

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 527 073**

51 Int. Cl.:

G06F 3/01

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.03.2009 E 09723879 (4)**

97 Fecha y número de publicación de la concesión europea: **01.10.2014 EP 2277098**

54 Título: **Sistema y procedimiento para la evaluación de gestos multidimensional**

30 Prioridad:

26.03.2008 DE 102008015730

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

19.01.2015

73 Titular/es:

**MICROCHIP TECHNOLOGY GERMANY II GMBH &
CO. KG (100.0%)
Friedrichshafener Strasse 3
82205 Gilching, DE**

72 Inventor/es:

**RICHTER, WOLFGANG y
AUBAUER, ROLAND**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 527 073 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

Sistema y procedimiento para la evaluación de gestos multidimensional

Campo de la invención

5 La invención se refiere a un sistema y a un procedimiento para la generación de señales de control mediante gestos, en particular gestos manuales y/o dactilares que lleva a cabo como tales un usuario enfrente de un equipo de detección.

Antecedentes de la invención

10 Por la solicitud internacional de patente PCT/EP2007/003555, que tiene su origen en el solicitante, publicada como documento WO2007/121977, es conocido un sistema de control para un vehículo que comprende una disposición de electrodos mediante la cual puede registrarse el movimiento espacial de una mano del conductor o copiloto, y conforme a este movimiento pueden inducirse operaciones de conexión y entrada, por ejemplo para el manejo de un aparato de navegación.

15 Por la solicitud de patente DE 10 2007 016 408.6, que también tiene su origen en el solicitante, es conocido un teléfono móvil que está dotado de un equipo de sensor mediante el cual pueden efectuarse sin contacto operaciones de entrada con un dedo del usuario.

Además, por la solicitud de patente DE 10 2007 039 163.5, que tiene su origen en el solicitante, es conocido un ratón de ordenador que posibilita efectuar funciones de control del ratón y del menú mediante gestos manuales enfrente del ratón del ordenador.

20 Por la solicitud de patente DE 10 2007 036 636.3, que tiene su origen en el solicitante, es conocida una muñeca que está provista como tal de varios electrodos mediante los cuales se puede registrar una detección de los gestos de un niño que está jugando enfrente de la muñeca y, en consecuencia, se puede llevar a cabo una activación de funciones de reproducción del habla de la muñeca así como, eventualmente, una activación de componentes mecatrónicos de la muñeca. Mediante esta técnica se hace posible llevar a cabo un diálogo que parece particularmente realista y no monótono con el juguete.

25 En los sistemas anteriormente mencionados, el registro del movimiento espacial de una mano se realiza enfrente del sistema de detección sobre la base de efectos de interacción de campos eléctricos. Para esto, el sistema de detección está dotado de electrodos que están integrados como tales en redes LC y proporcionan señales que dan información acerca de la distancia de la mano al electrodo. Mediante estas señales puede calcularse el curso de trayectoria de un gesto. A partir del curso de trayectoria calculado, a su vez, pueden determinarse gestos.

30 El cálculo de informaciones de distancia y de curso puede hacerse mediante un circuito electrónico en el que están depositados cursos de gestos característicos en una memoria electrónica.

Por la patente estadounidense US 1.661.058 es conocido un procedimiento y un dispositivo para la generación de sonidos. Por el documento US 6.137.042 es conocido un dispositivo de visualización para música que se generó con un dispositivo de este tipo.

Objetivo de la invención

35 La invención se basa en el objetivo de conseguir soluciones mediante las cuales se haga posible hacer un registro y una interpretación fiables de gestos con una complejidad de hardware lo menor posible y procesar de manera fiable un espectro amplio de gestos de entrada.

Solución de acuerdo con la invención

40 Este objetivo se resuelve de acuerdo con la reivindicación 1 independiente. Las reivindicaciones dependientes caracterizan otras configuraciones. De acuerdo con distintas formas de realización, puede facilitarse un procedimiento y un sistema para llevar a cabo un procesamiento de señal de señales de entrada que se generan como tales mediante gestos manuales que se ejecutan por parte de un usuario enfrente de un sistema de detección.

45 De este modo se hace posible de manera ventajosa efectuar también un reconocimiento de gestos y una evaluación recurriendo a sistemas de circuito y evaluación probados y desarrollados para el ámbito del reconocimiento del habla. En este caso, son posibles aplicaciones en las que los conceptos y conjuntos de chip desarrollados para el reconocimiento del habla se usan exclusivamente para la interpretación de gestos. Además, es posible dotar sistemas de reconocimiento del habla adicionalmente de la función del reconocimiento de gestos, pudiendo actuar la entrada de micrófono también directamente como entrada para las señales gestuales. El concepto de acuerdo con la invención de la transformación de cursos de trayectoria de gestos en secuencias a modo de sonidos puede realizarse en sistemas de ordenadores extendidos y otros sistemas en los que puede implementarse reconocimiento del habla, en particular microcontroladores, con relativamente poca complejidad de hardware adicional. Las secuencias a modo de sonidos generadas de acuerdo con la invención pueden generarse de tal manera que las mismas se acoplen en sistemas

correspondientes, eventualmente en superposición con señales de micrófono con un conector intermedio en la entrada acústica del sistema. Mediante este concepto se hace posible ampliar programas de reconocimiento del habla de manera particularmente ventajosa con funciones de interpretación gestual. Los electrodos de registro previstos para la grabación de gestos o para el registro del movimiento espacial de la mano que realiza el gesto o, eventualmente, también solo de dedos, pueden estar integrados en particular en unos auriculares, un teclado, una alfombrilla para el ratón, un ratón de ordenador o también un marco de pantalla. Los electrodos de registro también pueden estar integrados en otras estructuras que son adecuadas como punto de referencia típico para un registro de gestos. En este caso, en el ámbito de los vehículos son adecuadas en particular estructuras en el ámbito del tablero de instrumentos, el volante, la consola central y los revestimientos de puertas.

- 5
- 10 El equipo de sensor de acuerdo con la invención también puede estar diseñado de tal modo que mediante el mismo puedan registrarse gestos principalmente estáticos, por ejemplo formas de mano estáticas, y convertirse en secuencias de sonido suficientemente claras para esto. Los gestos estáticos de este tipo pueden realizarse, por ejemplo, como "mano abierta", "puño con pulgar extendido", "símbolo de la victoria". En estos gestos estáticos puede generarse una secuencia de sonido determinada mediante las señales de sensor generadas durante la fase estática.
- 15 También es posible reproducir con sonidos el procedimiento de movimiento típico para la composición y la descomposición de los gestos estáticos y tomarlo en consideración en el reconocimiento de estos gestos.

Mediante la conversión de las señales de curso de trayectoria que están correlacionadas con el gesto en secuencias de señal a modo de sonido también se posibilita una transmisión de las informaciones gestuales mediante sistemas de telecomunicaciones y sistemas VOIP dentro del intervalo de frecuencias de datos sonoros. Con esto, las informaciones gestuales puede facilitarse usando sistemas de transmisión del habla también a un lugar alejado del operador, es decir, del usuario que gesticula. Mediante el concepto de acuerdo con la invención también puede realizarse una conversión de un habla de gestos manuales a una lengua con sonido.

20

Pueden calcularse y depositarse para la evaluación patrones de reconocimiento para los gestos individuales mediante procedimientos de enseñanza o también para determinados cursos de trayectoria. De acuerdo con la invención, los gestos que proceden de seres vivos (preferentemente seres humanos) se registran con ayuda de un detector y se convierten en series de tonos (secuencias de sonidos). Las mismas pueden analizarse y evaluarse entonces con procedimientos actuales de procesamiento de series de habla o de tonos. Las órdenes verbalizadas pueden, eventualmente, paralelamente a esto, llegar al mismo procesamiento mediante el mismo canal y pueden evaluarse con los gestos de manera conjunta, individual o uno tras otro.

25

Preferentemente, las secuencias de tonos generadas por el circuito de reconocimiento de gestos se generan en la misma banda de frecuencia que la entrada de habla, de modo que también pueden usarse los procedimientos de filtro existentes en el procesamiento del habla. Sin embargo, también es posible escoger otro intervalo de frecuencias, por ejemplo desplazado hacia el intervalo marginal del intervalo de frecuencias procesado. De manera similar o exactamente como las órdenes de habla, los gestos pueden entonces entrenarse en el sistema y, al repetirlos, realizarse en funciones que se guardaron por ejemplo en una tabla. De este modo, puede usarse el mismo hardware y software que en el procesamiento del habla a excepción del transformador de series de tonos en gestos, lo cual significa una ventaja económica y técnica en términos del circuito con respecto a los sistemas que trabajan por separado. A este respecto, los respectivos gestos pueden llevarse a cabo delante de un equipo de registro tridimensionalmente en un determinado intervalo de tiempo y son en gran medida independientes del usuario.

30

35

Fundamentalmente, el concepto de acuerdo con la invención es adecuado para las más diversas técnicas del registro de gestos, tales como técnicas ópticas, capacitivas o de procesamiento de imágenes. Es ventajoso que la determinación de los gestos se realice en secuencias y para cada secuencia pueda generarse un determinado tono.

40

Por lo tanto, un gesto sencillo o complejo genera un serie de tonos diferente más o menos duradera. Como una persona no puede introducir nunca el mismo gesto dos veces exactamente igual, el sistema presenta preferentemente un reconocimiento tolerante. Esto sucede preferentemente basado en un programa. Un software adecuado para esto puede facilitarse de acuerdo con la invención usando los procedimientos de reconocimiento y de interpretación previstos en sistemas de procesamiento del habla, en particular cuando la serie de tonos que está correlacionada de acuerdo con la invención con el gesto, en particular con la trayectoria gestual, tiene una forma similar a una palabra o una frase hablada.

45

50 Recepción de las señales gestuales

La recepción de las señales gestuales se realiza preferentemente mediante uno o varios sensores capacitivos (de campo eléctrico). Estos sensores están estructurados preferentemente de tal modo que los mismos detectan modificaciones en un campo eléctrico generado artificialmente y, en consecuencia, proporcionan señales que están correlacionadas de manera suficientemente estrecha con la posición o el movimiento de la mano o de un dedo. Preferentemente se realiza un procesamiento previo de las señales gestuales. Este procesamiento previo puede estar ajustado de manera ventajosa a un gesto ya reconocido. El registro de gestos puede realizarse con división/desmultiplexado de la señal proporcionada por el respectivo sensor o en el uso preferente de varios sensores, de las señales.

55

La interpretación de los gestos se realiza preferentemente al generarse señales de sensor para al menos dos, preferentemente tres coordenadas espaciales. Estas señales de sensor se obtienen preferentemente en un sistema de medición multicanal, pudiendo registrarse las señales de sensor de manera multicanal paralelamente o en un múltiplex temporal. También es posible una excitación del sistema en un múltiplex temporal o incluso múltiplex de frecuencia.

5 La conversión de las señales moduladas por portador (recepción de electrodos) de los distintos canales de sensor puede realizarse en particular mediante detección de curvas envolventes o rectificación y filtración de pasabajos (frecuencia de movimiento < frecuencia de pasabajos < frecuencia portadora) en señales de recepción de baja frecuencia, cuya amplitud es proporcional a la aproximación/alejamiento del sensor de recepción. Es posible reconocer un principio de gesto y un final de gesto de la siguiente manera:

10 a. Detección de aproximación: como gestos se reconocen las señales de los sensores de campo eléctrico en los cuales al menos una señal de sensor excede o queda por debajo de un nivel determinado o que se ajusta de manera adaptativa.

b. Detección de movimientos: como gestos se reconocen señales de los sensores de campo eléctrico con una velocidad de modificación mínima. Para esto es ventajoso formar la derivada matemática de estas señales. Como alternativa, también es posible llevar a cabo una filtración de pasabajos que se corresponde con la derivada.

15 c. Las medidas anteriormente mencionadas de acuerdo con a. y b. también pueden llevarse a cabo en combinación.

De acuerdo con una forma de realización particularmente preferente de la invención se realiza una normalización de las señales gestuales, por ejemplo mediante la sustracción del valor medio temporal o de la parte de señal igual de todas las señales de sensor de las respectivas señales de sensor individuales.

Además, es posible efectuar una transformación de coordenadas de las señales de sensor detectadas, de modo que se reproduzcan los gestos correspondientes en un plano de referencia espacial.

Preferentemente, este plano de gestos de referencia se encuentra esencialmente paralelo con respecto a las superficies de electrodos y perpendicular con respecto al eje de sensor, de modo que los gestos correspondientes se reproducen siempre a una distancia determinada con respecto al origen del sensor. El centro de gravedad de la superficie abarcada mediante los gestos se encuentra ventajosamente en el eje de sensor.

La conversión de las señales de sensor multidimensionales transformadas de este modo se realiza preferentemente mediante una conversión de frecuencia de tensión (por ejemplo, VCO) en series de tonos adecuadas que se encuentran en el intervalo de frecuencias del habla de típicamente 50 Hz - 8000 Hz o 300-3400 Hz (intervalo de banda telefónica), de modo que las distintas señales de sensor (típicamente 3) se trasladan a una única señal gestual que debe evaluarse (1 canal). La misma se le facilita al reconocedor de gestos en la siguiente etapa para la evaluación.

Las señales de sensor pueden generarse como señales de múltiplex temporal. La conversión o transformación de las señales de sensor mediante cambio de VCO conduce preferentemente a señales en bandas de frecuencia respectivamente diferentes en el intervalo de frecuencias del habla.

35 Reconocimiento de gestos

El reconocimiento de los gestos se realiza preferentemente usando un reconocedor de patrones de DTW (alineación temporal dinámica, Dynamic Time Warping) para reconocer gestos completos, de manera análoga al reconocimiento de palabras en el caso de reconocedores del habla. Estos tipos de reconocedores se caracterizan por una selectividad suficiente para numerosos casos de aplicación con respecto a gestos realizados de manera similar y, con esto, una tasa de reconocimiento suficiente con un espectro gestual más pequeño con gestos relativamente marcados.

Como alternativa al planteamiento anteriormente mencionado del reconocimiento de patrones de DTW también es posible usar un procedimiento conocido también como concepto de reconocimiento del habla del reconocedor del modelo oculto de Markov (HMM). En este tipo de reconocimiento del habla, las palabras se descomponen en fonemas, sonidos casi estacionarios atómicos de la lengua. De manera análoga, los gestos se descomponen en fragmentos de gestos que se representan en el HMM mediante estados. Como fragmentos de gestos pueden usarse partes discretionales de un gesto, es decir, gestos discretionales, preferentemente dedo/mano/brazo. En particular, los movimientos rectilíneos o curvados, modificaciones de la orientación (por ejemplo giro de la mano), modificaciones de la forma (como abrir la mano), así como las formas respectivamente derivadas e integradas de forma discrecional de estas modificaciones (longitudes, velocidades de movimiento, etc.) Estos fragmentos gestuales se reconocen individualmente uno tras otro y el reconocedor de HMM los asigna de nuevo a los gestos completos definidos (en un procedimiento de entrenamiento o de descripción) (para este ejemplo: gesto L).

Además, las señales generadas como secuencias de señal a modo de sonidos y correlacionadas con el gesto manual de acuerdo con la invención también pueden evaluarse mediante un reconocedor de redes neuronales y clasificador de fragmentos gestuales de manera análoga al reconocimiento del habla (clasificador de fonemas).

Las medidas anteriormente mencionadas, así como procedimientos de reconocimiento del habla extendidos de alguna otra forma también pueden realizarse en combinación para evaluar las señales gestuales generadas a partir de las secuencias de señal a modo de sonido de acuerdo con la invención e interpretarse a este respecto.

Diferenciación de fase de entrenamiento y de reconocimiento

- 5 Los gestos que deben reconocerse se entrenan preferentemente al menos una vez, eventualmente también repetidas veces, y se archivan en el reconocedor de gestos como patrones de referencia. Los gestos del mismo enunciado que, sin embargo, también se reconocen como demasiado diferentes en el entrenamiento, preferentemente se desestiman. Los gestos no desestimados, suficientemente parecidos y, por tanto, claros, pueden resumirse mediante formación de valor medio de las correspondientes señales de referencia.
- 10 Como alternativa al procedimiento de entrenamiento, también puede realizarse una entrada gráfica de gestos de referencia (flecha de derecha a izquierda, de arriba a abajo, círculo, etc.), por ejemplo basada en ratón/menú. Para estos gestos introducidos pueden determinarse las señales de sensor esperadas correspondientemente y depositarse en el reconocedor de gestos como patrón de referencia.

Fase de reconocimiento

- 15 Durante la fase de reconocimiento se comparan preferentemente las señales gestuales con los patrones de señal de referencia archivados en el procedimiento de entrenamiento/descripción y se emite el gesto más probable. Si las señales gestuales difieren demasiado de todas las señales de referencia, se emite un "no reconocidas". En este caso, dependiendo de la diferencia, puede definirse un valor umbral con el que pueden adaptarse la probabilidad de detecciones erróneas (el gesto reconocido se asigna erróneamente) y la desestimación equivocada (en inglés, "miss", es decir, un gesto no se reconoce y se rechaza) en función de la aplicación. Con un reconocedor de HMM se realiza la búsqueda del mejor gesto ventajosamente mediante un algoritmo de Viterbi.
- 20

Mediante el concepto de acuerdo con la invención se hace posible suministrar señales gestuales a un sistema de reconocimiento del habla y evaluarlas de manera sorprendentemente eficaz mediante los procedimientos de reconocimiento ya desarrollados para el reconocimiento del habla. La sensórica con uso de campos eléctricos de baja frecuencia (sensórica de campo eléctrico) en el intervalo de hasta aproximadamente 100 kHz posibilita un reconocimiento seguro de movimientos a la distancia de típicamente hasta 50 cm de los correspondientes electrodos de detección.

Para una representación espacial del gesto se usan preferentemente al menos tres electrodos que proporcionan, por ejemplo, tres aportaciones de tonos que están correlacionadas respectivamente con la dinámica de la modificación de la distancia. Para reconocer en particular también la forma de la mano y aumentar en total la resolución y crear, eventualmente, una cierta redundancia, es ventajoso usar más de tres electrodos y, con esto, crear un sistema, eventualmente, sobredeterminado con respecto a las coordenadas.

Breve descripción de las figuras

A partir de la siguiente descripción junto con el dibujo resultan otras particularidades y características de la invención. Muestra:

La **Figura 1**, una representación esquemática para ilustrar el concepto de reconocimiento de gestos de acuerdo con la invención en el que se registra el movimiento espacial de una mano mediante tres electrodos de emisión y un electrodo de recepción y la señal registrada mediante el electrodo de recepción se genera o modifica, de tal manera que la misma puede reconocerse de manera suficientemente clara mediante un sistema de reconocimiento del habla;

La **Figura 2**, otra representación esquemática para ilustrar el concepto de reconocimiento de gestos de acuerdo con la invención;

La **Figura 3**, una representación fotográfica para ilustrar un circuito de registro de acuerdo con la invención mediante el cual pueden generarse señales a modo de sonidos en correlación con un gesto efectuado enfrente de una disposición de electrodos.

La **Figura 4**, un esquema de conexiones para ilustrar la estructura del circuito de registro de acuerdo con la Figura 3;

La **Figura 5**, un diagrama de bloques para ilustrar el depósito de gestos que deben reconocerse mediante entradas gráficas;

La **Figura 6**, un diagrama de bloques para ilustrar un sistema de ejemplo con varios emisores y un receptor en el modo de aprendizaje;

La **Figura 7**, un diagrama de bloques para ilustrar un sistema de ejemplo con varios receptores y un emisor en el modo de aprendizaje;

La **Figura 8**, un diagrama de bloques para ilustrar otro sistema de ejemplo con varios emisores y un receptor en el modo de reconocimiento;

La **Figura 9**, un diagrama de bloques para ilustrar otro sistema de ejemplo con un emisor y varios receptores en el modo de reconocimiento;

5 La **Figura 10**, un esbozo para ilustrar un sistema informático con un equipo de reconocimiento de gestos con el cual el acoplamiento de las señales gestuales se realiza en una entrada de micrófono y el reconocimiento de gestos se hace mediante un programa de reconocimiento del habla con vocabulario ampliado;

10 La **Figura 11**, una representación esquemática para ilustrar otras particularidades del concepto de acuerdo con la invención de la generación de señales a modo de sonidos y que están correlacionadas con el perfil gestual y la dinámica gestual que se identifican mediante un sistema de reconocimiento del habla con vocabulario ligeramente ampliado.

Descripción detallada de las figuras

15 La Figura 1 ilustra una primera variante de un circuito de acuerdo con la invención. Un componente electrónico que es fundamentalmente adecuado y está dotado para el procesamiento del habla (por ejemplo, el componente RSC 4128) contiene un controlador programable. Al mismo se le ordena por firmware generar una señal cambiante en al menos un puerto de E/S (serie 1-0). Esta señal puede emitirse ópticamente, por ejemplo mediante diodos luminosos, o también capacitivamente, por ejemplo sobre una superficie de acoplamiento. En el último caso, genera un campo alterno eléctrico. Un detector dotado preferentemente de un intensificador de transimpedancia puede recibir este campo o una cantidad de luz que procede del diodo luminoso. El mismo se cambia a una tensión continua que acciona un oscilador VCO controlado por tensión. Una modificación de la intensidad de luz o del campo eléctrico, por ejemplo introduciendo extremidades humanas en la zona de registro, modifica la altura de sonido. La misma puede descender o aumentar. Si el pin de controlador correspondiente solo se activa por poco tiempo (por ejemplo, 40 milisegundos), también la modificación de tonos tendrá lugar solo en este tiempo. Finalmente, el respectivo VCO puede encenderse o apagarse mediante otro pin de E/S, de modo que en punto muerto no pueda oírse ningún tono. La serie de tonos generada por aproximación se suministra al circuito de análisis previsto realmente para el procesamiento del habla, que típicamente está compuesto por una combinación de hardware y software, a saber, preferentemente en el mismo chip, que también genera la señal de campo o luminosa. Si se quieren registrar varias dimensiones, se prevén, en caso necesario, varios electrodos o diodos luminosos que presentan un campo. Esto puede suceder mediante otros pines de E/S del mismo controlador, que pueden activarse uno tras otro (=secuencia).

20 El procesamiento de señales y la evaluación se realizan de acuerdo con la invención mediante técnicas de análisis en sí extendidas para el procesamiento del habla. Estas técnicas de reconocimiento del habla se usan de acuerdo con la invención para el reconocimiento de gestos al formarse al principio, en correlación con los gestos, secuencias de señales a modo de sonidos similares al habla. La evaluación de estas secuencias de señales a modo de sonidos es comparativamente fiable, ya que los tonos pueden generarse de tal modo que los mismos aparezcan, por ejemplo, más bien como una serie de vocales y puedan faltar consonantes y sibilancias (pero no tienen por qué). Con esto también es posible mezclar entre sí órdenes lingüísticas y órdenes gestuales y realizarlas al mismo tiempo o una tras otra. Con ayuda de redes neuronales u otros algoritmos de aprendizaje es posible entrenar un sistema de este tipo y ajustar umbrales de tolerancia.

30 En el componente electrónico representado en la Figura 1 están acoplados en sus puertos de E/S, por ejemplo, tres electrodos que ceden campo (por ejemplo, superficies de cobre). Cerca de los mismos se encuentra un electrodo E-in que absorbe campo. El mismo está representado en un búfer (por ejemplo, intensificador de transimpedancia), un circuito de rectificador siguiente que está compuesto por diodo y condensador y un oscilador controlado por tensión (VCO, en este caso un generador de picos de impulsos transitorios parásitos). Mediante este circuito es posible generar series de tonos (secuencias de tonos) que se modifican en su altura mediante aproximación, preferentemente de extremidades humanas. Mediante un circuito secuencial (software o hardware) se activan ahora las superficies de electrodos individuales E1 a E3 una tras otra, a saber, con un campo alterno de 1-0 series de longitud temporal, por ejemplo 100 kHz para respectivamente 20 milisegundos. Se generan campos alternos eléctricos. La introducción de extremidades puede debilitar el campo con respecto a la superficie de electrodos de entrada Ke o actuar como puente. Ambos efectos modifican la altura de sonido del VCO encendido. Su salida está devuelta al chip en el que está integrado el procesamiento del habla. Con esto se puede entrenar y evaluar fácilmente una serie de movimientos. En la misma (o en otra entrada) también puede conectarse un micrófono que procesa órdenes lingüísticas de la misma manera.

35 La Figura 2 muestra la realización sencilla del detector con transistores. También es posible integrar este circuito en el chip y colocar de manera externa únicamente, por ejemplo, el condensador de carga si es necesario. Un VCO también puede realizarse como función de software en el chip.

La presente invención puede realizarse de manera extraordinariamente económica y es adecuada en particular para la integración en juguetes para controlar los mismos con habla y gestos y reforzar constantemente el atractivo lúdico. El usuario siempre puede entrenar nuevas órdenes de habla y/o gestos.

Sin embargo, la invención es adecuada también para fines técnicos así como para el ámbito del procesamiento de datos y textos, así como para aparatos de comunicación. Por ejemplo, puede colocarse en un electrodoméstico, por ejemplo una secadora, una disposición de este tipo dentro del tambor. Al introducir ropa se generan otras series de señales diferentes a cuando, por ejemplo, un niño entra en el tambor y se mueve en él (o un animal). Adicionalmente, mediante la invención también se presta atención también a ruidos. Por lo tanto, es posible reconocer seres vivos en electrodomésticos antes de que los mismos sufran daños por el funcionamiento. Mediante este circuito también es posible (en caso de funcionamiento temporal) despertar a los mismos de un estado de sueño y/o activarlos solo mediante una determinada orden de habla o un gesto determinable (circuito de activación).

La disposición de electrodos descrita puede integrarse también en muebles, por ejemplo para analizar posturas sentadas o gestos en unión con estos muebles o reaccionar a órdenes lingüísticas, por ejemplo, camillas para masajes regulables, etc. Puede hacerse que los aparatos de navegación sean más fáciles de manejar con ayuda de la invención. La introducción y el control por gestos permiten también funciones de zoom que solo se pueden realizar difícilmente por pantalla táctil justamente durante un trayecto. Cuando varias personas desean manejar un sistema de este tipo, es posible prever o analizar órdenes que dependen del usuario a partir de cuya dirección se realizan gestos introducidos.

La invención puede emplearse ventajosamente en todo lugar donde, debido a una elevada presión de coste, no puedan usarse circuitos individuales más complejos.

Los diagramas de bloques de acuerdo con las Figuras 5 a 9 son comprensibles por sí mismos. La Figura 5 muestra en este caso un diagrama de bloques para ilustrar el depósito de gestos que deben reconocerse mediante entradas gráficas. La Figura 6 muestra un diagrama de bloques para ilustrar un sistema de ejemplo con varios emisores y un receptor en el modo de aprendizaje. La Figura 7 muestra un diagrama de bloques para ilustrar un sistema de ejemplo con varios receptores y un emisor en el modo de aprendizaje. La Figura 8 muestra un diagrama de bloques para ilustrar otro sistema de ejemplo con varios emisores y un receptor en el modo de reconocimiento. La Figura 9 muestra un diagrama de bloques para ilustrar otro sistema de ejemplo con un emisor y varios receptores en el modo de reconocimiento.

Como alternativa a la técnica de detección descrita anteriormente con varios emisores y un receptor, o varios receptores y un emisor, también es posible accionar los electrodos individuales alternativamente como emisor y receptor y emitir el grado de puenteo entre los electrodos que actúan respectivamente como emisores o receptores, o emitir la modificación del grado de puenteo en forma de una señal a modo de sonido modulada correspondientemente.

En la Figura 10 está ilustrado un sistema de PED que comprende un ratón de ordenador 1 dotado de un sensor de gestos, así como unos auriculares 2 dotados también de un sensor de gestos. Tanto las señales por parte del ratón de ordenador 1 como por parte de los auriculares 2 e indicativas con respecto al movimiento de una mano H en las tres direcciones del espacio x, y, z se convierten en señales a modo de sonido y se dan directamente a la entrada de una tarjeta de sonido 3 de un PC. En este caso, el conector de ratón 4 está realizado como conector intermedio, de modo que en el mismo también se puede enchufar el conector de auriculares 5. Además, al conector de ratón 4 también está unido un conector de USB 6 mediante el cual pueden suministrarse los movimientos del ratón de manera en sí conocida al sistema de PC 7.

Como alternativa al cableado mostrado en el presente documento para una comprensión más fácil, la comunicación con el sistema de PC 7 puede realizarse también de manera inalámbrica. Los equipos de registro de gestos de acuerdo con la invención también pueden estar integrados en sistemas periféricos del PC de otro tipo o, por ejemplo en el caso de un portátil directamente en el mismo, por ejemplo en la pantalla o el panel táctil.

El sistema puede estar ajustado de tal modo que el mismo determine mediante qué equipo de entrada de gestos está realizada una entrada de los gestos. Al respectivo equipo de entrada de gestos pueden estar asignadas determinadas funciones. Por ejemplo, mediante el equipo de sensor de gestos del ratón puede llevarse a cabo principalmente la navegación dentro de un documento o dentro de archivos. Mediante el equipo de sensor de gestos de los auriculares pueden coordinarse en particular el formato de carácter, así como signos de puntuación y medidas de corrección. Eventualmente, ambos equipos de sensor de gestos pueden manejarse al mismo tiempo. El equipo de sensor de gestos del ratón puede dirigirse por ejemplo con la mano más ejercitada en motricidad fina (principalmente, la mano derecha). Con la mano libre restante puede dirigirse el segundo equipo de sensor de gestos, por ejemplo el equipo de sensor de gestos de los auriculares. Determinados gestos, por ejemplo gestos sencillos en forma de una modificación de distancia rectilínea enfrente del equipo de sensor de gestos del ratón pueden transmitirse como señal de tono duradero; los gestos más complejos con características de curso de trayectoria y de dinámica pronunciadas pueden transmitirse con un determinado retraso como "palabra única" cuando se reconoce una característica de fin de gesto. Eventualmente, el reconocimiento de la característica de fin de gesto también puede hacerse mediante funciones adicionales del equipo de convertidor antepuesto de acuerdo con la invención al sistema de reconocimiento del habla. La característica de fin de gesto puede consistir, por ejemplo, en que dentro de un espacio de tiempo del 10 % de una fase que había sucedido tiempo atrás causada por el gesto de entrada con dinámica pronunciada no aparece una dinámica de señal particular.

Mediante el concepto de acuerdo con la invención se hace posible procesar señales gestuales de manera sinérgica con señales de habla. En aplicaciones en las que no es necesaria una entrada de habla puede recurrirse a la herramienta de reconocimiento del habla exclusivamente para la interpretación gestual.

5 Sobre la base del concepto de acuerdo con la invención puede realizarse un sistema de reconocimiento gestual mediante juegos de chips o ASICS que se desarrollaron en sí para tareas de reconocimiento del habla de manera particularmente eficiente.

La presente invención puede usarse en particular en combinación con las técnicas que se infieren de las solicitudes de patente mencionadas al principio, que tienen su origen en el solicitante.

REIVINDICACIONES

1. Procedimiento para llevar a cabo un procesamiento de señales electrónico, de señales de entrada que se generan como tales mediante gestos manuales,
 - a) ejecutándose aquellos gestos manuales por parte de un usuario enfrente de un sistema de detección (E; electrodos Tx, Rx),
 - b) generándose las señales de entrada de tal modo que las mismas representan señales que están correlacionadas con el movimiento espacial de una mano o sus dedos,
 - c) transformándose en este caso las señales de tal manera que la banda de frecuencia de las señales resultantes está contenida en la banda de frecuencia de las señales de sonido accesibles para un sistema de reconocimiento del habla,
- 5 **caracterizado por que:**
 - d) se realiza una interpretación de los gestos manuales mediante el sistema de reconocimiento del habla sobre la base de aquellas señales resultantes usando procedimientos de análisis del habla.
- 10 2. Procedimiento de acuerdo con la reivindicación 1 en el que las señales de entrada se generan mediante un circuito de Theremin.
- 15 3. Procedimiento de acuerdo con la reivindicación 1 o 2, presentando el sistema de detección:
 - tres electrodos de emisión (E_{a1} , E_{a2} , E_{a3}) que generan respectivamente un campo alterno eléctrico en los que puede influirse mediante los gestos manuales y que detecta un electrodo de recepción (E_{in}), estando acoplado el electrodo de recepción (E_{in}) con un oscilador controlado por tensión (VCO) que genera una señal de entrada del sistema de reconocimiento del habla.
- 20 4. Procedimiento de acuerdo con la reivindicación 3, rectificándose la señal recibida por el electrodo de recepción (E_{in}) y controlando la señal rectificada el oscilador controlado por tensión (VCO).
- 25 5. Procedimiento de acuerdo con la reivindicación 3 o 4, sometándose los tres electrodos de emisión (E_{a1} , E_{a2} , E_{a3}) secuencialmente a impulsos de señal de corriente alterna.
6. Procedimiento de acuerdo con la reivindicación 5, teniendo los impulsos de señal de corriente alterna una frecuencia de 100 kHz y una duración de 20 ms.
7. Procedimiento de acuerdo con una de las reivindicaciones 1 a 6, aumentando en caso de un alejamiento de un punto de detección la altura de sonido de la señal para este punto de referencia.
8. Procedimiento de acuerdo con una de las reivindicaciones 1 a 6, descendiendo en caso de una aproximación a un punto de detección la altura de sonido de la señal para este punto de referencia.
- 30 9. Procedimiento de acuerdo con una de las reivindicaciones 1 a 8, formándose las señales de tal manera que las mismas describen la modificación de la distancia de la mano que realiza el gesto con respecto a un sistema de referencia que contiene al menos tres puntos de referencia.
10. Procedimiento de acuerdo con una de las reivindicaciones 1 a 9, reconociéndose un principio de gesto y un final de gesto de un gesto excediendo o quedando por debajo de un determinado nivel de señal al menos un sensor de campo eléctrico del sistema de detección (E; electrodos Tx, Rx).
- 35 11. Procedimiento de acuerdo con una de las reivindicaciones 1 a 10, reconociéndose un principio de gesto y un final de gesto de un gesto mediante una velocidad de modificación mínima de señales de sensor de campo eléctrico.
- 40 12. Sistema para llevar a cabo un procesamiento de señales electrónico de señales de entrada que se generan como tales mediante gestos manuales, que presenta:
 - un dispositivo de detección de gestos (E; electrodos Tx, Rx) para generar al menos una señal de entrada que está correlacionada con el movimiento espacial de una mano o de sus dedos;
 - un dispositivo de transformación (VCO) para transformar la al menos una señal de entrada en señales resultantes, estando contenida la banda de frecuencia de las señales resultantes en la banda de frecuencia de señales de sonido accesibles para un sistema de reconocimiento del habla,
- 45 **caracterizado por:**
 - un sistema de reconocimiento del habla al que se le suministran las señales resultantes para la interpretación de los gestos manuales usando procedimientos de análisis del habla.
- 50 13. Sistema de acuerdo con la reivindicación 12, presentando el sistema de detección (E; electrodos Tx, Rx) tres electrodos de emisión (E_{a1} , E_{a2} , E_{a3}) para generar campos alternos eléctricos en los que pueden influir los gestos manuales y un electrodo de recepción (E_{in}), estando acoplado el electrodo de recepción con un oscilador controlado

por tensión (VCO) que genera una señal de entrada del sistema de reconocimiento del habla.

14. Sistema de acuerdo con la reivindicación 13 con un rectificador (CA/CC) dispuesto entre electrodo de recepción y oscilador controlado por tensión (VCO).

5 15. Sistema de acuerdo con una de las reivindicaciones anteriores 12-14, estando configurado el sistema para realizar el procedimiento de acuerdo con una de las reivindicaciones 1-11.

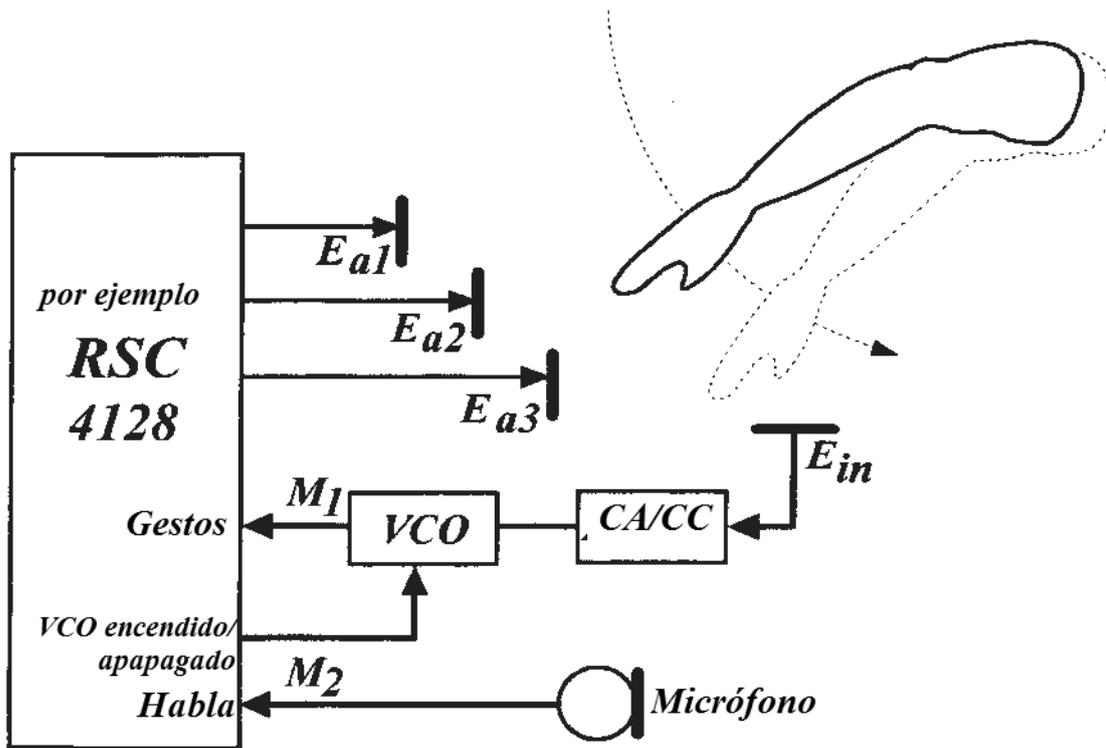
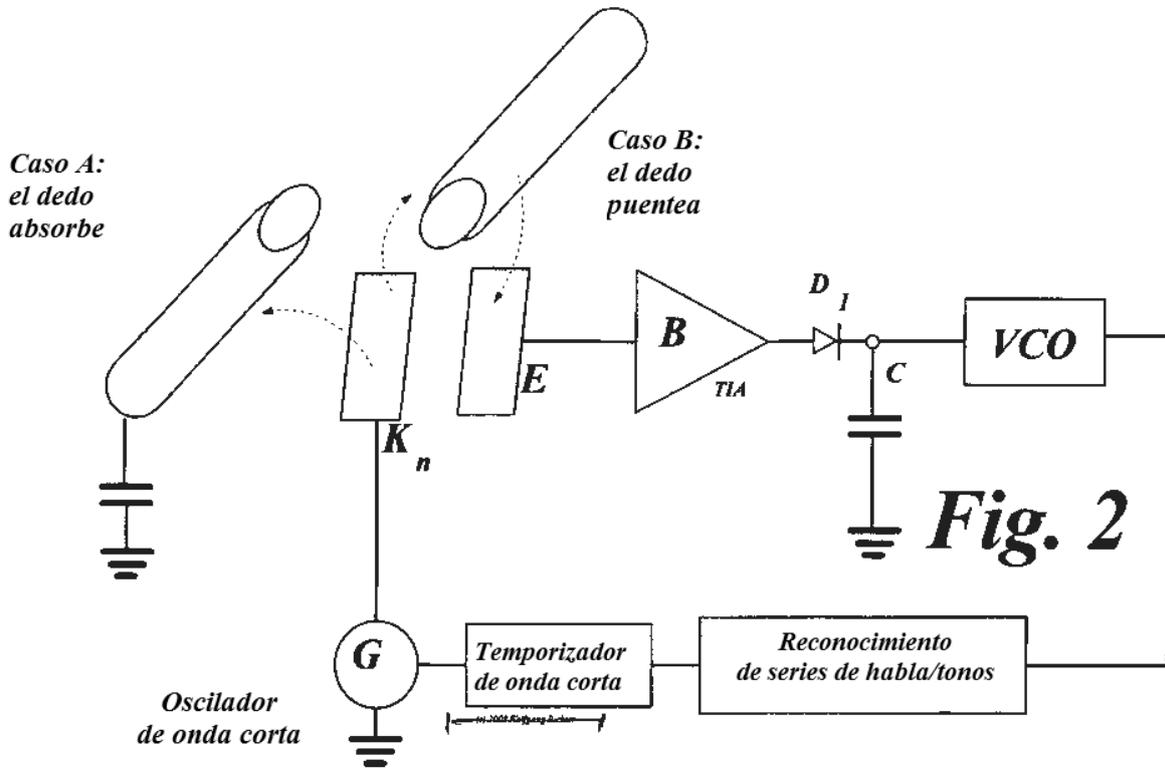


Fig. 1



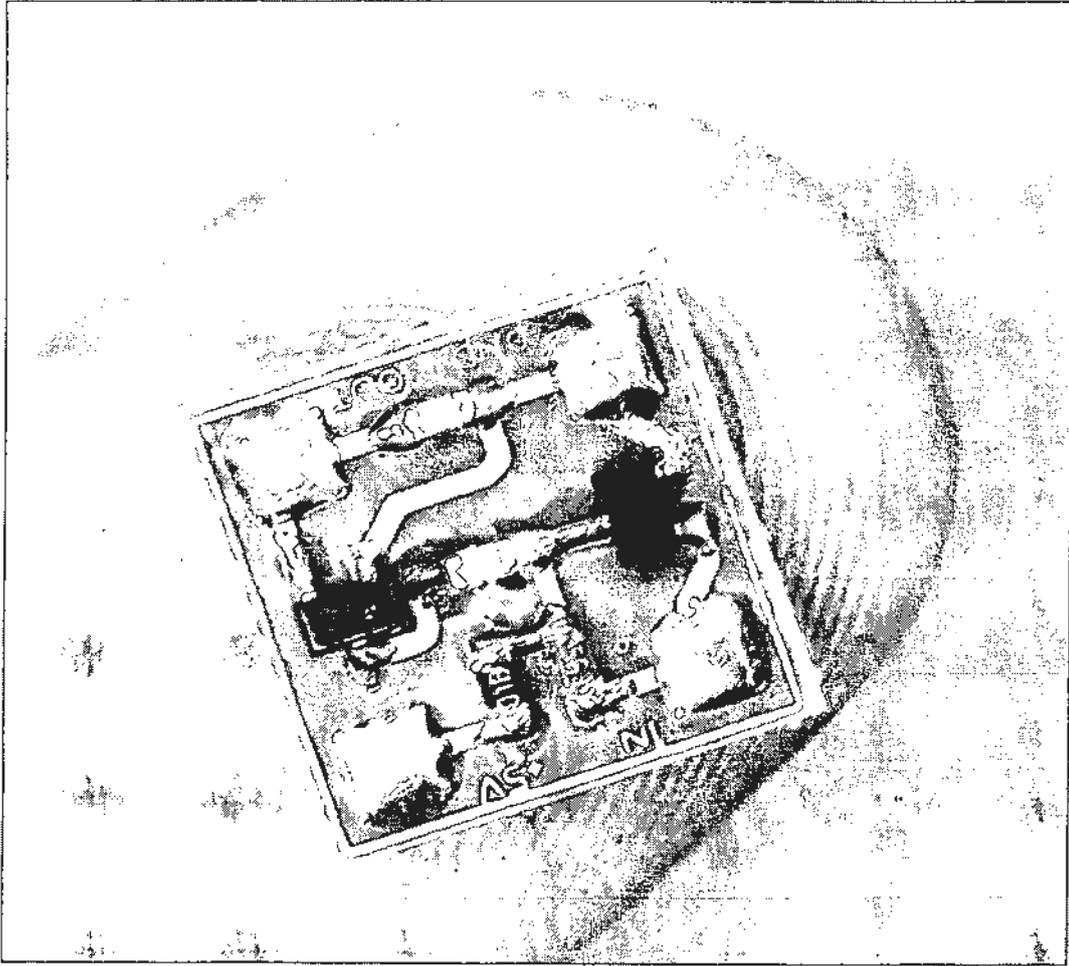


Fig. 3

Depósito de gestos que deben reconocerse mediante entrada gráfica



Fig. 5

Sistema de ejemplo con varios emisores y un receptor en el modo de aprendizaje

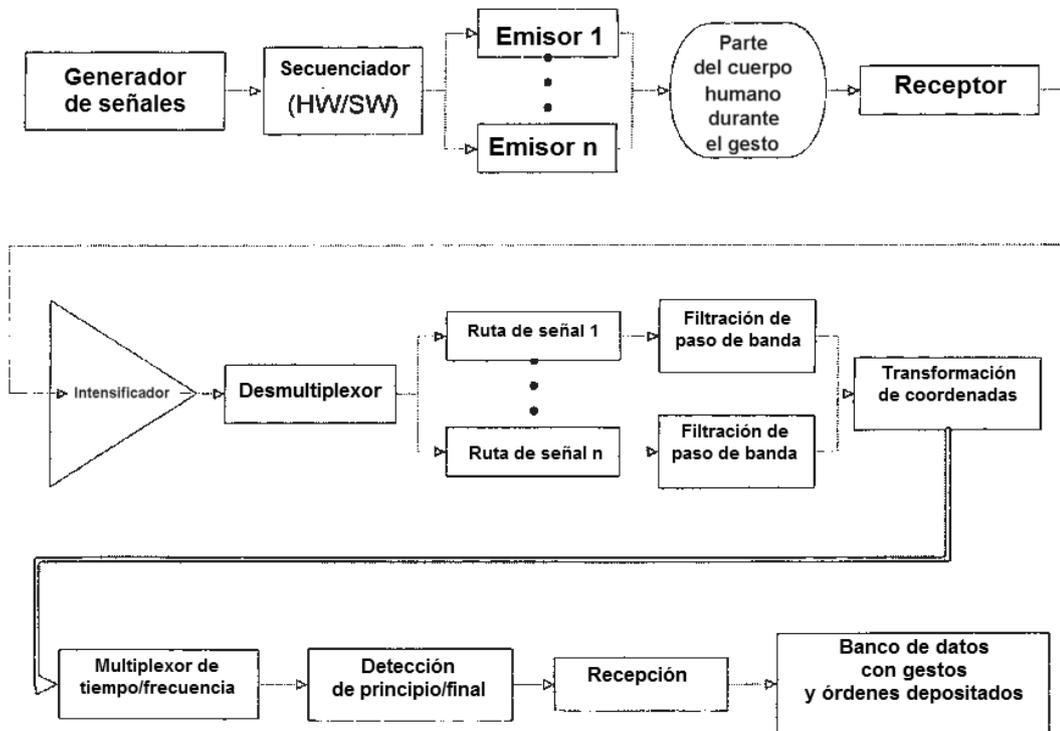


Fig. 6

Sistema de ejemplo con un emisor y varios receptores en el modo de aprendizaje

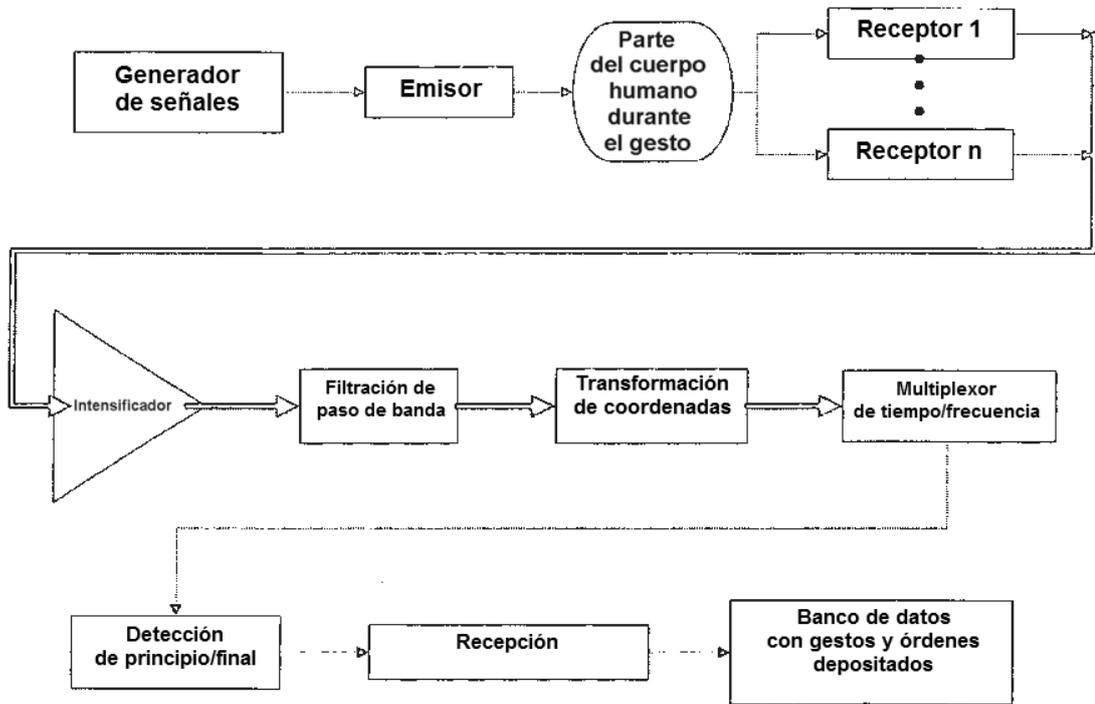


Fig. 7

Sistema de ejemplo con varios emisores y un receptor en el modo de reconocimiento

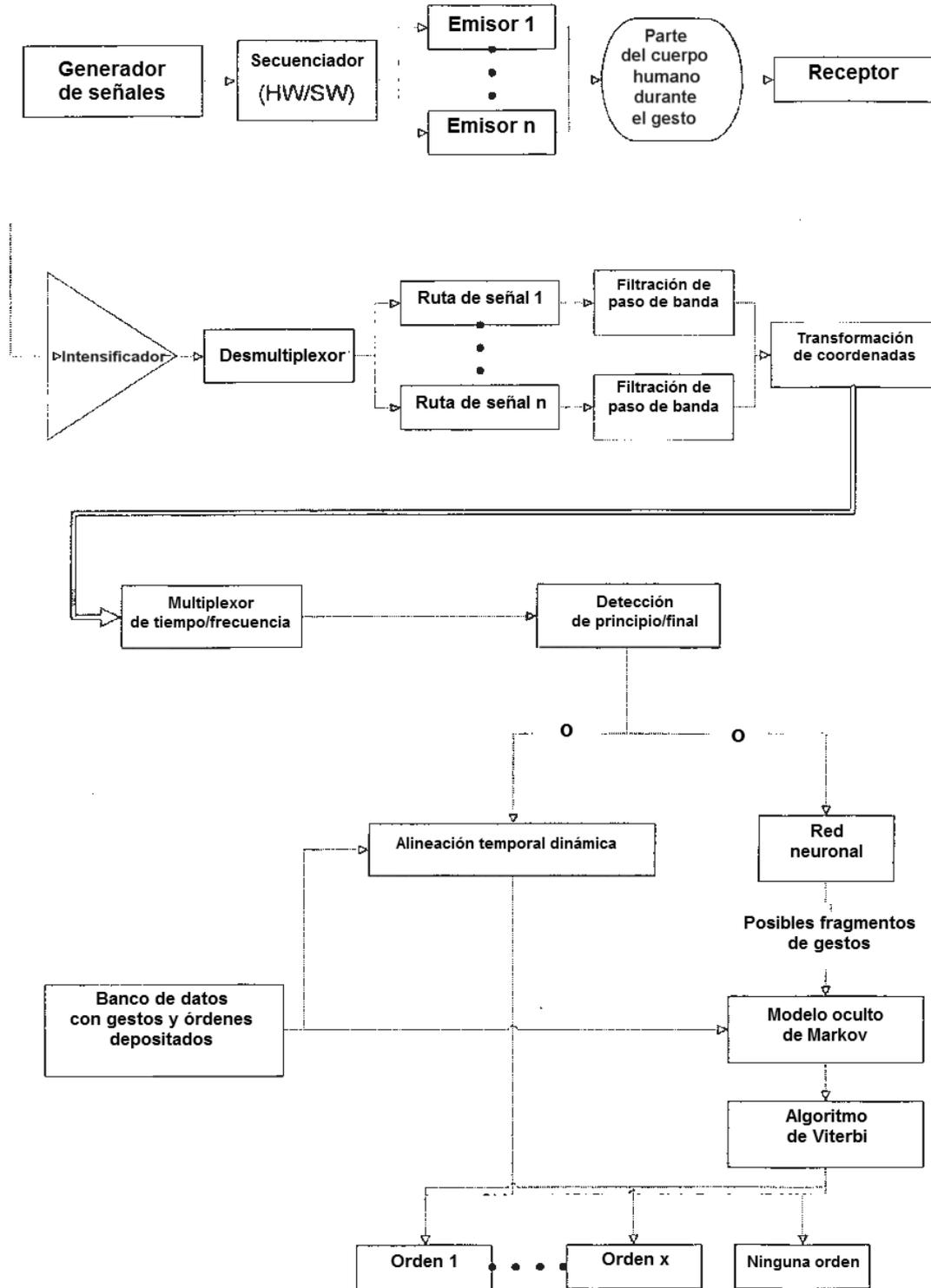


Fig. 8

Sistema de ejemplo con un emisor y varios receptores en el modo de reconocimiento

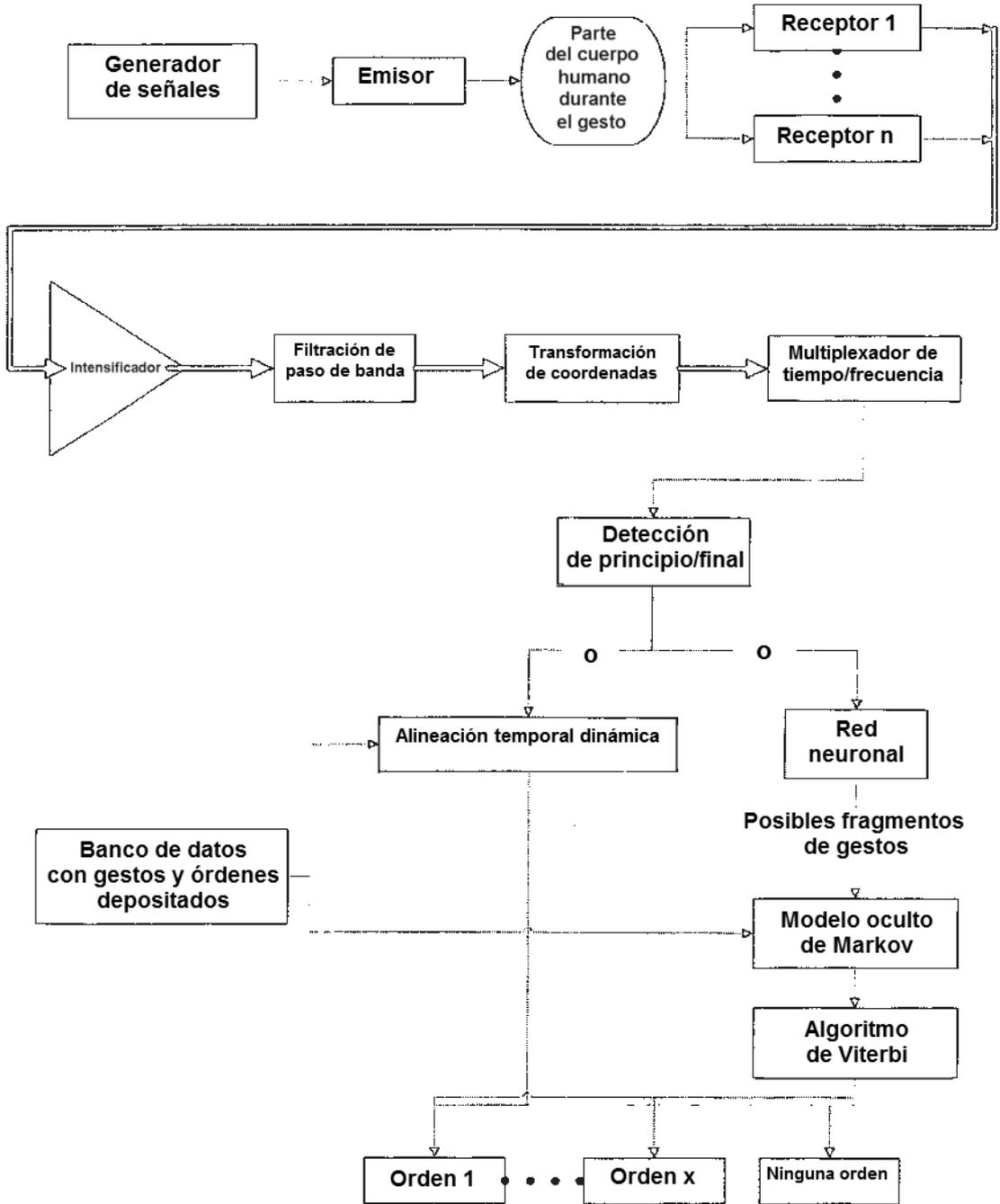


Fig. 9

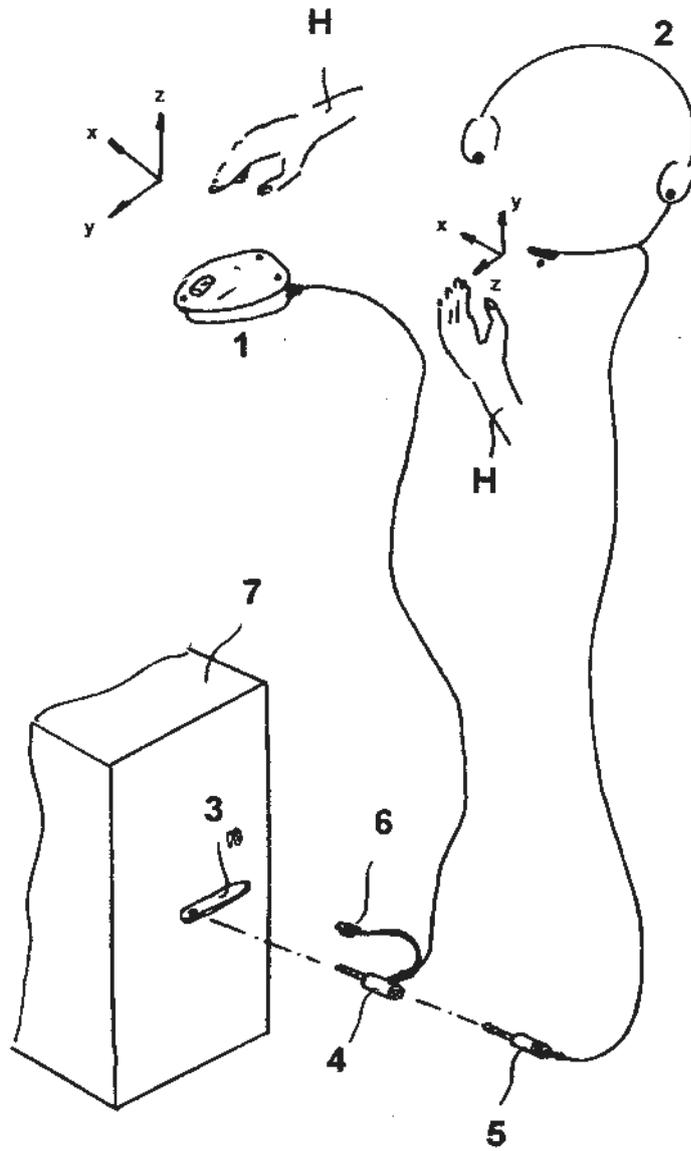


Fig. 10

Fig. 11

