

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 600 756**

51 Int. Cl.:

G06T 5/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.03.2003** **E 03005431 (6)**

97 Fecha y número de publicación de la concesión europea: **03.08.2016** **EP 1388816**

54 Título: **Sistema y procedimiento para facilitar la compresión de imágenes de documento utilizando una máscara**

30 Prioridad:

25.04.2002 US 133558

25.04.2002 US 133842

25.04.2002 US 133939

26.06.2002 US 180771

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

10.02.2017

73 Titular/es:

MICROSOFT TECHNOLOGY LICENSING, LLC

(100.0%)

One Microsoft Way

Redmond, WA 98052, US

72 Inventor/es:

SIMARD, PATRICE Y.;

RENSHAW, ERIN L. y

RINKER, JAMES RUSSELL

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 600 756 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 facilitar la compresión de imágenes de documento utilizando una máscara

Referencia a solicitudes relacionadas

5 La presente solicitud es una continuación parcial de la Solicitud de Utilidad Estados Unidos N.º de Serie 10/133.842 que se presentó el 25 de abril de 2002, titulada ACTIVITY DETECTOR, de la Solicitud de Utilidad N.º de Serie 10/133.558 que se presentó el 25 de abril de 2002, titulada CLUSTERING y de la Solicitud de Utilidad Estados Unidos N.º de Serie 10/133.939 que se presentó el 25 de abril de 2002, titulada LAYOUT ANALYSIS.

Campo técnico

10 La presente invención se refiere en general al procesamiento de imágenes de documentos y, más particularmente, a un sistema y procedimiento que facilita la compresión de imágenes de documentos utilizando una máscara que divide un primer plano de una imagen de un documento a partir de un plano posterior.

Antecedentes de la invención

15 La cantidad de información disponible a través de los ordenadores se ha incrementado dramáticamente con la amplia proliferación de redes de ordenadores, Internet y los medios de almacenamiento digital. Con la cantidad creciente de información ha surgido la necesidad de transmitir información rápidamente y de almacenar eficientemente la información. La compresión de datos es una manera por la que los documentos pueden transmitirse y/o almacenarse más efectivamente.

20 Los sistemas de compresión de datos convencionales han utilizado varios enfoques de compresión, por ejemplo, coincidencia de símbolos. Sin embargo, los enfoques de compresión típicos que funcionan de modo efectivo para documentos que tienen imágenes no funcionan bien, por ejemplo, para documentos que tienen texto y/o escritura manual.

25 La compresión de datos reduce el espacio necesario para representar la información. La compresión puede usarse para cualquier tipo de información. Sin embargo, la compresión de la información digital, incluyendo imágenes, texto, audio y vídeo se ha convertido en más importante. Típicamente, la compresión de datos se usa con sistemas de ordenador estándar. Sin embargo, otras tecnologías hacen uso de la compresión de datos, tal como, pero sin limitarse a, televisión digital y por satélite así como teléfonos celulares/digitales.

30 La compresión de datos es importante por varias razones. La compresión de datos permite que la información se almacene en menos espacio que los datos sin comprimir. Cuando se incrementa la demanda de grandes cantidades de información, puede requerirse la compresión de los datos para suministrar las grandes cantidades de información. El tamaño de los dispositivos de almacenamiento se ha incrementado significativamente, sin embargo la demanda de información ha sobrepasado estos incrementos de tamaño. Por ejemplo, una imagen sin comprimir puede necesitar hasta 5 megabytes de espacio mientras que la misma imagen puede comprimirse y necesitar solo 2,5 megabytes de espacio. Además, la compresión de datos permite la transferencia de cantidades de información comprimida más grandes que la información sin comprimir. Incluso con el incremento de las velocidades de transmisión, tales como la banda ancha, DSL, cable módem e Internet y similares, los límites de la transmisión pueden alcanzarse fácilmente con la información sin comprimir. Por ejemplo, la transmisión de una imagen sin comprimir a través de una línea DSL puede llevar diez minutos. Sin embargo, con la compresión de datos, la misma imagen puede transmitirse en aproximadamente un minuto.

40 En general, hay dos tipos de compresión, sin pérdida y con pérdida. La compresión sin pérdida permite que se recuperen los datos originales exactos tras la compresión, mientras que la compresión con pérdida permite que los datos originales difieran de los datos sin comprimir. La compresión con pérdida permite una mejor relación de compresión debido a que elimina datos del original. La compresión sin pérdida puede usarse, por ejemplo, cuando se comprime texto crítico, debido que el fallo en reconstruir exactamente los datos puede afectar seriamente a la calidad y legibilidad del texto. La compresión con pérdida puede usarse con imágenes o texto no crítico en el que una cierta cantidad de distorsión o ruido es o bien aceptable o bien imperceptible por nuestros limitados sentidos.

45 La compresión de datos es especialmente aplicable a documentos digitales. Los documentos digitales o imágenes de documentos digitales son representaciones digitales de documentos. Típicamente, los documentos digitales incluyen texto, imágenes y/o texto e imágenes. Además de usar menos espacio de almacenamiento para datos digitales actuales, el almacenamiento compacto sin degradación significativa de la calidad promoverá la digitalización de las copias en papel actuales haciendo más factibles las oficinas sin papel. El esfuerzo hacia dichas oficinas sin papel es un objetivo importante que han de tener las empresas, debido a que las oficinas sin papel proporcionan muchos beneficios, tales como permitir un fácil acceso a la información, reducir los costes medioambientales, reducir los costes de almacenamiento y otros similares. Además, la disminución de tamaño de los archivos de los documentos digitales a través de la compresión permite un uso más eficiente del ancho de banda de Internet, permitiendo así una transmisión más rápida de más información y la reducción de la congestión de la red. La reducción del almacenamiento requerido para la información, el movimiento hacia oficinas sin papel más

eficientes, y el incremento de la eficiencia del ancho de banda de Internet, son solo algunos de los muchos beneficios significativos de la tecnología de compresión.

5 La compresión de datos de los documentos digitales tiene un cierto número de metas para hacer más atractivo el uso de los documentos digitales. Primero, la compresión de datos debería ser capaz de comprimir y descomprimir grandes cantidades de información en una pequeña cantidad de tiempo. En segundo lugar, la compresión de datos debería ser capaz de reproducir con precisión el documento digital.

10 Además, la compresión de datos de documentos digitales debería hacer uso de la finalidad de un documento. Algunos documentos digitales se usan para rellenar o proporcionar copias impresas. Otros documentos pueden revisarse y/o editarse. La compresión de datos actual fracasa al manejar la recolocación de texto y/o imágenes cuando se ve, y fracasa al proporcionar medios eficientes y efectivos para permitir que la tecnología de compresión reconozca caracteres y los recoleque en procesadores de textos, asistentes digitales personales (PDA), teléfonos celulares y similares. Por lo tanto, si se escanean en forma digital documentos de oficina en papel, la tecnología de compresión actual puede hacer difícil, si no imposible, actualizar, modificar, o cambiar en general el documento digitalizado.

15 La compresión de imágenes compuestas con un enfoque multicapa de contenido de barrido fusionado se explica en "Optimizing Block-Thresholding Segmentation for Multilayer Compression of Compound Images" por R.L. de Queiroz et ál., IEEE Transactions on Image Compression, IEEE Inc. Nueva York, US, vol. 9, n.º 9, septiembre de 2000, páginas 1461-1471, ISSN: 1057-7149.

20 Las representaciones basadas en zonas de imagen y video se presentan en "Region-Based Representations of Image and Video: Segmentation Tools for Multimedia Services" por P.Salembier et ál., IEEE Transactions on Circuits and Systems for Video Technology, IEEE Inc. Nueva York, Estados Unidos, vol. 9, n.º 8, diciembre de 1999, páginas 1147-1169, ISSN: 1051-8215.

Sumario de la invención

25 La invención comprende un componente separador de máscara de acuerdo con la reivindicación 1, un procedimiento para la generación de una máscara que divide una imagen de un documento en un plano posterior y un primer plano de acuerdo con la reivindicación 12, un sistema de compresión de imágenes de documento de acuerdo con la reivindicación 19, un medio legible por ordenador de acuerdo con la reivindicación 23, una fotocopiadora de acuerdo con la reivindicación 25, un escáner de documentos de acuerdo con la reivindicación 26, un sistema de reconocimiento óptico de caracteres de acuerdo con la reivindicación 27, un asistente digital personal de acuerdo con la reivindicación 28, una máquina de fax de acuerdo con la reivindicación 29, una cámara digital de acuerdo con la reivindicación 30, una cámara de video digital de acuerdo con la reivindicación 31 y un sistema de videojuegos de acuerdo con la reivindicación 32. En algunas realizaciones favorables se describen en las reivindicaciones dependientes.

35 A continuación se presenta un sumario simplificado de la invención para proporcionar una comprensión básica de algunos aspectos de la invención. Este sumario no es una visión extensa de la invención. No se pretende que identifique elementos clave/críticos de la invención o delimite el alcance de la invención. Su única finalidad es presentar algunos conceptos de la invención de una forma simplificada como un prelude a una descripción más detallada que se presenta más adelante.

40 La presente invención se refiere en general a un sistema y procedimiento que facilita la compresión de imágenes de documentos utilizando una partición con máscara de un primer plano de una imagen de un documento respecto a un plano posterior. De acuerdo con un aspecto de la presente invención, un componente separador de máscara recibe una imagen del documento (por ejemplo, una representación binaria, RGB y/o YUV del (de los) documento(s)) como una entrada. El componente separador de máscara procesa la imagen del documento y produce la salida de una máscara (por ejemplo binaria) que indica si cada píxel de la imagen del documento pertenece al primer plano y/o a un plano posterior. Mediante la separación del primer plano (por ejemplo, la información textual) del plano posterior (por ejemplo información gráfica), el primer plano y/o el plano posterior pueden comprimirse más efectivamente, disminuyendo así el tamaño del archivo y/o el tiempo de transmisión. La máscara y/o la imagen del documento pueden procesarse entonces por otra(s) parte(s) de un sistema de compresión (por ejemplo, para conseguir una compresión mejorada de la imagen del documento). Por ejemplo, el sistema y/o procedimiento de la presente invención puede utilizarse en un sistema de imagen en capas segmentado global que facilite la identificación y/o compresión de texto, escritura manual, dibujos y similar.

55 De acuerdo con un aspecto particular de la invención, el componente separador de máscara incluye un componente de energía de píxel, un componente de fusión de regiones y un componente de almacenamiento de máscara. El componente de energía de píxel está adaptado para calcular la energía de píxel (por ejemplo, varianzas) para una(s) región(es) de una imagen del documento para minimizar las variaciones de energía del primer plano y/o plano posterior. La energía (por ejemplo, se mide la energía basándose en una suma del cuadrado de las distancias) se usa como una estimación de la compresión que se obtendría para el primer plano y/o el plano posterior. Sin embargo, para simplificar la sobrecarga de cálculo, la imagen del documento puede dividirse en regiones (por

ejemplo, de dos píxeles por dos píxeles) y determinarse un primer plano y plano posterior para cada región (por ejemplo, basándose al menos en parte, en la minimización de las variaciones de energía en el plano posterior y/o en el primer plano). En otras palabras, cada región se divide en sí misma en dos conjuntos: los píxeles que pertenecen al primer plano, y los píxeles que pertenecen al plano posterior. Para minimizar además la sobrecarga de cálculo, el componente de energía de píxel puede, al menos temporalmente, almacenar información de cálculo para su uso por el componente de fusión de regiones y/o el componente de almacenamiento de máscara.

El componente de fusión de regiones está adaptado para intentar fusionar pares de regiones de la imagen del documento basándose, al menos en parte, en la determinación de si las energías de un nuevo primer plano y/o un nuevo plano posterior de las regiones fusionadas potenciales son menores que una primera energía de umbral. El componente de fusión de regiones puede utilizar información de cálculo almacenada por el componente de energía de píxel. El resultado de una fusión es una región más grande que se caracterizará por su propia partición de primer plano y plano posterior. Los píxeles que fueron primer plano previamente a la fusión pueden acabar en el plano posterior de la región fusionada y viceversa. El componente de fusión de regiones puede determinar una partición primer plano/plano posterior adecuada de la región fusionada, por ejemplo, basándose al menos en parte en la minimización de las energías de un nuevo plano posterior y un nuevo primer plano.

El componente de fusión de regiones puede continuar intentando fusionar sucesivamente regiones mayores hasta que la energía de umbral se supere y/o sustancialmente toda la imagen del documento se haya fusionado. Por ejemplo, el componente de fusión de regiones puede fusionar regiones de dos por dos horizontalmente adjuntas en una región de dos por cuatro. Posteriormente, el componente de fusión de regiones puede fusionar verticalmente regiones en una región de cuatro por cuatro. Generalmente, el primer valor de energía de umbral puede seleccionarse para mitigar las situación(es) potencial(es) en las que el intento de fusión(s) dividiría varios niveles de gris en el primer plano o en el plano posterior, con una pérdida potencial de detalles importantes, tal como texto (por ejemplo cuando hay más de dos colores en una región). Por ello, puede capturarse una máscara que capture la mayor parte del texto y/o línea(s) gráfica(s) asociada(s) con una imagen de un documento.

Una vez se ha completado la fusión para una región, la partición de primer plano y plano posterior para esta región constituye la máscara, por ejemplo, el (los) píxel(es) que pertenecen al primer plano pueden tener asignado un "1" en la máscara, mientras que el (los) píxel(es) que pertenecen al plano posterior pueden tener asignado un "0". Desafortunadamente, mantener un seguimiento de las particiones de primeros planos y planos posteriores durante la operación de fusión puede ser caro desde el punto de vista computacional. Una alternativa (por ejemplo, más efectiva de modo computacional) es calcular una media de sustancialmente todos los píxeles de la región fusionada y asignar píxel(es) que tengan un nivel de gris mayor que la media al primer plano siendo asignados el (los) píxel(es) restante(s) al plano posterior. Alternativamente, el (los) píxel(es) que tengan un valor de nivel de gris mayor que la media puede estar asignados al plano posterior siendo asignados el (los) píxel(es) restante(s) al primer plano. Las dos alternativas pueden producir máscaras visualmente indiferenciables.

Posteriormente, el componente de almacenamiento de máscara está adaptado para almacenar la información asociada con la partición del primer plano y el plano posterior en la máscara. Por ello, la máscara indica si cada píxel de la imagen del documento pertenece al primer plano y/o al plano posterior.

De acuerdo con otro aspecto de la presente invención, para minimizar la sobrecarga computacional, puede calcularse la energía para una pequeña región (por ejemplo cuatro píxeles por cuatro píxeles) por el componente de energía de píxel. Si la energía es menor que una segunda energía de umbral, sustancialmente todos los píxeles pueden estar asignados al primer plano o al plano posterior estando los otros sustancialmente vacíos. Si la energía es mayor que o igual a la segunda energía de umbral, la partición puede proseguir como se ha descrito anteriormente. Para imágenes de documentos relativamente limpios (por ejemplo, que tengan área(s) constantes), puede conseguirse un incremento significativo en la velocidad computacional.

De acuerdo con otro aspecto de la presente invención, para minimizar el tamaño de la máscara, si una región final (por ejemplo una región que no puede fusionarse sin exceder el primer umbral), tiene una diferencia entre el primer plano medio y el plano posterior medio que es más alta que un tercer umbral, toda la región se declara primer plano o se declara plano posterior, dependiendo de si una media global para la región es mayor o menor que el valor medio del nivel de gris (por ejemplo, 127 si los valores de nivel de gris están entre 0 y 255). Para documentos en color que tengan un ligero tramado, la máscara para esta región parecería similar a una entremezcla sin esta optimización y tendría un alto coste de compresión. El tercer umbral se elige de modo que no se pierdan texto importante, mientras se eliminan los muchos casos de ligero tramado vistos en documentos impresos escaneados (por ejemplo, muchas impresoras tienen solo 4 a 6 colores y deben usar tramado para generar toda la paleta de colores). En un ejemplo, el valor de 40 es una buena elección para el tercer umbral.

En otro aspecto más de la presente invención proporciona el componente de energía de píxel para utilizar una regresión polinómica para describir el primer plano y/o el plano posterior.

Otro aspecto de la compresión de imágenes de documento proporciona un sistema de separación de la imagen de un documento que tiene un componente separador de máscara y un segmentador primer plano/plano posterior. El componente separador de máscara puede procesar una imagen de un documento (por ejemplo, que comprenda

texto y/o escritura manual) y almacenar información en relación a qué píxeles están en el primer plano y cuáles están en el plano posterior en una máscara. Posteriormente, el segmentador primer plano/plano posterior puede recibir la máscara y la imagen del documento y separar la imagen del documento en una imagen de primer plano y una imagen de plano posterior.

5 De acuerdo con otro aspecto más de la presente invención, se proporciona un sistema de compresión de imágenes de documento que tiene un componente de transformación de la imagen de un documento, un componente de separación de máscara y un componente de separación primer plano/plano posterior. Opcionalmente, el sistema de compresión de imágenes de documento puede incluir un codificador de máscara, un codificador del primer plano y/o un codificador del plano posterior.

10 Otro aspecto más de la compresión de imágenes de documento proporciona un sistema de imágenes segmentadas en capas que tiene un componente de energía de píxel y un componente de almacenamiento de máscara. El sistema de imágenes segmentado en capas puede emplearse en una amplia variedad de aplicaciones de imágenes de documentos, incluyendo, pero sin limitarse a, fotocopadoras, escáneres de documentos, sistemas de reconocimiento óptico de caracteres, asistentes digitales personales, máquinas de fax, cámaras digitales, cámaras de vídeo digitales y/o sistemas de videojuego.

15 Otros aspectos de la compresión de imágenes de documento proporciona procedimientos, metodologías para, generando una máscara, un medio legible por ordenador que tenga funciones utilizables por ordenador para un componente de separación de máscaras y un paquete de datos adaptado para transmitirse entre dos o más procesos de ordenador que comprenden información asociada con una máscara, asignando la máscara píxeles a al menos uno de un primer plano y un plano posterior de una imagen del documento, estando basada la máscara, al menos en parte, en el cálculo de la minimización de la energía de los píxeles en una región de la imagen del documento.

20 Para el cumplimiento de los fines precedente y otros relacionados, ciertos aspectos ilustrativos de la invención se describen en la presente memoria en conexión con la descripción siguiente y los dibujos adjuntos. Estos aspectos son indicativos, sin embargo, de unas pocas de las diversas formas en las que pueden emplearse los principios de la invención y se pretende que la presente invención incluya todos los dichos aspectos y sus equivalentes. Otras ventajas y características novedosas de la invención pueden hacerse evidentes a partir de la descripción detallada a continuación de la invención cuando se considera en conjunto con los dibujos.

Breve descripción de los dibujos

30 La Fig. 1 es un diagrama de bloques de un componente separador de máscara de acuerdo con un aspecto de la presente invención.

La Fig. 2 es una imagen de un documento ejemplar de acuerdo con un aspecto de la presente invención.

La Fig. 3 es una máscara asociada con la imagen del documento ejemplar de la Fig. 2 de acuerdo con un aspecto de la presente invención.

35 La Fig. 4 es un plano posterior asociado con la imagen del documento ejemplar de la Fig. 2 y la máscara de la Fig. 3 de acuerdo con un aspecto de la presente invención.

La Fig. 5 es una región de dos píxeles por dos píxeles ejemplar de una imagen de un documento de acuerdo con un aspecto de la presente invención.

40 La Fig. 6 es una región fusionada potencial de cuatro píxeles por cuatro píxeles ejemplar de acuerdo con un aspecto de la presente invención.

La Fig. 7 es una región fusionada potencial de cuatro píxeles por cuatro píxeles ejemplar de acuerdo con un aspecto de la presente invención.

La Fig. 8 es un diagrama de flujo que ilustra una metodología para la generación de una máscara de acuerdo con un aspecto de la presente invención.

45 La Fig. 9 es un diagrama de flujo que ilustra además la metodología de la Fig. 8.

La Fig. 10 es un diagrama de bloques de un sistema de separación de la imagen de un documento de acuerdo con un aspecto de la presente invención.

La Fig. 11 es un diagrama de bloques de una compresión de imágenes de documento de acuerdo con un aspecto de la presente invención.

50 La Fig. 12 es un diagrama de bloques de una compresión de imágenes de documento de acuerdo con un aspecto de la presente invención.

La Fig. 13 es un diagrama de bloques de un sistema de imagen segmentado en capas de acuerdo con un aspecto de la presente invención.

La Fig. 14 es un diagrama de bloques esquemático de un entorno operativo ejemplar para un sistema configurado de acuerdo con la presente invención.

55 La Fig. 15 es un diagrama de bloques esquemático de un entorno de comunicación ejemplar de acuerdo con la presente invención.

Descripción detallada de la invención

La presente invención se describe ahora con referencia a los dibujos, en los que se usan números de referencia

similares para referirse a elementos similares a todo lo largo. En la descripción siguiente, con finalidades de explicación, se exponen numerosos detalles específicos para proporcionar una comprensión global de la presente invención. Puede ser evidente, sin embargo, que la presente invención puede ponerse en práctica sin estos detalles específicos. En otros casos, estructuras y dispositivos bien conocidos se muestran en forma de diagramas de bloques para facilitar la descripción de la presente invención.

Tal como se usan en la presente solicitud, los términos “componente” y “sistema” se pretende que se refieran a una entidad basada en ordenador, sea en hardware, una combinación de hardware y software, software, o en software en ejecución. Por ejemplo, un componente puede ser, pero sin limitarse a serlo, un proceso ejecutado en un procesador, un procesador, un objeto, un ejecutable, un hilo de ejecución, un programa y/o un ordenador. A modo de ilustración, tanto una aplicación que se ejecuta en un servidor como el servidor puede ser un componente. Uno o más componentes pueden residir dentro de un proceso y/o hilo de ejecución y un componente puede localizarse en un ordenador y/o distribuirse entre dos o más ordenadores.

Además, la “imagen de un documento” se pretende que se refiera a una representación digital de documentos que comprenden uno o más colores (por ejemplo, binarios, por ejemplo blanco/negro), escala de grises y/o documentos en color). Además, una imagen de un documento puede tener imágenes, texto y/o texto con imágenes, con superposición potencial del texto e imágenes. Una imagen de un documento pueden ser representaciones binarias, RGB y/o YUV de documentos. Una imagen de un documento RGB se representa en componentes rojo, verde y azul. Una imagen de un documento YUV se representa usando un componente de luminancia indicado por Y y componentes de crominancia indicados por U y V. Pueden usarse menos bits para representar los componentes de crominancia U y V sin sacrificar significativamente la calidad visual de la imagen YUV. La representación YUV es, en general una representación más compacta y fácil de utilizar que una representación RGB. Una imagen de un documento comprende elementos de foto comúnmente denominados como “píxeles”. Una imagen de un documento puede basarse en un documento de una o múltiples páginas de cualquier forma o tamaño.

En referencia a la Fig. 1, se ilustra un componente 100 separador de máscara con un aspecto de la presente invención. El componente 100 separador de máscara recibe una imagen 110 de un documento (por ejemplo, basada en un documento a ser archivado y/o transmitido). Por ejemplo, el componente 100 separador de máscara puede ser parte de un sistema de compresión de documentos (no mostrado). La imagen 110 de un documento puede ser una representación binaria, RGB y/o YUV del (de los) documento(s). El componente 100 separador de máscara procesa la imagen 110 del documento y produce la salida de la máscara 120 (por ejemplo binaria) que indica si cada píxel de la imagen 110 del documento pertenece al primer plano y/o al plano posterior. La máscara 120 y/o la imagen 110 del documento pueden procesarse entonces por otra(s) parte(s) del sistema de compresión (no mostrado) para efectuar la compresión de la imagen 110 del documento.

Pasando brevemente a la Fig. 2, se ilustra una imagen de un documento ejemplar. La imagen del documento comprende las letras “C” y “O” junto con una barra. La Fig. 3 ilustra una máscara asociada con la imagen del documento ejemplar de la Fig. 2 de acuerdo con un aspecto de la presente invención. La Fig. 4 ilustra un plano posterior asociado con la imagen del documento ejemplar de la Fig. 2 y la máscara de la Fig. 3 de acuerdo con un aspecto de la presente invención. Las líneas discontinuas representan los límites de los píxeles de “importancia” del plano posterior; los píxeles que comprenden las líneas de puntos y los píxeles dentro de las líneas discontinuas son “no importantes” en el plano posterior dado que cuando la imagen del documento se vuelve a montar el primer plano se colocará sobre el plano posterior basándose, al menos en parte, en la información de reconstrucción almacenada en la máscara. En el caso en el que las letras “C” y “O” y/o la barra son de color(es) constante(s) y/o tienen transición(es) de color suaves, la compresión efectiva del primer plano puede conseguirse usando una de una variedad de técnicas de suavizado y/o compresión. La compresión efectiva del plano posterior puede conseguirse mediante la sustitución de los píxeles “no importantes” con valores de píxeles que permiten transiciones más suaves. Un algoritmo simple ejemplar para rellenado de los píxeles “no importantes” es procesar el plano posterior con un filtro paso bajo, y a continuación restaurar los píxeles importantes. Después de pocas iteraciones de estas dos etapas, los píxeles “no importantes” acaban con valores que permiten transiciones suaves, y que se comprimirán bien. Un refinamiento simple del este algoritmo es iniciar con un filtro paso muy bajo e incrementar la frecuencia de corte del filtro paso bajo en cada iteración. Puede usarse un algoritmo similar para llenar el primer plano.

Volviendo a la Fig. 1, el componente 100 separador de máscara incluye un componente 130 de energía de píxel, un componente 140 de fusión de regiones y un componente 150 de almacenamiento de máscara.

El componente 130 de energía de píxel está adaptado para calcular la energía de píxel para regiones de la imagen 110 del documento (por ejemplo, varianzas). Por ejemplo, en el caso en el que la imagen 110 del documento es una representación YUV, el componente 130 de energía de píxel calcula las varianzas de energía de píxel basándose en el componente Y y/o la combinación adecuada de los componentes YUV de la representación YUV. Para finalidades de cálculo, el primer plano y el plano posterior pueden suponerse constantes a través de una región. Si se desea calcular una máscara 120 que minimice la varianza alrededor de estas constantes. La varianza se usa como una estimación de la compresión que se obtendría para el primer plano y/o el plano posterior. Alternativamente, la(s) región(es) podría(n) comprimirse y el número de bits podría medirse cuantitativamente; sin embargo, la sobrecarga computacional sería prohibitivamente cara. En consecuencia, el cálculo de la varianza, que es también una medición

de la energía basada en una suma de las distancias al cuadrado, es una estimación aceptable del tamaño del primer plano y el plano posterior tras la compresión.

Suponiendo que una región es un conjunto S de N píxeles, y que un primer plano P y un plano posterior F son una partición de S de modo que $P \cup F = S$ y $P \cap F = \emptyset$. Si $f(x)$ es el valor de la imagen en la localización del píxel x , $x \in S$, la varianza del primer plano y el plano posterior son respectivamente:

$$v_P = \sum_{x \in P} (f(x) - \mu_P)^2 \quad (1)$$

$$v_F = \sum_{x \in F} (f(x) - \mu_F)^2 \quad (2)$$

en las que $\mu_P = \frac{1}{N_P} \sum_{x \in P} f(x)$ y $\mu_F = \frac{1}{N_F} \sum_{x \in F} f(x)$ son, respectivamente, el valor de píxel medio del primer plano y del plano posterior, y N_P y N_F son, respectivamente, el número de píxeles en el primer plano y el plano posterior. Obsérvese que estas varianzas pueden expresarse también como:

$$v_P = \sum_{x \in P} f(x)^2 - N_P \mu_P^2 \quad (3)$$

$$v_F = \sum_{x \in F} f(x)^2 - N_F \mu_F^2 \quad (4)$$

A continuación, se determina una partición adecuada en P y F de S , basándose al menos en parte en la minimización de la energía del primer plano y/o el plano posterior (por ejemplo, varianzas) por el componente 130 de energía de píxel (por ejemplo, que minimice la suma $E = v_P + v_F$). Sin embargo, hallar una partición adecuada en P y F de S puede ser computacionalmente intensivo dado que hay 2^N máscaras posibles.

Para simplificar el cálculo, la imagen del documento puede dividirse en regiones, por ejemplo regiones de dos píxeles por dos píxeles. Pasando brevemente a la Fig. 5, se ilustra una región de dos píxeles por dos píxeles de una imagen de un documento de acuerdo con un aspecto de la presente invención. Los cuatro píxeles tienen valores V_1, V_2, V_3 y V_4 . Para cada región de dos píxeles por dos píxeles, hay solo $2^4 = 16$ posibles máscaras. De acuerdo con cada una de estas regiones de dos píxeles por dos píxeles, es posible hallar los P y F óptimos, que minimicen $E = v_P + v_F$ mediante el cálculo de E para todas las 16 combinaciones y utilizando aquella con la energía más pequeña.

Sin embargo, utilizando un algoritmo de agrupamiento de K medias, en el que $K=2$, dado que la imagen del documento es una función escalar, pueden clasificarse los valores $f(x)$ que conducen a una solución que pueda calcularse eficientemente. Suponiendo que el orden clasificado es $V_1 V_2 V_3 V_4$, el algoritmo de agrupamiento de K medias, en el que $K=2$ conduce a tres posibles particiones:

Primer plano potencial	Plano posterior potencial
V_1	$V_2 V_3 V_4$
$V_1 V_2$	$V_3 V_4$
$V_1 V_2 V_3$	V_4

Puede verse que sustancialmente todas las otras combinaciones tendrán energía igual o más alta. Esto es intuitivo dado que siempre debería haber una agrupación de valores contiguos que tenga una varianza más baja que una agrupación de valores no contiguos. Si el orden de clasificación fuera diferente, el píxel puede siempre re-etiquetarse de modo que se clasifiquen $V_1 V_2 V_3 V_4$. Es entonces directo determinar cuál de las tres posibles particiones de primer plano y plano posterior conducen a la energía más baja. Significativamente, el componente 130 de energía de píxel puede almacenar la suma parcial $\sum_P f(x)$, $\sum_P f(x)^2$, $\sum_F f(x)$ y/o $\sum_F f(x)^2$ para minimizar la sobrecarga computacional. Además, el componente 130 de energía de píxel puede, al menos temporalmente, almacenar al menos parte de las sumas parciales $\sum_P f(x)$, $\sum_P f(x)^2$, $\sum_F f(x)$ y/o $\sum_F f(x)^2$ junto con N_P y/o N_F (por ejemplo, para su uso en el componente 140 de fusión de regiones y/o el componente 150 de almacenamiento de máscara).

Además, para minimizar la sobrecarga computacional, puede calcularse la energía para una región pequeña (por ejemplo, cuatro píxeles por cuatro píxeles). Si la energía es menor que una cantidad de umbral, todos los píxeles pueden asignarse al primer plano o al plano posterior siendo el otro vacío. Si la energía es mayor a o igual que la energía de umbral, la partición puede proseguir como se ha descrito anteriormente. Aunque esta partición pudiera no ser óptima, no se observan efectos adversos si el umbral es suficientemente pequeño. Para imágenes de documento(s) relativamente limpias (por ejemplo, que tienen área(s) constante(s)), puede conseguirse un incremento significativo en la velocidad computacional.

Además, una(s) región(es) que es (son) sustancialmente constante(s) (por ejemplo, primer plano puro o plano posterior puro) también pueden fijarse después de que se haya calculado la separación de la máscara del primer plano y el plano posterior. Por ejemplo, si la diferencia entre el primer plano medio y el plano posterior medio es menor que un cierto umbral, que puede determinarse experimentalmente (por ejemplo puede usarse un valor de 40 en comparación con la escala completa de niveles de gris que van de 0 a 255), la región entera se fija o bien al primer plano o bien al plano posterior (dependiendo de si la media está más próxima a 0 o a 255).

Volviendo a referirnos a la Fig. 1, la partición de la imagen 110 del documento en regiones de dos píxeles por dos píxeles puede dar como resultado regiones que tengan distintos primeros planos y planos posteriores lo que podría recoger ruido de píxel. Esto puede conducir a una máscara 120 que parezca como entremezclado que sería inconsistente con la meta de ser capaz de capturar texto y/o líneas de gráficos en la máscara 120. Por ello, el componente 140 de fusión de regiones se adapta para intentar fusionar pares de regiones de la imagen 110 del documento basándose, al menos en parte, en una determinación de si las energías de un primer plano y/o un plano posterior de las regiones fusionadas potenciales son menores que la primera energía de umbral. El componente 140 de fusión de regiones puede utilizar las sumas parciales $\sum_P f(x)$, $\sum_P f(x)^2$, $\sum_F f(x)$ y/o $\sum_F f(x)^2$ junto con N_P y/o N_F calculados y almacenados por el componente 130 de energía de píxel.

Después de cada fusión, estas cantidades deben recalcularse, pero afortunadamente, esto se realiza también en tiempo constante solamente sumando esas cantidades de acuerdo con la combinación primer plano y plano posterior. También obsérvese que la suma $\sum f(x)^2$ en todas las regiones es constante para cada partición, y no necesita ser calculada con la finalidad de seleccionar la partición óptima. Sin embargo, esta cantidad aún se necesitará para decidir cuándo no fusionar regiones.

En referencia brevemente a la Fig. 6, se ilustra una fusión potencial de una primera región que tiene un primer plano P_1 y un plano posterior F_1 con una segunda región que tiene un primer plano P_2 y un plano posterior F_2 . En la determinación de si las energías de las regiones a ser fusionadas son menores que la primera energía de umbral, el componente 150 de fusión de regiones puede calcular agrupaciones de un nuevo primer plano y un nuevo plano posterior. Las varianzas de la energía dentro de las dos regiones tienen siete posibles agrupaciones de un nuevo primer plano y un nuevo plano posterior:

Nuevo primer plano de región fusionada potencial	Nuevo plano posterior de región fusionada potencial
P_1	$F_1 P_2 F_2$
$P_1 F_1 P_2$	F_2
$P_1 F_2 P_2$	F_1
P_2	$P_1 F_1 F_2$
$P_1 P_2$	$F_1 F_2$
$P_1 F_1$	$P_2 F_2$
$P_1 F_2$	$P_2 F_1$

Si al menos una de las posibles agrupaciones proporciona energías de plano posterior y/o primer plano menores que la primera energía de umbral, el componente 140 de fusión de regiones puede determinar una partición primer plano/plano posterior adecuada de la región fusionada, por ejemplo, basándose al menos en parte en la minimización de las energías de plano posterior y primer plano (por ejemplo, $E = v_P + v_F$). Si ninguna de estas agrupaciones proporciona una energía más baja que la primera energía de umbral, la fusión no tiene lugar, y estas regiones no se consideran además para fusión. Por defecto P_1 y P_2 se usarán para calcular los píxeles del primer plano, mientras que F_1 y F_2 se usarán para calcular los píxeles del plano posterior.

El componente 140 de fusión de regiones puede continuar intentando fusionar regiones mayores hasta que la primera energía de umbral fuera superada y/o sustancialmente toda la imagen 110 del documento se haya fusionado. Por ejemplo, el componente 140 de fusión de regiones puede fusionar regiones de dos por dos horizontalmente adjuntas en una región de dos por cuatro tal como se ilustra en la Fig. 6. Posteriormente, el componente 140 de fusión de regiones puede fusionar regiones verticalmente en una región de cuatro por cuatro tal como se ilustra en la Fig. 7. En general, el primer valor de energía de umbral puede seleccionarse para mitigar situaciones potenciales en las que las fusiones intentadas dividirían varios niveles de gris en el primer plano o en el plano posterior, con una pérdida potencial de detalles importantes, tal como texto (por ejemplo, cuando hay más de dos colores en una región). Por ejemplo, si una primera región tiene texto escrito en gris sobre blanco, y una segunda región es principalmente negra, la fusión de las dos regiones puede conducir a gris y blanco yendo al primer plano y el negro al plano posterior de la región fusionada resultante, dando como resultado una pérdida de sustancialmente toda la información textual a partir de la máscara 120. Sin embargo, siempre que se fusionan dos colores o bien en primer plano o bien en plano posterior, tiene lugar un brusco incremento de energía para esa región, dado que una constante ya no es un buen modelo para esta región.

Además, como una alternativa al cálculo de la energía resultante para sustancialmente todas las siete combinaciones, puede clasificarse la media en primeros planos y planos posteriores y puede considerarse la partición con respecto a las medias clasificadas. Como para la clasificación de los valores $V_1 V_2 V_3 V_4$, esto reduce el número de particiones a 3 (clasificar $P_1 F_1 P_2 F_2$ por la media y considerar las particiones con respecto al orden).

Adicional y/o alternativamente, el componente 120 de fusión de regiones puede evaluar un subconjunto restringido de combinaciones de primeros planos y planos posteriores de las dos regiones, basándose en una aproximación f sobre las regiones dadas. El componente 120 de fusión de regiones puede seleccionar la combinación con una energía menos acumulativa.

5 Una vez una región ya no puede fusionarse más debido a que dicha fusión incrementaría la energía más allá del primer umbral, los píxeles en esta región pueden dividirse en primer plano y plano posterior. Dicha partición puede realizarse junto con cada fusión, pero esto sería computacionalmente caro. Alternativamente, el componente 140 de fusión de regiones puede calcular una media de sustancialmente todos los valores de píxeles de la región fusionada y asignar píxel(es) que tengan un valor mayor que la media al primer plano siendo asignados el (los) restante(s) píxel(es) al plano posterior. Alternativamente, el (los) píxel(es) que tengan un valor mayor que la media pueden estar asignados al plano posterior siendo asignados el (los) restante(s) píxel(es) al primer plano.

10 Volviendo a referirnos a la Fig. 1, una vez se ha agotado la fusión por el componente 140 de fusión de regiones, el componente 150 de almacenamiento de máscara está adaptado para almacenar información asociada con la partición del primer plano y plano posterior realizada por el componente 130 de energía de píxel y/o el componente 140 de fusión de regiones en la máscara 120. De ese modo, la máscara 120 indica si cada píxel de la imagen 110 del documento pertenece al primer plano y/o al plano posterior.

15 En un ejemplo, para minimizar la sobrecarga computacional, puede calcularse la energía para una región pequeña (por ejemplo, cuatro píxeles por cuatro píxeles) por el componente 130 de energía de píxel. Si la energía es menor que una segunda energía de umbral, sustancialmente todos los píxeles pueden asignarse al primer plano o al plano posterior quedando el otro sustancialmente vacío. Si la energía es mayor que o igual a la segunda energía de umbral, la partición puede proseguir como se ha descrito anteriormente. Para imagen(es) de documentos relativamente limpios (por ejemplo que tienen áreas constantes), puede conseguirse un incremento significativo en la velocidad computacional.

20 En otro ejemplo, para minimizar el tamaño de la máscara, si una región final (por ejemplo una región que no puede mezclarse sin superar el primer umbral), tiene una diferencia entre el primer plano medio y el plano posterior medio que es más alta que un tercer umbral, toda la región se declara primer plano o se declara plano posterior, dependiendo de si la media global para la región es mayor o menor que el valor de nivel de gris medio (por ejemplo, 127 si los valores de nivel de gris están entre 0 y 255). Para documentos en color que tienen un ligero tramado, la máscara para esta región parecería igual a entremezclado sin esta optimización y tendría un alto coste de compresión. El tercer umbral se elige de modo que no se pierda texto importante, y continúe eliminando los muchos casos de ligero tramado vistos en documentos impresos escaneados (por ejemplo, muchas impresoras tienen solo 4 a 6 colores y deben usar un tramado para generar toda la paleta de colores). Por ejemplo, un valor de 40 puede ser una buena elección para el tercer umbral.

25 El componente 100 separador de máscara se ha descrito con relación a la suposición de que el primer plano y el plano posterior de la cada uno son generalmente constantes. Sin embargo, de acuerdo con un aspecto de la presente invención, puede usarse una regresión polinómica por el componente 130 de energía de píxel para describir el primer plano y/o el plano posterior. Por ejemplo, si los polinomios del primer plano y/o el plano posterior son planos de ecuación $ax + \beta y + \mu$, la energía se definiría mediante:

$$v_P = \sum_{x,y \in P} (f(x,y) - \alpha_P x + \beta_P y + \mu_P)^2$$

$$40 \quad v_F = \sum_{x,y \in F} (f(x,y) - \alpha_F x + \beta_F y + \mu_F)^2$$

en la que x , y indexan las localizaciones del píxel, y α_P , β_P y μ_P son escalares que minimizan v_P y α_F , β_F y μ_F son escalares que minimizan v_F . Tómese nota de que α_P , β_P y μ_P puedan resolverse en tiempo constante usando las cantidades $\sum f(x,y)^2$, $\sum f(x,y)x$, $\sum f(x,y)y$ y $\sum f(x,y)$ que es un sistema lineal de tres incógnitas y tres ecuaciones. Igualmente, α_F , β_F y μ_F puedan resolverse de una forma similar. Como se ha descrito anteriormente con relación a un primer plano y/o plano posterior generalmente constantes, el componente 130 de energía de píxel prosigue para calcular energías de píxel para regiones pequeñas que dividen la región en un primer plano y plano posterior basándose en la minimización de la energía. Posteriormente, las regiones pequeñas se intenta que sean sucesivamente fusionadas por el componente 140 de fusión de regiones basándose, al menos en parte, en la minimización de la energía (E) en cada fusión intentada. Sin embargo, los primeros planos y planos posteriores no pueden clasificarse por media, y por lo tanto todas las 7 combinaciones deben comprobarse para hallar qué combinación minimiza E . Para facilitar las fusiones, las cantidades $\sum f(x,y)^2$, $\sum f(x,y)x$, $\sum f(x,y)y$ y $\sum f(x,y)$ y N pueden almacenarse para cada región para el primer plano y el plano posterior.

De nuevo para minimizar la sobrecarga computacional, puede calcularse la energía para una región pequeña (por ejemplo, cuatro píxeles por cuatro píxeles) por el componente 130 de energía de píxel. Sin embargo, el componente 130 de energía de píxel puede utilizar un modelo basado en una constante sobre la región y/o utilizar una regresión polinómica.

Adicional y/o alternativamente, el componente 110 de energía de píxel puede calcular la energía usando regresión plana. La energía del primer plano v_P y del plano posterior v_F , se definen como:

$$v_P = \sum_{x,y \in P} (f(x,y) + A_P x + B_P y + C_P)^2$$

$$v_F = \sum_{x,y \in F} (f(x,y) + A_F x + B_F y + C_F)^2$$

5 en las que P es el primer plano, F es el plano posterior, $f(x,y)$ es el valor del píxel en la localización x,y . Además, A_P , B_P , C_P se eligen para minimizar la energía del primer plano v_P , y A_F , B_F , C_F se eligen para minimizar la energía del plano posterior v_F . Por ejemplo, la minimización de la energía del primer plano v_P puede obtenerse mediante la resolución del sistema de ecuaciones (3 incógnitas, 3 ecuaciones):

$$\frac{\partial v_P}{\partial A_P} = 0, \quad \frac{\partial v_P}{\partial B_P} = 0, \quad \frac{\partial v_P}{\partial C_P} = 0$$

10 en la que, por ejemplo:

$$\frac{\partial v_P}{\partial A_P} = \sum_{x,y \in P} 2(f(x,y) + A_P x + B_P y + C_P)x = 0$$

pueden resolverse ecuaciones similares para la energía del plano posterior v_F .

15 Mientras la Fig. 1 es un diagrama esquemático que ilustra componentes del componente 100 separador de máscara, se ha de apreciar que el componente 100 separador de máscara puede implementarse como uno o más componentes, tal como se define ese término en la presente memoria. Por ello, se ha de apreciar que los componentes ejecutables por ordenador operativos para implementar el componente 100 separador de máscara pueden almacenarse en un medio legible por ordenador que incluye, pero sin limitarse a, un ASIC (circuito integrado de aplicación específica), CD (disco compacto), DVD (disco de video digital), ROM (memoria solo de lectura), disco flexible, disco duro, EEPROM (memoria solo de lectura programable y borrable eléctricamente) y lápices de memoria de acuerdo con la presente invención.

20 A la vista de los sistemas ejemplares mostrados y descritos anteriormente, una metodología que puede implementarse de acuerdo con la presente invención se apreciará mejor con referencia a los diagramas de flujo de las Figs. 8 y 9. Mientras que, por razones de simplicidad de explicación, la metodología se muestra y describe como una serie de bloques, se ha de entender y apreciar que la presente invención no está limitada por el orden de los bloques, dado que algunos bloques pueden, de acuerdo con la presente invención, suceder en diferentes órdenes y/o simultáneamente con otros bloques respecto a lo mostrado y descrito en la presente memoria. Más aún, no todos los bloques ilustrados pueden requerirse para implementar una metodología de acuerdo con la presente invención.

30 La invención puede describirse en el contexto general de las instrucciones ejecutables por ordenador, tales como módulos de programa, ejecutados por uno o más componentes. Generalmente, los módulos de programa incluyen rutinas, programas, objetos, estructuras de datos, etc. que realizan tareas particulares o implementan tipos de datos abstractos particulares. Típicamente la funcionalidad de los módulos de programa puede combinarse o distribuirse según se desee en diversas realizaciones.

35 Volviendo a las Figs. 8 y 9, se ilustra una metodología 800 para la generación de una máscara de acuerdo con un aspecto de la presente invención. En 810, se calcula la energía de píxel (por ejemplo varianzas) para una región de una imagen de un documento. Por ejemplo, la energía de píxel calculada pueden ser varianzas que pueden basarse, al menos en parte, en una regresión polinómica de la región. Además, las varianzas de energía de píxel calculadas pueden calcularse utilizando un valor de píxel medio calculado para un primer plano y un valor de píxel medio calculado para un plano posterior empleando en una suma de cuadrados de distancias para sustancialmente todos los píxeles en la región. Alternativamente, las varianzas de energía de píxel calculadas pueden calcularse utilizando una suma de valores de píxel para un plano posterior y una suma de valores de píxel calculadas para un primer plano.

45 A continuación, en 820, se calcula una partición de píxeles para la región para minimizar la energía de píxeles de un primer plano y/o un plano posterior. En 830, se realiza la determinación de si sustancialmente todas las regiones de la imagen del documento se han dividido. Si la determinación en 830 es NO, el procesamiento continúa en 810. Si la determinación en 830 es SÍ, el procesamiento continúa en 840.

A continuación, en 840, se intenta que se fusionen pares de regiones adyacentes. En 850, se realiza la determinación de si la fusión intentada daría como resultado que se excediera una energía de umbral en un nuevo primer plano y/o un nuevo plano posterior. Si la determinación en 850 es SÍ, el procesamiento continúa en 860. Si la

determinación en 850 es NO, en 870, las regiones se fusionan. En 880, se calcula una nueva partición de primer plano y plano posterior de la región fusionada. En 885, se realiza la determinación de si se ha intentado que se fusionen sustancialmente todas las regiones de la imagen del documento. Si la determinación en 885 es SÍ, no tiene lugar ningún procesamiento adicional. Si la determinación 885 es NO, el procesamiento continúa en 840.

- 5 En 860, se realiza una determinación de si se ha intentado que se mezclen sustancialmente todas las regiones de la imagen del documento. Si la determinación en 860 es NO, en 890, el foco sobre las fusiones intentadas se mueve a la siguiente región sin fusionar (por ejemplo, la región de dos píxeles por dos píxeles). Si la determinación en 860 es SÍ, no tiene lugar ningún procesamiento adicional.

10 A continuación, en referencia a la Fig. 10, se ilustra un sistema 1000 para la separación de la imagen del documento de acuerdo con un aspecto de la presente invención. El sistema 1000 incluye un componente 100 separador de máscara y un segmentador 160 primer plano/plano posterior. El componente 100 separador de máscara incluye un componente 130 de energía de píxel, un componente 140 de fusión de regiones y un componente 150 de almacenamiento de máscara.

15 Como se ha descrito anteriormente, de acuerdo con un aspecto de la presente invención, el componente 100 separador de máscara recibe una imagen 110 de un documento como una entrada. El componente 100 separador de máscara procesa la imagen del documento para generar una máscara 120 como una salida.

20 El segmentador 160 primer plano/plano posterior recibe la máscara 120 y la imagen 110 del documento como entradas. Basándose, al menos en parte, en la máscara 120, el segmentador 160 primer plano/plano posterior está adaptado para separar la imagen 110 del documento en una imagen 170 del primer plano y una imagen 180 del plano posterior. Por ejemplo, sustancialmente el (todos los) píxel(es) representado(s) por un "1" en la máscara 120 pueden ir a la imagen 170 del primer plano y sustancialmente el (todos los) píxel(es) representado(s) por un "0" en la máscara 120 pueden ir a la imagen 180 del plano posterior. A la inversa, como un ejemplo, sustancialmente el (todos los) píxel(es) representado(s) por un "0" en la máscara 120 pueden ir a la imagen 170 del primer plano y sustancialmente el (todos los) píxel(es) representado(s) por un "1" en la máscara 120 pueden ir a la imagen 180 del plano posterior.

25 Por ejemplo, el componente 100 separador de máscara puede procesar una imagen 110 del documento que comprende texto mediante la separación de píxeles (por ejemplo, asociados con el texto) en un primer plano y almacenamiento de la información en relación a qué píxeles están en el primer plano en una máscara 120. Posteriormente, el segmentador 160 primer plano/plano posterior puede recibir la máscara 120 y la imagen 110 del documento. El segmentador 160 primer plano/plano posterior puede separar la imagen 110 del documento en la imagen 170 del primer plano y la imagen 180 del plano posterior.

30 Pasando a la Fig. 11, se ilustra un sistema 1100 para la compresión de imágenes de documento de acuerdo con un aspecto de la presente invención. El sistema 1100 incluye un componente 1110 de transformación de la imagen del documento, un componente 100 de separación de máscara y un componente segmentador 160 primer plano/plano posterior. El segmentador 160 primer plano/plano posterior recibe la máscara 120 y la imagen 1150 del documento como entradas. Basándose, al menos en parte, en la máscara 120, el segmentador 160 primer plano/plano posterior está adaptado para separar la imagen 1150 del documento en una imagen 170 del primer plano y una imagen 180 del plano posterior.

35 Como se ilustra en la Fig. 12, el sistema 1100 puede, opcionalmente, incluir un codificador 1120 de máscara, un codificador 1130 del primer plano y/o un codificador 1140 del plano posterior. El componente 100 de separación de máscara incluye un componente 130 de energía de píxel, un componente 140 de fusión de regiones y un componente 150 de almacenamiento de máscara. Opcionalmente, el sistema 110 puede incluir un procesador 1170 de la imagen del primer plano y/o un procesador 1180 de la imagen del plano posterior.

40 El componente 1110 de transformación de la imagen del documento está adaptado para recibir una imagen 1150 del documento y producir la salida de una representación transformada de la imagen 1160 del documento. Por ejemplo, el componente 1110 de transformación de la imagen del documento puede recibir una imagen del documento RGB y producir la salida de una representación YUV de la imagen del documento RGB.

45 El codificador 1120 de máscara está adaptado para codificar la máscara 120. Por ejemplo, dado que la máscara 120 es típicamente binaria, el codificador 1120 de máscara puede utilizar una(s) técnica(s) de compresión binaria convencional(es) para conseguir una compresión efectiva de la máscara. El codificador 1120 de máscara produce la salida de un flujo de bits de máscara.

50 El codificador 1130 del primer plano está adaptado para codificar la imagen 170 del primer plano. El primer plano es una imagen compuesta de los píxeles del primer plano, y píxeles "no importantes" (por ejemplo, píxeles que originalmente pertenecían al plano posterior). El procesador 1170 de la imagen del primer plano puede usarse para rellenar los píxeles "no importantes" con valores que facilitan la compresión y proporcionan la imagen de primer plano alterada al codificador 1130 del primer plano. Por ejemplo, en el caso de que la imagen 170 del primer plano comprenda generalmente información textual en color negro, los píxeles "no importantes" pueden también rellenarse en negro, de modo que toda la imagen del primer plano sea negra. El codificador 1130 del primer plano puede

utilizar una técnica(s) de compresión efectiva(s) para compresión de imagen, tal como JPEG, ondículas, o cualesquiera otros algoritmos de compresión de imagen. El codificador 1130 del primer plano produce la salida de un flujo de bits de primer plano.

5 El codificador 1140 de plano posterior está adaptado para codificar la imagen 180 del plano posterior. El plano posterior es una imagen compuesta de los píxeles del plano posterior, y píxeles “no importantes” (por ejemplo, píxeles que originalmente pertenecían al primer plano). El procesador 1180 de la imagen del plano posterior puede usarse para rellenar los píxeles “no importantes” con valores que facilitan la compresión y proporciona la imagen del plano posterior alterada al codificador 1140 del plano posterior. Por ejemplo, en el caso de que la imagen 180 del plano posterior comprenda una página blanca lisa, los píxeles “no importantes” que se localizan en donde estaba el texto pueden rellenarse con blanco, de modo que toda la imagen del plano posterior sea blanca. El codificador 1140 del plano posterior puede utilizar técnicas de compresión efectiva para la compresión de imagen, tales como JPEG, ondículas, o cualesquiera otros algoritmos de compresión de imagen. El codificador 1140 del plano posterior produce la salida de un flujo de bits de plano posterior.

15 Por ejemplo, un algoritmo simple para rellenar los píxeles “no importantes” es procesar la imagen con un filtro paso bajo, y a continuación restaurar los píxeles importantes. Después de unas pocas interacciones de estas dos etapas, los píxeles “no importantes” acaban con valores que permiten transiciones suaves, y que se comprimirán bien. Un simple refinamiento de este algoritmo es comenzar con un filtro de paso muy bajo e incrementar la frecuencia de corte del filtro paso bajo en cada iteración.

20 Además, el codificador 1130 del primer plano y/o el codificador 1140 del plano posterior pueden utilizar la máscara 120 para mejorar la compresión del primer plano y/o el plano posterior. Se ha de apreciar que se contemplan numerosos codificadores y/o decodificadores que utilizan una máscara que se basa, al menos en parte, en una partición de una imagen del documento basándose, al menos en parte, en la minimización de las varianzas de energía de píxel de al menos uno de entre un primer plano y un plano posterior en conexión con la invención sujeto. Cualquier codificador y/o decodificador de ese tipo adecuado para su empleo en conexión con la presente invención se pretende que caiga dentro del alcance de las reivindicaciones adjuntas.

25 El flujo de bits de la máscara, el flujo de bits del primer plano y/o el flujo de bits del plano posterior pueden combinarse en un único flujo de bits y/o enviarse individualmente a, por ejemplo, un sistema de decodificación (no mostrado). El sistema de decodificación puede decodificar el flujo de bits de la máscara para obtener la máscara 120. Alternativamente, el sistema de decodificación puede recibir la máscara 120. El sistema de decodificación puede utilizar la máscara 120 para recombinar el flujo de bits del primer plano y/o el flujo de bits del plano posterior en una imagen del documento.

30 Se ha de apreciar que el sistema y/o procedimiento de la presente invención pueden utilizarse en un sistema de imagen en capas segmentado global que facilita la identificación y/o compresión de texto, escritura manual, dibujos y similares. Además, los expertos en la materia reconocerán que el sistema y/o procedimiento de la presente invención puede emplearse en una amplia variedad de aplicaciones de imágenes de documentos, incluyendo, pero sin limitarse a, fotocopiadoras, escáneres de documentos, sistemas de reconocimiento óptico de caracteres, PDA, máquinas de fax, cámaras digitales, cámaras de vídeo digital y/o sistemas de videojuegos.

Pasando a la Fig. 13, se ilustra un sistema 1300 de imagen segmentada en capas. El sistema 1300 incluye un componente 130 de energía de píxel y un componente 150 de almacenamiento de máscara.

40 El componente 130 de energía de píxel está adaptado para calcular las varianzas de energía de píxel para una región de una imagen del documento. Además, el componente 130 de energía de píxel está adaptado además para calcular una partición de la región basándose al menos en parte, en la minimización de la energía de píxel (por ejemplo, varianzas) de al menos uno de entre un primer plano y un plano posterior.

45 El componente 150 de almacenamiento de máscara está adaptado para almacenar información asociada con la partición en una máscara. Por ejemplo, el sistema 1300 puede emplearse en una amplia variedad de aplicaciones de imágenes de documentos, que incluyen, pero sin limitarse a, fotocopiadoras, escáneres de documentos, sistemas de reconocimiento óptico de caracteres, PDA, máquinas de fax, cámaras digitales, cámaras de vídeo digital y/o sistemas de videojuegos.

50 Para proporcionar un contexto adicional para varios aspectos de la presente invención, la Fig. 14 y la explicación a continuación se pretende que proporcionen una descripción breve, general de un posible entorno 1410 informático adecuado en el que pueden implementarse varios aspectos de la presente invención. Se ha de apreciar que el entorno 1410 informático no es más que un posible entorno informático y no se pretende que limite los entornos informáticos en los que puede emplearse la presente invención. Aunque la invención se ha descrito anteriormente en el contexto general de instrucciones ejecutables por ordenador que pueden ejecutarse sobre uno o más ordenadores, se ha de reconocer que la invención también puede implementarse en combinación con otros módulos de programa y/o como una combinación de hardware y software. Generalmente, los módulos de programa incluyen rutinas, programas, componentes, estructuras de datos, etc., que realizan tareas particulares o implementan tipos de datos abstractos particulares. Más aún, se apreciará que los procedimientos inventivos pueden ponerse en práctica

con otras configuraciones de sistemas de ordenador, incluyendo sistemas de ordenador de procesador único o multiprocesador, miniordenadores, ordenadores centrales, así como ordenadores personales, y dispositivos de cálculo portátiles, electrónica de consumo basada en microprocesador o programable y similares, cada uno de los cuales puede acoplarse operativamente a uno o más dispositivos asociados. Los aspectos ilustrados de la invención pueden también ponerse en práctica en entornos de cálculo distribuido en donde ciertas tareas se realizan mediante dispositivos de procesamiento remoto que se enlazan a través de una red de comunicaciones. En un entorno de cálculo distribuido, los módulos de programa pueden localizarse en dispositivos de almacenamiento de memoria tanto locales como remotos.

La Fig. 14 ilustra una posible configuración de hardware para soportar los sistemas y procedimientos descritos en la presente memoria. Se ha de apreciar lo que se ilustra una estructura independiente, que puede emplearse en cualquier entorno informático adecuado de acuerdo con la presente invención. Por ejemplo, pueden emplearse de acuerdo con la presente invención arquitecturas de cálculo que incluyen, pero sin limitarse a, independiente, multiprocesador, distribuida, cliente/servidor, miniordenador, ordenador central, superordenador, digitales y analógicas.

Con referencia a la Fig. 14, un entorno 1410 de ejemplo para la implementación de diversos aspectos de la invención incluye un ordenador 1412, que incluye una unidad 1414 de procesamiento, una memoria 1416 del sistema y un bus 1418 del sistema que conecta diversos componentes del sistema incluyendo la memoria del sistema a la unidad 1414 de procesamiento. La unidad 1414 de procesamiento puede ser cualquiera de diversos procesadores comercialmente disponibles. También pueden usarse microprocesadores dobles y otras arquitecturas multiprocesador como la unidad 1414 de procesamiento.

El bus 1418 del sistema puede ser cualquiera de diversos tipos de estructuras de bus que incluyen un bus de memoria o controlador de memoria, un bus periférico, y un bus local que usan cualquiera de una variedad de arquitecturas de bus comercialmente disponibles. La memoria 1416 del ordenador incluye memoria 1420 solo de lectura (ROM), y memoria 1422 de acceso aleatorio (RAM). Se almacena en la ROM 1420 un sistema de entrada/salida básico (BIOS), que contiene las rutinas básicas que ayudan a transferir información entre elementos dentro del ordenador 1412, tal como durante el arranque.

El ordenador 1412 puede incluir además una unidad 1424 de disco duro, una unidad 1426 de disco magnético, por ejemplo, para leer desde o escribir a un disco 1428 extraíble, y una unidad 1430 de disco óptico, por ejemplo, para lectura de un disco 1432 CD ROM o para ver desde o escribir en otros medios ópticos. La unidad 1424 de disco duro, la unidad 1426 de disco magnético y la unidad 1430 de disco óptico se conectan al bus 1418 del sistema mediante una interfaz 1434 de unidad de disco duro, una interfaz 1436 de unidad de disco magnético y una interfaz 1438 de unidad óptica, respectivamente. El ordenador 1412 incluye típicamente al menos alguna forma de medio legible por ordenador. Un medio legible por ordenador puede ser cualquier medio disponible al que pueda accederse por parte del ordenador 1412. A modo de ejemplo, y no de limitación, un medio legible por ordenador pueden comprender medios de almacenamiento informático y medios de comunicación. El medio de almacenamiento informativo incluye medios volátiles y no volátiles, extraíbles y no extraíbles implementados en cualquier procedimiento o tecnología para almacenamiento de información tales como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento informático incluyen, pero sin limitarse a, RAM, ROM, EEPROM, memoria flash u otra tecnología de memoria, CD ROM, discos versátiles digitales (DVD) u otros dispositivos de almacenamiento magnético, o cualquier otro medio que pueda usarse para almacenar la información deseada y a la que pueda accederse por parte del ordenador 1412. El medio de comunicación incluye típicamente 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 incluye cualquier medio de entrega de información. La expresión "señal de datos modulada" significa una señal que tiene una o más de sus características ajustadas o cambiadas de tal manera que codifique informaciones de la señal. A modo de ejemplo, y no de limitación, los medios de comunicación incluyen medios cableados tales como una red cableada o conexión directa por cable, y medios inalámbricos tal como acústicos, de RF, infrarrojos y otros medios inalámbricos. Las combinaciones de cualquiera de los anteriores deberían también incluirse dentro del alcance de los medios legibles por ordenador.

Un cierto número de módulos de programa puede almacenarse en las unidades y en RAM 1422, incluyendo un sistema operativo 1440, uno o más programas 1442 de aplicación, otros módulos 1444 de programas, y datos 1446 no de interrupción de programas. El sistema operativo 1440 en el ordenador 1412 puede ser cualquiera de un cierto número de sistemas operativos disponibles comercialmente.

Un usuario puede introducir órdenes e información dentro del ordenador 1412 a través de un teclado 1448 y dispositivo de puntero tal como un ratón 1450. Otros dispositivos de entrada (no mostrados) pueden incluir un micrófono, un control remoto por IR, una palanca de juegos, una alfombrilla de juegos, una antena de satélite, un escáner o similares. Estos y otros dispositivos de entrada se acoplan frecuentemente a la unidad 1414 de procesamiento a través de una interfaz 1452 de puerto serie que se acopla al bus 1418 del sistema, pero que puede conectarse mediante otras interfaces, tales como un puerto paralelo, un puerto de juegos, un bus serie universal ("USB"), una interfaz de IR, etc. Un monitor 1454, u otro tipo de dispositivo de visualización, se conectan también al bus 1418 del sistema mediante una interfaz, tal como un adaptador 1456 de video. Además del monitor, un

ordenador incluye típicamente otros dispositivos de salida periféricos (no mostrados), tales como altavoces, impresoras, etc.

El ordenador 1412 puede operar en un entorno de red usando conexiones lógicas y/o físicas a uno o más ordenadores remotos, tal como un ordenador 1458 remoto. El ordenador 1458 remoto puede ser una estación de trabajo, un ordenador servidor, un enrutador, un ordenador personal, una aplicación de diversión basada en microprocesador, un dispositivo parejo u otro nodo de red común, e incluye típicamente muchos o todos los elementos descritos con relación al ordenador 1412, aunque, por razones de brevedad, solo se ilustra un dispositivo 1460 de almacenamiento de memoria. Las conexiones lógicas representadas incluyen una red 1462 de área local (LAN) y una red 1464 de área grande (WAN). Dichos entornos de red son comunes en oficinas, redes de ordenadores de ámbito de empresa, intranets y la Internet.

Cuando se usa en un entorno de red LAN, el ordenador 1412 se conecta a la red 1462 local a través de una interfaz o adaptador 1466 de red. Cuando se usa en un entorno de red WAN, el ordenador 1412 incluye típicamente un módem 1468, o se conecta a un servidor de comunicaciones sobre la LAN, o tiene otros medios para el establecimiento de comunicaciones sobre la WAN 1464, tal como la Internet. El modo en 1468, que puede ser interno o externo, se conecta al bus 1418 del sistema a través de la interfaz 1452 de puerto serie. En un entorno en red, los módulos de programas representados con relación al ordenador 1412, o partes de los mismos, pueden almacenarse en el dispositivo 1460 de almacenamiento de memoria remoto. Se apreciará que las conexiones de red mostradas son ejemplares y que pueden usarse otros medios de establecimiento de un enlace de comunicaciones entre los ordenadores.

La Fig. 15 es un diagrama de bloques esquemático de un entorno 1500 informático de muestra con el que puede interactuar la presente invención. El sistema 1500 incluye uno o más cliente(s) 1510. El (los) cliente(s) 1510 pueden ser hardware y/o software (por ejemplo, hilos, procesos, dispositivos de cálculo). El sistema 1500 incluye también uno o más servidor(es) 1530. El (los) servidor(es) 1530 pueden ser hardware y/o software (por ejemplo, hilos, procesos, dispositivos de cálculo). El (los) servidor(es) 1530 pueden alojar hilos para realizar transformaciones mediante el empleo de la presente invención, por ejemplo. Una comunicación posible entre un cliente 1510 y servidor 1530 puede estar en la forma de un paquete de datos adaptado para ser transmitido entre dos o más procesos de ordenador. El sistema 1500 incluye una estructura 1550 de comunicación que puede emplearse para facilitar las comunicaciones entre el (los) cliente(s) 1510 y el (los) servidor(es) 1530. El (los) cliente(s) 1510 se conecta(n) operativamente a uno o más almacén(es) 1560 de datos de cliente que pueden emplearse para almacenar información local para el (los) cliente(s) 1510. De modo similar, el (los) servidor(es) 1530 se conecta(n) operativamente a uno o más almacén(es) 1540 de datos de servidor(es) que pueden emplearse para almacenar información local para los servidores 1530.

Lo que se ha descrito anteriormente incluye ejemplos de la presente invención. No es posible, naturalmente, describir cada combinación de componentes o metodologías concebible para las finalidades de descripción de la presente invención, pero un experto en la materia reconocerá que son posibles muchas combinaciones y permutaciones adicionales de la presente invención. Además, en el grado en el que se usa el término "incluye" en cualquier de entre la descripción detallada o en las reivindicaciones, dicho término se pretende que sea inclusivo en una manera similar al término "comprendiendo" dado que "comprendiendo" se interpreta cuando se emplea como una palabra transitoria en una reivindicación.

REIVINDICACIONES

1. Un componente (100) separador de máscara, que comprende:

5 un componente (130) de energía de píxel adaptado para calcular la energía de píxel para una región de una imagen (110) de un documento, adaptado además el componente de energía de píxel para calcular una partición de la región basándose al menos en parte en la energía de píxel de al menos uno de entre un primer plano y un plano posterior;

caracterizado porque comprende además:

10 un componente (140) de fusión de regiones adaptado para fusionar pares de regiones adyacentes de la imagen del documento basándose al menos en parte en la determinación de si las energías de un primer plano y/o un plano posterior de las regiones fusionadas potenciales son menores que una primera energía de umbral, estando adaptado además el componente de fusión para dividir la región fusionada en un nuevo primer plano y nuevo plano posterior; y,
un componente (150) de almacenamiento de máscara adaptado para almacenar información asociada con la partición en una máscara (120).

15 2. El componente (100) separador de máscara de la reivindicación 1, adaptado además el componente de energía de píxel para calcular una partición de la región basándose al menos en parte en la minimización de la energía de píxel de al menos uno de entre un primer plano y un plano posterior; y/o comprendiendo la imagen (110) del documento al menos una de entre una representación YUV, RGB y binaria de un documento.

20 3. El componente (100) separador de máscara de la reivindicación 1, siendo la máscara (120) una representación binaria de información asociada con la partición del primer plano y el plano posterior.

4. El componente (100) separador de máscara de la reivindicación 1, calculando el componente (130) de energía de píxel varianzas de energía de píxel para una región utilizando al menos en parte un algoritmo de agrupamiento de K medias, en el que K=2.

25 5. El componente (100) separador de máscara de la reivindicación 1, calculando el componente (130) de energía de píxel varianzas de energía utilizando las ecuaciones:

$$v_P = \sum_{x \in P} (f(x) - \mu_P)^2$$

$$v_F = \sum_{x \in F} (f(x) - \mu_F)^2$$

30 en las que P es el primer plano,
F es el plano posterior,
f(x) es el valor del píxel,

$$\mu_P = \frac{1}{N_P} \sum_{x \in P} f(x) \text{ es el valor de píxel medio del primer plano,}$$

$$\mu_F = \frac{1}{N_F} \sum_{x \in F} f(x) \text{ es el valor de píxel medio del plano posterior,}$$

35 N_P es el número de píxeles en el primer plano, y
 N_F es el número de píxeles en el plano posterior
para el cálculo de las varianzas de energía de píxel de la región.

6. El componente (100) separador de máscara de la reivindicación 1, almacenando el componente (130) de energía de píxel, al menos temporalmente, al menos algo de la información asociada con el cálculo de la energía de píxel.

40 7. El componente (100) separador de máscara de la reivindicación 1, calculando el componente (130) de energía de píxel la energía de píxel para la región basándose al menos en parte en una regresión polinómica de la región.

8. El componente (100) separador de máscara de la reivindicación 1, basándose la partición del componente (140) de fusión de regiones al menos en parte en una media de los valores de píxel de la región fusionada.

9. El componente (100) separador de máscara de la reivindicación 8, siendo asignado un valor de píxel mayor que la media al primer plano.

45 10. El componente (100) separador de máscara de la reivindicación 8, siendo asignado un valor de píxel mayor que la media al plano posterior.

11. El componente (100) separador de máscara de la reivindicación 1, siendo la imagen (110) del documento una

representación YUV y utilizando el componente de energía de píxel al menos el componente Y en el cálculo de la energía de píxel.

12. Un procedimiento de generación de una máscara (120) que divide una imagen (110) del documento en un plano posterior y un primer plano, que comprende:

5 calcular la energía de píxel para una región (810) de la imagen (110) del documento;
 dividir la región basándose al menos en parte en la energía de píxel calculado de al menos uno de entre un primer plano y un plano posterior de la región;
 caracterizado porque comprende:

10 fusionar pares adyacentes de regiones (870) tras una determinación de si las energías de un primer plano y/o un plano posterior de las regiones fusionadas potenciales son menores que una primera energía de umbral;
 dividir (880) la región fusionada en un nuevo primer plano y un nuevo plano posterior; y
 almacenar la información de partición en una máscara.

13. El procedimiento de la reivindicación 12, comprendiendo además el cálculo de la energía de píxel para una región:

15 calcular una suma de valores de píxel para un primer plano;
 calcular una suma de valores de píxel para un plano posterior; y
 calcular la suma de cuadrados de distancias para sustancialmente todos los píxeles en la región.

14. El procedimiento de la reivindicación 12, basándose el cálculo de la energía de píxel para una región, al menos en parte, en una regresión polinómica de la región.

20 15. El procedimiento de la reivindicación 12, comprendiendo la partición de la región además minimizar la energía de píxel de al menos uno de entre un primer plano y un plano posterior.

16. Procedimiento de la reivindicación 12, comprendiendo además la partición de la región fusionada el cálculo de una media de valores de píxel de la región fusionada.

25 17. Procedimiento de la reivindicación 12, comprendiendo además la partición de la región fusionada las siguientes actuaciones:

 calcular un valor de píxel medio del primer plano de la región fusionada;
 calcular un valor de píxel medio del plano posterior de la región fusionada;
 asignar píxeles a uno de entre un primer plano y un plano posterior si la diferencia entre el valor de píxel medio del primer plano y el valor de píxel medio del plano posterior es mayor que un tercer umbral.

30 18. El procedimiento de la reivindicación 12, comprendiendo además la partición de la región las siguientes actuaciones:

 calcular la energía para la región; y
 asignar píxeles de la región a uno de entre el primer plano y el plano posterior si la energía es menor que una segunda energía de umbral;

35 19. Un sistema (1000) de compresión de imágenes de documento, que comprende:

 un componente de transformación del documento adaptado para recibir una imagen (110) del documento y producir la salida de una representación transformada de la imagen del documento; y el componente (100) de separador de máscara de la reivindicación 1.

40 20. El sistema (1000) de compresión de imágenes de documento de la reivindicación 19, que comprende además un componente (160) segmentador de primer plano/plano posterior adaptado para separar la representación transformada en una imagen (170) del primer plano y una imagen (180) del plano posterior basándose, al menos en parte, en información almacenada en la máscara (120).

45 21. El sistema (1000) de compresión de imágenes de documento de la reivindicación 19 que comprende además al menos uno de entre un codificador del primer plano, un codificador del plano posterior y un codificador de la máscara.

22. El sistema (1000) de compresión de imágenes de documento de la reivindicación 19, estando adaptado el componente de transformación del documento para recibir una imagen (110) del documento en RGB y producir la salida de una representación YUV de la imagen del documento RGB.

50 23. Un medio legible por ordenador que tiene componentes utilizables por ordenador adaptados para implementar el componente (100) separador de máscara de la reivindicación 1.

24. El componente (100) separador de máscara de la reivindicación 1 que comprende además:
medios para calcular una partición de la región basándose, al menos en parte, en la minimización de la energía de al menos uno de dos planos; y
medios para la partición de píxeles de las regiones en un primer plano y un plano posterior basándose, al menos en parte, en la que minimizar la energía de los píxeles que comprende al menos uno de entre el primer plano y el plano posterior.
25. Una fotocopiadora que emplea uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.
26. Un escáner de documentos que emplea uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.
27. Un sistema de reconocimiento óptico de caracteres que emplea uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.
28. Un asistente digital personal uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.
29. Una máquina de fax que emplea uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.
30. Una cámara digital que emplea uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.
31. Una cámara de video digital que emplea uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.
32. Un sistema de video juegos que emplea uno o más de lo siguiente:
un componente (100) de separador de máscara de cualquiera de las reivindicaciones 1-11 o 24, el procedimiento de cualquiera de las reivindicaciones 12-18, el sistema (1000) de compresión de imágenes de documento de cualquiera de las reivindicaciones 19-22 y el medio legible por ordenador de la reivindicación 23.

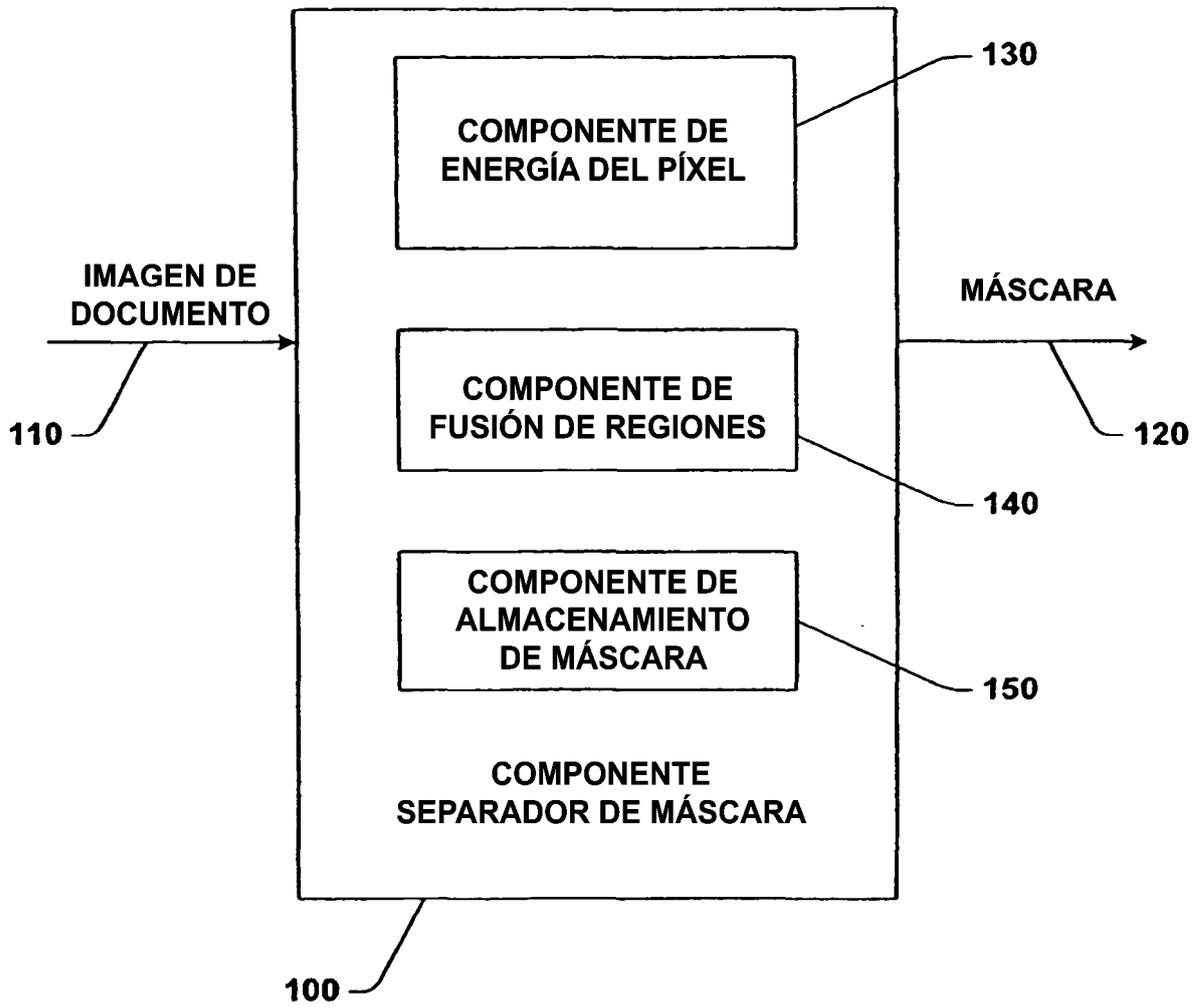
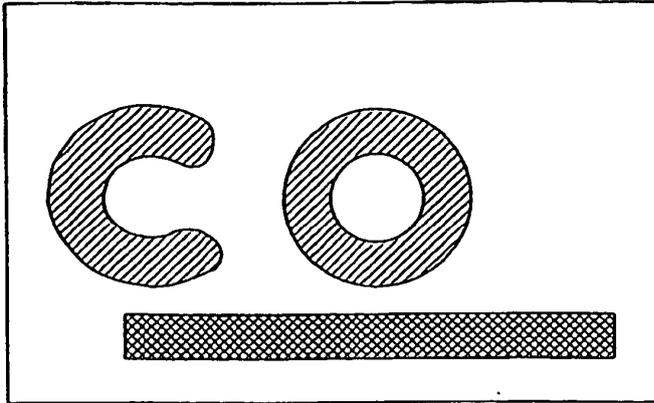


FIG. 1



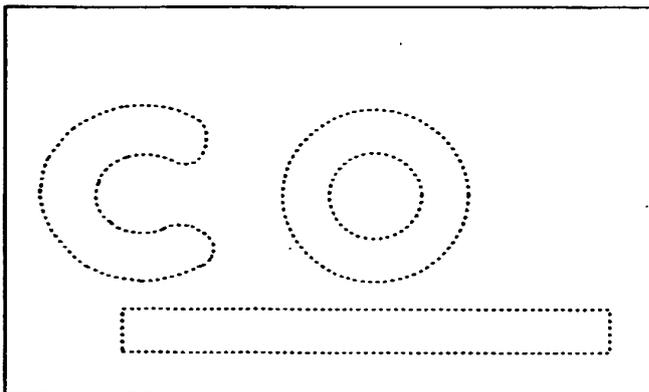
ORIGINAL

FIG. 2



MÁSCARA BINARIA

FIG. 3



PLANO POSTERIOR

FIG. 4

V_1	V_2
V_3	V_4

FIG. 5

P_1	F_1	P_2	F_2
F_1	P_1	P_2	F_2

FIG. 6

P_1, F_1		P_2, F_2	
P_3, F_3		P_4, F_4	

FIG. 7

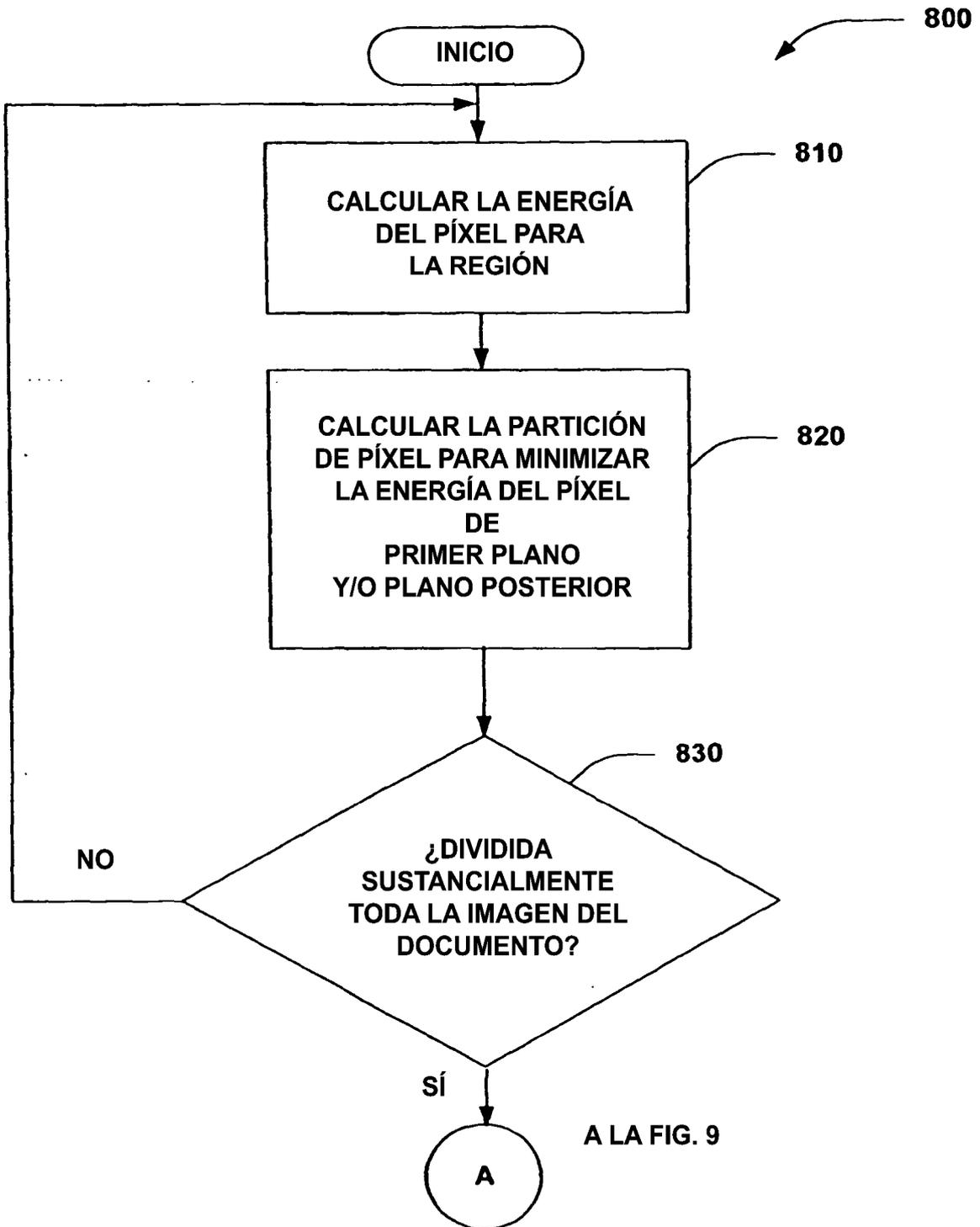


FIG. 8

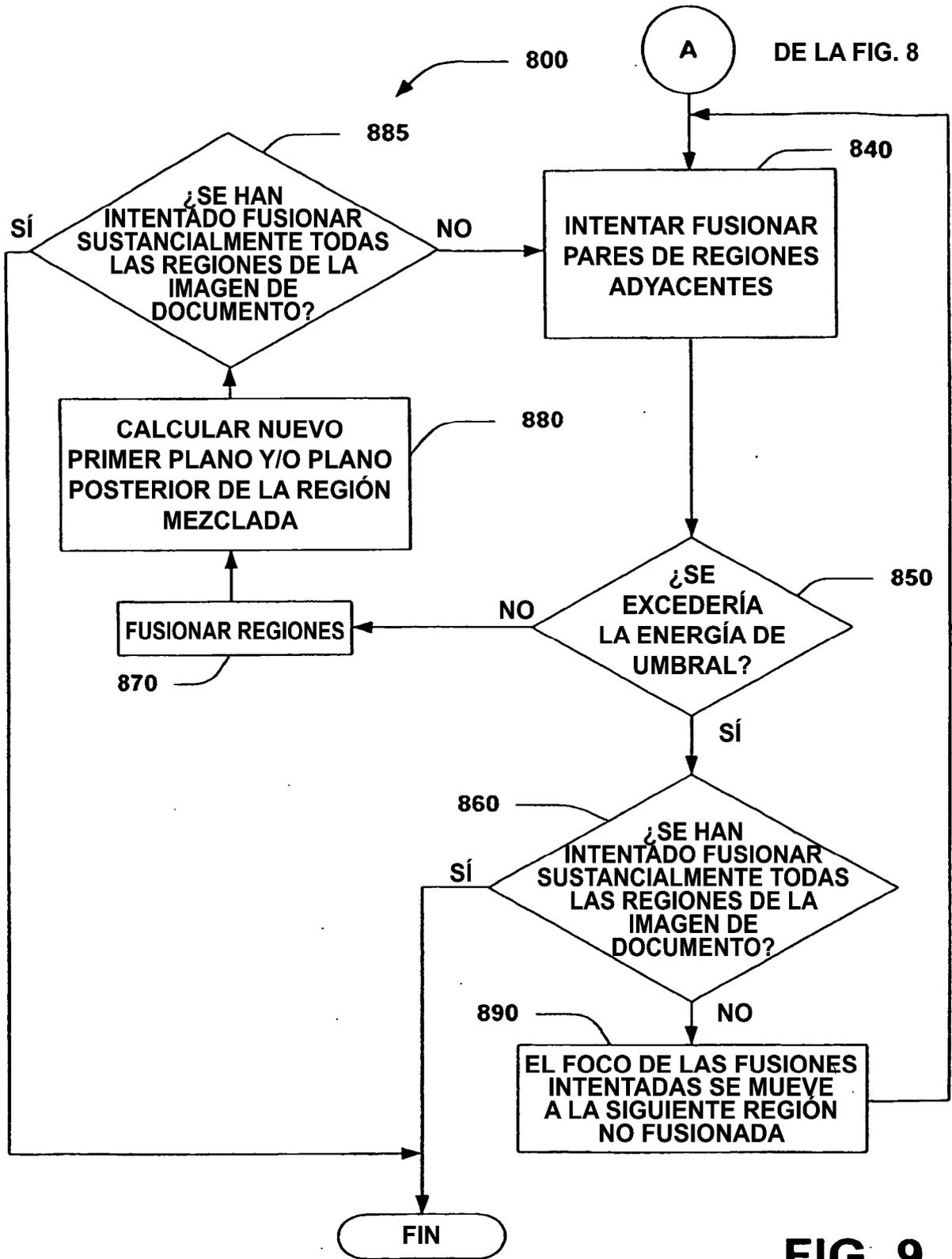


FIG. 9

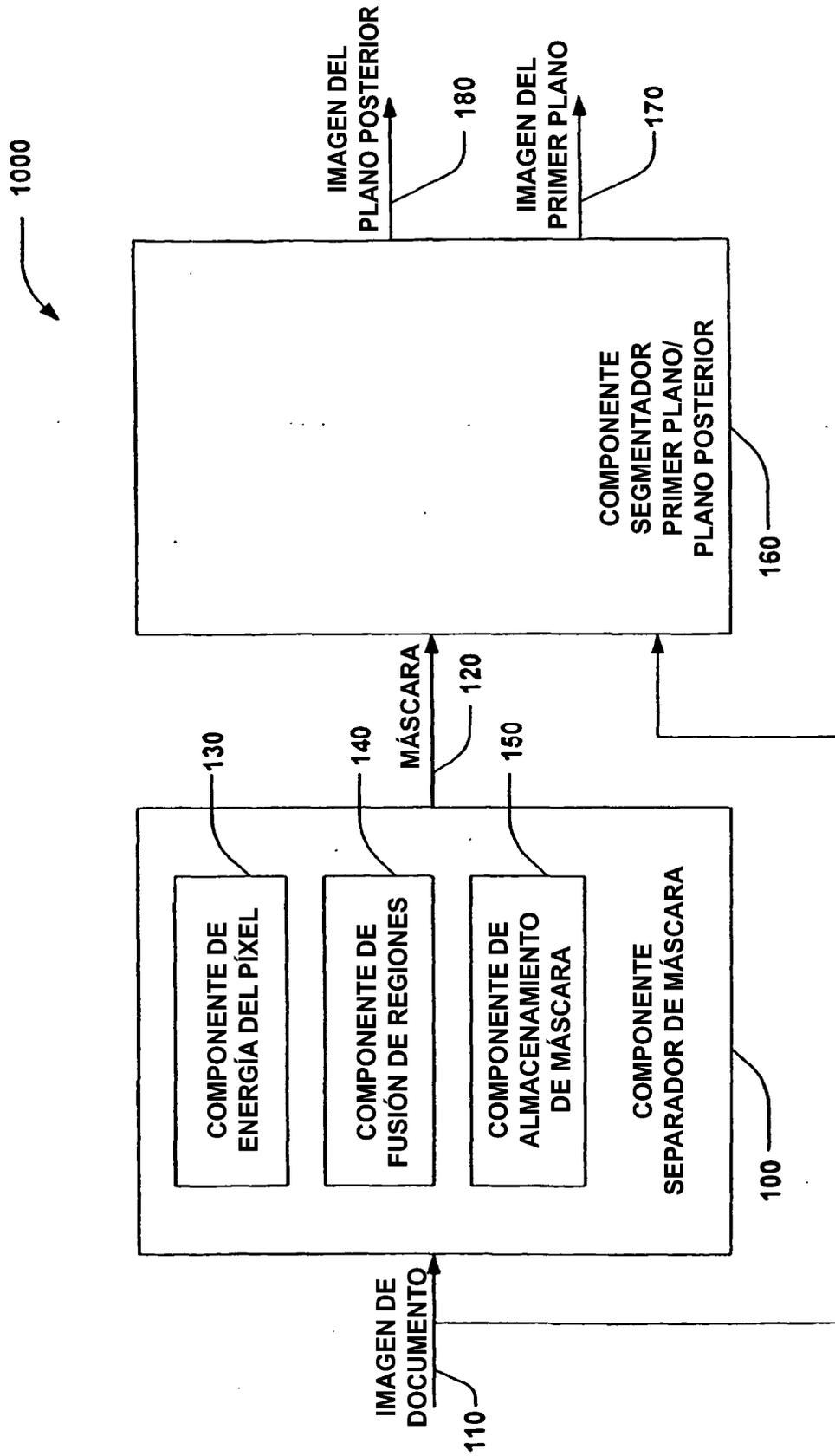


FIG. 10

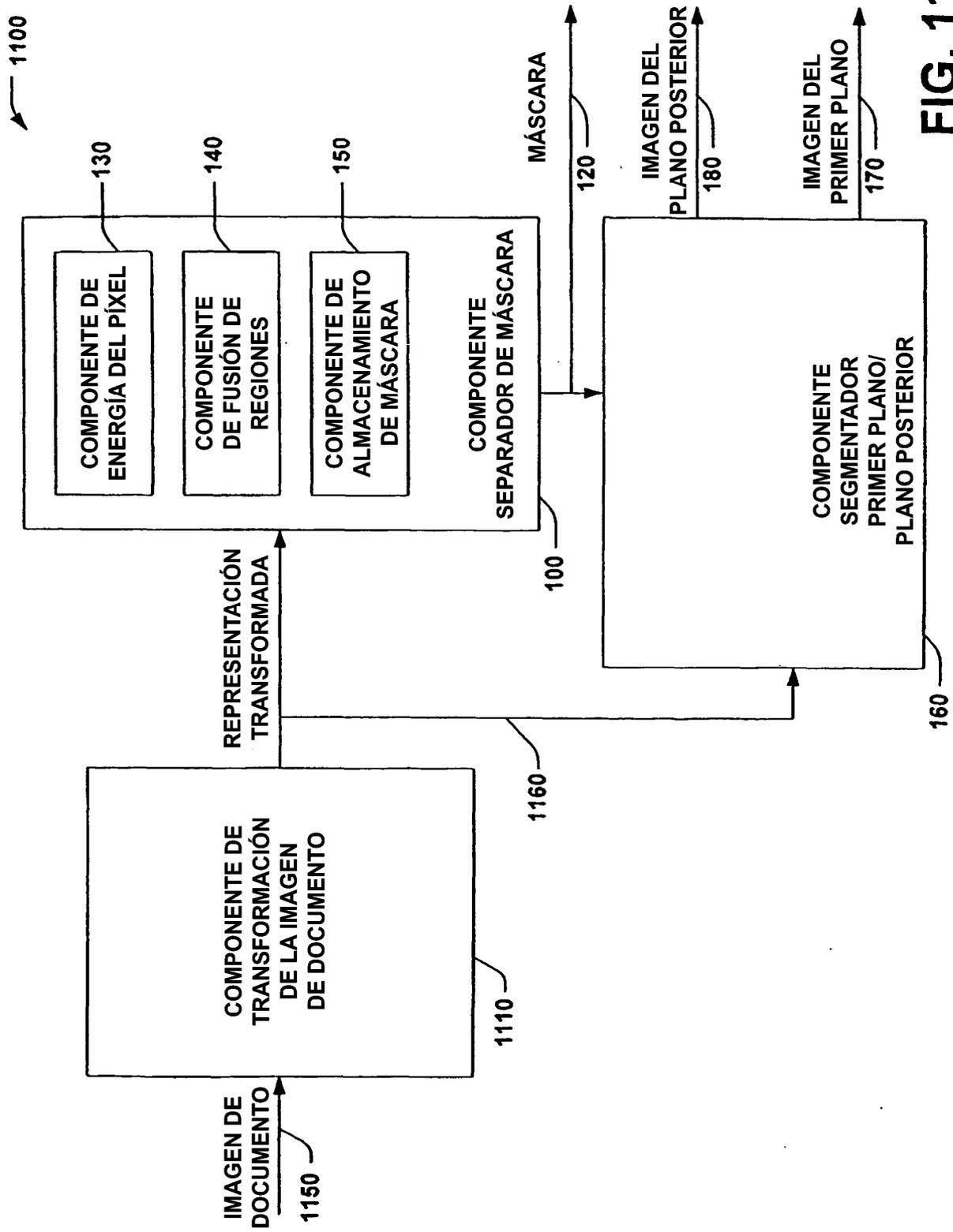


FIG. 11

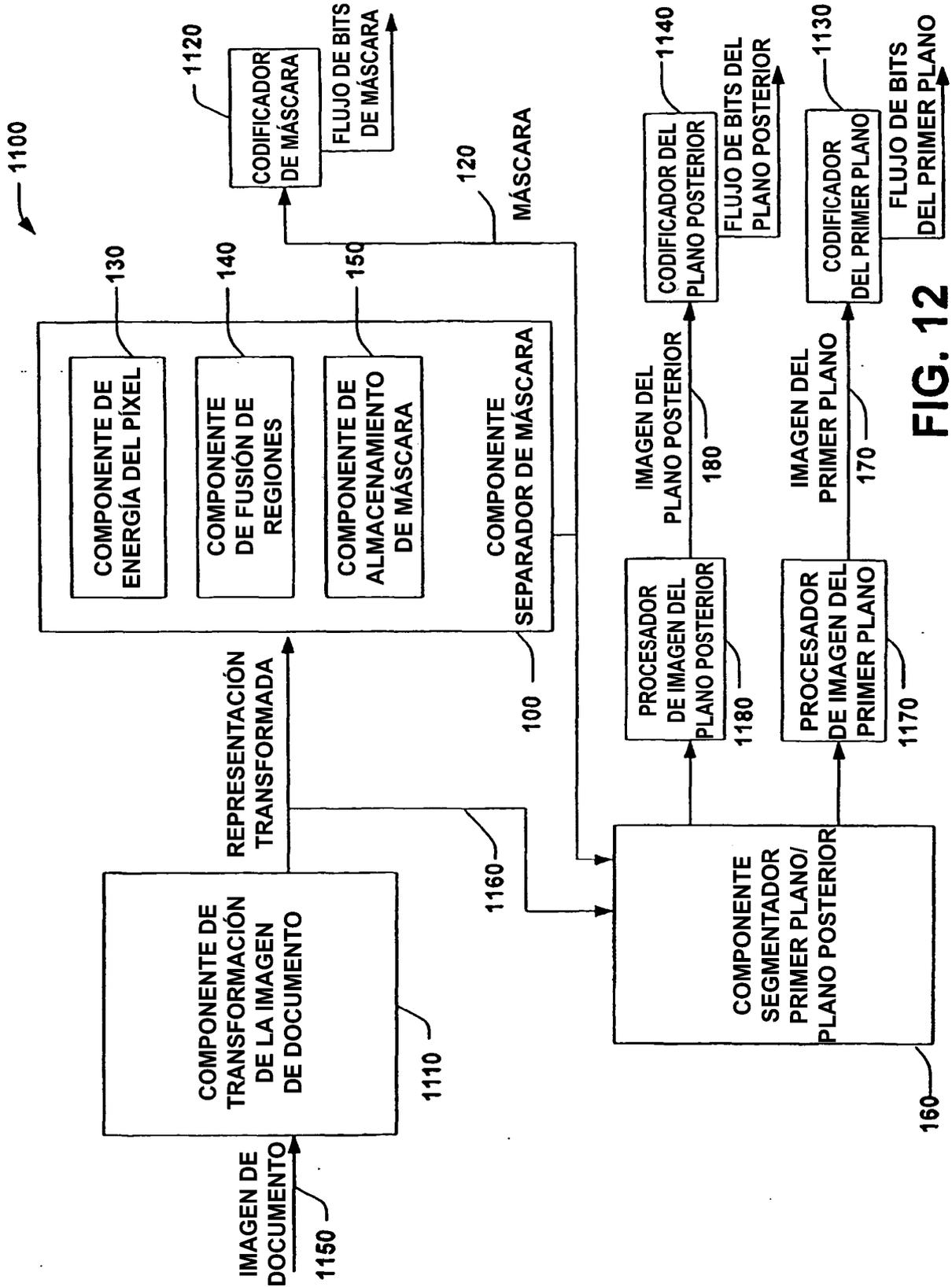


FIG. 12

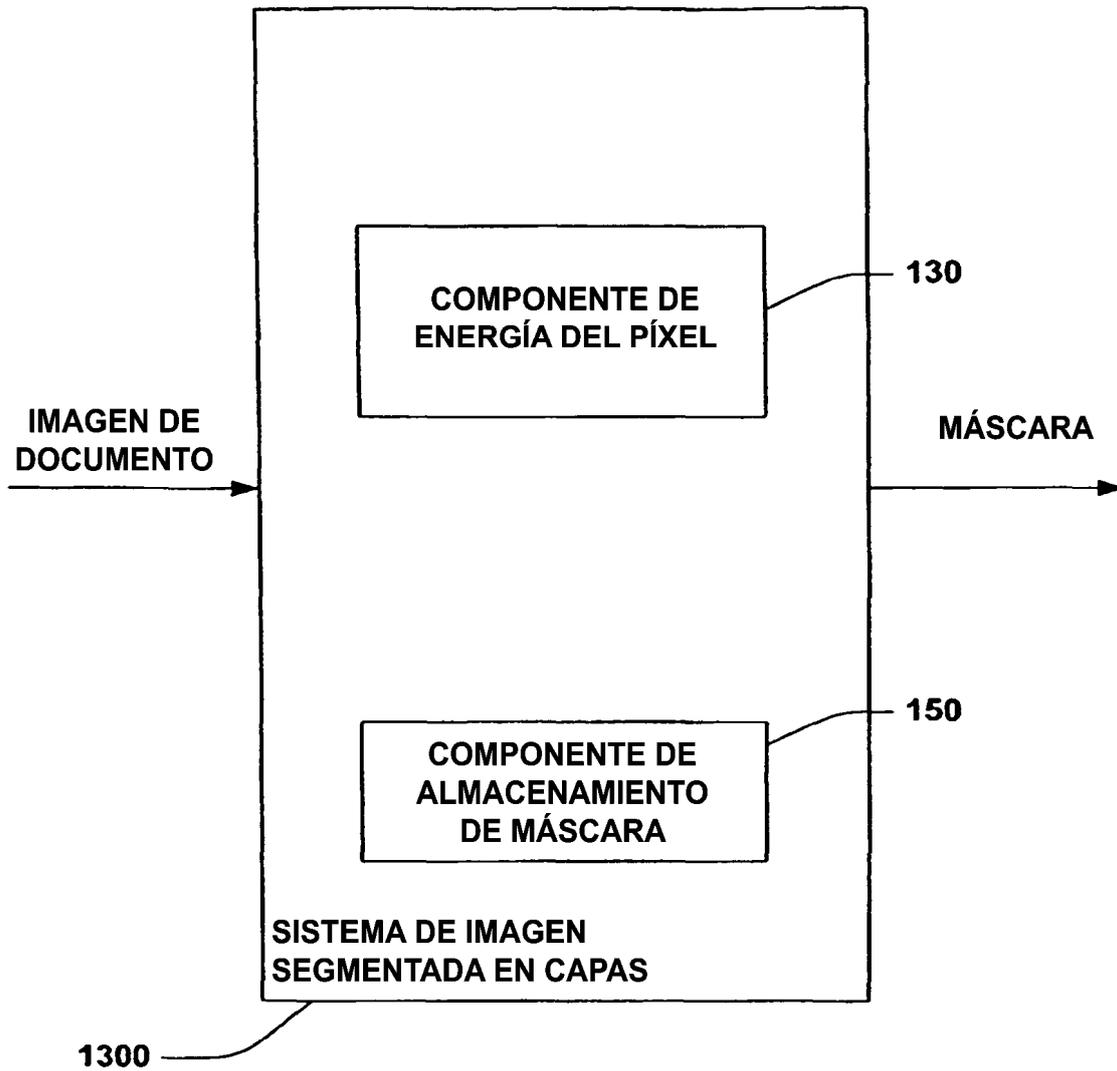


FIG. 13

FIG. 14

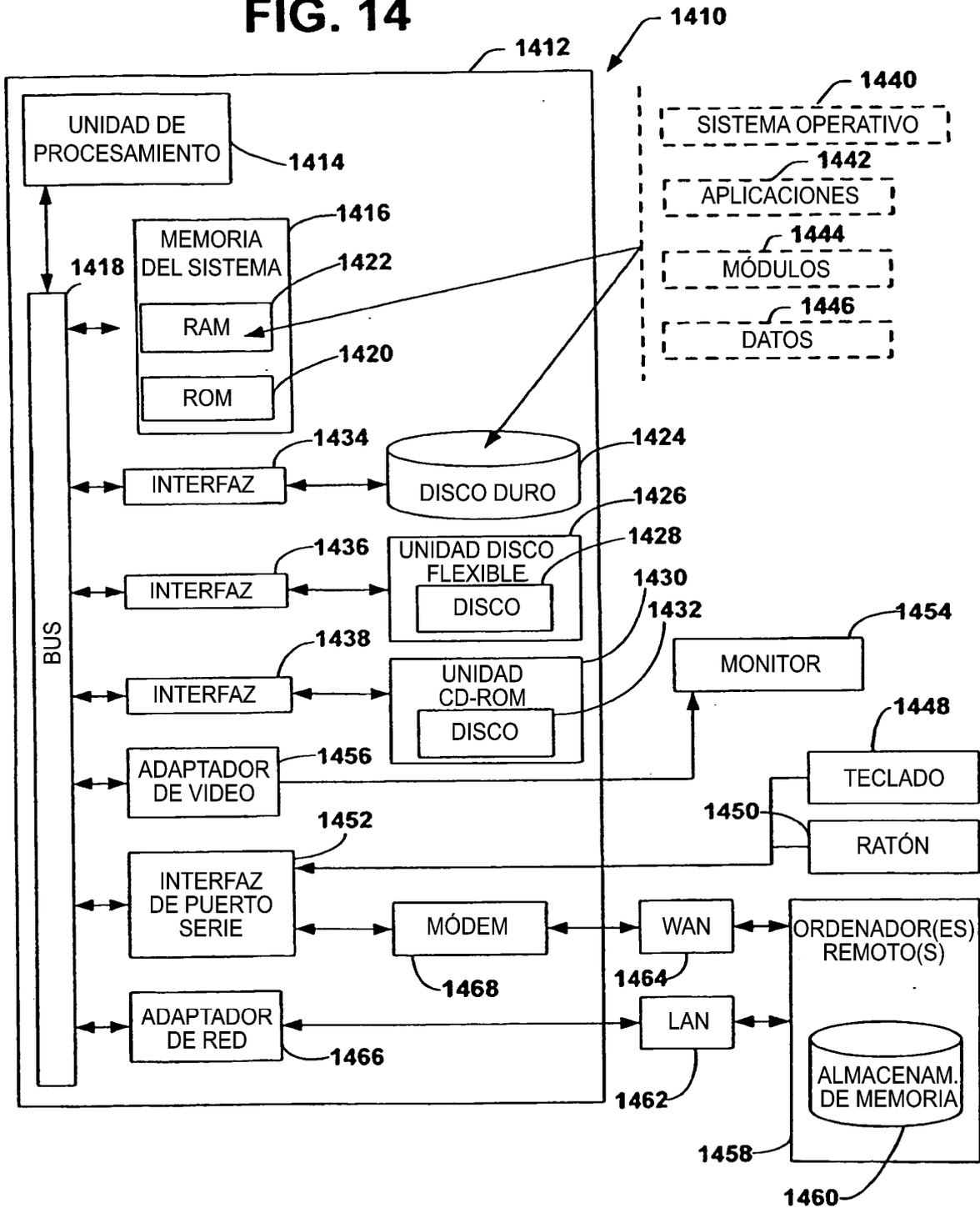


FIG. 15

