

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 994 029**

51 Int. Cl.:

G06T 7/00

(2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.10.2021** **PCT/CN2021/126393**

87 Fecha y número de publicación internacional: **04.05.2023** **WO23070312**

96 Fecha de presentación y número de la solicitud europea: **26.10.2021** **E 21958589 (0)**

97 Fecha y número de publicación de la concesión europea: **02.10.2024** **EP 4209993**

54 Título: **Método de procesamiento de imágenes**

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
16.01.2025

73 Titular/es:

**CONTEMPORARY AMPEREX TECHNOLOGY
(HONG KONG) LIMITED (100.00%)
Level 19, China Building, 29 Queen's Road
Central
Central, Central And Western District, HK**

72 Inventor/es:

**JIANG, GUANNAN y
HAN, LILI**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 994 029 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

Método de procesamiento de imágenes

5 **Campo técnico**

La presente solicitud se refiere al campo del procesamiento de imágenes y, en particular, a un método de procesamiento de imágenes como se especifica en las reivindicaciones 1 a 6 y a un medio de almacenamiento no volátil especificado en la reivindicación 7.

10 **Técnica anterior**

Durante la generación industrial, un producto puede tener algún defecto debido a razones de proceso y equipo. Por lo tanto, es necesario detectar si el producto es defectuoso, para mejorar la tasa de rendimiento de los productos. En los métodos convencionales de detección de defectos, la detección manual presenta desventajas, tales como una baja eficiencia de detección. Por medio de tecnologías de imágenes visuales, se puede mejorar la eficiencia de detección y también se puede reducir en gran medida el coste de la intervención humana. Los métodos actuales de detección visual normalmente emplean una cámara CCD para obtener imágenes bidimensionales (2D), que a continuación se usan para la detección. Sin embargo, un método de detección de este tipo que usa imágenes 2D no es lo suficientemente preciso, lo que puede conducir a detecciones fallidas o falsas de los defectos. Juan Li ET AL: "FUSION OF LIDAR 3D POINTS CLOUD WITH 2D DIGITAL CAMERA IMAGE", Tesis, 1 de enero de 2015 (01-01-2015), documento XP055722335, menciona datos de fusión de múltiples sensores. Guan Limin ET AL: "Real-Time Vehicle Detection Framework Based on the Fusion of LiDAR and Camera", Electronics, vol. 9, n.º 3, 7 de marzo de 2020 (07-03-2020), página 451, documento XP093094214, DOI: 3390/electronics9030451, menciona una detección de vehículos mediante el uso de LiDAR y cámara. El documento CN209432721U menciona un equipo de detección de defectos superficiales de batería de litio.

25 **Sumario de la invención**

30 En vista de los problemas descritos anteriormente, la presente solicitud propone un método de procesamiento de imágenes como se especifica en las reivindicaciones 1 a 6 y un medio de almacenamiento no volátil como se especifica en la reivindicación 7.

35 Para este fin, un primer aspecto de la presente invención proporciona un método de procesamiento de imágenes, donde el método de procesamiento de imágenes incluye:

obtener una imagen bidimensional y una imagen tridimensional de un objeto, en donde el objeto se selecciona de uno cualquiera de los siguientes: una superficie de batería, una soldadura de la cubierta superior de la batería y un pasador de sellado de la batería;

40 extraer información de luminancia e información de profundidad a partir de la imagen tridimensional para generar un mapa de luminancia y un mapa de profundidad para el objeto;

45 aplicar gris a la imagen bidimensional para obtener una imagen en escala de grises de un único canal;

seleccionar al menos los tres puntos de ubicación iguales de cada uno de la imagen en escala de grises y el mapa de luminancia; calcular una matriz de transformación de coordenadas entre la imagen en escala de grises y el mapa de luminancia basándose en las coordenadas de los al menos tres puntos de ubicación iguales en cada uno de la imagen en escala de grises y el mapa de luminancia; y

50 alinear la imagen bidimensional con el mapa de luminancia y el mapa de profundidad basándose en la matriz de transformación de coordenadas, para obtener datos de imágenes fusionadas, donde los datos de imágenes fusionadas incluyen información de color de la imagen bidimensional y la información de luminancia e información de profundidad de la imagen tridimensional, en donde los datos de imágenes fusionadas del objeto se introducen en un modelo de algoritmo de aprendizaje profundo actualizado para obtener un resultado de determinación de defectos del objeto.

55 En la presente invención, se aplica gris a la imagen bidimensional para obtener la imagen en escala de grises de un único canal. A continuación, se resuelve la matriz de transformación de coordenadas entre la imagen bidimensional y la imagen tridimensional. Las características de la imagen bidimensional se alinean con las de la imagen tridimensional, de modo que los datos de características de la imagen bidimensional se alinean con los de la imagen tridimensional, para obtener los datos de la imagen fusionada. De esta manera, se implementa la fusión de imágenes de la imagen bidimensional con la imagen tridimensional. Tales datos de imágenes fusionadas pueden proporcionar más información para el análisis del modelo de aprendizaje profundo posterior, obteniendo de esta manera un resultado de detección más preciso.

El propósito de la aplicación de gris es facilitar la alineación de las características de la imagen bidimensional con las de la imagen tridimensional. En algunas realizaciones, se puede usar un método de media ponderada para aplicar gris a una imagen en color con tres canales (RGB). Específicamente, un valor de escala de grises se puede calcular de acuerdo con la siguiente fórmula:

$$p = w_1 R + w_2 G + w_3 B$$

donde p representa un valor en escala de grises de un punto de coordenadas, R representa un valor de rojo del punto de coordenadas, w_1 representa un peso del valor de rojo, G representa un valor de verde del punto de coordenadas, w_2 representa un peso del valor de verde, B representa un valor de azul del punto de coordenadas y w_3 representa un peso del valor de azul. Las investigaciones muestran que los ojos humanos son más sensibles al verde, seguido del rojo, y son menos sensibles al azul. Por lo tanto, en algunas realizaciones, puede establecerse $w_2 > w_1 > w_3$. En algunas realizaciones, w_1 puede establecerse a 0,299, w_2 puede establecerse a 0,587 y w_3 puede establecerse a 0,114. Esto se debe a que los experimentos y las teorías demuestran que usando tales ajustes de peso se pueden obtener escalas de grises psicológicas que mejor se adaptan a la percepción de los ojos humanos. La información de rojo, verde y azul de la imagen se conserva completamente en los valores de escala de grises obtenidos usando el método de media ponderada. Debe entenderse que, la aplicación de gris se puede realizar usando otros métodos, tales como un método de componentes, un método de máximos, un método de promedio y un método de corrección gamma.

En algunas realizaciones, los al menos tres puntos de ubicación iguales son puntos del objeto con características. Estos puntos con características son ubicaciones particulares en patrones que están dentro de la imagen y pueden obtenerse mediante el análisis de los patrones. Por lo tanto, la selección de los al menos tres puntos de ubicación iguales puede implementarse usando algoritmos de inteligencia artificial.

En algunas realizaciones, los al menos tres puntos de ubicación iguales incluyen tres puntos de referencia no colineales. Cuando los tres puntos de referencia son colineales, la matriz de transformación de coordenadas no se puede obtener a partir de las coordenadas de los tres puntos de referencia. Por lo tanto, se requiere que los al menos tres puntos de ubicación iguales incluyan tres puntos de referencia no colineales.

En algunas realizaciones preferidas, un ángulo interior máximo de un triángulo formado por los tres puntos de referencia no es mayor que 160°. Cuanto mayor sea el ángulo interior máximo del triángulo formado por los tres puntos de referencia, más tenderán los tres puntos de referencia a ser colineales. Pequeños errores de muestreo pueden provocar que los tres puntos de referencia tiendan a ser colineales, de modo que la matriz de transformación de coordenadas obtenida basándose en las coordenadas de los tres puntos de referencia en cada una de las imágenes en escala de grises y el mapa de luminancia sea ineficaz. Por lo tanto, el ángulo interior máximo del triángulo formado por los tres puntos de referencia se establece no mayor que 160°, lo que mejora la fiabilidad de la matriz de transformación de coordenadas.

En algunas realizaciones específicas, el método de procesamiento de imágenes incluye:

alinear el mapa de luminancia y el mapa de profundidad con la imagen bidimensional basándose en la matriz de transformación de coordenadas usando las coordenadas de la imagen bidimensional como referencia; y

usar, como información de tres canales respectivamente, la información de escala de grises de la imagen en escala de grises,

y la información de luminancia y la información de profundidad que están alineadas con las coordenadas en la imagen bidimensional, para generar una imagen fusionada que contiene la información de los tres canales.

Debe entenderse que, una realización de la presente solicitud es alinear el mapa de luminancia y el mapa de profundidad obtenidos a partir de la imagen tridimensional con la imagen bidimensional usando las coordenadas de la imagen bidimensional como referencia. Los expertos en la materia pueden concebir que alinear la imagen bidimensional con el mapa de luminancia, es decir, alinear la imagen bidimensional con la imagen tridimensional usando coordenadas en el mapa de luminancia como referencia, implementa la alineación de la imagen bidimensional con la imagen tridimensional y, además, completa la alineación de datos de las dos imágenes. Sin embargo, la resolución de la imagen tridimensional es baja, por lo que el mapa de luminancia y el mapa de profundidad se alinean preferiblemente con la imagen bidimensional. Además, debe entenderse que únicamente es necesario alinear la información de la imagen bidimensional con la de la imagen tridimensional, y almacenar información de color, información de luminancia e información de profundidad, etc. para cada coordenada en la imagen, en lugar de sintetizar la imagen fusionada que contiene la información de los tres canales.

En algunas realizaciones, la información de color, la información de luminancia y la información de profundidad pueden apilarse en canales para obtener una imagen que tiene información de múltiples canales. En algunas otras

realizaciones, la información de color, la información de luminancia y la información de profundidad pueden fusionarse mediante ponderación para obtener una imagen que tiene información de un único canal o de múltiples canales.

En algunas realizaciones específicas, antes de que la imagen bidimensional se alinee con el mapa de luminancia y el mapa de profundidad, los valores de profundidad en el mapa de profundidad se normalizan para generar un mapa de profundidad normalizado. Los valores de profundidad en el mapa de profundidad generado pueden ser muy grandes, por ejemplo, diez mil o más, por lo tanto, por medio de normalización, los valores de profundidad se pueden limitar entre 0 y 1, lo que facilita el cálculo y el procesamiento posteriores. En algunas realizaciones, los valores de profundidad en el mapa de profundidad pueden normalizarse de acuerdo con la siguiente fórmula:

$$d_s = \frac{d - d_{\min}}{d_{\max} - d_{\min}}$$

donde J representa un valor de profundidad de un punto de coordenadas en el mapa de profundidad, $d_{\min}$ representa un valor de profundidad mínimo en el mapa de profundidad, $d_{\max}$ representa un valor de profundidad máximo en el mapa de profundidad y d_s representa un valor de profundidad normalizado del punto de coordenadas en el mapa de profundidad.

En la presente invención, el objeto descrito anteriormente se selecciona de uno cualquiera de los siguientes: una superficie de batería, una soldadura de cubierta superior de la batería y un pasador de sellado de la batería. Las realizaciones de la presente solicitud son aplicables a una situación donde se requieren tanto una imagen bidimensional como una imagen tridimensional para la detección de defectos. Durante la producción de baterías, los procedimientos de detección de rayones en la superficie de la batería, defectos en la soldadura de la cubierta superior de la batería y defectos en la soldadura del pasador de sellado son todos importantes para garantizar la seguridad de una batería después de salir de la fábrica. La imagen bidimensional no puede reflejar directamente la información de profundidad de las partes relacionadas de la batería y, por lo tanto, no se puede implementar una formación de imágenes efectiva de las características sensibles a la profundidad, mientras que la imagen tridimensional tiene baja resolución y no puede cubrir todas las características objetivo. Por lo tanto, los rayones en la superficie de la batería, los defectos en la soldadura de la cubierta superior de la batería y los defectos en la soldadura del pasador de sellado se pueden detectar usando el método de procesamiento de imágenes de acuerdo con la presente solicitud, para obtener datos de imágenes fusionadas para la detección de características correspondientes.

En algunas realizaciones, la imagen bidimensional y la imagen tridimensional descritas anteriormente pueden ser imágenes adquiridas en diferentes posiciones de una línea de producción. Sin embargo, en algunas otras realizaciones, la imagen bidimensional y la imagen tridimensional descritas anteriormente pueden ser imágenes adquiridas en la misma posición de una línea de producción.

En algunas realizaciones, la imagen bidimensional y la imagen tridimensional descritas anteriormente pueden ser imágenes del mismo objeto tomadas en un ángulo consistente. Sin embargo, en algunas otras realizaciones, la imagen bidimensional y la imagen tridimensional descritas anteriormente pueden ser imágenes del mismo objeto tomadas desde diferentes ángulos.

Otro aspecto de la presente solicitud proporciona un medio de almacenamiento no volátil que almacena instrucciones, donde las instrucciones, cuando son ejecutadas por un procesador, hacen que el procesador realice el método de procesamiento de imágenes del primer aspecto descrito anteriormente.

Otra realización de la presente solicitud proporciona un método de detección de apariencia, que incluye:

obtener datos de imágenes fusionadas de una pluralidad de objetos de acuerdo con el método de procesamiento de imágenes del primer aspecto descrito anteriormente;

obtener resultados de determinación de defectos de la pluralidad de objetos;

usar los datos de imágenes fusionadas y los resultados de determinación de defectos de la pluralidad de objetos para entrenar iterativamente un modelo de algoritmo de aprendizaje profundo, para actualizar el modelo de algoritmo de aprendizaje profundo;

adquirir una imagen bidimensional y una imagen tridimensional de un objeto objetivo;

fusionar la imagen bidimensional y la imagen tridimensional del objeto objetivo de acuerdo con el método de procesamiento de imágenes del primer aspecto descrito anteriormente, para obtener datos de imágenes fusionadas del objeto objetivo; y, como se define en el primer aspecto, introducir los datos de imágenes

fusionadas del objeto objetivo en un modelo de algoritmo de aprendizaje profundo actualizado, para obtener un resultado de determinación de defectos del objeto objetivo.

En realizaciones de la presente solicitud, el resultado de determinación de defectos del objeto objetivo se obtiene sustituyendo los datos de imágenes fusionadas del objeto objetivo en el modelo entrenado. Por lo tanto, se mejora la precisión de la detección y se reducen los fallos de detección y las detecciones falsas (falsos positivos) de defectos. Al mismo tiempo, dado que los datos de imágenes fusionadas se sustituyen en un modelo para el cálculo y el análisis, no hay necesidad de construir dos modelos o sustituir respectivamente datos de diferentes imágenes en los dos modelos para el cálculo y el análisis, ahorrando de esta manera recursos informáticos y aumentando la velocidad de detección. Además, dado que se usan únicamente datos de imágenes fusionadas, únicamente es necesario entrenar con antelación un modelo en lugar de dos.

En algunas realizaciones, los datos de imágenes fusionadas del objeto objetivo incluyen información de escala de grises, información de luminancia e información de profundidad del objeto objetivo, y el modelo de algoritmo de aprendizaje profundo descrito anteriormente incluye un canal de entrada de datos de escala de grises, un canal de entrada de datos de luminancia y un canal de entrada de datos de profundidad. En las realizaciones de la presente solicitud, el modelo de algoritmo de aprendizaje profundo puede incluir tres canales de entrada de datos, y el modelo detecta, basándose en la información de escala de grises de entrada, la información de luminancia y la información de profundidad, si hay un defecto.

En algunas realizaciones, el modelo de algoritmo de aprendizaje profundo descrito anteriormente incluye un canal de entrada de datos de rojo, un canal de entrada de datos de verde, un canal de entrada de datos de azul, un canal de entrada de datos de luminancia y un canal de entrada de datos de profundidad. En las realizaciones de la presente solicitud, el modelo de algoritmo de aprendizaje profundo puede incluir cinco canales de entrada de datos, y el modelo detecta, basándose en la información de color de entrada, la información de luminancia y la información de profundidad, si hay un defecto. Debe entenderse que, en la práctica, los expertos en la materia pueden elegir el número de canales de datos del modelo de acuerdo con los requisitos reales.

En algunas realizaciones, la imagen bidimensional adquirida del objeto objetivo incluye múltiples imágenes bidimensionales.

En algunas realizaciones, la imagen tridimensional adquirida del objeto objetivo incluye múltiples imágenes tridimensionales.

Otra realización de la presente solicitud proporciona un medio de almacenamiento no volátil que almacena instrucciones, donde las instrucciones, cuando son ejecutadas por un procesador, hacen que el procesador realice el método de detección de apariencia.

Breve descripción de los dibujos

Para ilustrar más claramente las soluciones técnicas de la presente solicitud, a continuación, se describirán brevemente los dibujos requeridos en la descripción de las realizaciones de la presente solicitud. Obviamente, los dibujos descritos a continuación son simplemente algunas realizaciones de la presente solicitud, y para los expertos en la materia, también se pueden obtener otros dibujos a partir de estos dibujos sin ningún esfuerzo creativo. En los dibujos:

La Figura 1 ilustra un diagrama de flujo de un método de procesamiento de imágenes de acuerdo con una realización de la presente solicitud;

La Figura 2 ilustra un diagrama de flujo de un método de procesamiento de imágenes de acuerdo con una realización de la presente solicitud;

Las Figuras 3a a 3d muestran fotos de una imagen bidimensional, un mapa de luminancia, un mapa de profundidad y una imagen fusionada de acuerdo con una realización de la presente solicitud; y

La Figura 4 ilustra un diagrama de flujo de un método de detección de apariencia de acuerdo con una realización de la presente solicitud.

Descripción detallada de realizaciones

A continuación, se describirán con más detalle realizaciones de la presente solicitud con referencia a los dibujos. Las siguientes realizaciones simplemente se pretenden para ilustrar más claramente las soluciones técnicas de la presente solicitud, por lo que simplemente sirven como ejemplos, pero no se pretenden para limitar el alcance de protección de la presente solicitud.

A menos que se defina de otra manera, todos los términos técnicos y científicos usados en el presente documento tienen los mismos significados que los comúnmente entendidos por los expertos en la materia a la que pertenece la presente solicitud. Los términos usados en el presente documento tienen únicamente el propósito de describir implementaciones específicas, pero no pretenden limitar la presente solicitud. Las expresiones "que comprende" y "que tiene" y cualquier variación de los mismos en la descripción y las reivindicaciones de la presente solicitud, así como la breve descripción de los dibujos adjuntos descritos anteriormente, tienen por objeto cubrir la inclusión no exclusiva.

El término "realización" mencionado en el presente documento significa que los rasgos, estructuras o características específicas descritas junto con la realización pueden abarcarse en al menos una realización de la presente solicitud. La expresión en diversas ubicaciones en la descripción no se refiere necesariamente a la misma realización, o a una realización independiente o alternativa exclusiva de otra realización. Los expertos en la materia entienden explícita o implícitamente que la realización descrita en el presente documento puede combinarse con otra realización.

Si las etapas se indican en secuencia en la descripción o en las reivindicaciones, esto no significa necesariamente que una realización o aspecto esté limitado a la secuencia indicada. Por el contrario, se puede prever que las etapas se realicen en una secuencia diferente o en paralelo entre sí, a menos que una etapa se establezca sobre otra, lo que requiere absolutamente que la etapa establecida se realice posteriormente (esto quedará claro en un caso individual). Por lo tanto, la secuencia indicada puede ser una realización preferida.

Los inventores de la presente solicitud han descubierto en la práctica que, dado que una imagen 2D no puede reflejar directamente la información de profundidad de cada parte de un producto y no puede cubrir completamente las formas de defectos en el producto, los métodos actuales para detectar defectos del producto usando una imagen 2D conducen fácilmente a la detección fallida o falsa de defectos. Además, los inventores también han reconocido que una imagen 3D adquirida usando una cámara tridimensional (3D) tiene baja resolución y tampoco puede cubrir completamente las formas de defectos en el producto, lo que fácilmente conduce a una detección fallida o falsa de defectos. Los inventores han concebido que, dado que la imagen 2D tiene suficiente resolución, puede usarse para detectar preliminarmente defectos de producto, y, a continuación, se usa la imagen 3D para volver a detectar los defectos de producto, para eliminar las detecciones falsas (falsos positivos) y añadir los defectos fallados.

Además, los inventores se han dado cuenta de que, si se extraen características de la imagen 2D y de la imagen 3D y se sustituyen respectivamente en dos modelos de redes neuronales diferentes para su análisis, es necesario realizar dos análisis y es necesario entrenar con antelación los dos modelos de redes neuronales. Basándose en esto, los inventores han concebido fusionar la imagen 2D con la imagen 3D para obtener datos de imágenes fusionadas y sustituir los datos en un modelo de red neuronal de señal para su análisis. Por lo tanto, se reducen los recursos computacionales requeridos y se mejoran la velocidad y la eficiencia de análisis.

La Figura 1 ilustra un diagrama de flujo de un método de procesamiento de imágenes 100 de acuerdo con un aspecto de implementación de la presente solicitud. Como se muestra en la Figura 1, en la etapa 102, se obtienen una imagen bidimensional y una imagen tridimensional de un objeto. En la etapa 104, se extraen a continuación la información de luminancia y la información de profundidad de la imagen tridimensional para generar un mapa de luminancia y un mapa de profundidad. En la etapa 106, se aplica gris a la imagen bidimensional para obtener una imagen en escala de grises de un único canal, para facilitar la alineación de las características de la imagen bidimensional con las de la imagen tridimensional. En la etapa 108, después de que se obtiene el mapa de luminancia y la imagen en escala de grises, se seleccionan al menos tres puntos de ubicación iguales a partir de cada uno de la imagen en escala de grises y del mapa de luminancia. En la etapa 110, se calcula una matriz de transformación de coordenadas entre la imagen en escala de grises y el mapa de luminancia basándose en las coordenadas de los al menos tres puntos de ubicación iguales en cada uno de la imagen en escala de grises y el mapa de luminancia.

Debe entenderse por los expertos en la materia que el mismo punto característico en imágenes adquiridas usando dos cámaras diferentes tiene diferentes coordenadas de puntos de imagen. Dado que las formas de producto en imágenes tomadas usando cámaras tienen la misma planitud y paralelismo, una transformación de coordenadas entre las coordenadas de puntos de imagen de las diferentes cámaras es una transformación afín, con la siguiente ecuación de transformación de matrices:

$$\begin{bmatrix} qx \\ qy \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} px \\ py \\ 1 \end{bmatrix}$$

donde (py, py) representa las coordenadas de puntos de imagen de una cámara, (qx, qy) representa las coordenadas de puntos de imagen de la otra cámara y $a_1, b_1, c_1, a_2, b_2, y c_2$ representa coeficientes de transformación en la matriz de transformación de coordenadas. Tal transformación afín tiene seis coeficientes de transformación, es decir, seis grados de libertad. Por lo tanto, se requieren tres grupos de puntos para calcular los coeficientes de transformación en la matriz de transformación de coordenadas.

En la etapa 112, después de que se calcula la matriz de transformación de coordenadas, la imagen bidimensional se alinea con el mapa de luminancia y el mapa de profundidad de acuerdo con la matriz de transformación de coordenadas, para obtener datos de imágenes fusionadas. En la presente invención, se aplica gris a la imagen bidimensional para obtener la imagen en escala de grises de un único canal. A continuación, se resuelve la matriz de transformación de coordenadas entre la imagen bidimensional y la imagen tridimensional. Las características de la imagen bidimensional se alinean con las de la imagen tridimensional, de modo que los datos de características de la imagen bidimensional se alinean con los de la imagen tridimensional, para obtener los datos de la imagen fusionada. De esta manera, se implementa la fusión de imágenes de la imagen bidimensional con la imagen tridimensional. Tales datos de imágenes fusionadas pueden proporcionar más información para el análisis del modelo de aprendizaje profundo posterior, obteniendo de esta manera un resultado de detección más preciso.

La Figura 2 ilustra un diagrama de flujo de un método de procesamiento de imágenes de acuerdo con una realización de la presente solicitud. Como se muestra en la Figura 2, en la etapa 202, se obtiene una imagen tridimensional de un objeto. En la etapa 206, se extraen información de luminancia e información de profundidad de la imagen tridimensional para generar un mapa de luminancia y un mapa de profundidad, como se muestra en las Figuras 3b y 3c. A continuación, en la etapa 210, los valores de profundidad en el mapa de profundidad se normalizan para generar un mapa de profundidad normalizado. En algunas realizaciones, los valores de profundidad en el mapa de profundidad pueden normalizarse de acuerdo con la siguiente fórmula:

$$d_s = \frac{d - d_{\min}}{d_{\max} - d_{\min}}$$

donde J representa un valor de profundidad de un punto de coordenadas en el mapa de profundidad, $d_{\min}$ representa un valor de profundidad mínimo en el mapa de profundidad, $d_{\max}$ representa un valor de profundidad máximo en el mapa de profundidad y d_s representa un valor de profundidad normalizado del punto de coordenadas en el mapa de profundidad. Los valores de profundidad en el mapa de profundidad generado en la etapa 206 pueden ser muy grandes, por ejemplo, diez mil o más, por lo tanto, por medio de normalización, los valores de profundidad se pueden limitar entre 0 y 1, lo que facilita el cálculo y el procesamiento posteriores.

En la etapa 204, se obtiene una imagen bidimensional del objeto, como se muestra en la Figura 3a. A continuación, en la etapa 208, se aplica gris a la imagen en color bidimensional con tres canales (RGB) mediante ponderación para obtener una imagen en escala de grises de un único canal. Específicamente, un valor de escala de grises se puede calcular de acuerdo con la siguiente fórmula:

$$p = w_1 R + w_2 G + w_3 B$$

donde p representa un valor en escala de grises de un punto de coordenadas, R representa un valor de rojo del punto de coordenadas, w_1 representa un peso del valor de rojo, G representa un valor de verde del punto de coordenadas, w_2 representa un peso del valor de verde, B representa un valor de azul del punto de coordenadas y w_3 representa un peso del valor de azul. Las investigaciones muestran que los ojos humanos son más sensibles al verde, seguido del rojo, y son menos sensibles al azul. Por lo tanto, en algunas realizaciones, puede establecerse $w_2 > w_1 > w_3$. En algunas realizaciones, w_1 puede establecerse a 0,299, w_2 puede establecerse a 0,587 y w_3 puede establecerse a 0,114. Esto se debe a que los experimentos y las teorías demuestran que usando tales ajustes de peso se pueden obtener escalas de grises psicológicas que mejor se adaptan a la percepción de los ojos humanos. Obtener el valor de escala de grises usando un método de media ponderada se implementa fácilmente usando algoritmos con una cantidad de cálculo relativamente pequeña. Además, la información de rojo, verde y azul de la imagen se conserva completamente. Debe entenderse que, la aplicación de gris se puede realizar usando otros métodos, tales como un método de componentes, un método de máximos, un método de promedio y un método de corrección gamma.

Debe entenderse que, como se muestra en la Figura 2, la etapa 202, la etapa 206 y la etapa 210 pueden realizarse antes, después o cuando se realizan la etapa 204 y la etapa 208. En la etapa 212, después de que se completan las etapas 206 y 208, se pueden seleccionar al menos los tres puntos de ubicación iguales con características de cada uno de la imagen en escala de grises y del mapa de luminancia. Estos puntos con características son ubicaciones particulares en patrones que están dentro de la imagen y pueden obtenerse mediante el análisis de los patrones. Por lo tanto, la selección de los tres puntos de ubicación iguales con características puede implementarse usando algoritmos de inteligencia artificial. Las Figuras 3a a 3d ilustran una realización de detección de defectos de soldadura de un pasador de sellado de batería de litio, de acuerdo con la presente solicitud. En esta realización, en las figuras, dos puntos centrales de ángulos entre una soldadura recta y una soldadura circular, y el punto central en el pasador de sellado pueden seleccionarse como los tres puntos de ubicación iguales con características. Los tres puntos de ubicación iguales no son colineales, para cumplir con los requisitos de cálculo de coeficientes de transformación en una matriz de transformación de coordenadas.

En algunas realizaciones preferidas, un ángulo interior máximo de un triángulo formado por los tres puntos de ubicación iguales no es mayor que 160° . Cuanto mayor sea el ángulo interior máximo del triángulo formado por los tres puntos de ubicación iguales, más tenderán los tres puntos de referencia a ser colineales. Pequeños errores de muestreo pueden provocar que los tres puntos de ubicación iguales tiendan a ser colineales, de modo que la matriz de transformación de coordenadas obtenida basándose en las coordenadas de los tres puntos de ubicación iguales en cada uno de las imágenes en escala de grises y el mapa de luminancia sea ineficaz. Por lo tanto, el ángulo interior máximo del triángulo formado por los tres puntos de ubicación iguales se establece no mayor que 160° , lo que mejora la fiabilidad de la matriz de transformación de coordenadas. Cuando los tres puntos de ubicación iguales no cumplen el requisito de que los puntos no sean colineales o el requisito de que el ángulo interior máximo del triángulo formado no sea mayor que 160° , se pueden seleccionar más puntos de ubicación, siempre que tres puntos de referencia de los puntos de ubicación iguales seleccionados cumplan cualquiera de los requisitos anteriores.

En la etapa 214, se calcula una matriz de transformación de coordenadas entre la imagen en escala de grises y el mapa de luminancia basándose en las coordenadas de los tres puntos de ubicación iguales en cada uno de la imagen en escala de grises y el mapa de luminancia. La etapa 214 es similar a la etapa 110 en la Figura 1, y el contenido relacionado no se repetirá en el presente documento.

En la etapa 216, después de que se calcula la matriz de transformación de coordenadas, el mapa de luminancia y el mapa de profundidad normalizada se alinean con la imagen bidimensional basándose en la matriz de transformación de coordenadas. De esta manera, la alineación de datos de los puntos de ubicación iguales se puede implementar en la dimensión de píxeles de la imagen bidimensional y, por lo tanto, se garantiza la efectividad y precisión de la fusión de información.

Finalmente, en la etapa 218, el mapa de luminancia transformado por coordenadas y el mapa de profundidad se apilan en canales con la imagen en escala de grises, para formar una imagen fusionada que tiene información de tres canales, como se muestra en la Figura 3d. Además, debe entenderse que únicamente es necesario alinear la información de la imagen bidimensional con la de la imagen tridimensional en coordenadas, y almacenar información de color, información de luminancia e información de profundidad, etc. para cada coordenada en la imagen, en lugar de sintetizar la imagen fusionada que contiene la información de los tres canales. En algunas otras realizaciones, la información de color, la información de luminancia y la información de profundidad pueden fusionarse, como alternativa, mediante ponderación para obtener una imagen que tiene información de un único canal o de múltiples canales.

El método de procesamiento de imágenes de la presente solicitud es aplicable a una situación en la que se requieren tanto una imagen bidimensional como una imagen tridimensional para la detección de defectos. Durante la producción de baterías, los procedimientos de detección de rayones en la superficie de la batería, defectos en la soldadura de la cubierta superior de la batería y defectos en la soldadura del pasador de sellado son todos importantes para garantizar la seguridad de una batería después de salir de la fábrica. La imagen bidimensional no puede reflejar directamente la información de profundidad de las partes relacionadas de la batería y, por lo tanto, no se puede implementar una formación de imágenes efectiva de las características sensibles a la profundidad, mientras que la imagen tridimensional tiene baja resolución y no puede cubrir todas las características objetivo. Por lo tanto, los rayones en la superficie de la batería, los defectos en la soldadura de la cubierta superior de la batería y los defectos en la soldadura del pasador de sellado se pueden detectar usando el método de procesamiento de imágenes de acuerdo con la presente solicitud, para obtener datos de imágenes fusionadas para la detección de características correspondientes. Las Figuras 3a a 3d ilustran una realización de detección de defectos de soldadura de un pasador de sellado de batería de litio, de acuerdo con la presente solicitud. Sin embargo, debe entenderse que el objeto en el método de procesamiento de imágenes de la presente solicitud puede ser, como alternativa, una superficie de batería, una soldadura de la cubierta superior de la batería, etc. En la realización ilustrada en las Figuras 3a a 3d, la imagen bidimensional de la Figura 3a y la imagen tridimensional correspondiente a las Figuras 3b y 3c son imágenes adquiridas en diferentes posiciones de una línea de producción. Sin embargo, en algunas otras realizaciones, la imagen bidimensional y la imagen tridimensional pueden ser imágenes adquiridas en la misma posición de una línea de producción.

La Figura 4 ilustra un diagrama de flujo de un método de detección de apariencia de acuerdo con la presente solicitud. Como se muestra en la Figura 4, en la etapa 402, se obtienen datos de imágenes fusionadas de una pluralidad de objetos de acuerdo con el método de procesamiento de imágenes de acuerdo con la presente solicitud. En la etapa 404, se obtienen los resultados de determinación de defectos de la pluralidad de objetos. Los resultados de determinación de defectos obtenidos en la etapa 404 se determinan por inspectores visuales. A continuación, en la etapa 406, los datos de imágenes fusionadas y los resultados de determinación de defectos de la pluralidad de objetos se usan para entrenar iterativamente un modelo de algoritmo de aprendizaje profundo y actualizar el modelo de algoritmo de aprendizaje profundo. En la etapa 408, después de que se obtiene un modelo de algoritmo de aprendizaje profundo útil a través del entrenamiento, se adquieren una imagen bidimensional y una imagen tridimensional de un objeto objetivo. En la etapa 410, la imagen bidimensional y la imagen tridimensional del objeto objetivo se fusionan de acuerdo con el método de procesamiento de imágenes descrito anteriormente, para obtener datos de imágenes fusionadas del objeto objetivo. Finalmente, en la etapa 412, los datos de imágenes fusionadas del objeto objetivo se

introducen en un modelo de algoritmo de aprendizaje profundo actualizado para obtener un resultado de determinación de defectos del objeto objetivo.

En la realización ilustrada en la Figura 4, el resultado de determinación de defectos del objeto objetivo se obtiene sustituyendo los datos de imágenes fusionadas del objeto objetivo en el modelo entrenado. Por lo tanto, se mejora la precisión de la detección y se reducen los fallos de detección y las detecciones falsas (falsos positivos) de defectos. Al mismo tiempo, dado que los datos de imágenes fusionadas se sustituyen en un modelo para el cálculo y el análisis, no hay necesidad de construir dos modelos o sustituir respectivamente datos de diferentes imágenes en los dos modelos para el cálculo y el análisis, ahorrando de esta manera recursos informáticos y aumentando la velocidad de detección. Además, dado que se usan únicamente datos de imágenes fusionadas, únicamente es necesario entrenar con antelación un modelo en lugar de dos.

En algunas realizaciones, los datos de imágenes fusionadas del objeto objetivo descrito anteriormente incluyen información de escala de grises, información de luminancia e información de profundidad del objeto objetivo, y el modelo de algoritmo de aprendizaje profundo descrito anteriormente incluye un canal de entrada de datos de escala de grises, un canal de entrada de datos de luminancia y un canal de entrada de datos de profundidad. En esta realización, el modelo de algoritmo de aprendizaje profundo puede incluir tres canales de entrada de datos, y el modelo detecta, basándose en la información de escala de grises de entrada, la información de luminancia y la información de profundidad, si hay un defecto.

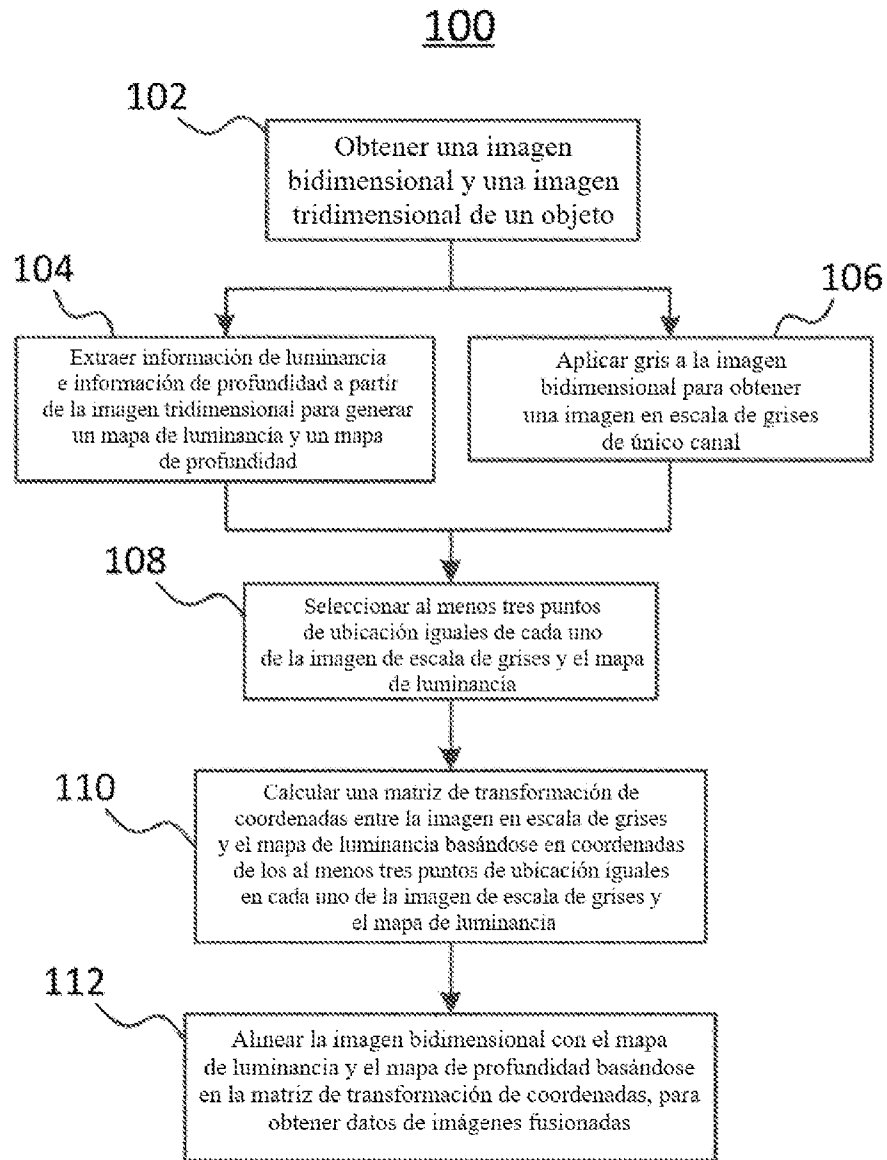
En algunas realizaciones, el modelo de algoritmo de aprendizaje profundo descrito anteriormente incluye un canal de entrada de datos de rojo, un canal de entrada de datos de verde, un canal de entrada de datos de azul, un canal de entrada de datos de luminancia y un canal de entrada de datos de profundidad. En esta realización, el modelo de algoritmo de aprendizaje profundo puede incluir cinco canales de entrada de datos, y el modelo detecta, basándose en la información de color de entrada, la información de luminancia y la información de profundidad, si hay un defecto. Debe entenderse que, en la práctica, los expertos en la materia pueden elegir el número de canales de datos del modelo de acuerdo con los requisitos reales.

La presente invención no se limita a las realizaciones específicas divulgadas en el presente documento, sino que incluye todas las soluciones técnicas que caen dentro del alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un método de procesamiento de imágenes implementado por ordenador, que comprende:

- 5 obtener una imagen bidimensional y una imagen tridimensional de un objeto (102; 408), en donde el objeto se selecciona de uno cualquiera de los siguientes: una superficie de batería, una soldadura de la cubierta superior de la batería y un pasador de sellado de la batería;
extraer información de luminancia e información de profundidad a partir de la imagen tridimensional para generar un mapa de luminancia y un mapa de profundidad para el objeto (104);
10 aplicar gris a la imagen bidimensional para obtener una imagen en escala de grises de un único canal (106);
seleccionar al menos los tres puntos de ubicación iguales de cada uno de la imagen en escala de grises y del mapa de luminancia (108, 212);
calcular una matriz de transformación de coordenadas entre la imagen en escala de grises y el mapa de luminancia basándose en las coordenadas de los al menos tres puntos de ubicación iguales en cada uno de la imagen en escala de grises y el mapa de luminancia (110, 214); y
15 alinear la imagen bidimensional con el mapa de luminancia y el mapa de profundidad basándose en la matriz de transformación de coordenadas, para obtener datos de imágenes fusionadas (112; 410), en donde los datos de imágenes fusionadas comprenden información de color de la imagen bidimensional y la información de luminancia e información de profundidad de la imagen tridimensional,
20 en donde los datos de imágenes fusionadas del objeto se introducen en un modelo de algoritmo de aprendizaje profundo entrenado para obtener un resultado de determinación de defectos del objeto (412).
2. El método de procesamiento de imágenes de acuerdo con la reivindicación 1, en donde los al menos tres puntos de ubicación iguales son puntos del objeto con características (212).
- 25 3. El método de procesamiento de imágenes de acuerdo con la reivindicación 1 o 2, en donde los al menos tres puntos de ubicación iguales comprenden tres puntos de referencia no colineales.
4. El método de procesamiento de imágenes de acuerdo con la reivindicación 3, en donde un ángulo interior máximo de un triángulo formado por los tres puntos de referencia no es mayor que 160°.
- 30 5. El método de procesamiento de imágenes de acuerdo con una cualquiera de las reivindicaciones 1 a 4, que comprende:
alinear el mapa de luminancia y el mapa de profundidad con la imagen bidimensional basándose en la matriz de transformación de coordenadas usando las coordenadas de la imagen bidimensional como referencia; y
35 usar, como información de tres canales respectivamente, la información de escala de grises de la imagen en escala de grises, y la información de luminancia y la información de profundidad que están alineadas con las coordenadas en la imagen bidimensional, para generar una imagen fusionada que contiene la información de los tres canales.
- 40 6. El método de procesamiento de imágenes de acuerdo con una cualquiera de las reivindicaciones 1 a 5, antes de que la imagen bidimensional se alinee con el mapa de luminancia y el mapa de profundidad, los valores de profundidad en el mapa de profundidad se normalizan para generar un mapa de profundidad normalizado (210).
- 45 7. Un medio de almacenamiento no volátil, que almacena instrucciones, en donde las instrucciones, cuando son ejecutadas por un procesador, hacen que el procesador realice el método de procesamiento de imágenes de acuerdo con una cualquiera de las reivindicaciones 1 a 6.

*FIG. 1*

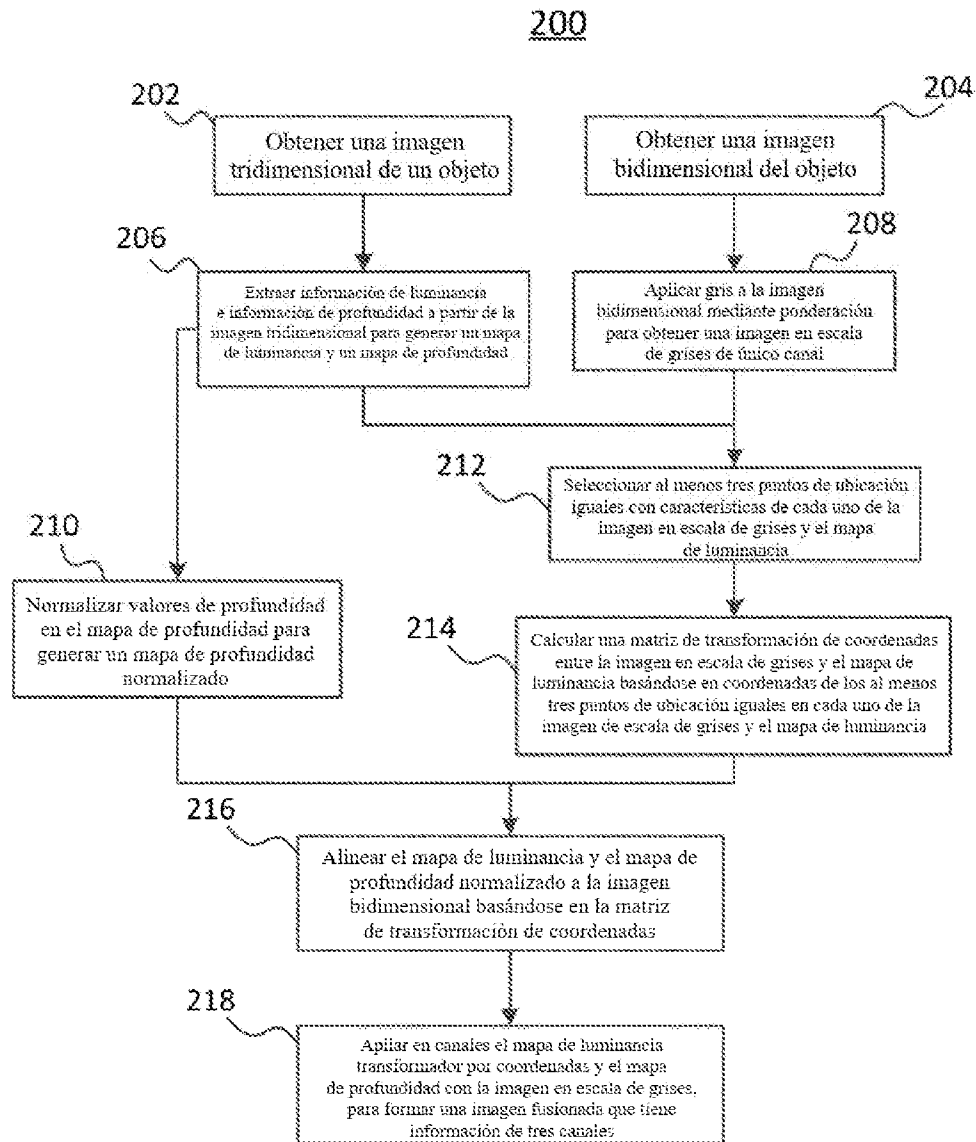


FIG. 2

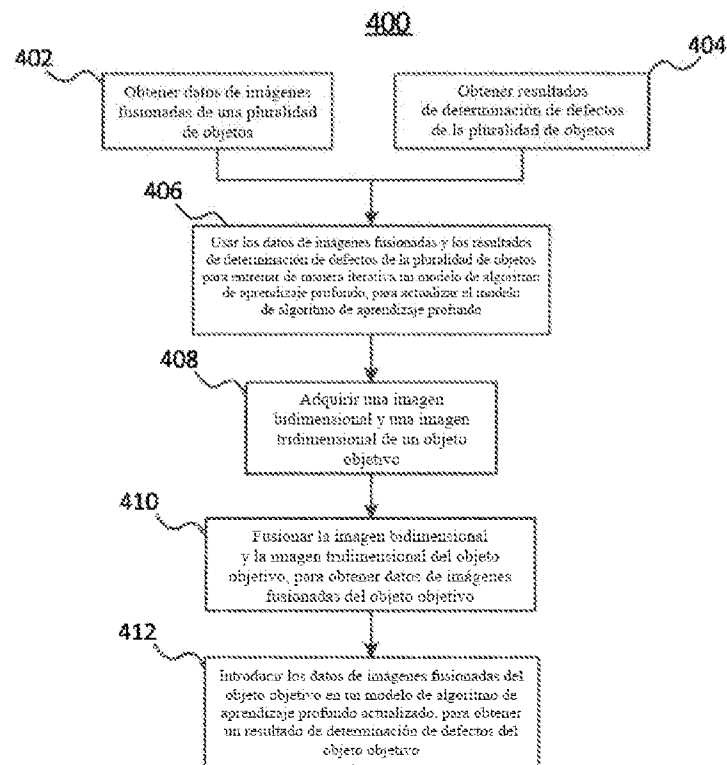
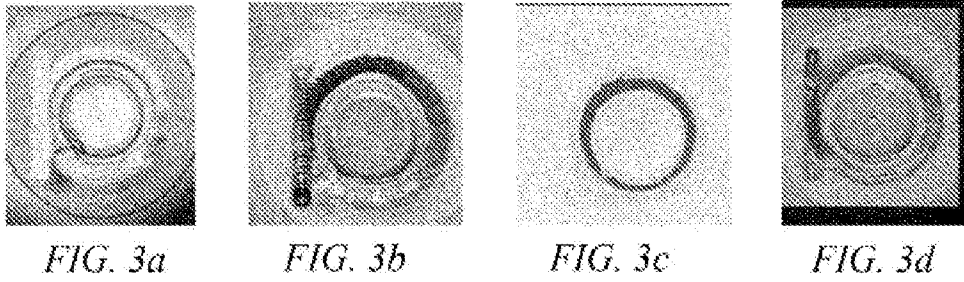


FIG. 4