, y expresiones o gestos. Según se muestra en la Figura 2, un servicio convencional de descripción de vídeo basada en audio puede incluir un narrador 4 que graba una pista 6 de audio de contenido de vídeo mientras lee un guion 2 de producción o bien ve una producción 1 en directo o pregrabada. Después se puede distribuir 12 la pista 6 de audio a los usuarios, por ejemplo, mediante difusión tradicional por TV, cable, satélite, telecomunicaciones, o a través de medios grabados tales como DVD, Blu-Ray o videocasete. Cuando se escucha un programa en particular en una ubicación 8 de visionado, se puede reproducir el programa conjuntamente con la pista 6 de audio. Así, los usuarios con discapacidad visual pueden percibir la información que está contenida en la parte de vídeo de un contenido multimedia.

55 Sin embargo, los servicios convencionales de descripción de vídeo basada en audio requieren mucho tiempo y sus procesos de producción son caros. Por ejemplo, para crear estas descripciones de vídeo basadas en audio, uno o varios narradores han de ver las producciones completas y grabar pistas de audio del contenido de vídeo. Además, a menudo se requiere que las entidades productoras utilicen los costosos servicios de narradores profesionales.

Además, las grabaciones hechas por narrador de pistas de audio, basándose en los guiones de producción o en programas pregrabados, agregan otro paso al proceso de producción, sumando tiempo y coste adicionales. Aun así, este procedimiento se utiliza con frecuencia, ya que permite a los narradores crear pistas de audio con mayor detalle y precisión, ya que ofrece la posibilidad de poner en marcha, parar, quitar, añadir, volver a grabar y editar de cualquier otra manera todas las pistas de audio o partes de ellas.

También se emplea a menudo el procedimiento alternativo de que los narradores graben las pistas de audio durante la producción en directo. Tal procedimiento añade más complejidad al proceso de producción en directo. Este procedimiento en directo también produce pistas de audio con menos detalles y precisión que el método alternativo, ya que puede ocurrir que los narradores dispongan de una sola oportunidad para grabar las pistas de audio durante la producción en directo.

Además, la distribución del servicio de descripción de vídeo basada en audio consume ancho de banda, que es escaso, de un canal de audio adicional. La transmisión de contenido de audio requiere un gran ancho de banda de aproximadamente 100-400 kilobits por segundo. Puede que no esté disponible un ancho de banda tan grande o, en caso de estar disponible, puede que sea costoso. Por el contrario, los subtítulos cerrados tienen asignado generalmente un ancho de banda de 9,6 kilobits por segundo, con un espacio de datos real de aproximadamente 1 kilobit por segundo. Así pues, si no está disponible el gran ancho de banda requerido para transmitir contenido de audio, los usuarios con discapacidad visual no pueden disponer del servicio de descripción de vídeo basada en audio.

La patente de EE. UU. US 5.900.908 describe un aparato y método para proporcionar servicios de televisión descritos mediante los cuales se ofrece al espectador una descripción de audio de aspectos no hablados de un programa de televisión. Un editor de descripción (una persona) introduce datos de descripción referentes a un programa audiovisual, y son insertados, por ejemplo, en la línea 21 del intervalo de supresión vertical del programa.

Compendio

La invención se define en las reivindicaciones independientes 1, 8 y 10. En las reivindicaciones dependientes se exponen realizaciones particulares de la invención.

Realizaciones de la presente invención proporcionan servicios de descripción de vídeo que ofrecen descripciones auxiliares de programa, basadas en texto, en particular para usuarios con discapacidad visual. Realizaciones de la presente invención proporcionan servicios de descripción de vídeo que consumen menos tiempo y son menos costosos en sus procesos de producción, al tiempo que mantienen o incluso mejoran el detalle y la precisión del contenido. Realizaciones de la presente invención proporcionan servicios de descripción de vídeo que, durante la distribución a los usuarios, consumen menos ancho de banda que servicios convencionales de descripción de vídeo basada en audio.

Según una realización ilustrativa de la presente invención, un método para generar indicaciones escénicas sincrónicas en el tiempo puede conllevar: crear, a partir de un guion de producción, una lista secuencial en el tiempo de flujos de diálogo y una lista secuencial en el tiempo de indicaciones escénicas, correlacionar la lista secuencial en el tiempo de flujos de diálogo con una lista de flujos de subtítulos cerrados, en donde se asocian códigos de tiempo para diálogo con los flujos de subtítulos cerrados, asignar los códigos de tiempo para diálogo asociados con los flujos de subtítulos cerrados a los flujos de diálogo correlacionados, calcular códigos de tiempo para indicación escénica entre códigos de tiempo para diálogo consecutivos, sobre la base de al menos uno de los códigos de tiempo para diálogo consecutivos y de la cantidad de indicaciones escénicas entre flujos de diálogo consecutivos, y asignar los códigos de tiempo para indicación escénica calculados a la lista secuencial en el tiempo de indicaciones escénicas.

En una variante ilustrativa de esta realización, el paso de correlacionar incluye realizar la correlación por medio de un algoritmo correlacionador de ajuste no estrecho.

En una variante ilustrativa de esta realización, el paso de calcular incluye determinar, para dos códigos de tiempo para diálogo consecutivos, una diferencia entre los códigos de tiempo para diálogo consecutivos, dividir la diferencia por uno más la cantidad N de indicaciones escénicas entre flujos de diálogo consecutivos asociados con los dos códigos de tiempo para diálogo consecutivos, con el fin de calcular un intervalo de código de tiempo para indicación escénica, multiplicar el intervalo de código de tiempo para indicación escénica por una z^{ésima} indicación escénica entre los flujos de diálogo consecutivos, con el fin de calcular un intervalo de código de tiempo para la z^{ésima} indicación escénica, y sumar el intervalo de código de tiempo para la z^{ésima} indicación escénica a un código de tiempo más antiguo de los códigos de tiempo para diálogo consecutivos.

En una realización ilustrativa del método, el método puede conllevar además distribuir las indicaciones escénicas sincrónicas en el tiempo como flujos de texto mediante al menos uno de sistemas de supresión vertical, de sistemas de emisión en continuo (en inglés, "streaming") por Internet y de autoría de DVD.

En una variante ilustrativa de esta realización, el al menos uno de sistemas de supresión vertical y de sistemas de emisión en continuo por Internet incluye al menos uno de un sistema de subtitulación cerrada CEA-608, de un

sistema de subtítulos cerrado CEA-708, de un sistema de teletexto de sistema mundial (WST, del inglés "World-System Teletext"), de un sistema de difusión de vídeo digital (DVB, del inglés "Digital Video Broadcasting") y de un sistema de texto temporizado para multimedia (SMIL).

5 En una realización ilustrativa del método, el método puede conllevar además recibir las indicaciones escénicas sincrónicas en el tiempo y descodificar las indicaciones escénicas sincrónicas en el tiempo utilizando un sistema basado en texto.

En una variante ilustrativa de esta realización, el sistema basado en texto incluye al menos uno de un dispositivo Braille, de un sintetizador de texto a voz y de una visualización de letra grande en pantalla.

10 Según una realización ilustrativa de la presente invención, un método para proporcionar descripciones de indicaciones escénicas puede conllevar: crear descripciones textuales de las indicaciones escénicas de un guion de producción y transmitir las descripciones textuales a un sistema de representación.

En una realización ilustrativa del método, el método puede conllevar además visionar al menos una de una producción en directo y de una producción pregrabada, incluidos el diálogo y las indicaciones escénicas, en donde las descripciones textuales de las indicaciones escénicas están basadas en el visionado de la producción.

15 En una realización ilustrativa del método, el método puede conllevar además crear descripciones textuales del diálogo y transmitir las descripciones textuales del diálogo al sistema de representación, en donde: o bien se transmiten al sistema de representación a través de canales de transmisión distintos las descripciones textuales de las indicaciones escénicas y las descripciones textuales del diálogo, o bien se incluye un código junto con las descripciones textuales de las indicaciones escénicas, efectuando el sistema de representación la distinción entre
20 las descripciones textuales de las indicaciones escénicas y las descripciones textuales del diálogo basándose en el código.

En una variante ilustrativa de esta realización, las descripciones textuales son transmitidas como flujos de texto mediante al menos uno de sistemas de supresión vertical, de sistemas de paquetes de datos y de autoría de DVD.

25 En una variante ilustrativa de esta realización, el al menos uno de sistemas de supresión vertical y de sistemas de paquetes de datos incluye al menos uno de un sistema de subtítulos cerrado CEA-608, de un sistema de subtítulos cerrado CEA-708, de un sistema de teletexto de sistema mundial (WST), de un sistema de difusión de vídeo digital (DVB) y de un sistema de texto temporizado para multimedia (SMIL).

30 En una realización ilustrativa del método, el método puede conllevar además recibir las descripciones textuales junto con un flujo de texto de subtítulos cerrados, omitir, basándose en el flujo de texto de subtítulos cerrados, las descripciones textuales de entre los subtítulos cerrados visualizados en la pantalla, y representar las descripciones textuales.

En una variante ilustrativa de esta realización, las descripciones textuales son representadas mediante al menos uno de un dispositivo Braille, de un sintetizador de texto a voz y de una visualización de letra grande en pantalla.

35 Según una realización ilustrativa de la presente invención, un método para proporcionar descripciones textuales de indicaciones escénicas puede conllevar al menos uno de a) leer un guion de producción, y b) visionar al menos una de una producción en directo y de una producción pregrabada, incluidos el diálogo y las indicaciones escénicas, crear descripciones de audio de las indicaciones escénicas y convertir las descripciones de audio en descripciones textuales de las indicaciones escénicas utilizando un sistema de reconocimiento de voz.

40 En una realización ilustrativa del método, el método puede conllevar además distribuir las descripciones textuales como flujos de texto mediante al menos uno de sistemas de supresión vertical, de sistemas de emisión en continuo de paquetes de datos y de autoría de DVD.

45 En una variante ilustrativa de esta realización, el al menos uno de sistemas de supresión vertical y de sistemas de emisión en continuo de paquetes de datos incluye al menos uno de un sistema de subtítulos cerrado CEA-608, de un sistema de subtítulos cerrado CEA-708, de un sistema de teletexto de sistema mundial (WST), de un sistema de difusión de vídeo digital (DVB) y de un sistema de texto temporizado para multimedia (SMIL).

En una realización ilustrativa del método, el método puede conllevar además recibir las descripciones textuales de las indicaciones escénicas y descodificar las descripciones textuales a través de un sistema de representación.

En una variante ilustrativa de esta realización, el sistema de representación incluye al menos uno de un dispositivo Braille, de un sintetizador de texto a voz y de una visualización de letra grande en pantalla.

50 Según una realización ilustrativa de la presente invención, un método para utilizar descripciones textuales con el fin de proporcionar descripciones auxiliares de programa a consumidores con discapacidad visual puede conllevar: transmitir un flujo de descripciones textuales de indicaciones escénicas para al menos una de una producción en directo y de una producción pregrabada, y representar el flujo de descripciones textuales de indicaciones escénicas a través de al menos uno de un dispositivo Braille, de un sintetizador de texto a voz y de una visualización de letra

grande en pantalla.

Breve descripción de los dibujos

La Figura 1 es un diagrama que ilustra componentes de un sistema convencional de subtítulos cerrada.

5 La Figura 2 es un diagrama que ilustra componentes de un servicio convencional de descripción de vídeo basada en audio.

La Figura 3 es un diagrama que ilustra componentes de un servicio de descripción de vídeo basada en texto que utiliza algoritmos desglosadores basados en guion y algoritmos sincronizadores, según una realización ilustrativa de la presente invención.

10 La Figura 4 es un diagrama que ilustra componentes detallados adicionales del servicio de descripción de vídeo basada en texto de la Figura 3, según una realización ilustrativa de la presente invención.

La Figura 5 es un diagrama que ilustra componentes de un servicio de descripción de vídeo basada en texto en paralelo con un sistema de subtítulos cerrada, según una realización ilustrativa alternativa de la presente invención.

15 La Figura 6 es un diagrama que ilustra componentes de un servicio de descripción de vídeo basada en texto que utiliza un sistema de reconocimiento de voz, según otra realización ilustrativa alternativa de la presente invención.

La Figura 7 es un diagrama que ilustra componentes de un sistema de distribución para un servicio de descripción de vídeo basada en texto, según una realización ilustrativa de la presente invención.

La Figura 8 es un diagrama que ilustra componentes de una ubicación de visionado para un servicio de descripción de vídeo basada en texto, según una realización ilustrativa de la presente invención.

20 Descripción detallada de realizaciones ilustrativas

La Figura 3 ilustra componentes de un servicio de descripción de vídeo basada en texto que utiliza algoritmos desglosadores basados en guion y algoritmos sincronizadores, según una realización ilustrativa de la presente invención. Un guion 2 de producción incluye generalmente dos elementos principales: diálogo e indicaciones escénicas. El diálogo es aquella parte del guion 2 de producción que representa los diálogos que los actores van a mantener durante un programa. Ello puede incluir palabras habladas tales como, por ejemplo, "Te quiero", pronunciadas por un actor, según se muestra en la Figura 4. Las indicaciones escénicas son aquella parte del guion 2 de producción que representa las instrucciones que los actores y/o las estructuras de escena deben seguir durante un programa. Ello puede incluir las acciones de un actor tales como, por ejemplo, la indicación [Tom besa a Molly], según se muestra en la Figura 4. Un guion 2 de producción puede incluir otros elementos, tales como indicaciones técnicas para cámaras, iluminación, etc.; sin embargo, aquí no se discutirán con mayor detalle esos otros elementos.

Los guiones 2 de producción siguen generalmente ciertas convenciones de formato en la escritura de guiones. Por ejemplo, para indicar un diálogo puede aparecer el nombre de un actor, seguido de dos puntos (p. ej., Tom:) y centrado en la página, inmediatamente antes del diálogo para ese actor. Se pueden seguir esta y otras convenciones de formato para indicar los diálogos y las indicaciones escénicas. En particular, los diálogos y las indicaciones escénicas se pueden señalar mediante diversas convenciones de formato dentro de un guion 2 de producción, con el fin de diferenciar entre sí estos elementos. Si se reconocen las convenciones de formato utilizadas en un guion 2 de producción, se puede emplear un algoritmo desglosador 21 para desglosar una copia electrónica de un guion 2 de producción, dando un guion desglosado 22 que tiene dos listas separadas, una que contiene los diálogos y otra que contiene las indicaciones escénicas.

Según se ha comentado más arriba, los subtítulos cerrados 5 incluyen flujos de diálogo que han sido grabados por un taquígrafo 3 que escuchaba una producción 1 en directo o pregrabada. Además, partes del flujo de diálogo de los subtítulos cerrados 5 pueden estar asociadas con un código de tiempo. El código de tiempo puede establecer el orden secuencial de todos los flujos de diálogo de los subtítulos cerrados 5.

45 Se puede emplear un algoritmo sincronizador 23 para sincronizar la lista de diálogos del guion desglosado 22 con los flujos de diálogo de los subtítulos cerrados 5. Dado que es posible que los flujos de diálogo de los subtítulos cerrados 5 no se correlacionen exactamente con los diálogos del guion desglosado 22 debido a diversas causas, por ejemplo cambios, improvisaciones y otras causas, el algoritmo sincronizador 23 puede utilizar una correlación "aproximada" entre el guion desglosado 22 y los subtítulos cerrados 5. De este modo, el algoritmo sincronizador 23 crea una línea temporal 24 del programa.

Se pueden sincronizar las indicaciones escénicas del guion desglosado 22 con la línea temporal 24 del programa, dando como resultado indicaciones escénicas sincronizadas 25. Las indicaciones escénicas sincronizadas 25 son descripciones basadas en texto, que después son distribuidas 7 a los usuarios, como se describirá con mayor detalle más adelante, en relación con la Figura 7. Además, cuando los usuarios reciben las indicaciones escénicas

sincronizadas 25 en una ubicación 8 de visionado, las descripciones basadas en texto pueden ser descodificadas utilizando cualquier sistema 11 basado en texto, adecuado, como se describirá con mayor detalle más adelante, en relación con la Figura 8.

5 La Figura 4 ilustra componentes más detallados del servicio de descripción de vídeo basada en texto de la Figura 3, según una realización ilustrativa de la presente invención. Un algoritmo desglosador 21 puede desglosar una copia electrónica de un guion 2 de producción en diálogos 22A e indicaciones escénicas 22B. Un guion 2 de producción y sus subtítulos cerrados 5 asociados pueden incluir cualquier cantidad de diálogos e indicaciones escénicas entremezclados entre sí. Según se muestra en la Figura 4, los diálogos 22A pueden incluir, por ejemplo, el diálogo A, el diálogo B y el diálogo C. Además, las indicaciones escénicas 22B pueden incluir, por ejemplo, la indicación A' y la indicación B'. La indicación A' se representa situada entre los diálogos A y B, y la indicación B' se representa situada entre los diálogos B y C. Además, según se muestra en la Figura 4, los subtítulos cerrados 5 pueden incluir, por ejemplo, el subtítulo A en el código de tiempo A, el subtítulo B en el código de tiempo B y el subtítulo C en el código de tiempo C.

15 Un algoritmo sincronizador 23 puede sincronizar los diálogos 22A con los subtítulos cerrados 5, correlacionando cada diálogo A, B, C con su subtítulo asociado A, B, C. Según se ha explicado con mayor detalle más arriba, dado que es posible que los diálogos 22A no se correlacionen exactamente con los subtítulos cerrados 5 debido a diversas causas, se puede utilizar una correlación "aproximada". Se puede utilizar cualquier algoritmo correlacionador de ajuste no estrecho, apropiado, para encontrar las correlaciones aproximadas. Al estar cada diálogo A, B, C correlacionado con su subtítulo asociado A, B, C, a cada diálogo se le asigna también el código de tiempo 24A de su subtítulo asociado. Por ejemplo, al estar el diálogo A correlacionado con el subtítulo A, al diálogo A se le asigna el código de tiempo A; al estar el diálogo B correlacionado con el subtítulo B, al diálogo B se le asigna el código de tiempo B; al estar el diálogo C correlacionado con el subtítulo C, al diálogo C se le asigna el código de tiempo C. Mediante la correlación de los diálogos 22A con los subtítulos cerrados 5, se crea una línea temporal 24 del programa.

25 Se pueden sincronizar las indicaciones escénicas 22B con la línea temporal 24 del programa. En el ejemplo que se muestra en la Figura 4, dado que la indicación A' se produce entre los diálogos A y B, a la indicación A' se le asigna un código de tiempo A' entre los códigos de tiempo A, B para los diálogos A, B. Además, dado que la indicación B' se produce entre los diálogos B y C, a la indicación B' se le asigna un código de tiempo B' entre los códigos de tiempo B, C para los diálogos B, C. Además, para este paso, se puede emplear un algoritmo de inserción en el tiempo para determinar los códigos de tiempo para las indicaciones escénicas 22B. Para el caso en que exista una única indicación escénica A' entre dos códigos de tiempo para diálogo A y B, el algoritmo de inserción en el tiempo se puede expresar, por ejemplo, como:

Código de tiempo para indicación escénica A' =

$$\text{Código de tiempo para diálogo A} + \left(\frac{\text{Código de tiempo para diálogo B} - \text{Código de tiempo para diálogo A}}{2} \right),$$

35 donde 2 es la cantidad de unidades de tiempo en las que se expresan los códigos de tiempo. Como se puede ver en el precedente algoritmo de inserción en el tiempo, se inserta una única indicación escénica A', que se encuentra entre los diálogos A y B, en un código de tiempo A' de indicación escénica que está a medio camino entre los correspondientes códigos de tiempo para diálogo A y B.

40 Se puede ampliar la ecuación precedente para que tenga aplicación en casos en donde existan una o varias indicaciones escénicas entre dos códigos de tiempo para diálogo A y B. La ecuación ampliada se puede expresar como:

Código de tiempo para indicación escénica A'_z =

$$\text{Código de tiempo para diálogo A} + z \left(\frac{\text{Código de tiempo para diálogo B} - \text{Código de tiempo para diálogo A}}{N+1} \right),$$

45 donde N es la cantidad de indicaciones escénicas que existen entre los diálogos A y B, y z es una en particular de las N indicaciones escénicas. Por ejemplo, en un caso en el que existen tres indicaciones escénicas entre el diálogo A y el diálogo B, a N se le asigna un valor de 3. A continuación, utilizando el precedente algoritmo de inserción en el tiempo para determinar el código de tiempo para indicación escénica de una en particular de las tres indicaciones escénicas entre los diálogos A y B, a z se le asigna un valor de 1, 2 o 3. Así, para la primera indicación escénica A'₁ entre los diálogos A y B, a z se le asigna el valor de 1. Análogamente, para la segunda indicación escénica A'₂ entre los diálogos A y B, a z se le asigna el valor de 2, y para la tercera indicación escénica A'₃ entre los diálogos A y B, a z se le asigna el valor de 3. Empleando el precedente algoritmo de inserción en el tiempo, se pueden asignar a las N indicaciones escénicas códigos de tiempo que están igualmente espaciados entre los diálogos A y B, entre los cuales se da la totalidad de las N indicaciones escénicas.

Como se puede apreciar al analizar los dos algoritmos precedentes, el primer algoritmo, más simple, es un caso especial del segundo algoritmo, donde los valores de N y z son ambos 1. Para el ejemplo que se muestra en la Figura 4, con el precedente algoritmo de inserción en el tiempo se pueden asignar a las indicaciones A', B' códigos de tiempo para indicación escénica A', B', lo que origina indicaciones escénicas 25 que están sincronizadas con respecto a los subtítulos A, B, C. Específicamente, a la indicación A' se le asigna el código de tiempo A' que se encuentra entre los subtítulos A y B, y a la indicación B' se le asigna el código de tiempo B' que se encuentra entre los subtítulos B y C.

En una realización ilustrativa alternativa, se puede utilizar el diálogo como contenido de los subtítulos cerrados, en cuyo caso se puede omitir el paso inicial de correlacionar los diálogos con los subtítulos cerrados.

Las indicaciones escénicas sincronizadas 25 son descripciones basadas en texto, que se distribuyen 7 a los usuarios, como se describirá con mayor detalle más adelante, en relación con la Figura 7. Además, cuando los usuarios reciben las indicaciones escénicas sincronizadas 25 en una ubicación 8 de visionado, se pueden descodificar las descripciones basadas en texto utilizando cualquier sistema 11 basado en texto adecuado, como se describirá con mayor detalle más adelante, en relación con la Figura 8.

Por lo tanto, en una realización ilustrativa de la presente invención se desglosa mediante un algoritmo desglosador un guion de producción, para dar un guion desglosado que incluye listas separadas de diálogos y de indicaciones escénicas. A continuación, un algoritmo sincronizador sincroniza el guion desglosado con los subtítulos cerrados, produciendo una línea temporal de programa. Después se sincronizan el guion desglosado y la línea temporal de programa, dando como resultado indicaciones escénicas basadas en texto que están sincronizadas con el contenido de audio del programa. Entonces se distribuyen las indicaciones escénicas sincronizadas resultantes a una ubicación receptora, donde se pueden descodificar las descripciones basadas en texto y representarlas como se describirá más adelante.

Los algoritmos desglosadores y sincronizadores de las Figuras 3 y 4, descritos en lo que antecede, pueden ejecutarse en tiempo real mientras se difunde una producción a los usuarios. Como alternativa, se pueden ejecutar los algoritmos antes de la difusión, utilizando un programa de subtítulos cerrados pregrabados. Además, los algoritmos generalmente se pueden ejecutar en las instalaciones de una emisora, que es donde se encuentra generalmente el guion de producción. No obstante, los algoritmos se pueden ejecutar en cualquier lugar en donde estén disponibles una copia electrónica del guion de producción y los subtítulos cerrados.

La Figura 5 ilustra componentes de un servicio de descripción de vídeo basada en texto, en paralelo con un sistema de subtitulación cerrada, según una realización ilustrativa alternativa que no forma parte de la presente invención. Según se ha descrito más arriba, en un sistema tradicional de subtitulación cerrada, un taquígrafo 3 escucha una producción en directo o pregrabada 1 y escribe el flujo 5 de subtítulos cerrados. El flujo 5 de subtítulos cerrados se puede distribuir entonces a los usuarios mediante sistemas de subtitulación cerrada y, cuando se visiona un programa en particular en una ubicación 8 de visionado, se puede mostrar 9 el programa conjuntamente con los subtítulos cerrados en la pantalla.

Según se muestra en la Figura 5, el sistema de subtitulación cerrada existente puede mejorarse añadiendo a un canal de subtítulos cerrados, p. ej., cc-2, un canal adicional y separado, p. ej., cc-4, para transmitir descripciones de vídeo. Después, el receptor puede dar salida a estos datos por separado, a través de dispositivos separados. En un sistema mejorado de este tipo, el taquígrafo 3 no solo escucharía, sino que también vería la producción en directo o pregrabada 1. Mientras el taquígrafo 3 escucha y ve un programa, el taquígrafo 3 puede escribir una salida separada 26 de servicio de descripción de vídeo (VDS, del inglés "Video Description Service"), que describe los eventos, acciones, ubicaciones, expresiones y otro contenido de vídeo, es decir, indicaciones escénicas, en paralelo al flujo 5 de subtítulos cerrados y de manera simultánea con el mismo.

En lugar de añadir un canal separado para el servicio de descripción de vídeo, se puede integrar la salida VDS con la salida de subtítulos cerrados, pero se puede distinguir de la salida de subtítulos cerrados mediante el uso de códigos distintos, p. ej., códigos de introducción distintos antes de cada entrada o antes de cada primera entrada después de un cambio entre salida de subtítulos cerrados y salida VDS. El dispositivo de representación en el extremo de usuario puede determinar así cuál de los textos es una descripción de vídeo, basándose en las etiquetas VDS.

La salida VDS incluye descripciones basadas en texto, que luego son distribuidas 7 a los usuarios, como se describirá con mayor detalle más adelante, en relación con la Figura 7. Sin embargo, cuando los usuarios reciben en una ubicación 8 de visionado la salida VDS 26, se puede omitir esta de la visualización normal en pantalla de los subtítulos cerrados 9 y, en lugar de ello, se puede redirigir la salida VDS a un sistema 11 basado en texto. Se pueden descodificar las descripciones basadas en texto utilizando cualquier sistema 11 basado en texto adecuado, como se describirá con mayor detalle más adelante, en relación con la Figura 8. Además, si la salida VDS 26 se ha grabado utilizando un canal separado de los subtítulos cerrados 9, los usuarios pueden emitir directamente la salida VDS 26 mediante un dispositivo de salida distinto del utilizado para la salida de los subtítulos cerrados 9. Por ejemplo, se puede dar salida a los subtítulos cerrados 9 hacia un altavoz y enviar la salida VDS 26 hacia otro altavoz separado, de forma que un usuario pueda escuchar ambas salidas simultáneamente.

La Figura 6 ilustra componentes de un servicio de descripción de vídeo basada en texto que utiliza un sistema de reconocimiento de voz, según otra realización ilustrativa alternativa que no forma parte de la presente invención. Según se ha descrito más arriba, en un servicio de descripción basado en audio, un narrador 4 crea una pista 6 de audio de contenido de vídeo mientras lee un guion 2 de producción o bien ve una producción en directo o pregrabada 1. Después se puede distribuir 12 la pista 6 de audio a los usuarios mediante métodos tradicionales de difusión de audio y, cuando se escucha un programa en particular en una ubicación 8 de visionado, entonces se puede reproducir el programa junto con la pista 6 de audio.

Según se muestra en la Figura 6, se puede mejorar el servicio de descripción basada en audio existente convirtiendo la pista 6 de audio en descripciones 27 basadas en texto, utilizando *software* y/o sistemas de reconocimiento de voz. Para la conversión de descripciones basadas en audio a descripciones basadas en texto se puede utilizar cualquier *software* o sistema de reconocimiento de voz. Según se ha expuesto con mayor detalle más arriba, las descripciones basadas en texto resultantes requieren menos ancho de banda para su distribución que las descripciones originales basadas en audio. Además, se puede recurrir con mayor facilidad a narradores no profesionales, lo que reduce los costes de producción. Así mismo, aunque se recurra a narradores profesionales para generar las descripciones basadas en audio, los costes de tal producción pueden seguir siendo menores, ya que no se distribuyen a los usuarios las voces reales grabadas de los narradores.

Las descripciones basadas en audio, convertidas, son descripciones 27 basadas en texto, que luego son distribuidas 7 a los usuarios, como se describirá con mayor detalle más adelante, en relación con la Figura 7. Además, cuando los usuarios reciben las descripciones 27 basadas en texto en una ubicación 8 de visionado, las descripciones basadas en texto pueden ser descodificadas utilizando cualquier sistema 11 basado en texto adecuado, como se describirá con mayor detalle más adelante, en relación con la Figura 8. La temporización de la salida puede basarse en la temporización que se ha dado a las descripciones basadas en audio durante su conversión a descripciones basadas en texto.

La Figura 7 ilustra componentes de un sistema de distribución para un servicio de descripción de vídeo basada en texto, según una realización ilustrativa de la presente invención. Las descripciones 25, 26, 27 basadas en texto pueden estar insertadas en el flujo de servicio de texto de flujos de supresión vertical o de tipo de datos. Los ejemplos de sistemas que proporcionan tales flujos incluyen sistemas 31 de subtítulos cerrada CEA-608 (T-2), sistemas 32 de subtítulos cerrada CEA-708 (T-2), teletexto 33 de sistema mundial (WST) utilizado internacionalmente, flujos 34 de subtítulos en difusión de vídeo digital (DVB) o sistemas 36 de texto temporizado utilizados en Internet 37 para multimedia tales como SMIL. Los sistemas 31 y 32 de subtítulos cerrada, el WST 33 y la DVB 34 son sistemas 35 de distribución basados en televisión que se pueden distribuir por televisión, cable u otros métodos de transmisión. Los sistemas 36 de texto temporizado son sistemas 38 de distribución basados en ordenador que se pueden distribuir a través de Internet 37 o por otros métodos. Además, las descripciones 25, 26, 27 basadas en texto pueden ser sometidas a un proceso 39 de autoría de DVD u otro proceso de producción de medios de almacenamiento, y distribuirse a través de un DVD 40 u otro medio de almacenamiento.

En el ejemplo de los sistemas 31 de subtítulos cerrada CEA-608, el T-2 (servicio de texto, canal 2) es transportado en el campo-1, línea-21 del VBI (intervalo de supresión vertical, por sus siglas en inglés), junto con los subtítulos cerrados, e intercalado en su flujo de datos de aproximadamente 120 bits por segundo, lo que resulta adecuado para transmitir texto de la naturaleza aquí descrita. Así pues, el servicio de texto consume mucho menos ancho de banda que el audio. De hecho, la transmisión de CEA-608, por ejemplo, tiene ancho de banda permanentemente asignado tanto en VBI como en dominios comprimidos MPEG, se utilice o no. Por lo tanto, no se requiere ancho de banda adicional para transmitir un servicio de descripción de vídeo basada en texto dentro de este ancho de banda permanentemente asignado.

La Figura 8 ilustra componentes de una ubicación de visionado para un servicio de descripción de vídeo basada en texto, según una realización ilustrativa de la presente invención. Las descripciones 25, 26, 27 basadas en texto pueden recibirse a través de sistemas 35 de distribución basados en televisión, sistemas 38 de distribución basados en ordenador, medios 40 de almacenamiento u otros sistemas de distribución. A continuación, del flujo de datos recibido se extraen las descripciones basadas en texto recibidas (a las descripciones extraídas se las designa con la referencia numérica 40 en la Figura 8). Por ejemplo, utilizando métodos de recepción tales como televisión aérea, descodificador de aparato de televisión para cable (en inglés, "set-top-box") u otros métodos similares, o bien un descodificador de aparato de televisión puede recibir internamente el flujo de datos o bien se puede utilizar una señal de vídeo, tal como como una banda de base analógica NTSC, para extraer las descripciones 41 basadas en texto.

A continuación, las descripciones 41 basadas en texto recibidas y extraídas pueden ser descodificadas por diversos dispositivos para "presentar" el texto a usuarios con discapacidad visual. Por ejemplo, las descripciones 41 basadas en texto pueden ser procesadas por un dispositivo Braille 42 que presenta al usuario con discapacidad visual las descripciones 41 basadas en texto en el formato de puntos en relieve. Así, el usuario puede leer las descripciones 41 basadas en texto en el dispositivo Braille mientras escucha también el contenido de audio normal del programa. Dependiendo de su estructura, las descripciones 41 basadas en texto pueden tener prefijos de puntuación u otros caracteres compatibles con texto que pueden no ser óptimos para la conversión de texto a voz, pero sí apropiados para dicho dispositivo Braille.

De manera adicional o alternativa, las descripciones 41 basadas en texto pueden ser procesadas mediante un sintetizador 43 de texto a voz que convierte las descripciones 41 basadas en texto en palabras habladas. Así, el usuario puede oír la versión hablada convertida de las descripciones 41 basadas en texto mientras también escucha el contenido de audio normal del programa. Dependiendo de su estructura, se pueden seleccionar algunos caracteres adicionales o la ortografía fonética de las descripciones 41 basadas en texto para que sean objeto de un uso específico por parte de un sistema de conversión de texto a voz, a fin de pronunciar con mayor precisión nombres o palabras confusas.

Además, para usuarios con visión parcial, se pueden procesar las descripciones 41 basadas en texto y reproducirlas en un dispositivo 44 de visualización de letra grande en pantalla. Así, un usuario con visión parcial puede leer el texto de letra grande visualizado de las descripciones 41 basadas en texto, mientras también escucha el contenido de audio normal del programa.

Cuando se utiliza un sintetizador 43 de texto a voz, puede ocurrir que las palabras habladas convertidas desde las descripciones 41 basadas en texto sean emitidas de forma auditiva al mismo tiempo que el contenido de audio normal del programa. En general, se debe evitar esta situación para que no se superpongan las dos salidas de audio una sobre otra, aunque un usuario podría ser capaz de escuchar salidas simultáneas del sintetizador 43 de texto a voz y del contenido de audio normal del programa. Sin embargo, cuando se utiliza un dispositivo Braille 42 o un dispositivo 44 de visualización de letra grande en pantalla, no existe tal problema potencial de superposición de audios, ya que el usuario puede leer desde un dispositivo Braille 42 o en un dispositivo 44 de visualización de letra grande en pantalla, mientras también escucha el contenido de audio normal del programa.

Una realización ilustrativa de la presente invención está dirigida a uno o varios procesadores, que se pueden implementar utilizando cualquier circuito y dispositivo de procesamiento convencional, o combinación de estos, por ejemplo una unidad central de procesamiento (CPU, del inglés "Central Processing Unit") de un ordenador personal (PC, de "Personal Computer"), para ejecutar código suministrado, p. ej., en un medio legible por ordenador implementado en *hardware*, a fin de poner en práctica los métodos ilustrativos descritos en lo que antecede, o partes de los mismos. El medio legible por ordenador puede incluir cualquier circuito de memoria convencional permanente y/o transitoria, o sus combinaciones, en donde una lista no exhaustiva de los mismos incluye una memoria de acceso aleatorio (RAM, de "Random Access Memory"), una memoria de solo lectura (ROM, de "Read-Only Memory"), discos compactos (CD, de "Compact Disc"), un disco versátil digital (DVD, de "Digital Versatile Disc") y una cinta magnética.

Una realización ilustrativa de la presente invención está dirigida a uno o varios medios legibles por ordenador implementados en *hardware* que contienen almacenadas instrucciones que, cuando son ejecutadas, hacen que uno o varios procesadores pongan en práctica los métodos ilustrativos descritos en lo que antecede, o partes de los mismos.

Una realización ilustrativa de la presente invención está dirigida a un método para transmitir instrucciones ejecutables por uno o varios procesadores, instrucciones que, cuando son ejecutadas, hacen que el o los procesadores pongan en práctica los métodos ilustrativos descritos en lo que antecede, o partes de los mismos.

A partir de la descripción precedente, los expertos en la materia pueden apreciar que la presente invención puede implementarse de diversas maneras.

REIVINDICACIONES

1. Un método para proporcionar indicaciones escénicas, que comprende los pasos de:
 desglosar, mediante un procesador de ordenador, diálogo e indicaciones escénicas de un guion de producción para dar una lista secuencial en el tiempo de flujos de diálogo y una lista secuencial en el tiempo de indicaciones escénicas, en donde la lista secuencial en el tiempo de indicaciones escénicas incluye una descripción textual de las indicaciones escénicas;
 correlacionar, utilizando el procesador, la lista secuencial en el tiempo de flujos de diálogo con una lista de flujos de subtítulos cerrados, en donde se asocian códigos de tiempo para diálogo con los flujos de subtítulos cerrados;
 asignar, utilizando el procesador, los códigos de tiempo para diálogo asociados con los flujos de subtítulos cerrados a los flujos de diálogo correlacionados;
 calcular, utilizando el procesador, códigos de tiempo para indicación escénica entre códigos de tiempo para diálogo consecutivos, sobre la base de los códigos de tiempo para diálogo consecutivos y de la cantidad de indicaciones escénicas entre los flujos de diálogo consecutivos;
 asignar, utilizando el procesador, los códigos de tiempo para indicación escénica calculados a indicaciones escénicas respectivas de la lista secuencial en el tiempo de indicaciones escénicas; y
 transmitir la lista secuencial en el tiempo de indicaciones escénicas a un sistema de representación en sincronía con una representación de un programa, estando basada la sincronía en los códigos de tiempo para indicación escénica asignados.
2. El método según la reivindicación 1, en donde el paso de calcular incluye:
 determinar, para dos códigos de tiempo para diálogo consecutivos, una diferencia entre los códigos de tiempo para diálogo consecutivos;
 dividir la diferencia entre uno más la cantidad N de indicaciones escénicas entre flujos de diálogo consecutivos asociados con los dos códigos de tiempo para diálogo consecutivos, con el fin de calcular un intervalo de código de tiempo para indicación escénica;
 multiplicar el intervalo de código de tiempo para indicación escénica por un intervalo de código de tiempo para la z^{ésima} indicación escénica; y
 sumar el intervalo de código de tiempo para la z^{ésima} indicación escénica a un código de tiempo más antiguo de los códigos de tiempo para diálogo consecutivos.
3. El método según la reivindicación 1, en donde:
 o bien la lista secuencial en el tiempo de indicaciones escénicas y la lista secuencial en el tiempo de flujos de diálogo son transmitidas al sistema de representación a través de canales de transmisión distintos;
 o bien el sistema de representación distingue, basándose en un código, entre la lista secuencial en el tiempo de indicaciones escénicas y la lista secuencial en el tiempo de diálogos.
4. El método según la reivindicación 1, que comprende además el paso de:
 transmitir la lista secuencial en el tiempo de indicaciones escénicas como flujos de texto mediante al menos uno de sistemas de supresión vertical, de sistemas de emisión en continuo por Internet y de autoría de DVD.
5. El método según la reivindicación 4, en donde el al menos uno de supresión vertical y de sistemas de emisión en continuo por Internet incluye al menos uno de un sistema de subtitulación cerrada CEA-608, de un sistema de subtitulación cerrada CEA-708, de un sistema de teletexto de sistema mundial (WST), de un sistema de difusión de vídeo digital (DVB) y de un sistema de texto temporizado para multimedia (SMIL).
6. El método según la reivindicación 1, que comprende además los pasos de:
 recibir la lista secuencial en el tiempo de indicaciones escénicas en el sistema de representación;
 recibir un flujo de texto de subtítulos cerrados;
 omitir, basándose en el flujo de texto de subtítulos cerrados, la lista secuencial en el tiempo de indicaciones escénicas de entre los subtítulos cerrados visualizados en la pantalla; y
 representar las indicaciones escénicas de la lista secuencial en el tiempo utilizando un sistema basado en texto.
7. El método según la reivindicación 6, en donde el sistema basado en texto incluye al menos uno de un dispositivo Braille, de un sintetizador de texto a voz y de una visualización de letra grande en pantalla.
8. Un medio legible por ordenador implementado en *hardware* que contiene almacenadas instrucciones ejecutables por un procesador, instrucciones que, cuando son ejecutadas por el procesador, hacen que el procesador ponga en práctica un método, comprendiendo el método los pasos de:
 desglosar diálogo e indicaciones escénicas de un guion de producción para dar una lista secuencial en el tiempo de flujos de diálogo y una lista secuencial en el tiempo de indicaciones escénicas, en donde la lista secuencial en el tiempo de indicaciones escénicas incluye una descripción textual de las indicaciones escénicas;
 correlacionar la lista secuencial en el tiempo de flujos de diálogo con una lista de flujos de subtítulos cerrados, en donde se asocian códigos de tiempo para diálogo con los flujos de subtítulos cerrados;
 asignar los códigos de tiempo para diálogo asociados con los flujos de subtítulos cerrados a los flujos de diálogo correlacionados;
 calcular códigos de tiempo para indicación escénica entre códigos de tiempo para diálogo consecutivos, sobre la

- base de los códigos de tiempo para diálogo consecutivos y de la cantidad de indicaciones escénicas entre flujos de diálogo consecutivos;
asignar los códigos de tiempo para indicación escénica calculados a indicaciones escénicas respectivas de la lista secuencial en el tiempo de indicaciones escénicas; y
- 5 transmitir la lista secuencial en el tiempo de indicaciones escénicas a un sistema de representación en sincronía con una representación de un programa, estando basada la sincronía en los códigos de tiempo para indicación escénica asignados.
9. El medio legible por ordenador implementado en *hardware* según la reivindicación 8, en donde el procesador es un componente de un dispositivo que incluye al menos uno de un dispositivo Braille, de un sintetizador de texto a voz
- 10 y de una visualización de letra grande en pantalla.
10. Un sistema para proporcionar descripciones auxiliares de programa a consumidores con discapacidad visual, comprendiendo el sistema:
un dispositivo que incluye un procesador configurado para:
desglosar diálogo e indicaciones escénicas de un guion de producción para dar una lista secuencial en el tiempo de
- 15 indicaciones escénicas, en donde la lista secuencial en el tiempo de indicaciones escénicas incluye una descripción textual de las indicaciones escénicas;
correlacionar la lista secuencial en el tiempo de flujos de diálogo con una lista de flujos de subtítulos cerrados, en donde se asocian códigos de tiempo para diálogo con los flujos de subtítulos cerrados;
- 20 asignar los códigos de tiempo para diálogo asociados con los flujos de subtítulos cerrados a los flujos de diálogo correlacionados;
calcular códigos de tiempo para indicación escénica entre códigos de tiempo para diálogo consecutivos, sobre la base de los códigos de tiempo para diálogo consecutivos y de la cantidad de indicaciones escénicas entre flujos de diálogo consecutivos;
- 25 asignar los códigos de tiempo para indicación escénica calculados a indicaciones escénicas respectivas de la lista secuencial en el tiempo de indicaciones escénicas; y
transmitir la lista secuencial en el tiempo de indicaciones escénicas a un sistema de representación en sincronía con una representación de un programa, estando basada la sincronía en los códigos de tiempo para indicación escénica asignados.

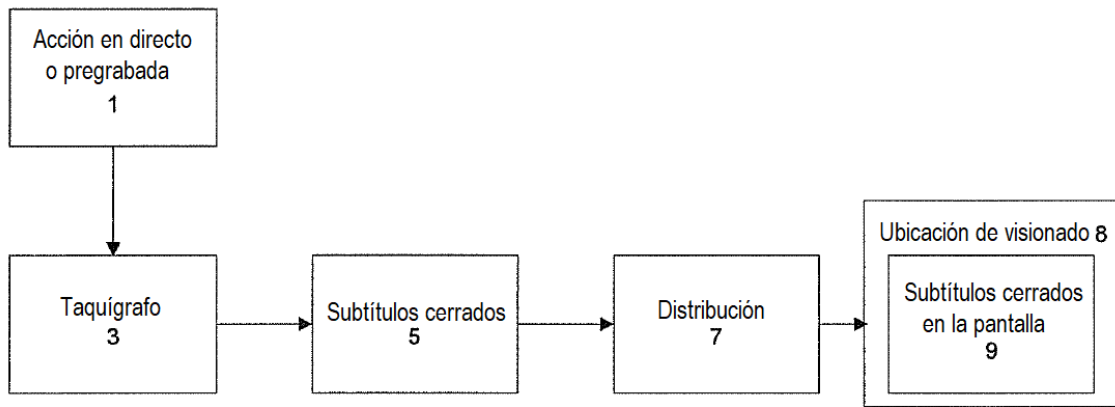


FIG. 1
Técnica anterior

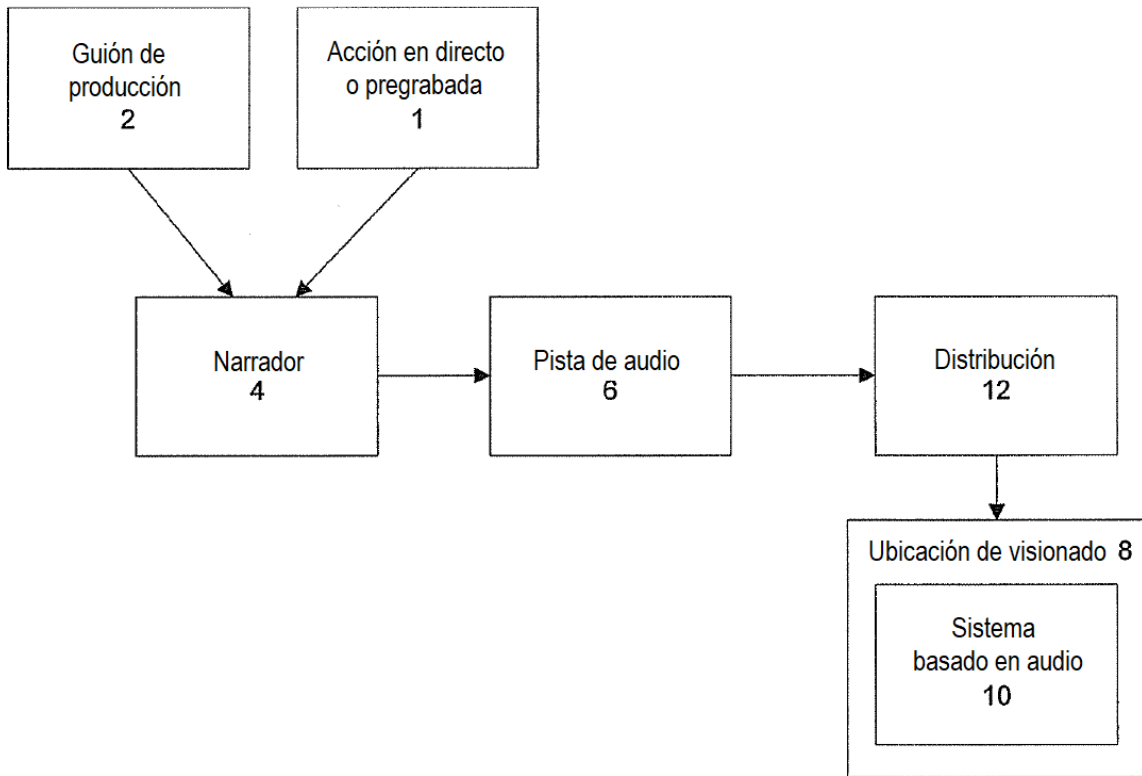


FIG. 2
Técnica anterior

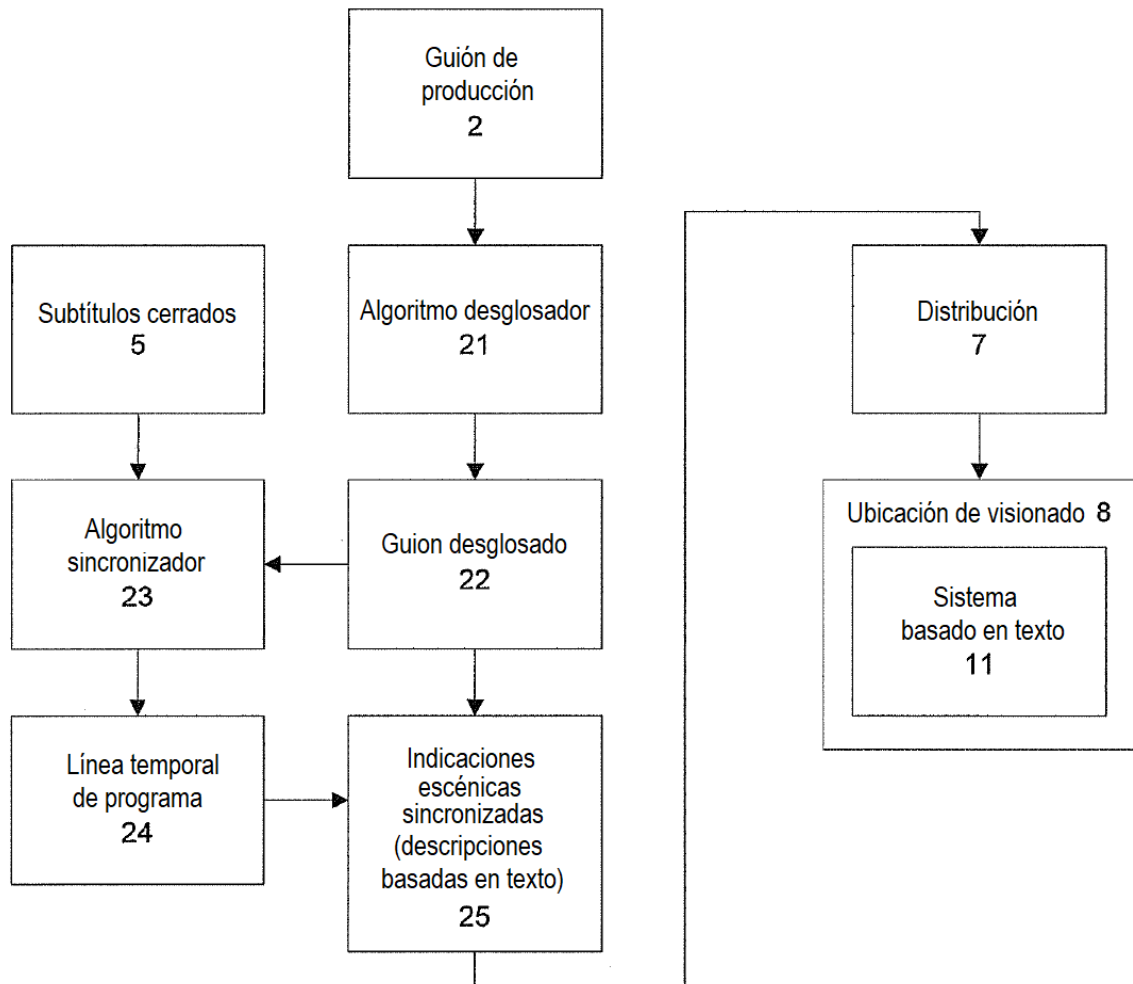


FIG. 3

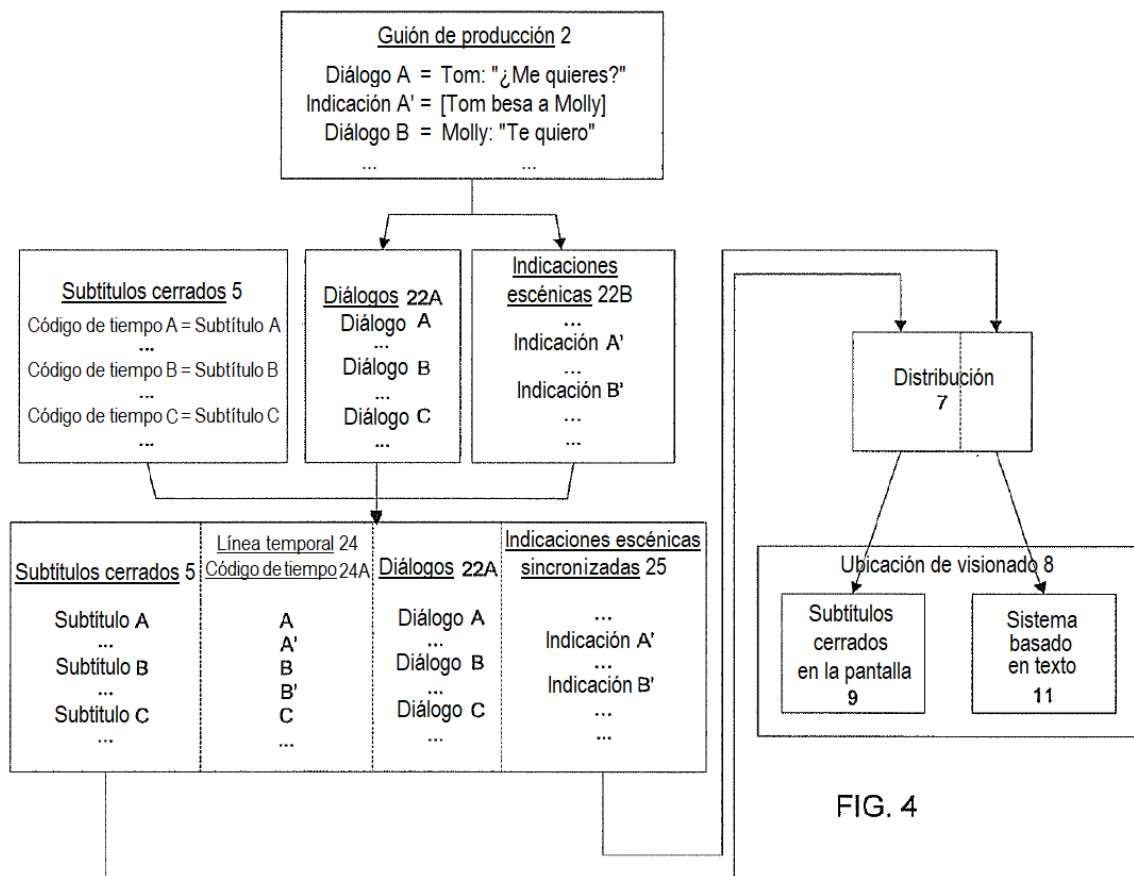


FIG. 4

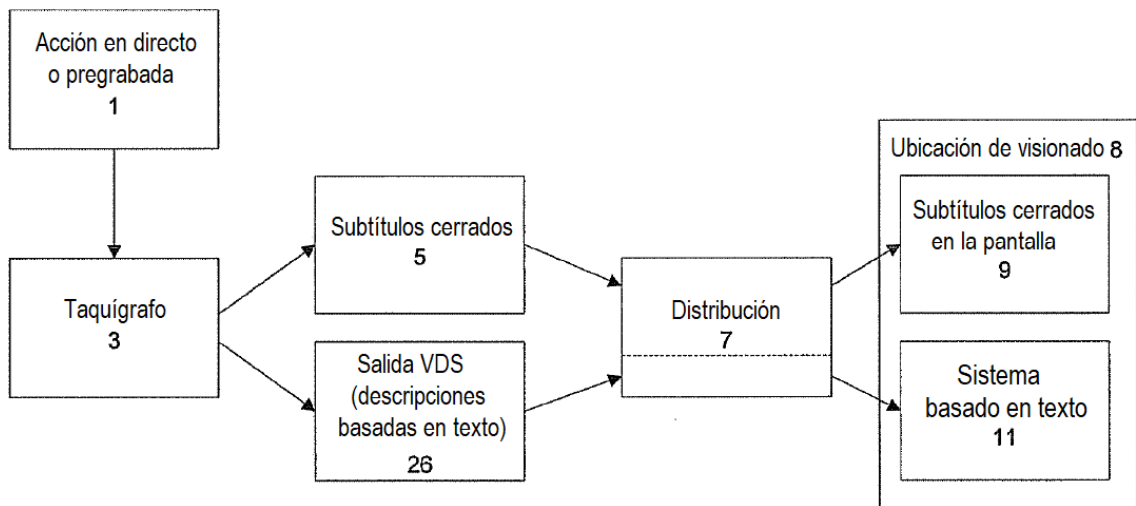


FIG. 5

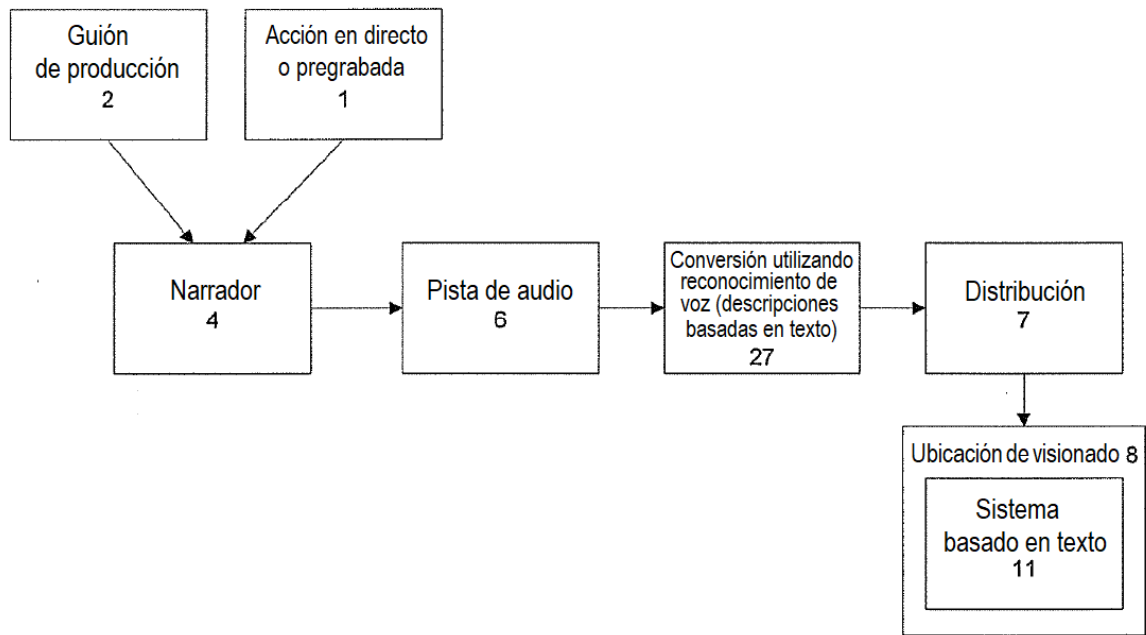


FIG. 6

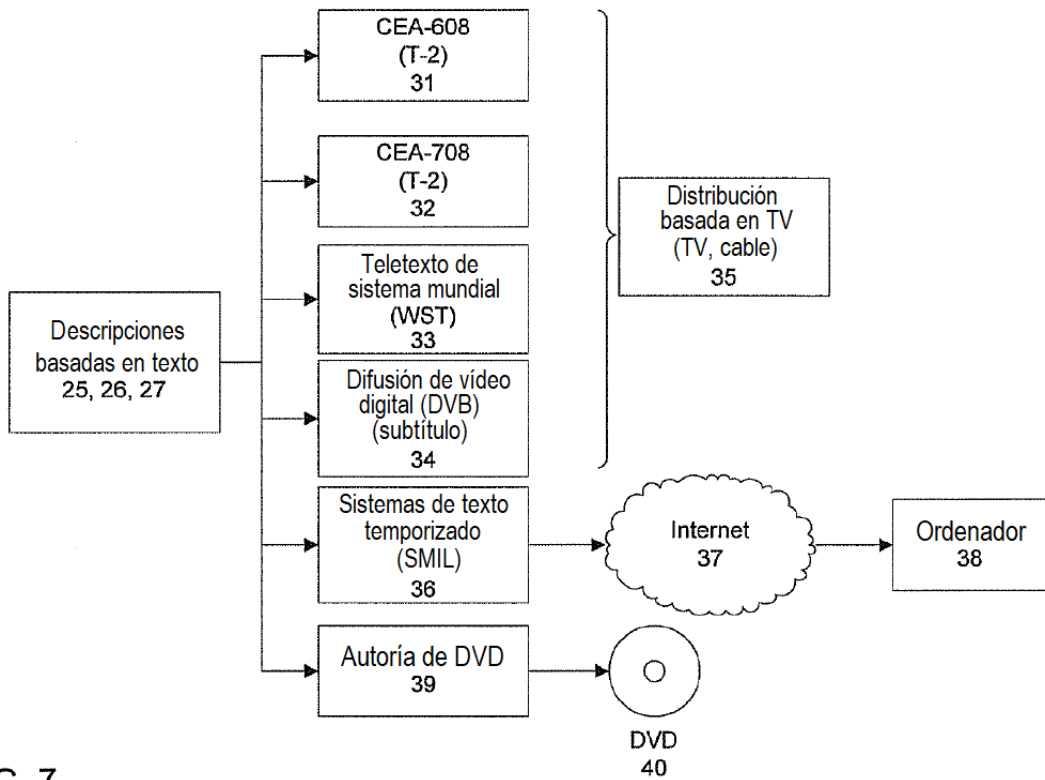


FIG. 7

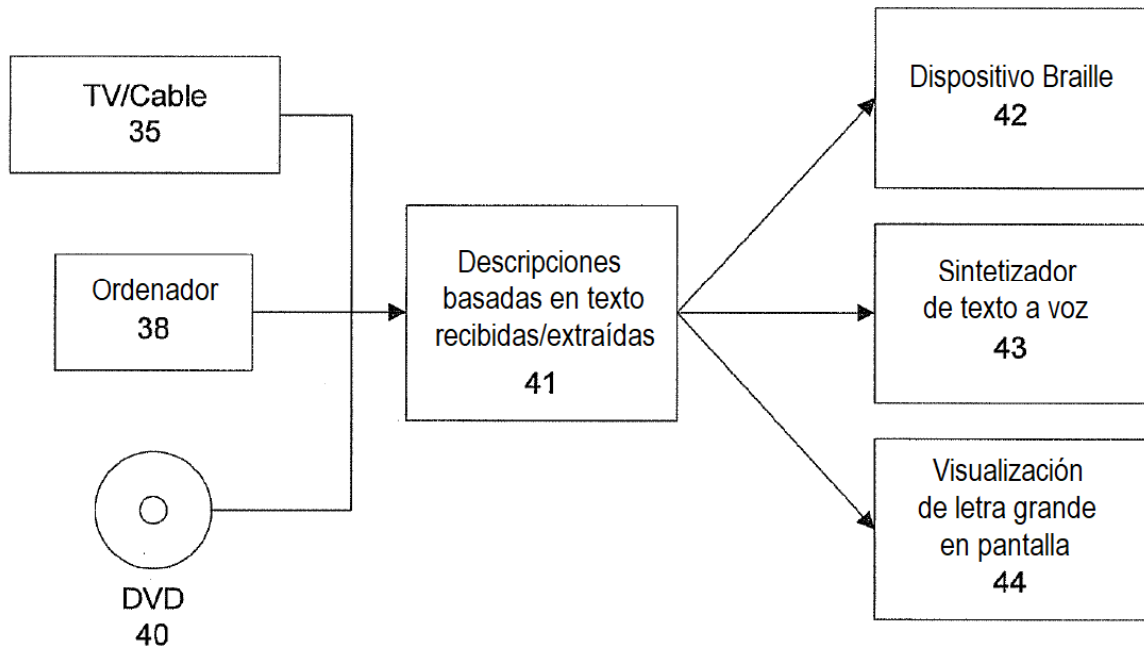


FIG. 8