

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 680 825**

51 Int. Cl.:

**G06F 3/01** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.03.2015 PCT/US2015/020178**

87 Fecha y número de publicación internacional: **26.11.2015 WO15179008**

96 Fecha de presentación y número de la solicitud europea: **12.03.2015 E 15760529 (6)**

97 Fecha y número de publicación de la concesión europea: **25.04.2018 EP 3123283**

54 Título: **Seguimiento de mirada ocular basado en correlación de homografía adaptativa**

30 Prioridad:

**26.03.2014 US 201414226467**

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

**11.09.2018**

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC**

**(100.0%)**

**One Microsoft Way**

**Redmond, WA 98052-6399, US**

72 Inventor/es:

**ZHANG, ZHENGYOU;**

**CAI, QIN;**

**LIU, ZICHENG y**

**HUANG, JIA-BIN**

74 Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 680 825 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**

Seguimiento de mirada ocular basado en correlación de homografía adaptativa

**Antecedentes**

- 5 A medida que han evolucionado los ordenadores incluyendo los dispositivos móviles, los dispositivos de mano y la tecnología relacionada, tales como los visualizadores, los mecanismos de entrada de datos humanos han avanzado de manera similar. Las interfaces naturales de usuario tales como basadas en reconocimiento de habla, seguimiento de la cabeza y del esqueleto y detección de gestos están llegando a estar más extendidas para complementar o, en algunos casos, sustituir al teclado, dispositivo de apuntamiento (ratón o lápiz) y/o la entrada de símbolo/escritura reconocida. La detección de la mirada ocular (seguimiento ocular) es otra tecnología natural de interfaz de usuario.
- 10 Un tipo de tecnología de seguimiento ocular se conoce como seguimiento ocular basado en relación cruzada (CR). Esta tecnología explota la invariancia de una proyectividad plana para permitir la estimación remota de la mirada de un sujeto usando una única cámara en una configuración no calibrada. En general, se proyecta luz infrarroja hacia un usuario, con reflexiones de la córnea del ojo del usuario (reflejos) detectadas por la cámara y procesadas para seguir la mirada.
- 15 El documento US 2004/174496 describe un método y un sistema informático para seguimiento de la mirada ocular. En el sistema, una cámara se enfoca sobre un ojo de un sujeto que ve un punto de mirada en una pantalla mientras que se dirige la luz hacia el ojo. Se muestrea los datos de la mirada ocular pertenecientes a un reflejo y la imagen de la pupila del ojo en un plano de imagen de la cámara. Los parámetros de la mirada ocular se determinan a partir de los datos de la mirada ocular. Los parámetros de la mirada ocular determinados incluyen: proyecciones ortogonales de un vector de desplazamiento pupila-reflejo, una relación de una dimensión del semieje mayor a una dimensión del semieje menor de una elipse que se ajusta a la imagen de la pupila en el plano de la imagen, una orientación angular de la dimensión del semieje mayor en el plano de la imagen, y coordenadas mutuamente ortogonales del centro del reflejo en el plano de imagen. El punto de mirada se estima a partir de los parámetros de la mirada ocular.
- 20

**Compendio**

- 25 Este Compendio se proporciona para introducir una selección de conceptos representativos de una forma simplificada que se describen aún más a continuación en la Descripción detallada. Este Compendio no se pretende que identifique las características clave o las características esenciales de la materia objeto reivindicada, ni se pretende que sea usado de ninguna forma que limitase el alcance de la materia objeto reivindicada.
- 30 Como resultado de las suposiciones de simplificación, el rendimiento de los seguidores de la mirada ocular basados en CR conocidos se deteriora significativamente a medida que el sujeto se aleja de una posición de calibración inicial (fija). Al mismo tiempo, no es práctico implementar un sistema en el cual un sujeto necesita calibrar el seguimiento ocular en cada una de las muchas posiciones posibles de la cabeza x, y y z que ocurren en el uso del mundo real. De esta manera, se desea una tecnología mejorada para los seguidores de la mirada ocular basados en CR y otros.
- 35 Brevemente, diversos aspectos de la materia objeto descrita en la presente memoria se dirigen hacia la correlación de homografía adaptativa para lograr la detección de la mirada. En uno o más aspectos, al menos cuatro fuentes de luz generan reflexiones de la córnea como reflejos del ojo de un sujeto, y una cámara está configurada para capturar una imagen actual que contiene los reflejos. Un modelo de correlación de homografía adaptativa aprendido a través de variables, incluyendo variables representativas de las ubicaciones de la cabeza en relación con una posición de calibración y/o direcciones de la mirada, está configurado para hacer coincidir datos de características correspondientes a los reflejos, datos relacionados con la pupila y/o datos de la mirada para emitir información de la mirada indicativa de donde está mirando actualmente el ojo del sujeto.
- 40
- 45 Uno o más aspectos se dirigen hacia el uso de un modelo de correlación de homografía adaptativa para la detección de la mirada, en el que el modelo de correlación de homografía adaptativa se entrena para compensar los errores de la mirada que varían espacialmente y los errores dependientes de la postura de la cabeza en relación con una posición de calibración. Los datos actuales de reflejo y los datos relacionados con las pupilas se capturan en una imagen y se procesan a partir de la imagen como características proporcionadas al modelo de correlación de homografía adaptativa. Los datos se reciben del modelo de correlación de homografía adaptativa en base a las características que corresponden a la información actual de la mirada.
- 50 Uno o más aspectos se dirigen hacia capturar una imagen que incluye el ojo de un sujeto a partir de la cual se extraen datos de reflejo y datos relacionados con la pupila como características y usar las características como entrada para un modelo de correlación de homografía adaptativa para determinar una dirección de la mirada. El modelo de correlación de homografía adaptativa se puede aprender usando al menos algunos datos simulados correspondientes a valores de corrección de desviación predichos en diferentes posiciones de la cabeza. El modelo de correlación de homografía adaptativa se puede aprender obteniendo una primera variable de predicción que comprende un vector de movimiento correspondiente a una posición relativa de la cabeza, y obteniendo una segunda variable de predicción correspondiente a una dirección de la mirada. El aprendizaje puede incluir minimizar
- 55

una función objetivo basada en datos correspondientes a una pluralidad de posiciones de la cabeza y direcciones de la mirada. En general, el modelo de correlación de homografía adaptativa usa escalado y translación para predicción y homografía para su corrección.

5 Otras ventajas pueden llegar a ser evidentes a partir de la siguiente descripción detallada cuando se toma en conjunto con los dibujos.

**Breve descripción de los dibujos**

La presente invención se ilustra a modo de ejemplo y no está limitada en las figuras anexas en las que números de referencia iguales indican elementos similares y en las que:

10 La FIG. 1 es un diagrama de bloques que ilustra componentes de ejemplo que incluyen un modelo de correlación de homografía adaptativa aprendido que se puede usar en la detección de la mirada ocular, según una o más implementaciones de ejemplo.

La FIG. 2 es una representación de cómo se captura un reflejo para su uso en la detección de la mirada para su uso como una característica para un modelo de correlación de homografía adaptativa aprendido para la detección de la mirada, según una o más implementaciones de ejemplo.

15 La FIG. 3 es una representación de cómo se usan los reflejos y los datos relacionados con la pupila (por ejemplo, el centro de la pupila) para obtener información de la mirada a partir de un modelo de correlación de homografía adaptativa aprendido, según una o más implementaciones de ejemplo.

La FIG. 4 es una representación de cómo las transformaciones basadas en la relación cruzada se pueden usar para entrenar un modelo de correlación de homografía adaptativa, según una o más implementaciones de ejemplo.

20 La FIG. 5 es una representación de entrenamiento de un modelo de correlación de homografía adaptativa en diversas posiciones de la cabeza, según una o más implementaciones de ejemplo.

Las FIG. 6A y 6B, y 7A y 7B son representaciones de ejemplo de cómo los datos de entrenamiento simulados se pueden basar en variaciones suaves de escala y translación debidas al movimiento de la cabeza, según una o más implementaciones de ejemplo.

25 La FIG. 8 es un diagrama de flujo que ilustra pasos de ejemplo que se pueden tomar para obtener información de la mirada a partir de un modelo de correlación de homografía adaptativa aprendido, según una o más implementaciones de ejemplo.

30 La FIG. 9 es un diagrama de bloques que representa un sistema informático o un entorno de operación ejemplar no limitante, en forma de un dispositivo informático y/o de comunicaciones móvil y/o de mano, en el cual se pueden implementar uno o más aspectos de diversas realizaciones descritas en la presente memoria.

**Descripción detallada**

Diversos aspectos de la tecnología descrita en la presente memoria se dirigen generalmente hacia correlación de homografía adaptativa para lograr la predicción de la mirada con una precisión más alta en la posición de calibración y más robustez bajo movimientos de la cabeza. Esto se logra con una tecnología basada en aprendizaje para compensar errores de la mirada que varían espacialmente y errores dependientes de la postura de la cabeza simultáneamente en un marco unificado. En uno o más aspectos, el modelo de homografía adaptativa se puede entrenar fuera de línea usando datos simulados, ahorrando tiempo y esfuerzo significativos en la recopilación de datos; en otras palabras, no se necesita requerir que el sujeto realice la calibración en las muy diversas posiciones posibles de la cabeza. Por ejemplo, los términos de escalado y los términos de translación para  $x$ ,  $y$  cambian suavemente en la práctica, y de esta manera los datos simulados pueden incluir predicciones en cuanto a cómo la homografía de corrección de desviación cambia a una nueva posición de la cabeza, para su uso como (al menos parte de) los datos sobre el terreno.

45 Como se entenderá, la correlación de homografía adaptativa se basa en variables de predicción que capturan el movimiento de la cabeza en relación con la posición de calibración y la posición de la mirada en la pantalla. Los datos sobre el terreno para el entrenamiento de la correlación de homografía adaptativa se pueden recopilar a través de una serie de calibración independiente del sujeto en diversas posiciones de la cabeza, incluyendo el uso de simulación/datos simulados en las posiciones.

50 Durante la operación en línea, el modelo entrenado se usa para corregir de manera adaptativa la desviación inducida a partir de los errores de la mirada que varían espacialmente y los errores dependientes de la postura de la cabeza. En la práctica, esta operación en tiempo real es más precisa y más robusta para el movimiento de la cabeza que otras tecnologías conocidas de la mirada ocular.

Con este fin, el planteamiento de adaptación basado en aprendizaje compensa simultáneamente los errores que varían espacialmente y los errores inducidos a partir de los movimientos de la cabeza usando la transformación de

reflejo, por ejemplo, la distancia entre reflejos y/o la variación de tamaño del patrón de reflejo considerando la transformación geométrica entre los patrones de reflejo. El modelo resultante no solamente compensa las variaciones de profundidad de un sujeto, sino los movimientos paralelos al plano de la pantalla. Obsérvese que mientras que la función de adaptación se puede obtener a través de un proceso de aprendizaje entrenado sobre datos simulados, no obstante, cualquier conocimiento previo acerca de la configuración del sistema (si está disponible) se puede incorporar fácilmente en el sistema.

Se debería entender que cualquiera de los ejemplos en la presente memoria no son limitantes. Por ejemplo, aunque se ejemplifican cuatro fuentes de luz y una cámara, cualquier número de cámaras y fuentes de luz (que proporcionen un patrón de reflejo adecuado) se puede colocar de cualquier número de formas. Además, los algoritmos y similares usados para detectar la mirada ocular son solamente ejemplos, y la tecnología descrita en la presente memoria es independiente de y no está limitada a ninguna particular, y además es capaz de ser adaptada a medida que se desarrollen nuevos algoritmos. Por tanto, la presente invención no está limitada a ninguna realización, aspecto, concepto, estructura, funcionalidad o ejemplo particular descrito en la presente memoria. Más bien, cualquiera de las realizaciones, aspectos, conceptos, estructuras, funcionalidades o ejemplos descritos en la presente memoria son no limitantes, y la presente invención se puede usar de diversas formas que proporcionan beneficios y ventajas en la detección de la mirada ocular en general.

La FIG. 1 es un diagrama de bloques general que ilustra los componentes de ejemplo que se pueden usar para realizar la detección de la mirada ocular. En la FIG. 1, se muestran un dispositivo informático 102 y un visualizador 104. El visualizador 104 puede ser un visualizador externo acoplado al dispositivo informático o un visualizador incorporado dentro del dispositivo informático, por ejemplo, su alojamiento.

Como se muestra en la FIG. 1, se muestra una pluralidad de fuentes de luz IR 106(1) -106 (m), junto con una o más cámaras sensibles a la luz IR 108(1) -108 (n). Obsérvese que para la detección de la mirada ocular basada en relación cruzada, típicamente es suficiente una única cámara; no obstante, si está presente, las imágenes de múltiples cámaras se pueden procesar y combinar de alguna forma (por ejemplo, promediada) tal como para reducir los efectos del ruido.

Las fuentes de luz pueden ser fuentes de luz individuales tales como diodos emisores de luz (LED) de láser, y/o LED o similares que proyectan a través de un elemento óptico que difracta/refleja la luz, proporcionando por ello una pluralidad de fuentes de luz. Obsérvese que cualquiera o todas las cámaras sensibles a la luz IR se pueden combinar con cámaras de luz visible. Obsérvese además que la cámara (o cámaras) se pueden unir al dispositivo, por ejemplo, insertar en un borde (por ejemplo, la cámara 208 de la FIG. 2 representada por la X dentro del círculo) o acoplar físicamente al dispositivo, o pueden ser externas al dispositivo (por ejemplo, la cámara 408 de la FIG. 4), o una combinación de ambas.

Como se entiende en seguimiento ocular basado en relación cruzada, se necesitan al menos cuatro fuentes de luz para proporcionar los reflejos para calcular la homografía, y estas fuentes de luz están dispuestas de manera que haya al menos tres direcciones diferentes entre cualquiera de ellas y las otras proporcionando un cuadrilátero, por ejemplo, un patrón rectangular de fuentes 222 - 225 como en la FIG. 2 es una disposición típica. Sin embargo, son factibles otras disposiciones que incluyan más fuentes de luz y, por ejemplo, pueden proporcionar beneficios tales como proporcionar al menos cuatro reflejos cuando no se detecta uno de los otros reflejos.

Un controlador 110 se puede usar para controlar la operación de las fuentes de luz IR 106(1) -106(m) y/o las cámaras sensibles a la luz IR 108(1) -108(n), aunque en una o más implementaciones las fuentes de luz y las cámaras pueden estar "siempre encendidas", por lo que no se necesita ningún "controlador" distinto de una fuente de alimentación presumiblemente con capacidades de encendido/apagado. Obsérvese que a luz IR se usa porque no es perceptible para los humanos, no obstante, en ciertas situaciones puede ser deseable usar luz visible, tal como con los ojos del sujeto que usa lentes de contacto que bloquean la longitud de onda de la luz visible particular que se usa. De esta manera, como se usa en la presente memoria, "fuente de luz" no está limitada a longitudes de onda de IR.

En general, la una o más cámaras 108(1) - 108(n) capturan imágenes que se alimentan a un componente de procesamiento de imagen 112, que incluye un detector de mirada ocular 114, que está acoplado a o incorpora un corrector de desviación de la mirada relacionada con la posición de la cabeza 116; como se describe en la presente memoria, el corrector de desviación incluye un componente de correlación de homografía adaptativa entrenado. El componente de procesamiento de imágenes 112 proporciona una salida de detección de mirada ocular 118, tal como coordenadas de mirada representativas de dónde está mirando actualmente el usuario en el marco dado o similar que se procesa. Tal salida 118 se puede almacenar temporalmente, tal como para su uso con otra entrada (por ejemplo, pulsaciones del ratón o gestos), se puede consumir por un sistema operativo (por ejemplo, para mover un cursor), se puede usar por una aplicación (por ejemplo, para resaltar un elemento de menú) y/o similares. En general, la información actual de la mirada se puede usar para tomar acciones con respecto al cambio de un estado de una interfaz de usuario. La detección de la mirada ocular se puede usar para otros cambios de estado, por ejemplo, para encender un visualizador a un estado activo desde un estado de espera o apagado (o viceversa), posiblemente en combinación con otra detección (por ejemplo, gesto) tal como una patrón de parpadeo ocular.

Con respecto al detector de la mirada ocular 114, cualquier técnica existente o a ser desarrollada (tal como la tecnología de relación cruzada) se puede emplear para convertir reflejos detectados, datos de la pupila y cualquier otra característica capturada en la salida de información de la mirada ocular 118. Las decisiones de una o más técnicas se pueden combinar (por ejemplo, promediar) para hacer la salida final. Como se describe en la presente memoria, el corrector de desviación de la mirada relacionada con la posición de la cabeza 116 proporciona el algoritmo de detección de la mirada ocular 114 con corrección de desviación que es más preciso y robusto que otras técnicas de corrección de desviación.

En general, los sistemas remotos de seguimiento de la mirada operan usando las fuentes de luz infrarroja para generar reflexiones de la córnea, conocidas como reflejos, que se capturan como parte de las imágenes del ojo del sujeto. Las imágenes capturadas se procesan para extraer características informativas que son invariantes a la iluminación y el punto de vista, tales como el centro de la pupila, las reflexiones de la córnea (por ejemplo, indicativas de la posición del globo ocular) y/o el contorno del limbo.

Obsérvese que en la FIG. 2, el concepto de reflejos reflejados desde la fuente de luz IR 225 se muestra como que se captura por la cámara 208 mientras el usuario está mirando a una ubicación actual de la mirada 226 en la pantalla. Como se puede apreciar fácilmente, los reflejos de las otras fuentes de luz 222-224 se capturan de manera similar al mismo tiempo (aunque solamente un conjunto tal de flechas al ojo/reflejadas a la cámara se muestra en la FIG. 2).

Como se entiende, la posición de la cabeza del sujeto 228 y la ubicación de la mirada 226 influyen en las posiciones y los tamaños de los reflejos  $g_1 - g_4$  que se capturan. Esta información, junto con otra información tal como el centro de la pupila  $u_p$ , corresponde a datos de rasgos 304 extraídos de la imagen 302, que se alimenta al corrector de desviación de la mirada relacionada con la posición de la cabeza 116 aprendida. A partir de allí, se obtiene información de la mirada 310 tal como coordenadas de pantalla y se proporciona a un programa 312.

Como se describe en la presente memoria, los métodos basados en homografía para la corrección de la desviación por estimación de la mirada pueden, en algunas circunstancias, aumentar la precisión y/o la robustez de la estimación de la mirada. El método basado en homografía para la corrección de la desviación puede implementar una transformación de homografía de corrección de desviación. La transformación de homografía de corrección de desviación se puede calcular resolviendo el problema de registro del conjunto de puntos a partir de los puntos de la mirada predichos mediante el método básico de relación cruzada para objetivos sobre el terreno en la pantalla durante una fase de entrenamiento de calibración.

En general, los métodos basados en homógrafos generalmente funcionan bien en la posición de calibración porque modelan de manera eficaz los desplazamientos de los ejes óptico y visual, como se representa de manera general en la FIG. 4. No obstante, debido al error del modelo de la suposición de planitud en el centro de la pupila y el plano formado por reflejos, surgen errores que varían espacialmente. Para una predicción precisa, la correlación de homografía de corrección de desviación necesita depender de la dirección de la mirada del sujeto.

El rendimiento de los métodos basados en homografía se degrada significativamente cuando el sujeto se aleja de la posición de calibración porque la homografía de corrección de desviación óptima es una función de las posiciones de la cabeza. Uno o más aspectos descritos en la presente memoria están dirigidos a correlación de homografía adaptativa, que se dirige hacia "predecir" cómo cambia la homografía de corrección de desviación en una nueva posición de la cabeza, de manera que el rendimiento del seguidor de la mirada será como si estuviera calibrado en esa nueva posición de la cabeza. Se describe un esquema para predecir la variación de la homografía de corrección de desviación calculada en la posición de calibración basada en los cambios relativos entre la posición actual de la cabeza y la posición de calibración y la dirección actual de la mirada.

Con respecto a la relación cruzada de correlación de homografía con la corrección de desviación basada en homografía, como se muestra de manera general en la FIG. 4, denotamos  $L_i$  como las fuentes de luz puntuales situadas en las cuatro esquinas de la pantalla ( $1 \leq i \leq 4$ ),  $G_i$  como las reflexiones de la córnea correspondientes y  $g_i$  como las imágenes de  $G_i$ .  $P$  es el centro de la pupila en 3D y  $p$  como su proyección en la imagen. Aunque se muestran cuatro fuentes de luz puntuales en las esquinas de la pantalla, se pueden usar múltiples fuentes de luz de muy diferentes estructuras, números, colocaciones según sea adecuado y se muestran aquí cuatro luces de esquina con propósitos de ejemplo. El método de relación cruzada supone que cada uno del grupo ( $L_i, G_i, g_i$ ) es coplanar, denotado como plano  $\Pi_L, \Pi_G, \Pi_g$  respectivamente. La transformación entre los planos  $\Pi_L, \Pi_G, \Pi_g$  se puede describir a través de homografías. Bajo el supuesto de que el centro de la pupila  $P$  se encuentra en  $\Pi_G$ , la predicción del punto de vista se da por:

$$P \circ R_{CR} = H_{GL} \left( H_{gG}(p) \right) = H_{CR}(p) \quad (1)$$

donde  $H_{gG}$  correlaciona el plano  $\Pi_g$  con el plano  $\Pi_G$ ,  $H_{GL}$  correlaciona el plano  $\Pi_G$  con el plano  $\Pi_L$ , y  $H_{CR}$  es la transformada combinada de  $H_{GL}$  y  $H_{gG}$ . No obstante, debido a que estos supuestos de simplificación no son válidos en la práctica, se observa una desviación grande de la estimación de la mirada.

Las técnicas basadas en homografía aplican otra transformación de homografía para corregir esta desviación de estimación de la mirada. En una técnica, los reflejos en las imágenes se correlacionan primero en un espacio normalizado (por ejemplo, un cuadrado unitario  $\Gamma_N$ ) con la homografía de corrección de desviación usada para correlacionar los puntos de la mirada estimados en el espacio normalizado con los puntos de la mirada esperados en el espacio de pantalla  $\Gamma_L$ . La predicción del punto de mirada mediante predicción basada en homografía se da por:

$$PoR_{HOM} = \mathbf{H}_{NL}(\mathbf{H}_{CR}^N(p)) \quad (2)$$

donde  $\mathbf{H}_{CR}^N$  correlaciona el espacio de imagen con el espacio normalizado y  $\mathbf{H}_{NL}$  correlaciona el espacio normalizado con el espacio de pantalla. Denotando  $v$  como el índice de la posición objetivo en la pantalla,  $V$  como el conjunto del índice objetivo y  $t_v$  como la posición objetivo en el espacio de pantalla, una meta de la calibración dependiente del sujeto es encontrar la homografía de corrección de desviación óptima  $\mathbf{H}_{NL}^*$  que minimice los errores de reproyección:

$$\mathbf{H}_{NL}^* = \underset{\mathbf{H}_{NL}}{\operatorname{argmin}} \sum_{v \in V} ||t_v - \mathbf{H}_{NL}(\mathbf{H}_{CR}^N(p_v))||_2^2, \quad (3)$$

donde  $p_v$  es la posición del centro de la pupila 2D en la imagen cuando se mira al objetivo  $v$ .

Se describe en la presente memoria la correlación de homografía adaptativa, que modela la variación de la homografía de corrección de desviación  $\mathbf{H}_{NL}$  usando otra correlación de homografía  $\mathbf{H}_A$ . El punto de vista por la homografía adaptativa se da por:

$$PoR_{AH} = \mathbf{H}_{NL}(\mathbf{H}_A(\mathbf{H}_{CR}(p))) \quad (4)$$

Obsérvese que en la Ecuación (4), la homografía de corrección de desviación  $\mathbf{H}_{NL}$  se calcula mediante el mismo proceso de minimización en la Ecuación (3) en la calibración y permanece sin cambios para el mismo sujeto. Por el contrario, la correlación de homografía adaptativa  $\mathbf{H}_A$  necesita variar de manera adaptativa para la posición actual de la cabeza en relación con la posición de calibración, así como la dirección de la mirada. En uno o más aspectos, la homografía adaptativa se describe en la presente memoria como un problema de regresión. Es decir, dadas las variables de predicción que describen la posición relativa de la cabeza y la dirección de la mirada, el sistema quiere predecir los valores en  $\mathbf{H}_A$ .

Se pueden usar diferentes tipos de variables de predicción, incluyendo, sin limitación, movimiento (correspondiente a la posición de la cabeza) y dirección de la mirada,  $x = [x_m, x_g]^T$ . En primer lugar, se capturan los movimientos de la cabeza en relación con la posición de calibración usando la transformación geométrica entre el cuadrilátero de los reflejos almacenado en la posición de calibración y el cuadrilátero de los reflejos actual. En la práctica, la transformación afin o de similitud se puede usar para codificar el movimiento relativo. Por ejemplo, cuando el sujeto se mueve hacia la pantalla después de la calibración, el término de escala de la transformación será mayor que uno. La transformación de homografía es otra técnica adecuada descrita de manera general anteriormente con respecto a la corrección de desviación en general.

El primer tipo de variable de predicción  $x_m$  se obtiene vectorizando los parámetros de movimiento. Hay un vector de seis dimensiones para  $x_m$  cuando se usa transformación afin o un vector de cuatro dimensiones para  $x_m$  cuando se usa transformación de similitud. Además, para codificar la dirección de la mirada para correlación que varía espacialmente, los datos relacionados con la pupila se usan como una de las características, por ejemplo, la posición del centro de la pupila en el espacio normalizado  $x_g = \mathbf{H}_{CR}(p - p_0)$ , donde  $p_0$  es la posición central de la pupila cuando se mira al centro de la pantalla.

Con estas variables de predicción, la homografía adaptativa se puede modelar como una regresión polinómica de grado dos (es decir, regresión cuadrática):

$$\mathbf{H}_{A,x} = f(x, \beta) \quad (5)$$

En la regresión cuadrática, los valores de la homografía adaptativa son lineales con las variables de predicción, que contienen un término constante, términos lineales, términos cuadráticos, así como los términos de interacción.

Se puede lograr compensación de error para variación de la profundidad escalando adaptativamente los vectores de corrección de translación usando el tamaño relativo de los cuadriláteros de reflejo en la posición de calibración y la posición actual. En una o más implementaciones, la tecnología descrita en la presente memoria considera un conjunto más rico de transformaciones que la escala para la predicción y usa homografía (en lugar de solamente la translación) para la corrección. Los valores de las homografías de corrección de desviación óptimas son dependientes de los movimientos de la cabeza.

Obsérvese que en lugar de tener una mirada de calibración del sujeto en cada posición posible de la cabeza, en una o más implementaciones, se puede lograr la compensación de error aprendiendo primero la adaptación a través de datos de simulación, y luego prediciendo el vector actual de translación usando esos datos de entrenamiento simulados para aprender la adaptación. Usar esta metodología puede ahorrar tiempo y esfuerzo significativos de calibración del sujeto, así como mejorar la velocidad de cálculo. Además, el uso de simulación permite usar un modelo más complejo que simplemente la translación para la predicción.

La FIG. 5 resume el proceso general de una implementación. Durante el entrenamiento, cada una de las diversas transformaciones (por ejemplo, transformaciones afines A) para las posiciones de la cabeza  $H_0$  a  $H_n$  y posiciones de la mirada son conocidas, y se pueden representar como  $X = \{(A_1, g_1), (A_2, g_2), \dots, (A_n, g_n)\}$ . Las posiciones correspondientes de la cabeza mostradas en la FIG. 5 como  $H_0$  a  $H_n$  se conocen durante el entrenamiento, de manera que Y se puede aprender,  $Y = \{H_0^{-1} H_1, H_0^{-1} H_2, \dots, H_0^{-1} H_n\}$ . Obsérvese que los datos de entrenamiento sobre el terreno pueden ser datos simulados, al menos en parte.

De esta manera, como se ha descrito anteriormente, la operación de entrenamiento obtiene los datos que se usan para aprender una función de regresión  $f: X \rightarrow Y$  (polinomio de grado 2). Estas posiciones aprendidas de la cabeza y/o función de regresión basadas en los datos entrenados se pueden almacenar de cualquier manera apropiada y pueden ser accesibles para el sistema de seguimiento de la mirada de la FIG. 1 tal como en el almacenamiento local y/o almacenamiento en la nube.

Durante el uso por el usuario actual que puede estar dentro de una aplicación, un entorno en línea, dentro de un sistema operativo base de un dispositivo informático, etc., el registro afín A de un conjunto actual de reflejos correspondiente a las posiciones de reflejo capturadas se obtiene a través de la captura de imagen. Una vez que éste se obtiene junto con la dirección de la mirada  $g$ , usar la regresión:  $y = f(x)$  con  $x = (A, g)$  proporciona la homografía de corrección de desviación:

$$H = H_0 y.$$

Con respecto a la adaptación de homografía de aprendizaje, denotamos  $u$  como la posición de la cabeza en tres dimensiones y  $U$  como el conjunto de posiciones muestreadas de la cabeza. Una función objetivo adecuada se define como:

$$\mathcal{L}(\beta) = \sum_{u \in U} \sum_{v \in V} \|t_{u,v} - \mathbf{H}_{NL}^* (\mathbf{H}_{A,x} (\mathbf{H}_{CR}^N (p_{u,v})))\|_2^2, \quad (6)$$

donde  $H_{A,x} = f(x, \beta)$  (ecuación (5)) es el modelo de regresión cuadrática para homografía adaptativa. Una meta del aprendizaje de homografía adaptativa es encontrar una matriz de coeficientes que minimice los errores de reproyección sumando los errores cuadráticos entre las posiciones de la mirada predicha y los sobre el terreno en la pantalla cuando los sujetos simulados están situados en las posiciones muestreadas de la cabeza.

Para minimizar la función objetivo definida en la Ecuación (6), se puede usar un planteamiento de dos pasos. Primero, la función de predicción se puede estimar minimizando un error algebraico. En cada posición de la cabeza  $u$ , calculamos la homografía de corrección de desviación óptima  $H_{NL}^u$  realizando una calibración dependiente del sujeto en la posición  $u$ . Idealmente hasta un factor de escala,  $H_{NL}^u = H_{NL}^* H_{A,x}$ . El proceso puede minimizar, de esta manera, los errores algebraicos entre la predicción  $H_{A,x} = f(x_{u,v}, \beta)$  y la diferencia de la homografía de corrección de desviación  $(H_{NL}^*)^{-1} (H_{NL}^u)$  (con el último elemento normalizado a 1), donde la  $H_{NL}^*$  es la homografía de corrección de desviación calculada en la posición de calibración por defecto. La minimización de error algebraico, se puede formular de esta manera como:

$$\beta^a = \underset{\beta}{\operatorname{argmin}} \sum_{u \in U} \sum_{v \in V} \|(\mathbf{H}_{NL}^*)^{-1} (\mathbf{H}_{NL}^u) - f(x_{u,v}, \beta)\|_2^2, \quad (7)$$

donde  $\beta^a$  es la matriz estimada de coeficientes después de minimizar los errores algebraicos.

En segundo lugar, para minimizar los errores de reproyección en la Ecuación (6), el proceso puede comenzar con la solución inicial usando  $\beta^a$ , y realizar optimización no lineal de mínimos cuadrados usando el algoritmo de Levenberg-Marquardt.

Volviendo a una implementación de ejemplo del proceso de calibración, se usa un conjunto de sujetos de entrenamiento, cada uno invitado a mirar a un patrón de cuadrícula regular  $n \times n$ ,  $n \in \{2, 3, 4, 5\}$  que se distribuye uniformemente sobre la pantalla. En un escenario de entrenamiento, se usó una cuadrícula  $5 \times 5$  uniformemente distribuida en la pantalla. Para la recopilación de los datos de entrenamiento, definamos el plano de la pantalla como el plano  $x - y$  y la profundidad de la pantalla como el eje  $z$  en el sistema de coordenadas del mundo. Se puede muestrear un espacio de trabajo típico frente a la pantalla usando una cuadrícula de  $5 \times 5 \times 5$  con rangos de -200 mm a 200 mm, centrada en la posición  $[0, 0, 600]$  mm. En cada posición de la cabeza  $u$ , realicemos la calibración

dependiente del sujeto en la Ecuación (3), por ejemplo, usando un patrón de calibración  $n \times n$  (por ejemplo,  $5 \times 5$ ) en la pantalla. Para dar cuenta de los sujetos con diferentes parámetros oculares, muestreamos aleatoriamente algún número de (por ejemplo, cincuenta) sujetos virtuales usando distribuciones Gaussianas con medias de parámetros oculares típicos y desviaciones estándar del diez por ciento de los valores del parámetro. Por ejemplo, el tamaño típico del radio de la córnea es 7,8 mm. Luego, el proceso extrae muestras aleatorias usando una distribución Gaussiana con una media de 7,8 y una desviación estándar de 0,78.

Por ejemplo, comenzando con los parámetros oculares típicos (radio de la córnea  $R_c = 7,8$  mm, distancia desde el centro de la córnea al centro de la pupila  $K = 4,2$  mm, la desviación angular horizontal y vertical es de 5,0 grados y 1,5 grados, el proceso varía el valor de cada parámetro ocular con  $[-30, 30]$  % de los valores originales.

En lugar de o además de los datos reales de los sujetos, se pueden usar datos simulados. Por ejemplo, las FIG. 6A y 6B muestran gráficos de valores de la homografía de corrección de desviación óptima calculada en diferentes posiciones de la cabeza a lo largo del eje de profundidad para la escala en  $x$ ,  $y$ . Las FIG. 7A y 7B muestran gráficos similares para translación en  $x$ ,  $y$ . Obsérvese que el último elemento de cada homografía está normalizado a uno. Como se puede ver, los gráficos son suaves. De esta manera, se pueden predecir valores óptimos como datos simulados sobre el terreno.

Obsérvese que el uso de datos simulados en lugar de datos de calibración reales para un sujeto no está limitado a la tecnología de relación cruzada. Otras soluciones de detección de la mirada ocular, tales como los métodos basados en modelos (que estiman un vector de mirada 3D y calculan puntos de interés 2D cruzando rayos 3D con el plano de la pantalla 2D), también pueden usar simulación para la calibración.

La FIG. 8 es un diagrama de flujo generalizado que muestra pasos de ejemplo en el uso real del modelo de homografía adaptativa aprendido. El paso 802 captura la imagen, que se procesa (paso 804) en los datos de reflejo y los datos relacionados con la pupila para su uso como características (paso 808). El modelo entrenado usa los datos de características para determinar los datos de corrección de la posición de la cabeza usados para calcular la información de la mirada corregida, por ejemplo, las coordenadas (o el identificador general de cuadrícula) donde el ojo del sujeto está mirando en la pantalla que se puede emitir a un almacenador temporal o similar para su consumo por el entorno de operación del sistema de seguimiento de la mirada, tal como una aplicación, entorno en línea, sistema operativo, etc. Los resultados de la información de la mirada se pueden usar en muchos escenarios diferentes incluyendo para interacciones naturales de interfaz de usuario, determinación de la atención para interpretación de interés del usuario, etc. El paso 810 repite el proceso para otro cuadro; la tasa de cuadros o alguna operación de suavizado se pueden usar para evitar demasiados saltos alrededor. En respuesta a los cambios de las coordenadas de la mirada, el sistema de seguimiento de la mirada puede desencadenar otra acción o respuesta del sistema de seguimiento de la mirada dependiendo del cambio en la mirada del usuario, por ejemplo, desencadenar o detener o iniciar una interacción natural de interfaz de usuario diferente, indicar una determinación de atención diferente para interpretación de interés del usuario, etc.

Una homografía adaptativa, tal como la descrita en la metodología descrita anteriormente, proporciona precisión más allá de los métodos conocidos basados en homografía porque además de corregir las desviaciones del movimiento de la cabeza, la homografía adaptativa también da cuenta de los errores de la mirada que varían espacialmente predichos por la posición de la pupila en el espacio normalizado  $x_g$ .

La tecnología anterior se puede combinar con otras tecnologías de seguimiento de la mirada ocular. Por ejemplo, la tecnología descrita en la presente memoria se puede combinar en un sistema con otra tecnología basada en dos ojos, tal como se describe en la Solicitud de patente de EE.UU. en tramitación titulada "EYE GAZE TRACKING USING BINOCULAR FIXATION CONSTRAINTS" presentada conjuntamente con la presente memoria, expediente del agente N° 340834.01.

Como se puede ver, se proporciona un sistema que comprende, al menos cuatro fuentes de luz y una cámara, en el que las fuentes de luz configuradas para generar reflexiones de la córnea como reflejos del ojo de un sujeto, y la cámara está configurada para capturar una imagen actual que contiene los reflejos. Un modelo de correlación de homografía adaptativa aprendido a través de variables, incluyendo variables representativas de las ubicaciones de la cabeza en relación con una posición de calibración y/o direcciones de la mirada, está configurado para hacer coincidir datos de características correspondientes a los reflejos, datos relacionados con la pupila y/o datos de la mirada para emitir información de la mirada indicativa de donde está mirando actualmente el ojo del sujeto.

En uno o más aspectos, las variables representativas de las ubicaciones de la cabeza en relación con una posición de calibración y las posiciones de la mirada se pueden basar, al menos en parte, en datos simulados. Los datos simulados se pueden usar para representar datos sobre el terreno para entrenar la correlación de homografía adaptativa a través de calibración para obtener las variables de predicción en diversas posiciones de la cabeza. Los datos sobre el terreno modelan la homografía adaptativa como una regresión polinómica.

En uno o más aspectos, las variables representativas de las ubicaciones de la cabeza en relación con una posición de calibración corresponden a movimientos de la cabeza relativos entre las diversas posiciones de la cabeza



codificadas por transformaciones afines, transformaciones de similitud o transformaciones de homografía. Las variables representativas de las direcciones de la mirada se codifican por datos relacionados con la pupila.

5 Uno o más aspectos se dirigen hacia el uso de un modelo de correlación de homografía adaptativa para la detección de la mirada, en el que el modelo de correlación de homografía adaptativa se entrena para compensar errores de la mirada que varían espacialmente y errores dependientes de la postura de la cabeza con respecto a una posición de calibración. Los datos actuales de reflejo y los datos relacionados con la pupila se capturan en una imagen, y procesan a partir de la imagen como características proporcionadas al modelo de correlación de homografía adaptativa. Los datos se reciben del modelo de correlación de homografía adaptativa basado en las características que corresponden a información actual de la mirada.

10 Uno o más aspectos se dirigen hacia el aprendizaje del modelo de correlación de homografía adaptativa, incluyendo el uso de la pluralidad de conjuntos de datos de posición y datos relacionados con la pupila como variables de predicción para modelar la homografía adaptativa como una regresión cuadrática. El uso de la pluralidad de conjuntos de datos de posición y datos de posición de la pupila puede comprender usar al menos algunos datos simulados, por ejemplo, prediciendo valores de corrección de desviación en diferentes escalas y/o translaciones de posición de la cabeza.

15 Uno o más aspectos se dirigen hacia capturar una imagen que incluye el ojo de un sujeto a partir del cual se extraen datos de reflejo y datos relacionados con la pupila como características, y usando las características como entrada a un modelo de correlación de homografía adaptativa para determinar una dirección de la mirada. El modelo de correlación de homografía adaptativa se puede aprender usando al menos algunos datos simulados correspondientes a valores de corrección de desviación predichos en diferentes posiciones de la cabeza. El modelo de correlación de homografía adaptativa se puede aprender obteniendo una primera variable de predicción que comprende un vector de movimiento correspondiente a una posición relativa de la cabeza, y obteniendo una segunda variable de predicción correspondiente a una dirección de la mirada. El aprendizaje puede incluir la minimización de una función objetivo basada en datos correspondientes a una pluralidad de posiciones de la cabeza y direcciones de la mirada. En general, el modelo de correlación de homografía adaptativa usa escalado y translación para predicción y homografía para corrección.

### Ejemplo de entorno de operación

La FIG. 9 ilustra un ejemplo de un dispositivo móvil 900 adecuado sobre el que se pueden implementar aspectos de la materia objeto descrita en la presente memoria. El dispositivo móvil 900 es solamente un ejemplo de un dispositivo y no se pretende que sugiera ninguna limitación en cuanto al alcance del uso o la funcionalidad de los aspectos de la materia objeto descrita en la presente memoria. Tampoco se debería interpretar el dispositivo móvil 900 como que tiene ninguna dependencia o requisito en relación con una cualquiera o combinación de componentes ilustrados en el dispositivo móvil 900 de ejemplo. El dispositivo móvil puede comprender un dispositivo de mano tal como un teléfono inteligente, tableta, ordenador portátil, etc. Alternativamente, se puede usar un ordenador personal, por ejemplo, con una cámara o cámaras y fuentes de luz montadas en el visualizador.

El dispositivo móvil 900 de ejemplo se puede usar en unas gafas, gafas de protección o sombreros, u otros dispositivos posibles, tales como dispositivos de tipo reloj de pulsera, incluyendo ordenadores externos, son todos entornos adecuados. Obsérvese que aunque las gafas y sombreros se usan en la cabeza, se pueden usar en diferentes posiciones en relación con la cabeza y, de esta manera, puede ser adecuada una corrección de desviación de la posición de la cabeza.

Con referencia a la FIG. 9, un dispositivo de ejemplo para implementar aspectos de la materia objeto descrita en la presente memoria incluye un dispositivo móvil 900. En algunas realizaciones, el dispositivo móvil 900 comprende un teléfono celular, un dispositivo de mano que permite comunicaciones de voz con otros, algún otro dispositivo de comunicaciones de voz o similar. En estas realizaciones, el dispositivo móvil 900 puede estar equipado con una cámara para tomar imágenes, aunque esto puede no ser requerido en otras realizaciones. En otras realizaciones, el dispositivo móvil 900 puede comprender un asistente digital personal (PDA), un dispositivo de mano de juegos, un ordenador de agenda, una impresora, un aparato que incluye un decodificador, un centro multimedia u otro aparato, otros dispositivos móviles, o similares. En otras realizaciones más, el dispositivo móvil 900 puede comprender dispositivos que se consideran generalmente no móviles, tales como ordenadores personales, servidores o similares.

Los componentes del dispositivo móvil 900 pueden incluir, pero no están limitados a, una unidad de procesamiento 905, memoria del sistema 910 y un bus 915 que acopla diversos componentes del sistema incluyendo la memoria del sistema 910 a la unidad de procesamiento 905. El bus 915 puede incluir cualquiera de varios tipos de estructuras de bus incluyendo un bus de memoria, controlador de memoria, un bus periférico, y un bus local usando cualquiera de una variedad de arquitecturas de bus, y similares. El bus 915 permite que los datos sean transmitidos entre diversos componentes del dispositivo móvil 900.

El dispositivo móvil 900 puede incluir una variedad de medios legibles por ordenador/legibles por máquina. Tales medios pueden ser cualquier medio disponible al que se pueda acceder mediante el dispositivo móvil 900 e incluyen

tanto medios volátiles como no volátiles, y medios extraíbles y no extraíbles. A modo de ejemplo, y no de limitación, los medios legibles por ordenador pueden comprender medios de almacenamiento informático y medios de comunicación. Los medios de almacenamiento informático incluyen medios volátiles y no volátiles, extraíbles y no extraíbles implementados en cualquier método o tecnología para el almacenamiento de información tal como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento informático incluyen, pero no se limitan a, RAM, ROM, EEPROM, memoria rápida u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento en disco óptico, casetes magnéticos, cinta magnética, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que se pueda usar para almacenar la información deseada y al que se pueda acceder por el dispositivo móvil 900.

Los medios de comunicación típicamente incorporan instrucciones legibles por ordenador, estructuras de datos, módulos de programa, u otros datos en una señal de datos modulada tal como una onda portadora u otro mecanismo de transporte e incluyen cualquier medio de suministro de información. El término “señal de datos modulada” significa una señal que tiene una o más de sus características establecidas o cambiadas de tal manera que codifican información en la señal. A modo de ejemplo, y no de limitación, los medios de comunicación incluyen medios cableados tales como una red cableada o conexión cableada directa, y medios inalámbricos tales como acústicos, RF, Bluetooth®, USB inalámbrico, infrarrojos, Wi-Fi, WiMAX y otros medios inalámbricos. Combinaciones de cualquiera de los anteriores también se deberían incluir dentro del alcance de los medios legibles por ordenador.

La memoria del sistema 910 incluye medios de almacenamiento informático en forma de memoria volátil y/o no volátil y puede incluir memoria de sólo lectura (ROM) y memoria de acceso aleatorio (RAM). En un dispositivo móvil tal como un teléfono celular, el código del sistema operativo 920 se incluye algunas veces en la ROM aunque, en otras realizaciones, esto no se requiere. De manera similar, los programas de aplicaciones 925 a menudo se colocan en la RAM aunque de nuevo, en otras realizaciones, los programas de aplicaciones se pueden colocar en ROM o en otra memoria legible por ordenador. La pila 930 proporciona memoria para el estado asociado con el sistema operativo 920 y los programas de aplicaciones 925. Por ejemplo, el sistema operativo 920 y los programas de aplicaciones 925 pueden almacenar variables y estructuras de datos en la pila 930 durante sus operaciones.

El dispositivo móvil 900 también puede incluir otra memoria extraíble/no extraíble, volátil/no volátil. A modo de ejemplo, la FIG. 9 ilustra una tarjeta de memoria rápida 935, una unidad de disco duro 936 y una tarjeta de memoria 937. La unidad de disco duro 936 se puede miniaturizar para caber en una ranura de memoria, por ejemplo. El dispositivo móvil 900 puede interactuar con estos tipos de memoria extraíble no volátil a través de una interfaz de memoria extraíble 931, o se puede conectar a través de un bus serie universal (USB), IEEE 9394, uno o más del puerto o los puertos cableados 940, o la antena o las antenas 965. En estas realizaciones, los dispositivos de memoria extraíbles 935 - 937 pueden interactuar con el dispositivo móvil a través del módulo o de los módulos de comunicaciones 932. En algunas realizaciones, no todos estos tipos de memoria se pueden incluir en un único dispositivo móvil. En otras realizaciones, uno o más de estos y otros tipos de memoria extraíble se pueden incluir en un único dispositivo móvil.

En algunas realizaciones, la unidad de disco duro 936 puede estar conectada de tal forma que se una más permanentemente al dispositivo móvil 900. Por ejemplo, la unidad de disco duro 936 se puede conectar a una interfaz tal como una tecnología avanzada paralela de conexión (PATA), tecnología avanzada en serie de conexión (SATA) o de otra manera, que puede estar conectada al bus 915. En tales realizaciones, la extracción del disco duro puede implicar eliminar una cubierta del dispositivo móvil 900 y la extracción de tornillos u otros sujetadores que conectan el disco duro 936 a estructuras de soporte dentro del dispositivo móvil 900.

Los dispositivos de memoria extraíble 935 - 937 y sus medios de almacenamiento informático asociados, discutidos anteriormente e ilustrados en la FIG. 9, proporcionan almacenamiento de instrucciones legibles por ordenador, módulos de programa, estructuras de datos y otros datos para el dispositivo móvil 900. Por ejemplo, el dispositivo o dispositivos de memoria extraíble 935 - 937 pueden almacenar imágenes tomadas por el dispositivo móvil 900, grabaciones de voz, información de contacto, programas, datos para los programas, etc.

Un usuario puede introducir comandos e información en el dispositivo móvil 900 a través de dispositivos de entrada tales como un teclado 941 y el micrófono 942. En algunas realizaciones, el visualizador 943 puede ser una pantalla sensible al tacto y puede permitir a un usuario introducir comandos e información en el mismo. El teclado 941 y el visualizador 943 se pueden conectar a la unidad de procesamiento 905 a través de una interfaz de entrada de usuario 950 que está acoplada al bus 915, pero también se pueden conectar por otras estructuras de interfaz y de bus, tales como el módulo o los módulos de comunicaciones 932 y el puerto o puertos cableados 940. La detección de movimiento 952 se puede usar para determinar gestos hechos con el dispositivo 900.

Como se describe en la presente memoria, los reflejos de los ojos y otros datos relacionados con los ojos se pueden capturar y procesar para su introducción. El procesamiento se puede realizar en software, en lógica de hardware o en una combinación de software y lógica de hardware.

Un usuario puede comunicarse con otros usuarios a través de hablar al micrófono 942 y a través de mensajes de texto que se introducen en el teclado 941 o un visualizador sensible al tacto 943, por ejemplo. La unidad de audio

955 puede proporcionar señales eléctricas para accionar el altavoz 944 así como recibir y digitalizar señales de audio recibidas desde el micrófono 942.

5 El dispositivo móvil 900 puede incluir una unidad de video 960 que proporciona señales para accionar una cámara 961. La unidad de video 960 también puede recibir imágenes obtenidas por la cámara 961 y proporcionar estas imágenes a la unidad de procesamiento 905 y/o la memoria incluida en el dispositivo móvil 900. Las imágenes obtenidas por la cámara 961 pueden comprender video, una o más imágenes que no forman un video, o alguna combinación de las mismas.

10 El módulo o los módulos de comunicación 932 pueden proporcionar señales a y recibir señales desde una o más antenas 965. Una de las antenas 965 puede transmitir y recibir mensajes para una red de telefonía celular. Otra antena puede transmitir y recibir mensajes Bluetooth®. Otra antena más (o una antena compartida) puede transmitir y recibir mensajes de red a través de un estándar de red Ethernet inalámbrica.

Aún más, una antena proporciona información basada en localización, por ejemplo, señales de GPS a una interfaz y mecanismo de GPS 972. A su vez, el mecanismo de GPS 972 pone a disposición los datos de GPS correspondientes (por ejemplo, tiempo y coordenadas) para su procesamiento.

15 En algunas realizaciones, se puede usar una única antena para transmitir y/o recibir mensajes para más de un tipo de red. Por ejemplo, una única antena puede transmitir y recibir mensajes de voz y de paquetes.

20 Cuando se opera en un entorno de red, el dispositivo móvil 900 puede conectarse a uno o más dispositivos remotos. Los dispositivos remotos pueden incluir un ordenador personal, un servidor, un encaminador, un PC de red, un teléfono celular, un dispositivo de reproducción multimedia, un dispositivo igual u otro nodo de red común, y típicamente incluyen muchos o todos los elementos descritos anteriormente en relación con el dispositivo móvil 900.

25 Los aspectos de la materia objeto descrita en la presente memoria son operacionales con otros numerosos entornos o configuraciones de sistemas informáticos de propósito general o propósito especial. Ejemplos de sistemas informáticos, entornos y/o configuraciones bien conocidos que pueden ser adecuados para su uso con aspectos de la materia objeto descrita en la presente memoria incluyen, pero no se limitan a, ordenadores personales, ordenadores servidores, dispositivos de mano o portátiles, sistemas multiprocesador, sistemas basados en microcontrolador, decodificadores, electrónica de consumo programable, PC en red, miniordenadores, ordenadores centrales, entornos informáticos distribuidos que incluyen cualquiera de los sistemas o dispositivos anteriores, y similares.

30 Los aspectos de la materia objeto descrita en la presente memoria se pueden describir en el contexto general de las instrucciones ejecutables por ordenador, tales como módulos de programa, que se ejecutan por un dispositivo móvil. En general, los módulos de programa incluyen rutinas, programas, objetos, componentes, estructuras de datos, etc., que realizan tareas particulares o implementan tipos de datos abstractos particulares. Los aspectos de la materia objeto descrita en la presente memoria también se pueden poner en práctica en entornos informáticos distribuidos donde las tareas se realizan mediante dispositivos remotos de procesamiento que están enlazados a través de una red de comunicaciones. En un entorno informático distribuido, los módulos de programa se pueden situar tanto en medios de almacenamiento informático local como remoto incluyendo dispositivos de almacenamiento de memoria.

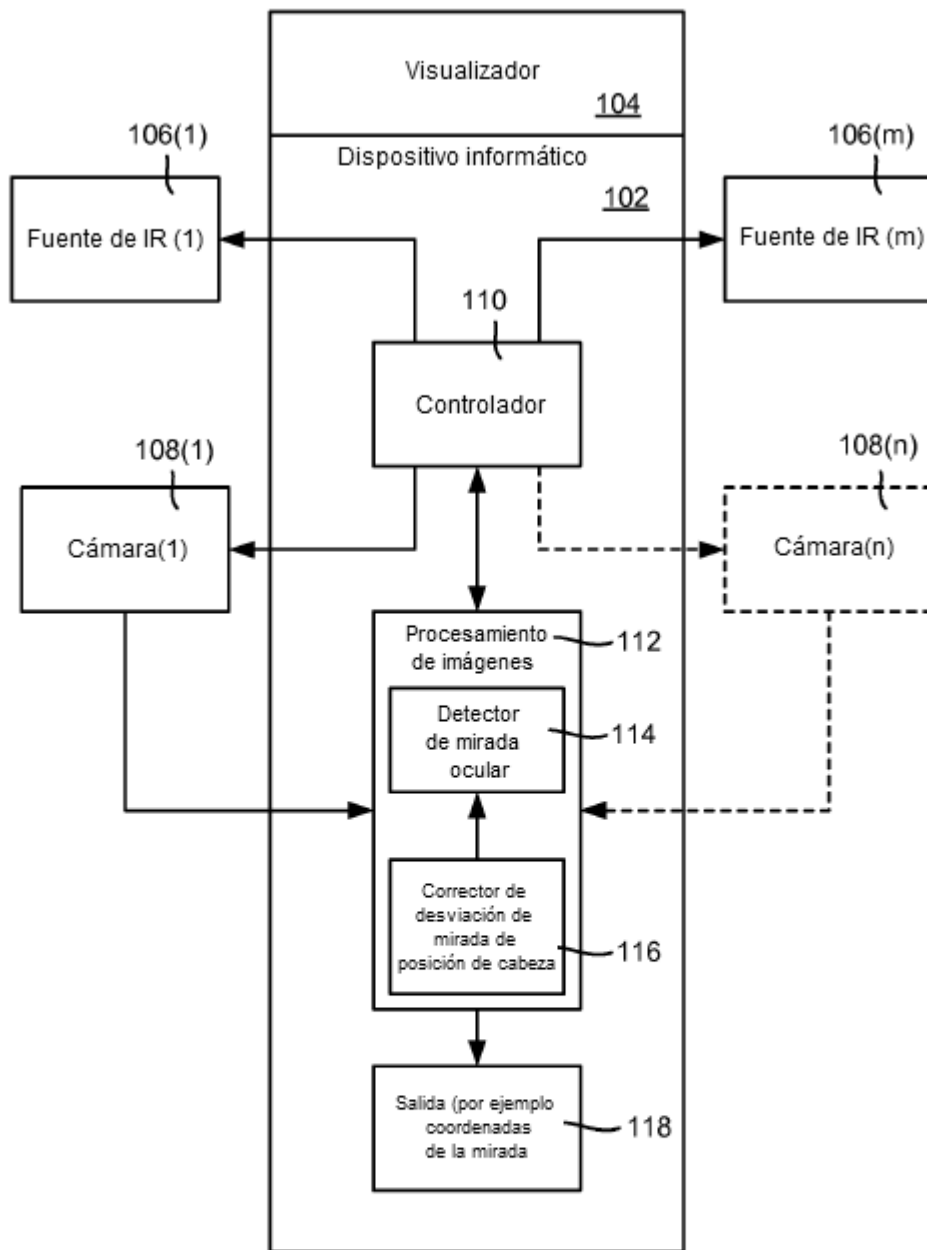
35 Además, aunque el término servidor se puede usar en la presente memoria, se reconocerá que este término también puede abarcar un cliente, un conjunto de uno o más procesos distribuidos en uno o más ordenadores, uno o más dispositivos de almacenamiento autónomos, un conjunto de uno o más de otros dispositivos, una combinación de uno o más de los anteriores y similares.

40

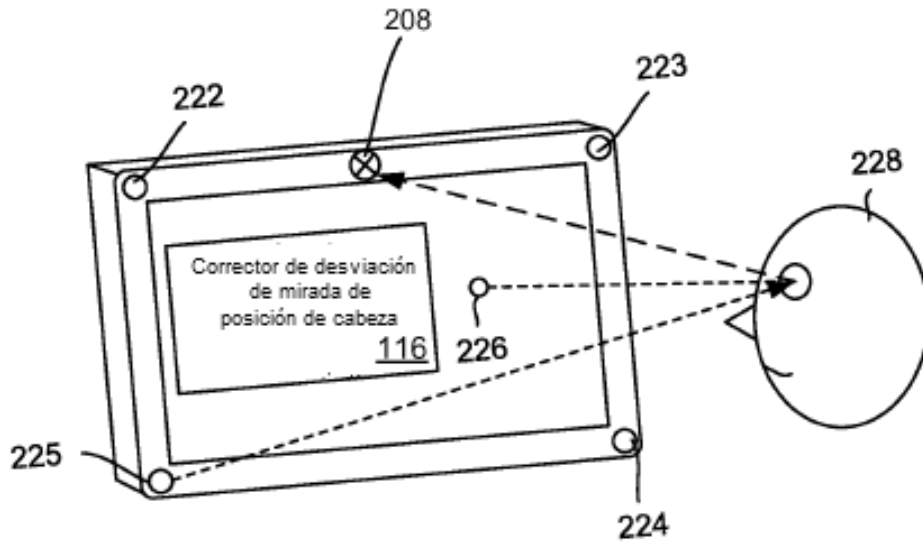
**REIVINDICACIONES**

1. Un sistema que comprende:
  - al menos cuatro fuentes de luz (106(1) - 106(m), 222-225) configuradas para generar reflexiones de la córnea como reflejos ( $g_1 - g_4$ ) del ojo de un sujeto (228); una cámara (108(1), 208) configurada para capturar una imagen (302) actual que contiene los reflejos; y
  - un detector de la mirada ocular (114) que incorpora o acoplado a un corrector de desviación de la mirada relacionada con la posición de la cabeza (116), el detector de la mirada ocular (114) configurado para recibir la imagen actual que contiene los reflejos y estimar una mirada ocular del ojo del sujeto,
  - el corrector de desviación de la mirada relacionada con la posición de la cabeza (116) configurado para recibir la estimación de la mirada ocular de los ojos del sujeto y corregir una desviación en la mirada ocular estimada haciendo coincidir los datos de las características correspondientes a los reflejos y los datos relacionados con la pupila del sujeto para emitir información de la mirada corregida indicativa de donde está mirando actualmente el ojo del sujeto,
  - en donde el corrector de desviación de la mirada relacionada con la posición de la cabeza está configurado para corregir la desviación en la mirada ocular estimada usando una transformación de correlación de homografía adaptativa entrenada minimizando una función objetivo basada en datos correspondientes a una pluralidad de posiciones de la cabeza y direcciones de la mirada para compensar errores de la mirada que varían espacialmente o errores dependientes de la postura de la cabeza en relación con una posición de calibración, o ambos.
2. El sistema de la reivindicación 1 en donde la correlación de homografía adaptativa se aprende en base a una o más variables representativas de las ubicaciones de la cabeza con respecto a una posición de calibración y las posiciones de la mirada basadas al menos en parte en datos simulados.
3. El sistema de la reivindicación 2 en donde los datos simulados están configurados para recoger datos sobre el terreno para entrenar la correlación de homografía adaptativa a través de calibración para obtener las variables de predicción en diversas posiciones de la cabeza.
4. El sistema de la reivindicación 1 en donde la correlación de homografía adaptativa se aprende en base a una o más variables representativas de las ubicaciones de la cabeza con respecto a una posición de calibración corresponden a movimientos de la cabeza relativos entre las diversas posiciones de la cabeza codificadas al menos en parte por transformaciones afines, transformaciones de similitud y/o transformaciones de homografía.
5. Un método para estimar la mirada, que comprende:
  - usar (806) una transformación de correlación de homografía adaptativa para corrección de desviación, en la que la transformación de correlación de homografía adaptativa se entrena minimizando una función objetivo basada en datos correspondientes a una pluralidad de posiciones de la cabeza y direcciones de la mirada para compensar errores de la mirada que varían espacialmente o errores dependientes de la postura de la cabeza en relación con una posición de calibración, o ambos, incluyendo:
  - capturar (802) datos actuales de reflejo ( $g_1 - g_4$ ) y datos relacionados con la pupila en una imagen (302) usando una pluralidad de fuentes de luz (106(1) - 106(m), 222-225) y una cámara (108(1), 208), y
  - proporcionar (804) los datos actuales de reflejo y datos relacionados con la pupila procesados a partir de la imagen como características para obtener datos dependientes de la postura de la cabeza, basados en la transformación de correlación de homografía adaptativa aprendida, que se usan para determinar (808) información de la mirada actual.
6. El método de la reivindicación 5 en donde usar la pluralidad de conjuntos de datos de posición y datos de posición de la pupila comprende usar al menos algunos datos simulados incluyendo predecir los valores de corrección de desviación en diferentes escalas de posición de la cabeza y/o diferentes translaciones de posición de la cabeza.
7. El método de la reivindicación 5 que además comprende:
  - emitir la información de la mirada actual; y
- usar la información actual de la mirada para tomar acciones con respecto a cambiar un estado de una interfaz de usuario.
8. El método de la reivindicación 5 en donde el aprendizaje del modelo de correlación de homografía adaptativa comprende codificar los movimientos relativos de la cabeza mediante transformaciones afines, mediante transformaciones de similitud o transformaciones de homografía.

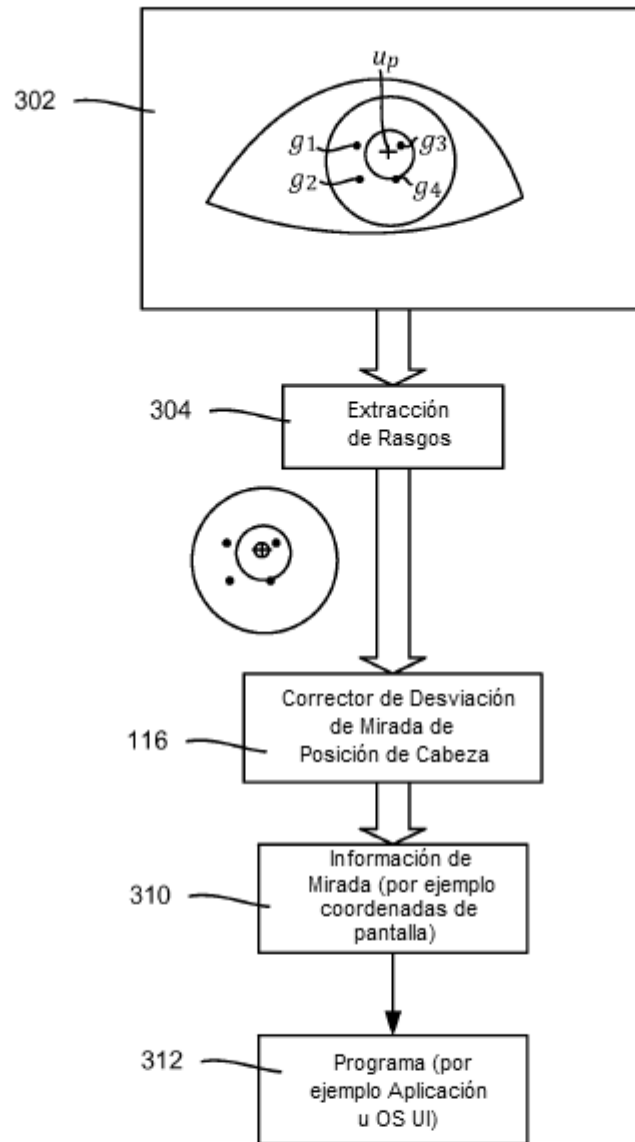
9. El método de la reivindicación 5 en donde el aprendizaje del modelo de correlación de homografía adaptativa comprende codificar datos relacionados con la pupila representativos de las direcciones de la mirada como una o más características.
- 5 10. Uno o más medios de almacenamiento legibles por máquina o lógica de hardware que tienen instrucciones ejecutables, las instrucciones ejecutables que hacen que un procesador realice operaciones que comprenden:  
capturar (802) una imagen (302) incluyendo el ojo de un sujeto (228) a partir del cual se extraen datos de reflejo ( $g_1 - g_4$ ) y datos relacionados con la pupila (804) como características, los datos de reflejo y los datos relacionados con la pupila basados en reflejos de la córnea del ojo del sujeto generados por una pluralidad de fuentes de luz (106(1) - 106(m), 222-225); y
- 10 estimar la mirada ocular, incluyendo usar (806) una transformación de correlación de homografía adaptativa para corrección de desviación, en la que la transformación de correlación de homografía adaptativa se entrena minimizando una función objetiva basada en datos correspondientes a una pluralidad de posiciones de la cabeza y direcciones de la mirada para compensar errores de la mirada que varían espacialmente o errores dependientes de la postura de la cabeza en relación con una posición de calibración, o ambos, incluyendo:
- 15 proporcionar los datos actuales de reflejo y los datos relacionados con la pupila extraídos de la imagen como características para obtener datos dependientes de la postura de la cabeza, basados en la transformación de correlación de homografía adaptativa aprendida, que se usan para determinar (808) una dirección actual de la mirada.



**FIG. 1**

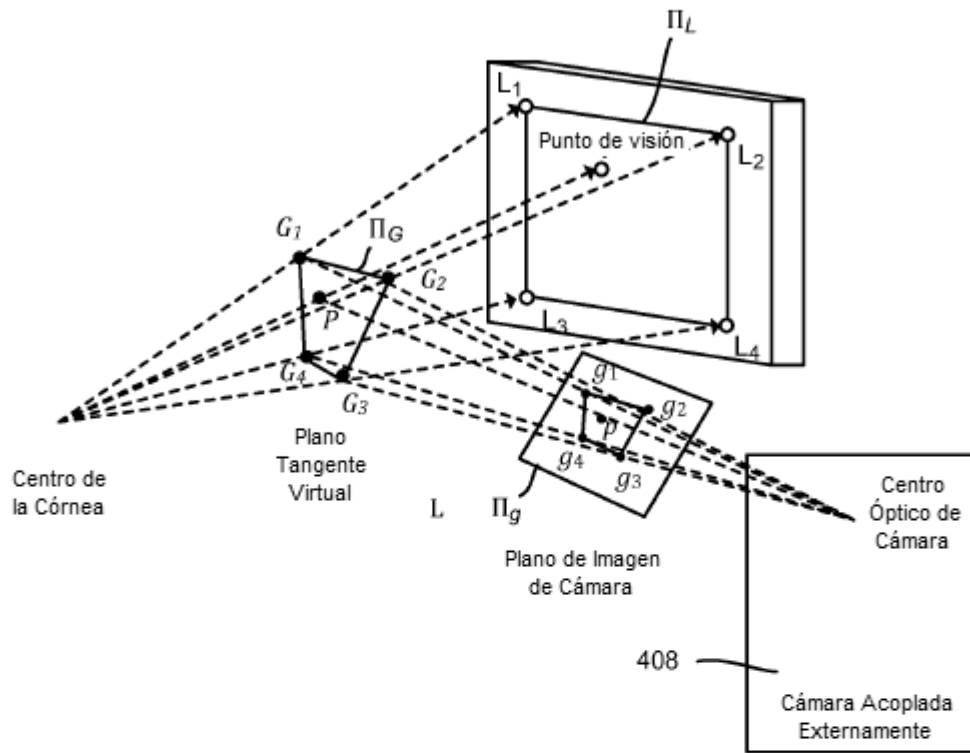


**FIG. 2**

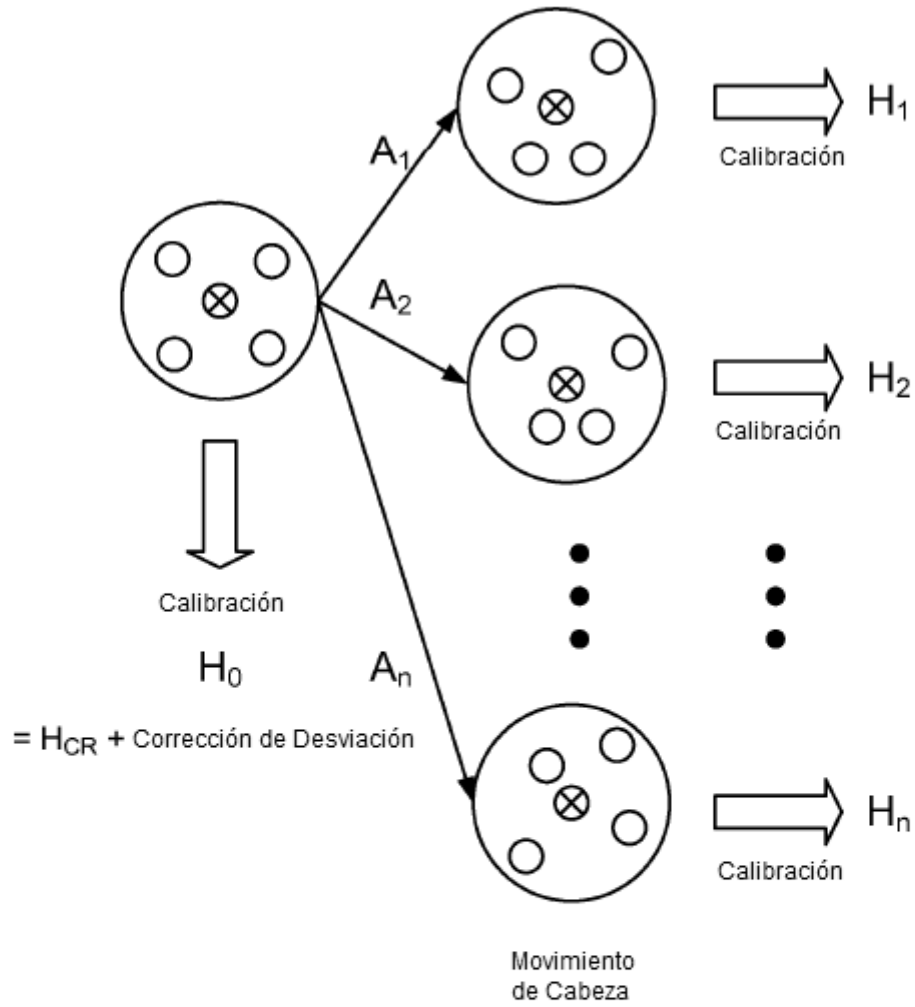


**FIG. 3**

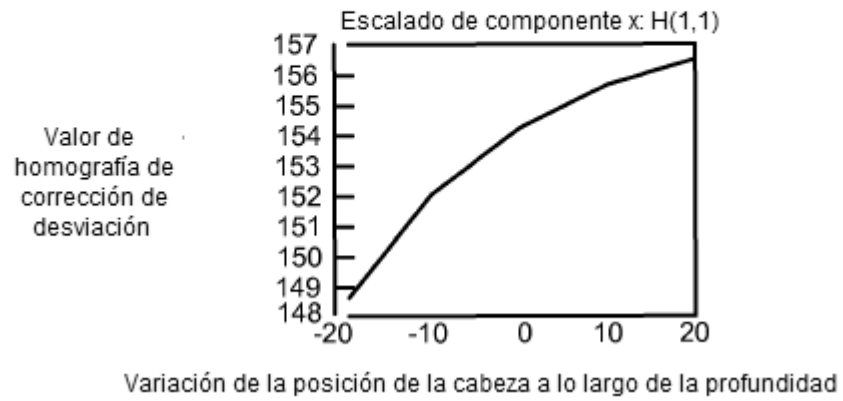




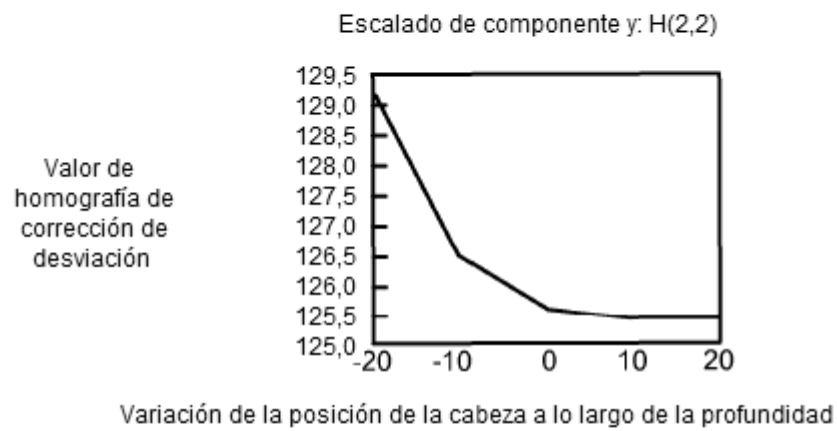
**FIG. 4**



**FIG. 5**



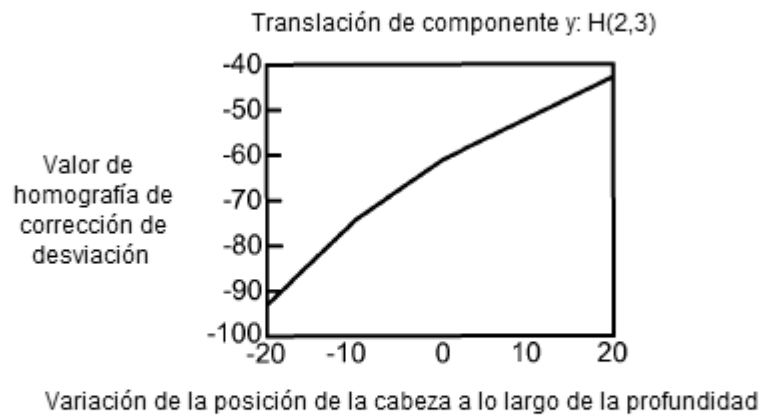
**FIG. 6A**



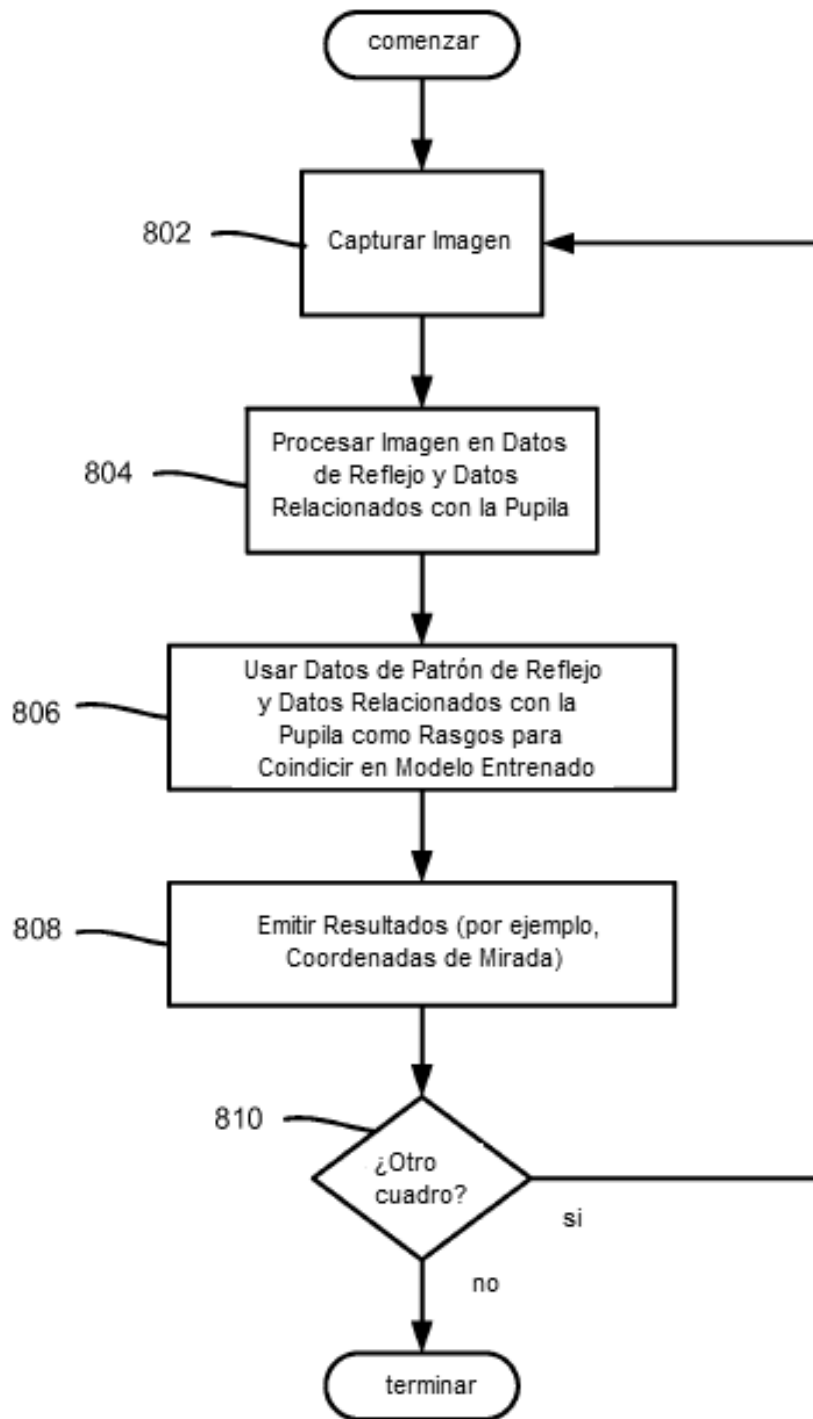
**FIG. 6B**



**FIG. 7A**



**FIG. 7B**



**FIG. 8**

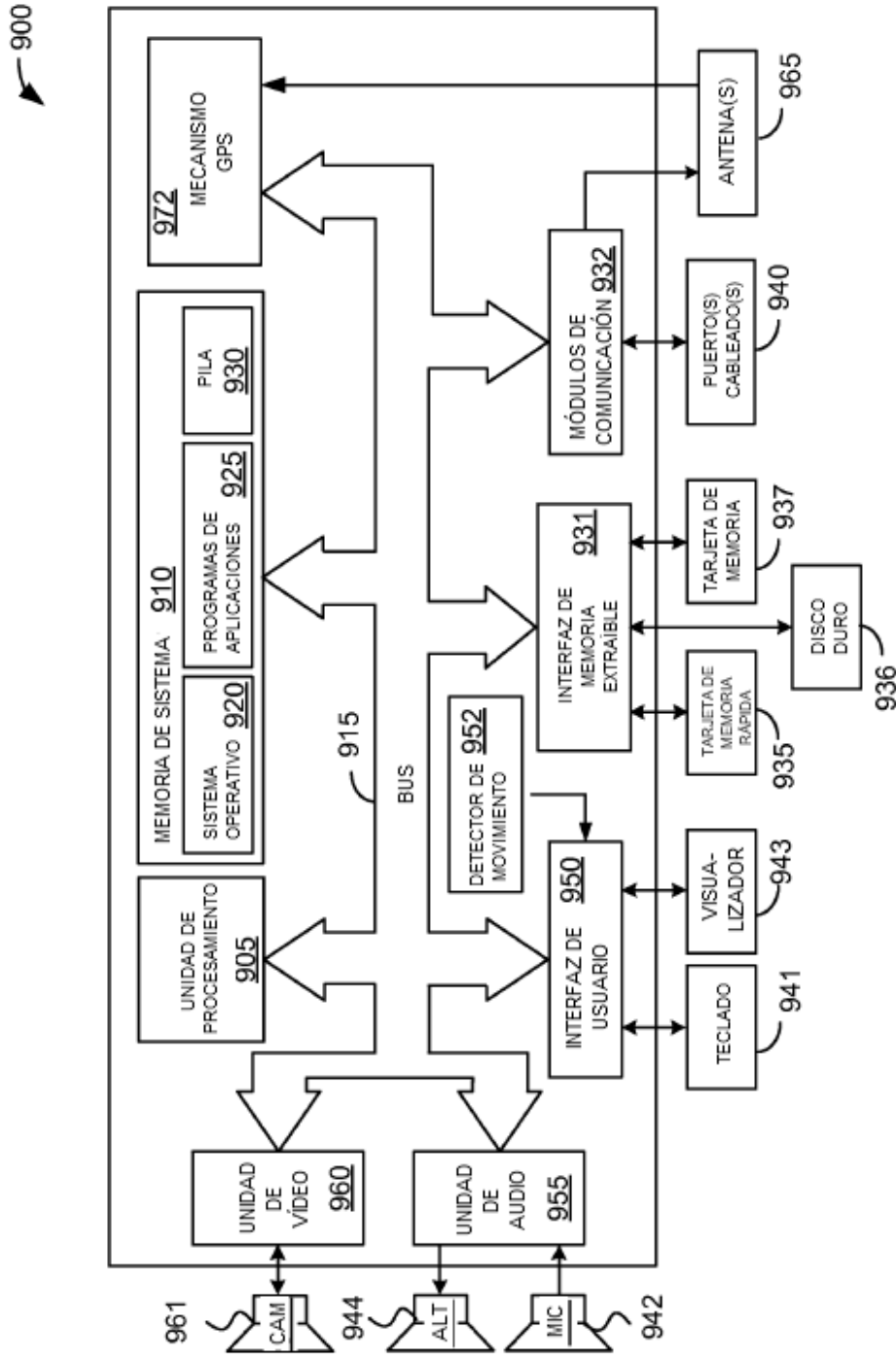


FIG. 9