

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 768 712**

51 Int. Cl.:

**H04N 5/232** (2006.01)

**H04N 5/235** (2006.01)

**G06K 9/46** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.03.2015 PCT/US2015/022418**

87 Fecha y número de publicación internacional: **08.10.2015 WO15153218**

96 Fecha de presentación y número de la solicitud europea: **25.03.2015 E 15716630 (7)**

97 Fecha y número de publicación de la concesión europea: **23.10.2019 EP 3127320**

54 Título: **Sistema y procedimiento para la formación de imágenes de enfoque múltiple**

30 Prioridad:

**03.04.2014 US 201414244595**

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

**23.06.2020**

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)  
International IP Administration, 5775 Morehouse  
Drive  
San Diego, CA 92121-1714, US**

72 Inventor/es:

**NAGARAJA, ADARSH HOSAAGRAHARA;  
LIU, WEILIANG y  
LIU, SHIZHONG**

74 Agente/Representante:

**FORTEA LAGUNA, Juan José**

ES 2 768 712 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 formación de imágenes de enfoque múltiple

## 5 ANTECEDENTES

Campo de la invención

10 [0001] Los sistemas y procedimientos divulgados en el presente documento se refieren en general a la formación de imágenes de enfoque múltiple.

Descripción de la técnica relacionada

15 [0002] Para tomar imágenes de alta calidad con una cámara, los niveles de exposición y de enfoque deben ajustarse bien a los objetos dentro de la escena de interés. Sin embargo, en muchas circunstancias, puede ser difícil obtener una exposición óptima y un nivel de enfoque para todos los objetos dentro de la escena. Por ejemplo, puede ser difícil mantener los objetos enfocados que están dispersos a diferentes distancias a lo largo de un eje (comúnmente el eje z) que define la profundidad de la pantalla que se extiende desde la lente de la cámara hasta el infinito. La profundidad de campo es típicamente el área del eje z en la cual están enfocados los objetos de la escena de interés. Cuando la lente de la cámara se amplía, la profundidad de campo es menor que cuando la lente de la cámara se aleja. Por esta razón, puede ser difícil mantener el enfoque en varios objetos localizados en diferentes posiciones a lo largo de la profundidad del eje z del campo.

25 [0003] El documento US 2012/120269 se refiere a la detección de rostros en imágenes, en particular al reconocimiento de rostros desenfocados en una imagen. La divulgación es que, durante la captura de una imagen, un componente MEMS ajusta una óptica en tiempo real a una primera posición de enfoque para enfocar un primer rasgo característico, y el componente MEMS ajusta en tiempo real a una segunda posición de enfoque para enfocarse en un segundo rasgo característico dentro de la escena. Se divulga que se puede formar una imagen compuesta que incluya el primer rasgo característico de una primera imagen cuando el componente MEMS ajusta la óptica para enfocar en el primer rasgo característico, e incluir el segundo rasgo característico de una segunda imagen cuando el componente MEMS ajusta la óptica para enfocar en el segundo rasgo característico.

35 [0004] El documento US 2011/267530 se refiere al uso de un terminal móvil con una pantalla táctil para capturar una imagen en la cual se visualice en la pantalla táctil para determinar el área de enfoque de la imagen.

## SUMARIO

40 [0005] Los sistemas, procedimientos y dispositivos de la divulgación tienen cada uno varios aspectos innovadores, ninguno de los cuales es el único responsable de los atributos deseables divulgados en el presente documento. Las combinaciones de las innovaciones, aspectos y rasgos característicos descritos en el presente documento se pueden incorporar en diversos modos de realización de sistemas, procedimientos y dispositivos y dichas combinaciones no están limitadas por los ejemplos de modos de realización descritos en el presente documento.

45 [0006] Los modos de realización e innovaciones descritos en el presente documento se refieren a sistemas y procedimientos que se pueden ejecutar en un procesador para un dispositivo electrónico para obtener una imagen que tiene múltiples regiones de interés enfocadas. Se presenta una experiencia de usuario intuitiva para permitir al usuario enfocar en diferentes áreas a diferentes profundidades dentro de la escena dentro de una imagen individual.

50 [0007] En algunos modos de realización, la cámara se selecciona y apunta a una escena que contiene múltiples regiones de interés (ROI). El mecanismo de enfoque automático comienza a ejecutarse y se pueden sugerir varias ROI al usuario en una vista previa de la imagen. Estas ROI se pueden determinar usando diversos procedimientos conocidos, que incluyen, pero no se limitan a, (a) detección de rostros, (b) seguimiento de objetos o (c) entrada directa del usuario. El usuario puede seleccionar múltiples ROI durante la vista previa tocando o presionando la pantalla para seleccionar y ajustar la selección deseada. En base al tacto del usuario, el enfoque automático calcula los parámetros para cada región de interés y almacena esta información.

60 [0008] En algunos modos de realización, cuando el usuario presiona el botón del obturador para capturar la imagen, se capturan múltiples imágenes en base al número de ROI seleccionadas (o, de manera equivalente, el número de parámetros de enfoque automático calculados y almacenados). Cada una de las múltiples imágenes contiene una de las ROI enfocadas. A continuación se genera una imagen individual a partir de las múltiples imágenes, apareciendo la imagen individual formada a la perfección a partir de las múltiples imágenes. La imagen individual resultante contiene todas las múltiples regiones de interés enfocadas.

65 [0009] Un aspecto se refiere a un sistema para sugerir múltiples regiones de interés dentro de una escena y generar una imagen de enfoque múltiple a partir de las múltiples regiones de interés. El sistema incluye un módulo de control configurado para determinar regiones de interés candidatas dentro de una escena, indicar sugerencias de regiones

de interés a partir de las posibles regiones de interés dentro de una escena en una pantalla de visualización, recibir entrada de datos que indique la selección de múltiples regiones de interés, determinar parámetros de enfoque automático asociados con cada una de las múltiples regiones de interés seleccionadas, capturar una imagen de cada región de interés usando los parámetros de enfoque automático asociados con la región de interés y generar una imagen de enfoque múltiple que incluya datos de imagen de la imagen que incluya cada región de interés. El sistema puede incluir además un dispositivo de formación de imágenes que incluya un sensor en el que el módulo de control esté configurado para capturar una imagen de cada región de interés usando el dispositivo de formación de imágenes y los parámetros de enfoque automático asociados con la región de interés. El módulo de control se puede configurar además para hacer funcionar el dispositivo de formación de imágenes en un modo de vista previa. El módulo de control se puede configurar además para visualizar gráficamente las múltiples regiones de interés sugeridas para el usuario en el modo de vista previa. El módulo de control se puede configurar además para determinar las regiones de interés dentro de una escena usando un algoritmo de detección de rostros o un algoritmo de seguimiento de objetos. El módulo de control se puede configurar además para recibir entrada de datos con respecto a la determinación de regiones de interés candidatas dentro de una escena. El dispositivo de formación de imágenes puede incluir además una pantalla táctil, y el módulo de control se puede configurar además para recibir entrada de datos que indique una selección de múltiples regiones de interés al registrar un evento táctil en la pantalla táctil. El sistema puede incluir además un dispositivo móvil en el que el módulo de control puede ser un componente de una aplicación de cámara del dispositivo móvil.

**[0010]** En otro aspecto, un procedimiento para sugerir múltiples regiones de interés dentro de una escena y generar una imagen de enfoque múltiple a partir de las múltiples regiones de interés incluye los pasos para determinar las posibles regiones de interés dentro de una escena, indicar sugerencias de regiones de interés de las posibles regiones de interés dentro de una escena en una pantalla de visualización, recibir entrada de datos con respecto a la selección de múltiples regiones de interés, determinar los parámetros de enfoque automático asociados con cada una de las múltiples regiones de interés seleccionadas, capturar una imagen de cada región de interés usando los parámetros de enfoque automático asociados con la región de interés, y generar una imagen de enfoque múltiple que incluya datos de imagen de la imagen que incluya cada región de interés. El procedimiento puede incluir además el paso de hacer funcionar un dispositivo de formación imágenes en un modo de vista previa. En algunos modos de realización, la determinación de regiones de interés candidatas dentro de una escena incluye además realizar uno o más de detección de rostros y de seguimiento de objetos en rostros u objetos dentro de la escena. El procedimiento puede incluir además el paso de mostrar gráficamente las múltiples regiones de interés sugeridas para el usuario en el modo de vista previa. En algunos modos de realización, aceptar la entrada de usuario con respecto a la selección de las múltiples regiones de interés incluye además registrar un evento táctil en una pantalla táctil del dispositivo de formación de imágenes. En algunos modos de realización, el dispositivo de formación de imágenes está configurado para funcionar como parte de una aplicación de cámara para un dispositivo móvil.

**[0011]** En otro aspecto más, un aparato para sugerir múltiples regiones de interés dentro de una escena y generar una imagen de enfoque múltiple a partir de las múltiples regiones de interés incluye medios para determinar regiones de interés candidatas dentro de una escena, medios para sugerir múltiples regiones de interés a partir de las posibles regiones de interés dentro de una escena para un usuario, medios para aceptar la entrada de usuario con respecto a la selección de las múltiples regiones de interés, medios para determinar los parámetros de enfoque automático asociados con cada una de las múltiples regiones de interés, medios para capturar una imagen individual de cada región de interés usando los parámetros de enfoque automático asociados con la región de interés, y medios para generar una imagen individual de enfoque múltiple a partir de la imagen individual de cada región de interés. El aparato puede incluir además medios para hacer funcionar un dispositivo de formación de imágenes en un modo de vista previa. En algunos modos de realización, los medios para determinar regiones de interés candidatas dentro de una escena incluyen además medios para realizar uno o más de detección de rostros y de seguimiento de objetos en rostros u objetos dentro de la escena. En algunos modos de realización, el aparato puede incluir además medios para visualizar gráficamente las múltiples regiones de interés sugeridas para el usuario en el modo de vista previa. En algunos modos de realización, los medios para aceptar la entrada de usuario con respecto a la selección de las múltiples regiones de interés incluyen además medios para registrar un evento táctil en una pantalla táctil del dispositivo de formación de imágenes.

**[0012]** En otro aspecto, un medio no transitorio legible por ordenador almacena instrucciones que, cuando se ejecutan, causan que al menos un procesador físico informático realice un procedimiento para sugerir múltiples regiones de interés dentro de una escena y generar una imagen de enfoque múltiple a partir de las múltiples regiones de interés. El procedimiento incluye los pasos de determinar múltiples regiones posibles o de interés candidatas dentro de una escena, sugerir múltiples regiones de interés a partir de las posibles regiones de interés dentro de una escena a un usuario, aceptar la entrada de usuario con respecto a la selección de las múltiples regiones de interés, determinar los parámetros de enfoque automático asociados con cada una de las múltiples regiones de interés, capturar una imagen individual de cada región de interés usando los parámetros de enfoque automático asociados con la región de interés, y generar una imagen individual de enfoque múltiple a partir de la imagen individual de cada región de interés. En algunos modos de realización, el procedimiento incluye el paso de hacer funcionar un dispositivo de formación de imágenes en un modo de vista previa. En algunos modos de realización, la determinación de múltiples regiones de interés posibles o candidatas dentro de una escena incluye además realizar uno o más de detección de rostros y de seguimiento de objetos en rostros u objetos dentro de la escena. En algún modo de realización, el

procedimiento incluye además el paso de visualizar gráficamente las múltiples regiones de interés sugeridas para el usuario en el modo de vista previa. En algunos modos de realización, aceptar la entrada de usuario con respecto a la selección de las múltiples regiones de interés incluye además registrar un evento táctil en una pantalla táctil del dispositivo de formación de imágenes. En algunos modos de realización, el dispositivo de formación de imágenes está configurado para funcionar como parte de una aplicación de cámara para un dispositivo móvil.

**BREVE DESCRIPCIÓN DE LOS DIBUJOS**

[0013] Los aspectos divulgados se describirán a continuación en el presente documento junto con los dibujos adjuntos, proporcionados para ilustrar y no para limitar los aspectos divulgados, en los que designaciones similares indican elementos similares.

La Figura 1 es un diagrama de bloques que representa un sistema que implementa algunos elementos operativos para obtener una imagen que tiene múltiples regiones de interés enfocadas.

La Figura 2 es un diagrama de flujo que ilustra un proceso para obtener una imagen que tenga múltiples regiones de interés enfocadas.

La Figura 3 es un ejemplo de una interfaz de usuario que representa sugerencias para regiones de interés enfocadas y regiones de interés seleccionadas por el usuario.

La Figura 4 es un ejemplo de una imagen de enfoque múltiple que tiene múltiples regiones de enfoque.

**DESCRIPCIÓN DETALLADA**

[0014] Típicamente, las cámaras en dispositivos inteligentes, tales como teléfonos móviles, tablets u otros dispositivos móviles, no pueden capturar imágenes de objetos con claridad a diferentes profundidades. La mayoría de los dispositivos actuales no pueden capturar y ni enfocar en múltiples objetos de interés en una imagen individual. Algunas cámaras se pueden dirigir para enfocar múltiples objetos de interés a diferentes profundidades, pero la posición de la lente para adquirir esta imagen "compromete" el enfoque de cada objeto en la escena de interés en lugar de producir una imagen individual con todos los objetos de interés enfocados. Los procedimientos y sistemas analizados a continuación proporcionan soluciones para obtener una imagen individual y enfocada de múltiples objetos localizados a diferentes profundidades de enfoque.

[0015] En algunos modos de realización, el proceso de formación de imágenes de enfoque múltiple propuesto dirige un procesador para sugerir varias regiones de interés al usuario durante un modo de vista previa. Estas sugerencias se pueden hacer en base a la detección de rostros, al seguimiento de objetos o a otros procedimientos de identificación de rasgos característicos u objetos. El usuario puede seleccionar una o más regiones de interés para incluirlas en la imagen final de enfoque múltiple. A continuación, el procesador calcula los parámetros de enfoque automático para cada región de interés y almacena esta información. Cuando el usuario presiona el botón del obturador para capturar la imagen, se capturan múltiples imágenes en base al número de ROI seleccionadas (o, de manera equivalente, el número de parámetros de enfoque automático calculados y almacenados). Cada una de las múltiples imágenes contiene de forma deseable una de las ROI enfocadas. A continuación, se genera una imagen individual a partir de las imágenes múltiples, la imagen individual aparece como formada sin problemas a partir de las múltiples imágenes. La imagen individual resultante contiene de forma deseable todas las múltiples regiones de interés enfocadas.

[0016] Cabe destacar que los ejemplos se pueden describir como un proceso, que se representa como un organigrama, un diagrama de flujo, un diagrama de estados finitos, un diagrama estructural o un diagrama de bloques. Aunque un organigrama puede describir las operaciones como un proceso secuencial, muchas de las operaciones se pueden realizar en paralelo o simultáneamente, y el proceso se puede repetir. Además, el orden de las operaciones se puede disponer. Un proceso termina cuando se completan sus operaciones. Un proceso puede corresponder a un procedimiento, una función, un proceso, una subrutina, un subprograma, etc. Cuando un proceso corresponde a una función de software, su finalización corresponde al retorno de la función a la función de llamada o a la función principal.

[0017] Los modos de realización se pueden implementar en sistema en chip (SoC) o en hardware externo, software, firmware o en cualquier combinación de los mismos. Los expertos en la técnica entenderán que la información y las señales se pueden representar usando cualquiera entre varias tecnologías y técnicas diferentes. Por ejemplo, los datos, instrucciones, comandos, información, señales, bits, símbolos y segmentos que se pueden haber mencionado a lo largo de la descripción anterior se pueden representar mediante tensiones, corrientes, ondas electromagnéticas, campos o partículas magnéticas, campos o partículas ópticas o cualquier combinación de los mismos.

[0018] En la siguiente descripción, se dan detalles específicos para proporcionar una comprensión exhaustiva de los ejemplos. Sin embargo, un experto en la técnica entenderá que se pueden llevar a la práctica los ejemplos sin estos detalles específicos. Por ejemplo, los componentes/dispositivos eléctricos se pueden mostrar en diagramas de

bloques con el fin de no oscurecer los ejemplos con detalles innecesarios. En otros casos, dichos componentes, otras estructuras y técnicas se pueden mostrar en detalle para explicar mejor los ejemplos.

Visión general del sistema

5 **[0019]** La **Figura 1** ilustra un ejemplo de un sistema de formación de imágenes 100 configurado para obtener una imagen individual que tenga múltiples regiones de interés enfocadas. El modo de realización ilustrado no pretende ser limitante, sino que es más bien ilustrativo de determinados componentes en algunos modos de realización. El sistema 100 puede incluir una variedad de otros componentes para otras funciones que no se muestran para mayor claridad de los componentes ilustrados.

15 **[0020]** El sistema de formación de imágenes 100 puede incluir un dispositivo de formación de imágenes 110 y una pantalla electrónica 130. Determinados modos de realización de la pantalla electrónica 130 pueden ser cualquier tecnología de pantalla de panel plano, tal como un LED, un LCD, un plasma o una pantalla de proyección. La pantalla electrónica 130 puede estar acoplada al procesador 120 para recibir información para la visualización visual a un usuario. Dicha información puede incluir, pero no se limita a, representaciones visuales de archivos almacenados en una localización de memoria, aplicaciones de software instaladas en el procesador 120, interfaces de usuario y objetos de contenido accesibles en red.

20 **[0021]** El dispositivo de formación de imágenes 110 puede emplear uno o una combinación de sensores de formación de imágenes. El sistema de formación de imágenes 100 puede incluir además un procesador 120 conectado al dispositivo de formación de imágenes 110. Una memoria de trabajo 135, una pantalla electrónica 130 y una memoria de programa 140 también están en comunicación con el procesador 120. El sistema de formación de imágenes 100 puede ser un dispositivo móvil, tal como una tablet, un ordenador portátil o un teléfono móvil.

25 **[0022]** El procesador 120 puede ser una unidad de procesamiento de uso general o puede ser un procesador diseñado especialmente para aplicaciones de formación de imágenes para un dispositivo electrónico manual. Como se muestra, el procesador 120 está conectado a, y en comunicación de datos con, la memoria de programa 140 y una memoria de trabajo 135. En algunos modos de realización, la memoria de trabajo 135 se puede incorporar en el procesador 120, por ejemplo, la memoria caché. La memoria de trabajo 135 también puede ser un componente separado del procesador 120 y acoplado al procesador 120, por ejemplo, uno o más componentes RAM o DRAM. En otras palabras, aunque la Figura 1 ilustra dos componentes de memoria, incluyendo el componente de memoria 140 que comprende varios módulos y una memoria separada 135 que comprende una memoria de trabajo, un experto en la técnica reconocerá varios modos de realización que usan diferentes arquitecturas de memoria. Por ejemplo, un diseño puede usar la memoria ROM o RAM estática para el almacenamiento de instrucciones del procesador que implementen los módulos contenidos en la memoria 140. Las instrucciones del procesador se pueden cargar a continuación en la RAM para facilitar la ejecución mediante el procesador. Por ejemplo, la memoria de trabajo 135 puede ser una memoria RAM, con instrucciones cargadas en la memoria de trabajo 135 antes de su ejecución por el procesador 120.

35 **[0023]** En el modo de realización ilustrado, la memoria de programa 140 almacena un módulo de captura de imágenes 145, un módulo de sugerencia de región de interés (ROI) 150, un módulo de determinación de parámetros de enfoque automático 155, un módulo de generación de imagen individual 160, un sistema operativo 165 y un módulo de interfaz de usuario 170. Estos módulos pueden incluir instrucciones que configuren el procesador 120 para realizar diversas tareas de procesamiento de imágenes y de gestión de dispositivos. La memoria de programa 140 puede ser cualquier medio de almacenamiento adecuado legible por ordenador, tal como un medio no transitorio de almacenamiento. La memoria de trabajo 135 se puede usar por el procesador 120 para almacenar un conjunto de trabajo de instrucciones de procesador contenidas en los módulos de la memoria 140. De forma alternativa, la memoria de trabajo 135 también se puede usar por el procesador 120 para almacenar datos dinámicos creados durante el funcionamiento del sistema de formación de imágenes 100.

40 **[0024]** Como se mencionó anteriormente, el procesador 120 está configurado por varios módulos almacenados en la memoria 140. En otras palabras, el proceso 120 puede ejecutar instrucciones almacenadas en módulos en la memoria 140. El módulo de captura de imágenes 145 puede incluir instrucciones que configuren el procesador 120 para obtener imágenes del dispositivo de formación de imágenes. Por lo tanto, el procesador 120, junto con el módulo de captura de imágenes 145, el dispositivo de formación de imágenes 110 y la memoria de trabajo 135, representan un medio para obtener datos del sensor de imagen de una o más regiones de interés.

45 **[0025]** Todavía en referencia a la Figura 1, la memoria 140 también puede contener el módulo de sugerencia de ROI 150. El módulo de sugerencia de ROI 150 puede incluir instrucciones que configuren el procesador 120 para sugerir regiones de interés para el usuario, usando la detección de rostros, el seguimiento de objetos u otros algoritmos, y para visualizar las regiones de interés para el usuario en la pantalla electrónica 130 durante un modo de vista previa del dispositivo de formación de imágenes 110, como se describirá con más detalle a continuación. Por lo tanto, el procesador 120, junto con el módulo de sugerencia de ROI 150, la memoria de trabajo 135 y la pantalla electrónica 130 representan un medio para sugerir y mostrar una o más regiones de interés para un dispositivo electrónico de usuario.

**[0026]** La memoria 140 también puede contener el módulo de determinación de parámetros de enfoque automático 155. El módulo de determinación de parámetros de enfoque automático 155 puede incluir instrucciones que configuren el procesador 120 para realizar una función de enfoque automático y calcular y almacenar los parámetros de enfoque automático para cada una de las regiones de interés identificadas. Por ejemplo, si se seleccionan tres regiones de interés, el procesador 120 puede recibir instrucciones del módulo de determinación de parámetros de enfoque automático 155 para calcular los parámetros de enfoque automático correspondientes a cada región de interés y almacenar los parámetros de enfoque automático en la memoria de trabajo 135 o el dispositivo de almacenamiento 125. A continuación, el procesador 120 puede recibir instrucciones del módulo de captura de imágenes 145 para, usando el dispositivo de formación de imágenes 110, capturar imágenes individuales de cada región de interés en base a los parámetros de enfoque automático de cada región de interés. Por lo tanto, el procesador 120, junto con el módulo de sugerencia de ROI 150, el módulo de determinación de parámetros de enfoque automático 155 y la memoria de trabajo 135 representan un medio para calcular y almacenar parámetros de enfoque automático para cada región de interés seleccionada.

**[0027]** La memoria 140 también puede contener un módulo de generación de imagen individual 160. El módulo de generación de imagen individual 160 ilustrado en la Figura 1 puede incluir instrucciones que configuren el procesador 120 para generar una imagen individual de enfoque múltiple que contenga las múltiples regiones de interés identificadas por el módulo de sugerencia de ROI 150. Por ejemplo, si se seleccionan tres regiones de interés, los parámetros de enfoque automático se calculan y almacenan para cada región de interés, y se toman múltiples imágenes en las cuales cada región de interés está enfocada en una de las múltiples imágenes, el procesador 120 puede recibir instrucciones del módulo de generación de imagen individual 160 para generar una imagen individual de enfoque múltiple integrando sin problemas las múltiples imágenes. Por lo tanto, el procesador 120, junto con el módulo de sugerencia de ROI 150, el módulo de generación de imagen individual 160 y la memoria de trabajo 135 representan un medio para generar una imagen individual a partir de múltiples imágenes de regiones de interés.

**[0028]** La memoria 140 también puede contener el módulo de interfaz de usuario 170. El módulo de interfaz de usuario 170 ilustrado en la Figura 1 puede incluir instrucciones que configuren el procesador 120 para proporcionar una colección de objetos en pantalla y controles suaves que permitan al usuario interactuar con el dispositivo, tales como permitir que el usuario seleccione regiones de interés identificadas y visualizadas en un modo de vista previa del dispositivo de formación de imágenes. El módulo de interfaz de usuario 170 también permite que las aplicaciones interactúen con el resto del sistema. Un módulo de sistema operativo 165 también puede residir en la memoria 140 y funcionar con el procesador 120 para gestionar la memoria y los recursos de procesamiento del sistema 100. Por ejemplo, el sistema operativo 165 puede incluir controladores de dispositivo para gestionar recursos de hardware tales como la pantalla electrónica 130 o el dispositivo de formación de imágenes 110. En algunos modos de realización, las instrucciones contenidas en el módulo de sugerencia de ROI 150 y en el módulo de determinación de parámetros de enfoque automático 155 pueden no interactuar directamente con estos recursos de hardware, sino que, en lugar de eso, interactúan a través de API o subrutinas estándar localizadas en el sistema operativo 165. Las instrucciones dentro del sistema operativo 165 pueden interactuar a continuación directamente con estos componentes de hardware.

**[0029]** El procesador 120 puede escribir datos en el módulo de almacenamiento 125. Si bien el módulo de almacenamiento 125 se representa gráficamente como una unidad de disco tradicional, los expertos en la técnica comprenderían que múltiples modos de realización podrían incluir un dispositivo de almacenamiento basado en disco o uno de varios otros tipos de medios de almacenamiento, incluyendo un disco de memoria, una unidad USB, una unidad flash, un medio de almacenamiento conectado de forma remota, un controlador de disco virtual o similar.

**[0030]** Aunque la Figura 1 representa un dispositivo que tiene componentes independientes para incluir un procesador, un sensor de formación de imágenes y la memoria, un experto en la técnica reconocerá que estos componentes separados se pueden combinar de una variedad de formas para lograr objetivos de diseño particulares. Por ejemplo, en un modo de realización alternativo, los componentes de memoria se pueden combinar con componentes de procesador para ahorrar costes y mejorar el rendimiento.

**[0031]** Adicionalmente, aunque la Figura 1 ilustra dos componentes de memoria, incluyendo el componente de memoria 140 que comprende varios módulos y una memoria separada 135 que comprende una memoria de trabajo, un experto en la técnica reconocerá varios modos de realización que usen diferentes arquitecturas de memoria. Por ejemplo, un diseño puede usar la memoria ROM o RAM estática para el almacenamiento de instrucciones del procesador que implementen los módulos contenidos en la memoria 140. De forma alternativa, las instrucciones del procesador se pueden leer durante el arranque del sistema desde un dispositivo de almacenamiento en disco que está integrado en el sistema de formación de imágenes 100 o conectado por medio de un puerto de dispositivo externo. Las instrucciones del procesador se pueden cargar a continuación en la RAM para facilitar la ejecución mediante el procesador. Por ejemplo, la memoria de trabajo 135 puede ser una memoria RAM, con instrucciones cargadas en la memoria de trabajo 135 antes de su ejecución por el procesador 120.

Visión general del procedimiento

[0032] Los modos de realización de la invención se refieren a un proceso para la formación de imágenes de enfoque múltiple, que puede incluir recomendar una pluralidad de regiones de interés candidatas dentro de una única escena de interés para el usuario, recibir entrada de datos con respecto a la determinación de la región de interés candidata dentro de una escena, recibir entrada de datos sobre la selección de las regiones de interés deseadas que se capturarán dentro de la escena individual, obtener múltiples imágenes en base a parámetros de enfoque automático asociados con las regiones de interés deseadas o seleccionadas, e integrar sin problemas las múltiples imágenes en una imagen individual de enfoque múltiple. Es decir, por ejemplo, una imagen individual que tiene múltiples regiones en diferentes planos de objetos (o profundidades), cada uno de los cuales está enfocado en la imagen integrada. Los ejemplos se pueden describir como un proceso, que se representa como un organigrama, un diagrama de flujo, un diagrama de estados finitos, un diagrama estructural o un diagrama de bloques. Aunque un organigrama puede describir las operaciones como un proceso secuencial, muchas de las operaciones se pueden realizar en paralelo o simultáneamente, y el proceso se puede repetir. Además, el orden de las operaciones se puede rediseñar. Un proceso termina cuando se completan sus operaciones. Un proceso puede corresponder a un procedimiento, una función, un proceso, una subrutina, un subprograma, etc. Cuando un proceso corresponde a una función de software, su finalización corresponde al retorno de la función a la función de llamada o a la función principal.

[0033] La **Figura 2** ilustra un modo de realización de un proceso 200 para configurar un dispositivo electrónico que tenga un dispositivo de formación de imágenes (tal como el dispositivo de imágenes 110 mostrado en la Figura 1) para realizar la formación de imágenes de enfoque múltiple que se puede implementar en uno o más de los módulos representados en la Figura 1. En algunos ejemplos, el proceso 200 se puede ejecutar en un procesador, por ejemplo, el procesador 120 (Figura 1), y en otros componentes ilustrados en la Figura 1 que están almacenados en la memoria 140 o que están incorporados en otro hardware o software.

[0034] El proceso 200 comienza en el bloque de arranque 202 y pasa al bloque 204 en el que el dispositivo de formación de imágenes 110 se hace funcionar en un modo de vista previa. En algunos modos de realización, un modo de vista previa puede incluir visualizar en una pantalla electrónica, por ejemplo, la pantalla electrónica 130 (Figura 1) la escena como la ve la lente del dispositivo de formación de imágenes 110. El dispositivo electrónico puede ser un dispositivo de comunicación manual, por ejemplo, un teléfono móvil o "teléfono inteligente", o un asistente móvil de datos personales (PDA) que incluye una tablet. Por ejemplo, el usuario puede hacer funcionar la cámara de un dispositivo móvil en un modo de vista previa para determinar si capturar o no una imagen. En otro ejemplo, el usuario puede hacer funcionar la cámara del dispositivo móvil en un modo de vista previa para determinar qué regiones de interés le gustaría capturar en una imagen de enfoque múltiple.

[0035] El proceso 200 pasa a continuación al bloque 206, en el que el procesador, tal como el procesador 120, recibe instrucciones para determinar posibles regiones de interés dentro de la escena actual. En algunos modos de realización, el procesador puede usar la detección de rostros o el seguimiento de objetos para identificar posibles regiones de interés. En un modo de realización, se puede usar un algoritmo de detección de rostros, por ejemplo, el algoritmo de detección de rostros OMRON™. Sin embargo, el sistema descrito es preferentemente insensible al algoritmo de detección de rostros usado siempre que el algoritmo seleccionado sea capaz de detectar rostros y genere una ventana precisa de región de interés (ROI) alrededor de los rostros detectados. En algunos modos de realización, el algoritmo de seguimiento de objetos podría ser similar a uno usado en casos de uso táctil para rastrear (T2T) desarrollado en el procesador Snapdragon™ de QUALCOMM Inc.

[0036] En un ejemplo, la detección de rostros puede determinar que aparezcan tres rostros a diversas profundidades dentro de la escena. En otro ejemplo, el seguimiento de objetos podría identificar dos o más elementos dentro de una escena, por ejemplo, un cubo, un castillo de arena y un pájaro en la playa. Una vez que se han determinado las posibles regiones de interés, el proceso 200 pasa al bloque 208.

[0037] En el bloque 208, el procesador recibe instrucciones de sugerir regiones de interés para el usuario y visualizar estas regiones de interés en el modo de vista previa del dispositivo de formación de imágenes. Para continuar con los ejemplos analizados anteriormente, si el procesador detecta una pluralidad de rostros dentro de la escena, tal como los tres rostros a diferentes profundidades de escena, el procesador podría recibir instrucciones para indicar cada rostro como una posible región de interés visualizando un rectángulo, un círculo, unos paréntesis u otro indicador alrededor de cada rostro. El tamaño de la región de interés puede depender de si el sistema sugiere la ROI o si el usuario selecciona la ROI. Si el sistema sugiere la ROI, por ejemplo, a través de una operación de detección de rostros o de detección de objetos, la ventana de la ROI será la forma, tal como un rectángulo, un círculo, una elipse, etc., que puede contener completamente el objeto sugerido. En algunos modos de realización, la ROI podría ser un rectángulo dibujado alrededor del rostro. En algunos modos de realización, el tamaño de la ROI podría ser tan pequeño como unos pocos píxeles o tan grande como una pantalla de visualización completa, dependiendo del área de la imagen ocupada por la ROI detectada. En algunos modos de realización, si el usuario selecciona la ROI, la ROI podría ser un rectángulo u otra forma de un tamaño fijo. En otros modos de realización, el usuario podría indicar una ROI seleccionando dos o más puntos que rodean la ROI para definir una forma que rodea la ROI. En otro ejemplo, el procesador podría recibir instrucciones para visualizar un puntero o un indicador manual en cada rostro o región de interés para indicar claramente al usuario estas posibilidades de formación de imágenes de enfoque múltiple. El proceso 200 pasa a continuación al bloque 210, en el que el procesador recibe instrucciones para aceptar la entrada de usuario con respecto a una o más regiones de interés. Por ejemplo, para un dispositivo electrónico configurado

con una pantalla táctil, el usuario puede seleccionar uno o más rostros visualizados en un modo de vista previa seleccionando cada rostro en la pantalla táctil. El usuario también puede usar un ratón, un puntero o un botón para seleccionar una pluralidad de regiones de interés. En la Figura 3 se muestra un ejemplo de una interfaz de usuario que ilustra a un usuario que selecciona múltiples regiones de interés. Como se muestra en la Figura 3, el usuario puede dirigir un puntero a cada una de las múltiples regiones de interés. En este ejemplo, el usuario ha seleccionado el texto "Enfocar", el texto "Arrastrar", la pieza de ajedrez del caballo y las flores.

**[0038]** El procesador 200 pasa a continuación al bloque 212, en el que el procesador recibe instrucciones para determinar los parámetros de enfoque automático relacionados con las ROI seleccionadas. En algunos modos de realización, el procesador puede recibir instrucciones para realizar una función de enfoque automático del dispositivo de formación de imágenes 110 para determinar los parámetros de enfoque automático de modo que cada una de las regiones de interés seleccionadas esté enfocada. Los parámetros de enfoque automático son el conjunto de comandos enviados a un controlador de sensor del procesador 120. El controlador de sensor incluye configuraciones o instrucciones para mover la posición de la lente del dispositivo de formación de imágenes 110. En algunos modos de realización, el controlador de sensor podría incluir parámetros para ajustar el enfoque automático, el balance de blancos automático y la configuración de exposición automática del dispositivo de formación de imágenes 110. Los parámetros de enfoque automático se determinan preferentemente por el algoritmo de ajustes de enfoque automático subyacente, balance de blancos automático y exposición automática y se pueden basar en el dispositivo de formación de imágenes 100. Por ejemplo, la función de enfoque automático podría calcular los parámetros de enfoque automático para cada uno de los rostros identificados como regiones de interés dentro de la escena y seleccionadas por el usuario como regiones de interés deseadas para una imagen de enfoque múltiple. El proceso 200 pasa a continuación al bloque 214, en el que el procesador recibe instrucciones para capturar una o más imágenes, conteniendo cada imagen una de las regiones de interés enfocadas, usando los parámetros de enfoque automático asociados con cada región de interés. Por ejemplo, el usuario puede capturar una imagen separada de cada uno de los tres rostros identificados por el usuario como regiones de interés.

**[0039]** El proceso 200 pasa a continuación al bloque 216, en el que el procesador recibe instrucciones para generar una imagen individual a partir de una o más imágenes capturadas de las múltiples regiones de interés. Un ejemplo de una imagen individual única de enfoque múltiple se muestra en la Figura 4. En esta ilustración, las regiones de interés seleccionadas como se muestra en la Figura 3 se visualizan enfocadas en la imagen individual mostrada en la Figura 4. Se puede generar una imagen individual combinando dos o más imágenes capturadas por el dispositivo de formación de imágenes 110 como se analiza anteriormente y "uniendo" las imágenes juntas usando cualquier algoritmo de combinación de imágenes conocido. Por ejemplo, se puede adquirir una imagen individual aplicando técnicas de alineación de imágenes, de registro de imágenes y de combinación de imágenes, como se conoce en la técnica, a las dos o más imágenes capturadas por el dispositivo de formación de imágenes 110 con diferentes configuraciones de enfoque. Después de que se genera la imagen individual, el proceso 200 pasa al bloque 218 y termina.

#### Aclaraciones sobre terminología

**[0040]** A menos que se indique de otro modo, cualquier divulgación de una operación de un aparato que tenga un rasgo característico particular también está expresamente prevista para divulgar un procedimiento que tenga un rasgo característico análogo (y viceversa), y cualquier divulgación de una operación de un aparato de acuerdo con una configuración particular también está prevista expresamente para divulgar un procedimiento de acuerdo con una configuración análoga (y viceversa). El término "configuración" se puede usar en referencia a un procedimiento, aparato y/o sistema según lo indicado por su contexto particular. Los términos "procedimiento", "proceso" y "técnica" se usan de forma genérica e intercambiable a menos que se indique de otro modo por el contexto particular. Los términos "aparato" y "dispositivo" también se usan de forma genérica e intercambiable a menos que se indique de otro modo por el contexto particular. Los términos "elemento" y "módulo" se usan típicamente para indicar una parte de una configuración mayor. A menos que esté limitado expresamente por su contexto, el término "sistema" se usa en el presente documento para indicar cualquiera de sus significados comunes, incluyendo "un grupo de elementos que interactúan para servir a un propósito común". Cualquier incorporación por referencia de una parte de un documento también debe entenderse como que incorpora definiciones de términos o variables a las que se hace referencia dentro de la parte, donde dichas definiciones aparecen en otra parte del documento, así como cualquier figura referenciada en la parte incorporada.

**[0041]** Los expertos en la técnica apreciarán además que los diversos bloques lógicos, módulos, circuitos y etapas de algoritmo ilustrativos descritos en relación con las implementaciones divulgadas en el presente documento se pueden implementar como hardware electrónico, software informático o combinaciones de ambos. Para ilustrar claramente esta intercambiabilidad de hardware y software, anteriormente se han descrito diversos componentes, bloques, módulos, circuitos y etapas ilustrativos, en general, en términos de su funcionalidad. Que dicha funcionalidad se implemente como hardware o software depende de la aplicación y de las restricciones de diseño particulares impuestas al sistema global. Los expertos en la técnica pueden implementar la funcionalidad descrita de formas distintas para cada solicitud particular, pero no debería interpretarse que dichas decisiones de implementación suponen apartarse del alcance de la presente invención. Un experto en la técnica reconocerá que una porción, o una

parte, puede comprender algo menos, o igual que, un todo. Por ejemplo, una parte de una colección de píxeles se puede referir a una subcolección de esos píxeles.

5 **[0042]** Los diversos bloques lógicos, módulos y circuitos ilustrativos descritos en relación con las implementaciones divulgadas en el presente documento se pueden implementar o realizar con un procesador de uso general, con un procesador de señales digitales (DSP), con un circuito integrado específico de la aplicación (ASIC), con una matriz de puertas programables por campo (FPGA) o con otro dispositivo de lógica programable, lógica de puertas discretas o de transistores, componentes de hardware discretos o con cualquier combinación de los mismos diseñada para realizar las funciones descritas en el presente documento. Un procesador de uso general puede ser un  
10 microprocesador, pero, como alternativa, el procesador puede ser cualquier procesador, controlador, microcontrolador o máquina de estados convencional. Un procesador también se puede implementar como una combinación de dispositivos informáticos, por ejemplo, una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores, uno o más microprocesadores junto con un núcleo de DSP o cualquier otra configuración de este tipo.

15 **[0043]** Los pasos de un procedimiento o algoritmo descrito en relación con las implementaciones divulgadas en el presente documento se pueden realizar directamente en hardware, en un módulo de software ejecutado por un procesador o en una combinación de los dos. Un módulo de software puede residir en una memoria RAM, una memoria flash, una memoria ROM, una memoria EPROM, una memoria EEPROM, registros, un disco duro, un disco extraíble, un CD-ROM o en cualquier otra forma de medio de almacenamiento no transitorio conocida en la técnica. Un medio ejemplar de almacenamiento legible por ordenador está acoplado al procesador de modo que el procesador pueda leer información de, y escribir información en, el medio de almacenamiento legible por ordenador. De forma  
20 alternativa, el medio de almacenamiento puede estar integrado en el procesador. El procesador y el medio de almacenamiento pueden residir en un ASIC. El ASIC puede residir en un terminal de usuario, en una cámara o en otro dispositivo. Como alternativa, el procesador y el medio de almacenamiento pueden residir como componentes discretos en un terminal de usuario, una cámara u otro dispositivo.

25 **[0044]** Los títulos se incluyen en el presente documento para referencia y para ayudar a la localización de diversas secciones. Estos títulos no están previstos para limitar el alcance de los conceptos descritos con respecto a los mismos. Dichos conceptos pueden tener aplicabilidad a lo largo de toda la memoria descriptiva.  
30

**REIVINDICACIONES**

- 5           1. Un procedimiento para sugerir múltiples regiones de interés dentro de una escena y generar una imagen de enfoque múltiple a partir de las múltiples regiones de interés, que comprende:
- recibir datos de imágenes digitales de una escena (204);
- determinar (206) regiones de interés candidatas dentro de la escena;
- 10          indicar (208) las regiones de interés candidatas en una pantalla de visualización que está visualizando la escena;
- recibir entrada de datos (210) indicando la selección de múltiples regiones de interés que pueden incluir una o más de las regiones de interés candidatas y otras regiones de interés;
- 15          determinar (212) parámetros de enfoque automático asociados con cada región de interés seleccionada, en el que determinar los parámetros de enfoque automático incluye realizar el enfoque automático para determinar los parámetros de enfoque automático para cada región de interés seleccionada donde la región de interés seleccionada está enfocada;
- 20          almacenar los parámetros de enfoque automático en un componente de memoria;
- después de recibir una entrada de datos para capturar imágenes, capturar una imagen de cada región de interés seleccionada usando los parámetros de enfoque automático almacenados asociados con la región de
- 25          interés (214) y
- generar una imagen de enfoque múltiple (216) que incluya datos de imagen de la imagen capturada de cada región de interés seleccionada.
- 30          2. El procedimiento de la reivindicación 1, que comprende además hacer funcionar un dispositivo de formación de imágenes en un modo de vista previa.
3. El procedimiento de la reivindicación 1, que comprende además determinar regiones de interés dentro de una escena usando un algoritmo de detección de rostros o un algoritmo de seguimiento de objetos.
- 35          4. El procedimiento de la reivindicación 2, que comprende además visualizar gráficamente las múltiples regiones de interés sugeridas para el usuario en el modo de vista previa.
5. El procedimiento de la reivindicación 1, en el que recibir entrada de datos (210) que indique la selección de múltiples regiones de interés comprende además registrar un evento táctil en una pantalla táctil.
- 40          6. El procedimiento de la reivindicación 2, en el que el dispositivo de formación de imágenes es un componente de una aplicación de cámara de un dispositivo móvil.
- 45          7. Un aparato (100) para sugerir múltiples regiones de interés dentro de una escena y generar una imagen de enfoque múltiple a partir de las múltiples regiones de interés, que comprende:
- medios (150) para determinar regiones de interés candidatas dentro de una escena;
- 50          medios (130) para indicar las regiones de interés candidatas en una pantalla de visualización que está visualizando la escena;
- medios para recibir entrada de datos indicando la selección de múltiples regiones de interés que pueden incluir una o más de las regiones de interés candidatas y otras regiones de interés;
- 55          medios (155) para determinar los parámetros de enfoque automático asociados con cada región de interés seleccionada, en el que determinar los parámetros de enfoque automático incluye realizar el enfoque automático para determinar los parámetros de enfoque automático para cada región de interés seleccionada donde la región de interés seleccionada está enfocada;
- 60          medios (125) para almacenar los parámetros de enfoque automático en un componente de memoria;
- medios (145) para capturar una imagen de cada región de interés seleccionada usando los parámetros de enfoque automático almacenados asociados con la región de interés; y
- 65

medios (160) para generar una imagen de enfoque múltiple que incluya datos de imagen de la imagen capturada de cada región de interés seleccionada.

- 5 **8.** El aparato de la reivindicación 7, que comprende además un dispositivo de formación de imágenes (110) que comprende un sensor, en el que dichos medios para capturar una imagen de cada región de interés usan el dispositivo de formación de imágenes (110) y los parámetros de enfoque automático asociados con la región de interés.
- 10 **9.** El aparato de la reivindicación 8, que comprende además medios para hacer funcionar el dispositivo de formación de imágenes en un modo de vista previa.
- 15 **10.** El aparato de la reivindicación 7, que comprende además medios para determinar las regiones de interés dentro de una escena usando un algoritmo de detección de rostros o un algoritmo de seguimiento de objetos.
- 20 **11.** El aparato de la reivindicación 9, que comprende además medios para visualizar gráficamente las múltiples regiones de interés sugeridas para el usuario en el modo de vista previa.
- 25 **12.** El aparato de la reivindicación 7, que comprende además una pantalla táctil, y en el que el aparato comprende además medios para recibir entrada de datos que indique una selección de múltiples regiones de interés registrando un evento táctil en la pantalla táctil.
- 13.** El aparato de la reivindicación 7, que comprende además medios para recibir entrada de datos con respecto a la determinación de regiones de interés candidatas dentro de una escena.
- 14.** Un medio no transitorio legible por ordenador que comprende instrucciones que, cuando se ejecutan, causan que al menos el procesador físico informático cause que un aparato de acuerdo con la reivindicación 7 realice el procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 6.

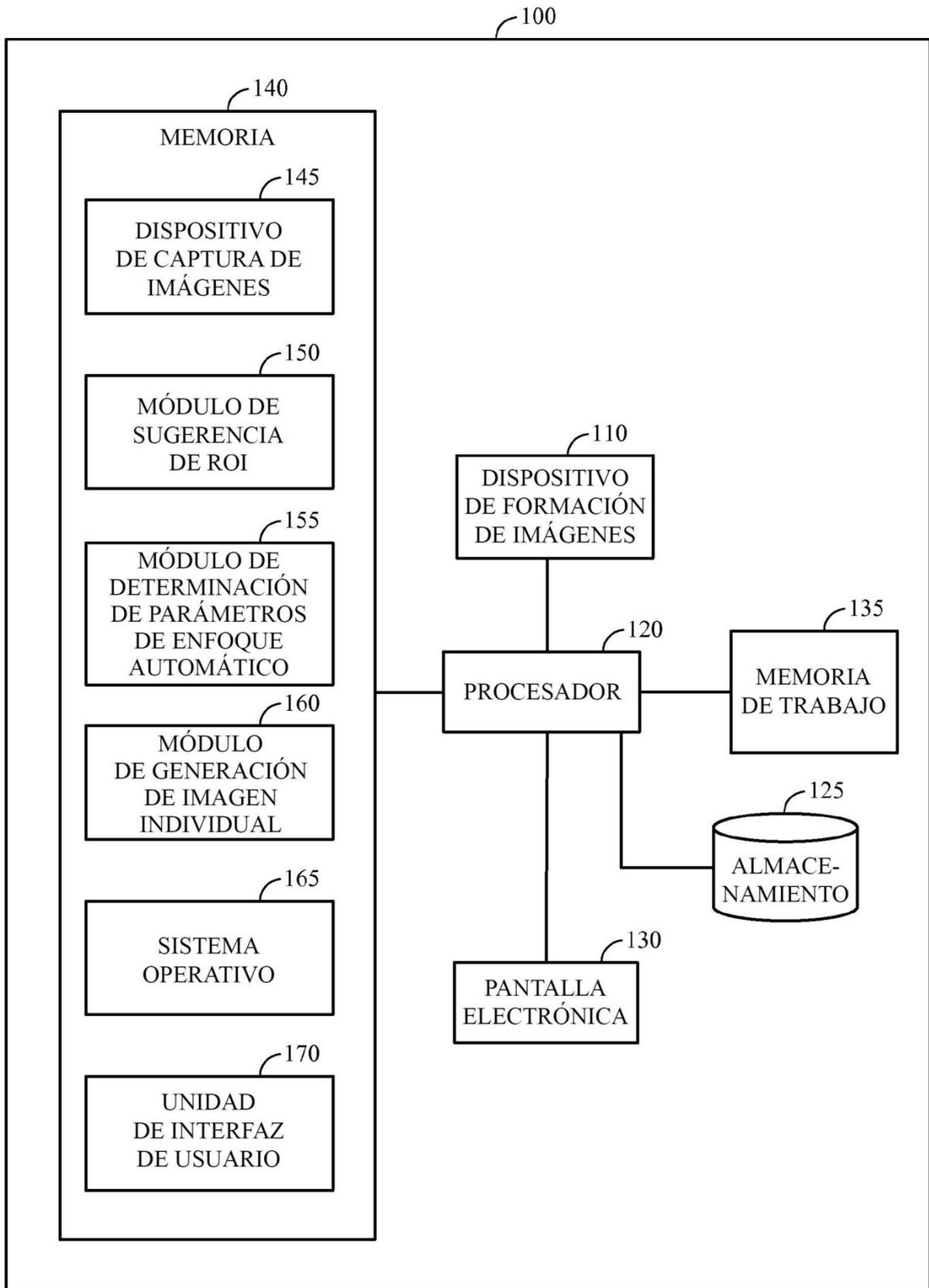


FIG. 1

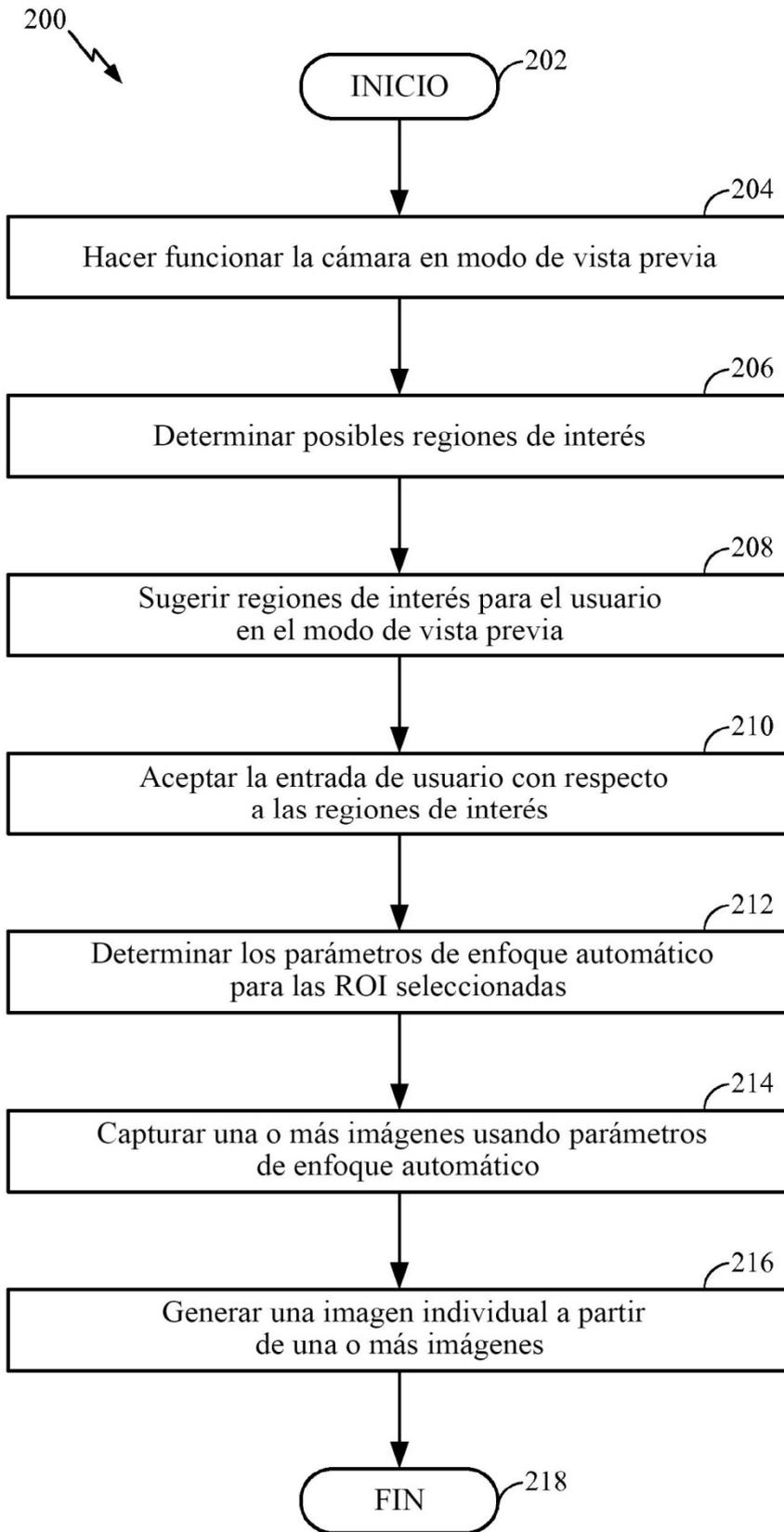


FIG. 2

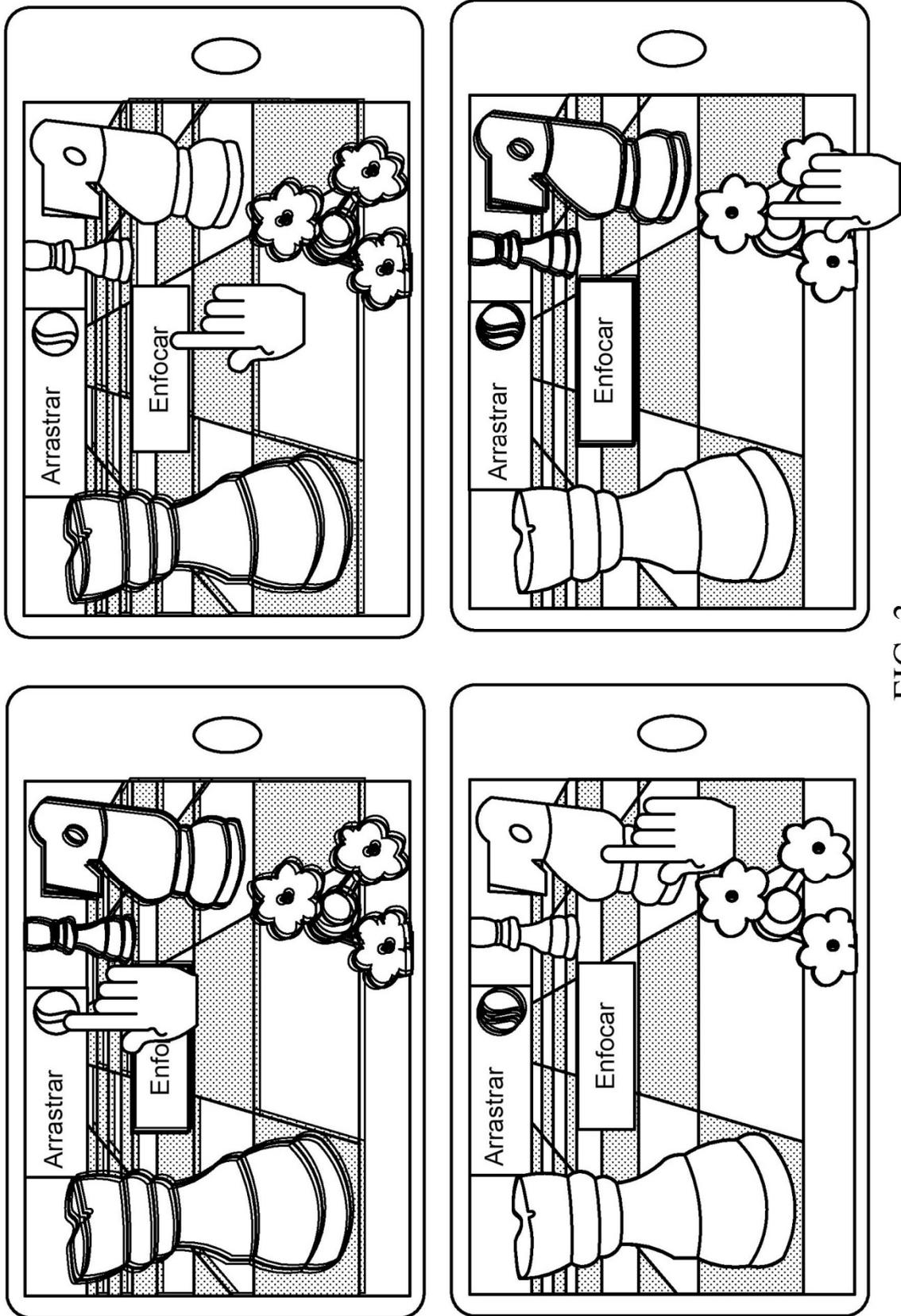


FIG. 3

