

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 788 872**

51 Int. Cl.:

G06F 3/01 (2006.01)

G06F 3/023 (2006.01)

G06F 3/0484 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.09.2015 PCT/US2015/048154**

87 Fecha y número de publicación internacional: **10.03.2016 WO16036862**

96 Fecha de presentación y número de la solicitud europea: **02.09.2015 E 15763752 (1)**

97 Fecha y número de publicación de la concesión europea: **18.03.2020 EP 3189398**

54 Título: **Sistemas y métodos de entrada de texto basados en mirada**

30 Prioridad:

02.09.2014 US 201462044817 P

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

23.10.2020

73 Titular/es:

**TOBII AB (100.0%)
Box 743
182 17 Danderyd, SE**

72 Inventor/es:

**KRISTENSSON, PER OLA;
VERTANEN, KEITH y
MJELDE, MORTEN**

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

ES 2 788 872 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

Sistemas y métodos de entrada de texto basados en mirada

5 La presente invención se refiere de manera general a sistemas y a métodos para entrada de texto basada en mirada, y en particular, a sistemas y a métodos para usar la mirada para proporcionar una entrada de texto basada en la mirada del usuario que indica un deseo de seleccionar una letra.

Antecedentes de la invención

10 Proporcionar entrada de texto a un dispositivo informático es un requisito clave para una interacción eficiente entre ordenadores y personas. Tradicionalmente se proporciona entrada de texto por medio de un teclado, o en el caso de dispositivos informáticos que usan pantallas táctiles puede usarse un teclado en pantalla. Todos estos planteamientos requieren contacto físico por parte de un usuario sobre un dispositivo.

15 En algunos casos es deseable proporcionar un medio para introducir texto en un dispositivo informático mediante el cual no se requiera contacto físico con el dispositivo u otros medios de introducción. Esto puede usarse, por ejemplo, cuando un usuario no posee control de sus manos o cuando un usuario no puede tener contacto con un dispositivo porque está usando sus manos para otro fin. Una solución posible es usar software de reconocimiento de voz que permite que un usuario comunique de manera audible sus entradas.

20 Los problemas con este enfoque incluyen que puede ser demasiado lento como para resultar conveniente, puede requerir entrenamiento extenso de un algoritmo de reconocimiento de voz antes de usar el mismo, o puede no estar disponible para un usuario que también tiene un trastorno de habla, hasta e incluyendo mudéz, que impide que tal software funcione apropiadamente o sea en absoluto una opción. Por tanto, un software de reconocimiento de voz puede no proporcionar un método útil y/o fiable para proporcionar una entrada de texto a un dispositivo informático.

25 El documento US 2014/0002341 A1 da a conocer un método para introducir texto en un dispositivo informático por medio de mecanografiado ocular. Por ejemplo, una realización proporciona un método que incluye recibir un conjunto de datos que incluye una pluralidad de muestras de mirada, incluyendo cada muestra de mirada una ubicación de mirada y un instante correspondiente en el tiempo. El método comprende además procesar la pluralidad de muestras de mirada para determinar uno o más términos probables representados por el conjunto de datos.

30 Aunque están disponibles métodos básicos para usar una entrada de texto basada en mirada, pueden presentar lentitud o inexactitud. La presente invención proporciona soluciones para al menos algunos de los problemas anteriormente mencionados.

Breve descripción de la invención

40 En una realización, se proporciona un método para introducir texto en un dispositivo informático usando entrada de mirada de un usuario según la reivindicación 1.

45 En otra realización, se proporciona un sistema para introducir texto en un dispositivo informático usando entrada de mirada de un usuario según la reivindicación 10.

50 En otra realización, se proporciona un medio legible por máquina no transitorio que tiene instrucciones almacenadas en el mismo para introducir texto en un dispositivo informático usando entrada de mirada de un usuario según la reivindicación 14.

55 En la siguiente descripción, se considera que las realizaciones y/o ejemplos que no están cubiertos por las reivindicaciones adjuntas no forman parte de la presente invención.

Breve descripción de los dibujos

60 La presente invención se describe junto con las figuras adjuntas:

65 la figura 1 es una representación de un entorno en el que pueden implementarse sistemas y métodos de la invención;

la figura 2 es una pantalla de visualización de diversas realizaciones de la invención para introducir texto en un dispositivo informático por medio de una entrada de mirada de un usuario;

la figura 3 es otra pantalla de visualización de diversas realizaciones de la invención para corregir texto introducido por medio de la pantalla de visualización de la figura 2;

la figura 4 es un diagrama de flujo de un método de la invención para introducir texto en un dispositivo informático por medio de una entrada de mirada de un usuario;

5 la figura 5 es un diagrama de flujo de un método de la invención para determinar si proporcionar una pantalla de visualización para corregir texto anteriormente introducido;

la figura 6 es un diagrama de flujo de un método de la invención para corregir texto anteriormente introducido;

10 la figura 7 es un diagrama de bloques de un sistema informático a modo de ejemplo que puede usarse en al menos alguna parte de los aparatos o los sistemas de la presente invención, o que implementa al menos alguna parte de los métodos de la presente invención;

15 la figura 8 es un diagrama de flujo de un método de la invención para introducir texto en un dispositivo informático por medio de múltiples tipos de entradas de mirada de un usuario, de tal manera que los tipos indican un nivel diferente de importancia de entrada; y

la figura 9 es un diagrama de flujo de un método de la invención para determinar si debe añadirse texto anteriormente introducido a un modelo de lenguaje para aumentar la exactitud de texto introducido adicional.

20 En las figuras adjuntas, componentes y/o características similares pueden tener la misma etiqueta de referencia numérica. Además, diversos componentes del mismo tipo pueden distinguirse siguiendo la etiqueta de referencia por una letra que distingue entre los componentes y/o las características similares. Si sólo se usa la primera etiqueta de referencia numérica en la memoria descriptiva, la descripción puede aplicarse a uno cualquiera de los componentes y/o las características similares que tienen la misma primera etiqueta de referencia numérica independientemente del sufijo de letra.

Descripción detallada de la invención

30 La siguiente descripción sólo proporciona realizaciones a modo de ejemplo, y no se pretende que limite el alcance, la aplicabilidad o la configuración de la divulgación. En vez de eso, la siguiente descripción de las realizaciones a modo de ejemplo proporcionará a los expertos en la técnica una descripción reproducible para implementar una o más realizaciones a modo de ejemplo. Se entiende que pueden realizarse diversos cambios en la función y la disposición de elementos sin apartarse del alcance de la invención tal como se expone en las reivindicaciones adjuntas.

35 Por ejemplo, cualquier detalle comentado con respecto a una realización puede o puede no estar presente en todas las versiones contempladas de esa realización. De manera similar, cualquier detalle comentado con respecto a una realización puede o puede no estar presente en todas las versiones contempladas de otras realizaciones comentadas en el presente documento. Finalmente, la ausencia de comentarios de cualquier detalle con respecto a una realización en el presente documento será un reconocimiento implícito de que tal detalle puede o puede no estar presente en cualquier versión de cualquier realización comentada en el presente documento.

40 Se dan detalles específicos en la siguiente descripción para proporcionar un entendimiento riguroso de las realizaciones. Sin embargo, un experto habitual en la técnica entenderá que las realizaciones pueden ponerse en práctica sin estos detalles específicos. Por ejemplo, circuitos, sistemas, redes, procesos, y otros elementos en la invención pueden mostrarse como componentes en forma de diagrama de bloques con el fin de no complicar las realizaciones con detalles innecesarios. En otros casos, circuitos, procesos, algoritmos, estructuras y técnicas bien conocidos pueden mostrarse sin detalles innecesarios con el fin de evitar complicar las realizaciones.

45 Además, se observa que pueden describirse realizaciones individuales como un proceso que se representa como organigrama, un diagrama de flujo, un diagrama de flujo de datos, un diagrama de estructura o un diagrama de bloques. Aunque un organigrama puede describir las operaciones como un proceso secuencial, muchas de las operaciones pueden realizarse en paralelo o de manera simultánea. Además, el orden de las operaciones puede reorganizarse. Un proceso puede terminarse cuando se completan sus operaciones, pero puede tener etapas adicionales no comentadas o incluidas en una figura. Además, puede que no todas las operaciones en cualquier proceso descrito particularmente se produzcan en todas las realizaciones. Un proceso puede corresponder a un método, una función, un procedimiento, una subrutina, un subprograma, etc. Cuando un proceso corresponde a una función, su terminación corresponde a un retorno de la función a la función de llamada o la función principal.

50 El término "medio legible por máquina" incluye, pero no se limita a, dispositivos de almacenamiento transitorios y no transitorios, portátiles o fijos, dispositivos de almacenamiento ópticos, canales inalámbricos y varios otros medios que pueden almacenar, contener o portar instrucción/instrucciones y/o datos. Un segmento de código o instrucciones que pueden ejecutarse por máquina pueden representar un procedimiento, una función, un subprograma, un programa, una rutina, una subrutina, un módulo, un paquete de software, una clase o cualquier combinación de instrucciones, estructuras de datos o sentencias de programa. Puede acoplarse un segmento de código a otro segmento de código o un circuito de hardware pasando y/o recibiendo información, datos, argumentos,

parámetros o contenido de memoria. Puede pasarse, reenviarse o transmitirse información, argumentos, parámetros, datos, etc. por medio de cualquier medio adecuado incluyendo uso compartido de memoria, paso de mensajes, paso de testigo, transmisión en red, etc.

5 Además, pueden implementarse realizaciones de la invención, al menos en parte, o bien manual o bien automáticamente. Pueden ejecutarse, o al menos asistirse, implementaciones manuales o automáticas por medio del uso de máquinas, hardware, software, firmware, software intermedio, microcódigo, lenguajes de descripción de hardware o cualquier combinación de los mismos. Cuando se implementa en software, firmware, software intermedio o microcódigo, el código de programa o segmentos de código para realizar las tareas necesarias pueden almacenarse en un medio legible por máquina. Un(os) procesador(es) puede(n) realizar las tareas necesarias.

10 Pasando ahora a la figura 1, se muestra un sistema 100 que puede implementar realizaciones de la invención. El sistema 100 puede incluir un dispositivo 110 informático, un dispositivo 120 de visualización, dispositivos 130 de entrada tradicionales (teclado 130a y ratón 130b) y un dispositivo 140 de seguimiento ocular.

15 El dispositivo 110 informático, aunque se muestra en esta realización como un ordenador de escritorio, puede ser cualquier tipo de dispositivo informático en otras realizaciones. Meramente a modo de ejemplo, el dispositivo 110 informático puede ser un ordenador portátil, un miniordenador portátil, un ordenador de tipo tableta, un asistente de datos personal, un dispositivo móvil, un teléfono móvil, un reloj inteligente u otro dispositivo electrónico ponible, etc. En algunas realizaciones, varios otros componentes del sistema 100 pueden estar integrados con el dispositivo 110 informático. Por ejemplo, un ordenador de tipo tableta puede tener un dispositivo 120 de visualización integrado, así como una versión de pantalla táctil del teclado 130a.

20 El dispositivo 120 de visualización puede ser capaz de visualizar información tal como se indica por el dispositivo 110 informático. Tal como se comentó anteriormente, el dispositivo 120 de visualización puede estar integrado con el, o ser independiente del, dispositivo 110 informático. El dispositivo 140 de seguimiento ocular puede ser cualquier forma de dispositivo de seguimiento ocular conocido que permite la determinación de la dirección de mirada de un usuario. El dispositivo 140 de seguimiento ocular puede usar el dispositivo 110 informático para procesar información para determinar la dirección de mirada de un usuario, o el dispositivo 140 de seguimiento ocular puede realizar todo el procesamiento directamente.

25 Un ejemplo de un dispositivo 140 de seguimiento ocular adecuado es un dispositivo basado en infrarrojos en el que se emite luz infrarroja hacia el ojo o los ojos de un usuario y un sensor de imagen captura imágenes del ojo o los ojos del usuario que comprenden una reflexión de la luz infrarroja. El procesamiento de estas imágenes puede determinar la dirección de mirada de un usuario a partir de la posición y la orientación de la reflexión o las reflexiones de luz infrarroja. Un experto en la técnica entenderá esto fácilmente. El dispositivo 140 de seguimiento ocular puede estar configurado para seguir la dirección de mirada de un usuario con respecto al dispositivo 120 de visualización.

30 Pasando ahora a la figura 2, en una realización de la invención, el dispositivo 110 informático puede hacer que el dispositivo 120 de visualización visualice una interfaz 200 de introducción de texto para un usuario. La interfaz 200 puede incluir una representación de letras en el elemento de visualización, por ejemplo, un teclado 210 en pantalla, una zona 220 de terminación de entrada y una ventana 230 de mensaje. En términos amplios los sistemas y métodos de la invención pueden funcionar permitiendo que el usuario mire a o cerca de las letras visualizadas en el teclado 210 en pantalla en el orden en el que el usuario desea deletrear una palabra o palabras. En este ejemplo, el usuario ha mirado recientemente en un patrón a las teclas de tal manera que "This is a test." ("Esto es una prueba") se ha introducido previamente, y "How are you?" ("¿Qué tal estás?") se ha introducido más recientemente. El texto introducido recientemente puede anotarse en cualquiera de varias maneras, quizá destacándolo, y se muestra en este caso como recuadros 240 discontinuos.

35 Para terminar la entrada de una palabra o palabras, que a continuación en el presente documento se entiende que incluye números, signos de puntuación y cualquier otra entrada textual, entonces el usuario puede mirar a la zona 220 de terminación, indicando de ese modo al dispositivo 110 informático que el usuario desea terminar la entrada de texto (al menos de momento). A continuación se describirán detalladamente alternativas adicionales a que el usuario mire a la zona 220 de terminación como medio para terminar la entrada. A medida que el usuario mira a las letras, o después de que el usuario haya mirado a la zona 220 de terminación, la palabra o las palabras deletreadas por el usuario pueden visualizarse en la ventana 230 de mensaje.

40 La interfaz 200 también puede incluir un icono 240 de ajustes, un icono 250 de entrada de reconocimiento de voz, un icono 260 de retroceso, un icono 270 de caracteres alternativos, un icono 280 de confirmación y una visualización del punto 290 de mirada actual del usuario. Cada uno de los iconos puede activarse por la persistencia de la mirada del usuario en el mismo para provocar un efecto determinado. El icono 240 de ajustes puede hacer que se presente una interfaz de ajustes para permitir que el usuario ajuste el funcionamiento de realizaciones de la invención. El icono 250 de entrada de reconocimiento de voz puede permitir que un usuario active un paquete de software de reconocimiento de voz para la entrada de texto. El icono 260 de retroceso puede permitir que un usuario borre la última letra o palabra introducida. El icono 270 de caracteres alternativos puede permitir que el usuario cambie los

caracteres visualizados por el teclado 210 en pantalla. El icono 280 de confirmación puede permitir que un usuario confirme que las palabras introducidas recientemente (destacadas por medio de recuadros 240 discontinuos) son en efecto las que pretendía introducir el usuario. La visualización del punto 290 de mirada actual del usuario, aunque en este caso se muestra como una X discontinua, puede presentarse de otras maneras dependiendo de la realización, y puede visualizarse siempre o no, quizá dependiendo de cuánto tiempo ha permanecido la mirada del usuario en un punto o zona particular de la interfaz 200.

En la figura 3 se muestra una segunda interfaz 300 de diversas realizaciones de la invención. La interfaz 300 de corrección de error puede activarse dirigiendo la mirada del usuario a una palabra o palabras en la ventana 230 de mensaje. Una vez que el dispositivo 110 informático determina que se detecta un acontecimiento de mirada de este tipo, puede visualizarse la interfaz 300 para permitir que el usuario corrija la introducción de texto o bien introducida de manera errónea por el usuario o bien interpretada de manera errónea por el dispositivo 110 informático.

La interfaz 300 puede incluir una palabra o palabras 310 que estaban en la ubicación, o alrededor de la ubicación, en la que el usuario miraba a las mismas en la ventana 230 de mensaje de la interfaz 200. Pueden visualizarse signos 320 "+" entre las palabras 310 para proporcionar al usuario un lugar para mirar para indicar al dispositivo 110 informático que desea introducir palabras adicionales en esa ubicación. Debajo de una o más de las palabras 310, pueden visualizarse posibles palabras 330 de sustitución, tal como se presentan como hipótesis de manera inteligente por el dispositivo 110 informático, de manera que el usuario puede seleccionarlas por medio de su mirada para sustituir las palabras 310 originales. Un icono 340 de borrado (simbolizado por una "X" u otro símbolo) puede permitir que un usuario borre una palabra 310 dada.

También puede proporcionarse un icono 350 de cancelación de manera que el usuario puede cancelar, con su mirada, la entrada de cualquier corrección y volver a la interfaz 200. Un icono 360 de aceptación puede permitir que un usuario acepte todos los cambios realizados y vuelva a la interfaz 200.

Pasando ahora a las figuras 4-6, se comentará un método 400, 500, 600 a modo de ejemplo de la invención que emplea el sistema 100 y las interfaces 200, 300. Tras la visualización de la interfaz 200, el método 400, en el bloque 410, puede recibir datos de mirada continuos asociados con la mirada de un usuario tal como se proporcionan por el dispositivo 110 informático y el dispositivo 140 de seguimiento ocular. En el bloque 420, puede visualizarse el punto de mirada al igual que se visualiza el punto 290 de mirada en la figura 2. En el bloque 430, pueden visualizarse una o más "teclas" (letra o símbolo en el teclado 210 en pantalla) para indicar al usuario que el dispositivo 110 informático determina que la mirada del usuario está ubicada en esa posición.

En el bloque 440, se monitoriza la ubicación de la mirada del usuario en el dispositivo 120 de visualización para determinar la aparición de un acontecimiento de observación. Un acontecimiento de observación se produce cuando el dispositivo 110 informático (y/o el dispositivo 140 de seguimiento ocular, que se supondrá de aquí en adelante cuando se haga referencia al dispositivo 110 informático) determina que la mirada del usuario se ha fijado en o cerca de un carácter en el teclado 210 en pantalla. Un acontecimiento de observación puede producirse basándose en una cualquiera o más de las siguientes características: la fijación de la mirada del usuario en o cerca de un carácter o caracteres; una aproximación de la región de la fijación, las coordenadas x, y de una fijación; la duración de la fijación; el tamaño de la fijación; otro sistema de coordenadas basado en mediciones de píxeles o distancia estándar o métrica, etc.; otra característica; cualquier combinación de los anteriores.

En el bloque 450 se graba el acontecimiento de observación y se combina de una manera en serie con acontecimientos de observación grabados anteriormente, incluyendo posiblemente cada acontecimiento almacenado cualquiera de las características anteriores o más asociadas con el acontecimiento de observación. Siempre y cuando no se detecte ningún acontecimiento de terminación en el bloque 460, este proceso se repetirá de tal manera que se graban continuamente acontecimientos de observación detectados hasta que se produce un acontecimiento de terminación finalizando la secuencia de observación.

Se produce un acontecimiento de terminación en el bloque 460 basado en uno cualquiera o más acontecimientos posibles, dependiendo de la realización. Algunos ejemplos, no a modo de limitación, incluyen: la detección de la mirada de un usuario dentro de la zona 220 de terminación durante un periodo de término predeterminado; que el usuario termine implícitamente al transcurrir el tiempo (la mirada del usuario ha permanecido estática en cualquier lugar en la visualización durante una cantidad de tiempo fija); que el usuario mire hacia otro lado mirando fuera de la visualización; que el usuario parpadee, o bien un único parpadeo o bien dos o más parpadeos; que el usuario realice un gesto facial tal como sonreír u otra señal afectiva; que el usuario articule un gesto de mirada específico (por ejemplo, una combinación de movimientos de mirada) con el fin predeterminado/predefinido de provocar la terminación; que el dispositivo 110 informático deduzca una terminación, por ejemplo, determinando que la articulación de mirada del usuario ha terminado calculando continuamente la probabilidad de este acontecimiento después de haberse recibido cada observación de mirada (es decir, realizando una determinación de que se ha producido una probabilidad de terminación); la activación de un dispositivo de entrada táctil tal como un teclado, ratón, pantalla táctil, etc.; un sonido verbal u otro realizado por el usuario; un gesto manual, un gesto de cabeza o gesto con otra parte del cuerpo u objeto; una señal cerebral detectada por medio de un electroencefalograma o similares.

Tras la terminación en el bloque 460, el método 400 continúa hasta el bloque 470 en el que un módulo/algoritmo de decodificación presenta como hipótesis una cadena (frase, oración, etc.) a partir de la secuencia de observación. Tal como se comentará a continuación, el módulo de decodificación puede presentar como hipótesis la cadena usando varios métodos, incluyendo un decodificador de paso de testigo. Tras determinar una o más hipótesis, pueden visualizarse una o más cadenas que se correlacionan en la ventana 230 de mensaje en el bloque 480. En el caso de dos hipótesis que tienen una probabilidad razonable de reflejar la cadena pretendida por el usuario, pueden visualizarse ambas cadenas para la selección por el usuario (mediante fijación de mirada u otro mecanismo de entrada).

Después, puede recibirse una entrada de texto de mirada adicional según el método 400, pero en el bloque 510, a medida que se reciben nuevamente datos de mirada, y se visualizan en los bloques 520, 530 (similares a los bloques 420, 430), se espera una confirmación de que la cadena visualizada es indicativa de la cadena deseada del usuario en el bloque 540. Esto puede producirse fijando el usuario su mirada en el icono 280 de confirmación. En ese caso, el método vuelve al bloque 410 para esperar una introducción de texto adicional o la terminación de la interfaz de introducción de texto. Sin embargo, si el usuario fija su mirada en una palabra, letra u otro símbolo particular en la ventana 230 de mensaje, entonces se visualiza en el bloque 550 la interfaz 300 de corrección de error.

En la interfaz 300 de corrección de error, la palabra, carácter u otro símbolo en el que se fija el usuario en la interfaz 200 se visualiza junto con palabras circundantes de la cadena. Otras palabras que el módulo de decodificación puede presentar como hipótesis como otras palabras, caracteres o símbolos probables pretendidos como entrada correcta por el usuario pueden visualizarse debajo de las palabras, caracteres o símbolos presentados como hipótesis original visualizados. Tras la recepción de datos de mirada en el bloque 610, visualizados en los bloques 620, 630 (similares a los bloques 420, 430), se espera un acontecimiento de fijación en el bloque 640. Un usuario puede fijarse durante un período predeterminado en cualquier opción de palabra con el fin de seleccionar las palabras más deseadas. En el bloque 650, tales palabras/caracteres/símbolos se cambian tras la producirse el acontecimiento. Si el usuario desea borrar una palabra, carácter o símbolo, el usuario se fija en el icono 340 de borrado asociado con (debajo de) la palabra, carácter o símbolo que se desea borrar. Si el usuario desea insertar una nueva palabra, el usuario se fija en el icono de inserción. El usuario puede añadir palabras usando los signos 320 "+". La fijación de la mirada del usuario en los signos 320 "+" puede hacer que se proporcione una interfaz igual que o similar a la interfaz 200 para la introducción de texto. Alternativamente o de manera simultánea, puede usarse un dispositivo de entrada alternativo tal como un ratón, un teclado, etc., para introducir o corregir palabras siempre que se describa la adición y/o corrección en el presente documento. Estas correcciones pueden continuar produciéndose hasta que se produce un tipo diferente de acontecimiento de fijación en el bloque 660.

En el bloque 660, se espera una fijación de la mirada del usuario en el icono 350 de cancelación o el icono 360 de aceptación. Si no se produce una fijación de este tipo, entonces el método 600 se repite hasta que el usuario ha realizado todas las correcciones deseadas por el usuario. Si se produce la fijación en el icono 350 de cancelación, entonces el método vuelve al bloque 410 sin corrección de la cadena. Si se produce la fijación en el icono 360 de aceptación, entonces el método vuelve al bloque 410 habiéndose completado las correcciones realizadas por el usuario en la cadena y visualizándose en la ventana 230 de mensaje.

Centrándose ahora en el módulo de decodificador, tal módulo es un programa de software que toma como entrada una secuencia de observación compuesta de al menos un acontecimiento de observación. El decodificador emite una serie de texto, que forma lo más habitualmente una oración o similar, que representa una estimación de una serie de texto pretendida por el usuario.

En una realización de la presente invención el decodificador se implementa como decodificador de paso de testigo mediante lo cual un testigo comprende tres elementos de información: una probabilidad en un espacio logarítmico acumulada hasta ahora (prob. log.), un contexto de modelo de lenguaje de palabras y caracteres, y la posición del testigo en la secuencia de observación. Los expertos en la técnica estarán familiarizados con tales decodificadores de paso de testigo y el funcionamiento de los mismos con respecto a las realizaciones de la invención también resultará ahora evidente. Puede encontrarse información adicional con respecto a decodificadores de paso de testigo en Young, S., Russel, N.H. y Thornton, J.H.S. 1989, "Passing: A Simple Conceptual Model for Connected Speech Recognition Systems", informe técnico, departamento de ingeniería de la Universidad de Cambridge.

Cada testigo representa una hipótesis: una secuencia de letras que puede ser el texto pretendido del usuario. Un testigo se termina cuando el testigo se ha sometido a todas las observaciones en la secuencia de observación. El decodificador emitirá, por medio del conjunto de testigos terminados, un conjunto de hipótesis. La hipótesis más probable corresponde a la secuencia de letras que es más probable que correspondan al texto pretendido del usuario.

Para calcular una probabilidad pueden usarse dos modelos, el primero es un modelo de teclado, que calcula la probabilidad de una observación dada una letra presentada como hipótesis. El modelo de teclado asigna una o varias distribuciones, por ejemplo, distribuciones gaussianas bidimensionales, distribuciones no paramétricas, procesos gaussianos en cada tecla. La distribución de modelo de teclado se usa para calcular la probabilidad dadas

las ubicaciones de X e Y en la observación. El sistema actualmente implementado asigna distribuciones gaussianas bidimensionales en cada tecla.

5 Un diseño de sistema alternativo (no actualmente implementado) puede estimar probabilidades usando una regla de decisiones, un árbol de decisiones, una variante de un bosque aleatorio, una máquina de vectores de soporte u otra técnica de aprendizaje automático.

10 El segundo modelo posible es un modelo de lenguaje de caracteres y palabras. El modelo de lenguaje de caracteres calcula la probabilidad anterior de una letra presentada como hipótesis dadas letras anteriores en la hipótesis actual (tal como se sigue por un testigo en la búsqueda). El modelo de lenguaje de palabras se usa para calcular la probabilidad anterior de una palabra presentada como hipótesis dadas palabras anteriores.

15 El decodificador usa un algoritmo de decodificación para propagar testigos en cada etapa a lo largo de la formación de la hipótesis. Con el fin de restringir la búsqueda para mantener un tamaño gestionable, pueden fusionarse testigos redundantes cuando dos o más testigos han alcanzado la misma conclusión (posiblemente provisional). Otra técnica posible que puede implementarse por el módulo de decodificación se conoce como eliminación de haces, que restringe la búsqueda usando una anchura de haz configurable cuando la hipótesis de testigo no es lo suficientemente alta como para justificar mantenerla.

20 La regla de propagación de testigos usada por el algoritmo de decodificación puede incluir:

25 [A] (SUSTITUCIÓN) Dada la observación actual: calcular la probabilidad de esa observación para cada tecla usando el modelo de teclado; el testigo antiguo crea un nuevo testigo para cada tecla en el teclado; la probabilidad para el nuevo testigo es: la prob. log. de la observación dada la tecla específica + la probabilidad del modelo de lenguaje de caracteres (ajustada a escala mediante un factor de escala) + la prob. log. acumulada hasta ahora por el testigo antiguo; y opcionalmente, si el nuevo testigo está generando un separador de palabras, por ejemplo un carácter de barra espaciadora, entonces añadir adicionalmente una prob. log. basada en el modelo de lenguaje de palabras. Si la secuencia de letras no se modela por el vocabulario del modelo de lenguaje de palabras entonces evaluar la secuencia de letras usando un modelo de palabras desconocidas.

30 [B] (BORRADO) Dada la observación actual: ignorar la observación y el testigo se propaga a la siguiente observación en la secuencia de observación y acumula una penalización de prob. log. fijada.

35 [C] (INSERCIÓN) Dada la observación actual: propagar testigos a todas las teclas posibles en el teclado; la probabilidad de un nuevo testigo es: la prob. log. acumulada hasta ahora por el testigo antigua + una prob. log. nueva basada en el modelo de lenguaje de caracteres (ajustada a escala mediante un factor de escala) + una penalización de prob. log. adicional que es diferente dependiendo de si se está insertando la misma tecla de teclado que antes o no; opcionalmente, si el nuevo testigo está generando un separador de palabras, por ejemplo un carácter de barra espaciadora, añadir entonces adicionalmente una prob. log. basada en el modelo de lenguaje de palabras. Si la secuencia de letras no se modela por el vocabulario del modelo de lenguaje de palabras entonces evaluar la secuencia de letras usando un modelo de palabras desconocidas.

45 [D] (ELIMINACIÓN) Después de los casos de sustitución, inserción y borrado, antes de propagar un testigo, se aplica una eliminación de haces de la siguiente manera: para cada índice de observación en la secuencia de observación, el sistema almacena la mejor prob. log. generada por cualquier testigo hasta ahora para ese índice de observación; un nuevo testigo sólo se propaga a un índice de observación nuevo si la diferencia entre su prob. log. generada y la mejor prob. log. generada por un testigo para ese índice de observación está dentro de un umbral configurable (es decir la anchura de haz). El umbral es un compromiso entre velocidad-exactitud.

50 [E] (TERMINACIÓN) Todos los testigos no eliminados se comprueban para ver si están en el último índice de observación. Si lo están: tomar la prob. log. existente del testigo; añadir una prob. log. del modelo de lenguaje de caracteres que modela el símbolo de fin de oración (ajustada a escala mediante un factor de escala); opcionalmente, añadir una prob. log. basada en el modelo de lenguaje de palabras. Si la secuencia de letras no se modela por el vocabulario del modelo de lenguaje de palabras entonces evaluar la secuencia de letras usando un modelo de palabras desconocidas.

[F] (FUSIÓN) Se fusionan testigos con contexto de moldeo de lenguaje idéntico.

60 La salida del algoritmo de decodificación es la secuencia de letras modelada por la hipótesis del testigo terminado con la mayor probabilidad. Estrategias de decodificación alternativas también pueden ser adecuadas para la presente invención, tales como:

65 [1] Coincidencia basada en cadenas: añadir todas las letras que se identifican por medio de mirada, por ejemplo, fijándose sobre las mismas, en una memoria intermedia. Aplicar una regla basada en cadena que comprende palabras válidas, retirar cualquier letra que es improbable que forme palabras válidas. Por tanto, las letras restantes representan el texto pretendido del usuario que comprende una o más palabras.

[2] Coincidencia de trayectoria espacial: definir palabras como trayectorias espaciales mapeándolas en un teclado visualizado. Conectar todas las fijaciones u otra métrica basada en mirada de manera conjunta en trayectorias espaciales que abarcan una o más palabras. Calcular coincidencias realizando comparaciones de trayectorias espaciales en combinación con un método de segmentación de trayectorias espaciales que abarca posiblemente muchas palabras a trayectorias espaciales que definen palabras individuales. Puede encontrarse información adicional en Kristensson, P.O. y Zhai, S. 2004, "SHARK2: A Large Vocabulary Shorthand Writing System for Pen-Based Computers", en Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST 2004), ACM Press: 43-52.

En algunas realizaciones de la invención, los siguientes aspectos también pueden presentar las siguientes mejoras: cuando un testigo ha formado posiblemente una palabra, puede consultarse un modelo de lenguaje basado en palabras con el fin de determinar la probabilidad de que la palabra posible sea correcta; y opcionalmente el sistema puede generar una cuadrícula que representa el espacio que se ha explorado por los testigos anteriores. La cuadrícula puede volverse a puntuar mediante lo cual las posiciones de testigos anteriores pueden volver a calcularse basándose en un modelo de lenguaje.

El sistema también puede generar redes de confusión de palabras (WCN). Las mismas pueden o bien estimarse basándose en una N mejor búsqueda o bien, si está disponible una cuadrícula, crearse a partir de la cuadrícula. Las WCN son ventajosas dado que es más fácil trabajar con las mismas que con una cuadrícula y son particularmente útiles para interfaces de usuario de corrección de errores. Véase Hakkani-Tür, D., Béchet, F., Riccardi, G. y Tur, G. "Beyond ASR 1-best: Using Word Confusion Networks in Spoken Language Understanding. J. Computer Speech and Language 20, 4 (2006), 495-514.; Ogata, J. y Goto, M. "Speech Repair: Quick Error Correction Just by Using Selection Operation for Speech Input Interfaces" Proc. ICSLP 2005, ISCA (2005), 133-136; Vertanen, K. y Kristensson, P.O. 2009. "Parakeet: A Continuous Speech Recognition System for Mobile Touch-Screen Devices" en Proceedings of the 14th ACM International Conference on Intelligent User Interfaces (IUI 2009). ACM Press: 237-246.

Como alternativa a las distribuciones gaussianas bidimensionales como modelos para ubicaciones de fijación, pueden usarse procesos gaussianos (GP), véase Rasmussen, CE. y Williams, C.K.I. 2006. "Gaussian Processes for Machine Learning" MIT Press. Los GP tienen una mejor capacidad para aprender la forma de distribuciones sobre las teclas, lo que dará como resultado estimaciones de probabilidad más exactas.

El decodificador puede modelar duraciones de fijaciones para mejorar la exactitud. Las duraciones pueden modelarse usando una función de etapa en la que las duraciones de fijación por debajo de un umbral establecido se penalizan por una función lineal. Alternativamente, las duraciones pueden modelarse por una función de parámetro bajo tal como una función logística, que puede parametrizarse usando sólo dos variables. Alternativamente, la duración puede modelarse usando una distribución paramétrica (tal como una de Poisson) o una distribución no paramétrica.

El sistema puede comprender además una función de aprendizaje adaptativa mediante la cual los resultados del algoritmo de decodificación se almacenan y usan en un algoritmo de aprendizaje para adaptar modelos de lenguaje a usuarios individuales. A continuación se describe un ejemplo de una realización de este tipo, con respecto a la figura 9. Opcionalmente, si el sistema usa procesos gaussianos para el modelo de teclado entonces los parámetros pueden aprenderse a partir de datos tanto sin conexión como en línea usando una estrategia similar tal como se expuso anteriormente. Pueden encontrarse enfoques algorítmicos para aprender los parámetros de funciones de covarianza, por ejemplo, en: Rasmussen, C. E., y Williams, C. K. I. "Gaussian Processes for Machine Learning" MIT Press, 2006. Referencia adicional: Weir, D., Rogers, S., Murray-Smith, R., y Löchtfeld, M. "A User-Specific Machine Learning Approach for Improving Touch Accuracy on Mobile Devices" Proc. UIST '12. 465-476.

A modo de ejemplo y no de limitación, la figura 7 es un diagrama de bloques que representa un dispositivo 702 informático de ejemplo para implementar algunas realizaciones comentadas en el presente documento. El dispositivo 702 informático puede incluir un procesador 704 que está acoplado de manera comunicativa a una memoria 706 y que ejecuta instrucciones de programa ejecutables por ordenador y/o accede a información almacenada en la memoria 706. El procesador 704 puede comprender un microprocesador, un circuito integrado específico de aplicación ("ASIC"), una máquina de estados u otro dispositivo de procesamiento. El procesador 704 puede incluir cualquiera de varios dispositivos de procesamiento informáticos, incluyendo uno. Un procesador de este tipo puede incluir o puede estar en comunicación con un medio legible por ordenador que almacena instrucciones que, cuando se ejecutan por el procesador 704, hacen que el procesador realice las etapas descritas en el presente documento.

El dispositivo 702 informático también puede incluir un bus 708. El bus 708 puede acoplar de manera comunicativa uno o más componentes del sistema 702 informático. El dispositivo 702 informático también puede incluir y/o estar acoplado manera comunicativa a varios dispositivos externos o internos, tales como dispositivos de entrada o salida. Por ejemplo, el dispositivo 702 informático se muestra con una interfaz 710 de entrada/salida ("I/O"), un dispositivo 712 de visualización, dispositivo(s) 714 de entrada y dispositivo(s) 715 de salida.

Los ejemplos no limitativos de un dispositivo 712 de visualización incluyen una pantalla integrada con el dispositivo 702 informático, un monitor externo al, y acoplado con el, sistema informático, etc. Los ejemplos no limitativos de dispositivos 714 de entrada incluyen dispositivos de detección de mirada, pantallas táctiles, paneles táctiles, dispositivos de ratón externos, micrófonos y/u otros dispositivos mencionados en el presente documento, etc. Un ejemplo no limitativo de un dispositivo 715 de salida es un altavoz de audio. En algunas realizaciones, el dispositivo 712 de visualización, el/los dispositivo(s) 714 de entrada y el/los dispositivo(s) 715 de salida pueden ser dispositivos independientes. En otras realizaciones, el dispositivo 712 de visualización y al menos alguno del/de los dispositivo(s) 714 de entrada pueden estar integrados en el mismo dispositivo. Por ejemplo, un dispositivo 712 de visualización puede ser una pantalla y un dispositivo 714 de entrada puede ser uno o más componentes que proporcionan funciones de seguimiento ocular y/o de pantalla táctil para el dispositivo de visualización, tales como emisores para emitir luz y/o cámaras para obtener imágenes del/de los ojo(s) de un usuario y/o una zona táctil, etc. La pantalla, componentes de dispositivo de entrada y cualquier componente de dispositivo de salida pueden estar integrados dentro del mismo alojamiento o en otras configuraciones integradas.

El dispositivo 702 informático puede modificar, acceder a o usar de otro modo contenido electrónico. El contenido electrónico puede residir en cualquier medio legible por ordenador no transitorio adecuado y ejecutarse en cualquier procesador adecuado. En una realización, el contenido electrónico puede residir en la memoria 706 en el sistema 702 informático. En otra realización, el sistema 702 informático puede acceder al contenido electrónico por desde un proveedor de contenido remoto por medio de una red de datos.

La memoria 706 puede incluir cualquier medio legible por ordenador no transitorio adecuado. Un medio legible por ordenador puede incluir, pero no se limita a, un dispositivo de almacenamiento electrónico, óptico, magnético u otro que puede proporcionar a un procesador instrucciones legibles por ordenador u otro código de programa. Otros ejemplos comprenden, pero no se limitan a, un disquete, CD-ROM, DVD, disco magnético, chip de memoria, ROM, RAM, un ASIC, un procesador configurado, almacenamiento óptico, cinta magnética u otro almacenamiento magnético, o cualquier otro medio a partir del cual un procesador informático puede leer instrucciones. Las instrucciones pueden comprender instrucciones específicas de procesador generadas por un compilador y/o un intérprete a partir de código escrito en cualquier lenguaje de programación informático adecuado, incluyendo, por ejemplo, C, C++, C#, Visual Basic, Java, Python, Perl, JavaScript y ActionScript.

Un módulo 716 de gráficos almacenado en la memoria 706 puede configurar el procesador 704 para preparar contenido electrónico para representar en una interfaz gráfica y/o representar el contenido electrónico en la interfaz gráfica. En algunas realizaciones, el módulo 716 de gráficos puede ser una aplicación autónoma ejecutada por el procesador 704. En otras realizaciones, el módulo 716 de gráficos puede ser un módulo de software incluido en o accesible por una aplicación independiente ejecutada por el procesador 704 que está configurado para modificar, acceder o usar de otro modo el contenido electrónico.

Mecanografiado de letras bloqueadas e interacción manual

En una realización alternativa de la invención, el módulo de decodificación puede crear y seleccionar diferentes hipótesis basadas en la aparición de diferentes tipos de acontecimientos de observación. Tal como se describió anteriormente, un acontecimiento de observación puede producirse basándose en diferentes características, incluyendo, por ejemplo y sin limitación, la fijación de la mirada del usuario en o cerca de un carácter o caracteres y/o la duración de la fijación. También pueden usarse otras características para diferenciar un acontecimiento de observación, tal como se describe a continuación. Un primer tipo de acontecimiento de observación, tal como una duración de fijación corta, puede indicar una probabilidad moderada de que el usuario pretendía seleccionar el carácter; un segundo tipo de acontecimiento, tal como una duración de fijación larga, puede indicar una probabilidad fuerte de que el usuario pretendía seleccionar el carácter. Pueden usarse tipos de acontecimientos según se necesite por el usuario para introducir texto de mayor o menor importancia para el usuario, tal como introducir una contraseña complicada frente a escribir un correo electrónico informal. Los caracteres introducidos con alta importancia indicados pueden denominarse "bloqueados" o "letras bloqueadas".

En una realización adicional, puede presentarse entonces una cadena introducida usando letras bloqueadas para corrección de errores, tal como mediante un método de corrección descrito anteriormente. Entonces el usuario puede ajustar la cadena introducida usando el método de corrección. Una cadena que se ha introducido usando al menos una letra bloqueada y/o se ha ajustado usando un método de corrección de errores puede denominarse "con interacción manual" por el usuario. Una cadena con interacción manual puede recibirse entonces por el sistema como adición a un modelo de lenguaje (es decir, una palabra que es probable que vuelva a usarse), permitiendo al usuario personalizar un vocabulario usado por el módulo de decodificador. En una realización adicional, el sistema puede determinar opcionalmente una medición de interacción manual con la cadena basándose en interacciones de usuario. Tales interacciones comprenden el número y/o la proporción de letras bloqueadas indicadas mientras se introduce la cadena, y el número y/o la proporción de caracteres que se han editado usando una interfaz de corrección de errores. La medición puede comprender un valor de confianza determinado basado en las interacciones, y el valor de confianza puede tener un umbral configurable. En aún una realización adicional, un modelo de lenguaje puede usar el valor de confianza para determinar si o con qué frecuencia el modelo de lenguaje

debe adaptarse al texto introducido del usuario, es decir, si el usuario corrige frecuentemente las hipótesis seleccionadas por el sistema, el sistema puede adaptarse y seleccionar hipótesis que son más similares al texto introducido del usuario. Valores de confianza mayores pueden dar como resultado una adaptación más fuerte, permitiendo que el sistema incorpore palabras que son importantes para el usuario. Aumentar el vocabulario del modelo de lenguaje permite que el módulo de decodificador cree hipótesis futuras basadas en las palabras añadidas, conduciendo a hipótesis más exactas.

Pasando ahora a la figura 8, se comentará un método 800 a modo de ejemplo de una realización alternativa de la invención. De manera comparable al método 400, el método 800 puede recibir datos de mirada continuos en el bloque 810. El punto de mirada puede visualizarse en el bloque 820, y una o más teclas de teclado en pantalla pueden visualizarse en el bloque 830. A medida que el usuario interactúa con el dispositivo, se monitoriza el punto de mirada en el bloque 840 tanto para determinar la aparición de un acontecimiento de observación como para determinar el tipo de acontecimiento de observación usado. A medida que se producen acontecimientos de observación, se graba de una manera en serie una secuencia de acontecimientos y tipos. La aparición de un primer tipo de acontecimiento de observación (por ejemplo, un tiempo de fijación de aproximadamente 200 milisegundos en un carácter particular) va seguida en el bloque 850 por la grabación en la secuencia tanto del primer acontecimiento de observación como del tipo de acontecimiento. La aparición de un segundo tipo de acontecimiento de observación (por ejemplo, un tiempo de fijación de aproximadamente 1000 milisegundos en un carácter particular) va seguida en el bloque 860 por la grabación en secuencia del segundo acontecimiento de observación y el tipo de acontecimiento. Los tipos acontecimiento de ejemplo descritos en este caso sólo son para ilustración, y pueden usarse otros tipos de acontecimiento sin apartarse de la realización de invención, incluyendo acontecimientos de combinación. Por ejemplo, el segundo acontecimiento puede indicarse al fijarse el usuario en un carácter durante cualquier cantidad de tiempo y pronunciar la palabra "bloquear" durante la fijación.

La grabación y la combinación de cada acontecimiento y tipo de observación particular puede producirse para cualquier número de acontecimientos. Las etapas de método se repetirán a partir del bloque 810, de tal manera que acontecimientos de observación detectados y sus tipos respectivos se graban continuamente en los bloques 850 ó 860, hasta que se produce un acontecimiento de terminación. La comprobación de un acontecimiento de terminación se produce en el bloque 870; si se produce uno, el método procede al bloque 880; de lo contrario, el método vuelve al bloque 810.

Tras el acontecimiento de terminación en 870, el método puede aplicar un módulo/algoritmo de decodificación en el bloque 880 para presentar como hipótesis una cadena a partir de la secuencia de observaciones y tipos. En esta realización, el módulo de decodificador puede presentar como hipótesis la cadena usando la información adicional del tipo de cada acontecimiento de observación grabado. Tal como se describió anteriormente, una hipótesis representa una secuencia de letras que puede ser el texto pretendido del usuario. El módulo de decodificador emitirá un conjunto de hipótesis, y a cada hipótesis se le asigna una probabilidad de corresponder al texto pretendido del usuario, de tal manera que la probabilidad asignada se basa al menos parcialmente en el tipo de acontecimiento de observación usado para introducir cada carácter. Los caracteres introducidos con el segundo tipo de acontecimiento de observación se interpretan por el decodificador como letras bloqueadas, y el decodificador asigna una probabilidad alta de que el usuario ha introducido el carácter pretendido. En algunas realizaciones, a las letras bloqueadas se les asigna una probabilidad del 100% (es decir, el decodificador supone que el usuario introdujo indudablemente el carácter pretendido); en otras realizaciones, la probabilidad asignada es de menos del 100%. Tras determinarse una o más hipótesis, pueden visualizarse una o más cadenas de correlación en la ventana 230 de mensaje en el bloque 890.

Tras la visualización de la cadena, la realización permite que el usuario confirme la cadena introducida usando una interfaz de confirmación, tal como la interfaz 200. Si el usuario fija su mirada en un símbolo particular en la ventana de mensaje, entonces se visualiza una interfaz de corrección de errores, tal como la interfaz 300. La realización permite entonces que el usuario corrija la cadena visualizada usando cualquier método apropiado.

En la figura 9 se representa un método a modo de ejemplo de una función de aprendizaje adaptativa. Tras la cancelación o confirmación de una cadena corregida, la realización puede comprender opcionalmente además una función de aprendizaje adaptativa mediante la cual se almacenan los resultados del algoritmo decodificado corregido y se usan en un algoritmo de aprendizaje para adaptar modelos de lenguaje a usuarios individuales. Si la cadena se corrigió, y la cadena corregida no se conoce en el modelo de lenguaje, y el usuario interactuó manualmente con la cadena, la palabra se añade al modelo. De lo contrario, no se añade la palabra y el sistema espera una entrada de mirada adicional.

Comenzando en el bloque 910, el método determina si se confirmó cualquier corrección de la cadena; si no es así, el método finaliza. Si es así, el método determina además en el bloque 920 si la palabra corregida u otra cadena no se conoce en un modelo de lenguaje usado por el algoritmo de decodificación; si no es así, el método finaliza. Si la cadena corregida se desconoce, entonces el método determina además en el bloque 930 si el usuario había interactuado manualmente con la cadena desconocida, es decir, la cadena se había introducido usando al menos una letra bloqueada y/o se había corregido al menos un error. Si no se había producido ninguna interacción manual, el método finaliza. Si se había interactuado con la cadena desconocida, entonces el método añade la cadena al

modelo de lenguaje como palabra de vocabulario conocida en el bloque 940, y entonces el método finaliza. Después de que el método 900 finalice en cualquier punto, la realización vuelve a un método para recibir una entrada de mirada, tal como el método 800 de la figura 8.

5 En determinadas realizaciones (no representadas en la figura 9), el método puede determinar una medición de interacción manual basándose en interacciones de usuario, tales como el número y/o la proporción de letras bloqueadas indicadas mientras se introduce la cadena, y el número y/o la proporción de caracteres que se han corregido. La medición puede comprender un valor de confianza, tal como en una escala arbitraria de 0 a 100, y el valor de confianza puede determinarse basándose en las interacciones. Como ejemplo, y no a modo de limitación,
10 una cadena que se introdujo con el 50% de los caracteres bloqueados y también se corrigió dos veces puede tener un valor de confianza determinado mayor que una cadena introducida con el 20% de los caracteres bloqueados y sin correcciones. La medición determinada puede basarse adicionalmente en otros factores, tales como si la cadena se había introducido o corregido anteriormente. La medición determinada puede necesitar cumplir un determinado umbral para añadir la palabra al modelo de lenguaje, y el umbral puede tener un valor por defecto que puede
15 configurarse por el usuario.

En determinadas realizaciones adicionales (tampoco representadas en la figura 9), el método puede usar el valor de confianza para determinar si o con qué frecuencia el modelo de lenguaje debe adaptarse al texto introducido del usuario. La determinación puede realizarse basándose en el valor de confianza solo o en combinación con otros
20 factores. Valores de confianza mayores pueden dar como resultado una adaptación más fuerte. Por ejemplo, y no a modo de limitación, si el usuario corrige más del 30% de las hipótesis seleccionadas por el módulo de decodificador, y el valor de confianza determinado de cada hipótesis corregida es relativamente alto, la realización puede adaptarse y seleccionar hipótesis que son más similares al texto introducido del usuario.

25 La invención se ha descrito ahora en detalle con fines de claridad y comprensión. Sin embargo, se apreciará que determinados cambios y modificaciones pueden ponerse en práctica dentro del alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario, en el que el método comprende:
- 5 hacer que un dispositivo (120; 712) de visualización visualice una interfaz (200) de introducción de texto que incluye una representación visual de una pluralidad de letras;
 recibir (410) información de mirada que identifica un movimiento de la mirada del usuario en la representación visual;
 10 grabar (450) una secuencia de observación de uno o más acontecimientos de observación determinados (440) que se producen durante el movimiento de la mirada del usuario en la representación visual;
 recibir (460) información de mirada que indica un acontecimiento de terminación;
 15 tras producirse el acontecimiento de terminación, proporcionar la secuencia de observación a un módulo de decodificador, en el que el módulo de decodificador determina (470) al menos una palabra a partir de la secuencia de observación que representa una estimación de un texto pretendido del usuario;
 hacer que el dispositivo (120; 712) de visualización visualice en una ventana (230) de mensaje de la interfaz (200) de introducción de texto la estimación del texto pretendido; y
 20 recibir (510) información de mirada adicional y determinar (520, 530, 540) que la información de mirada adicional indica que el usuario está fijando su mirada en una palabra, letra u otro símbolo particular de la estimación del texto pretendido en la ventana (230) de mensaje de la interfaz (200) de introducción de texto;
 estando el método caracterizado por,
 25 en respuesta a la petición de corrección, hacer que el dispositivo de visualización visualice una interfaz (300) de corrección que comprende la estimación del texto pretendido e iconos (320) de inserción entre palabras del texto pretendido, en el que un acontecimiento de observación adicional determinado (620, 630, 640) en un icono de inserción hace que la ubicación del icono de inserción se establezca para introducir una palabra adicional.
- 30 2. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 1, en el que un acontecimiento de observación comprende una determinación de que la mirada del usuario se ha fijado en o cerca de una letra.
- 35 3. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 2, en el que el acontecimiento de observación está asociado con uno o más de los siguientes:
- 40 coordenadas X, Y en el dispositivo (120; 712) de visualización de una fijación;
 una duración de la fijación;
 un tamaño de la fijación; o
 la letra más cerca de la fijación.
- 45 4. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 1, en el que el acontecimiento de terminación es una selección de un grupo que consiste en:
- 50 detección de la mirada del usuario en una zona (220) predeterminada del dispositivo (120; 712) de visualización durante un primer periodo de tiempo predeterminado;
 detección de que la mirada del usuario permanece estática durante un segundo periodo de tiempo predeterminado;
 detección de que la mirada del usuario no está ubicada en el dispositivo (120; 712) de visualización;
 55 detección de uno o más parpadeos del usuario;
 determinación de un gesto de mirada predeterminado del usuario; y
 una determinación automática, basada en alguna parte de la secuencia de observación, de que se ha cumplido o superado una probabilidad particular de que el usuario desee provocar el acontecimiento de terminación.
- 60 5. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 1, en el que la interfaz (300) de corrección permite además al menos uno de:
- 65 borrado de una palabra en la estimación del texto pretendido;
 corrección de una palabra en la estimación del texto pretendido.

6. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 1, en el que el módulo de decodificador comprende uno o más algoritmos que comprenden un decodificador de paso de testigo.
- 5 7. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 6, en el que el decodificador de paso de testigo emplea un modelo de teclado para calcular una probabilidad de una observación dada una letra presentada como hipótesis.
- 10 8. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 6, en el que el decodificador de paso de testigo emplea un modelo de lenguaje de palabras para calcular una probabilidad de una letra presentada como hipótesis dadas letras anteriores en una hipótesis actual.
- 15 9. Método (400; 500; 600; 800) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 1, en el que el módulo de decodificador comprende una función de aprendizaje adaptativa mediante la cual se analizan resultados del módulo de decodificación por el módulo de decodificación para adaptar el módulo de decodificación a usuarios individuales.
- 20 10. Sistema (100) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario, en el que el sistema (100) comprende:
- un dispositivo (140) de seguimiento ocular;
un dispositivo (120; 712) de visualización; y
- 25 el dispositivo (110; 702) informático configurado para realizar el método según la reivindicación 1.
11. Sistema (100) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 10, en el que un acontecimiento de observación comprende una determinación de que la mirada del usuario se ha fijado en o cerca de una letra.
- 30 12. Sistema (100) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 10, en el que el módulo de decodificador comprende uno o más algoritmos que comprenden un decodificador de paso de testigo.
- 35 13. Sistema (100) para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario según la reivindicación 10, en el que el módulo de decodificador comprende una función de aprendizaje adaptativa mediante la cual resultados del módulo de decodificación se analizan por el módulo de decodificación para adaptar el módulo de decodificación a usuarios individuales.
- 40 14. Medio legible por máquina no transitorio que tiene instrucciones almacenadas en el mismo para introducir texto en un dispositivo (110; 702) informático usando entrada de mirada de un usuario, pudiéndose ejecutar las instrucciones por uno o más procesadores (704) para realizar el método según la reivindicación 1.
- 45 15. Medio legible por máquina no transitorio según la reivindicación 14, en el que un acontecimiento de observación comprende una determinación de que la mirada del usuario se ha fijado en o cerca de una letra.
16. Medio legible por máquina no transitorio según la reivindicación 14, en el que el módulo de decodificador comprende uno o más algoritmos que comprenden un decodificador de paso de testigo.
- 50 17. Medio legible por máquina no transitorio según la reivindicación 14, en el que el módulo de decodificador comprende una función de aprendizaje adaptativa mediante la cual se analizan resultados del módulo de decodificación por el módulo de decodificación para adaptar el módulo de decodificación a usuarios individuales.
- 55

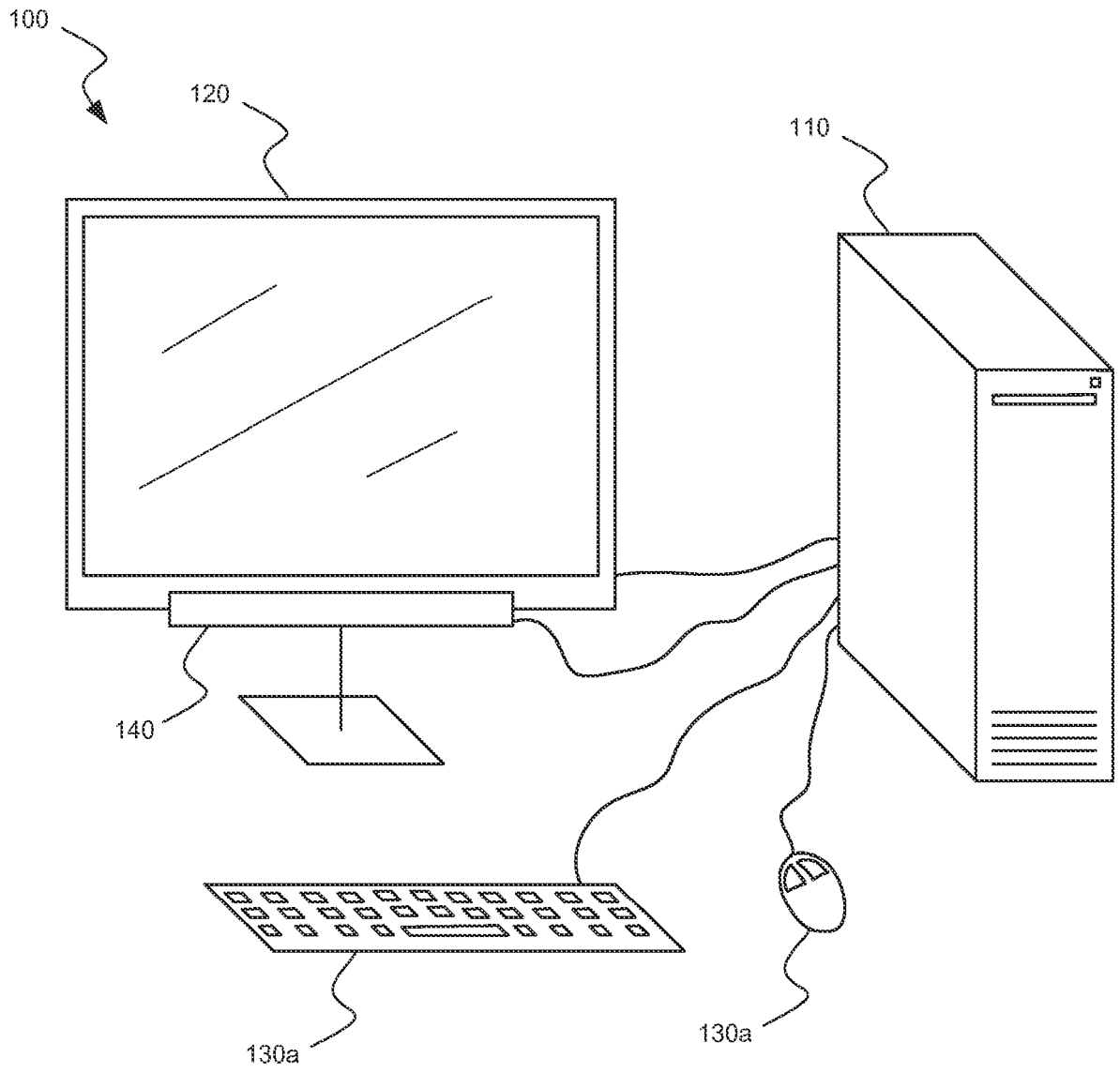


Fig. 1

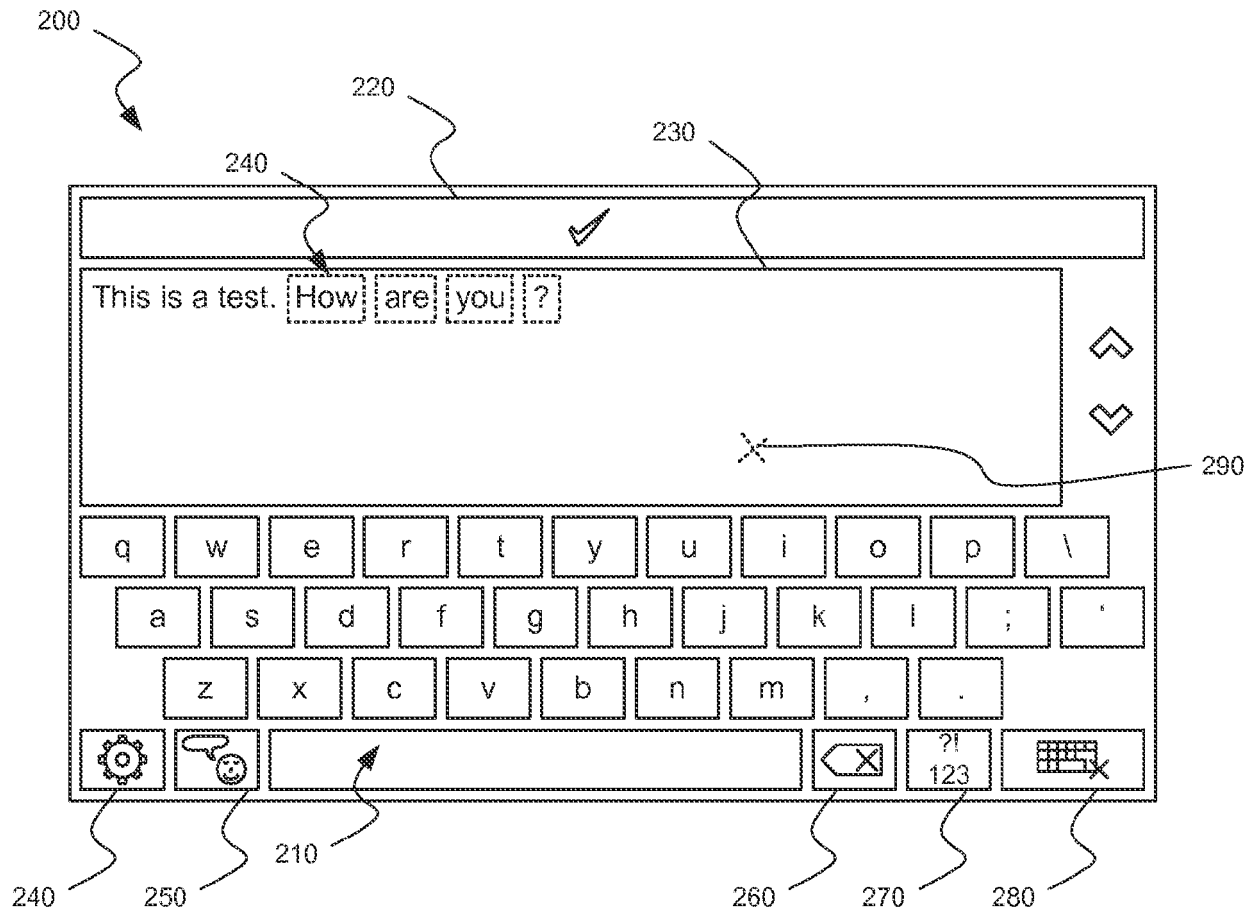


Fig. 2

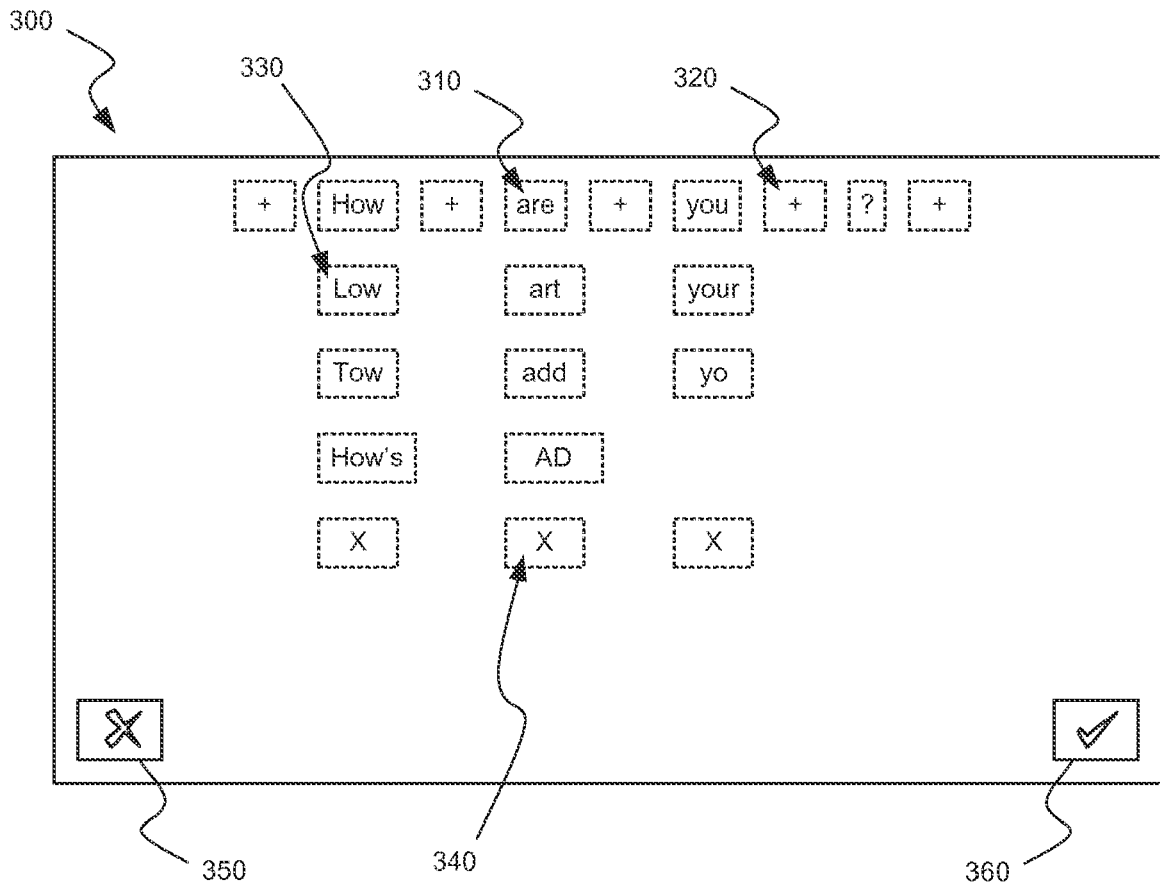


Fig. 3

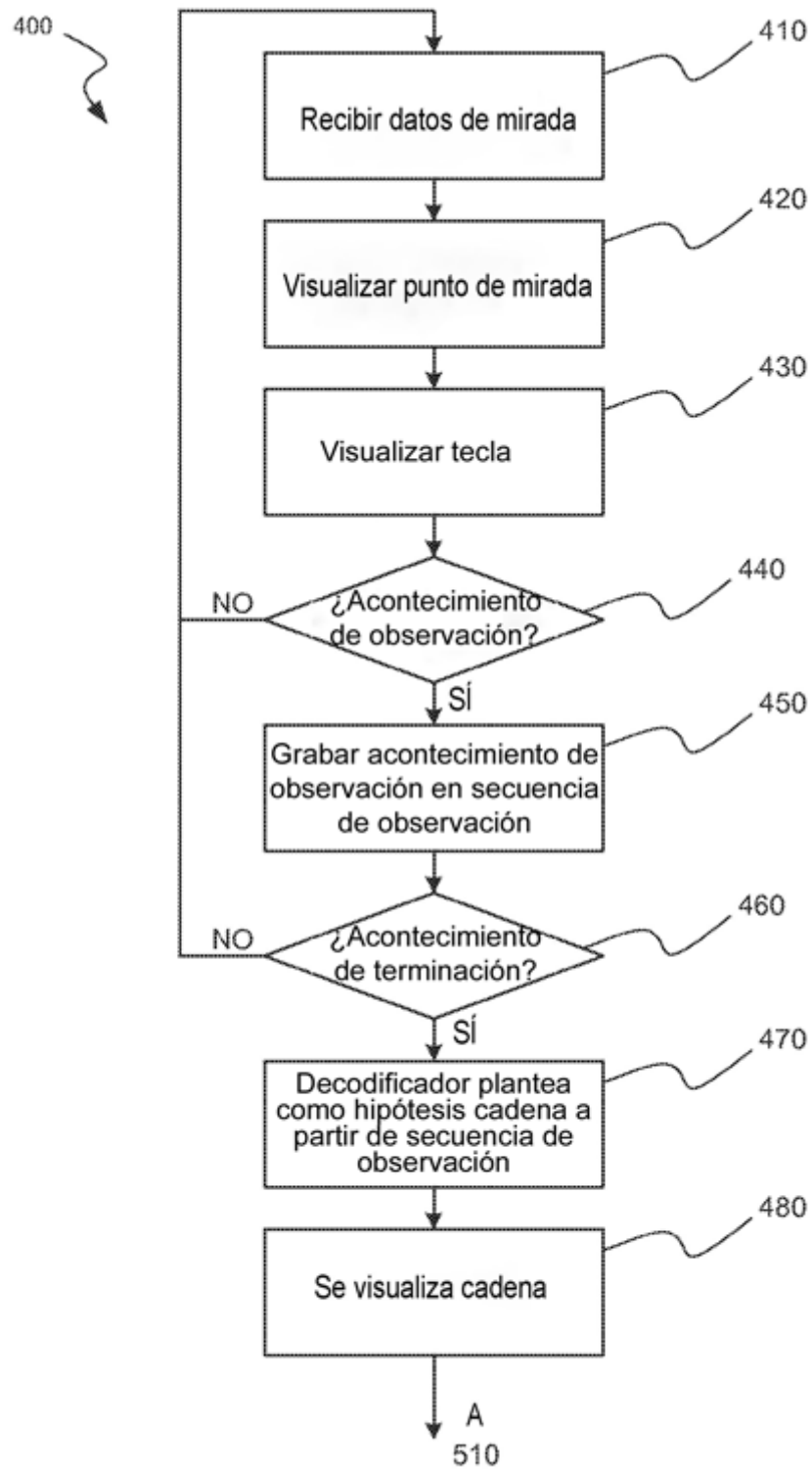


Fig. 4

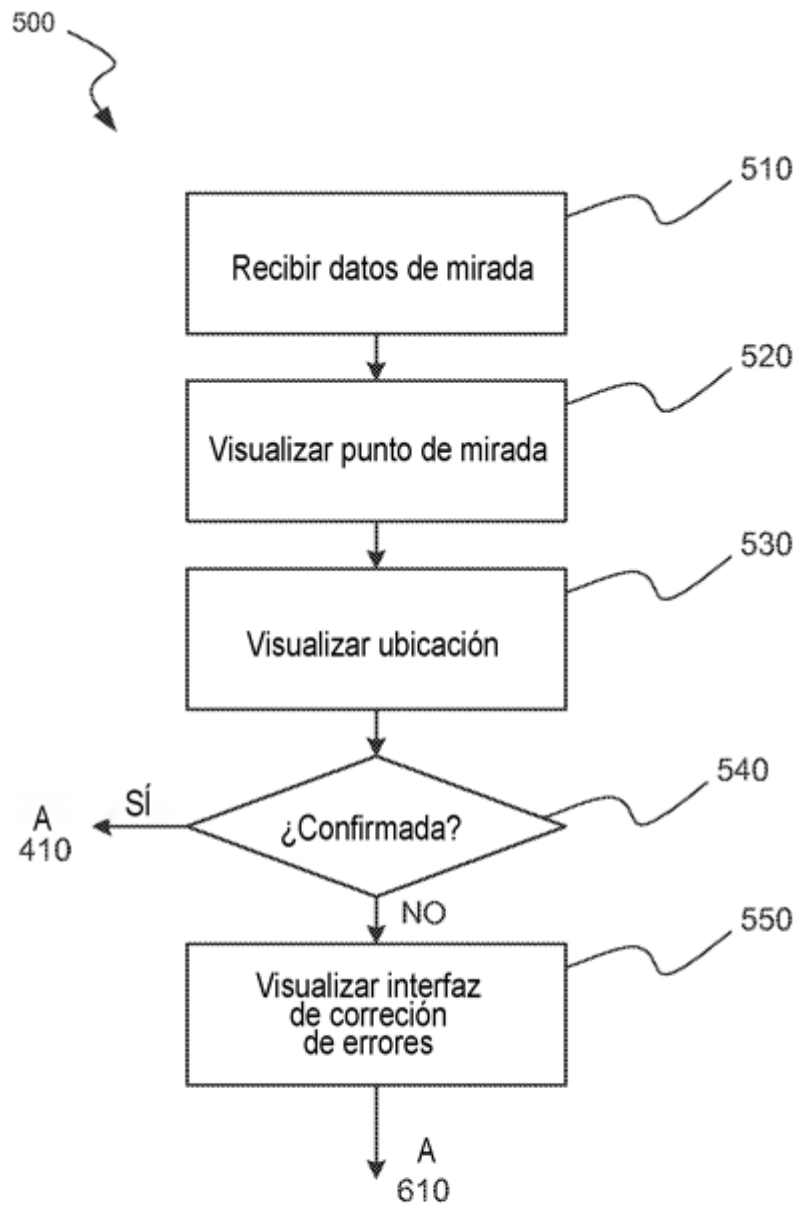


Fig. 5

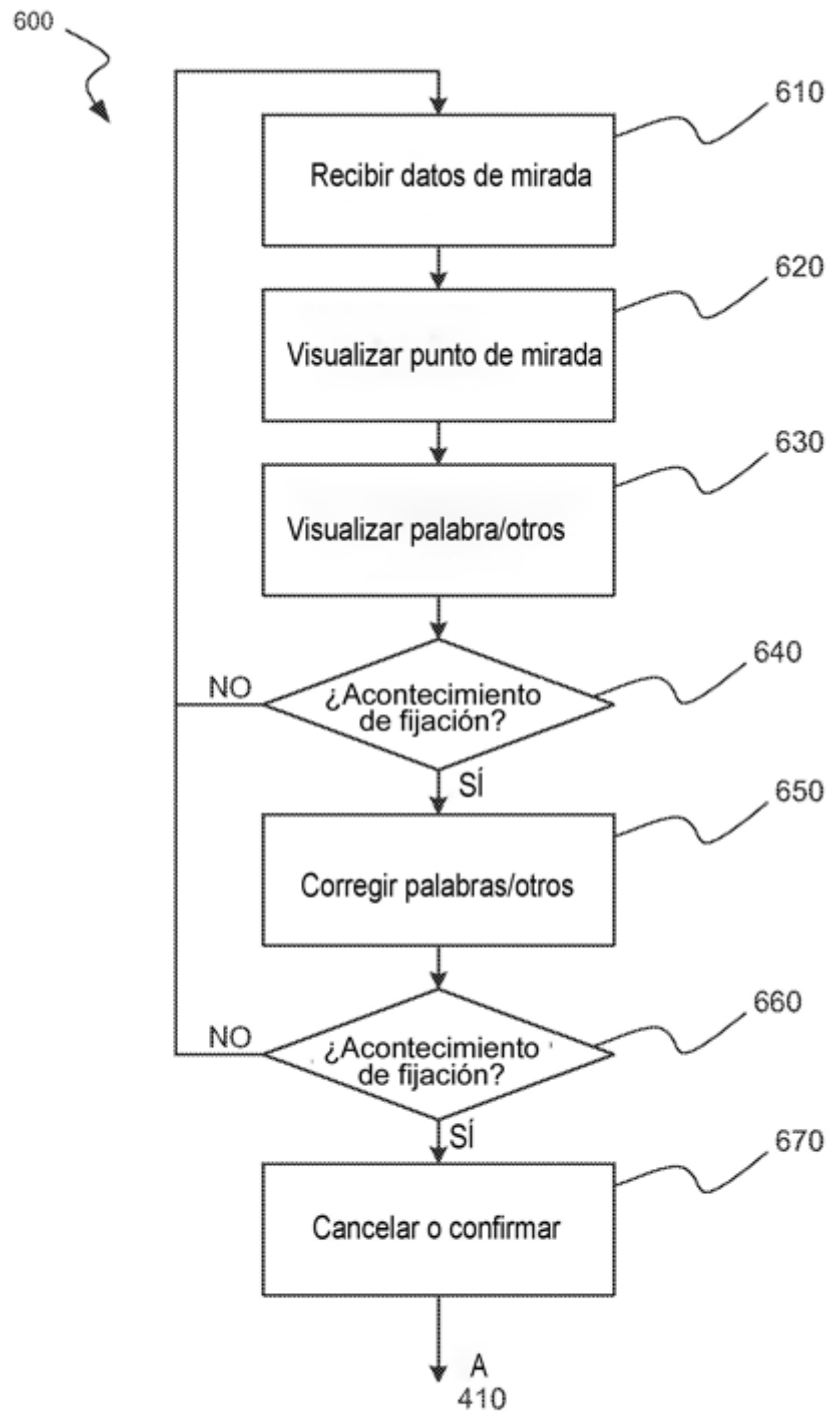


Fig. 6

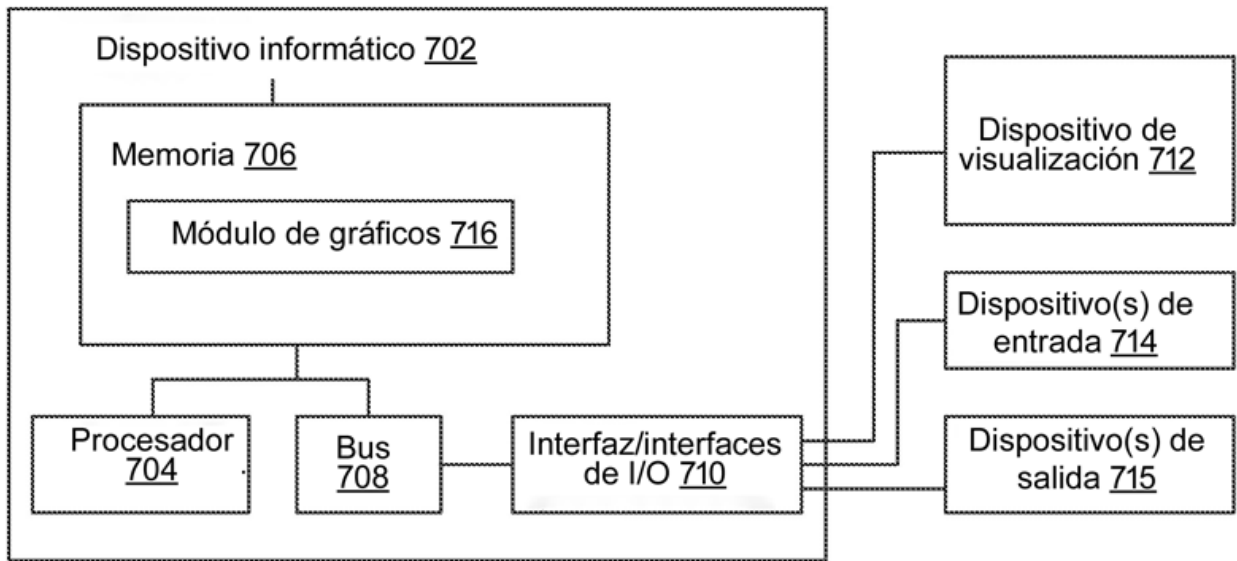


Fig. 7

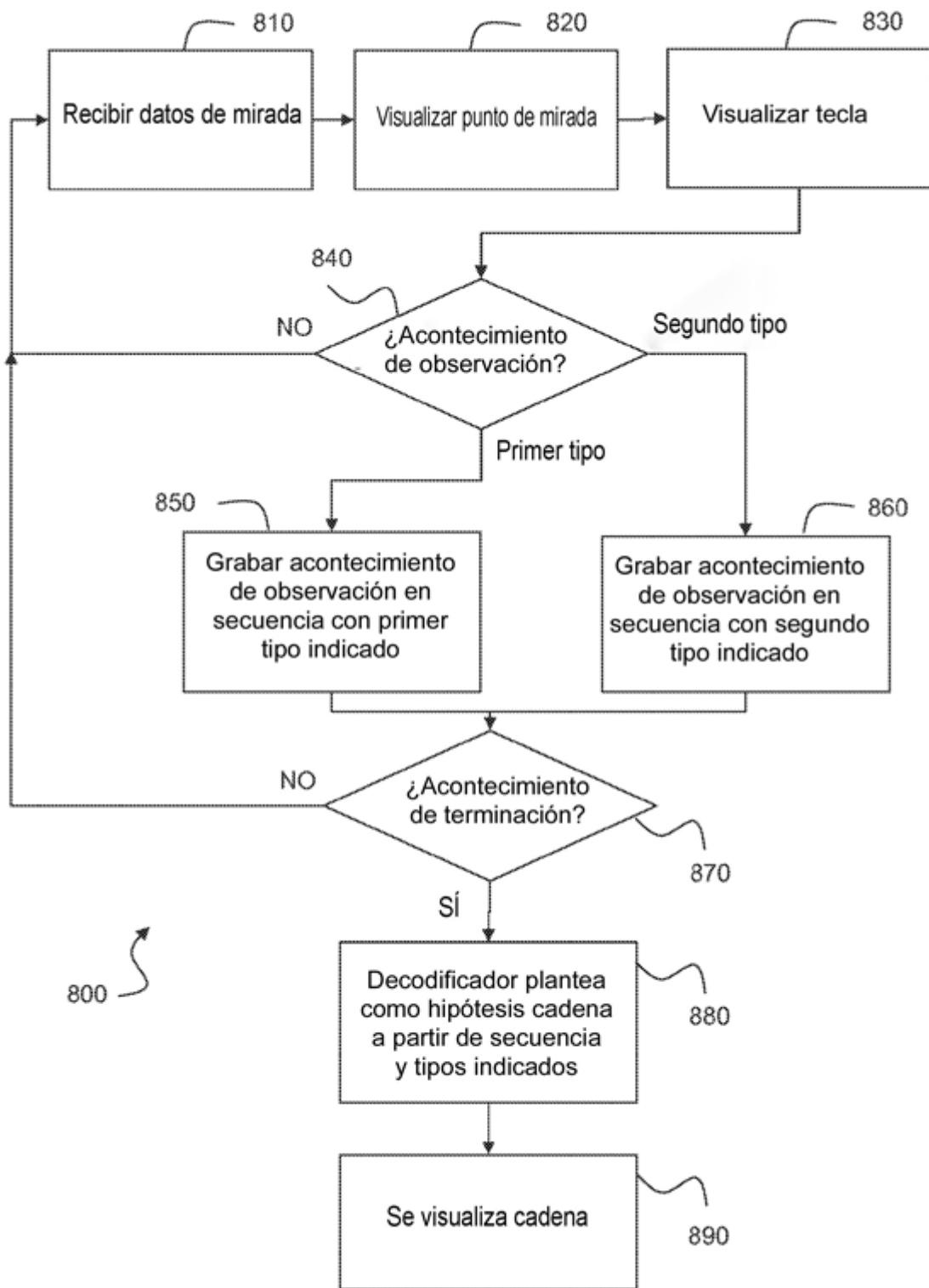


Fig. 8

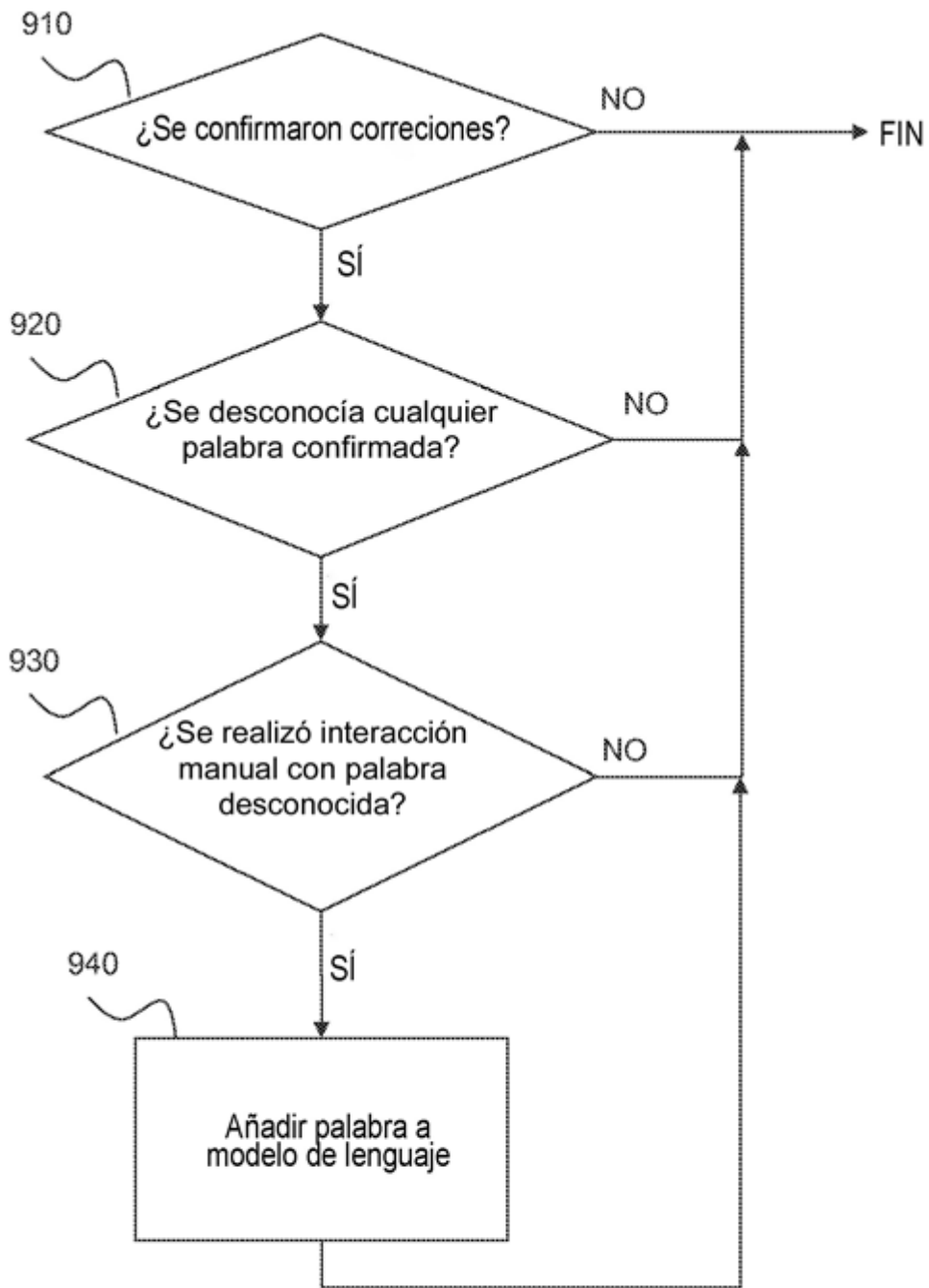


Fig. 9