

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 688 399**

51 Int. Cl.:

G06K 9/56	(2006.01)
G06K 9/00	(2006.01)
G06K 9/32	(2006.01)
G06T 7/70	(2007.01)
G06T 7/90	(2007.01)
H04L 29/06	(2006.01)
G06T 7/32	(2007.01)
H04N 5/225	(2006.01)
H04N 5/232	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **09.09.2015 PCT/US2015/049195**
- 87 Fecha y número de publicación internacional: **17.03.2016 WO16040487**
- 96 Fecha de presentación y número de la solicitud europea: **09.09.2015 E 15781779 (2)**
- 97 Fecha y número de publicación de la concesión europea: **27.06.2018 EP 3192008**

54 Título: **Sistemas y métodos para análisis de vida**

30 Prioridad:

09.09.2014 US 201414480802

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

02.11.2018

73 Titular/es:

**EYEVERIFY INC. (100.0%)
1740 Main Street, Suite 100
Kansas City, MO 64108, US**

72 Inventor/es:

HIRVONEN, DAVID

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 688 399 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

Sistemas y métodos para análisis de vida

5 REFERENCIA A APLICACIONES RELACIONADAS

Esta solicitud es una continuación de, y reivindica la prioridad para, la solicitud pendiente de patente de EE. UU. N° 14/480.802, presentada el 9 de septiembre de 2014, titulada "Systems and Methods for the Liveness Analysis", cuya descripción se incorpora en el presente documento como referencia en su totalidad.

10 CAMPO TÉCNICO

La presente invención se refiere, en general, al análisis de imágenes y, en particular, a técnicas de procesamiento de imágenes para detectar si un sujeto representado en una imagen está vivo.

15 ANTECEDENTES

A menudo es deseable restringir el acceso a la propiedad o recursos a individuos concretos. Los sistemas biométricos se pueden utilizar para autenticar la identidad de un individuo para otorgar o negar el acceso a un recurso. Por ejemplo, los escáneres de iris pueden ser utilizados por un sistema de seguridad biométrica para identificar a un individuo en base a estructuras únicas en el iris del individuo. Sin embargo, dicho sistema puede autorizar erróneamente a un impostor si el impostor presenta para escaneo una imagen o un video pregrabado de la cara de una persona autorizada. Tal imagen o video falso puede ser mostrado en un monitor tal como una pantalla de tubo de rayos catódicos (CRT – Cathode Ray Tube, en inglés) o de pantalla de cristal líquido (LCD – Liquid Crystal Display, en inglés), en fotografías con brillo, etc., que se encuentra frente a una cámara utilizada para escanear. Algunos de los llamados sistemas de detección de falsificación pueden detectar una imagen falsa mediante la determinación del movimiento del ojo. Pero, dicho sistema puede no ser efectivo para identificar un video falso que incluye el movimiento esperado del ojo. Por lo tanto, se necesitan sistemas y métodos mejorados para determinar de manera eficiente imágenes y videos falsos de aquellos proporcionados en vivo por las personas autorizadas.

30 COMPENDIO

En diversas implementaciones descritas en la presente memoria, las diferencias en las propiedades reflectantes de caras reales / auténticas y caras de impostores se utilizan para distinguir caras y/u ojos vivos y auténticos de imágenes / videos impostores. Esto se logra, en parte, mediante la representación de una señal de referencia en una pantalla que se encuentra delante de un objetivo, que puede ser una cara real o una imagen falsa, mediante la grabación de un reflejo de la señal de referencia por parte del objetivo, y mediante el cálculo de una o más correlaciones entre las señales reflejadas y las renderizadas.

En consecuencia, en un aspecto, un método implementado por ordenador para determinar si una imagen viva se presenta para autenticación incluye renderizar en un dispositivo de visualización una primera señal variable en el tiempo que incluye varias señales respectivas diferentes que están separadas en fase entre sí. El método incluye asimismo capturar, durante la renderización, varias imágenes de un objetivo que es iluminado con la primera señal renderizada, y aplicar temporalmente un filtro de paso de banda a las imágenes para generar una pluralidad de imágenes filtradas. El método incluye además extraer una segunda señal de las imágenes filtradas, y generar una primera medida basada en, al menos, una correlación temporal de la primera señal y la segunda señal. Además, el método incluye, para cada ubicación de píxel en varias ubicaciones de píxeles, extraer una señal respectiva para la ubicación del píxel en base a los cambios en un valor respectivo de la ubicación de píxeles en el tiempo en una serie de imágenes filtradas, y calcular una puntuación de correlación de la ubicación de píxeles respectiva para cada una de las ubicaciones de píxeles en base a una correlación de la señal extraída respectiva de la ubicación de píxeles para la primera señal. El método incluye además generar una segunda medida basada en, al menos, varias de las puntuaciones de correlación de ubicación de píxeles, y aceptar o rechazar el objetivo en base, por lo menos, a las medidas primera y segunda.

Cada señal respectiva de la primera señal puede tener un color diferente, y cada uno de los colores puede ser representado utilizando una misma frecuencia. En algunas implementaciones, cada señal respectiva de la primera señal es una señal monocromática diferente, y la primera señal puede ser sinusoidal. Cada señal respectiva de la primera señal puede ser una senoide distinta, y las sinusoides se pueden superponer en la primera señal. Las señales respectivas de la primera señal pueden ser generadas al azar. En algunas implementaciones, un valor respectivo concreto de una ubicación de píxel puede ser un color.

Cada imagen puede incluir una serie de imágenes respectivas que han sufrido cada una una transformación respectiva, que está en una resolución respectiva diferente, o que incluye una banda de frecuencia espacial respectiva diferente que corresponde a un fenómeno de iluminación seleccionado. La primera medida puede estar basada más en si una fase de la primera señal coincide con una fase de la segunda señal. La extracción de una segunda señal de las imágenes filtradas puede incluir extraer la segunda señal de un valor de color dominante respectivo de cada una de las imágenes filtradas. En algunas implementaciones, el método incluye además estabilizar el objetivo en varias de las imágenes capturadas y/o procesadas antes de aplicar el filtro de paso de banda. El filtro de paso de banda puede ser aplicado en el dominio de la frecuencia o en el dominio del tiempo.

En algunas implementaciones, generar la segunda medida en base, al menos, a un número de puntuaciones de correlación de ubicación de píxeles incluye combinar las puntuaciones de correlación de ubicación de píxeles para generar la segunda medida. El objetivo puede ser un rostro humano y las puntuaciones de la correlación de la ubicación de píxeles combinados pueden ser para ubicaciones de píxeles de una región particular de la cara. La región particular de la cara puede ser determinada utilizando uno o más de: (i) análisis dinámico de imágenes para evitar, al menos en parte, una o más porciones de la cara que están ocluidas o sobreexpuestas en la pluralidad de imágenes, e (ii) una máscara o un mapa de pesos que representa el conocimiento acerca de las características de la cara que es probable que reflejen la primera señal emitida.

Cada ubicación de píxel puede representar varios elementos de datos de imagen, pudiendo tener, algunos o todos, distintas resoluciones. Cada ubicación de píxel puede ser una combinación ponderada de los elementos de datos de imagen respectivos de la ubicación del píxel. Las varias imágenes capturadas pueden representar una pirámide gaussiana o una pirámide laplaciana. Una imagen filtrada concreta de las imágenes filtradas puede representar una combinación ponderada de varios niveles de la pirámide. Otras realizaciones de este aspecto incluyen sistemas, aparatos y programas informáticos correspondientes.

Las implementaciones concretas del objeto descrito en esta memoria descriptiva pueden realizar una o más de las siguientes ventajas. Por ejemplo, la técnica de detección depende de una señal de patrón de múltiples espectros que se representa mientras se realiza la detección. Es poco probable que las imágenes de la cara y/o el ojo de una persona provenientes de una autenticación preprogramada de video / imagen para la autenticación se correlacionen con la señal de patrón multiespectral proporcionada durante la detección de la vida. Además, cualquier reflexión de la señal de patrón multiespectral de una pantalla que representa dicho video / imagen es probablemente de naturaleza diferente a la reflexión de la cara y/o el ojo de una persona viva. Diversas implementaciones descritas en este documento pueden detectar estas anomalías, tal como se explica a continuación y, de este modo, pueden ser más robustas para distinguir a una persona autorizada viva de videos y/o imágenes falsos.

Los detalles de una o más realizaciones del tema descrito en esta especificación se establecen en los dibujos adjuntos y la descripción que siguen. Otras características, aspectos y ventajas del objeto se pondrán de manifiesto a partir de la descripción, los dibujos y las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

El archivo de la patente o la solicitud contiene, al menos, un dibujo ejecutado en color. Las copias de esta publicación de patente o solicitud de patente con dibujos a color serán proporcionadas por la Oficina tras la solicitud y el pago de la tarifa necesaria.

La figura 1 ilustra un procedimiento de ejemplo para determinar dos medidas de vida.

Las figuras 2(a) - 2(c) representan un fotograma de video capturado de ejemplo, un fotograma de video normalizado y estabilizado correspondiente, y un fotograma de filtrado en paso de banda temporal correspondiente, respectivamente.

Las figuras 3(a) - 3(c) representan ejemplos de señales RGB recuperadas correspondientes a una secuencia de fotogramas de video capturados, a una respuesta de filtro de paso de banda de las señales RGB recuperadas, y a las correspondientes señales RGB de referencia representadas, respectivamente.

Las figuras 4(a) - 4(c) describen, de arriba a abajo, los diagramas de periodo de la transformada rápida de Fourier (FFT) de las señales representadas en las figuras 3(a) - 3(c), respectivamente.

La figura 4(d) representa un filtro Butterworth temporal utilizado para generar la respuesta del filtro de paso de banda representada en la figura 3(b).

La figura 5(a) representa un ejemplo de fotograma de imagen normalizado y estabilizado promediado.

Las figuras 5(b) - 5(d) representan una imagen de correlación bidimensional (2D) correspondiente, una imagen de correlación procesada y una imagen de saturación correspondiente, respectivamente.

La figura 5(e) representa una máscara facial de ejemplo.

Las figuras 5(f) - 5(k) representan fotogramas de video capturados, correspondientes a un ciclo completo de una señal RGB de referencia, utilizando la imagen de correlación 2D mostrada en la figura 5(b).

La figura 6 representa una configuración de ejemplo de un monitor LCD que representa una imagen falsa y un teléfono que captura y analiza la imagen falsa.

La Figura 7 representa otra configuración de ejemplo de un monitor LCD que representa una imagen falsa y un teléfono que captura y analiza la imagen falsa.

Las figuras 8(a) - 8(k) representan un fotograma de imagen media falsa capturada desde un monitor LCD configurado tal como se muestra en la figura 6, la correspondiente imagen de correlación 2D y fotogramas de video, correspondientes a un ciclo completo de la señal RGB mediante cuya utilización se calcula la imagen de correlación 2D mostrada en la figura 8(b).

Las figuras 9(a) - 9(k) representan un fotograma de imagen media falsa capturada desde un monitor LCD configurado tal como se muestra en la figura 7, la correspondiente imagen de correlación 2D, y los fotogramas de video, correspondientes a un ciclo completo de la señal RGB, mediante cuya utilización se calcula la imagen de correlación 2D mostrada en la figura 9(b).

Las figuras 10(a) -10(c) representan las señales RGB recuperadas correspondientes a la secuencia de los cuadros de video falso capturados que se muestran en las figuras 9(f) - 9(k), respuesta de filtro de paso de banda de las señales RGB recuperadas, y las correspondientes señales RGB de referencia representadas, respectivamente.

5 Las figuras 11(a) -11(c) representan, de arriba a abajo, los diagramas de periodo de la transformada rápida de Fourier (FFT) de las señales representadas en las figuras 10(a) -10(c), respectivamente.

La figura 11(d) representa un filtro temporal de Butterworth utilizado para generar la respuesta del filtro de paso de banda representado en la figura 10(b).

La figura 12 representa patrones de Moiré asociados con una imagen falsa.

10 La figura 13 ilustra otro procedimiento de ejemplo para detectar la vida de un ojo.

Las figuras 14(a) y 14(b) representan un ojo de ejemplo que refleja un teléfono que captura una imagen del ojo, y la correspondiente imagen de correlación 2D, respectivamente.

La figura 15(a) muestra la imagen falsa representada en la figura 12 con una resolución más alta.

15 Las figuras 15(b) y 15(c) muestran una porción recortada de alta resolución de la imagen representada en la figura 15(a) y una correlación 2D local calculada de acuerdo con el procedimiento mostrado en la figura 13.

Los números de referencia y las designaciones iguales en los diversos dibujos indican elementos iguales.

DESCRIPCIÓN DETALLADA

20 La figura 1 ilustra un marco general para calcular dos medidas que pueden ser utilizadas para distinguir entre imágenes de un ojo obtenidas de una persona viva real y una falsa, por ejemplo, imágenes o videos previamente capturados de una persona viva). En la etapa 102, se representa un patrón multiespectral en el dispositivo de visualización, de modo que la cara (u "objetivo") de una persona es iluminada mediante el patrón. En algunas implementaciones, el patrón se muestra durante aproximadamente un segundo, pero otras duraciones son posibles.

25 El dispositivo de visualización puede ser el dispositivo de visualización de un aparato de procesamiento de datos tal como, por ejemplo, un teléfono inteligente, unas gafas inteligentes, un reloj inteligente, una tableta, un ordenador portátil, etc. Son posibles otros dispositivos de visualización. Las imágenes del objetivo iluminadas por el patrón multiespectral son capturadas por una cámara digital en la etapa 104. En algunas implementaciones, la cámara digital es una cámara digital frontal del aparato de procesamiento de datos. Se pueden utilizar otras cámaras digitales, incluidas cámaras digitales en otros dispositivos.

En diversas implementaciones, el patrón multiespectral incluye tres señales sinusoidales superpuestas. Por ejemplo, las sinusoides rojas, verdes y azules (RGB – Red Green Blue, en inglés) pueden ser utilizadas para hacer coincidir las sensibilidades de los filtros nativos para cada canal de color en las cámaras digitales comunes de patrón Bayer.

35 Las señales sinusoidales pueden ser renderizadas sustancialmente a una sola frecuencia, de manera que se puede utilizar un solo filtro de paso de banda para su posterior análisis (descrito a continuación). Además, las tres señales sinusoidales pueden estar separadas de manera uniforme en fase a través de los tres canales de color (por ejemplo, rojo = 0, verde = $(2 * \pi) / 3$, y azul = $(2 * \pi) * 2/3$), para mejorar la capacidad de separación de la señal recuperada para reducir las brechas de iluminación que pueden exacerbar los efectos de destello que pueden ser incómodos para algunos usuarios. En una implementación, se utiliza una frecuencia de aproximadamente 4 Hz, que es más baja que el umbral para la epilepsia fotosensible, aunque es lo suficientemente rápida para poder ser separada fácilmente del ruido típico de iluminación de baja frecuencia en un corto período de tiempo. Se pueden utilizar otros patrones multiespectrales además de las sinusoides RGB, que incluyen patrones que tienen menos o más señales componentes, una senoide roja y azul, por ejemplo.

45 En la etapa 104, se graba una señal de video que incluye imágenes capturadas por la cámara digital. En algunas implementaciones, la señal de video es un videoclip de 0,75 segundos aproximadamente a 25 Hz, es decir, 25 fotogramas / segundo. Otras duraciones y velocidades de fotogramas son posibles. En la etapa 106, cada fotograma en la señal de video grabada se puede etiquetar con el valor (por ejemplo, el valor RGB) del patrón representado en el dispositivo de visualización en la etapa 102 aproximadamente en el momento en que se capturó el fotograma de imagen. Los metadatos de archivos de imágenes intercambiables (EXIF) (u otros metadatos) también se pueden almacenar en la etapa 106, en general, para proporcionar una medida de la iluminación ambiental para el ajuste automático del umbral. Los metadatos pueden incluir brillo ambiental, tiempo de exposición, ajuste ISO y/o el valor de la apertura.

55 En algunas implementaciones, se puede llevar a cabo la estabilización de video (registro y deformación) en la señal de video grabada en la etapa 108 para asignar puntos en la escena al sistema de coordenadas común de referencia. Después de la estabilización y la deformación, los fotogramas pueden ser convertidos en un espacio de color RGB normalizado para reducir la sensibilidad a las sombras y otras aberraciones de iluminación en el entorno y, de este modo, se obtiene una señal de video estabilizada y normalizada en la etapa 108.

60 En la etapa 110, el video estabilizado y normalizado es procesado utilizando un filtro de paso de banda temporal que está sintonizado a la frecuencia de la senoide procesada, por ejemplo, 4 Hz en un ejemplo. A modo de ilustración, el filtro se puede aplicar a las pirámides gaussianas correspondientes a los fotogramas de video estabilizados y normalizados. El filtrado en paso de banda temporal se puede realizar para aislar, a partir de la señal normalizada obtenida en la etapa 108, una señal de respuesta correspondiente al patrón multiespectral representado en la etapa

102. Finalmente, la señal de video filtrada por paso de banda es comparada con el patrón multiespectral previamente renderizado, por ejemplo, a diferentes escalas, para obtener: (1) una correlación temporal basada en el fotograma global en la etapa 112, y/o (2) una correlación local de píxeles en la etapa 114, tal como se describe a continuación.

5 Con el fin de calcular una medida global de la correlación temporal, cada fotograma de la señal de respuesta filtrada obtenida en la etapa 110 se puede representar con un valor RGB dominante, en la etapa 122. Se espera que el valor RGB dominante asignado en la etapa 122 corresponda al color del patrón multiespectral RGB entregado, tal como se representa mediante los valores RGB etiquetados en la señal de video grabada en la etapa 106. A modo de
10 ilustración, los valores RGB dominantes se pueden calcular mediante un modo robusto a partir de un histograma de cromaticidad, o como una media ponderada de los valores de píxel para cada fotograma. Otras formas de determinar el valor RGB dominante son posibles.

15 Se calcula una imagen de saturación media a partir de la señal de respuesta filtrada (etapa 110), y se puede utilizar para proporcionar los pesos para el método de media ponderada (etapa 122). En algunas implementaciones, la imagen de saturación promedio es la distancia desde una imagen gris correspondiente al fotograma que se va a procesar. La imagen de saturación bidimensional (2D) resultante es proporcional a la intensidad del patrón multiespectral RGB reflejado. A continuación, en la etapa 124, se realiza una deducción de tendencia lineal de manera independiente en cada una de las señales estimadas de rojo, verde y azul, para eliminar cualquier
20 componente de rampa de los datos, lo que lo hace más adecuado para la comparación con la señal patrón multiespectral RGB de referencia. La tendencia lineal puede ser calculada utilizando un estimador m lineal, por ejemplo.

25 La figura 3(a) muestra un ejemplo de señal RGB global. La señal se denomina "global" porque representa los valores RGB dominantes correspondientes a un fotograma, y no a ningún píxel en particular en ese fotograma. En la etapa 126, esta señal global es procesada con un filtro de Butterworth de paso de banda temporal en el dominio de la frecuencia, para extraer la frecuencia apropiada correspondiente a la señal grabada. Las figuras 2(b) y 2(c) muestran la señal RGB filtrada y la señal de referencia renderizada (es decir, el patrón multiespectral RGB), respectivamente. Estas dos señales son comparadas en la etapa 128 utilizando una correlación cruzada normalizada, y el valor
30 resultante, designado como n_{corr} , indica una primera medida de vida. En una implementación, se realiza una pequeña búsqueda temporal unidimensional (1D) en la etapa 128 para compensar la latencia en el controlador de la cámara, que puede causar un pequeño desplazamiento entre las señales RGB medidas y las renderizadas. La búsqueda es una búsqueda 1D, porque cada punto en la forma de onda combinada en la figura 2(a) representa un fotograma completo. Las figuras 4(a) - 4(c) representan los diagramas de periodo de la transformada rápida de
35 Fourier (FFT – Fast Fourier Transform, en inglés) de las señales representadas en las figuras 3(a) - 3(c), respectivamente.

Correlación local por píxeles

40 En la etapa 114, se calcula una media espacial de correlación cruzada normalizada temporal local calculada en cada ubicación de píxel en la respuesta de video filtrada (es decir, la señal obtenida en la etapa 110 mediante el filtrado de la señal registrada estabilizada y normalizada por medio del filtro de paso de banda temporal). La media espacial puede producir una imagen de correlación 2D (por ejemplo, en un rango [-1 ...+1], que puede indicar con qué precisión coincide cada píxel en la respuesta filtrada con la señal RGB renderizada. Por ejemplo, la figura 5(b) muestra una imagen de correlación correspondiente a una imagen registrada estabilizada y normalizada de ejemplo
45 representada en la figura 5(a). La figura 5(c) muestra una imagen de correlación 2D procesada obtenida, por ejemplo, seleccionando el máximo de las imágenes de correlación izquierda y derecha, tal como se describe a continuación. Para calcular una correlación 2D, se puede aplicar una máscara facial en la etapa 132, por ejemplo, para restringir el procesamiento a la parte de piel de la cara, y para eliminar las características oscuras de la cara con un albedo bajo y/o para eliminar el ruido del movimiento independiente de los ojos. La figura 5(e) representa una máscara de ejemplo. A continuación, se calcula la correlación local píxel por píxel en la etapa 134, por ejemplo, para
50 cada uno de los fotogramas de imagen mostrados en las figuras 5(f) - 5(k). Estas imágenes corresponden a un ciclo completo del patrón multiespectral RGB, y las respectivas correlaciones píxel por píxel pueden ser promediadas y procesadas para obtener la imagen final de correlación 2D que se muestra en la figura 5(c).

55 En algunas implementaciones, al calcular la correlación local, píxel por píxel, el retardo de fase recuperado de la correlación global anterior puede ser utilizado en la etapa 134 para evitar la necesidad de una búsqueda costosa de la correlación en los datos volumétricos correspondientes a los fotogramas estabilizados y normalizados obtenidos en la etapa 110. En algunas implementaciones, los valores promedio de la correlación cruzada espacial normalizada se calculan por separado, en las etapas 136, 138, respectivamente, para los lados izquierdo y derecho de la
60 máscara facial. El máximo de dos correlaciones espaciales puede ser seleccionado en la etapa 140. Esto puede proporcionar una medida más robusta de la correlación, que solo una media, ya que las condiciones de iluminación extrema a menudo se limitan a un solo lado de la cara. Alternativamente, la media global de todos los píxeles de la máscara facial puede ser utilizado si el valor del brillo ambiental de los metadatos EXIF es lo suficientemente bajo para hacer que la saturación sea poco probable, como la que se puede encontrar en la mayoría de los entornos de
65 interior. La figura 5(d) representa una imagen de saturación correspondiente a la imagen de correlación 2D que se

muestra en la figura 5(c). La medida de correlación local promediada final, designada como $nxcrr2$, puede ser una segunda medida de vida.

Típicamente, la piel de una cara real proporciona una reflexión relativamente difusa con un alto albedo y, por lo tanto, el valor de correlación en cada píxel puede ser alto. La imagen de correlación también tiende a ser bastante uniforme, con una varianza espacial relativamente baja. Por el contrario, cuando se utiliza un monitor de video para una reproducción impostora, el monitor tiende a comportarse como un espejo y, dependiendo del ángulo de reflexión de la luz emitida desde la pantalla en la que se representa el patrón multispectral RGB, la luz es reflejada principalmente de manera local en una pequeña porción de la imagen de la cara capturada en la pantalla (tal como se representa en la figura 6) o se refleja fuera de la pantalla de visualización, tal como se muestra en la figura 7.

Por ejemplo, la figura 8(a) representa una imagen impostora capturada que se muestra en una pantalla LCD mantenida frente al dispositivo al que se debe autorizar el acceso (por ejemplo, un teléfono), tal como se muestra en la figura 6. Las figuras 8(b) y 8(c) muestran las correspondientes imágenes de correlación 2D, la figura 8(d) muestra la imagen de saturación correspondiente, la figura 8(e) muestra la máscara aplicada, y las figuras 8(f) a 8(k) representan diversos fotogramas de imagen capturados que corresponden a un ciclo completo del patrón multispectral RGB proporcionado tal como se muestra en la etapa 102 en la figura 1. En este ejemplo, la segunda medición $nxcrr2$ es alta (alrededor de 0,63) debido a que la pantalla LCD es mantenida paralela al teléfono utilizado para capturar las imágenes, y debido a que la pantalla LCD actúa como un espejo. La primera medida $nxcrr$, es decir, la correlación global, es baja, no obstante, lo que indica que las imágenes capturadas probablemente no se obtienen de una fuente viva. Si la pantalla LCD que muestra las imágenes impostoras es mantenida en un ángulo relativo a la pantalla utilizada para representar el patrón multispectral RGB, tal como se muestra en la figura 7, por ejemplo, se espera que los valores de $nxcrr2$ y $nxcrr$ sean bajos, es decir, menores que un umbral seleccionado de 15, tal como 0,5, 0,4, 0,3, etc. Un ejemplo típico correspondiente a este caso, en el que la luz es reflejada lejos de la cámara, se muestra en las figuras. 9(a) - 9(k). En este caso, ni las medidas de correlación local global ni media corresponden a la señal RGB esperada, lo que, en general, hace que ambas medidas $nxcrr$ y $nxcrr2$ sean bajas. Por lo tanto, la señal de respuesta filtrada obtenida en la etapa 124 puede tener mucho ruido, tal como ilustra la señal 1D RGB mostrada en las figuras 10(a) a 10(c).

Además de aprovechar las propiedades similares a un espejo de muchas pantallas de reproducción de video, las medidas de correlación pueden reflejar otras anomalías de una reproducción de video, por ejemplo, aberraciones de muestreo tales como bandas verticales en las imágenes de salida filtradas de paso de banda temporal, tal como se puede ver en los últimos seis fotogramas en la figura 9. En una implementación, una FFT normalizada para cada señal de color representada en la señal de respuesta filtrada es un indicador fuerte de que el sujeto es un impostor, tal como se puede ver en la figura 11. Las tres filas superiores son los diagramas de periodo correspondientes a los canales de color rojo, verde y azul, obtenidos a partir de la señal de respuesta filtrada / obtenidas en la etapa 110, figura 1). La última fila es un filtro Butterworth de paso de banda temporal sintonizado al período esperado de la señal en el video grabado. Una relación baja de la señal de paso de banda filtrada con respecto a la energía total de la señal es otra medida que se puede utilizar para detectar casos de impostores.

El análisis de los reflejos desde la pantalla LCD mantenida delante del dispositivo de captura de imágenes (por ejemplo, una cámara de teléfono celular) puede ser utilizado para ayudar a la detección de un impostor cuando, por ejemplo, $nxcrr2$ es alta, pero $nxcrr$ es baja, tal como se describe haciendo referencia a las figuras 8(a) - 8(k). Por ejemplo, las figuras 12(a) - 12(c) muestran una imagen falsa que se muestra en una pantalla LCD que se encuentra delante de una cámara, una imagen recortada de la región de la cara cerca del ojo y una imagen de borde correspondiente a la imagen recortada, que representa un reflejo del teléfono que se utilizó para capturar la imagen falsa mostrada en la pantalla LCD. Otra aberración son los patrones muaré del monitor, que son visibles en la imagen de correlación 2D, tal como se puede ver en la figura 12(d). Un clasificador 2D, tal como un clasificador Haar, se puede entrenar para identificar patrones en la imagen de correlación que son únicos para los casos de impostores. En general, en diversas implementaciones, se devuelve una clasificación de auténtica si y solo si tanto la correlación global ($nxcrr$) como la correlación local ($nxcrr2$) exceden un umbral predeterminado.

La figura 13 ilustra otra técnica de detección de impostores que aprovecha las propiedades de reflexión de un ojo típico. Específicamente, la etapa 1302 de renderizar un patrón multispectral RGB, la etapa 1304 de capturar una señal de video, la etapa 1306 de etiquetar cada fotograma con un valor RGB, y la etapa 1306 de estabilizar la señal de video grabada y etiquetada se realizan de manera similar a la descrita anteriormente haciendo referencia a la figura 1. Posteriormente, en la etapa 1308 se realiza una descomposición en paso de banda espacio temporal para aprovechar las propiedades reflectantes convexas del ojo. Se observa que un ojo tiene típicamente una superficie reflectante convexa, de modo que cada fotograma de imagen capturado en la etapa 1304 incluye una imagen reducida del espejo del entorno del ojo, que puede incluir una imagen compacta del patrón RGB renderizado en una pantalla de visualización en la etapa 1302.

En la etapa 1310 se aplican filtros temporales de paso de banda a una pirámide laplaciana correspondiente a señales marcadas estabilizadas. La pirámide laplaciana puede proporcionar una descomposición del paso de banda espacial del video de entrada para ayudar a aislar las frecuencias espaciales principalmente altas del patrón multispectral RGB reflejado desde el ojo.

A continuación, se produce una imagen de correlación 2D píxel por píxel por medio de la correlación cruzada normalizada temporal entre la señal de referencia y la salida de video filtrada en paso de banda, en la etapa 1312. Un promedio local en una proximidad pequeña del pico dominante puede ser utilizado como una medida adicional de vida. En general, este enfoque puede detectar la viveza de los ojos en lugar de detectar la vitalidad de la cara utilizando las primera y segunda medidas descritas anteriormente. En una correlación local píxel por píxel solo para la región del ojo de un ojo real vivo, se espera solamente una mancha brillante correspondiente a la reflexión de la señal RGB renderizada por la pupila del ojo, tal como se puede ver en las figuras 14(a) y 14(b). Si se ven múltiples puntos o no se detecta ningún punto, se determina que las imágenes capturadas probablemente están suministradas por un impostor.

Los sistemas y las técnicas que se describen en el presente documento pueden ser implementadas en un sistema informático que incluye un componente del lado del sistema operativo (back end, en inglés) (por ejemplo, tal como un servidor de datos), o que incluye un componente entre el sistema operativo y las aplicaciones (middleware, en inglés) (por ejemplo, un servidor de aplicaciones), o que incluye un componente del lado de las aplicaciones (front end, en inglés) (por ejemplo, un ordenador del cliente que tiene una interfaz gráfica de usuario o un navegador web con el que un usuario puede interactuar con una implementación de los sistemas y técnicas descritos en el presente documento), o cualquier combinación de dichos componentes back end, middleware o front end. Los componentes del sistema pueden estar interconectados por medio de cualquier forma o medio de comunicación de datos digitales (por ejemplo, una red de comunicación). Ejemplos de redes de comunicación incluyen una red de área local ("LAN" – Local Area Network, en inglés), una red de área extensa ("WAN" – Wide Area Network, en inglés) e Internet.

El sistema informático puede incluir clientes y servidores. Un cliente y un servidor, en general, son remotos, y pueden interactuar a través de una red de comunicación. La relación del cliente y el servidor surge en virtud de los programas informáticos que se ejecutan en los ordenadores respectivos y que tienen una relación de cliente a servidor entre sí. Se han descrito varias realizaciones. No obstante, se comprenderá que se pueden realizar diversas modificaciones sin apartarse del alcance de la invención.

Las realizaciones del objeto y las operaciones descritas en esta memoria descriptiva pueden ser implementadas en circuitos electrónicos digitales, o en software, firmware o hardware, que incluyen las estructuras descritas en esta memoria descriptiva y sus equivalentes estructurales, o en combinaciones de uno o más de ellos. Las realizaciones del objeto descrito en esta memoria descriptiva pueden ser implementadas como uno o más programas informáticos, es decir, uno o más módulos de instrucciones de programas informáticos, codificados en un medio de almacenamiento informático para su ejecución, o para controlar el funcionamiento del aparato de procesamiento de datos. Alternativa o adicionalmente, las instrucciones del programa pueden ser codificadas en una señal propagada generada artificialmente, por ejemplo, una señal eléctrica, óptica o electromagnética generada mediante una máquina, que se genera para codificar información para su transmisión a un aparato receptor adecuado para su ejecución mediante un aparato de procesamiento de datos. Un medio de almacenamiento informático puede ser, o estar incluido en, un dispositivo de almacenamiento legible por ordenador, un sustrato de almacenamiento legible por ordenador, un conjunto o dispositivo de memoria de acceso aleatorio o en serie, o una combinación de uno o más de ellos. Además, aunque un medio de almacenamiento informático no es una señal propagada, un medio de almacenamiento informático puede ser una fuente o destino de instrucciones de programa informático codificadas en una señal propagada generada artificialmente. El medio de almacenamiento informático también puede ser, o puede estar incluido en, uno o más componentes o medios físicos separados (por ejemplo, múltiples CD, discos u otros dispositivos de almacenamiento).

Las operaciones descritas en esta memoria descriptiva pueden ser implementadas como operaciones realizadas mediante un aparato de procesamiento de datos sobre datos almacenados en uno o más dispositivos de almacenamiento legibles por ordenador o recibidos de otras fuentes.

El término "aparato de procesamiento de datos" abarca todo tipo de aparatos, dispositivos y máquinas para procesar datos, que incluyen, a modo de ejemplo, un procesador programable, un ordenador, un sistema en uno o en múltiples chips, o combinaciones de los anteriores. El aparato puede incluir una circuitería lógica de propósito especial, por ejemplo, una FPGA (matriz de puertas programable por campo – Field Programmable Gate Array, en inglés) o un ASIC (circuito integrado específico para una aplicación – Application-Specific Integrated Circuit, en inglés). El aparato puede incluir, asimismo, además de hardware, código que crea un entorno de ejecución para el programa informático en cuestión, por ejemplo, código que constituye el firmware del procesador, una pila de protocolo, un sistema de administración de bases de datos, un sistema operativo, un entorno de ejecución multiplataforma, una máquina virtual o una combinación de uno o más de ellos. El aparato y el entorno de ejecución pueden realizar diversas infraestructuras de modelos informáticos diferentes, tales como servicios web, cálculo distribuido e infraestructuras de cálculo en malla.

Un programa informático (también conocido como programa, software, aplicación de software, secuencia de comandos o código) puede estar escrito en cualquier forma de lenguaje de programación, incluidos los lenguajes compilados o interpretados, los lenguajes declarativos o de procedimiento, y puede ser desplegado en cualquier forma, incluso como un programa independiente o como un módulo, componente, subrutina, objeto u otra unidad

adecuada para ser utilizada en un entorno informático. Un programa informático puede, pero no es necesario, corresponder a un archivo en un sistema de archivos. Un programa puede estar almacenado en una parte de un archivo que contiene otros programas o datos (por ejemplo, uno o más textos almacenados en un recurso de lenguaje de marcado), en un único archivo dedicado al programa en cuestión, o en múltiples archivos coordinados (por ejemplo, archivos que almacenan uno o más módulos, subprogramas o porciones de código). Se puede implementar un programa informático para ser ejecutado en un ordenador o en varios ordenadores ubicados en un sitio o distribuidas a través de múltiples sitios e interconectados por una red de comunicación.

Las realizaciones del tema descrito en esta especificación pueden ser implementadas en un sistema informático que incluye un componente back end, por ejemplo, tal como un servidor de datos, o que incluye un componente middleware, por ejemplo, un servidor de aplicaciones, o que incluye un componente front end, por ejemplo, un ordenador de un cliente que tiene una interfaz gráfica de usuario o un navegador web a través del cual un usuario puede interactuar con una implementación del objeto descrito en esta memoria descriptiva, o cualquier combinación de uno o más de dichos componentes back end, middleware o front end. Los componentes del sistema pueden estar interconectados por cualquier forma o medio de comunicación de datos digitales, por ejemplo, una red de comunicación. Ejemplos de redes de comunicación incluyen una red de área local ("LAN") y una red de área extensa ("WAN"), una red interconectada (por ejemplo, Internet) y redes de igual a igual (por ejemplo, redes de igual a igual ad hoc).

El sistema informático puede incluir clientes y servidores. Un cliente y un servidor, en general, son remotos y pueden interactuar a través de una red de comunicación. La relación de cliente y servidor surge en virtud de los programas de ordenador que se ejecutan en los respectivos ordenadores y que tienen una relación de cliente a servidor entre sí. En algunas realizaciones, un servidor transmite datos (por ejemplo, una página HTML) a un dispositivo del cliente (por ejemplo, con el fin de mostrar datos y recibir la entrada de usuario de un usuario que interactúa con el dispositivo del cliente). Los datos generados en el dispositivo del cliente (por ejemplo, un resultado de la interacción del usuario) pueden ser recibidos desde el dispositivo del cliente en el servidor.

Un sistema de uno o más ordenadores se puede configurar para realizar operaciones o acciones concretas en virtud de tener software, firmware, hardware o una combinación de ellos, instalado en el sistema, que en operación causa o hace que el sistema realice las acciones. Se pueden configurar uno o más programas informáticos para realizar operaciones o acciones concretas mediante la inclusión de instrucciones que, cuando son ejecutadas mediante un aparato de procesamiento de datos, hacen que el aparato realice las acciones.

Si bien esta memoria descriptiva contiene muchos detalles específicos de implementación, estos no deben ser interpretados como limitaciones en el alcance de cualquier invención o de lo que se puede reivindicar, sino más bien como descripciones de características específicas de realizaciones concretas de invenciones concretas. Ciertas características que se describen en esta memoria descriptiva en el contexto de realizaciones separadas pueden ser implementadas asimismo en combinación en una única realización. A la inversa, diversas características que se describen en el contexto de una única realización pueden ser implementadas asimismo en múltiples realizaciones, por separado o en cualquier combinación secundaria adecuada. Además, aunque las características se han podido describir anteriormente como que actúan en ciertas combinaciones e incluso se reivindican inicialmente como tales, una o más características de una combinación reivindicada pueden ser, en algunos casos, escindidas de la combinación, y la combinación reivindicada puede dirigirse a un combinación secundaria o variación de una combinación secundaria.

De manera similar, aunque las operaciones se representan en los dibujos en un orden particular, esto no se debe entender como que requiere que dichas operaciones se realicen en el orden particular mostrado o en orden secuencial, o que se realicen todas las operaciones ilustradas, para lograr resultados deseables. En determinadas circunstancias, la multitarea y el procesamiento paralelo pueden ser ventajosos. Además, la separación de diversos componentes del sistema en las realizaciones descritas anteriormente no se debe entender como que requiere dicha separación en todas las realizaciones, y se debe entender que los componentes y sistemas del programa descritos, en general, pueden estar integrados juntos en un solo producto de software o empaquetados en múltiples productos de software.

Por lo tanto, se han descrito realizaciones particulares del objeto. Otras formas están dentro del alcance de las siguientes reivindicaciones. En algunos casos, las acciones enumeradas en las reivindicaciones pueden ser realizadas en un orden diferente y, aun así, lograr los resultados deseables. Además, los procesos representados en las figuras adjuntas no requieren necesariamente el orden particular mostrado, o el orden secuencial, para lograr resultados deseables. En ciertas implementaciones, la multitarea y el procesamiento en paralelo pueden ser ventajosos.

REIVINDICACIONES

1. Un método implementado por ordenador que comprende:

5 renderizar (102) en un dispositivo de visualización una primera señal variable en el tiempo que comprende una pluralidad de diferentes señales respectivas que están separadas en fase entre sí; durante el renderizado, capturar (104) una pluralidad de imágenes de un objetivo que se ilumina mediante la primera señal renderizada;
 10 aplicar (110) un filtro de paso de banda temporal a las imágenes para generar una pluralidad de imágenes filtradas;
 extraer una segunda señal de las imágenes filtradas;
 generar una primera medida en base, al menos, a una correlación temporal de la primera señal y la segunda señal; para cada ubicación de píxel en una pluralidad de ubicaciones de píxel, extraer una señal respectiva para la ubicación del píxel en base a cambios en un valor respectivo de la ubicación del píxel a lo largo del tiempo en una pluralidad de imágenes filtradas;
 15 calcular una puntuación de correlación de ubicación de píxel respectiva para cada uno de las ubicaciones de píxel en base a una correlación de la señal extraída respectiva de la ubicación del píxel a la primera señal;
 generar una segunda medida en base, al menos, a una pluralidad de las puntuaciones de correlación de ubicación de píxel; y
 20 aceptar o rechazar el objetivo en base, al menos, a las medidas primera y segunda.

2. El método de la reivindicación 1, en el que cada señal respectiva de la primera señal es de un color diferente, en el que, opcionalmente, cada uno de los colores es renderizado utilizando una misma frecuencia.

25 3. El método de la reivindicación 1 en el que cada señal respectiva de la primera señal es una señal monocromática diferente.

4. El método de la reivindicación 1 en el que la primera señal es sinusoidal.

30 5. El método de la reivindicación 1, en el que cada señal respectiva de la primera señal es una sinusoidal distinta, y en el que las sinusoides se superponen en la primera señal.

6. El método de la reivindicación 1, en el que las señales respectivas de la primera señal se generan aleatoriamente.

35 7. El método de la reivindicación 1, en el que un valor respectivo concreto de una ubicación de píxel es un color.

8. El método de la reivindicación 1, en el que cada imagen comprende una pluralidad de imágenes respectivas que se transforman cada una, respectivamente, en una resolución respectiva diferente, o comprenden una banda de frecuencia espacial respectiva diferente que corresponde a un fenómeno de iluminación seleccionado.

40 9. El método de la reivindicación 1, en el que la primera medida se basa además en si una fase de la primera señal coincide con una fase de la segunda señal.

45 10. El método de la reivindicación 1, en el que extraer una segunda señal de las imágenes filtradas comprende extraer la segunda señal de un valor de color dominante respectivo de cada una de las imágenes filtradas.

50 11. El método de la reivindicación 1, que comprende además estabilizar el objetivo en la pluralidad de imágenes antes de aplicar el filtro de paso de banda, y/o en el que el filtro de paso de banda es aplicado en el dominio de la frecuencia o en el dominio del tiempo.

55 12. El método de la reivindicación 1 en el que generar la segunda medida en base, al menos, a una pluralidad de puntuaciones de correlación de ubicación de píxel comprende combinar las puntuaciones de correlación de ubicación de píxel para generar la segunda medida, en el que, opcionalmente el objetivo es un rostro humano, y en el que las puntuaciones de correlación de ubicación del píxel combinadas son para ubicaciones de píxel de una región concreta de la cara, y en el que, además, opcionalmente, se determina la región concreta de la cara utilizando al menos uno de: (i) análisis dinámico de imágenes para evitar, al menos en parte, que una o más porciones de la cara sean ocluidas o sobreexpuestas en la pluralidad de imágenes, e (ii) una máscara o un mapa de pesos que representa el conocimiento acerca de las características de la cara que es probable que reflejen la primera señal renderizada.

60 13. El método de la reivindicación 1, en el que cada ubicación de píxel representa una respectiva pluralidad de elementos de datos de imagen y en el que, opcionalmente, una pluralidad de elementos de datos de imagen están en diferentes resoluciones, y en el que, opcionalmente, cada ubicación de píxel es una combinación ponderada de los elementos de datos de imagen respectivos de la ubicación de píxel.

65

14. El método de la reivindicación 1, en el que la pluralidad capturada de imágenes representa una pirámide gaussiana o una pirámide laplaciana, en el que, opcionalmente una imagen filtrada concreta de las imágenes filtradas representa una combinación ponderada de una pluralidad de niveles de pirámide.
- 5 15. Un sistema que comprende un procesador y una memoria en comunicación electrónica con el procesador, estando programado el procesador para realizar operaciones de acuerdo con el método de cualquiera de las reivindicaciones 1 a 14.

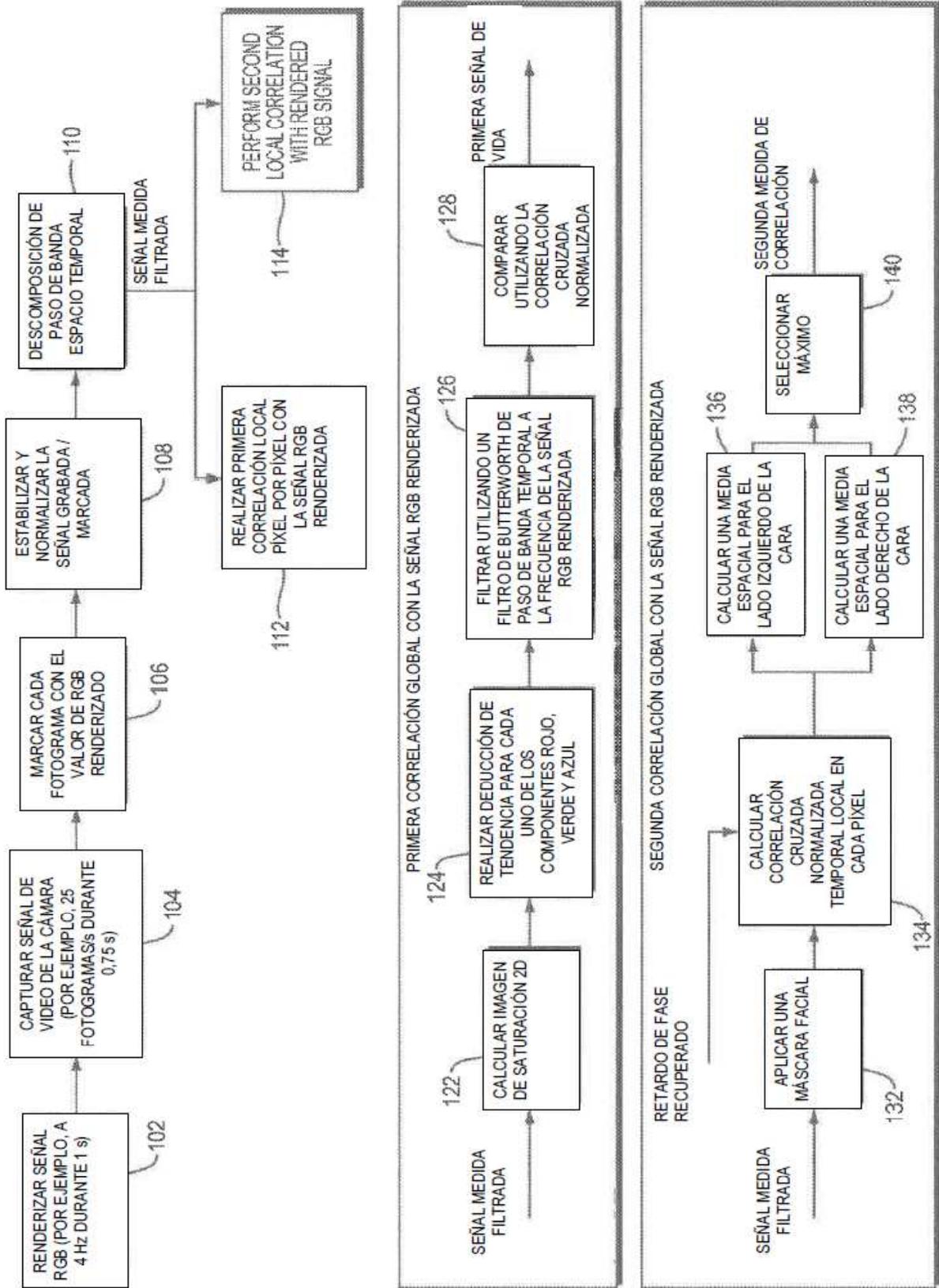


FIG.1

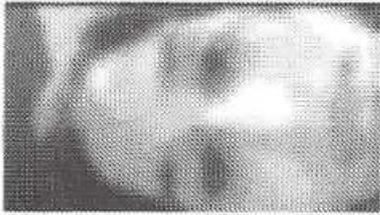


FIG. 2A

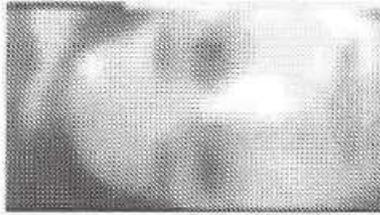


FIG. 2B

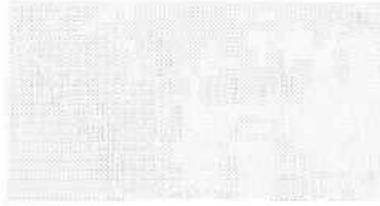


FIG. 2C

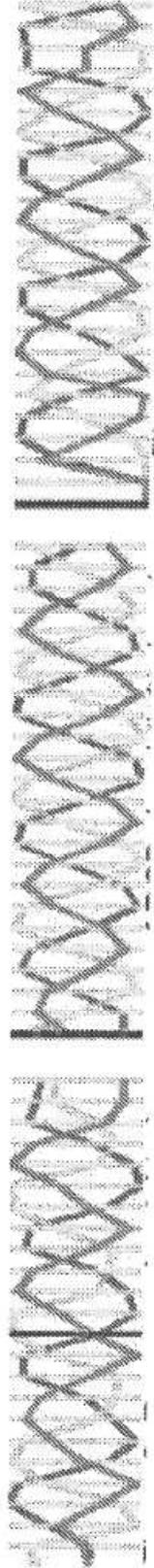


FIG. 3A

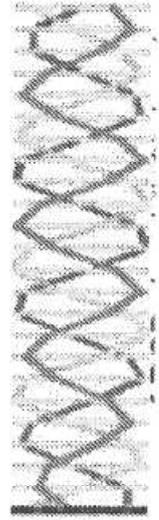


FIG. 3B



FIG. 3C



FIG. 4

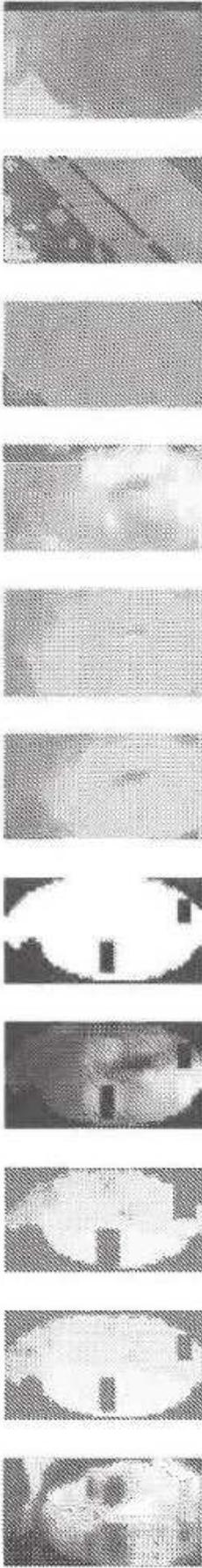


FIG. 5A FIG. 5B FIG. 5C FIG. 5D FIG. 5E FIG. 5F FIG. 5G FIG. 5H FIG. 5I FIG. 5J FIG. 5K

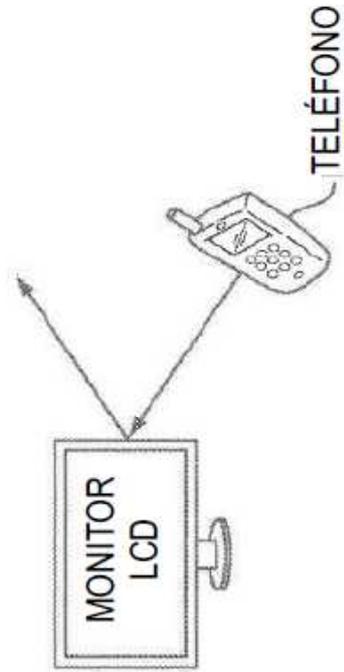


FIG. 7

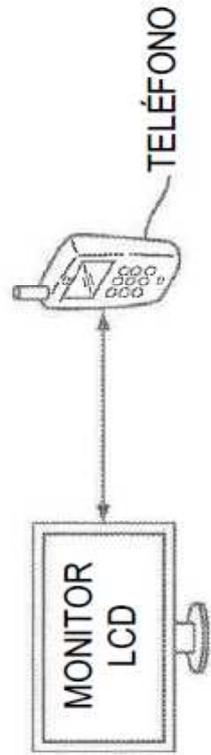


FIG. 6

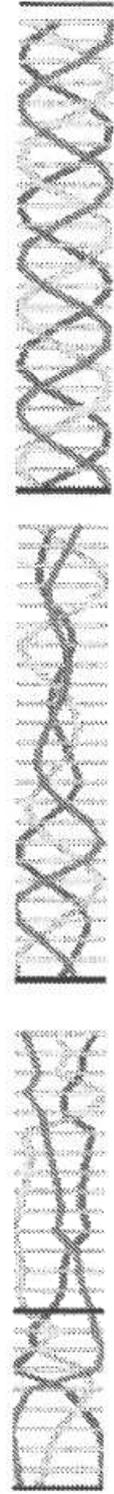
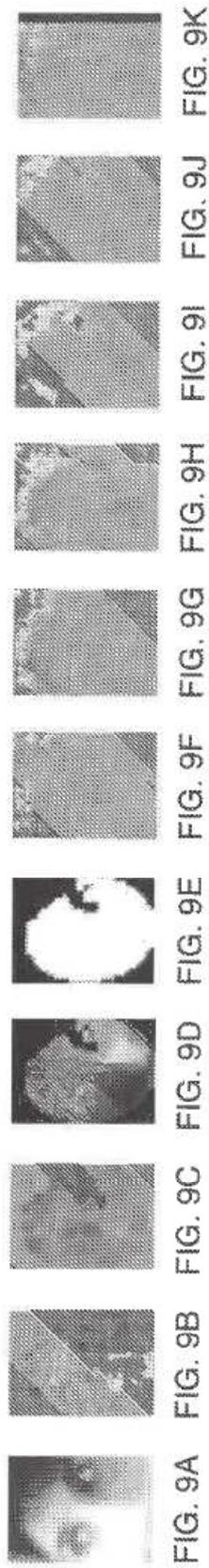
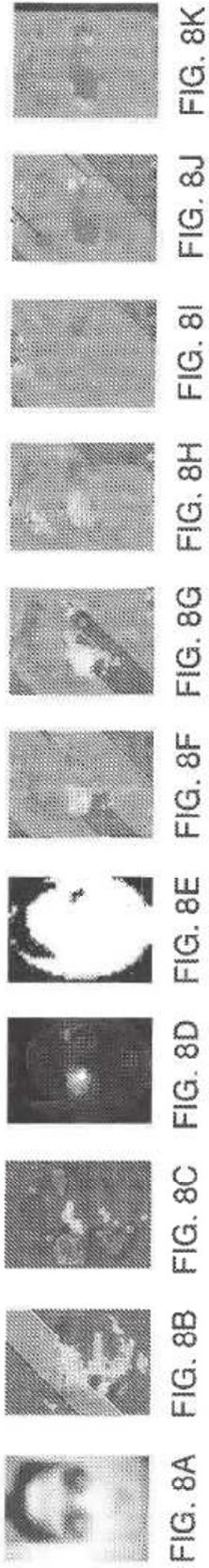


FIG. 11

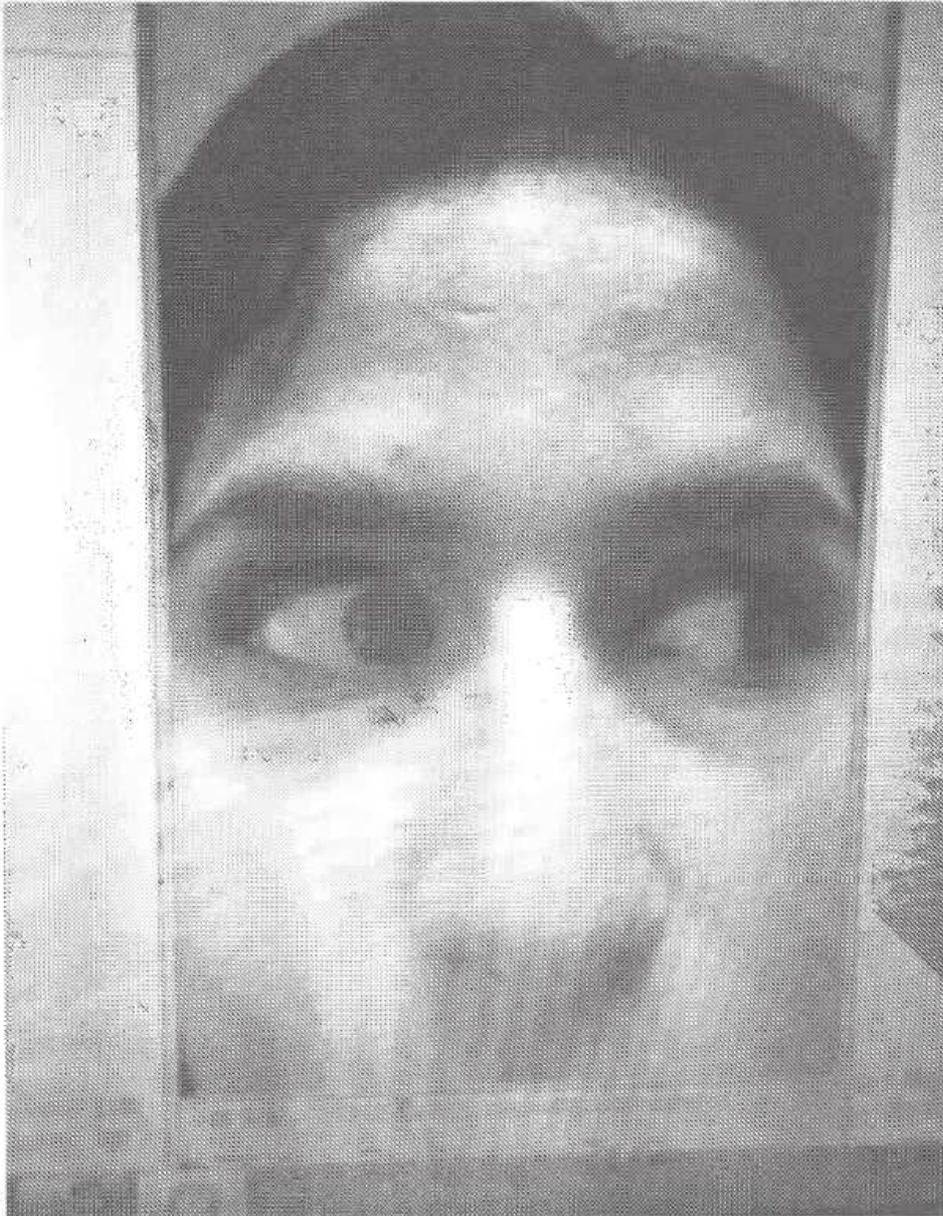


FIG. 12

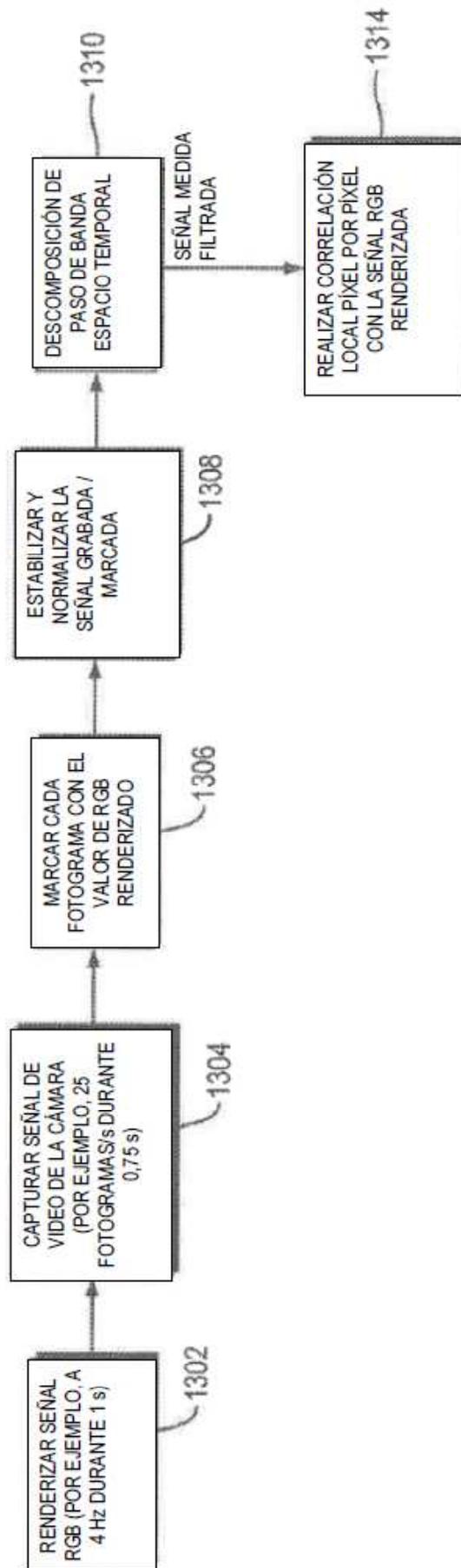


FIG. 13

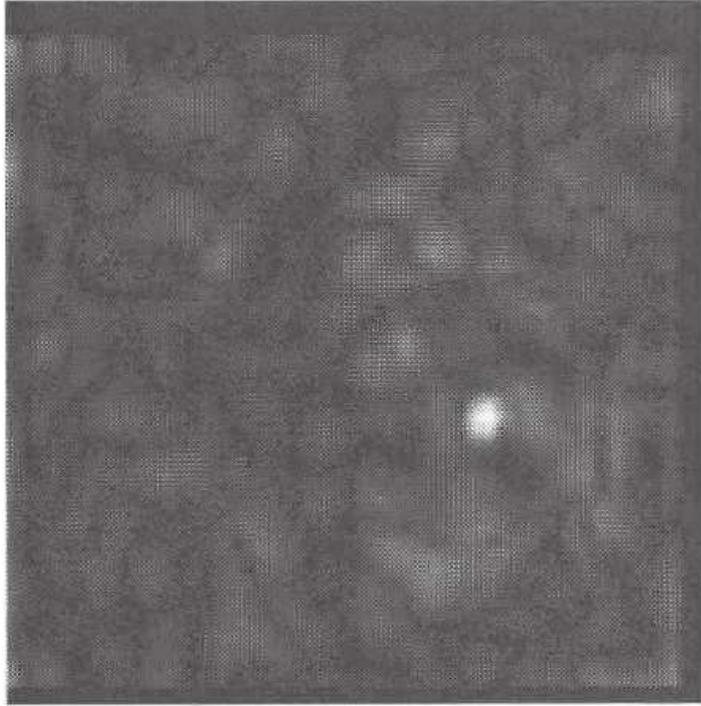


FIG. 14B

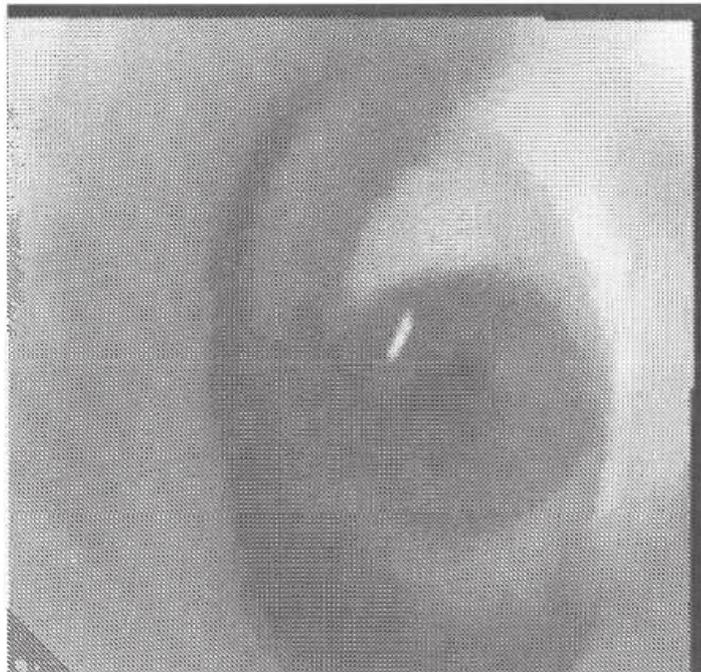


FIG. 14A

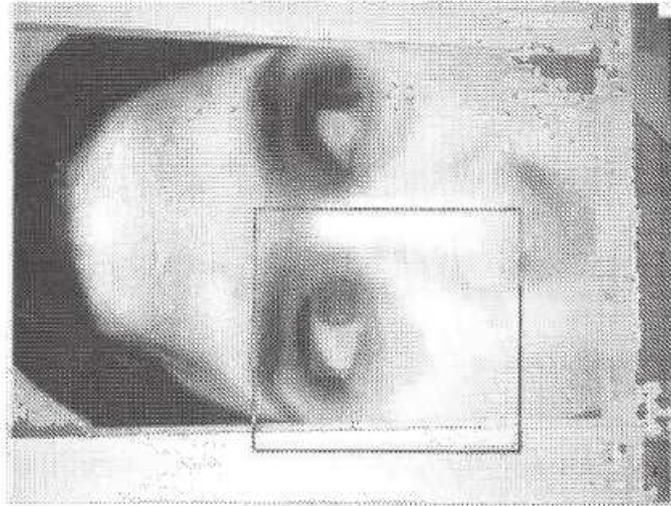


FIG. 15A

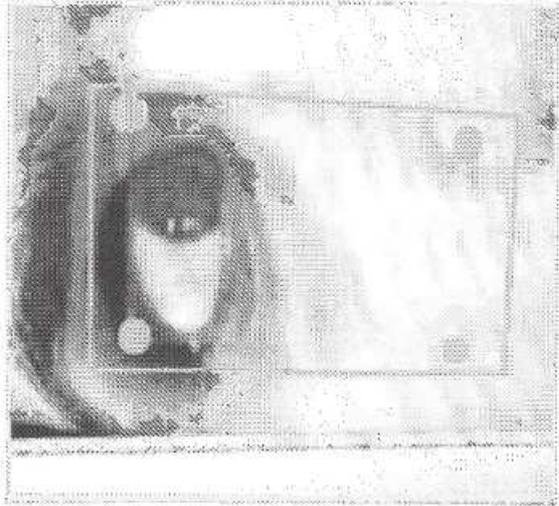


FIG. 15B

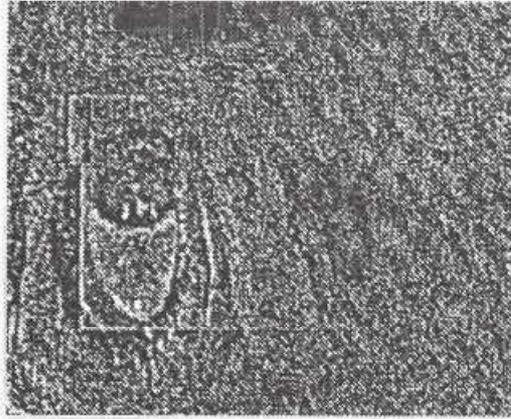


FIG. 15C