



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 361 228**

51 Int. Cl.:
G06T 19/00 (2006.01)
G06T 15/00 (2006.01)
G06T 17/05 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **08762451 .6**
96 Fecha de presentación : **23.06.2008**
97 Número de publicación de la solicitud: **2174297**
97 Fecha de publicación de la solicitud: **14.04.2010**

54 Título: **Renderización no-fotorrealista de una realidad aumentada.**

30 Prioridad: **29.06.2007 GB 0712690**

45 Fecha de publicación de la mención BOPI:
15.06.2011

45 Fecha de la publicación del folleto de la patente:
15.06.2011

73 Titular/es: **IMPERIAL INNOVATIONS LIMITED**
Level 12 Electrical and Electronic Engineering
Building Imperial College
London SW7 2AZ, GB

72 Inventor/es: **Yang, Guang-Zhong y**
Lerotic, Mirna

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 361 228 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

Renderización no-fotorrealista de una realidad aumentada.

5 La presente invención se refiere a un método de renderizar (interpretar) imágenes, en particular para proporcionar indicaciones de oclusión, por ejemplo presentaciones visuales de realidad médica aumentada.

10 La realidad aumentada (RA) se está convirtiendo en una herramienta valiosa en las operaciones quirúrgicas. El hecho de proporcionar datos pre-operativos grabados en tiempo real durante una operación quirúrgica elimina la necesidad de referirse a imágenes fuera de línea y ayuda a la grabación de éstas para el tejido real. La visualización de los objetos de interés resulta accesible a través de la visión de “penetración” que proporciona la RA.

15 En años recientes, están siendo utilizados de manera creciente robots médicos en Cirugía Mínimamente Agresiva (CMA). Con CMA robótica asistida, se mejora la destreza mediante muñecas mecánicas controladas por microprocesador, permitiendo la graduación del movimiento para reducir bastos movimientos manuales y el comportamiento de tareas de micro-escala que no son de otro modo posibles.

20 El establecimiento operacional único del robot quirúrgico proporciona una plataforma ideal para mejorar el campo visual con imágenes pre-operativas/intra-operativas o gráficos generados por ordenador. La efectividad y el beneficio clínico de la RA han sido bien reconocidos en neurocirugía y cirugía ortopédica. Su aplicación a cirugía cardiotorácica o gastrointestinal, sin embargo, permanece limitada, ya que la complejidad de la deformación del tejido impone retos significativos a la presentación de la RA.

25 La síntesis sin juntura de la RA depende de cierto número de factores relativos al modo en el que aparecen los objetos visuales e interactúan visualmente con una escena real. Uno de los mayores problemas de la RA es la manipulación correcta de la oclusión. Aunque la manipulación de oclusión parcial del entorno virtual y real puede ser conseguida mediante reconstrucción exacta en 3D de la escena quirúrgica, particularmente con el advenimiento de técnicas recientes para recuperación de deformación de tejido en 3D en tiempo real, la mayoría de las aplicaciones quirúrgicas de RA implican la superposición de estructuras anatómicas detrás de la superficie de tejido expuesta. Esto es importante, por ejemplo, para bypass coronario, para el cual la visualización anatómica y funcional mejorada permite más exacta navegación intra-operativa y extirpación de vasos. En prostatectomía, la visualización en 3D de la anatomía circundante puede dar lugar a preservación mejorada del haz neurovascular e índices de continencia y potencia mejorados.

35 Aunque proporciona una referencia útil en plano en entornos de visión estéreo, la tradicionalmente superpuesta RA adolece de percepción inexacta en profundidad. Incluso si el objeto es renderizado (interpretado) a la profundidad correcta, el cerebro percibe el objeto como flotando sobre la superficie (Véase, por ejemplo, Johnson LG, et al, La transparencia superficial hace impredecible las superposiciones estéreo: las implicaciones para realidad aumentada, *Studies in Health Technology and Informatics* 2003, 94:131-6; y Swan JE, et al, Juicios de Profundidad Egocéntricos en Óptica, Realidad Aumentada de Penetración, *IEEE Transactions on Visualization and Computer Graphics* 2007, 13(3):429-42). Esta impresión es mejorada por una renderización (interpretación) estilizada; véase por ejemplo Ficher et al, Realidad Aumentada Estilizada para Inmersión Mejorada, Proc. of VR 2005, IEEE, Bonn, Alemania, 12-16-3-2005, p. 195-202, 325, xp010836683.

45 Para objetos que se han de percibir como embebidos en el tejido, nuestros cerebros esperan algún grado de oclusión. Para enfrentarse al problema de la percepción profunda en RA, han sido desarrolladas técnicas de renderización (interpretación) y estrategias de presentación para permitir la percepción exacta de profundidad en 3D de las estructuras virtuales con respecto a la superficie de tejido expuesta. En Sielhorst T, et al, Percepción de Profundidad – Una Cuestión Importante en RA Médica: Estudio de Evaluación por Veinte Cirujanos, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006* 2006:364-72, ha sido estudiado el tejido de percepción en profundidad en RA médica. De acuerdo con las dos referencias citadas anteriormente, se encontró que la percepción de profundidad es pobre si el objeto de RA se renderiza (interpreta) de manera opaca cuando parece flotar sobre la superficie exterior del cuerpo incluso aunque se renderiza (interpreta) a la profundidad correcta detrás de ella. Se identificaron dos modos de mejorar la percepción de profundidad: renderizar (interpretar) tanto la superficie del cuerpo como el objeto de RA como transparentes o renderizar (interpretar) la superficie del cuerpo con una ventana definida dentro de él de tal manera que la ventana proporciona una pista de oclusión por la que el objeto de RA pueda ser visto dentro de la ventana, pero esté de otro modo ocluido por la superficie del cuerpo. Con respecto al primer enfoque [renderización (interpretación) transparente], aunque esto puede dar lugar a una percepción de profundidad mejorada para algunas superficies, en general renderizar (interpretar) dos superficies transparentes superpuestas da lugar a indicaciones visuales conflictivas desde la oclusión de tal manera que es pobre la percepción de profundidad (Véase, por ejemplo, Johnson et al citado anteriormente). El último enfoque (renderización (interpretación) de una ventana) tiene la desventaja de que se pierde toda la información acerca de la superficie del cuerpo dentro de la ventana.

En Ventana Virtual para Percepción de Profundidad Mejorada en RA Médica: C. Bichmeier, N. Navab, Taller Internacional sobre entornos para Formación de Imágenes Médicas y Cirugía ayudada por Ordenador (AMI-ARCS 2006), Copenhague, Dinamarca, octubre de 2006 (disponible en línea en <http://ar.in.tum.de/pub/bichmeier2006window/bichmeier2006window.pdf>) han sido estudiados varios enfoques para mejorar la percepción de profundidad obtenida con el enfoque de ventana mientras se mantiene información acerca de la superficie del cuerpo dentro de la ventana. Han sido considerados los siguientes enfoques: adaptar la forma de ventana a la forma de la superficie del cuerpo, renderizar (interpretar) la superficie de la ventana parecida a vidrio utilizando efectos de luz máxima debido a una fuente de luz virtual, mapear el plano de ventana con una textura estructurada única, simular una estructura finitamente dimensionada para la ventana y fijar el fondo de los objetos de RA ya sea transparentes u opacos. Una desventaja de todos los enfoques, excepto el último, es que debe ser conocido un modelo de 3D de la superficie del cuerpo para que el contorno o superficie de la ventana se pueda renderizar (interpretar) correspondientemente. Un tal modelo de 3D puede ser difícil de obtener de manera fiable, en particular si la superficie del cuerpo se está deformando o cambiando de otros modos durante la formación de imagen.

De acuerdo con un aspecto de la invención, se proporciona un método de renderizar (interpretar) una imagen digital capturada como se define en la reivindicación 1.

Ventajosamente, estableciendo valores de transparencia de una imagen correspondiente renderizada (interpretada) de manera no-fotorrealista (NPR: non-photorealistically rendered) basada en la propia imagen capturada, es posible definir una ventana parcialmente transparente que conserve algo de la estructura de la imagen para proporcionar indicaciones de oclusión tales que hagan posible la percepción de profundidad [en el caso de una renderización (interpretación) bidimensional de la imagen] o ayudada [en el caso de renderización (interpretación) tridimensional de la imagen]. Esta solución no requiere un modelo de la escena subyacente a la imagen según está basado en los propios datos de imagen.

Como se ha mencionado anteriormente, la imagen NPR puede ser renderizada (interpretada) como una vista bidimensional o una segunda imagen puede ser capturada para definir una vista estereoscópica. En cualquier caso, puede ponerse detrás de la imagen capturada un objeto virtual (por ejemplo la RA).

La asignación de los valores de transparencia se puede hacer dentro de una ventana de tal manera que la imagen NPR capturada permanezca opaca fuera de la ventana, ocluyendo el objeto cuando no se ve a través de la ventana. Para una apariencia más natural de la escena o para ayudar a fundir las dos ventanas en una vista estereoscópica, los valores de transparencia pueden ser mezclados gradualmente desde dentro de la ventana al exterior de la misma. La posición de la ventana puede ser definida con dependencia de la mirada del observador, siguiendo continuamente la mirada del observador o actualizando la posición o actualizando la posición sólo cuando se reciba una petición de actualización del observador. Naturalmente, en caso de visión estereoscópica, las ventanas en las dos imágenes (derecha e izquierda) de NPR pueden ser desplazadas en una magnitud determinada por las posiciones de la cámara y los parámetros de acuerdo con la vista estereoscópica.

Los valores de transparencia pueden ser determinados como una función de un gradiente de intensidad de imagen normalizado en lugares correspondientes de la imagen capturada. El cálculo del gradiente de intensidad de imagen puede incluir la determinación de una derivada parcial con respecto a una coordenada de imagen dividida por la intensidad de imagen en los correspondientes lugares.

La determinación de la transparencia de imagen de NPR como se ha expuesto anteriormente, puede ser vista como un ejemplo de un método de fijar los valores de transparencia definiendo un mapa de prioridad o saliencia (saliency map) para un área de la imagen NPR y asignando valores de transparencia como una función de valores de respectivos lugares correspondientes en el mismo mapa de prioridad. El mapa de prioridad puede estar dispuesto para capturar características sobresalientes de la imagen, por ejemplo características que son sobresalientes debido a que sobresalen del fondo de la escena subyacente o debido al contraste de color y/o intensidad. En particular, el mapa de prioridad puede ser definido como una función de inclinaciones locales en la escena subyacente a la imagen, por ejemplo según es estimada sobre la base de sombreado de la imagen. Las inclinaciones locales pueden ser estimadas como una función de respectivos gradientes de intensidad normalizados en la imagen.

Además de asignar valores de transparencia para conservar características sobresalientes como más o menos opacas y hacer el fondo dentro de una zona o ventana más o menos transparente, el mapa de prioridad puede ser usado también para asignar un valor de color a un píxel de la imagen NPR, por ejemplo usando una escala de colores. En particular, los valores de transparencia y de color pueden ser asignados de tal manera que un objeto renderizado (interpretado) detrás de la imagen NPR está siendo percibido como visto a través de la zona transparente (ventana) mientras está siendo ocluido por píxeles dentro de la zona que tiene elevados valores en el mapa de prioridad.

En una aplicación, el objeto virtual puede ser derivado de datos médicos de formación de imagen, por ejemplo imágenes de CT o MRI de un tumor. En particular, las imágenes pueden ser capturadas usando un endoscopio estereoscópico, por ejemplo durante cirugía torácica de ojo de cerradura. Sin embargo, se comprenderá que el método

de renderización (interpretación) descrito anteriormente no está limitado a aplicaciones médicas de RA, sino que es más generalmente aplicable a aplicaciones de RA en las que el objeto virtual está fuera de la vista normal del observador detrás de una escena capturada.

5 En un aspecto más de la invención, se proporciona un sistema para renderizar (interpretar) una imagen digital según se reivindica en la reivindicación 21.

Todavía en un aspecto más de la invención se proporciona una consola de cirugía robótica según se define en la reivindicación 41.

10 Otros aspectos de la invención se extienden a un programa de ordenador según se define en la reivindicación 42.

Para evitar dudas, la expresión imagen NPR (abreviatura de Non-Photorealisticaly Rendered imagen) se usa en esta memoria para designar la imagen digital capturada y tratada, aplicada por ejemplo como una textura a un plano correspondiente a un plano de imagen en un modelo de gráfico de ordenador en 3D que puede contener también el objeto de RA. Naturalmente, este modelo puede ser renderizado (interpretado) como una imagen en 2D o endoscópica en 3D.

20 A continuación se describirán realizaciones de la invención a modo de ejemplos solamente y haciendo referencia a los dibujos que se acompañan, en los cuales:

La figura 1 es un diagrama de flujo de un método para renderización (interpretación) no-fotorrealista de al menos parte de una imagen, por ejemplo para revelar un objeto renderizado (interpretado) detrás de la escena capturada en la imagen;

25 La figura 2 es un diagrama de flujo de un algoritmo para la renderización (interpretación) de una imagen NPR correspondiente y un objeto de RA en 2D ó 3D;

La figura 3 representa un sistema correspondiente;

La figura 4 representa un función de enmascaramiento utilizada en el tratamiento; y

30 La figura 5 representa un ejemplo de una vista de RA renderizada (interpretada) usando el método descrito.

La idea subyacente para el presente método de renderizar (interpretar) una imagen como aplicado a presentaciones visuales médicas de RA es renderizar (interpretar) una superficie anatómica expuesta como una capa traslúcida mientras se conservan detalles suficientes para ayudar en la navegación e indicación de profundidad. Una realización está basada en espacio-pq basado en renderización (interpretación) No-Fotorrealista (NPR) para proporcionar una visión de penetración del objeto virtual embebido mientras se mantienen detalles anatómicos sobresalientes de la superficie anatómica expuesta. A este fin, la geometría de superficie basada en representación de espacio-pq es primeramente derivada, donde p y q representan la pendiente de la superficie de la imagen a los largo de los ejes x, y, respectivamente. Por ejemplo, esto se puede conseguir con estéreo fotométrico introduciendo múltiples condiciones de iluminación. Para deformar tejido, sin embargo, el problema está mal planteado y no es posible la introducción de múltiples fuentes de luz en una disposición endoscópica. Sin embargo, el problema puede ser simplificado para casos en los que tanto la cámara como una fuente de luz estén cerca de la superficie en la que se forma imagen (véase Rashid HU, Burger P. "Algoritmo diferencial para la determinación de la forma a partir del sombreado usando una fuente de luz puntual". Image and Vision Computing 1992; 10(2):119 – 27), tal como broncoscopios y endoscopios. En tales casos, el valor de la intensidad de imagen en coordenadas x, y para una fuente de luz puntual próxima está dada por

$$E(x, y) = \frac{s_0 \rho(x, y) \cos \theta}{r^2} \quad (1)$$

50 donde s_0 es la constante de intensidad de la fuente de luz, $\rho(x, y)$ es el albedo o coeficiente de reflexión, r es la distancia entre la fuente de luz y el punto de la superficie (x, y, z) , y θ es el ángulo entre el rayo de luz incidente y la normal a la superficie \tilde{n} . En espacio de gradiente, el vector normal a la superficie es igual a

$$\tilde{n} = \frac{(p, q, -1)}{\sqrt{1 + p^2 + q^2}} \quad (2)$$

55 donde p y q representan inclinaciones de la superficie en direcciones x e y, respectivamente. Para una superficie Lambertiana lisa en la escena, la intensidad de imagen dada por la Ecuación 1 puede ser reducida a

$$E(x, y) = s_0 \rho_{promedio} \frac{(1 - p_0 x - q_0 y)^3}{Z_0^2 (1 - p_0 x_0 - q_0 y_0)^2 (1 + p_0^2 + q_0^2)^{1/2} (1 + x^2 + y^2)^{3/2}} \quad (3)$$

que define la relación entre la intensidad de imagen $E(x, y)$ en el punto (x, y) y el flujo radiante de la escena en el correspondiente punto de superficie $(x_0 Z_0, y_0 Z_0, Z_0)$ con una superficie normal $(p_0, q_0, -1)$, donde $\rho_{promedio}$ indica el albedo promedio en una pequeña proximidad de la superficie y s_0 es la constante de intensidad de la fuente de luz puntual. La superficie Lambertiana bajo iluminación de fuente puntual es un material de superficie idealizada que satisface dos condiciones: (1) parece igualmente brillante desde todas las direcciones de visión, y (2) refleja toda la luz incidente.

Utilizando derivadas parciales de la intensidad de imagen en la Ecuación 3, las derivadas parciales normalizadas en x/y en el lugar de la imagen (x, y) , $R_{x/y}$, pueden ser expresadas en términos sólo de coordenadas de imagen e inclinaciones locales:

$$R_x = \frac{1}{E} \frac{\partial E}{\partial x} = -3 \left\{ \frac{p_0}{(1 - p_0 x - q_0 y)} + \frac{x}{(1 + x^2 + y^2)} \right\}$$

$$R_y = \frac{1}{E} \frac{\partial E}{\partial y} = -3 \left\{ \frac{q_0}{(1 - p_0 x - q_0 y)} + \frac{y}{(1 + x^2 + y^2)} \right\} \quad (4)$$

Esto se puede expresar nuevamente como dos ecuaciones lineales en p_0 y q_0 (siendo las derivadas parciales normalizadas o gradiente normalizado $R_{x/y}$, en el lugar (x, y) de la imagen determinable a partir de la intensidad de la imagen en (x, y) y su proximidad):

$$\begin{aligned} A_1 p_0 + B_1 q_0 + C_1 &= 0 \\ A_2 p_0 + B_2 q_0 + C_2 &= 0 \end{aligned} \quad (5)$$

con

$$\begin{aligned} A_1 &= (-x_0 \cdot R_x + 3) \cdot (1 + x_0^2 + y_0^2) - 3 \cdot x_0^2 \\ B_1 &= -R_x (1 + x_0^2 + y_0^2) \cdot y_0 - 3 \cdot x_0 \cdot y_0 \\ C_1 &= R_x \cdot (1 + x_0^2 + y_0^2) + 3 \cdot x_0 \\ A_2 &= -R_y \cdot (1 + x_0^2 + y_0^2) \cdot x_0 - 3 \cdot x_0 \cdot y_0 \\ B_2 &= (-y_0 \cdot R_y + 3) \cdot (1 + x_0^2 + y_0^2) - 3 \cdot y_0^2 \\ C_2 &= R_y \cdot (1 + x_0^2 + y_0^2) + 3 \cdot y_0 \end{aligned} \quad (6)$$

que da las siguientes expresiones para p_0 y q_0 en cada punto (x, y) de la imagen:

$$p_0 = \frac{B_2 C_1 - B_1 C_2}{B_1 A_2 - A_1 B_2} \quad (7)$$

$$q_0 = \frac{-A_2 C_1 + A_1 C_2}{B_1 A_2 - A_1 B_2} \quad (8)$$

Los valores de p y q de la superficie con imagen captan detalles en 3D de la estructura anatómica expuesta y se usan para acentuar características sobresalientes (que son características que sobresalen de la superficie y por tanto que tienen un elevado gradiente) mientras hacen semi-transparente la superficie de fondo suavemente variable. Se

entenderá que los valores de p y q pueden ser calculados utilizando cualquier técnica apropiada. Para crear las indicaciones visuales deseadas, las superficies de la escena que son paralelas al plano de visión (bajas p , q) son renderizadas (interpretadas) como más o menos transparentes mientras las estructuras inclinadas (altos p y q) son renderizadas (interpretadas) más o menos opacas. Una medida de la inclinación de la superficie es generada a partir de valores- pq para cada punto de imagen (x, y) por

$$S(x, y) = \log(\text{abs}(p_0) + \text{abs}(q_0) + 1) \quad (9)$$

En donde valores altos de $S(x, y)$ corresponden a gradientes grandes. En efecto, esto proporciona un mapa de prioridad o imagen *sobresaliente*. El algoritmo agrupa valores elevados de p y q para limitar el intervalo dinámico para fines de presentación visual. Un mapa B de fondo liso es creado aplicando un amplio filtro Gaussiano al mapa de prioridad S , con lo que se suavizan las elevadas variaciones de frecuencia en la imagen propensas a representar ruido o variaciones menos importantes de superficie en lugar de características sobresalientes de "verdad". Las imágenes sobresalientes y de fondo se combinan usando una máscara de tal manera que valores de píxeles bajos de S son sustituidos con el valor de B en el píxel correspondiente de (x, y) , como se describe con detalle en lo que sigue.

Pasando a la aplicación práctica del mapa de prioridad anterior, con referencia a la figura 1, en un método 2 de producir una renderización (interpretación) no-realista de una imagen capturada (o una zona de la misma) para uso como una textura proyectada sobre un plano de imagen en un modelo de gráfico de ordenador, en el paso 4 se define una región de interés (ROI: región of interest) para tratamiento de NPR. Esto puede incluir la imagen completa o una subzona de la misma. En el paso 6 la región de interés es previamente tratada incluyendo convertir los valores de color de píxel (si está en color) a la escala del gris y aplicando una función de alisamiento suave tal como una Gaussiana de 3x3 píxeles. En el paso 8, el mapa de prioridad es calculado a partir de la imagen capturada en la ROI como se ha descrito anteriormente usando cualquier método conocido para calcular las derivadas parciales, por ejemplo diferenciando simplemente los valores de píxel del píxel en cuestión con un píxel vecino en la dirección pertinente (por ejemplo, x). El mapa es calculado a partir de la derivada parcial y de la intensidad de imagen (por ejemplo, valor de la escala del gris) en cada píxel o situación en la ROI.

En el paso 10, el mapa de prioridad es desprovisto de ruido combinándolo con el mapa de fondo liso B (por ejemplo B puede ser derivada de S aplicando un filtro Gaussiano amplio de 6x6 píxeles con una extensión de 7 píxeles). S y B se combinan de acuerdo con una función de enmascaramiento como

$$S(x, y) = \text{máscara}(x, y) \cdot S(x, y) + (1 - \text{máscara}(x, y)) \cdot B(x, y) \quad (10)$$

de tal manera que el mapa de prioridad es mezclado con el mapa de fondo, por lo que S domina donde S es alto y B domina donde S es bajo. La función de enmascaramiento puede ser definida usando estrías con unos pocos puntos de control (por ejemplo una estría Catmull Rom). En la figura 5 está representada una función de enmascaramiento apropiada como una función de $S(x, y)$. Naturalmente, se pueden utilizar también otras funciones de enmascaramiento, por ejemplo una función de etapa con un valor de umbral apropiado para $S(x, y)$. Otro ejemplo podría ser una función polinómica apropiada.

En el paso 12, a los píxeles dentro de la región de interés desprovista de ruido les son asignados valores de color de acuerdo con el mapa de prioridad $S(x, y)$ usando una escala de colores. Una escala de colores apropiada puede estar comprendida en el intervalo del negro para el valor mínimo de S , y a través de azul a blanco para el valor máximo. Estos colores artificiales se aplican a la imagen NPR de acuerdo con una función de ventana, por ejemplo una función de ventana radial $f(r)$ de tal manera que el color artificial se aplica dentro de la ventana y el color original de la imagen permanece fuera de la ventana. Un ejemplo de $f(r)$ es una función de etapa, en cuyo caso la ventana tendría un borde agudo que cambia discretamente del color artificial al color de imagen. Para conseguir una transición suave, $f(r)$ define una región de transición en otro ejemplo

$$f(r) = \frac{r^2}{r_0^2} \quad \text{para } r < r_0, \quad f(r) = 1 \quad \text{para } r \geq r_0 \quad (11)$$

donde $r^2 = (x - x_{\text{central}})^2 + (y - y_{\text{central}})^2$ y r_0 determina el tamaño de ventana. Definiendo una transición suave, se puede ayudar, si es aplicable, la fusión de las dos (izquierda y derecha) ventanas en una imagen estéreo (véase más abajo).

En el paso 14, la transparencia de píxel de imagen NPR dentro de la misma ventana es fijada de acuerdo con $S(x, y)$ y la función de ventana $f(r)$ de tal manera que el valor más bajo de S corresponde a la completa (o casi completa, por ejemplo 95%) transparencia y el valor máximo de S corresponde a la completa (o casi completa, por ejemplo 95%) opacidad y los valores de transparencia se mezclan con el resto de la imagen NPR usando $f(r)$, por ejemplo.

Para evitar dudas, se puede ver un ejemplo de la operación de mezcladura usando $f(r)$ como mezcladura de la Imagen Renderizada (interpretada) No-Fotorrealista (NPI) tratada con la Imagen Capturada (CI) para llegar a la imagen NPR como sigue:

$$5 \quad \text{Imagen NPR} = f(r) \cdot \text{CI} + (1 - f(r)) \cdot \text{NPI} \quad (12)$$

10 La situación de la ventana (es decir, el origen de $f(r)$ en el anterior ejemplo ($x_{central}$, $y_{central}$)) se puede fijar de numerosos modos. Por ejemplo, en algunas realizaciones la posición de la ventana es previamente definida basándose en la posición del objeto de RA que se ha de renderizar (interpretar) detrás del plano de imagen, en particular si existe sólo un objeto único que se ha de renderizar (interpretar) dentro del campo de vista. Alternativamente, en algunas realizaciones la posición de la ventana puede ser determinada sobre la base de una mirada del observador (por ejemplo detectada utilizando un dispositivo seguidor de los ojos), ya sea actualizada continuamente o de acuerdo con una fijación del observador cuando el observador emite una petición para que la posición de ventana sea actualizada (o, en algunas realizaciones, si no es actualmente presentada ventana, la ventana es presentada de acuerdo con las fijaciones del observador cuando es activada la ventana). Esta presentación dinámica de ventana puede ser particularmente útil si está siendo renderizada (interpretada) una escena completa de RA en lugar de un objeto único.

20 Con referencia a las figuras 2 y 3, se describe a continuación un sistema y un método para presentar una imagen capturada de una escena junto con un objeto de realidad aumentada. En el paso 16, es capturada una imagen de la escena por un dispositivo 22 de formación de imagen y transmitida a una unidad de tratamiento 24. La imagen capturada es tratada como una textura proyectada sobre un objeto del plano de la escena que se ha de presentar en un modelo de gráfico de ordenador en algunas realizaciones. El modelo de gráfico de ordenador puede ser ejecutado en la práctica de varios modos, por ejemplo utilizando la librería de gráficos de OpenGL en un programa de C++, en algunas realizaciones.

30 La perspectiva correcta (y la disparidad en el caso de renderización (interpretación) en 3D) está ya contenida en la imagen capturada (imagen capturada izquierda y derecha en el caso de un dispositivo de formación de imagen con cámara estéreo), pero la información de profundidad acerca del objeto de realidad aumentada en el modelo de gráfico de ordenador es importante para manejar compatiblemente la oclusión.

35 Para cámaras y presentaciones visuales en estéreo, la posición del objeto de realidad aumentada en la visa izquierda y derecha tiene que adaptar la disparidad en las imágenes capturadas. Esto se consigue por medio de calibración de cámara estéreo que proporciona transformaciones necesarias entre las vistas. En la escena combinada, las imágenes capturadas son presentadas en la misma posición para ambas vistas, izquierda y derecha, puesto que contienen ya la disparidad, mientras que el objeto de realidad aumentada es puesto en diferentes posiciones para las vistas izquierda y derecha de manera que se adapte la disparidad de las imágenes capturadas. Las transformaciones obtenidas como parte de una calibración de cámara estéreo son utilizadas usando transformaciones de OpenGL en algunas realizaciones.

40 En algunas ejecuciones prácticas ejemplares, se abre una ventana de OpenGL para cada presentación (izquierda y derecha) y la imagen NPR es presentada (por ejemplo, aplicada a un objeto en el plano de escena que actúa como una pantalla de proyección) en cada ventana a una profundidad apropiada (véase más adelante). El objeto de RA es entonces renderizado (interpretado) para cada ventana usando la calibración de cámara para la respectiva cámara (izquierda y derecha). Similarmente, la posición relativa de la ROI y/o ventana en cada una de las ventanas de OpenGL se fija usando los datos de calibración de cámara para que sea compatible con la vista estereoscópica, en algunas realizaciones. A este respecto, ventanas suavemente mezcladas ayudarán en la fusión de dos ventanas en una vista estereoscópica incluso si el desplazamiento de la ventana no está exactamente de acuerdo con la vista estereoscópica. Por ejemplo, una aproximación apropiada para una cámara de endoscopio estereoscópico es usar el desplazamiento de cámara como el desplazamiento de ventana en algunas ejecuciones prácticas.

55 Para asegurar que cada píxel sea correctamente coloreado, la coordenada z de la imagen NPR (el plano del objeto de escena) debe estar más próximo a la posición de la cámara en el modelo de gráfico de ordenador que la posición del objeto de RA en la estructura de cámara de tal manera que los píxeles adopten los valores de color del objeto de escena donde el objeto de escena ocluye el objeto de RA detrás del mismo (o donde partes de baja transparencia del objeto de escena dominan la coloración de los píxeles correspondientes). Puesto que las indicaciones de oclusión proporcionadas por el tratamiento no-fotorrealista explicado anteriormente son sólo necesarias cuando el objeto de RA se sitúa detrás de la superficie capturada por el dispositivo 22 de formación de imagen (de otro modo el objeto real sería visible en la imagen capturada), se supone para renderización (interpretación) en algunas realizaciones que el objeto de escena está siempre delante del objeto de RA y, por ejemplo, el plano frontal del tronco de visión del modelo de gráfico de ordenador se utiliza como el objeto de escena sobre el que se proyecta(n) la(s) imagen(es) capturada(s) en algunas ejecuciones prácticas. Esto asegura la correcta manipulación de la oclusión sin influenciar la perspectiva o disparidad de la(s) imagen(es) capturada(s) que son definidas por el posicionamiento real de la cámara.

Alternativamente, en algunas ejecuciones prácticas la imagen NPR se sitúa en o cerca del plano focal de la cámara o la profundidad del plano de escena podría ser recuperado usando el descubrimiento de intervalo para fijar la profundidad de la imagen NPR (véase, por ejemplo, el documento de Stoyanov D., et al, *Cirugía Asistida por Ordenador*, julio de 2005, 10(4): 199-208, incorporado a esta memoria como referencia). Técnicas que son utilizadas para recuperación de profundidad en algunas realizaciones incluyen el uso de la vergencia del observador a partir del seguimiento del ojo (Mylonas GP et al, *Procedimientos del segundo Taller International en Formación de Imágenes Médicas y Realidad Aumentada*, MIAR (2004), Pekín, 311-319), recuperación de profundidad desde estéreo tal como técnicas de dispersión, forma a partir de sombreado o combinación de ambos (Stoyanov D et al, *MICCAI (2) 2004*: 41-48) o el uso de luz estructurada (Koninckx TP y Van Gool L, *IEEE PAMI*, vol 28, no. 3, 2006), marcadores de referencia o descubrimiento de intervalo de láser (Mitsuhiro H et al, *Análisis de Imagen Médica 10(2006)* 509-519).

En el paso 18 se determina la posición del objeto de RA virtual en el modelo de gráfico de ordenador. Una entrada de objeto de RA proporciona datos de coordenadas del objeto de RA obtenidos a partir de la formación de imagen médica, por ejemplo a partir de datos de MRI o CT. Estos datos han de ser expresados en coordenadas en una estructura de referencia fija en el cuerpo de un paciente antes de que sean obtenidas las imágenes para convertir los datos de imagen desde el dispositivo de formación de imagen médica a una estructura del cuerpo de referencia.

Para presentación exacta del objeto de RA junto con la imagen de escena capturada, las coordenadas de objetos de RA de la estructura del cuerpo de referencia definida por los indicadores de referencia necesitan ser transformadas en estructura de cámara del dispositivo de formación de imagen de referencia a partir de la entrada 28 de estructura de referencia que sigue tanto la posición de la cámara como la orientación, así como los indicadores de referencia, por ejemplo siguiendo los indicadores de referencia usando un dispositivo de seguimiento de 3 dimensiones para determinar la posición y la orientación de la estructura de referencia del cuerpo con relación a la estructura de referencia de la cámara, como es bien sabido en la técnica.

Si se conoce la posición de la cámara (por ejemplo en el caso de una consola de cirugía robótica, en la que la colocación de la cámara puede ser derivada de la posición del brazo robótico que soporta la cámara), esta posición conocida en una estructura de coordenadas fijada al recinto de operación se usa juntamente con los datos del seguidor para realizar la necesaria transformación de coordenadas en algunas ejecuciones prácticas particularmente relevantes para aplicaciones en una consola de cirugía robótica. En otras disposiciones en las que las cámaras son más móviles (tales como en una disposición de visión de penetración de vídeo montado en la cabeza) es necesario que la posición de la cámara sea seguida para poder realizar la requerida transformación de coordenadas. Detalles de las mediciones y transformación requeridas se describen en el documento de Vorgt S. et al, *Aumento de Realidad para Operaciones Médicas: Estructura del Sistema, Seguimiento de Indicador de Cámara Única y Evaluación del Sistema*, *International Journal of Computer Vision* 2006, 70(2): 179-190.

En el paso 2 (se ha de entender que el orden de pasos no está limitado al mostrado en la figura 2, sino que los pasos pueden ser realizados en cualquier orden, sometido naturalmente a la limitación de que una imagen deba ser capturada antes de que pueda ser tratada), el tratamiento NPR de la imagen capturada es realizado como se ha descrito anteriormente con referencia a la figura 1, posiblemente tomando una entrada acerca de la posición de ventana desde la entrada 30 de posición de ventana que, en algunas realizaciones, incluye un seguidor de ojos e interfaz de usuario para activar o desactivar la ventana y seleccionar la posición de ventana, para la ventana descrita anteriormente.

Una vez que la imagen capturada ha sido tratada y aplicada como una imagen NPR al objeto de escena y ha sido definida la posición en 3D del objeto de RA como se ha descrito anteriormente, el objeto de escena y el objeto de RA son renderizado (interpretado)s para presentar en el dispositivo de presentación 32 en el paso 20. Esto puede ser una vista en 2D del correspondiente modelo de gráfico de ordenador en el que el dispositivo de presentación es un simple monitor o una vista en 3D que consiste en una imagen derecha e izquierda donde el dispositivo de presentación es capaz de estéreo.

En algunas realizaciones, el sistema y el método descritos anteriormente son incorporados en una consola de cirugía robótica mínimamente agresiva, por ejemplo la consola quirúrgica robótica Vinci por Intuitive Surgical, Inc. Mountain View, USA. La consola proporciona manipulación robótica para controlar a distancia herramientas quirúrgicas mínimamente agresivas y realimentación visual en estéreo a través de un dispositivo de presentación en estéreo de posición fija que proporciona respectivas imágenes izquierda y derecha desde un endoscopio estereoscópico para cada ojo del operador. Las imágenes capturadas son tratadas como se ha descrito anteriormente y pueden ser entonces renderizada (interpretada)s conjuntamente con un objeto de RA (por ejemplo, la representación de un tumor).

Las figuras 5 a y b muestran respectivas vistas de ojos izquierdo y derecho de tejidos de pulmón a partir de una lobectomía de pulmón asistida por robótica. Las figuras 5 c y d muestran la misma vista con una superposición de RA transparente de un objeto de RA en el que tanto la escena como el objeto son renderizado (interpretado)s de manera transparente como en la solución de vídeo de vista de penetración de Vogt et al, a que se ha hecho referencia

5 anteriormente. Las figuras 5 e y f representan respectivas vistas izquierda y derecha de la escena de las figuras 5 a y b tratadas usando el método descrito anteriormente junto con el objeto de RA y las figuras g y h muestran las vistas de a y b, y e y f, respectivamente, combinadas usando ventana radial lisa. Como se puede ver claramente, la ventana proporciona una renderización (interpretación) más natural de la escena tridimensional según es visto el objeto de RA a través de la ventana mientras las características preservadas en la ventana mantienen información de referencia para el cirujano y proporcionan también indicaciones de profundidad adicionales por oclusión.

10 Se ha de entender que son posibles muchas modificaciones en las realizaciones descritas anteriormente. Por ejemplo, puede ser presentado más de un objeto y puede ser alterado el orden de las operaciones del método según se ha descrito anteriormente dentro de las limitaciones de ciertos pasos previos. Además, el método anteriormente descrito para la determinación de valores de transparencia que preserva características sobresalientes de una imagen mientras proporciona visión de penetración para un objeto subyacente al mismo tiempo será aplicable a cualquier otro tipo de escenas, objetos y aplicaciones además de las descritas anteriormente.

15 Se entenderá, por supuesto, que aunque se acaban de describir realizaciones particulares, el objeto reivindicado no está limitado en su alcance a una realización o ejecución práctica particular. Por ejemplo, una realización puede ser en hardware, tal como se ejecuta en la práctica para operar sobre un dispositivo o combinación de dispositivos, por ejemplo, mientras otras realizaciones pueden estar en software, o como cualquier combinación de hardware, software y/o firmware, por ejemplo. Análogamente, aunque el objeto reivindicación no está limitado en su alcance a este respecto, una realización puede comprender uno o más artículos, tales como un portador o medio de almacenamiento o soporte de almacenamiento. El soporte de almacenamiento, tal como uno o más CD-ROMs, memoria de estado sólido, disco magneto-óptico y/o cintas o discos magnéticos, por ejemplo, puede tener instrucciones almacenadas en el mismo que, cuando son ejecutadas por un sistema, tal como un sistema de ordenador, plataforma de cálculo u otro sistema, por ejemplo, pueden dar lugar a una realización de un método que esté siendo ejecutado de acuerdo con el objeto reivindicado, tal como una de las realizaciones previamente descritas, por ejemplo. Una realización puede comprender una señal portadora en un medio de telecomunicaciones, por ejemplo una red de telecomunicaciones. Ejemplos de señales portadoras apropiadas incluyen una señal de radio-frecuencia, una señal óptica y/o una señal electrónica.

30 Como un ejemplo potencial, una plataforma de cálculo u ordenador puede incluir una o más unidades de tratamiento o procesadores, uno o más dispositivos de entrada/salida, tales como una pantalla de presentación, un teclado y/o un ratón, y/o una o más memorias, tales como memoria estática de acceso aleatorio, memoria dinámica de acceso aleatorio, memoria flash y/o disco duro.

35 En la descripción precedente han sido descritos varios aspectos del objeto reivindicado. Para fines de explicación, fueron expuestos números concretos, sistemas y/o configuraciones para proporcionar una comprensión completa del objeto reivindicado. Sin embargo, resultará evidente para un experto en la técnica que tiene el beneficio de esta descripción que el objeto reivindicado puede ser puesto en práctica sin los detalles concretos. En otros casos, fueron omitidas y/o simplificadas características bien conocidas para no oscurecer el objeto reivindicado. Aunque ciertas características han sido ilustradas y/o descritas en esta memoria, a los expertos en la técnica se les ocurrirán ahora muchas modificaciones, sustituciones, cambios y/o equivalentes. Por lo tanto, se ha de entender que las reivindicaciones adjuntas pretender cubrir todas las citadas modificaciones y/o cambios que caigan dentro del alcance del objeto reivindicado.

REIVINDICACIONES

- 5 1 Un método de renderizar (interpretar) una imagen digital capturada, captada por una cámara y definida por una pluralidad de píxeles de imagen como una imagen NPR renderizada (interpretada) de manera no-fotorrealista definida por una pluralidad de píxeles de imagen NPR cada uno de los cuales tiene un valor de transparencia NPR asociado, en el que el valor de transparencia de al menos algunos de los píxeles de imagen NPR es determinado con dependencia de correspondientes píxeles de la imagen capturada, incluyendo el método renderizar (interpretar) un objeto virtual detrás de la imagen NPR.
- 10 2. Un método según la reivindicación 1, que incluye renderizar (interpretar) una imagen adicional capturada como una imagen NPR adicional para definir una vista estereoscópica.
- 15 3. Un método según la reivindicación 1 o la 2, que incluye definir una ventana dentro de la imagen NPR y asignar los valores de transparencia dentro de la ventana con dependencia de los correspondientes píxeles de imagen capturada.
4. Un método según la reivindicación 3, que incluye definir una posición de ventana con dependencia con dependencia de la mirada de un usuario.
- 20 5. Un método según una cualquiera de las reivindicaciones precedentes, que incluye determinar los valores de transparencia como una función de un gradiente de intensidad de imagen normalizado en correspondientes lugares en la imagen capturada.
- 25 6. Un método según una cualquiera de las reivindicaciones precedentes, que incluye definir un mapa de prioridad o saliencia para una zona de la imagen NPR y asignar un valor de transparencia a un píxel de la imagen NPR como una función de un valor correspondiente del mapa de prioridad.
- 30 7. Un método según la reivindicación 6, que incluye asignar un valor de color a un píxel de la imagen NPR como una función de una valor correspondiente valor del mapa de prioridad, en el que los valores de transparencia y de color son asignados de tal manera que un objeto renderizado (interpretado) detrás de la imagen NPR está siendo percibido como siendo visto a través de la zona en la que valores del mapa de prioridad son inferiores mientras está siendo ocluido donde el mapa de prioridad tiene valores más altos.
- 35 8. Un sistema para renderizar (interpretar) una imagen digital capturada, captada por una cámara y definida por una pluralidad de píxeles de imagen como una imagen NPR definida por una pluralidad de píxeles de imagen NPR, cada uno de los cuales tiene un valor de transparencia asociado, incluyendo el sistema un calculador de transparencia dispuesto para calcular el valor de transparencia de al menos algunos de los píxeles de imagen NPR con dependencia de correspondientes píxeles de la imagen capturada, siendo además el sistema operable para renderizar (interpretar) un objeto virtual detrás de la imagen NPR.
- 40 9. Un sistema según la reivindicación 8, el cual está dispuesto para renderizar (interpretar) una imagen adicional capturada usando el calculador de transparencia como una imagen NPR adicional para definir una vista estereoscópica.
- 45 10. Un método según la reivindicación 8 o la 9, en el que el calculador de transparencia está dispuesto para asignar los valores de transparencia dentro de una ventana dentro de la imagen NPR.
- 50 11. Un sistema según una cualquiera de las reivindicaciones 8 a 10, que incluye un módulo de formación de ventana que tiene una entrada representativa de la mirada de un usuario para definir una posición de ventana con dependencia de la mirada del usuario.
- 55 12. Un sistema según una cualquiera de las reivindicaciones 8 a 11, en el que el calculador de transparencia está dispuesto para determinar los valores de transparencia como una función de un gradiente de intensidad de imagen normalizado en respectivos lugares correspondientes en la imagen capturada.
- 60 13. Un sistema según una cualquiera de las reivindicaciones 8 a 12, en el que el calculador de transparencia está dispuesto para definir un mapa de prioridad para una zona de la imagen capturada y para asignar un valor de transparencia a un píxel de la imagen NPR como una función de un valor correspondiente del mapa de prioridad.
14. Un sistema según la reivindicación 13, que incluye un calculador de color dispuesto para asignar un valor de color a un píxel de la imagen NPR capturada como una función de un valor correspondiente del mapa de prioridad, en el que el calculador de transparencia y el calculador de color están dispuestos para asignar los respectivos valores de tal manera que un objeto renderizado (interpretado) detrás de la imagen NPR capturada está siendo percibido como

siendo visto a través de la zona en la que el mapa de prioridad tiene valores inferiores mientras está siendo ocluido donde el mapa de prioridad tiene valores más altos.

- 5 15. Una consola de cirugía robótica para cirugía mínimamente agresiva que incluye un endoscopio estereoscópico para capturar imágenes estereoscópicas de una escena quirúrgica y una disposición de visión estereoscópica para ver las imágenes capturadas, incluyendo la consola un sistema según una cualquiera de las reivindicaciones 8 a 14 dispuesto para renderizar (interpretar) imágenes recibidas desde el endoscopio estereoscópico y presentarlas en la disposición de visión estereoscópica.
- 10 16. Un medio legible por ordenador o señal portadora que codifica un programa de ordenador que, cuando se ejecuta en un ordenador, realiza en la práctica un método según una cualquiera de las reivindicaciones 1 a 7.

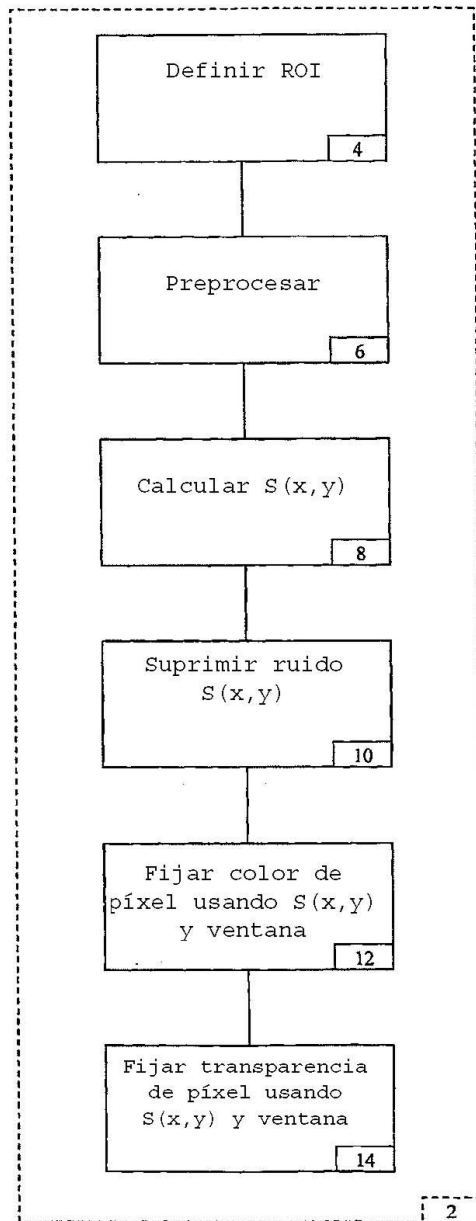


Fig 1

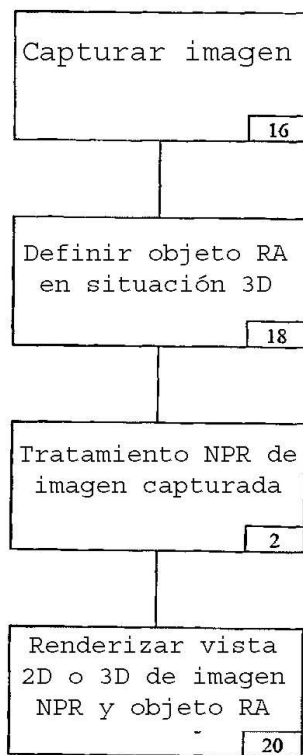


Fig 2

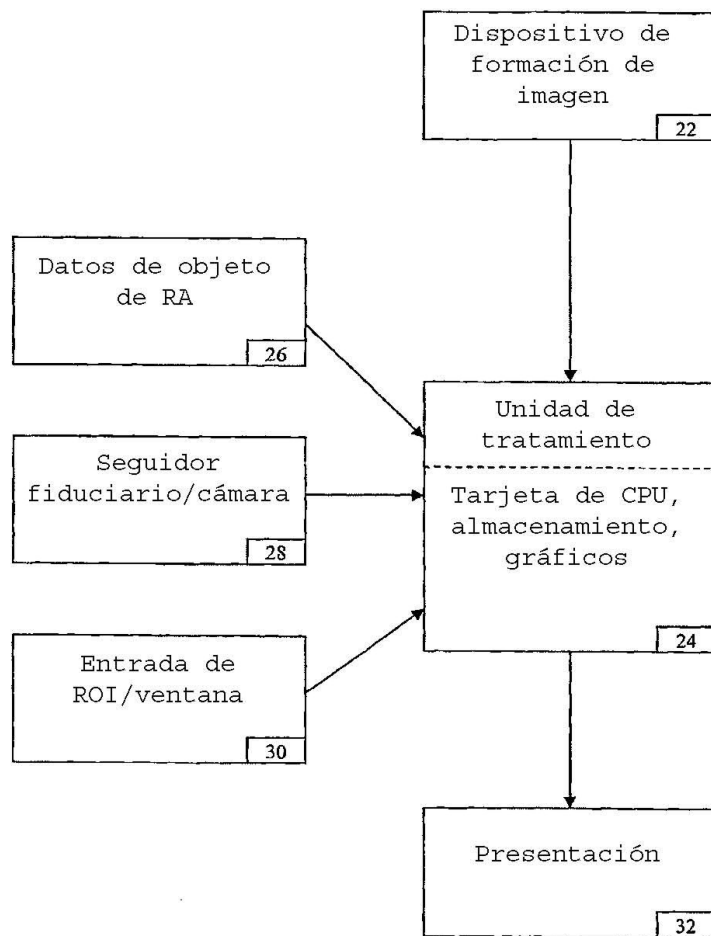


Fig 3

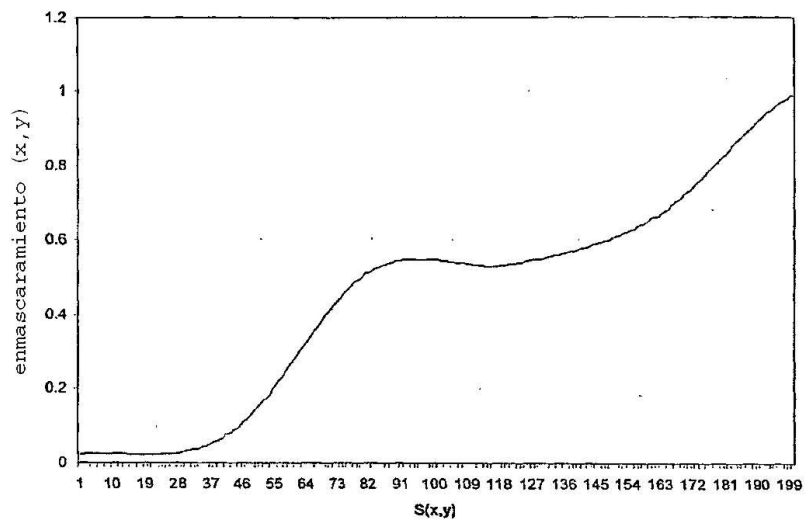


Fig 4

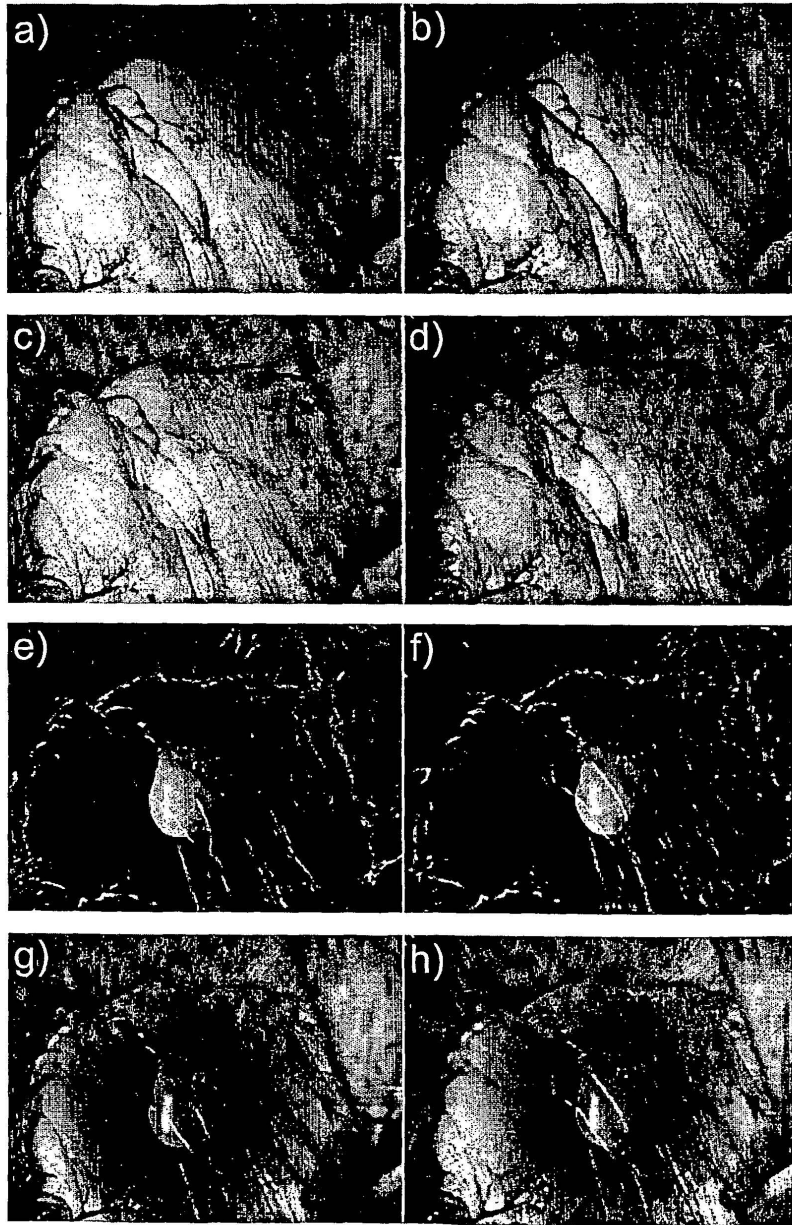


Fig 5