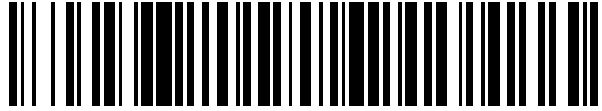


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 527 263**

51 Int. Cl.:

G08G 1/0968 (2006.01)

G01C 21/36 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.09.2005 E 12000135 (9)**

97 Fecha y número de publicación de la concesión europea: **12.11.2014 EP 2450867**

54 Título: **Sistemas y métodos para navegación automática de vehículo con voz, no incorporados**

30 Prioridad:

10.09.2004 US 608850 P
05.01.2005 US 29971

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
22.01.2015

73 Titular/es:

**SIRIUS XM CONNECTED VEHICLE SERVICES
INC. (100.0%)
8550 Freeport Parkway
Irving, TX 75063-2547 , US**

72 Inventor/es:

SCHALK, THOMAS

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 527 263 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 para navegación automática de vehículo con voz, no incorporados

Campo de la invención

5 La presente invención se relaciona en general con técnicas de navegación de vehículo, y en particular, con sistemas y métodos para navegación automática de vehículo con voz no incorporados.

Antecedentes de la invención

10 Los sistemas de navegación automotriz han estado disponibles durante numerosos años y se diseñan para guiar los operadores de vehículo a un destino específico. Un inconveniente principal de los sistemas de navegación convencional se relaciona con los métodos de ingresar los destinos objetivos. Es bien conocido que ocurre distracción del conductor cuando el operador del vehículo interactúa con un teclado o una pantalla táctil mientras maneja, de hecho, los usuarios típicamente se frustraron con los factores humanos y el necesario aprendizaje asociado para ingresar los destinos objetivo manualmente. Adicionalmente, los sistemas existentes le permiten a los usuarios ingresar el destino mientras se maneja, lo que ha mostrado que origina distracción del conductor. El ingreso de una dirección o punto de interés (POI) al utilizar métodos manuales típicamente requiere tiempo y concentración
15 de parte del operador del vehículo, y en particular, no se puede mirar el camino y manejar de manera segura. Existe un litigio pendiente en desarrollo que se relaciona con la distracción del conductor y el uso de sistemas de navegación mediante maneja.

20 Para la mayoría de los sistemas de navegación interna o que se encuentran en el vehículo, existen etapas secuenciales que ocurren durante el uso. El proceso inicia con la interacción del usuario donde el sistema de navegación determina primero la ubicación de partida, usualmente de la información GPS. El destino del blanco se ingresa típicamente como una dirección, una intersección de calles, o un punto de interés. Sería un avance sustancial en la técnica administrar sistemas de reconocimiento de voz automáticos que manejen un menú localizados en un centro de datos remoto que reconocería los destinos blanco hablados mientras se utiliza simultáneamente la información GPS transmitida desde el vehículo sobre un enlace inalámbrico a un centro de datos
25 remoto. También sería un avance significativo suministrar una interfaz de usuario por medio de voz que se diseñe para minimizar el tiempo de interacción del operador del vehículo y/o el tiempo de interacción del operador del centro de datos. Finalmente, sería un avance significativo si los destinos objetivo se pudieran determinar con una alta confiabilidad y eficiencia al utilizar la combinación de la tecnología de automatización con voz de información GPS, la ayuda del operador, y la ayuda del usuario para confirmar que es correcto el destino especificado. Cuando es necesario, el operador se involucraría en determinar el destino objetivo que se ha hablado, y el operador del vehículo (el usuario) confirmaría que el destino hablado es correcto antes de que se involucre el operador del centro de datos. Un reconocedor de habla automático, contexto para hablar de alta calidad, e información GPS juegan cada una un papel en el proceso final de determinar un destino objeto.

35 La WO 02/48650 divulga un sistema y un procedimiento para suministrar un servicio telefónico a un usuario con una descripción de la ruta a información acerca del objeto geográfico, por ejemplo una compañía, un tipo de compañía, una dirección o similar. El procedimiento incluye las etapas de que el usuario establezca una conexión por vía de un sistema de comunicación, preferiblemente un sistema telefónico móvil, con un dispositivo de servicio de habla que incluye un servidor de servicio de habla. El usuario establece/da un comando hablado al dispositivo de servicio hablado, cuyo comando incluye una inquietud con relación al tipo de objeto geográfico. El dispositivo del servicio de
40 habla deriva, de uno a la base de datos conectada al sistema de comunicación, la información relacionada con la posición asociada con dicho objeto, y presenta, mediante habla sintetizada, dicha información relacionada con la posición al usuario.

45 La US 2003/088359 divulga un método para suministrar información de navegación a un usuario de terminal de navegación. El método le permite a la terminal ITS recibir información de navegación de un centro de información por vía de un terminal de comunicación móvil y guía audiblemente la información de navegación a un usuario. El método comprende las etapas de transmitir información de posición al centro de información por vía de un canal de voz y recibir la información de navegación relacionada con la posición predeterminada del centro de información por vía de un canal de datos.

50 La WO 03/079310 divulga un sistema que incluye una unidad localizada en el vehículo y acoplada a una primera red, un centro de información de tráfico acoplado a una segunda red, y un servidor acoplado a la primera y segunda redes. Se genera un plan de viaje de acuerdo con la información de navegación, las coordenadas del vehículo y la información de solicitud de viaje recibida de la unidad, y la información del camino recibido del centro de información de tráfico. El plan de viaje se envía a la unidad para presentación.

La US2002156570 divulga un sistema de navegación de un coche para información de tráfico y otra información requerida en un coche, requiriendo a un proveedor de servicio de distribución de información la distribución, la información se compone de un dispositivo terminal en movimiento que se puede operar mediante voz que ingresa y que sale, un dispositivo de comunicación, un servidor interactivo de audio que puede ingresar y sacar voz y datos, está conectado a un dispositivo externo y la Internet sobre la base de interacciones de audio por medio de un proceso de reconocimiento de voz y un proceso de síntesis de voz, y ejecuta la transferencia de información, y uno o una pluralidad de proveedores de servicios de distribución de información conectados a la Internet para distribuir la información tal como información de mapa e información de tráfico.

Sumario de la invención

De acuerdo con esto, la presente invención está dirigida a un sistema y método para suministrar, o bajar la información de navegación desde una base de datos de un centro de datos remoto sobre un enlace inalámbrico a un vehículo. La información suministrada sería en respuesta a unos destinos objetivo reconocidos por voz hablados por el operador del vehículo. El sistema de reconocimiento de voz estaría localizado en el centro de datos remoto. La información suministrada, o bajada, podría ser el POI del destino objetivo, la intersección de calle, o dirección. El destino se determinaría a través de una interfaz de usuario con voz por medio de la cual cuatro componentes están involucrados en el proceso de automatización, que incluyen: tecnología de voz, información GPS del vehículo, el operador del centro de datos, y el operador del vehículo. La información suministrada, o bajada, podría ser la información de la ruta para el POI del destino objetivo, o la dirección, determinada a través de la interfaz de usuario con voz.

Las ventajas primarias del centro de datos remoto son la flexibilidad y la efectividad de costos. Se puede acceder a datos precisos, actualizados y la cantidad de datos puede ser muy grande en razón de la tecnología de memoria. En razón a que la plataforma de información no está incorporada, la aplicación se puede modificar fácilmente sin cambiar ni el hardware ni el software del vehículo. Tal flexibilidad permite la personalización del usuario y la incorporación de aplicación, en el cual un número de diferentes aplicaciones son accesibles a través de un menú principal de voz. En términos de costos, los recursos de reconocimiento de voz que se basan en el servidor se pueden compartir a través de un gran espectro de diferentes vehículos. Por ejemplo, 48 canales de reconocimiento de voz que se basan en servidor se pueden acomodar en mil vehículos de manera simultánea.

Los requisitos de la tecnología de voz para la invención incluyen texto a voz altamente inteligible, reconocimiento de voz, n-mejores resultados de búsqueda y niveles de confianza de reconocimiento asociada. El término "n-mejores resultados de búsqueda" se refiere a un formato de salida de reconocimiento de habla común que ordena por rango las hipótesis de reconocimiento basándose en la probabilidad. El texto para hablar se utiliza para representar lo que se reconoció automáticamente y se puede distinguir de la voz del operador del vehículo. Una base de datos de pronunciación, también denominada como una base de datos fonética, es necesaria para las pronunciaciones inteligibles correctas de los POI, ciudades, estados, y nombres de las calles. Para los casos en los cuales un reconocimiento resulta no tener una alta calificación de confianza, un registro de lo que se habló es reproducido al operador del vehículo para confirmación de que la representación de habla, o archivo de onda de audio, es correcto y reconocible por un humano, finalmente el operador del centro de datos. Por ejemplo, si un operador de vehículo dice una ciudad y estado, un nombre de calle, y un número de calle, entonces la aplicación repite lo que se ha hablado en una de tres maneras: una voz de ordenador pura (texto a habla), y una combinación de una voz de ordenador y la voz del operador del vehículo, o solamente la voz del operador del vehículo. En el último caso, el operador del centro de datos escucharía el habla y determinaría la dirección al escuchar y observar las n-mejores listas asociadas con cada parte de la dirección. En el caso anterior, el operador del centro de datos no estaría involucrado o no se requeriría; el proceso sería completamente automático. En el caso híbrido, el operador del centro de datos escucharía una parte de lo que se habló para determinar la dirección al escuchar y observar las n-mejores listas asociadas con la parte de la dirección no reconocida de manera automática. Sería típico para el operador escuchar o simplemente pinchar sobre la n-mejor selección que coincida con el componente de dirección en cuestión. Digitar el componente de dirección se requeriría si la n-mejor lista no contiene el componente de dirección correcto. Cuando se involucra, el operador del centro de datos puede seleccionar escuchar a cualquier componente de la dirección. Se utiliza una estrategia similar para determinar un POI hablado.

Para la entrada del POI, la interfaz de usuario de voz se puede diseñar para capturar una categoría POI (por ejemplo restaurante o ATM) y determinar si se desea la ubicación más cercana. Si es así, se completa la tarea de ingreso del destino hablado después de la confirmación con una respuesta de "sí". Si no se desea el sitio más cercano, una respuesta de "no" se manifiesta y el operador del vehículo se compele al decir el nombre del POI. De manera similar, si no se reconoce la categoría, esta se registra y pasa al operador del centro de datos además del nombre del POI, también grabado si no se reconoció, posterior a la confirmación del operador del vehículo de que la grabación está correcta. Para la determinación del POI, el GPS se puede utilizar para restringir la gramática activa del POI basándose en un radio específico con relación a la ubicación de vehículo.

Si el operador del vehículo dice una categoría POI y un nombre POI entonces la aplicación repite lo que se habló en una de tres formas: una voz de ordenador pura (texto a habla) una combinación de voz de ordenador y la voz del

operador del vehículo, o exclusivamente en solamente la voz del operador del centro de datos. En el último caso, el operador del centro de datos escucharía todo lo que se habló y determina el POI al escuchar y observar la n-mejor lista asociada con la categoría y nombre del POI. En el anterior caso, el operador no estaría involucrado o no se necesitaría; el proceso sería completamente automático. En el caso híbrido, el operador del centro de datos escucharía una parte de lo que se habló y determinaría el POI a través de escuchar y observar la n-mejor lista asociada con cualquier categoría o nombre del POI. Sería típico para el operador escuchar y simplemente pinchar sobre la n-mejor selección que coincida con el componente del POI en cuestión. Digitar el componente del POI se requeriría si la n-mejor lista no contiene el componente POI correcto. Cuando se involucra, el operador del centro de datos puede seleccionar escuchar cualquier componente del POI.

La invención descrita pretende estar integrada con un sistema de navegación a bordo capaz de procesar GPS en tiempo real para suministro de ruta. El sistema de navegación es una solución híbrida en el caso optimizado por que la ruta no se puede suministrar tan efectivamente en tiempo real desde un centro de datos remoto. Cuando direcciones vez por vez se suministran directamente desde el centro de datos remoto la información GPS que especifica la ubicación del vehículo puede perder sincronización con la posición efectiva del vehículo debido a latencias en la comunicación inalámbrica entre el vehículo y el centro de datos remoto. Por ejemplo, un indicativo generado por el sistema (por ejemplo, instrucción de voltear) puede experimentarse demasiado tarde por el operador del vehículo lo cual resulta en una desviación de la ruta. En resumen, la implementación ideal utiliza tecnología a bordo que incluye información de GPS en tiempo real para suministrar direcciones vez a vez mediante la voz en el ambiente del vehículo.

Breve descripción de los dibujos

Para un más completo entendimiento de la presente invención, y las ventajas de la misma, se hace referencia a la siguiente descripción tomada en conjunto con los dibujos que la acompañan, en los cuales:

La FIGURA 1A es un diagrama de bloque de un sistema de navegación automatizada de vehículo con voz, no incorporado que efectúa los principios de la presente invención;

Las FIGURAS 1B-1E muestran un diagrama de flujo que ilustra las operaciones de navegación de un vehículo automatizado con voz representativo implementado en el sistema mostrado en la FIGURA 1A.

La FIGURA 2 es un diagrama conceptual de una exhibición de un centro de datos representativo adecuado para implementar la ayuda del operador del centro de datos en reconocimiento del destino objetivo basándose en la información de punto de interés (POI);

La FIGURA 3 es un diagrama conceptual de una exhibición de un centro de datos representativo adecuado para implementar la ayuda del operador del centro de datos en reconocimiento del destino objetivo basándose en la información de ciudad y estado; y

Las FIGURAS 4 y 5 son diagramas conceptuales de exhibiciones de centros de datos representativos adecuados para implementar la ayuda del operador del centro de datos en el reconocimiento del destino objetivo basándose en la información de la ciudad, estado, y nombre de la calle.

Descripción detallada de la invención

Los principios de la presente invención y sus ventajas se entenderán mejor al referirse a las realizaciones ilustradas descritas en las FIGURAS 1-5 de los dibujos, en los cuales números similares designan partes similares.

La FIGURA 1A es un diagrama de un sistema de navegación automático con voz, no incorporado que representa los principios de la presente invención. La FIGURA 1B es un diagrama de flujo de un procedimiento 100 que ilustra las operaciones representativas del sistema 100, también representando los principios de la presente invención.

En referencia a las FIGURAS 1A y 1B, cuando el operador del vehículo 10 desea ingresar un destino objetivo con el fin de recibir la guía de ruta, se inicia un enlace de comunicaciones inalámbrico al centro de datos remoto 19 en el bloque 101 del procedimiento 100. El proceso se podría iniciar de numerosas maneras, tal como hablar un comando en el vehículo o preferiblemente al presionar un botón. Se establece la comunicación y el operador del vehículo 10 pronuncia comandos en el micrófono de manos libres 11, localizado en proximidad del operador del vehículo 10, en el bloque 102.

Los comandos hablados del operador del vehículo pasan por el enlace inalámbrico 25 a través del módulo de comunicación inalámbrico montado en el vehículo 14, a través de la antena inalámbrica montada en el vehículo 15, a través de la antena de la red inalámbrica 16 y de la estación base inalámbrica de la red 17, a través de una de muchas redes de comunicación 18 y en el centro de datos 19. Desde el centro de datos, la unidad de

reconocimiento de voz 20 interpreta el o los comandos hablados. Los comandos incluyen información con relación a la dirección, POI, o intersección de calle. Para una entrada de dirección, se pueden mencionar primero la ciudad y el estado.

5 La unidad de reconocimiento de voz 20 intenta, en el bloque 103 del procedimiento 100 de la FIGURA 1B, reconocer el ingreso hablado en el bloque 104 crea una n-mejor lista de la hipótesis superior donde n típicamente no excede cinco (esto es, la unidad de reconocimiento 20 genera hasta cinco representaciones de texto de posibles combinaciones de ciudad/estado, cada una con una probabilidad asociada de reconocimiento correcta). A cada hipótesis de reconocimiento se le asigna una calificación de confianza (probabilidad), en el bloque 105, que es normalizada a 1. Si a la elección más alta se le asigna una calificación de confianza por encima de un umbral especificado, en el bloque de decisión 106, el ingreso hablado se considera como reconocido, y el audio del habla de texto a habla ingresado en el ordenador es reproducido al operador del vehículo 10 (bloque 107) para confirmación (bloque 108). Si la confirmación es positiva en el bloque 111, entonces en los bloques 113 y 114 se genera una información de ruta automáticamente y se transmite a la unidad a bordo 13.

15 El audio de lo hablado se dirige a el o los parlantes del vehículo 12 en un ambiente de manos libres. El operador del vehículo 10 responde en el micrófono manos libres 11 a cada indicativo del sistema para especificar una dirección, diciendo de esta manera una ciudad, estado, nombre de calle, y número de calle. El operador del vehículo 10 escucha al o a los parlantes del vehículo 12 para oír la dirección de hipótesis representada por el audio de lo conversado que es 1) puramente generado por el ordenador, 2) puramente del operador del vehículo 12, o 3) una combinación de los dos tipos de audio de lo hablado.

20 La voz generada por el ordenador, utilizada en el bloque 107 del procedimiento 100, solo ocurre para declaraciones reconocidas (reconocimiento de la selección superior con alta confianza). Los componentes del destino (ciudad, estado, nombre de calle y número, POI, etc.) se identifican de otra forma auditivamente de manera individual en la propia voz del operador del vehículo 12 para confirmación cuando la calificación de confianza cae por debajo de un umbral. En particular, si alguno, o aún todos los componentes de destino hablados por el operador del vehículo tienen calificaciones de confianza por debajo del umbral en el bloque 106 del procedimiento 100, entonces al menos aquellos componentes con baja confianza son reproducidos al propietario del vehículo en la propia voz del operador del vehículo en el bloque 109, para confirmación en el bloque 110. Si el operador del vehículo confirma la reproducción del bloque 109, entonces en el bloque de decisión 112 el procedimiento 100 continúa al bloque 115 para ayuda del operador del centro de datos para la determinación del destino adecuado y la generación de unas guías de navegación adecuadas.

De otro lado, cuando la primera confirmación intentada produce un resultado negativo en el bloque 111 o en el bloque 112 del procedimiento 100, se efectúa la segunda reproducción en el bloque 117 y la segunda confirmación del propietario del vehículo se intenta en el bloque 118. Para el segundo intento de confirmación, todos los componentes del destino son reproducidos al operador del vehículo. Una confirmación positiva, en el bloque 118, concluye la experiencia del usuario para el ingreso del destino, después de lo cual el operador se involucra en el bloque 115, si es necesario. Se debe enfatizar que si el destino objetivo es dicho y grabado correctamente, no se requiere repetir de nuevo por el operador del vehículo 12; sin embargo, si el operador del vehículo aún no confirma los componentes de destino desde el segundo intento de confirmación, entonces el procedimiento 100, por ejemplo, regresa al menú principal y al operador del vehículo se le solicita repetir el destino deseado en el bloque 102.

40 Es crítico enfatizar que el operador del vehículo 10 confirma que el archivo de onda de audio almacenado es preciso antes de que se involucre la respuesta del operador del centro 23. Una confirmación de sí/no por vía de una unidad de reconocimiento de voz 20 se requiere para todos los destinos antes de que el operador del centro de datos 23 se involucre, si es necesario. Si la confirmación es negativa, otra elección de la n-mejor lista de entrada se selecciona en el bloque 106, para reproducción en el bloque 109 y se hace otro intento de confirmación en el bloque 110.

45 La FIGURA 2 representa una captura de pantalla de muestra de una estación de operador en vivo 22 que se diseña para ayudar al operador del centro de respuesta 23, en el bloque 115 del procedimiento 100, a determinar un destino objetivo. El ejemplo mostrado es para un POI específico, que incluye la correspondiente categoría POI. La FIGURA 2 ilustra dos n-mejores listas, una para la categoría POI y una para el correspondiente POI. Las calificaciones de confianza se listan luego en cada hipótesis de reconocimiento mostrada en las n-mejores listas, y sirve para indicar la probabilidad relativa de que la frase que se pronunció es la que se listó. Para la hipótesis, "complejo deportivo", el nivel de confianza mostrado es de .67199999, que es significativamente mejor que el nivel de confianza para la siguiente elección .01600000 (la frase de hipótesis hablada, "arrendamiento de coche"). Las dos casillas superiores contienen texto que coincide con las primeras elecciones de las n-mejores listas. El texto contenido en las dos casillas se puede modificar mediante o por el operador del centro de respuesta mediante una entrada de carácter 10 por carácter desde un teclado, al seleccionar la n-mejor entrada, también al utilizar un ratón (o teclado). A la derecha de cada casilla están los controles de audio (botones de radio) que permiten que los archivos de audio almacenados sean reproducidos y escuchados por el operador del centro de respuesta 23.

La capacidad del operador del centro de datos al reproducir las representaciones del archivo de fondo de audio de los componentes del destino pronunciados es crítica para el proceso completo. Para el ejemplo bajo consideración, existen dos componentes de destino: la categoría POI y el nombre POI. Si una frase diferente de la elección de voz se selecciona de la n-mejor lista, entonces el texto en la casilla superior correspondiente cambia automáticamente.

5 En el ejemplo mostrado, si se selecciona una categoría POI diferente por el operador del centro de respuesta 23, entonces una gramática posterior diferente se puede activar; la n-mejor lista para la POI cambia y una nueva elección superior se ingresa automáticamente en la casilla superior para el nombre del POI. Las calificaciones de confianza para la nueva n-mejor lista será muy diferente y se esperaría que fuera significativamente superior si el archivo de onda de audio almacenado coincide con la entrada de gramática. Para el ejemplo descrito aquí, el
10 operador del vehículo dice una categoría de POI. La categoría se reconoce, y al operador del vehículo se le solicita si el más cercano "complejo deportivo" es el destino deseado. Una respuesta positiva completa el ingreso de destino en el lado de la interfaz de usuario por que la información del GPS para la posición del vehículo es todo lo que se requiere para determinar la ruta en el bloque 113 del procedimiento 100. El GPS se utiliza como el punto de partida, y el POI más cercano se determina basándose en la selección de categoría y distancia.

15 La FIGURA 3 representa parte de una captura de pantalla de muestra de la estación del operador en vivo 22 que se diseña para ayudar a la respuesta del operador del centro 23, en el bloque 115 del procedimiento 100, al determinar un componente del destino objetivo. El ejemplo mostrado es para una ciudad y estado específicos, e incluye la n-mejor lista generada por la unidad de reconocimiento de voz 20 para la ciudad y estado que se pronunció por el operador del vehículo 10. Las calificaciones de confianza se listan luego en cada hipótesis de reconocimiento
20 mostrada en la n-mejor lista y sirven para indicar la probabilidad relativa de que la frase que se ha pronunciado es la que se listó. Para la hipótesis "Dallas Texas", la calificación de confianza mostrada es .96799999, que es significativamente mejor que la calificación de confianza para la siguiente mejor elección, .01899999 (la frase pronunciada de hipótesis, "Alice, Texas").

25 En relación de nuevo con la FIGURA 3, la casilla superior contiene texto que coincide con las primeras elecciones de las n-mejores listas. El texto contenido en cada casilla se puede modificar por el operador del centro de respuesta bien sea mediante una entrada de carácter por carácter desde el teclado, o al seleccionar una n-mejor entrada al utilizar un ratón. A la derecha de la casilla superior están los controles de audio que le permiten a los archivos de onda de audio almacenados ser reproducidos y escuchados por el operador del centro de respuesta 23. De nuevo, la capacidad de reproducir las representaciones del archivo de onda de audio del destino de los componentes del
30 destino pronunciado es crítico para el proceso completo. Si una frase diferente de la elección superior se selecciona de la n-mejor lista, entonces el texto en la casilla superior correspondiente cambia automáticamente. El archivo de onda de audio representa lo hablado suministrado por el operador del vehículo 10 (en este caso, una ciudad y un estado).

35 La FIGURA 4 representa otra captura de pantalla de la estación del operador en vivo 22 que se diseña para ayudar al operador del centro de respuesta 23 a determinar el destino objetivo. El ejemplo mostrado es para una ciudad, estado, y nombre de calle específico. La FIGURA 4 ilustra dos n-mejores listas, una para la ciudad y el estado y una para el nombre de la calle. Las calificaciones de confianza se listan luego a cada hipótesis de reconocimiento mostrada en las n-mejores listas y sirven para indicar la probabilidad relativa de que la frase que se ha pronunciado es la que se listó. Para la hipótesis "Winchester, California", la calificación de confianza mostrada es 0.18600000,
40 que no es significativamente mejor que la calificación de confianza para la siguiente mejor elección, 0.14499999 (la frase hablada de hipótesis, "Westchester, California"). En referencia a la FIGURA 4, las dos casillas superiores contienen textos que coinciden con las primeras elecciones de las dos n-mejores listas. El texto contenido en las dos casillas se puede modificar por el operador del centro de respuesta bien sea en una entrada carácter por carácter desde un teclado, o al seleccionar la n-mejor entrada, al utilizar un ratón (u otros medios). A la derecha de cada
45 casilla se encuentran los controles de audio que le permiten a los archivos de onda de audio almacenados ser reproducidos y escuchados por el operador del centro de respuesta 23.

La capacidad de reproducir las representaciones del archivo de onda de audio de los componentes de destino hablados es crítica para el proceso completo. Para el ejemplo en consideración, existen dos componentes de destino: la ciudad/estado y el nombre de la calle. Si una hipótesis diferente de la elección se selecciona de la n-mejor lista, entonces el texto en la casilla superior correspondiente cambia automáticamente. En el ejemplo
50 mostrado, si una ciudad/estado diferente se selecciona por el operador del centro de respuesta 23, entonces se activa una gramática posterior diferente; la n-mejor lista para el nombre de la calle cambia y la nueva selección superior es ingresada automáticamente en la casilla superior para el nombre de la calle. La FIGURA 5 ilustra el resultado que ocurre cuando "Lancaster, California" se selecciona por el operador del centro de respuesta 23. Las calificaciones de confianza para la nueva n-mejor lista del nombre de la calle son muy diferentes, y la calle de elección superior tiene una alta calificación de confianza, .996, que es cercana a una coincidencia perfecta. Se aprecia que la tarea del operador del centro de respuesta 23 para el ejemplo descrito aquí es: 1) escuchar el archivo de onda de audio de la ciudad/estado, 2) seleccionar la ciudad/estado correcta, 3) escuchar el archivo de onda de audio del nombre de la calle para confirmar que es correcto, 4) escuchar el archivo de onda de audio del número de
55 la calle para confirmar que es correcto (no ilustrado) y hacer cualquier corrección de digitación si se requiere antes del suministro final para el procesamiento relacionado con la navegación.
60

5 El nivel de archivos de onda de audio capturados se puede normalizar al aplicar un control de ganancia automático digital para mejorar la inteligibilidad humana y la consistencia de la interfaz de usuario durante la reproducción de audio de los componentes de destino. El audio capturado puede servir para indicar la calidad de las condiciones de la red al operador del vehículo. El audio capturado le enseña al operador del vehículo como hablar en el micrófono y lograr un reconocimiento óptimo.

10 Aunque la invención se ha descrito en referencia a realizaciones específicas, estas descripciones no deben interpretarse en un sentido limitativo. Diversas modificaciones de las realizaciones divulgadas, así como realizaciones alternativas de la invención, serán aparentes para las personas expertas en la materia al hacer referencia a la descripción de la invención. Los expertos en la materia deberían apreciar que la concepción y la realización específica divulgadas pueden utilizarse de inmediato como base para modificar o diseñar otras estructuras para llevar a cabo los mismo fines de la presente invención. Los expertos en la materia también deberían darse cuenta de que tales construcciones equivalentes no se apartan del alcance de la invención según lo expuesto en las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un método para suministrar información de navegación que comprende:

recibir en una interfaz de usuario de voz el habla (102) de un usuario de la interfaz de usuario de voz (11, 12, 14), incluyendo el habla una completa dirección de destino que es en su totalidad al menos uno de:

5 un punto de interés de destino objetivo;

una intersección de calle; y

un conjunto de dirección que incluye al menos un número de calle y un nombre de calle;

procesar el habla (102) en información de navegación con voz;

procesar el habla (102) en información de navegación con voz;

10 transmitir la información de navegación con voz por vía de un enlace inalámbrico (25) a un centro de datos remoto (19);

analizar (103) la información de navegación con voz con un sistema de reconocimiento de voz (20) en el centro de datos remoto (19) para reconocer los componentes de la dirección de destino completa pronunciada por el usuario; **caracterizado por**

15 generar (104) en el centro de datos remoto (19) una lista de los componentes reconocidos hipotéticos de la dirección de destino completa listada por los niveles de confianza tal como se calculó para cada componente de la dirección de destino completa analizada por el sistema de reconocimiento de voz (20);

desplegar la lista de los componentes reconocidos hipotéticos y los niveles de confianza en el centro de datos remoto para hacer una revisión selectiva por un operador humano del centro de datos (23);

20 seleccionar un conjunto de componentes hipotéticos reconocidos basándose en los niveles de confianza en la lista; confirmar (111, 112) la precisión del conjunto seleccionado de componentes hipotéticos reconocidos de la dirección de destino completa por vía de intercambios de voz interactivos entre el usuario y el centro de datos remoto (19) al reproducir un audio hablado de texto a voz generado por ordenador al usuario para confirmación;

determinar el destino desde componentes confirmados de la dirección de destino completa;

25 generar (113) la información de ruta al destino en el centro de datos remoto (19); y

transmitir (114) la información de ruta a la interfaz de usuario de voz desde el centro de datos remoto (19) por vía del enlace inalámbrico (25).

2. El método de acuerdo con la reivindicación 1, que comprende además llevar a cabo la etapa de confirmación de precisión al:

30 transmitir una representación generada por ordenador de al menos un componente hipotético reconocido de la dirección de destino completa al usuario por vía del enlace inalámbrico (25); e

indicar al usuario por vía del enlace inalámbrico (25) a confirmar oralmente la precisión del componente de la dirección de destino completa.

35 3. El método de acuerdo con la reivindicación 1, que comprende además llevar a cabo la etapa de confirmación de precisión al:

transmitir al menos un componente hipotético reconocido grabado de la dirección de destino completa pronunciada por el usuario al usuario por vía del enlace inalámbrico (25); e

indicar al usuario por vía del enlace inalámbrico (25) a confirmar oralmente la precisión del componente hipotético reconocido de la dirección de destino completa.

40 4. El método de acuerdo con la reivindicación 2, que comprende además llevar a cabo la etapa de confirmación de

precisión al:

determinar (106) si un nivel de confianza de un componente hipotético reconocido está por encima de un umbral seleccionado; y

5 generar por ordenador una representación del componente hipotético reconocido para transmisión al usuario (10) cuando el nivel de confianza está por encima del umbral seleccionado.

5. El método de acuerdo con la reivindicación 1, que comprende además llevar a cabo la etapa de determinación de destino al suministrar ayuda al operador humano del centro de datos (23) utilizando la lista desarrollada de los componentes hipotéticos reconocidos y los niveles de confianza para reconocer la dirección de destino completa.

10 6. El método de acuerdo con la reivindicación 1, que comprende además llevar a cabo la etapa de confirmación de precisión al transmitir representaciones orales de los componentes hipotéticos reconocidos de la dirección de destino completa al usuario (10), los componentes hipotéticos reconocidos de la dirección de destino completa seleccionados de representaciones orales de porciones de la dirección de destino completa del grupo que consiste en número de la dirección, nombre de la calle, ciudad, estado y punto de interés.

15 7. El método de acuerdo con la reivindicación 6, que comprende además llevar a cabo la etapa de proporcionar ayuda al operador al:

reproducir las representaciones grabadas de la dirección de destino completa pronunciada por el usuario (10) al operador del centro de datos para análisis del operador del centro de datos; y

recibir información del operador del centro de datos (23) que identifica el destino, lo que comprende ingresar una selección de la lista desplegada de los componentes hipotéticos reconocidos del operador del centro de datos.

20 8. El método de acuerdo con la reivindicación 1, que comprende además llevar a cabo la etapa de generación de información de ruta al generar información de ruta de la información del sistema de posicionamiento global recibida por el centro de datos remoto desde la interfaz de usuario de voz.

9. Un sistema para suministrar información de navegación que comprende:

25 una interfaz de usuario de voz (11, 12, 14) para procesar y transmitir por vía de enlace inalámbrico (25) el habla de un usuario (10) de la interfaz de usuario de voz (11, 12, 14) incluyendo el habla (102) una dirección de destino completa; y

un centro de datos (19) para procesar el habla (102) recibida por vía del enlace inalámbrico (25) en información de navegación con voz **caracterizado porque** el sistema es operable para:

30 efectuar procesamiento de reconocimiento de voz automático sobre la información de navegación con voz para reconocer los componentes de destino de la dirección de destino completa;

confirmar los componentes de destino reconocidos a través de intercambios de habla interactivos con el usuario (10) por vía del enlace inalámbrico (25) y la interfaz de usuario de voz al reproducir un audio del habla de texto a voz generado por ordenador al usuario para confirmación;

35 permitir de manera selectiva la intervención del operador humano del centro de datos (23) para ayudar a identificar los componentes de destino reconocidos seleccionados que tienen una confianza de reconocimiento por debajo de un valor de umbral seleccionado; y

bajar información de navegación perteneciente a la dirección de destino completa para la transmisión a la interfaz de usuario de voz derivada de los componentes de destino de reconocimiento confirmados.

40 10. El sistema de acuerdo con la reivindicación 9, en el que el centro de datos (19) es operable además para bajar la información de navegación en respuesta a la información de posición recibida de la interfaz de usuario de voz por vía del enlace inalámbrico (25).

11. El sistema de acuerdo con la reivindicación 9, en el que el centro de datos es operable además para:

generar (104) una lista de posibles componentes de destino que correspondan al habla (102) del usuario;

asignar una calificación de confianza para cada uno de los posibles componentes de destino en la lista;

determinar si un componente de destino posible con una más alta calificación de confianza tiene una calificación de confianza por encima de un umbral; y

generar por ordenador una representación oral de la dirección de destino para la transmisión a la interfaz de usuario de voz (11, 12, 14) para confirmación del usuario (10) de si la calificación de confianza está por encima del umbral.

- 5 12. El sistema de acuerdo con la reivindicación 9, en el que el centro de datos (19) es operable además para:

determinar (106) que al menos uno de los componente de destino del habla del usuario tiene un valor de confianza de reconocimiento por debajo de un umbral; y

- 10 reproducir (109) una grabación en la voz del usuario (10) del al menos un componente de destino con el valor de confianza de reconocimiento por debajo del umbral al usuario por vía de la interfaz de usuario de voz (11, 12, 14) para confirmación.

13. El sistema de acuerdo con la reivindicación 12, en el que el centro de datos (19) comprende además una instalación del operador del centro de datos (23) para reproducir los componentes del destino para ayudar a identificar la dirección de destino completa.

- 15 14. El sistema de acuerdo con la reivindicación 12, en el que el centro de datos (19) es operable además para presentar una lista de posibles destinos listados por las calificaciones de confianza al operador del centro de datos (23) para selección como la dirección de destino completa.

15. El sistema de acuerdo con la reivindicación 14, en el que el centro de datos (19) es operable además para permitirle al operador del centro de datos (23) variar un orden de los posibles destinos en la lista.

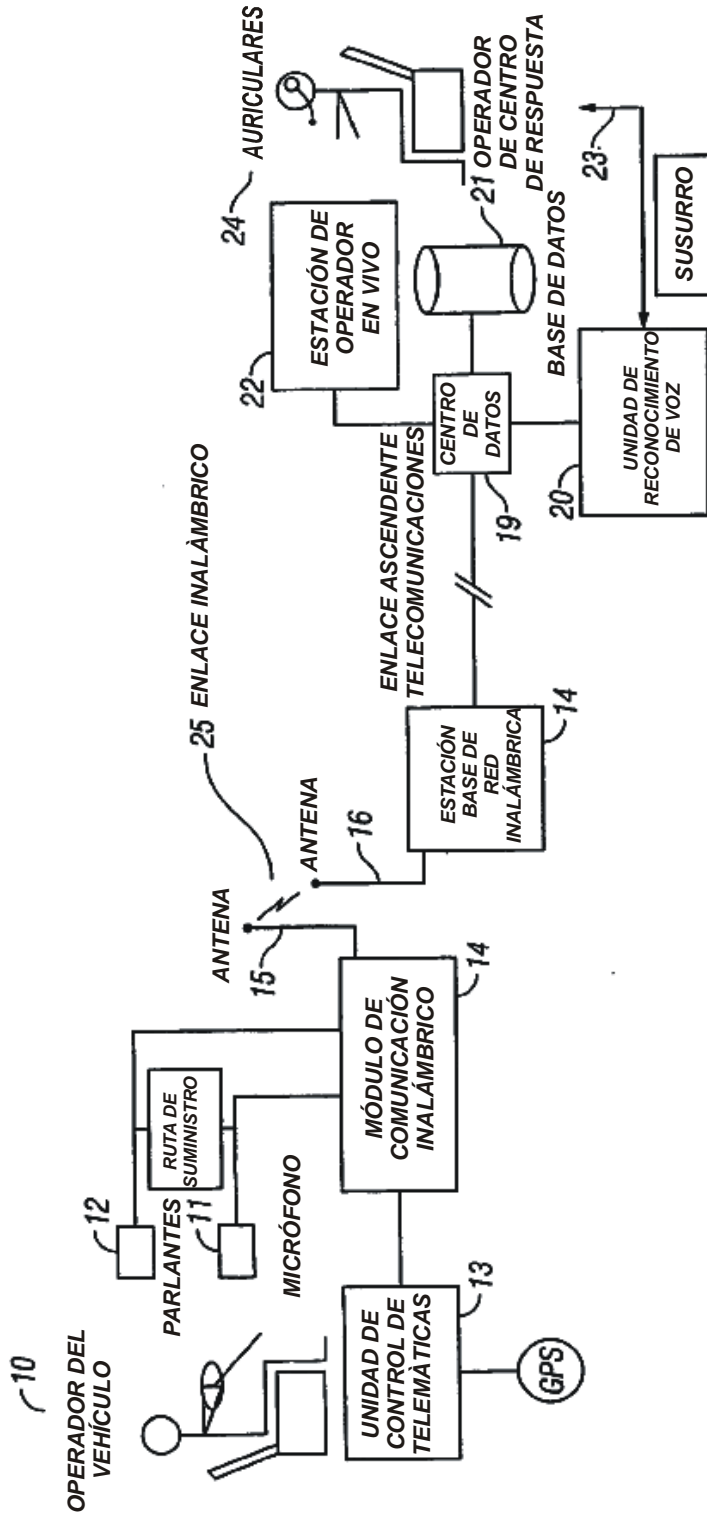


FIG. 1A

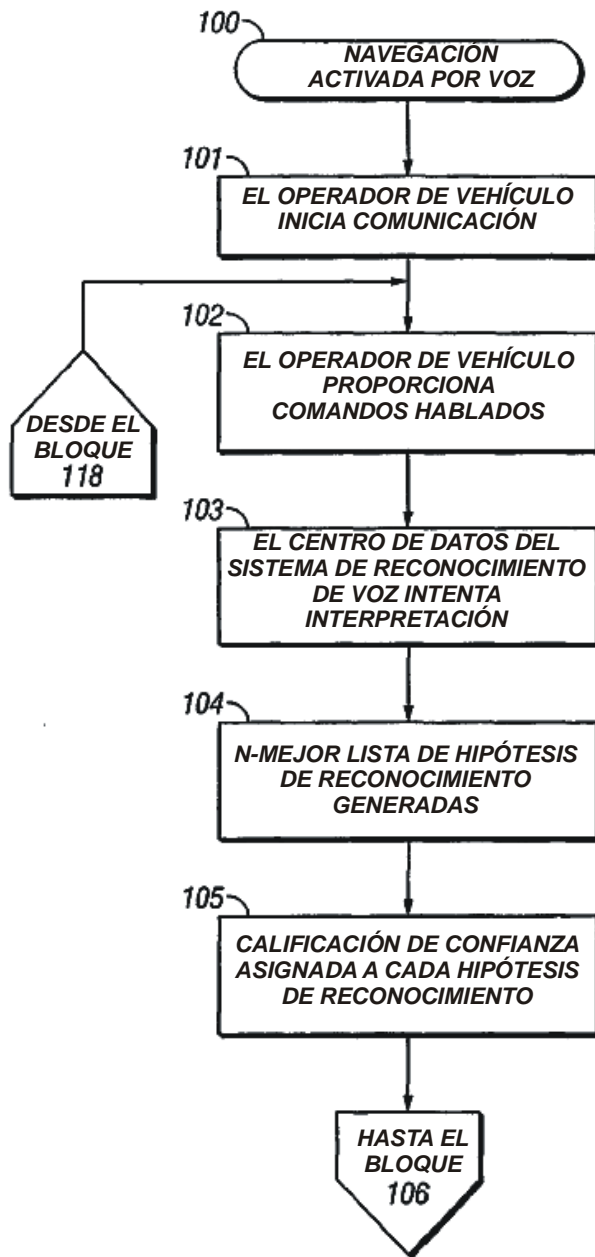


FIG. 1B

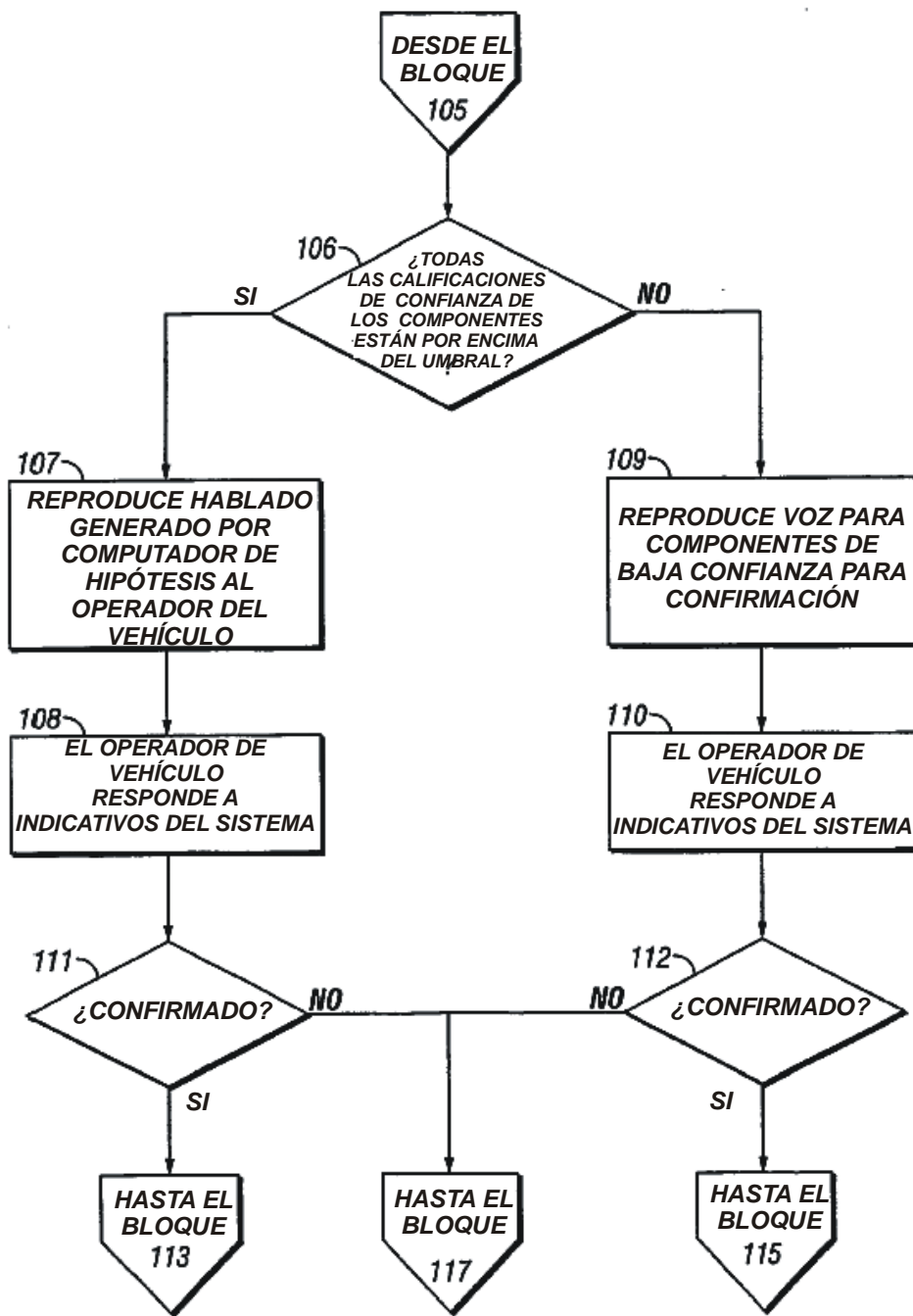


FIG. 1C

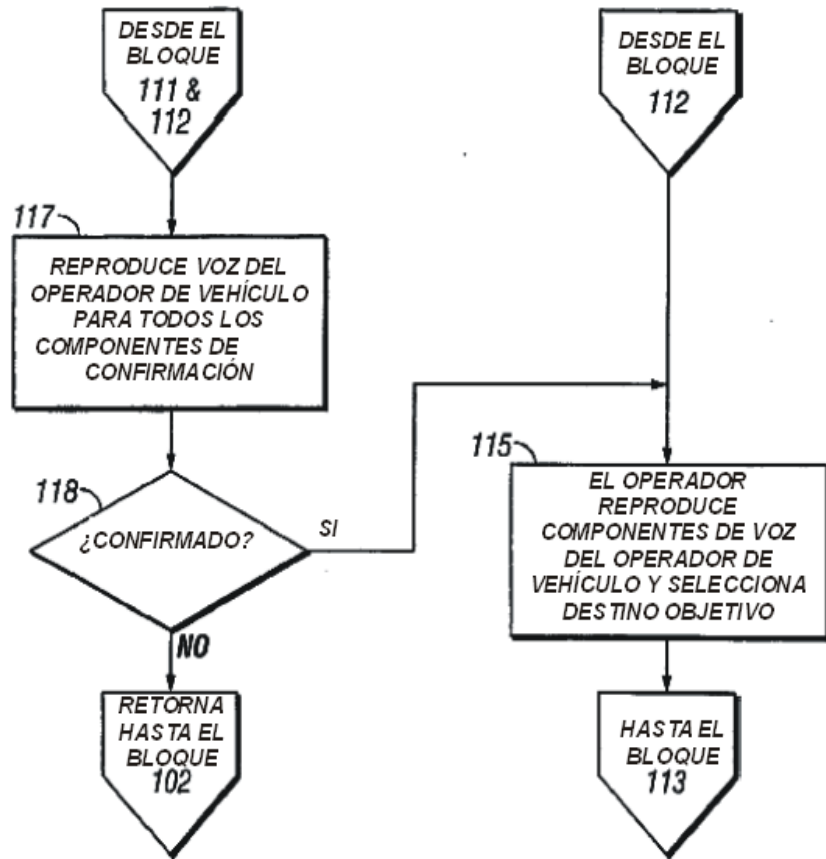


FIG. 1D

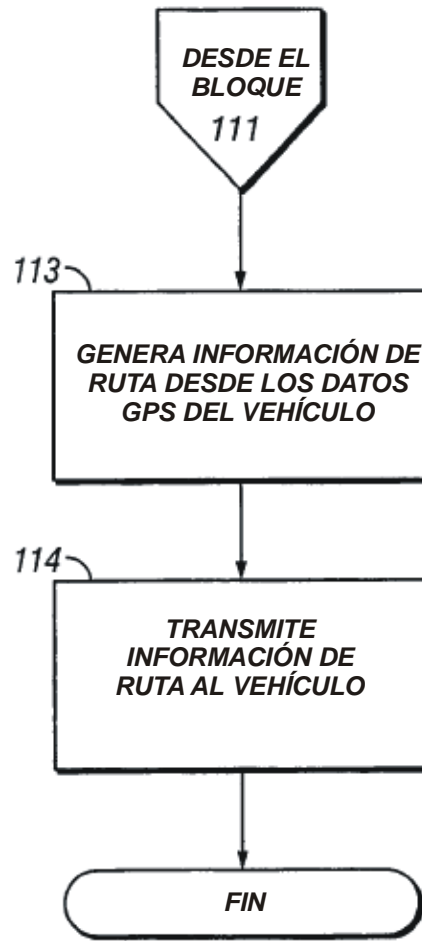


FIG. 1E

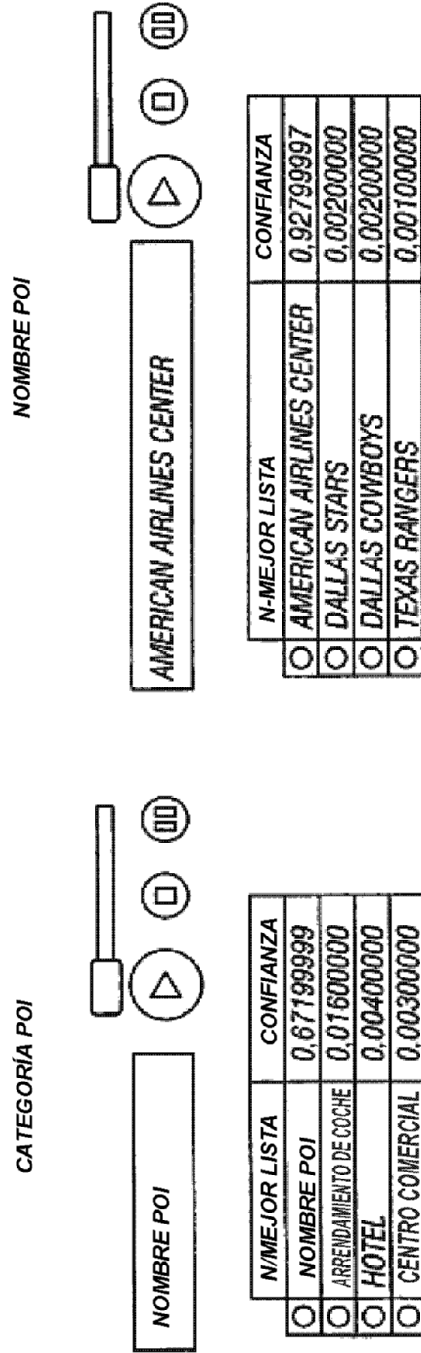
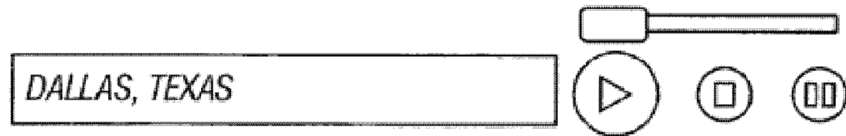


FIG. 2

CIUDAD ESTADO



	<i>N-MEJOR LISTA</i>	<i>CONFIANZA</i>
○	<i>DALLAS, TEXAS</i>	<i>0,96799999</i>
○	<i>ALICE, TEXAS</i>	<i>0,01899999</i>
○	<i>ATLAS, TEXAS</i>	<i>0,00600000</i>
○	<i>WALLACE, TEXAS</i>	<i>0</i>

FIG. 3

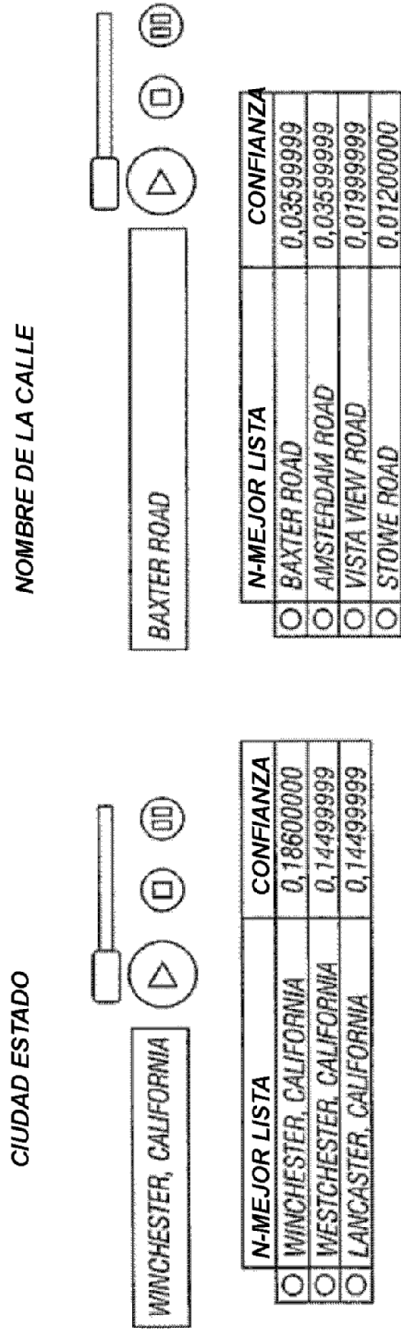


FIG. 4

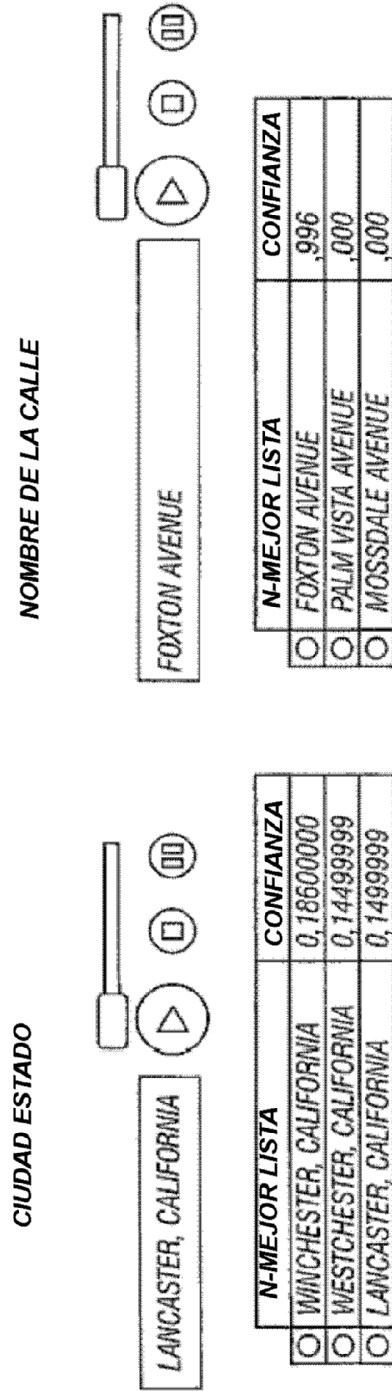


FIG. 5