

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 745 701**

51 Int. Cl.:

**G06F 11/07** (2006.01)

**G06F 11/30** (2006.01)

**G06N 7/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.05.2013 PCT/GB2013/051237**

87 Fecha y número de publicación internacional: **21.11.2013 WO13171474**

96 Fecha de presentación y número de la solicitud europea: **14.05.2013 E 13723913 (3)**

97 Fecha y número de publicación de la concesión europea: **17.07.2019 EP 2850522**

54 Título: **Identificación de estado anómalo de sistema**

30 Prioridad:

**15.05.2012 GB 201208542**

**10.10.2012 GB 201218216**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**03.03.2020**

73 Titular/es:

**UNIVERSITY OF LANCASTER (100.0%)  
Bailrigg, Lancaster  
Lancashire LA1 4YW, GB**

72 Inventor/es:

**ANGELOV, PLAMEN**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 745 701 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

Identificación de estado anómalo de sistema

La presente invención se refiere a procedimientos, aparatos y programas informáticos para identificar estados anómalos en sistemas, y en particular sistemas multisensor en tiempo real.

5 Hay una gran cantidad de sistemas en los que es deseable poder identificar de manera confiable el comportamiento anómalo. El comportamiento anómalo del sistema puede ser indicativo de una variedad de comportamientos diferentes del sistema, como una operación incorrecta, operación subóptima o simplemente una diferencia en la operación fuera de la norma. En sistemas complicados con muchas partes, puede ser difícil identificar exactamente qué parte está causando el comportamiento anómalo. También, en algunos sistemas complicados, puede ser muy difícil detectar que el comportamiento general del sistema es anómalo ya que el efecto de una o más partes en el sistema general puede ser muy sutil.

10 Aún más, puede ser muy difícil identificar comportamientos anómalos en sistemas complicados en tiempo real debido al gran número de partes y/o combinaciones de partes que pueden necesitar ser evaluadas. Si bien dichos sistemas se pueden evaluar en un modo de tiempo no real, es decir, recolectando grandes cantidades de datos y analizándolos, este enfoque puede ser de poca utilidad cuando se requiere la detección de anomalías en tiempo real, por ejemplo por razones de salud y seguridad.

15 La detección de novedad, tanto anomalía como valores atípicos, es una tarea difícil incluso si se realiza fuera de línea y es aún más difícil de hacer en tiempo real sin recursos de procesamiento de datos muy significativos. Sin embargo, en muchas aplicaciones sería preferible poder identificar de manera confiable anomalías en tiempo real y se han adoptado varios enfoques. Tradicionalmente, los enfoques estadísticos se han basado en la estimación de la densidad de datos y la estimación de la densidad del núcleo es uno de los enfoques más populares para esta tarea, un ejemplo del cual se describe en T. Hastie, R. Tibshirani y J. Friedman, Los elementos del aprendizaje estadístico: Minería de datos, Inferencia y predicción. Heidelberg, Alemania: Springer Verlag, 2001.

20 "Detección y diagnóstico de fallas en sistemas de fabricación: un enfoque de modelo de comportamiento", Fabricación integrada por computadora, 1990, Actas de la Segunda Conferencia Internacional de Rensselaer en Troy, NUEVA YORK, ESTADOS UNIDOS, 21-23 de mayo de 1990, páginas 252-259, El documento XP010016968 describe la detección y diagnóstico de fallas en línea en sistemas de fabricación automatizados basados en modelos de comportamiento de componentes del sistema desarrollados mientras se diseña el sistema. Los modelos caracterizan la respuesta de los dispositivos en el sistema a señales de entrada arbitrarias en el rango de parámetros operativos. Se captura el flujo esperado de señales a través del sistema desde las entradas de control hasta las salidas del sensor. Las señales de respuesta esperadas del sistema se pueden comparar en línea, en tiempo real con las señales reales del sensor del sistema. El diagnóstico de fallas se logra manteniendo un conjunto actual de supuestos operativos que identifican los componentes del sistema que podrían causar desviaciones del comportamiento esperado.

25 El documento US 2008/071501 describe procedimientos y aparatos para diagnosticar fallas en un sistema monitoreado que es monitoreado por sensores, por ejemplo para la detección de fugas en un tubo de caldera. Se genera un modelo empírico para un componente objetivo del sistema monitoreado. El modelo empírico está entrenado con una fuente de datos históricos que contiene observaciones de ejemplo de los sensores. Sustancialmente, las estimaciones en tiempo real se generan en función de datos instrumentados correspondientes al componente objetivo. Las estimaciones sustancialmente en tiempo real se comparan y diferencian con lecturas instrumentadas de los sensores para proporcionar valores residuales. Los valores residuales se analizan para detectar las fallas y determinar una ubicación de las fallas en el sistema monitoreado.

30 El documento US 2007/255442 describe un analizador de fallas de procedimiento que incluye una parte de edición de datos de procedimiento para extraer una cantidad característica de procedimiento de datos de procedimiento en una serie temporal almacenada en una parte de almacenamiento de datos de procedimiento. Una parte de almacenamiento de datos de la regla de análisis de fallas almacena una regla de análisis de fallas para realizar la detección de fallas en un producto fabricado en un sistema de fabricación y en equipos de fabricación, basado en la cantidad característica del procedimiento. Una parte que determina la falla determina la existencia/ausencia de una falla en un producto y en el equipo de fabricación en función de la cantidad característica del procedimiento. Se utiliza un modelo de regresión de mínimos cuadrados parciales (PLS) como modelo de estimación utilizado para la regla de análisis de fallas.

35 "Estimación de densidad para estadísticas y análisis de datos", 26 de julio de 2006, Monografías sobre estadística y probabilidad aplicada, Chapman y Hall, XP055363517, describe varios procedimientos de estimación de densidad.

Por lo tanto, existe la necesidad de un enfoque para la identificación de anomalías que sea manejable computacionalmente en tiempo real y que proporcione una identificación confiable de estados anómalos de un sistema.

40 Un primer aspecto de la invención proporciona un procedimiento en tiempo real para identificar un estado anómalo de un sistema de acuerdo con la reivindicación 1.

Por lo tanto, como el procedimiento utiliza cálculos recursivos, la carga computacional es muy baja. Los cálculos

recursivos utilizan solo un elemento de datos actual y datos almacenados que resumen, de manera estadística, la operación pasada del sistema. Por lo tanto, el procedimiento no necesita procesar todo, o una gran cantidad de, elementos de datos pasados o históricos.

5 La carga computacional es lo suficientemente baja como para que el procedimiento pueda operar en tiempo real, incluso para sistemas complicados que tienen cientos o incluso miles de salidas de sensor diferentes a procesar.

El procedimiento se puede aplicar a sistemas que incluyen una pluralidad de sensores, cada uno de los cuales emite datos de series temporales. Los flujos de datos de series temporales pueden relacionarse con una propiedad diferente del sistema.

10 Se puede calcular recursivamente una pluralidad de propiedades estadísticas. La pluralidad de propiedades estadísticas puede incluir la media actual de la densidad de datos actual y la varianza actual de la densidad de datos actual.

El elemento de datos actual se puede determinar para indicar un estado anómalo basado en el grado de diferencia entre la estimación de la densidad de datos actual y la media actual de la densidad de datos actual.

15 El elemento de datos actual se puede determinar para indicar un estado anómalo si la diferencia entre la estimación de la densidad de datos actual y la media actual de la densidad de datos actual es mayor que un número de desviaciones típicas de la media actual de la densidad de datos actual, por ejemplo uno, dos o preferentemente tres desviaciones típicas.

20 El procedimiento puede comprender además aplicar una prueba adicional o una pluralidad de pruebas adicionales para determinar si el elemento de datos actual indica un estado anómalo. La prueba o pruebas adicionales se pueden seleccionar de: una prueba basada en el tiempo; y una prueba basada en eventos. La prueba basada en el tiempo puede determinar si se ha producido un estado anómalo potencialmente identificado dentro de un período de tiempo específico desde un estado anómalo potencialmente identificado anterior. La prueba basada en eventos puede determinar si se ha producido un estado anómalo potencialmente identificado dentro de una secuencia u orden específico en relación con un estado anómalo potencialmente identificado anterior.

25 El procedimiento puede comprender además: calcular recursivamente un valor medio del elemento de datos utilizando el elemento de datos actual; y/o calcular recursivamente un producto escalar del elemento de datos sobre o consigo mismo.

El procedimiento puede comprender además: usar el valor medio del elemento de datos y el producto escalar del elemento de datos para calcular recursivamente la estimación de la densidad de datos actual.

30 El sistema puede incluir una pluralidad de sensores, cada uno de los cuales emite elementos de datos de series temporales relacionados con una propiedad diferente del sistema. El procedimiento puede aplicarse a los elementos de datos actuales recibidos respectivamente de cada uno de la pluralidad de sensores.

La determinación puede comprender determinar a partir de la al menos una propiedad estadística si un subconjunto de elementos de datos actuales de la pluralidad de elementos de datos indica un estado anómalo del sistema.

35 La determinación puede comprender determinar a partir de la al menos una propiedad estadística si todos los elementos de datos actuales de la pluralidad de elementos de datos indican un estado anómalo del sistema.

El sistema puede ser un sistema de vídeo. Un estado anómalo del sistema de vídeo puede corresponder a la detección de un nuevo objeto o un cambio en un objeto dentro del campo de visión del sistema de vídeo.

40 El sensor puede ser un sensor de imagen. Los datos de la serie temporal pueden ser datos de vídeo en color, como datos de vídeo rojos, verdes y azules, o datos de imagen en gris o en escala de grises.

La propiedad puede ser una subregión de una trama de datos de vídeo. Cada subregión o contenedor de una trama de datos de vídeo se puede analizar por separado para identificar estados anómalos del sistema de vídeo.

Un segundo aspecto de la invención proporciona un aparato de procesamiento de datos para identificar un estado anómalo de un sistema en tiempo real según la reivindicación 13.

45 Un tercer aspecto de la invención proporciona un sistema de acuerdo con la reivindicación 14.

El sistema puede ser cualquier sistema industrial, eléctrico o electromecánico. El sistema puede ser una máquina, un aparato, un vehículo, un motor, una planta, un fragmento de planta, una pieza de maquinaria, un dispositivo eléctrico o electrónico o similar.

50 Las realizaciones de la invención se describirán ahora en detalle, a modo de ejemplo solamente, y con referencia a los dibujos adjuntos, en los que:

la figura 1 muestra una representación gráfica de puntos de datos en un espacio de datos bidimensional;  
 la Figura 2 muestra un diagrama de bloques esquemático de un sistema de control de aeronave de acuerdo con la invención e incluye un aparato de procesamiento de datos de acuerdo con una primera realización de la invención;  
 5 la Figura 3 muestra una representación gráfica de una estructura de datos utilizada por la primera realización de la invención;  
 la Figura 4 muestra un diagrama de flujo que ilustra un procedimiento de procesamiento de datos de la primera realización de la invención;  
 la Figura 5 muestra un diagrama de flujo del procedimiento que ilustra una etapa de detección de anomalías del procedimiento ilustrado en la Figura 4 con mayor detalle;  
 10 la Figura 6 muestra una representación gráfica de la detección de anomalías según el procedimiento de la primera realización;  
 la Figura 7 muestra un diagrama de bloques esquemático de un sistema de vídeo de acuerdo con la invención e incluye un aparato de procesamiento de datos de acuerdo con una segunda realización de la invención;  
 15 la Figura 8 muestra una representación gráfica de trama de datos de vídeo que está compuesta por una pluralidad de contenedores de datos de vídeo;  
 la Figura 9 muestra una representación gráfica de una estructura de datos utilizada por la segunda realización de la invención;  
 la Figura 10 muestra un diagrama de flujo que ilustra un procedimiento de procesamiento de datos de la segunda realización de la invención;  
 20 la Figura 11 muestra un diagrama de flujo del procedimiento que ilustra una etapa de detección de anomalías del procedimiento ilustrado en la Figura 10 con mayor detalle; y  
 la figura 12 muestra un diagrama de bloques de un dispositivo de procesamiento de datos adecuado para implementar la primera o segunda realización de la invención.

25 Los elementos similares en diferentes figuras comparten signos de referencia comunes a menos que se indique lo contrario.

Hay una amplia variedad de sistemas en los que se puede usar la presente invención. A modo de ejemplo solo, a continuación se describen dos realizaciones de la invención en el contexto de un sistema de control de aeronaves (como un Sistema de Monitoreo de Condición de Aeronaves o ACMS) y un sistema de captura de vídeo. Sin embargo, se apreciará que la invención puede aplicarse a otros tipos de sistemas tanto más complejos como menos complejos, y en una amplia variedad de campos de la industria.

La invención se refiere particularmente a la identificación de comportamientos anómalos en un sistema en tiempo real y en función del procesamiento de datos obtenidos de uno o más sensores que miden una o más propiedades del sistema. La invención puede dar lugar a varios tipos diferentes de señales de salida como resultado de identificar o 35 detectar comportamientos anómalos. Por lo tanto, la invención generalmente cae dentro del campo de las tecnologías sensibles a la condición. La invención tiene una baja carga computacional y, por lo tanto, puede identificar o detectar anomalías en tiempo real. La invención es computacionalmente altamente eficiente ya que el estado  $N + 1^{\text{ésimo}}$  de un sistema se calcula a partir del punto  $N + 1^{\text{ésimo}}$  y los datos que describen el estado  $N^{\text{ésimo}}$  del sistema solamente, sin necesidad de cálculo utilizando todos los  $N$  puntos de datos anteriores individualmente. Esto hace que la técnica sea 40 particularmente adecuada para aplicaciones en tiempo real.

Se proporciona un mecanismo adaptativo que puede identificar condiciones nuevas en sistemas complejos. Dependiendo de la naturaleza del sistema y el comportamiento anómalo, la nueva condición identificada podría ser una falla o un desencadenante de alguna acción secundaria, como la grabación de datos. Es importante destacar que no se requiere conocimiento a priori del sistema. No es necesario conocer los rangos de datos del sensor, límites de funcionamiento para los datos del sensor o el significado de los datos del sensor. Más bien, después de un período de aprendizaje no supervisado (ya sea en tiempo real, o después de haber recibido datos históricos capturados), se construye una imagen o modelo de comportamiento "normal" del sistema basado en las propiedades estadísticas de los datos.

Este comportamiento "normal" puede incluir múltiples modos de funcionamiento normales, y estos modos se descubren automáticamente. Por ejemplo, los datos del sensor de una aeronave tomarán diferentes valores normales dependiendo de la fase del vuelo (por ejemplo, despegue, crucero y aterrizaje) y cada fase de vuelo puede tener su propio modo normal que es diferente de los modos normales respectivos de las otras fases de vuelo. Se puede generar alguna forma de señal cuando se detecta un comportamiento anómalo, que, por ejemplo, puede indicar un estado de alarma o ser una señal de activación, cuando los datos del sensor caen suficientemente estadísticamente fuera de uno de los modos de operación normales aprendidos.

Antes de describir las realizaciones de la invención con mayor detalle, se discutirá la base matemática para el procedimiento de la invención. Como se ha mencionado anteriormente, el procedimiento es un procedimiento recursivo que utiliza datos que resumen el estado anterior del sistema junto con datos de entrada de uno o más sensores para el estado actual del sistema.

60 El procedimiento de la invención usa densidad de datos,  $D$ , y no una función de densidad de probabilidad (pdf),  $p$  que

es muy similar. Sin embargo, la densidad de datos,  $D$ , y pdf,  $p$ , difieren porque si bien ambos satisfacen la condición:

$$0 \leq D \leq 1 ; \quad 0 \leq p < 1 \tag{1}$$

la densidad de datos no cumple la condición:

$$\int_{-\infty}^{\infty} p(\cdot) = 1 \tag{2}$$

5 En su lugar,  $\max_x D = 1 ; \max_x p < 1$

la densidad de datos se puede definir como la inversa de la suma de las distancias totales entre todos los puntos de datos o muestras. La función específica o núcleo de la densidad de datos puede ser, por ejemplo, de tipo Cauchy porque es conveniente para cálculos recursivos:

$$D(x_k) = \frac{1}{1 + \frac{1}{k} \sum_{i=1}^k \|x_k - x_i\|^2} \tag{3}$$

10 Donde  $D$  denota la densidad de datos global en la muestra de datos  $x_k$  y donde  $k$  es el instante de tiempo actual o el índice de muestra de datos. Otras funciones como Gaussiana, Epanechnikov, etc. también son posibles, pero son más difíciles de calcular de forma recursiva.

15 La figura 1 muestra una representación gráfica de un espacio de datos bidimensional, en el que el primer tipo de datos,  $x_1$ , es la presión de un sistema en el instante  $k$  y el segundo tipo de datos,  $x_2$ , es la temperatura del sistema en el instante  $k$ . Se apreciará que un sistema puede tener menos o más dimensiones de datos (es decir, desde 1 hacia arriba, por ejemplo miles o más). Sin embargo, la figura 1 ilustra, para un sistema de dos dimensiones de datos, el estado (en términos de temperatura y presión) para diez estados del sistema ( $k = 1$  a 10). El punto 106 representa el noveno estado del sistema (es decir,  $k = 9$ ) que tiene una alta densidad de datos (es decir, está cerca de muchos otros puntos de datos del sistema) y el punto 108 representa el décimo estado del sistema (es decir,  $k = 10$ ) que tiene una baja densidad de datos (es decir, estar lejos de muchos otros puntos de datos del sistema).

20 La densidad local (es decir, en términos de grupos de puntos de datos en lugar de globalmente en términos de todos los puntos de datos) se puede introducir de una manera similar:

$$d^j(x_k) = \frac{1}{1 + \frac{1}{M_j} \sum_{i=1}^{M_j} \|x_k - x_i\|^2} \quad j = [1, N] \tag{4}$$

25 Donde  $d$  denota la densidad local en la muestra de datos  $x_k$ ,  $M_j$  es la cantidad de muestras de datos asociadas con un grupo particular y  $N$  es la cantidad de grupos.

Ambos, la densidad global y local, pueden estimarse recursivamente, lo que lleva al concepto de estimación de densidad recursiva (RDE):

$$D(x_k) = \frac{1}{1 + \|x_k - \mu_k\|^2 + \zeta_k - \|\mu_k\|^2} \tag{5}$$

30 donde  $\mu$  denota la media y  $\zeta$  denota un producto escalar de todas las muestras de datos, si una densidad global,  $D$ , es usada. Si la densidad local,  $d$ , está siendo calculada, entonces  $\mu$  denota la media y  $\zeta$  denota un producto escalar de datos de un grupo.

La estimación de datos recursivos (RDE) de la ecuación (5) puede derivarse de la ecuación (3) de la siguiente manera.

De la ecuación (3) o (4):

$$D(x_k) = \frac{1}{1 + x_k^2 - 2x_k \frac{1}{k} \sum_{i=1}^k x_i + \frac{1}{k} \sum_{i=1}^k x_i^2} \quad (6)$$

Ecuación de reagrupación (6):

$$D(x_k) = \frac{1}{1 + x_k^2 - 2x_k \frac{1}{k} \sum_{i=1}^k x_i + \left(\frac{1}{k} \sum_{i=1}^k x_i\right)^2 - \left(\frac{1}{k} \sum_{i=1}^k x_i\right)^2 + \frac{1}{k} \sum_{i=1}^k x_i^2} \quad (7)$$

$$D(x_k) = \frac{1}{1 + \left\|x_k - \frac{1}{k} \sum_{i=1}^k x_i\right\|^2 + \frac{1}{k} \sum_{i=1}^k x_i^2 - \left(\frac{1}{k} \sum_{i=1}^k x_i\right)^2} \quad (8)$$

5 La media recursivamente actualizada,  $\mu$  y producto escalar,  $\zeta$ , se presentan de la siguiente manera:

$$\mu_k = \frac{k-1}{k} \mu_{k-1} + \frac{1}{k} x_k \quad \mu_1 = x_1 \quad (9)$$

$$\Sigma_k = \frac{k-1}{k} \Sigma_{k-1} + \frac{1}{k} \|x_k\|^2 \quad \Sigma = \|x_1\|^2 \quad (10)$$

y las ecuaciones (8) a (10) conducen a la ecuación (5).

De forma similar, la RDE local (por grupo) viene dada por:

$$d^j(x_k) = \frac{1}{1 + \|x_k - \mu_k^j\|^2 + \Sigma_k - \|\mu_k^j\|^2} \quad j = [1, N] \quad (11)$$

10 Basado en el valor de la densidad de datos calculada usando las ecuaciones RDE es posible un simple, pero eficaz, procedimiento de detección de anomalías atípicas. Tomando el caso global como ejemplo, la densidad de datos,  $D(x_k)$ , se calcula en tiempo real por muestra de datos. La densidad media de datos,  $\bar{D}(x_k)$  es dada por:

$$\bar{D}(x_k) = \frac{1}{k} \sum_{i=1}^k D(x_i) \quad (12)$$

y también se puede calcular en tiempo real (usando la notación  $\bar{D}_k$  en vez de  $\bar{D}(x_k)$ ) como:

$$\bar{D}_k = \frac{k-1}{k} \bar{D}_{k-1} + \frac{1}{k} D_k \quad \bar{D}_1 = D_1 \quad (13)$$

La varianza de la densidad de datos (no de los datos originales, x):

$$\sigma_k^2 = \left\| \bar{D}_k - D_k \right\|^2 \quad (14)$$

también se puede calcular de forma recursiva utilizando:

$$\sigma_k^2 = \frac{k-1}{k} \sigma_{k-1}^2 + \frac{1}{k} (D_k - \bar{D}_k)^2 \quad \sigma_1^2 = 1 \quad (15)$$

5 donde la suma es solo para el elemento k-ésimo.

Usando las ecuaciones (13) y (15), los datos atípicos o anómalos se pueden identificar o detectar utilizando alguna regla o umbral basado en estadísticas. Por ejemplo, un punto de datos que es mayor que, por ejemplo, dos o tres desviaciones típicas ( $\sigma_k$ ) lejos de la media de la densidad de datos,  $\bar{D}_k$ , puede identificarse como una anomalía. Se prefieren tres desviaciones típicas como, para una distribución de datos gaussiana, 99 % de los datos (y mayor que 10 8/9 de los datos para una distribución arbitraria) serán menos de tres desviaciones típicas de la media.

La media y la desviación típica se pueden determinar para cada dimensión de datos de un sistema por separado, por ejemplo, para temperatura y presión en el ejemplo anterior. Entonces se puede identificar un valor atípico si la desviación típica de la presión es más de tres desviaciones típicas de la media de la densidad de datos de presión o de manera similar para la temperatura. Esto puede proporcionar una visión útil de la causa específica de la falla o 15 problema. Además o alternativamente, la media y la densidad de datos se pueden calcular utilizando todo el grupo de dimensiones de datos y valores atípicos identificados globalmente para el sistema en su conjunto. Esto puede proporcionar una mejor visión global o general del rendimiento del sistema.

Adicionalmente o alternativamente, se pueden determinar las desviaciones típicas y medias para uno o una pluralidad de subgrupos de dimensiones de datos dentro del grupo de todas las dimensiones de datos para un sistema. En este 20 caso, se usa una estimación de la densidad de datos local (d véase la ecuación (11) anterior) y se puede determinar utilizando las siguientes ecuaciones:

$$\bar{d}^j(x_k) = \frac{1}{M_j} \sum_{i=1}^{M_j} d^j(x_k) \quad (16)$$

$$\bar{d}_k^j = \frac{M_j-1}{M_j} \bar{d}_{k-1}^j + \frac{1}{M_j} d_k^j \quad \bar{d}_1^j = d_1^j \quad (17)$$

$$\left( \sigma_{M_j}^j \right)^2 = \left\| \bar{d}_k^j - d_k^j \right\|^2 \quad (18)$$

$$\left( \sigma_{M_j}^j \right)^2 = \frac{M_j-1}{M_j} \left( \sigma_{M_{j-1}}^j \right)^2 + \frac{1}{M_j} \left( \bar{d}_k^j - d_k^j \right)^2 \quad \left( \sigma_1^j \right)^2 = 1 \quad (19)$$

Donde  $M_j$  es el número de elementos de datos asociados con el subgrupo j-ésimo.

25 Habiendo descrito la base matemática del procedimiento de la invención, ahora se describirán dos sistemas de ejemplo diferentes que incorporan la invención.

Con referencia a la figura 2, se muestra un diagrama de bloques esquemático de una primera realización de un sistema 200 según la invención que incluye un aparato 202 de procesamiento de datos también según la invención. El sistema 200 es parte de un avión e incluye una pluralidad de sensores 204, 206, 208, cada uno de los cuales mide una 30 propiedad de la aeronave. Los sensores 204, 206, 208 emiten datos que se comunican a un registrador 210 de datos de vuelo también conocido coloquialmente como un registrador de datos de "caja negra". El tercer sensor 208 se proporciona como parte de un subsistema 212 de la aeronave. El subsistema puede incluir múltiples componentes,

como un servo 214. Por ejemplo, el subsistema 212 puede ser parte del subsistema de control de vuelo de la aeronave y el servo 214 puede funcionar para ajustar las aletas de la aeronave. Aunque se ilustran tres sensores en la Figura 2, se apreciará que se proporcionará un número mucho mayor de sensores en la práctica. Por ejemplo, un avión comercial típico puede tener en cualquier parte de la región de dos a tres mil sensores diferentes.

5 El aparato 202 de procesamiento de datos incluye una unidad 220 de procesamiento de datos que incluye una o más unidades centrales de procesamiento, memoria local y otro hardware como se encuentra típicamente en una computadora programable electrónica convencional de uso general. La unidad 220 de procesamiento de datos está en comunicación con un almacén 222 de datos que puede tener la forma de una base de datos. La unidad 220 de procesamiento de datos tiene una pluralidad de salidas 224, 226, 228. Una primera salida 224 está en comunicación con una parte adicional del sistema 200, tal como una unidad 230 de visualización en la cabina del avión. El sistema 10 200 puede incluir una parte 232 adicional, tal como otro dispositivo informático o de procesamiento de datos al que la unidad 220 de procesamiento de datos puede suministrar una señal de salida. Finalmente, una tercera salida 228 está en comunicación con el subsistema 212 y, en particular, permite una ruta de señal al servo 214 de ala. Por lo tanto, la unidad 220 de procesamiento de datos puede emitir varias señales diferentes a diferentes partes del sistema para controlar o interactuar con otras partes del sistema 200.

La unidad 220 de procesamiento de datos almacena localmente el código del programa informático para implementar un procedimiento de procesamiento de datos también de acuerdo con un aspecto de la invención y que se describirá con mayor detalle a continuación. Por ejemplo, el código del programa informático puede almacenarse en forma compilada en una ROM local. También se proporciona una RAM local para proporcionar memoria de trabajo y almacenamiento para la unidad de procesamiento de datos con el fin de ejecutar las instrucciones del programa informático.

La figura 3 ilustra una estructura 300 de datos, en forma de mesa, que almacena varios elementos de datos utilizados por el procedimiento de procesamiento de datos ilustrado en la Figura 4. La estructura 300 de datos incluye una pluralidad de filas, por ejemplo, fila 302, y cada fila corresponde a un paso de tiempo (indexado k) en el que la unidad 220 de procesamiento de datos recibe un punto de datos actual desde los sensores 204, 206 y 208. La estructura de datos proporciona campos para almacenar valores calculados de la media de diferentes propiedades de la aeronave detectada por el sensor, como la velocidad (V) 304, altitud (A) 306 e inclinación (P) 308. La tabla también tiene campos para almacenar elementos de datos que representan los valores calculados recursivamente del producto escalar del elemento de datos en sí mismo 310, un valor 312 de densidad de datos estimado recursivamente, un valor 314 medio de densidad de datos y una desviación 316 típica de densidad de datos. El cálculo de estos elementos de datos se describe con mayor detalle a continuación. En una realización en la que la detección de anomalías se realiza en una dimensión de datos por base de dimensión de datos, luego la tabla 300 almacena los elementos 316, 314, 312 y 310 de datos estadísticos para cada dimensión de datos, por ejemplo, para cada uno de V, A y P.

La figura 4 muestra un diagrama de flujo que ilustra un procedimiento 400 de procesamiento de datos de identificación o detección de anomalías. El procedimiento 400 comienza en la etapa 402 con la inicialización del software y antes de recibir cualquier elemento de datos de los sensores, en  $k = 0$ . En la etapa 404, la unidad de procesamiento de datos recibe un primer conjunto de valores 406 de datos desde los sensores 204 a 208 en  $k = 1$ . En el ejemplo ilustrado, los datos proporcionados por los sensores se relacionan con las propiedades de la velocidad, altitud e inclinación de la aeronave. En la etapa 408, un valor medio de los datos del sensor se actualiza de forma recursiva utilizando la ecuación 9 para cada elemento de datos y se almacena en los campos 304, 306 y 308, respectivamente, de la tabla 300. Se apreciará que para el primer elemento de datos,  $k = 1$ , la media simplemente se establece en el primer valor de datos. Luego, en la etapa 410, el producto escalar se calcula de forma recursiva utilizando la ecuación 10 para cada uno de V, A y P y se almacena en el campo 310 de la tabla 300. En otras realizaciones, la covarianza y los otros datos estadísticos pueden calcularse a nivel de todo el sistema para todo el grupo de parámetros del sistema o para un subsistema de parámetros. Como podrá apreciarse, para el primer punto de datos,  $k = 1$ , la covarianza se establece en el cuadrado del primer punto de datos, como se indica en la ecuación 10 anterior.

Luego, en la etapa 412, la densidad de datos se estima recursivamente utilizando la ecuación 11. Por lo tanto, en la etapa 412, la densidad 312 de datos se estima de forma recursiva para cada una de velocidad, altitud e inclinación utilizando el producto escalar respectivo y los valores medios almacenados en la tabla 300 y se escriben en el campo 50 312.

A continuación, en la etapa 414, la densidad media de datos se calcula de forma recursiva utilizando la ecuación 13 y los resultados almacenados en el campo 314. Como se indica en la ecuación 13, para el primer elemento de datos,  $k = 1$ , la densidad 314 de datos media simplemente se establece en la densidad 312 de datos estimada. Luego, en la etapa 416, la varianza de la densidad de datos se calcula de forma recursiva utilizando la ecuación 15 y la desviación típica (raíz cuadrada de la varianza) se almacena en el campo 316 de la tabla 300. Como se indica en la ecuación 15, para el primer elemento de datos,  $k = 1$ , la varianza se establece igual a 1. En la etapa 418, se pueden aplicar varias pruebas utilizando los datos estadísticos calculados de forma recursiva almacenados en la tabla 300 para determinar si se ha identificado o detectado un estado anómalo del sistema. Donde solo hay un único punto de datos presente, no se puede identificar un comportamiento anómalo y, por lo tanto, el procesamiento pasa a la etapa 422 en la que el índice de datos k se incrementa y el procesamiento retorna, como se ilustra en la línea 424 de flujo de procedimiento, a la etapa 404 en la que se recibe un siguiente conjunto de datos de los sensores.



El procesamiento continúa como se describe anteriormente, para  $k = 2$ . Sin embargo, como se apreciará por la forma de las ecuaciones descritas anteriormente, cada uno de los elementos de datos estadísticos se calcula de forma recursiva, eso es usar solo los elementos de datos almacenados en la tabla para el punto de datos anterior (en este caso  $k = 1$ ) y no usar todos los puntos de datos recibidos previamente de los sensores. Por lo tanto, por ejemplo, para el conjunto de datos 1001 ( $k = 1001$ ) los elementos de datos estadísticos almacenados en la tabla 300 para el conjunto de datos 1000-ésimo ( $k = 1000$ ), y no todos los 1000 conjuntos de puntos de datos anteriores, se utilizan junto con el 1001 conjunto de elementos de datos para determinar la desviación 316 típica, densidad 314 de datos media y densidad 312 de datos sobre la base de la cual se lleva a cabo la identificación de anomalías en la etapa 418. Por lo tanto, el procedimiento es computacionalmente ligero y puede realizarse en tiempo real.

La Figura 5 muestra un diagrama de flujo que ilustra un procedimiento 500 de procesamiento de datos de detección de anomalías, correspondiente en general a la etapa 418, con más detalle. El procedimiento 500 de detección de anomalías incluye una primera prueba 502 en la que se determina si se considera que el sistema en general exhibe un comportamiento anómalo al evaluar los datos de todos los sensores. Por ejemplo, se puede aplicar una única regla estadística a los datos de todos los sensores. Por ejemplo, la regla puede ser que se identifique una anomalía si la densidad de datos difiere de la densidad de datos media en más de tres desviaciones típicas.

La Figura 6 ilustra esta regla estadística. La Figura 6 muestra una representación 600 gráfica del gráfico de densidad 602 de datos  $D$ , la densidad 604 de datos media  $D$  y  $D$  menos tres desviaciones típicas, 606, como función del índice 608 de datos  $k$ . Tomando, por ejemplo, los datos para ser la velocidad del avión,  $V$ , entonces se puede ver que para muestras de datos en  $k$  de aproximadamente 180, 190 y 1150, la densidad de datos de la velocidad 602 está a más de tres desviaciones típicas de la densidad 604 de datos media y, por lo tanto, se han identificado tres condiciones 610 anómalas. En otra parte, la densidad 602 de datos es inferior a tres desviaciones típicas de la densidad 604 de datos media y, por lo tanto, no se identifica ningún estado anómalo.

Por lo tanto, en la etapa 502, la prueba estadística de si el valor medio para el punto de datos actual para todo el sistema está dentro de tres desviaciones típicas de la densidad de datos media se aplica a cada elemento de datos. Si el valor promedio para cada elemento de datos excede tres desviaciones típicas de la densidad de datos media, entonces se identifica un estado anómalo del sistema en su conjunto en la etapa 502 y el procesamiento pasa a la etapa 504 en la que el aparato de procesamiento de datos marca el estado del sistema como operación anómala.

Independientemente de si se identifica un estado anómalo para el sistema en su conjunto en la etapa 502, el procesamiento pasa a la etapa 506 en la que se puede aplicar una prueba a un subconjunto de sensores (es decir, a menos de todos los sensores) para detectar el comportamiento anómalo de una parte o subsistema del sistema. Por ejemplo, podría ser que ese sensor en particular, o grupo de sensores, sea particularmente sensible a la identificación de comportamientos potencialmente anómalos. Se puede aplicar una prueba estadística diferente al uno o al grupo de sensores para identificar comportamientos anómalos. Por ejemplo, en la etapa 506, una prueba estadística de si el valor medio de la altitud es más de tres desviaciones típicas de la densidad de datos media para la altitud puede aplicarse solo al elemento de datos de altitud. Por ejemplo, durante un modo de vuelo de crucero, un cambio repentino de altitud puede verse como una operación anómala significativa de una aeronave. Por lo tanto, en 506, se aplica una regla estadística diferente a un subconjunto de datos del sensor para identificar comportamientos anómalos. Si se determina que el valor medio de altitud está a más de tres desviaciones típicas de la densidad de datos media en la etapa 506, entonces se identifica una anomalía diferente en la etapa 508. El procesamiento continúa hasta el final del procedimiento 500, con cualquier anomalía del sistema y/o subsistema identificada, y el procesamiento vuelve a la rutina 400 principal.

Si en la etapa 420, se considera que se identifica una anomalía, entonces la unidad 230 de procesamiento de datos puede emitir una o más señales en la etapa 426. La unidad 220 de procesamiento de datos puede emitir una variedad de señales de salida diferentes. Por ejemplo, se puede emitir una señal de control para controlar una parte del subsistema del sistema general. Por ejemplo, se puede emitir una señal de control a lo largo de la línea 228 al servo 214 para cambiar el ángulo de las aletas en respuesta a la detección de una condición de operación anómala. Además de las señales de control, el procesamiento de datos en 220 puede emitir una señal de alerta o alarma. Por ejemplo, puede emitirse una señal de alarma a lo largo de la línea 224 a la unidad 230 de visualización para operar una alarma visual en la cabina para resaltar al piloto, en el panel de instrumentación, la existencia de una condición anómala. Adicional o alternativamente, se puede emitir una señal de control o comando para indicar a una parte adicional del sistema que realice alguna operación. Por ejemplo, se puede emitir una señal de comando a lo largo de la línea 236 a un dispositivo 232 de procesamiento de datos adicional para llevar a cabo el procesamiento de datos adicional o para comenzar la recopilación y/o análisis de datos adicionales. Por lo tanto, la unidad 220 de procesamiento de datos puede emitir una amplia variedad de diferentes tipos de señales, dependiendo de la naturaleza de la condición anómala identificada y qué acción remediadora u otra acción puede ser requerida.

Con referencia a la Figura 7, se muestra una segunda realización de un sistema 700 en forma de un sistema de vídeo, en el que se puede usar la invención. El sistema 700 de vídeo incluye un aparato 702 de procesamiento de datos que incluye un aparato 720 de procesamiento de datos y un dispositivo 722 de almacenamiento de datos similar al descrito anteriormente. El sistema 700 de vídeo también incluye un dispositivo 704 de captura de imagen de vídeo y un dispositivo 706 de almacenamiento de datos de vídeo para almacenar permanentemente datos de vídeo capturados por el dispositivo 704 de captura de vídeo. El dispositivo 704 de captura de vídeo puede tener la forma de un dispositivo

de carga acoplada y puede incluir elementos ópticos para capturar imágenes 710 entrantes y generar tramas de datos de vídeo que se almacenan en el almacén 706 de datos de vídeo. El sistema 700 de vídeo puede incluir otras partes, componentes o subsistemas, como se ilustra en la parte 708.

5 El dispositivo 704 de captura de vídeo captura tramas secuenciales de datos de vídeo que comprenden una pluralidad de píxeles, cada píxel con un valor de datos rojo, verde y azul. Como alternativa a R, G, B también es posible usar H, S, V (tono, saturación y valor de brillo respectivamente) y que pueden derivarse de la misma trama de imagen y píxeles/contenedores. Esto captura efectivamente la misma información pero puede ser más robusto en algunas aplicaciones.

10 Por ejemplo, la figura 8 ilustra una trama 720 de datos de vídeo que comprende una pluralidad de filas y columnas de píxeles. La trama 720 de datos de vídeo se divide en doce subregiones, o contenedores, dispuestos en una matriz de 3x4. En otras realizaciones, un mayor (p.ej 16 o 24) o menor (p.ej nueve) número de subregiones se pueden utilizar, así como otras disposiciones de subregiones. Se ilustra un primer contenedor 722, solo como ejemplo, en la región superior izquierda de la trama 720 de imagen. Otras disposiciones de almacenamiento son posibles, por ejemplo una matriz de 3x3. Se ha encontrado que la separación de una trama de imagen en subregiones mejora la identificación de eventos anómalos en los datos de imagen capturados como se describirá con mayor detalle a continuación. En algunas realizaciones, el evento anómalo puede ser un tipo específico de hito.

15 La Figura 9 muestra una estructura 730 de datos similar a la ilustrada en la Figura 3. Sin embargo, en la estructura 730 de datos, los valores 732, 734, 736 promedio de los datos del sensor son para los elementos de datos de color rojo, verde y azul para las imágenes capturadas por el sistema 700 de vídeo. En otras realizaciones, los elementos de datos para Tono, Saturación y Valor del brillo se pueden usar, tal como se ha expuesto antes. En esta realización descrita, se puede considerar que el sensor es el CCD u otro sensor de formación de imágenes utilizado por el componente 704 de captura de vídeo. Además, la estructura 730 de datos puede considerarse esquemática en el sentido de que cada contenedor de la imagen tiene su propio conjunto de elementos de datos. Es decir, cada contenedor de la trama de imagen se considera una entidad independiente, para detección de eventos anómalos. Por lo tanto, una trama de datos de vídeo, en el ejemplo actual, comprende doce entidades, cada una con sus propios valores de datos estadísticos y RGB. Por lo tanto, la estructura 730 de datos almacena efectivamente doce conjuntos de elementos de datos específicos ilustrados en la Figura 9. Además, los elementos de datos estadísticos se calculan y almacenan por separado para cada uno de los elementos de datos R, G y B. Por lo tanto, hay tres valores de cada uno del producto 738 escalar, densidad 740 de datos estimada, densidad 742 de datos media y desviación 744 típica, uno para cada uno de R, G y B.

20 La figura 10 muestra un diagrama de flujo que ilustra un procedimiento 750 de detección de anomalías de procesamiento de datos de acuerdo con una segunda realización de la invención. El procedimiento 750 es similar al procedimiento 400 y las diferencias se resaltarán a continuación en el contexto del sistema 700 de vídeo. En el contexto del sistema 700 de vídeo, el comportamiento anómalo puede considerarse equivalente a la detección de novedad en las imágenes de vídeo, por ejemplo, una nueva entidad que aparece en una imagen de vídeo que no estaba presente previamente. Tomando, como ejemplo práctico, si el sistema de vídeo captura imágenes de una calle, entonces un evento anómalo puede ser la aparición de un automóvil en la imagen de vídeo que no estaba previamente presente.

35 El procedimiento se inicializa en la etapa 752 y en la etapa 754, la unidad 720 de procesamiento de datos recibe una trama 756 de datos de vídeo del dispositivo 704 de captura de vídeo. Un primer contenedor, por ejemplo, contenedor 722, está seleccionado y para los píxeles del primer contenedor. Un valor medio de los datos R, G y B para todos los píxeles dentro del contenedor seleccionado actualmente se calcula de forma recursiva y se almacena en la tabla 730.

40 Luego, en las etapas 760 a 766, el producto escalar, la densidad de datos, la densidad de datos media y la varianza 738, 740, 742, 744 de densidad de datos se calculan y almacenan recursivamente en la tabla 730. Luego, en la etapa 768, se aplica una rutina de detección de anomalías utilizando los datos estadísticos almacenados en la tabla 730 para identificar cualquier comportamiento anómalo en el contenedor actual. Si se determina que se ha detectado algún comportamiento anómalo en la etapa 768, entonces el procesamiento pasa a la etapa 770 en la que la unidad 720 de procesamiento de datos puede emitir una señal. Si no se detecta ninguna anomalía en la etapa 768, luego el procesamiento pasa a 772, y el flujo del procedimiento vuelve, como se ilustra en la línea 774, a la etapa 758 en la que se procesa un siguiente contenedor de la trama de imagen. El procesamiento continúa como se describió anteriormente para el siguiente contenedor y el procesamiento continúa en bucle hasta que se hayan procesado todos los contenedores de la trama actual. Luego, en la etapa 776, se selecciona una siguiente trama de imagen para procesar y devoluciones de procesamiento, como se ilustra procesando la línea de retorno 778 a la etapa 754 en la que la unidad 720 de procesamiento de datos recibe la siguiente trama de los datos de imagen desde el dispositivo 704 de captura de imagen.

55 Aplicando la detección de anomalías de contenedor a contenedor, en lugar de trama a trama, así, se ha encontrado que mejora la detección de anomalías en las imágenes. Si los valores promedios R, G y B se utilizan para la trama completa de datos de imagen, entonces los valores R, G y B promediados sobre toda la trama de la imagen tienden a eliminar las variaciones locales. Por ejemplo, tomando la trama como un todo, si un automóvil rojo ingresara a la trama de imagen desde la esquina inferior derecha, entonces el valor rojo promedio en toda la trama puede no cambiar mucho. Sin embargo, el valor rojo promedio para el contenedor inferior derecho cambiaría significativamente. Se ha

60

encontrado que al dividir la trama de la imagen en subregiones de manera que cada subregión tenga un tamaño correspondiente al tamaño típico de una entidad dentro de la imagen, esto mejora la detección de anomalías.

La Figura 11 muestra un diagrama de flujo del procedimiento que ilustra un procedimiento 800 de procesamiento de datos de detección de anomalías, correspondiente en general a la etapa 768 de la Figura 10. Como se ha ilustrado en la Figura 11, el procedimiento 800 de detección de anomalías, incluye un enfoque de reglas múltiples para la identificación de anomalías. En primer lugar, se aplica una regla estadística en la etapa 802, si la regla estadística no se pasa para identificar una anomalía, entonces termina el procesamiento. De otra manera, el procesamiento pasa a la segunda regla 804 que se basa en una prueba temporal. Si no se pasa la regla temporal, entonces no se identifica ninguna anomalía y el procesamiento finaliza. Si se pasa la regla temporal para indicar una anomalía potencial, luego el procesamiento continúa con la etapa 806 en la que se aplica una regla basada en eventos. Si no se pasa la regla basada en eventos, entonces el procesamiento termina. De otra manera, el procesamiento pasa a la etapa 810 en la que se puede considerar que una anomalía ha sido identificada de manera confiable.

Con más detalle, la regla estadística aplicada en la etapa 802 puede ser similar a la utilizada para la realización de la aeronave. Es decir, la regla estadística puede ser si, para el contenedor actual, la densidad de datos para los tres colores difiere de la densidad de datos media para los tres colores en más de tres desviaciones típicas. En caso negativo, entonces no se puede identificar ninguna anomalía.

Los valores para todos los colores de los canales de color (p. ej. RGB o HSV) se suman usando la distancia euclidiana de la siguiente manera:

$$D(x_k) = \frac{1}{1 + \frac{1}{k} \sum_{i=1}^k \|x_k - x_i\|^2} = \frac{1}{1 + \frac{1}{k} \sum_{j=1}^n (x_{kj} - x_{ij})^2} \quad (20)$$

En la etapa 804, la regla temporal aplica la prueba de si la anomalía potencial es más probable que sea ruido o un cambio genuino en la imagen que se captura. Por ejemplo, en una típica escena callejera, un automóvil puede moverse a aproximadamente 48,2 km/h y, por lo tanto, se esperaría que ocurriera una anomalía correspondiente a un automóvil que ingresa al campo de visión de la imagen en un gran número de tramas consecutivas debido al intervalo de muestreo de la cámara de vídeo. Es decir, una anomalía de imagen equivalente a un automóvil probablemente estaría presente en el mismo contenedor para múltiples muestras consecutivas. Sin embargo, si la anomalía de la imagen estuviera presente en un contenedor de una primera muestra pero ausente de una segunda muestra, ese cambio sería inconsistente con la velocidad que requeriría un automóvil para entrar y salir del campo de visión en el corto tiempo de muestra del sistema de vídeo. En este caso, la anomalía puede considerarse ruido más que un evento genuino. Por lo tanto, la regla temporal puede probar si se ha identificado la misma anomalía potencial durante un período de tiempo suficientemente largo. En caso negativo, entonces la anomalía puede descartarse como ruido. De otra manera, la regla temporal puede considerarse pasada.

El procesamiento continúa con la etapa 806 en la que se puede aplicar una regla basada en eventos a la anomalía potencial identificada por las etapas anteriores. Por ejemplo, se podría considerar que una anomalía de imagen genuina tendría una ruta consistente a través del campo de visión. Es decir, se esperaría que un automóvil que ingrese al campo de visión desde la derecha dé lugar a posibles anomalías en la imagen en los contenedores inferior derecho, inferior central derecho, inferior central izquierdo e inferior izquierdo en esa secuencia. Por lo tanto, una posible anomalía identificada en el contenedor inferior derecho, el central izquierdo medio y luego el superior derecho pueden ser inconsistentes con una nueva entidad de imagen genuina. Por lo tanto, también se puede aplicar una regla basada en eventos para confirmar la autenticidad de la entidad de imagen anómala potencialmente identificada. Como podrá apreciarse, anómalo, en este ejemplo, no significa necesariamente un error, sino simplemente una propiedad nueva o modificada de la imagen.

Se apreciará que en otras realizaciones, se pueden aplicar diferentes tipos de reglas. También, se pueden aplicar diferentes combinaciones de reglas y puede ser necesario pasar solo un subconjunto de todas las reglas para que una anomalía se considere identificada de manera confiable. Otras variaciones y modificaciones serán evidentes para una persona de habilidad ordinaria en la técnica a partir de la discusión anterior.

Si en la etapa 768, se ha detectado una anomalía, entonces en la etapa 770, la unidad 720 de procesamiento de datos puede emitir una o más señales de salida en la etapa 770. Por ejemplo, la unidad de procesamiento de datos puede emitir una señal de control al dispositivo 704 de captura de imagen para controlar su funcionamiento, por ejemplo para aumentar la tasa de captura de datos. Adicional o alternativamente, la unidad 720 de procesamiento de datos puede emitir un comando al almacén 706 de datos para simplemente marcar o marcar de otro modo los datos de imagen almacenados en el almacén 706 de datos como incluyendo una entidad de imagen anómala. Adicional o alternativamente, la unidad 720 de procesamiento de datos puede emitir una señal a su propio almacén 722 de datos para marcar los datos como relacionados con un evento anómalo. Adicional o alternativamente, la unidad 720 de procesamiento de datos puede emitir una señal de alerta o alarma a otra parte del sistema 708 de formación de imágenes, por ejemplo un dispositivo de visualización o alarma, para indicar a un usuario la identificación de un evento

anómalo. Se pueden proporcionar otros tipos de señal de salida, como será evidente para una persona de habilidad ordinaria en la técnica desde la discusión anterior.

5 Se apreciará que existen muchas aplicaciones diferentes para la detección de anomalías o novedades en el procesamiento de imagen. Por ejemplo, la invención se puede utilizar para identificar puntos de referencia en escenas capturadas por un robot móvil (por ejemplo, para fines de navegación) o para diarios de vídeo automáticos (por ejemplo, para proporcionar detección automática de ubicación).

10 Generalmente, las realizaciones de la presente invención, y en particular los procedimientos implicados en la identificación de estados anómalos del sistema emplean diversos procesos que implican datos procesados por, almacenados en o transferidos a través de uno o más dispositivos informáticos o de procesamiento de datos. Las realizaciones de la presente invención también se refieren a un aparato, que puede incluir uno o más dispositivos de procesamiento de datos individuales, para realizar estas operaciones. Este aparato puede estar especialmente construido para los fines requeridos, o puede ser una computadora de uso general o un dispositivo o dispositivos de procesamiento de datos, activado selectivamente o reconfigurado por un programa de computadora y/o estructura de datos almacenados en la computadora o dispositivos. Los procedimientos presentados en el presente documento no están inherentemente relacionados con ninguna computadora en particular u otro aparato. En particular, se pueden usar varias máquinas de fin general con programas escritos de acuerdo con las enseñanzas en el presente documento, o puede ser más conveniente construir un aparato más especializado para realizar las etapas requeridas del procedimiento.

20 Además, las realizaciones de la presente invención se refieren a medios legibles por computadora o productos de programas informático que incluyen instrucciones de programa y/o datos (incluyendo estructuras de datos) para realizar diversas operaciones implementadas por computadora. Los ejemplos de medios legibles por computadora incluyen, pero no se limitan a, medios magnéticos como discos duros, disquetes y cinta magnética; medios ópticos tales como discos CD-ROM; medios magnetoópticos; dispositivos de memoria de semiconductores y dispositivos de hardware que están especialmente configurados para almacenar y realizar instrucciones del programa, tales como dispositivos de memoria de solo lectura (ROM) y memoria de acceso aleatorio (RAM). Los datos y las instrucciones de programa de la presente invención también pueden realizarse en una onda portadora u otro medio de transporte. Los ejemplos de instrucciones de programa incluyen código de máquina, como los producidos por un compilador y los archivos que contienen código de nivel superior que la computadora puede ejecutar con un intérprete.

30 La Figura 12 ilustra un sistema informático típico que, cuando se configura o diseña adecuadamente, puede servir como un aparato de la presente invención. El sistema 900 informático incluye cualquier número de procesadores 902 (también conocidos como unidades centrales de procesamiento o CPU) que están acoplados a dispositivos de almacenamiento, incluido el almacenamiento 906 primario (típicamente una memoria de acceso aleatorio o RAM), almacenamiento 904 primario (generalmente una memoria de solo lectura o ROM). La CPU 902 puede ser de varios tipos, incluidos microcontroladores y microprocesadores, como dispositivos programables (por ejemplo, CPLD y FPGA) y dispositivos no programables como ASIC de matriz de puerta o microprocesadores de fin general. Como es bien sabido en la técnica, el almacenamiento 904 primario actúa para transferir datos e instrucciones unidireccionalmente a la CPU y el almacenamiento 906 primario se usa típicamente para transferir datos e instrucciones de manera bidireccional. Ambos dispositivos de almacenamiento primario pueden incluir cualquier medio legible por computadora adecuado, como los descritos anteriormente. Un dispositivo 908 de almacenamiento masivo también está acoplado bidireccionalmente a la CPU 902 y proporciona capacidad de almacenamiento de datos adicional y puede incluir cualquiera de los medios legibles por computadora descritos anteriormente. El dispositivo 908 de almacenamiento masivo puede usarse para almacenar programas, datos y similares, y normalmente es un medio de almacenamiento secundario, como un disco duro. Se apreciará que la información retenida dentro del dispositivo 908 de almacenamiento masivo, puede, en casos apropiados, incorporarse de manera estándar como parte del almacenamiento 406 primario como memoria virtual. Un dispositivo de almacenamiento masivo específico como un CD-ROM 914 también puede pasar datos unidireccionalmente a la CPU.

50 La CPU 902 también se puede acoplar a una interfaz 910 que se puede conectar a uno o más dispositivos de entrada/salida, como monitores de vídeo, bolas de desplazamiento, ratones, teclados, micrófonos, pantallas sensibles al tacto, lectores de tarjetas de transductores, lectores de cinta magnética o de papel, tabletas, estiletes, reconocedores de voz o de escritura a mano, u otros dispositivos de entrada bien conocidos como, por supuesto, otras computadoras. Finalmente, la CPU 902 opcionalmente se puede acoplar a un dispositivo externo como una base de datos o una red informática o de telecomunicaciones utilizando una conexión externa como se muestra generalmente en 912. Con tal conexión, se contempla que la CPU podría recibir información de la red, o podría enviar información a la red en el curso de realizar las etapas del procedimiento descritas en el presente documento.

55 Aunque lo anterior ha descrito generalmente la presente invención de acuerdo con procedimientos y aparatos específicos, la presente invención tiene una gama de aplicabilidad mucho más amplia. En particular, los aspectos de la presente invención no se limitan a ningún tipo específico de sistema industrial y se pueden aplicar a prácticamente cualquier tipo de sistema industrial en el que uno o más sensores estén disponibles para proporcionar datos de series temporales relacionados con una o más propiedades del sistema. Un experto en la materia reconocería otras variantes, modificaciones y alternativas a la luz de la discusión anterior.

60

**REIVINDICACIONES**

1. Un procedimiento (400, 750) en tiempo real para identificar un estado anómalo de un sistema, incluyendo el sistema un sensor que emite elementos de datos de series temporales relacionados con una propiedad del sistema, comprendiendo el procedimiento:
  - 5 recibir (404, 754) un elemento de datos actual del sensor; recuperar una estimación calculada recursivamente almacenada de una densidad de datos anterior para todos los elementos de datos de series temporales anteriores de una estructura (300) de datos, en el que la densidad de datos representa la distancia total entre elementos de datos; calcular recursivamente (412, 762) una estimación de una densidad de datos actual para los elementos de datos de series temporales usando el elemento de datos actual y dicha estimación de la densidad de datos anterior y almacenando dicha estimación de la densidad de datos actual en dicha estructura de datos;
  - 10 recuperar una propiedad estadística previa calculada recursivamente almacenada de la estimación de la densidad de datos anterior de la estructura (300) de datos; calcular recursivamente (414, 764) al menos una propiedad estadística de la estimación de la densidad de datos actual utilizando dicha propiedad estadística anterior de la estimación de la densidad de datos anterior y almacenar dicha al menos una propiedad estadística de la estimación de la densidad de datos actual en dicha estructura de datos;
  - 15 determinar (418, 768) a partir de la al menos una propiedad estadística si el elemento de datos actual indica un estado anómalo del sistema; y
  - 20 emitir una señal (426, 770) al sistema si se determina que el elemento de datos actual indica un estado anómalo del sistema, en el que la señal se selecciona de: una señal de control; una señal de retroalimentación; una señal de alarma; una señal de comando; una señal de advertencia; una señal de alerta; una señal servo; una señal de disparo; una señal de captura de datos; y una señal de adquisición de datos.
- 25 2. El procedimiento (400, 750) de la reivindicación 1, en el que la al menos una propiedad estadística incluye la media actual de la densidad de datos actual y la varianza actual de la densidad de datos actual.
3. El procedimiento (400, 750) de la reivindicación 2, en el que el elemento de datos actual se determina (418, 768) como anómalo en función del grado de diferencia entre la estimación de la densidad de datos actual y la media actual de la densidad de datos actual.
- 30 4. El procedimiento (400, 750) de la reivindicación 3, en el que el elemento de datos actual se determina (418, 768) como anómalo si la diferencia entre la estimación de la densidad de datos actual y la media actual de la densidad de datos actual es mayor que tres desviaciones típicas de la media actual de la densidad de datos actual.
5. El procedimiento (750) de cualquiera de las reivindicaciones anteriores, que comprende además aplicar una prueba (804, 806) adicional para determinar si el elemento de datos actual es anómalo.
- 35 6. El procedimiento (750) de la reivindicación 5, en el que la prueba adicional se selecciona de: una prueba (804) basada en el tiempo; y una prueba (806) basada en eventos.
7. El procedimiento (400, 750) de cualquier reivindicación anterior, y que comprende además:
  - calcular recursivamente (408, 758) un valor medio del elemento de datos utilizando el elemento de datos actual; y
  - calcular recursivamente (410, 760) un producto escalar para el elemento de datos utilizando el elemento de datos actual.
- 40 8. El procedimiento (400, 750) de la reivindicación 7, y que comprende además: utilizar el valor medio del elemento de datos y el producto escalar del elemento de datos para calcular recursivamente (412, 762) la estimación de la densidad de datos actual.
9. El procedimiento (400, 750) de cualquier reivindicación anterior, en el que el sistema incluye una pluralidad de sensores, emitiendo cada uno elementos de datos de series temporales relacionados con una propiedad diferente del sistema y en el que el procedimiento se aplica a elementos de datos actuales recibidos respectivamente de cada uno de la pluralidad de sensores.
- 45 10. El procedimiento (400, 750) de la reivindicación 9, en el que la determinación comprende determinar a partir de la al menos una propiedad estadística si un subconjunto de elementos de datos actuales de la pluralidad de elementos de datos indica un estado anómalo del sistema.
- 50 11. El procedimiento (400, 750) de la reivindicación 9, en el que determinar comprende determinar a partir de la al menos una propiedad estadística si todos los elementos de datos actuales de la pluralidad de elementos de datos indican un estado anómalo del sistema.
12. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que la estimación de la densidad de datos actual para los elementos de datos de series temporales es la inversa de la suma de las distancias totales entre todos

los puntos de datos en un espacio de datos para todos los elementos de datos recibidos.

13. Un aparato de procesamiento de datos para identificar un estado anómalo de un sistema en tiempo real, que comprende:

- 5 un dispositivo de procesamiento de datos; y  
un dispositivo de almacenamiento en comunicación con el dispositivo de procesamiento de datos, almacenando el dispositivo de almacenamiento el código de programa informático ejecutable por el dispositivo de procesamiento de datos para llevar a cabo el procedimiento de cualquiera de las reivindicaciones 1 a 12.

14. Un sistema, comprendiendo el sistema:

- 10 al menos una parte operativa;  
al menos un sensor para medir una propiedad de la parte operativa; y  
un aparato de procesamiento de datos según la reivindicación 13, en el que el aparato de procesamiento de datos está en comunicación con el sensor para recibir datos de series temporales del sensor y en el que el aparato de procesamiento de datos tiene una salida que está en comunicación con el sistema para emitir la señal al sistema.

Fig. 1

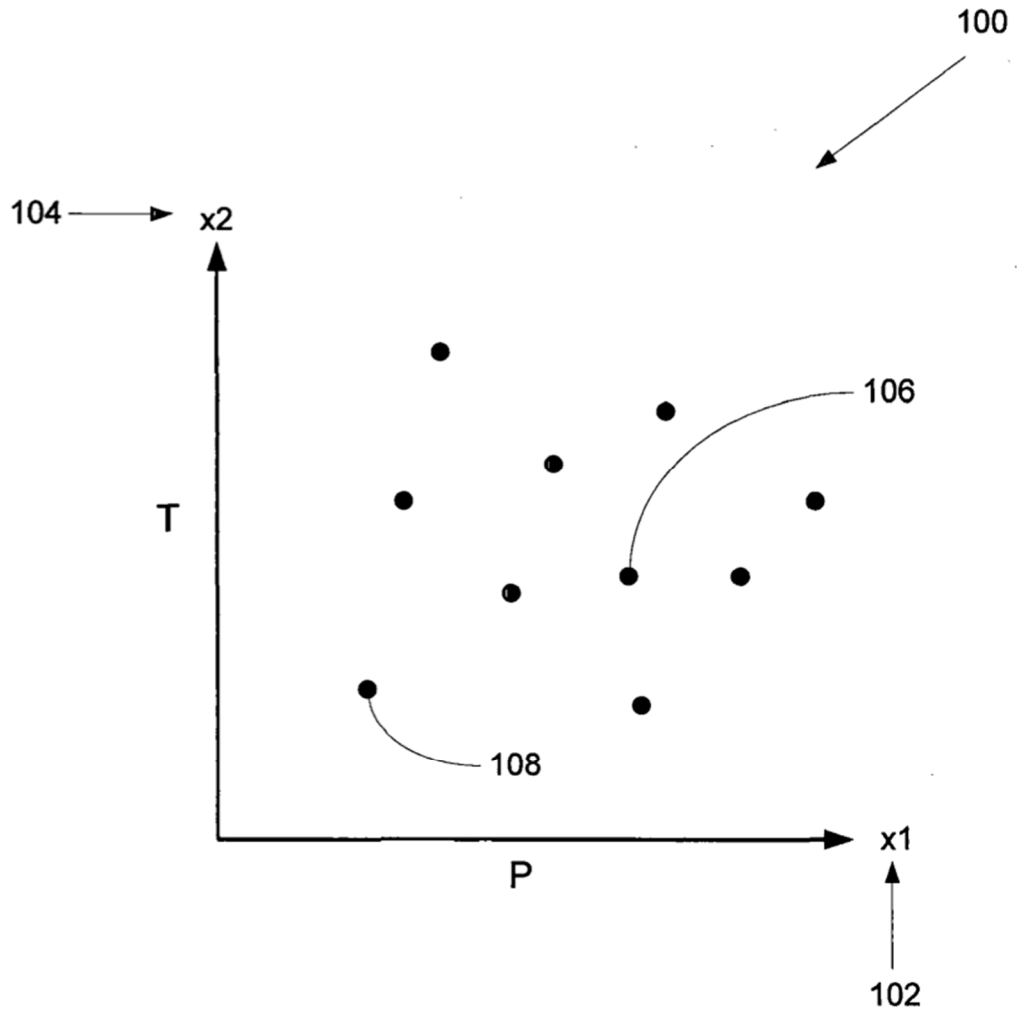


Fig. 2

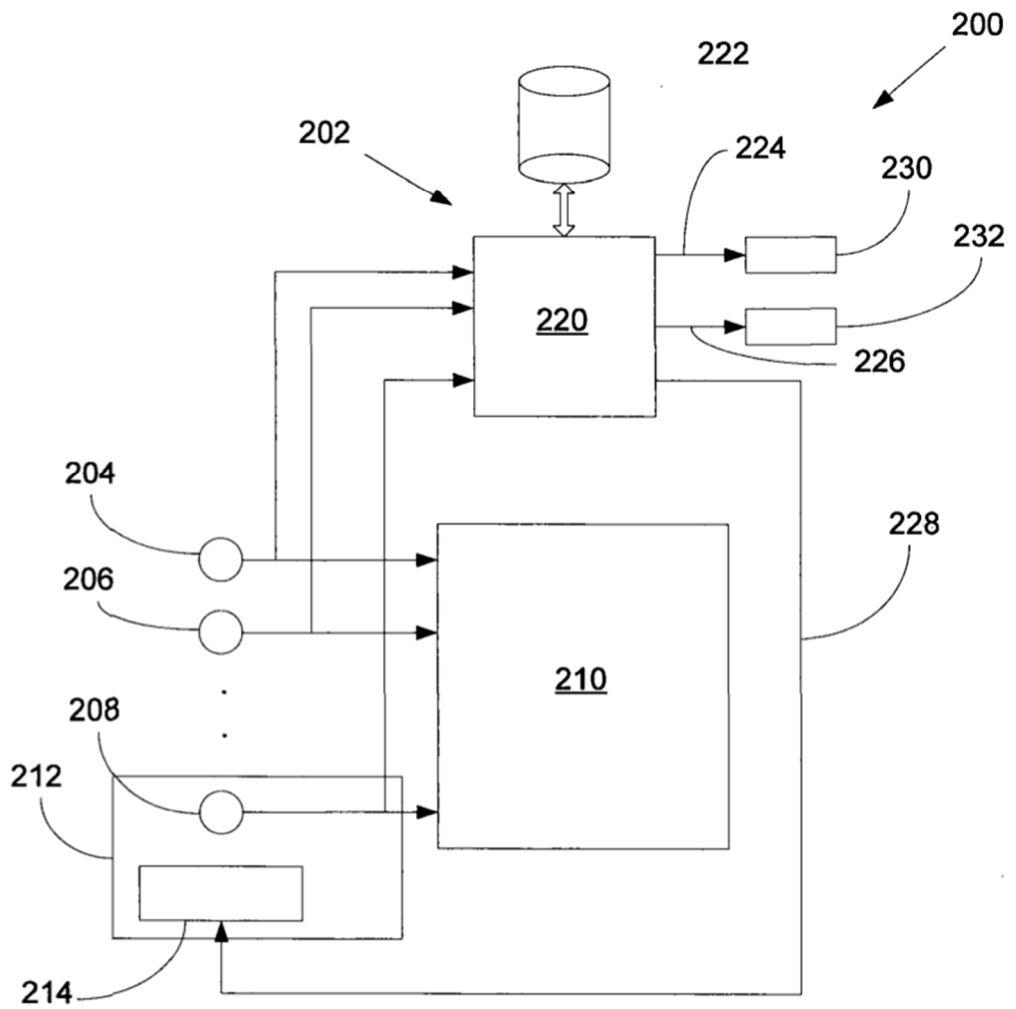




Fig. 3

$k$	$\sigma_D$	$\bar{D}$	$D$	$\Sigma$	$\bar{V}$	$\bar{A}$	$\bar{P}$
0							
1							
2							
·							
·							
·							

Fig. 4

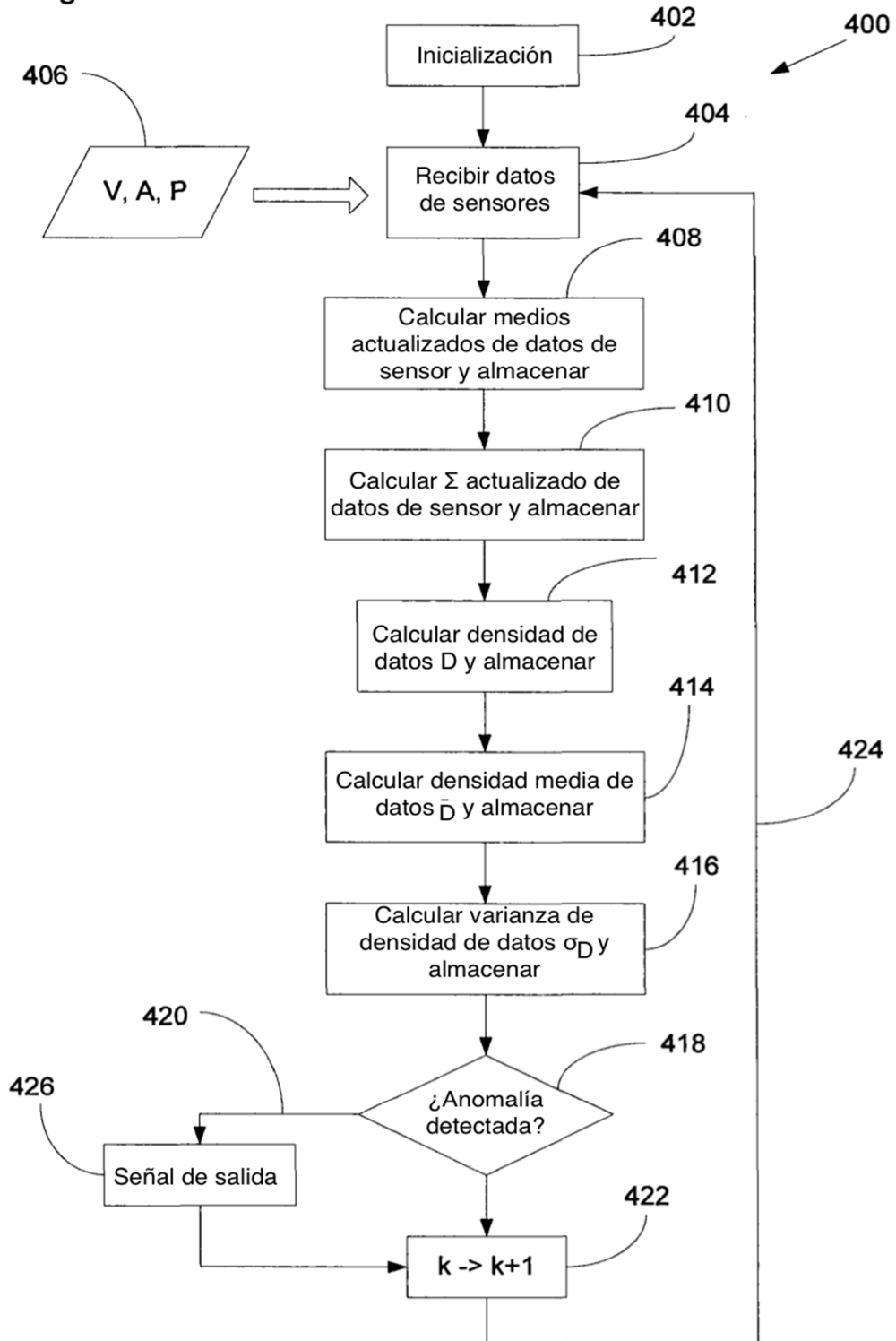


Fig. 5

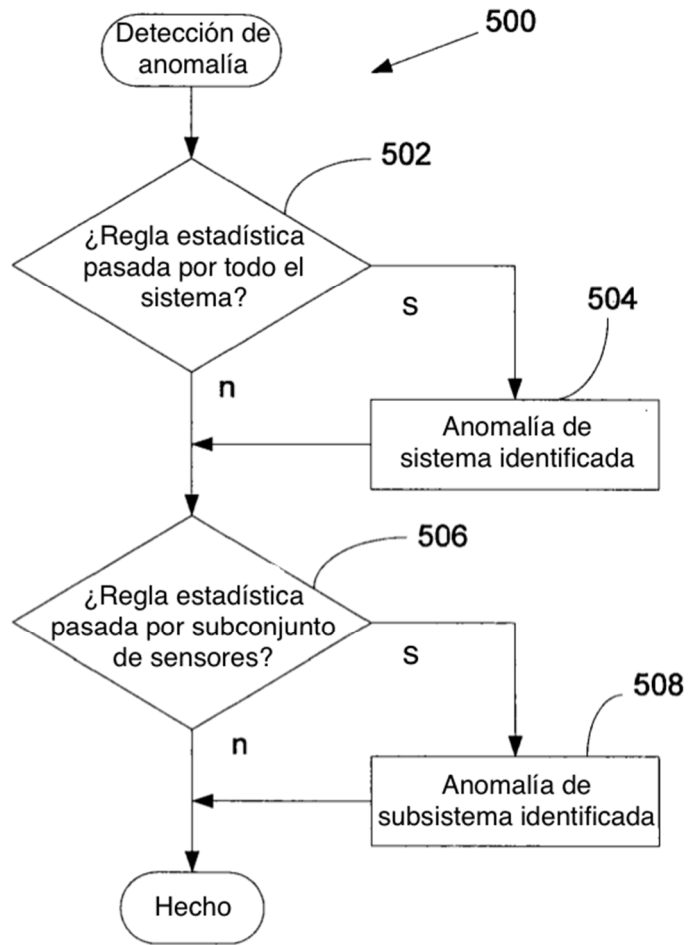


Fig. 6

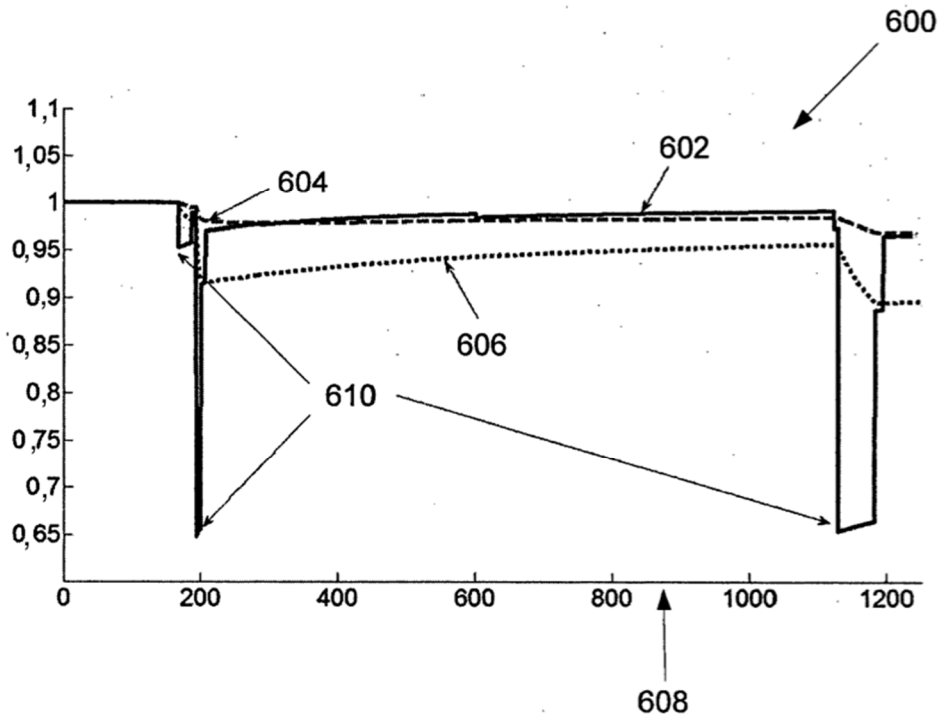


Fig. 7

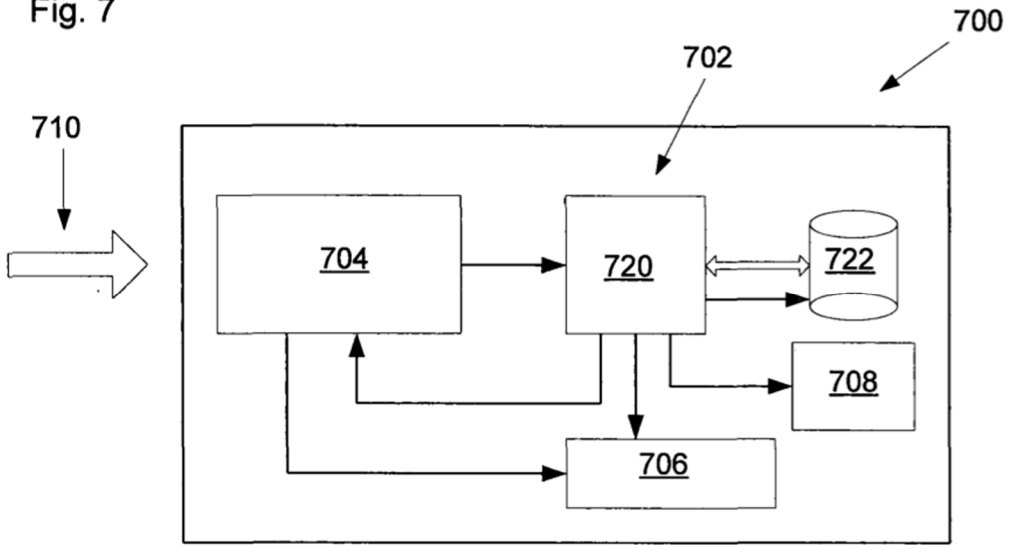


Fig. 8

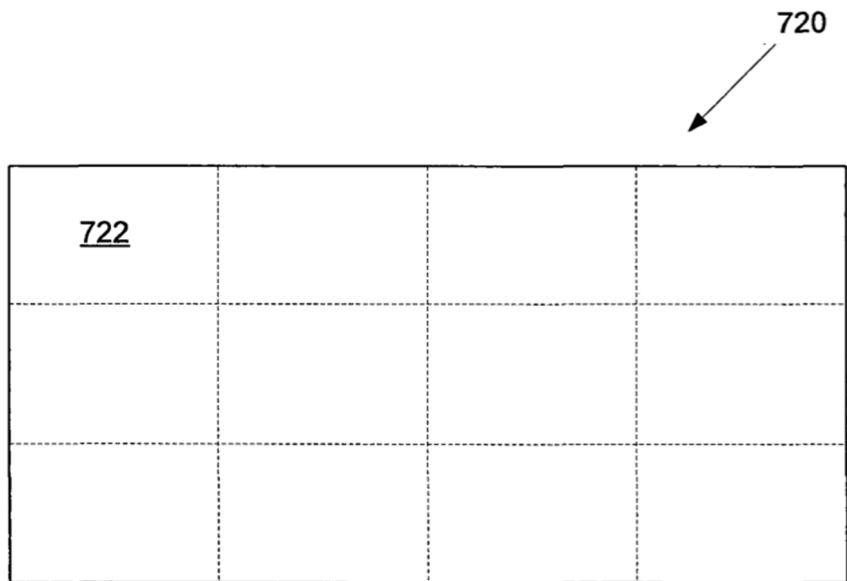


Fig. 9

744	742	740	738	732	734	736	730
k	$\sigma_D$	$\bar{D}$	D	$\Sigma$	$\bar{R}$	$\bar{G}$	$\bar{B}$
0							
1							
2							
.							
.							
.							

Fig. 11

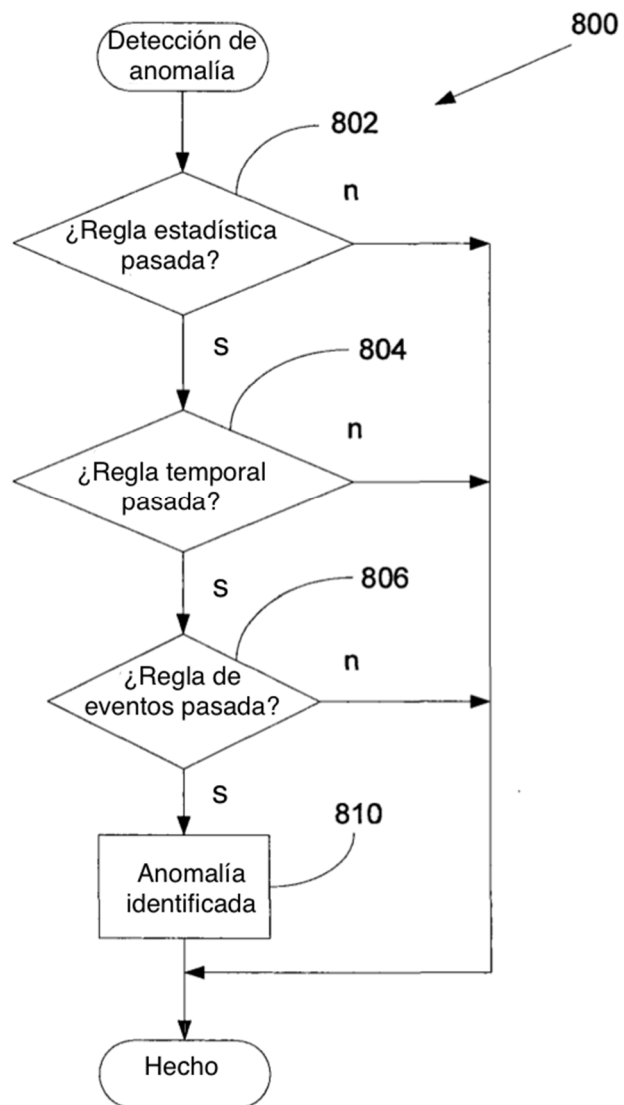


Fig. 10

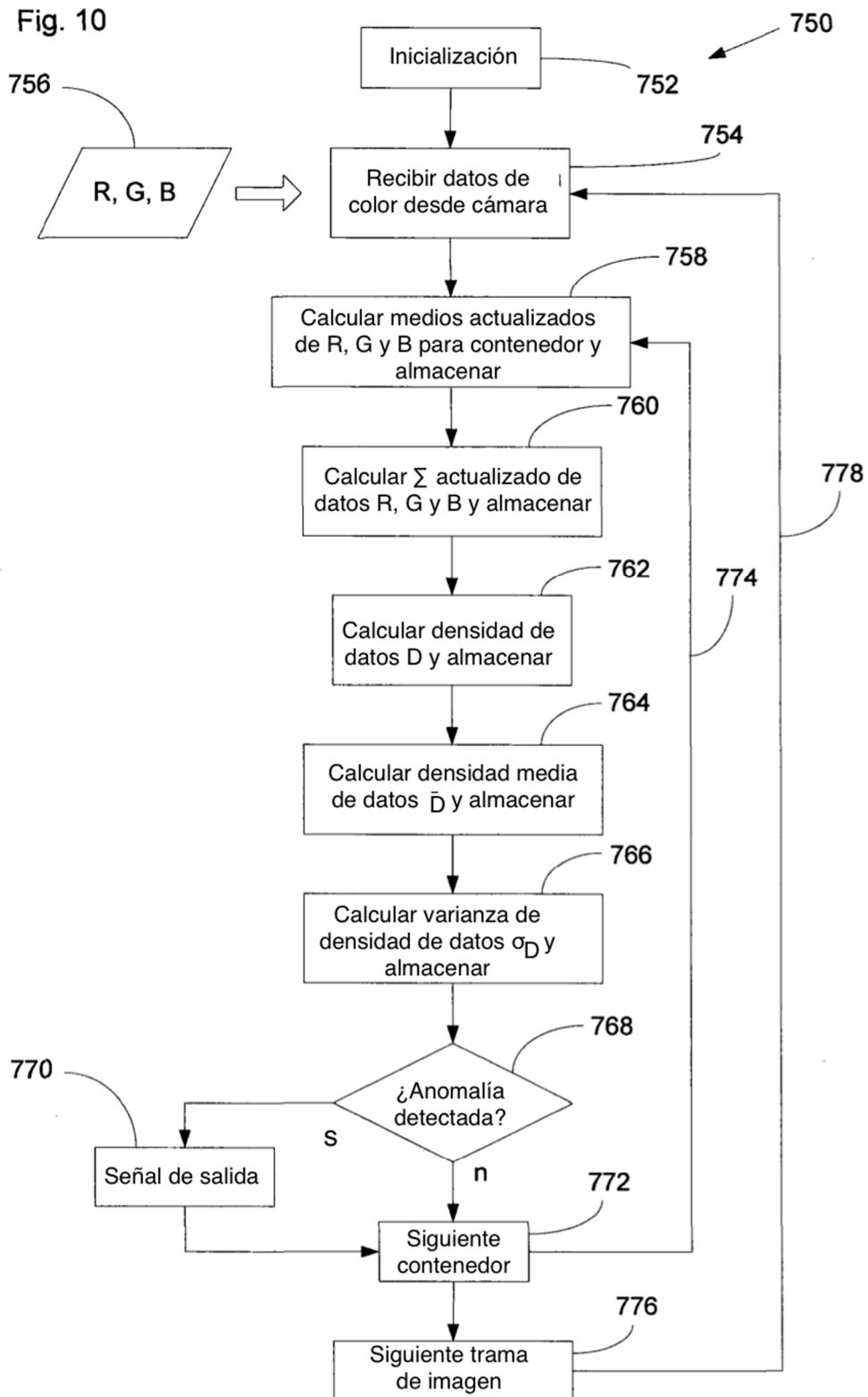


Fig. 12

