

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 017 789**

51 Int. Cl.:

H04N 7/18 (2006.01)

G08B 13/196 (2006.01)

H04N 21/8547 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.07.2017 E 24165302 (1)**

97 Fecha y número de publicación de la concesión europea: **22.01.2025 EP 4451665**

54 Título: **Método para importar y procesar imágenes de vídeo en tiempo real**

30 Prioridad:

07.11.2016 AT 510132016

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

13.05.2025

73 Titular/es:

TB-TRAXLER GMBH (100.00%)

Im Wirtschaftspark 4

3494 Gedersdorf, AT

72 Inventor/es:

TRAXLER, JOHANNES

74 Agente/Representante:

ELZABURU, S.L.P

ES 3 017 789 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 para importar y procesar imágenes de vídeo en tiempo real

Campo técnico

5 La presente invención se refiere a un método para leer y procesar imágenes de vídeo en tiempo real, en donde una o más cámaras de vídeo proporcionan imágenes, siendo las imágenes digitalizadas y almacenadas en una memoria, después de lo cual las imágenes son alimentadas a al menos un monitor.

Estado de la técnica

10 En el sector de la televisión se distingue entre imágenes (de vídeo) y señales de vídeo. La separación es necesaria, por una parte, para distinguir un fotograma de vídeo terminado de una señal de vídeo definida electrónicamente para su transmisión y, por otra parte, para separar las diferentes tareas de la imagen y la señal de vídeo. El término imagen se refiere a una imagen completa tal como se ve en la pantalla, mientras que la señal de vídeo es necesaria para transmitir la imagen electrónicamente.

15 Los sistemas clásicos de transmisión de imágenes eran analógicos, es decir, la imagen grabada por la cámara de vídeo se transmitía de forma analógica y se mostraba inmediatamente en una pantalla. Cualquier interrupción se veía inmediatamente en la pantalla, ya fuera debido a un problema con la cámara o a un problema en la cadena de transmisión. Por otra parte, se excluía cualquier desfase temporal entre la cámara y la pantalla, además del corto tiempo de propagación de la señal.

20 Las soluciones modernas para la transmisión de imágenes están diseñadas para producir la imagen de mayor calidad posible y eliminar las interferencias. Esto incluye los procesos de ajuste automático en la cámara de vídeo, la ruta de transmisión y el procesamiento de la imagen o imágenes delante de la pantalla. Según los requisitos de la televisión de entretenimiento, no deberían ser visibles los errores; es mejor que se muestre una imagen almacenada en la pantalla durante el estado de error. Por este motivo, todos los sistemas de transmisión modernos disponen de circuitos electrónicos bastante complejos con almacenamiento de imágenes, empezando por la cámara de vídeo, posiblemente en la vía de transmisión y finalmente en el procesamiento de imágenes. Sin embargo, cuando se trata del procesamiento y almacenamiento de imágenes digitales, no se puede descartar un retraso temporal entre la cámara y el monitor; en el caso de las transmisiones de televisión, esto es incluso muy común, como se puede observar cuando el mismo evento en vivo es transmitido por diferentes estaciones de televisión. A este respecto, son habituales retrasos de unos pocos segundos.

30 La visualización conjunta de varias imágenes de vídeo en un monitor se conoce desde hace décadas. Ejemplos típicos son los monitores de vigilancia de los supermercados, donde, por ejemplo, se muestran imágenes de cuatro cámaras en un solo monitor. La persona que monitoriza las imágenes puede mostrar la imagen de una cámara en pantalla completa en cualquier momento o cambiar otras cámaras al monitor si nota algo sospechoso.

35 Igualmente conocidos son los monitores de vigilancia del metro de Viena, que permiten al conductor ver cuándo han subido y bajado todos los pasajeros, así como despachar el tren y cerrar las puertas. A veces, las imágenes de dos cámaras se muestran en un monitor. En estas dos aplicaciones, las imágenes parciales simplemente se muestran una al lado de la otra y/o una encima de la otra sin transiciones.

En forma analógica, esto solo es posible si todas las cámaras de vídeo están sincronizadas entre sí, e incluso en ese caso solo con cámaras de vídeo especialmente adaptadas y una gran cantidad de circuitos.

40 El estado de la técnica consiste en digitalizar las cuatro imágenes individuales entrantes, guardarlas y luego crear la nueva imagen compuesta a partir de las imágenes almacenadas digitalmente disponibles y transmitirla a la pantalla. Si la transmisión a la pantalla es analógica, se deben regenerar las señales de sincronización necesarias.

El documento EP2341710 divulga un método en donde la latencia de una señal de cámara solo influye en la visualización de la imagen en el monitor asignado a la cámara y no tiene influencia en la visualización de las otras imágenes en las otras cámaras.

45 Además, se conocen los llamados espejos retrovisores electrónicos, en donde tres imágenes de cámara (correspondientes al espejo exterior izquierdo, al espejo exterior derecho y al espejo interior) se combinan de forma uniforme ("*stitching*") y se muestran al conductor en un monitor, de modo que este puede ver de un vistazo todo el tráfico que hay detrás de él y sin ángulos muertos. Se puede encontrar un ejemplo en Internet en <https://www.macnica.eu/de/node/219>.

50 La desventaja de estas soluciones es que no se pueden descartar retrasos importantes como consecuencia de la digitalización. Esto es inútil cuando se monitorean sistemas de tráfico porque la persona que mira la pantalla no puede detectar el desfase temporal. En dos segundos, por ejemplo, un tren que se acerca a la estación puede haber recorrido ya un metro, mientras que una situación crítica en el andén no se puede ver en la pantalla. Para un vehículo que viaja a 100 km/h, serían más de 50 m sin contacto visual. Lamentablemente ya se ha producido una víctima mortal como

consecuencia de ese desfase temporal.

Pero incluso sin procesar las imágenes (como "*stitching*"), las imágenes a menudo se almacenan en una memoria y se leen desde allí y se envían a un monitor; un ejemplo es la visualización de cuatro imágenes de cámara en un monitor mencionado anteriormente. Las señales de vídeo de todas las cámaras se almacenan digitalmente y luego solo cada segundo píxel de cada línea y solo cada segunda línea de cada cámara se envía al monitor, de modo que la imagen solo ocupa un cuadrante. También en este caso pueden producirse desfases temporales molestos, ya que siempre es posible que una cámara falle y deje de emitir señal. Luego, la última imagen permanece en la memoria y se muestra constantemente como una imagen fija.

Breve descripción de la invención

Un objetivo de la presente invención es proporcionar un método del tipo mencionado al principio en donde se evita la reproducción excesivamente retardada de imágenes de vídeo.

Este objetivo se resuelve según la invención mediante un método del tipo mencionado al principio, en el que todas las imágenes de vídeo están provistas de una marca de tiempo que se almacena en la memoria junto con las imágenes de vídeo, y en el que, a continuación, se lee la marca de tiempo de la imagen de vídeo o, cuando se reproducen varias imágenes de vídeo simultáneamente, la marca de tiempo de todas las imágenes de vídeo reproducidas y se compara con la hora actual, y la señal solo se suministra al al menos un monitor si la marca de tiempo se remonta a un máximo de un período de tiempo predeterminado o si todas las marcas de tiempo se remontan a un máximo de un período de tiempo predeterminado.

La invención se basa en la idea de que es mejor no mostrar una imagen que mostrar una imagen obsoleta, ya que esto hace evidente de inmediato que no contiene información útil. Si, por ejemplo, establece el período de tiempo especificado en 100 ms, entonces el desfase temporal máximo estará muy por debajo del tiempo de reacción humano típico y, por lo tanto, no causará ninguna perturbación; Si el desfase temporal es mayor, no se muestra ninguna imagen significativa (por ejemplo, solo negro, solo blanco, un patrón), por lo que la perturbación se reconoce inmediatamente.

El término "marca de tiempo" debe entenderse en sentido general. No importa qué momento se considera "cero", puede ser, por ejemplo, el encendido del sistema o una señal de "reinicio". Tampoco importa qué unidad se utilice para la medición. La unidad puede ser, por ejemplo, la velocidad de fotogramas (normalmente 40 ms), es decir, las imágenes digitalizadas se numeran simplemente. El número de la imagen y el momento del encendido dan entonces el momento "real" en donde se tomó la imagen. Sin embargo, este momento no tiene por qué calcularse. Es suficiente con poner en marcha un contador para fines de control, que se pone a cero cuando se enciende el sistema o cuando se recibe una señal de reinicio y que cuenta de 1 en 1 cada 40 ms. Si se compara este contador con el número de la imagen, se puede reconocer que una imagen está desactualizada si hay una determinada desviación de, por ejemplo, 3 o 5 imágenes (que corresponden a 120 ms o 200 ms).

Si se procesan las imágenes, las imágenes almacenadas se deben procesar después de haberles puesto una marca de tiempo, y la marca de tiempo se debe leer después del procesamiento y luego las imágenes procesadas se deben enviar al al menos un monitor. De esta forma también se registra el tiempo necesario para su tramitación.

Según una primera realización de la invención, las una o más cámaras de vídeo entregan las imágenes como una señal analógica y las imágenes procesadas se convierten en una señal analógica antes de alimentarse al al menos un monitor. En este caso, las imágenes de vídeo entrantes se proporcionan con una marca de tiempo y la marca de tiempo de la imagen de vídeo o, cuando se reproducen varias imágenes de vídeo simultáneamente, la marca de tiempo de todas las imágenes de vídeo reproducidas se lee y se compara cuando las imágenes procesadas se convierten en una señal analógica.

En la primera realización se intenta trabajar lo máximo posible con señales analógicas. La cámara de vídeo entrega señales analógicas (sin procesamiento digital previo), por lo que no se pueden producir retrasos. El monitor también funciona de forma analógica, por lo que muestra inmediatamente la señal de vídeo entrante. En este caso tampoco puede producirse ningún retraso.

Sin embargo, como se mencionó anteriormente, es difícil prescindir por completo de la digitalización, especialmente cuando se deben mostrar imágenes de varias cámaras de vídeo en un monitor. En este caso, la invención adopta el enfoque de que a cada imagen se le asigna una marca de tiempo durante la digitalización y que las imágenes almacenadas solo se utilizan si la marca de tiempo no es excesivamente antigua. Si se produce un retraso inesperado (por ejemplo, porque una imagen "permanece" en la memoria y no se sobrescribe con nuevas imágenes, por ejemplo, debido a un defecto en el convertidor analógico-digital), esto se reconoce y la imagen no se muestra en absoluto.

Según una segunda realización de la invención, la una o más cámaras de vídeo entregan las imágenes como una señal digital. En este caso, lo más seguro es que las imágenes de vídeo ya tengan una marca de tiempo en una o más cámaras de vídeo. Sin embargo, si una o más cámaras de vídeo entregan las imágenes como señales SDI, un desfase temporal es prácticamente imposible incluso en el caso de señales digitales: el sensor CCD utiliza un convertidor A/D para digitalizar los píxeles uno tras otro, y estos valores digitales se emiten en el protocolo SDI, ya sea a través de un cable coaxial o de una línea de dos hilos, e independientemente de los dispositivos receptores, de modo que también

en este caso es posible añadir únicamente la marca de tiempo a las imágenes de vídeo entrantes sin ningún riesgo de seguridad. (Aquí no hay prácticamente ninguna diferencia con las señales analógicas, solo el procesamiento en serie de los píxeles produce un pequeño pero constante desfase temporal del orden de 20 ms.) La ventaja de la segunda solución es que se pueden utilizar cámaras de vídeo SDI convencionales.

- 5 Por el contrario, las cámaras IP entregan datos a través de cables de red y tienen en cuenta todo el tráfico de datos en la red; y debido a que las redes pueden estar sobrecargadas y/o interrumpidas, pueden ocurrir retrasos impredecibles. Al utilizar cámaras IP, la marca de tiempo definitivamente debe generarse en la cámara.

10 Es útil que la imagen de vídeo se elimine de la memoria tan pronto como se haya procesado y enviado a un monitor. Esta es una medida de seguridad adicional para evitar que se muestren imágenes obsoletas. Incluso si una marca de tiempo se detecta accidentalmente como actual, aunque esté desactualizada, dicha imagen desactualizada se puede mostrar como máximo una vez, lo que difícilmente se puede notar (la frecuencia de actualización habitual es de 25 Hz). En el caso de la primera realización, la imagen de vídeo se borra de la memoria tan pronto como se ha convertido en una señal analógica.

15 En este caso, para señales analógicas, se deben utilizar al menos dos memorias, escribiéndose las señales de vídeo alternativamente en estas dos memorias, alternándose después de borrar cada imagen. Las señales de vídeo siempre se escriben en la última memoria eliminada. Si la memoria 1 se borra en el momento x , entonces las imágenes completas estarán disponibles en la memoria 1 en el momento $x + 40$ ms, y podrá comenzar la lectura y la conversión a una señal analógica. En este momento, la memoria 2 se borra y las señales de vídeo restantes se escriben en la memoria 2. Dado que la transferencia al monitor también tarda 40 ms, la memoria 1 se puede borrar en el momento $x + 80$ ms. En los siguientes 40 ms, la memoria 2 se convierte en una señal analógica y se transmite al monitor, y las señales de vídeo se almacenan nuevamente en la memoria 1.

20 Cuando se transmiten señales digitales al monitor, esta transmisión es extremadamente rápida en comparación con la frecuencia de refresco, de modo que una sola memoria es suficiente: porque si la transmisión y el borrado de imágenes son casi instantáneos, prácticamente están disponibles los 40 ms completos entre dos transmisiones de imágenes en donde se pueden almacenar las señales de vídeo, de modo que incluso cuando solo se utiliza una memoria, siempre se dispone de imágenes completas.

25 Adicional o alternativamente, es posible borrar las imágenes de vídeo de la memoria a intervalos regulares, por ejemplo, cada 200 ms.

30 En ambos casos, se puede aumentar la seguridad leyendo la imagen de vídeo después de eliminarla y, si no se ha eliminado por completo, borrándola nuevamente. De esta manera se reduce drásticamente la posibilidad de un borrado incorrecto.

35 Además de un desfase temporal no deseado, al mostrar imágenes de varias cámaras en un monitor también existe el riesgo de que se muestren las imágenes de la cámara incorrecta. Por ejemplo, en el metro de Viena está prevista una cámara para la mitad delantera del tren y una segunda cámara para la mitad trasera. Si debido a una conexión defectuosa, la imagen de la mitad trasera del tren se muestra accidentalmente dos veces en el monitor (o, peor aún, la imagen de un tren en el andén opuesto), esto no es inmediatamente evidente y existe el riesgo de que el conductor cierre las puertas mientras los pasajeros aún están subiendo o bajando.

40 Para evitar esto, otra característica de la invención prevé que las señales estén provistas de un código de identificación y que este código de identificación se compruebe para garantizar que se muestra la imagen correcta en el caso de una visualización en pantalla completa o que las imágenes individuales se muestran en el lugar correcto cuando las imágenes se muestran juntas en un monitor.

45 En el caso de la tecnología analógica, es decir en el caso de la primera realización de la invención, es habitual que las señales analógicas que llegan desde las cámaras de vídeo se introduzcan previamente en las entradas de una matriz de vídeo, la cual, en función del control, envía las señales analógicas a determinadas salidas. De esta manera, es posible conectar cualquier cámara de vídeo al monitor, o bien la imagen combinada de varias cámaras de vídeo, que también se alimenta a esta matriz de vídeo. Sin embargo, estos circuitos, que conmutan señales analógicas en función del control digital, pueden funcionar mal, especialmente en presencia de fuertes interferencias electromagnéticas. Por lo tanto, puede suceder que el conductor quiera encender la cámara de vídeo para la parte trasera del tren porque ha observado en la imagen combinada que allí puede haber una puerta bloqueada, pero después de cambiar de lugar ve en realidad una parte diferente del tren y supone erróneamente que todas las puertas están correctamente cerradas. Para evitar este tipo de fallos, una forma de realización de la invención prevé que este código de identificación se detecte después de la matriz de vídeo para comprobar si se ha conmutado la señal correcta.

55 Es más seguro si los códigos de identificación se modulan en las señales analógicas y se demodulan después de la matriz de vídeo. Como alternativa, los códigos de identificación podrían superponerse a la señal de vídeo, preferiblemente en un rango de frecuencia diferente, pero existe el riesgo de que los códigos de identificación se filtren.

Si el código de identificación ya está modulado en las cámaras de vídeo, entonces también es posible detectar si los cables de las cámaras de vídeo están conectados incorrectamente. Sin embargo, ya no se pueden utilizar cámaras estándares.

5 Si la señal de vídeo de las imágenes de vídeo es una señal de vídeo estándar con rellano, el código de identificación del rellano se puede imprimir en la señal de vídeo, preferiblemente después de la ráfaga de color (ráfaga de croma) en el caso de imágenes en color.

Si se modula un código de identificación, también se puede modular junto con él la marca de tiempo y decodificarla junto con él. Esto puede reducir potencialmente el esfuerzo del circuito.

10 En el caso de la tecnología digital, es decir, la segunda realización, es conveniente que la marca de tiempo y, si es necesario, el código de identificación se imprima en la imagen en forma de esteganograma. Esta es una técnica bien conocida, por ejemplo, se puede utilizar el LSB (bit menos significativo) de algunos píxeles de color para almacenar la marca de tiempo y el código de identificación. El hecho de que la imagen sea ligeramente más ruidosa como resultado solo se nota con cámaras de calidad extremadamente alta y la mejor iluminación.

Breve descripción de las figuras del dibujo

15 La presente invención se explica con más detalle mediante un ejemplo de realización. En los dibujos, la Fig. 1 muestra un circuito básico según una primera realización con el que se puede llevar a cabo el método según la invención; La Fig. 2 muestra una señal de televisión analógica con un código de identificación en relieve que se puede utilizar en la primera realización; y la Fig. 3 muestra un circuito esquemático de una segunda realización.

Descripción de los tipos de realización

20 Varias cámaras 11 de vídeo (en el ejemplo hay cuatro cámaras de vídeo) están conectadas a una matriz 13 de vídeo. Esta matriz 13 de vídeo es un interruptor electrónico controlado digitalmente que puede enviar una señal desde cualquier entrada a cualquier salida. Al menos un monitor 16 está conectado a esta matriz 13 de vídeo (en el ejemplo hay dos monitores). La matriz 13 de vídeo está controlada por un controlador 14. Por ejemplo, se puede controlar que el primer monitor 16 muestre la imagen de la cuarta cámara 11 de vídeo y el segundo monitor 16 muestre la imagen de la primera cámara 11 de vídeo. Sin embargo, ambos monitores 16 también pueden mostrar la misma imagen.

Tanto las cámaras 11 de vídeo como los monitores 16 son dispositivos analógicos, es decir, las cámaras 11 de vídeo entregan señales analógicas y los monitores 16 procesan señales analógicas sin que las imágenes sean digitalizadas en estos dispositivos. Por lo tanto, cuando un monitor 16 muestra la imagen de una cámara 11 de vídeo, no puede haber ningún retraso temporal.

30 Sin embargo, a menudo se desea que las imágenes de cuatro cámaras 11 de vídeo se muestren simultáneamente en un monitor 16 en cuatro cuadrantes (o las imágenes de dos cámaras 11 de vídeo en dos mitades). Para lograr esto, las señales analógicas de la matriz 13 de vídeo se alimentan a un convertidor 17 analógico-digital, donde se digitalizan y se alimentan a una FPGA 18 (FPGA = matriz de puerta programable en campo).

35 Con la ayuda de matrices de puertas programables en campo rápido, denominadas FPGA, se consiguen tiempos de respuesta del orden de nanosegundos. Esta reacción rápida no es posible con los procesadores de señales convencionales porque los procesadores de señales procesan las señales entrantes secuencialmente, según las especificaciones del programa. Las FPGA, por otro lado, son *hardware* programable; se pueden ejecutar hasta 100 procesos en paralelo a través del componente, independientemente de los pasos individuales, y el procesamiento es hasta 500 veces más rápido que en un procesador de señales.

40 Este FPGA 18 ahora almacena los datos suministrados por el convertidor 17 analógico-digital en una memoria 19 (RAM), de modo que se almacenan cuatro imágenes de cuatro cámaras 11 de vídeo. A partir de estas imágenes, la FPGA 18 ensambla nuevas imágenes (en el ejemplo hay dos imágenes), por ejemplo, las imágenes de las cuatro cámaras 11 de vídeo en los cuatro cuadrantes de una nueva imagen y las imágenes de dos de las cuatro cámaras 11 de vídeo en las dos mitades de otra nueva imagen. Estas nuevas imágenes, que están disponibles digitalmente en la FPGA, se convierten luego nuevamente en señales analógicas en un convertidor 20 digital-analógico y se alimentan a la matriz 13 de vídeo. De este modo, cada monitor 16 puede mostrar una imagen de cuatro cuadrantes, una imagen de dos mitades o la imagen de cada cámara 11 de vídeo como una imagen completa, dependiendo de qué señales la matriz 13 de vídeo aplique a la salida correspondiente.

50 La invención consiste ahora en que la FPGA 18 almacena una marca de tiempo para cada imagen al almacenar las imágenes y comprueba esta marca de tiempo cuando ensambla la nueva imagen. Si la marca de tiempo es demasiado antigua, la imagen guardada se ignora, es decir, la parte correspondiente de la nueva imagen (un cuadrante o la mitad) permanece oscura. (Por supuesto, también se puede mostrar un color o un patrón). Esto hace imposible que la nueva imagen se componga en base a datos obsoletos.

La invención evita que se muestre información obsoleta. Sin embargo, huelga mencionar que esto solo debería ocurrir en raras ocasiones. Por lo tanto, todo el sistema está diseñado de tal manera que los retrasos suelen ser muy escasos.

Como medida de seguridad adicional, la FPGA 18 borra los datos de la memoria 19 tan pronto como se han leído. De esta forma, cada imagen de cada cámara 11 de vídeo solo puede utilizarse una vez para construir una nueva imagen.

5 La matriz 13 de vídeo representa un riesgo de seguridad adicional, ya que se han reportado casos en donde dicha matriz de vídeo ha conectado incorrectamente las entradas y salidas cuando se han producido fuertes campos de interferencia. Esto puede provocar que el conductor vea en el monitor 16 la imagen de una cámara 11 de vídeo diferente a la que ha configurado.

10 Para eliminar este riesgo de seguridad, se proporcionan moduladores 12 delante de la matriz 13 de vídeo y demoduladores 15 detrás de la matriz 13 de vídeo. Los moduladores 12 imprimen un código de identificación en la señal de vídeo, que puede ser decodificado por los demoduladores 15. El controlador 14 ahora informa a cada demodulador 15 qué código de identificación se debe esperar en función del control de la matriz 13 de vídeo. Si el código de identificación no coincide con el valor objetivo, el demodulador 15 interrumpe la señal de vídeo. En este caso tampoco se muestra ninguna visualización en caso de error, lo que es definitivamente mejor que una visualización
15 falsa.

La superposición de señales adicionales en una señal de vídeo está muy extendida. Todos los sistemas de televisión en color utilizan este método para transmitir una ráfaga de color para la sincronización. Soluciones similares se describen en gran número en la literatura de patentes (grupo H04N7/025 de la clasificación internacional).

20 La superposición del código de identificación se explica en la Fig. 2. Este código de identificación está impreso en la señal de vídeo, es decir, en el rellano. Una señal FBAS clásica en modo de línea tiene un pulso 21 de sincronización de línea, que tiene un rellano 22 avanzado y un rellano 23 retardado. Los rellanos 22 y 23 se utilizan como referencia para una imagen negra, los valores se utilizan en la pantalla para la llamada sujeción. El contenido de la imagen está codificado en la posición 24. En el rellano 23 retardado hay una ráfaga 25 de color que transporta las señales de referencia para la modulación de color. El código 26 de identificación se escribe preferiblemente en la señal FBAS
25 después de la ráfaga 25 de color. El código 26 de identificación puede consistir en una codificación digital en el rango de nanosegundos porque los componentes modernos pueden procesar fácilmente estas velocidades.

Una solución específica utilizó una FPGA 18 moderna y de alto rendimiento que lee los datos digitales de las cuatro cámaras 11 de vídeo en paralelo y asigna una marca de tiempo a todas las señales. Durante la lectura, las imágenes se preparan para su posterior posición en la pantalla, se elimina el contenido innecesario de la imagen y la imagen de cada cámara 11 de vídeo se escribe en un área correspondiente de la memoria 19.
30

Luego, un grupo de circuitos en la FPGA 18 restaura los datos necesarios para una imagen combinada, siempre que las marcas de tiempo no sean demasiado antiguas. Para aumentar la seguridad y marcar fallos, la memoria de salida se borra inmediatamente después de leer y crear las imágenes compuestas, de modo que no se puedan mostrar imágenes estáticas.

35 Mediante las medidas descritas es posible limitar y comprobar el tiempo de procesamiento de las señales de vídeo dentro de la FPGA a un máximo de 100 ms (milisegundos) y comprobar el correcto cableado en la matriz 13 de vídeo. El tiempo de procesamiento de 100 ms es resultado del estado operativo asíncrono: una señal de vídeo analógica requiere 40 ms para toda la transmisión de la imagen (a 25 Hz). Dado que las cámaras 11 de vídeo no están sincronizadas, pueden pasar un máximo de 80 ms hasta que se guarden los datos de todas las cámaras 11 de vídeo.
40 La imagen puede permanecer en la memoria 19 durante este tiempo, con un margen de seguridad de 20 ms. Todo lo que permanezca en la memoria 19 durante un período más largo se considera obsoleto y ya no se puede mostrar. Este tiempo de 100 ms es como máximo la mitad del tiempo de reacción de una persona atenta. Esto significa que se considera que la transmisión es suficientemente oportuna para realizar los controles.

45 Para evitar adicionalmente que partes de imágenes de datos de imágenes anteriores permanezcan en la memoria 19, la memoria 19, en donde la FPGA 18 ensambla las imágenes en una imagen dual o una imagen de cuadrante, se borra después de la lectura y se escribe con "0". Este valor cero corresponde a una pantalla negra cuando se convierte a una señal de vídeo. En un paso inmediatamente siguiente, se lee nuevamente la memoria para garantizar que se haya borrado su contenido. Esta lectura para fines de control solo se procesa internamente; es posible que no haya más información disponible. En caso de mal funcionamiento, se transmitiría una segunda lectura en forma de imagen,
50 un contenido de imagen negra, en la pantalla, haciendo que el error fuera reconocible como tal incluso para un observador normal.

A diferencia de la primera realización, que se explicó con referencia a las figuras 1 y 2 y en donde se utiliza tecnología analógica siempre que es posible, la segunda realización, que se explica con referencia a la Fig. 3, hace uso de tecnología digital.

55 La Fig. 3 muestra el principio de un espejo retrovisor electrónico, en donde en el caso más sencillo se prevén tres cámaras 11, que corresponden al espejo exterior izquierdo, al espejo exterior derecho y al espejo interior. Estas tres imágenes se combinan y se muestran en un monitor 16.

La interfaz "SDI" sirve como base para la transmisión de imágenes digitales, en particular HD-SDI (SMPTE 292M 1,485 Gbit/s y 1,485/1,001 Gbit/s, 720p, 1080i (HDTV)) y 12G UHD-SDI (SMPTE ST 2082 12 Gbit/s, 4K). Esto significa que los datos de vídeo digital se transmiten desde los sensores de la cámara a una velocidad de hasta 12 Gbit/s a través de un cable coaxial con una impedancia de 75 ohmios o un par trenzado de dos hilos (CAT7).

- 5 Los valores de la matriz de imagen se convierten en un flujo de datos en serie digital, mediante el cual el valor de brillo y el valor de color de cada píxel se transmiten como valores digitales, normalmente se utilizan 20 bits para esto (10 bits para brillo y 10 bits para color).

10 Según la invención, el LSB (LSB= bit menos significativo) de algunos de estos valores no se utiliza para proporcionar información sobre el brillo o el color, sino para otra información, a saber, una marca de tiempo y, si es necesario, un código de identificación. Como con 10 bits se pueden representar 1024 valores, la información posible se reduce a 9 bits, es decir, 512 valores. Pero en la práctica esto es más que suficiente. Dado que el LSB utilizado para la marca de tiempo es "aleatorio" en términos de brillo o color, hay un mínimo de ruido de brillo o color, que, sin embargo, no se nota de ninguna manera durante la visualización normal. En la mayoría de los casos, el ruido del sensor es mayor de todos modos.

- 15 (Nota: los valores 0-3 y 1020-1023 están reservados y no se pueden utilizar para valores de brillo o color. Estrictamente hablando, no hay 1024, sino solo 1016 valores disponibles).

20 Según la invención, en las imágenes de las cámaras se imprime una marca de tiempo y, en su caso, un código de identificación; Como esto no se puede ver con la observación normal, se llama esteganografía. La ventaja es que el protocolo SDI normal sigue siendo utilizable sin el más mínimo cambio y que esta información se conserva automáticamente cuando la imagen se deserializa y se guarda.

25 Los datos digitales de varias cámaras se deserializan ahora en la unidad de procesamiento de imágenes, que a su vez está implementada por una FPGA 18, y se almacenan en una memoria 19. Así, las marcas de tiempo siguen estando disponibles. Si las imágenes en el caso de un espejo retrovisor digital (normalmente hay al menos tres imágenes) se corrigen ("*dewarping*") y se unen sin fisuras ("*stitching*"), se podría perder la información esteganográfica. Por lo tanto, es aconsejable extraer primero la información incrustada esteganográficamente y almacenarla en unos pocos bytes adicionales de cada imagen. Cuando la imagen compuesta se serializa y se transmite en el protocolo SDI al monitor 16 (es decir, al "espejo retrovisor electrónico"), se puede comprobar la marca de tiempo de las tres imágenes en ese momento; Si la marca de tiempo es demasiado antigua en comparación con el propio "reloj", se emite una imagen negra en el monitor 16, por ejemplo, de modo que resulta evidente inmediatamente que no hay ninguna imagen actual disponible. El código de identificación también se puede utilizar para comprobar si, por ejemplo, el contenido de la imagen que se muestra a la izquierda proviene realmente de la cámara montada en el extremo izquierdo.

30 Además, se proporciona un controlador 14 con el que se puede controlar la FPGA. Este control 14 se puede utilizar, por ejemplo, para especificar que solo se muestre en el monitor en pantalla completa la imagen de la cámara correspondiente al espejo interior o a uno de los espejos exteriores, lo que en ocasiones resulta útil al aparcar.

- 35 Aunque prácticamente no se pueden producir retrasos al visualizar una señal SDI en el monitor 16, la imagen compuesta también puede estar provista de una marca de tiempo, que luego puede comprobarse en el monitor 16. Si la imagen de una de las cámaras se muestra en pantalla completa, la marca de tiempo estará presente de todos modos automáticamente como resultado de la incrustación esteganográfica.

REIVINDICACIONES

1. Método de lectura y edición de imágenes en tiempo real con un espejo retrovisor electrónico,
presentando el espejo retrovisor al menos tres cámaras (11) de vídeo que suministran imágenes como señal digital en serie en forma de señal SDI,
- 5 correspondiendo las cámaras de vídeo a un espejo exterior izquierdo, un espejo exterior derecho y un espejo interior, almacenándose las imágenes en una memoria (19), ensamblándose y enviándose las imágenes a al menos un monitor (16), estando todas las imágenes de las cámaras (11) de vídeo provistas de una marca de tiempo en el bit menos significativo de la señal digital,
- 10 deserializándose las marcas de tiempo junto con las imágenes en una unidad de procesamiento de imágenes implementada por una FPGA (18) y almacenándose las marcas de tiempo en la memoria (19) junto con las imágenes, proporcionándose un controlador mediante el cual se puede controlar la FPGA (18), corrigiéndose y ensamblándose las imágenes sin fisuras y serializándose después la imagen ensamblada,
- 15 leyéndose a continuación la marca de tiempo de todas las imágenes renderizadas y comparándose con la hora actual y enviándose una señal para transmitir las imágenes al al menos un monitor solo al al menos un monitor (16) si todas las marcas de tiempo no se remontan a un período de tiempo máximo predefinido.
2. Método según la reivindicación 1, caracterizado por que las imágenes guardadas se editan después de proporcionarles la marca de tiempo, por que la marca de tiempo se lee después de la edición y por que las imágenes editadas se envían al al menos un monitor (16).
- 20 3. Método según cualquiera de las reivindicaciones 1 y 2, caracterizado por que las señales están provistas de un código (26) de identificación y por que dicho código (26) de identificación se verifica para garantizar que se muestre la imagen correcta en la visualización de pantalla completa y/o por que las imágenes individuales se muestren en sus posiciones correctas cuando las imágenes se muestran juntas en un monitor.

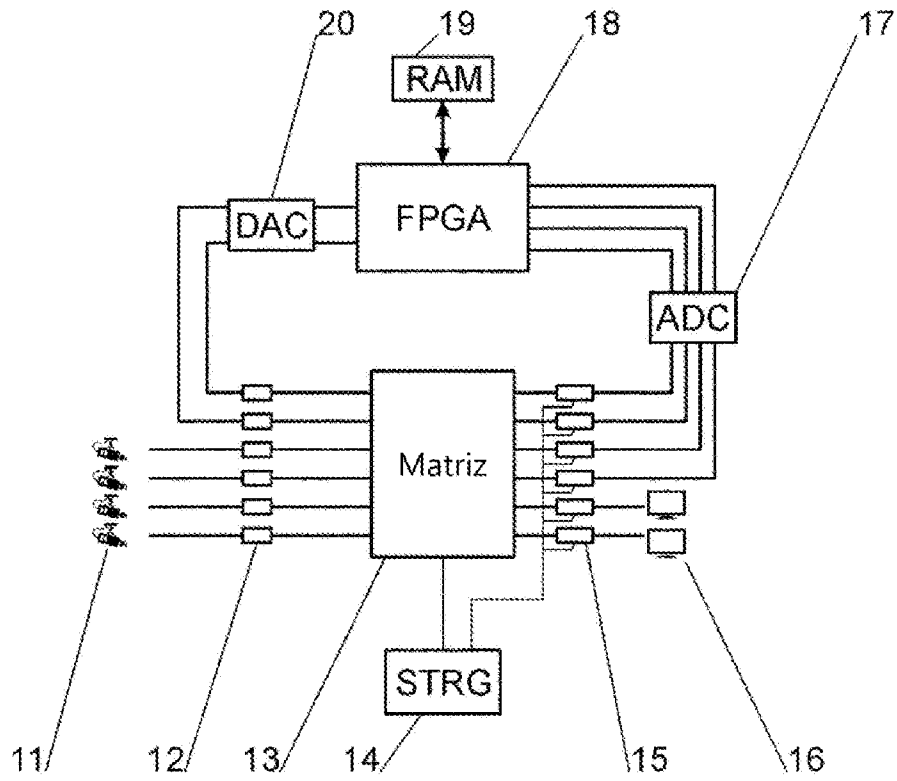


Fig. 1

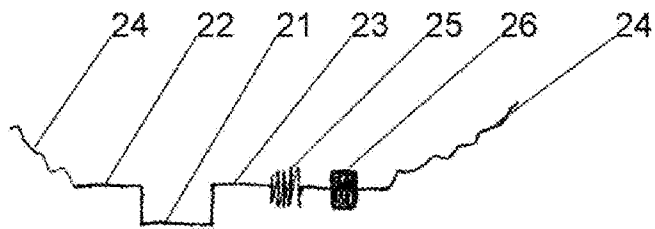


Fig. 2

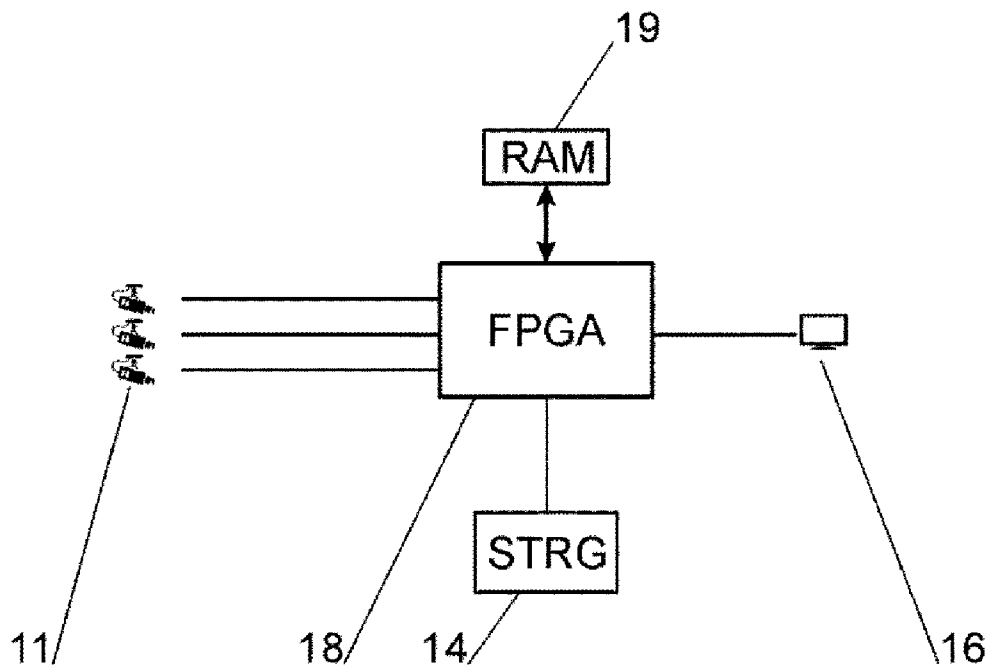


Fig. 3