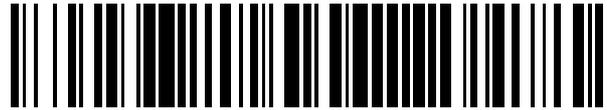


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 693 497**

51 Int. Cl.:

H04N 5/262 (2006.01)

H04N 5/272 (2006.01)

H04N 5/265 (2006.01)

H04N 5/341 (2011.01)

G06T 3/40 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.06.2015 E 15382315 (8)**

97 Fecha y número de publicación de la concesión europea: **22.08.2018 EP 3107279**

54 Título: **Procedimiento, aparato e instalación para componer una señal de video**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
12.12.2018

73 Titular/es:

COHERENT SYNCHRO, S.L. (100.0%)
Av. de Roma, 97
08029 Barcelona, ES

72 Inventor/es:

ARAGONÉS GAYA, JOSÉ MARÍA;
BASURÇO MANCISIDOR, SERGIO ANDRÉS;
MARÍ GÓMEZ, MANUEL y
MATILLA SARRATE, MARTÍN

74 Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

ES 2 693 497 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).

DESCRIPCION

Procedimiento, aparato e instalación para componer una señal de video

5 Sector técnico de la invención

La presente invención se refiere a un procedimiento, un aparato y una instalación para componer una señal de video de salida a partir de una pluralidad de señales de video de entrada.

10 Antecedentes de la invención

15 Son conocidos procedimientos para componer señales de video de salida a partir de una pluralidad de señales de video de entrada, por ejemplo una pluralidad de señales de video de entrada que provienen de diferentes cámaras de video y que visualizan partes de una escena. En estos procedimientos, los fotogramas de cada señal de vídeo de entrada se disponen en un espacio virtual, a modo de lienzo, mediante la técnica conocida como video stitching, de modo que usuarios puedan visualizar dicho lienzo a través de una ventana de visualización. El contenido visualizado por la ventana de visualización se codifica convenientemente para formar la señal de salida a la resolución predeterminada.

20 No obstante, en los procedimientos conocidos, tal como el descrito en el documento de patente US2007/0076099A1, la señal de salida se obtiene codificando todos los fotogramas de cada señal de vídeo de entrada, siendo necesario decodificar y recodificar todas las señales de video de entrada, con lo que se puede ser necesaria una alta potencia de procesado para decodificar y recodificar señales de entrada de alta resolución para evitar retrasos en la señal de salida.

25 El documento EP2037671 presenta un aparato de procesado de video y un método para mezclar una pluralidad de fotogramas de entrada en los que se han seleccionado previamente áreas en un fotograma de salida. Si áreas seleccionadas en los fotogramas de entrada cubren áreas que se superponen en el fotograma de salida, la imagen de salida se genera combinando los píxeles que se superponen.

30 El documento US6985638 presenta un método para componer imágenes que se superponen parcialmente y que contienen al menos un objeto común tomado desde un número de posiciones por un dispositivo captador de imágenes tal como una cámara digital o un escáner de imágenes, en una única imagen para incrementar la resolución en la captura de un objeto grande. La composición de las imágenes que se superponen parcialmente en una única imagen se realiza tras corregir por transformación de perspectiva las imágenes que se superponen parcialmente según unos parámetros de transformación de perspectiva que usan estructuras de datos piramidales.

35 Es por tanto un objetivo de la presente invención dar a conocer un procedimiento, aparato e instalación que permita generar señales de video en tiempo real a partir de una pluralidad de señales de video de entrada, y que permita utilizar señales de entrada a alta resolución sin que necesariamente conlleve acumular retrasos ni requerir de una gran potencia de computación.

Explicación de la invención

45 El procedimiento para componer una señal de video de salida de la presente invención, es de los que, a partir de una pluralidad de señales de video de entrada, que comprenden respectivas secuencias de fotogramas de entrada con un número de píxeles para cada fotograma de entrada asociado a una resolución de entrada permite componer una señal de video de salida, comprendiendo la señal de video de salida una secuencia de fotogramas de salida con un número de píxeles prestablecidos para cada fotograma de salida asociado a una resolución de salida. El procedimiento permite ventajosamente que a partir de diferentes señales de video, que por ejemplo representen diferentes partes de una escena, se forme una señal de video de salida a partir de partes de los fotogramas de las señales de entrada.

50 El procedimiento comprende una etapa de configuración en la que se determina la superficie espacial sobre un espacio virtual tridimensional formado por un conjunto de puntos con coordenadas tridimensionales asociada a cada píxel de cada fotograma de cada secuencia de fotogramas de entrada, establecer en el espacio virtual tridimensional un punto de visualización y una ventana de visualización, conocida también como "viewport", que determinan un tronco de visualización, conocido también como "frustrum", en el espacio virtual tridimensional, formándose fotogramas de salida a partir de las proyecciones de los fotogramas de las secuencias de entrada proyectadas en dicha ventana de visualización; calcular una densidad de píxel umbral asociada para los píxeles en la ventana de visualización, a partir de la resolución de salida y las dimensiones espaciales asociadas de dicha ventana de visualización; determinar combinaciones de fotogramas de entrada cuyas proyecciones quedan al menos parcialmente superpuestas en la ventana de visualización y, para cada combinación de fotogramas, seleccionar el fotograma de cada combinación de fotogramas que proyectado sobre la ventana de visualización tiene una densidad de píxel en toda la parte superpuesta mayor o igual a la densidad de píxel umbral o, en su defecto, seleccionar el fotograma que proyectado sobre la ventana de visualización tenga una mayor densidad de píxel para formar la parte de fotograma de salida que queda superpuesta. De esta manera, ventajosamente no se tiene que procesar el contenido de las señales de entrada de

mayor resolución que no vayan a ser utilizadas para codificar la señal de salida, puesto que ya existirá una señal de entrada de menor resolución que se podrá utilizar para codificar la señal de salida a la resolución de salida precisada, sin perder calidad.

5 En otra variante de realización, si en alguna de las combinaciones de fotogramas existe una pluralidad de fotogramas que proyectados sobre la ventana de visualización tienen una densidad de píxel en toda la parte superpuesta mayor o igual a la densidad de píxel umbral, se selecciona de entre esa pluralidad de fotogramas el fotograma que proyectado sobre la ventana de visualización tienen una densidad de píxel más próxima a la densidad de píxel umbral. De esta manera si hubiera una pluralidad de fotogramas que proporcionaran una resolución suficiente, solamente se toma
10 aquella inmediatamente superior a la densidad de píxel umbral, que es la que puede procesarse más fácilmente.

Se prevé que el punto de visualización y el tamaño de la ventana de visualización y por consiguiente el tronco de visualización, puedan modificarse mediante controles, por ejemplo controles PTZ recibidos de un usuario de modo conocido, de modo que la ventana de visualización pueda desplazarse y en ella puedan proyectarse otras regiones del espacio virtual tridimensional, actuando así a modo de cámara virtual de visualización del espacio virtual tridimensional. Naturalmente, cuando cambien punto de visualización del espacio virtual tridimensional, el tamaño de ventana o el tronco de visualización puede ser necesario tener que volver a seleccionar los fotogramas de las combinaciones de fotogramas para que así la señal de salida tenga la resolución adecuada. También se prevé que
15 pueda haber más de una cámara virtual de visualización del espacio virtual tridimensional, por ejemplo si más de un usuario tiene que visualizar el espacio virtual tridimensional. En este caso, el procedimiento para componer la señal de video de salida se tendrá que realizar para la señal de salida que se enviará a cada usuario.

Según una variante de realización, para formar la parte de fotograma de salida que queda superpuesta se realiza una mezcla entre el fotograma seleccionado de cada combinación de fotogramas y al menos otro fotograma de la misma combinación de fotogramas, por ejemplo cuando en la ventana de salida se proyectan dos o más fotogramas de entrada que solamente quedan parcialmente superpuestos, de modo que se consiga una transición, por ejemplo en forma de gradiente progresivo o máscara binaria entre los diferentes fotogramas de entrada que formarán el fotograma de salida. Esta transición puede realizarse solamente en los extremos de la parte que queda superpuesta.
25

30 Según otra variante de realización, para formar la parte de fotograma de salida que queda superpuesta se utiliza solamente el fotograma seleccionado de cada combinación de fotogramas, por ejemplo cuando el fotograma seleccionado de una combinación de fotogramas cubre toda la ventana de visualización, y por tanto no es necesario realizar una transición.

35 Según otra variante de realización, la densidad de píxel de cada fotograma proyectado sobre la ventana de visualización se calcula a partir de la proyección de agrupaciones de píxeles del fotograma, preferentemente las agrupaciones de píxeles pueden ser divisiones triangulares de cada fotograma.

40 Según otra variante de realización, el paso de seleccionar el fotograma de cada combinación de fotogramas se realiza evaluando progresivamente los fotogramas, empezando primero por el fotograma de menor resolución de entrada y acabando por el de mayor resolución de entrada, de modo que se tenga que procesar primero los fotogramas de menor resolución, que pueden procesarse más rápidamente.

45 Según otra variante de realización, las superficies espaciales asociadas a cada píxel de cada fotograma de cada secuencia de fotogramas de entrada en el espacio virtual tridimensional conforman el interior de una esfera, de modo que los fotogramas queden equidistantes del punto de visualización.

50 Según otra variante de realización, al menos una de las señales de video de entrada es una señal de video generada por una cámara de vídeo a partir de una escena, de modo que la señal de salida incorpore fotogramas tomados en tiempo real.

55 Según otra variante de realización, al menos dos de las señales de video están adaptadas para formar secuencias de fotogramas con al menos parte del mismo contenido a diferentes resoluciones, de modo que si una señal de vídeo representa una vista general de una escena y otra señal de vídeo un detalle de dicha escena a mayor resolución, cuando se desee que la señal de salida tenga una resolución similar a la de la vista general de la escena, por ejemplo cuando la ventana de visualización sea ancha, no sea necesario procesar la otra señal de vídeo con el detalle. Del mismo modo, cuando se desee visualizar solamente el detalle de dicha escena, ampliándolo, con lo que la ventana de visualización será más estrecha, será necesario utilizar la otra señal de vídeo con el detalle para formar la señal de salida.
60

Según otra variante de realización, al menos dos de las señales de video determinan secuencias de fotogramas con el mismo contenido a diferentes resoluciones, de modo que una señal de vídeo pueda sustituir completamente a la otra para formar la señal de salida.

65 Según otra variante de interés, al menos dos señales de video provienen de una misma cámara de vídeo, de modo que puedan utilizarse las salidas a diferente resolución de una cámara como señales de entrada.

Según otra variante de realización, al menos una de las señales de video está adaptada para formar fotogramas cuyo contenido es una representación virtual de al menos parte del contenido del fotograma de otra de las señales de video, de modo que esta representación virtual pueda generarse a partir de un conjunto de datos, unos elementos vectoriales o un modelo tridimensional, de modo que la resolución de dicha representación virtual pueda ser tan alta como se desee. De esta manera, si se conoce que parte del contenido de un fotograma de una señal de video que provenga de la captura de una cámara de video tendrá una imagen determinada, y esta parte puede ser modelada, esta parte modelada puede ser proporcionada en forma de otra señal de video, de modo que quede dispuesta y sustituir la parte del fotograma con el mismo contenido cuando no permita proporcionar la resolución de salida. Por ejemplo la representación virtual puede ser el contenido de una pantalla que se visualiza de forma general en otro fotograma de otra señal de entrada, de modo que si es necesario se pueda remplazar el contenido de la pantalla en la visualización general por la representación virtual.

El procedimiento de la presente invención se puede codificar en un programa de ordenador, estando dicho programa de ordenador adaptado para realizar el procedimiento cuando se ejecuta en un ordenador.

Se da a conocer también un aparato para procesar datos que comprende medios para realizar el procedimiento de la presente invención, pudiendo formar dicho aparato de una instalación que comprenda además al menos una cámara de video, de modo que la salida de la cámara de video represente una de las señales de entrada del aparato.

Breve descripción de los dibujos

Para complementar la descripción que se está realizando y con objeto de facilitar la comprensión de las características de la invención, se acompaña a la presente memoria descriptiva un juego de dibujos en los que, con carácter ilustrativo y no limitativo, se ha representado lo siguiente:

La Fig. 1 representa un esquema de funcionamiento de un sistema que implementa el procedimiento de la presente invención;

la Fig. 2 representa un espacio virtual tridimensional tras la fase de configuración con una ventana de visualización;

la Fig. 3 representa el contenido de la ventana de visualización de la Fig. 2;

las Figs. 4a y 4b representan la comparación entre la densidad de píxel umbral y la densidad de píxel de cada fotograma en la ventana de visualización de la Fig. 3;

la Fig. 5 representa la composición del fotograma de salida formado a partir del contenido de la ventana de visualización de la Fig. 3;

la Fig. 6 representa la visualización mediante otra ventana de visualización del espacio virtual tridimensional de la Fig. 2;

la Fig. 7 representa el contenido de la ventana de visualización de la Fig. 6;

las Figs. 8a y 8b representan la comparación entre la densidad de píxel umbral y la densidad de píxel de cada fotograma en la ventana de visualización de la Fig. 7;

la Fig. 9 representa la composición del fotograma de salida formado a partir del contenido de la ventana de visualización de la Fig. 7;

la Fig. 10 representa otro espacio virtual tridimensional tras la fase de configuración con una ventana de visualización;

la Fig. 11 representa el contenido de la ventana de visualización de la Fig. 10; y

la Fig. 12 representa la composición del fotograma de salida formado a partir del contenido de la ventana de visualización de la Fig. 11.

Descripción detallada de los dibujos

La Fig. 1 muestra un esquema de funcionamiento de un sistema que implementa el procedimiento de la presente invención, que a partir de una pluralidad de señales de video de entrada $S_{11}, S_{12}, S_{13} \dots S_{1m}$, que comprenden respectivas secuencias de fotogramas de entrada $S_{11}[n], S_{12}[n], S_{13}[n] \dots S_{1m}[n]$, en diferentes espacios temporales, con un número de píxeles para cada fotograma de entrada asociado a una resolución de entrada $R_{11}, R_{12}, R_{13} \dots R_{1m}$ permite componer una señal de video de salida S_o , que comprende una secuencia de fotogramas de salida $S_o[n]$ con un número de píxeles preestablecidos para cada fotograma de salida asociado a una resolución de salida R_o . Las resoluciones de entrada $R_{11}, R_{12}, R_{13} \dots R_{1m}$ pueden ser diferentes a la resolución de salida R_o .

En el ámbito de la invención, el término resolución se considera el número máximo de píxeles, normalmente indicados como píxeles verticales y píxeles horizontales, que pueden obtenerse de un fotograma de las secuencias de fotogramas de una señal de video.

En función de cuál sea el número de píxeles de un fotograma por unidad de longitud cuando este se represente, por ejemplo en una pantalla, se puede determinar otro parámetro, la densidad de píxel. Usualmente, las unidades que se utilizan para designar la densidad de píxel son los píxeles por pulgada (ppi), aunque cualquier otra medida de longitud real o virtual puede utilizarse. La densidad de píxel será por tanto inversamente proporcional al tamaño de píxel, que serían las unidades de longitud que ocuparía dicho píxel. Naturalmente esta densidad de píxel y tamaño de píxel se puede calcular cuando el fotograma se represente en un espacio virtual tridimensional, como se verá más adelante.

En el ejemplo representado, se indican una pluralidad de señales de video de entrada $S_{i1}, S_{i2}, S_{i3} \dots S_{im}$, que comprenden respectivas secuencias de fotogramas de entrada $S_{i1}[n], S_{i2}[n], S_{i3}[n] \dots S_{im}[n]$ con un número de píxeles para cada fotograma. Naturalmente se contempla que cualquier otro número de señales de video de entrada podría utilizarse y presentar varias resoluciones.

Mediante el procedimiento de la presente invención, que se implementa por el sistema mostrado en la Fig. 1, se consigue que partiendo de señales de video de entrada $S_{i1}, S_{i2}, S_{i3} \dots S_{im}$, que comprenden respectivas secuencias de fotogramas de entrada $S_{i1}[n], S_{i2}[n], S_{i3}[n] \dots S_{im}[n]$ con diferentes resoluciones, se consiga codificar una señal de video de salida S_o , cuyos fotogramas comprenderán un número de píxeles, verticales y horizontales, predeterminados que determinan una resolución de salida R_o . Esta resolución de salida R_o puede ser predeterminada por un operador o podría ser seleccionada durante el funcionamiento del sistema. Ejemplos de las posibles resoluciones, tanto de entrada $R_{i1}, R_{i2}, R_{i3} \dots R_{im}$ como de salida R_o serían las conocidas en el estado de la técnica, usualmente preparadas para su presentación en pantallas de misma resolución, tanto las consideradas como alta o baja definición, tales como 480i (243×243 píxeles), 576i (288×288 píxeles), 480p (720×480 píxeles), 576p (720 × 576), 720p (1,280 × 720 píxeles), 1080i o 1080p (1,920 × 1,080 píxeles), 2160p (3,840 × 2,160 píxeles) -4K UHD-, 4320p (7,680 × 4,320 píxeles) -8K UHD- así como cualquier otra resolución.

Como se puede observar en la Fig. 2, durante la etapa de configuración se determina la superficie espacial sobre un espacio virtual tridimensional 1 formado por un conjunto de puntos con coordenadas tridimensionales X,Y,Z asociada a cada píxel de cada fotograma de cada secuencia de fotogramas de entrada. A modo de ejemplo, la superficie espacial que ocupan los fotogramas de las secuencias de fotogramas de entrada mostrados se encuentran dispuestos sobre un plano, para facilitar la explicación de la invención. Naturalmente también se podrían disponer siguiendo otras formas geométricas, tal y como se verá más adelante. Esta forma podría por ejemplo reseguir el interior de una esfera, estando el punto de visualización dispuesto en el centro de la esfera, o la forma podría venir dada por un modelo virtualizado de un recinto, tal como una habitación, de modo que si las diferentes señales de video de entrada provienen de partes de dicha habitación, los fotogramas puedan disponerse sobre las partes de dicha habitación virtualizadas, de modo que un operario pueda desplazarse a modo de cámara virtual por la habitación virtualizada, ajustándose convenientemente el ángulo de visualización de cada uno de las secuencias de fotogramas que provengan de las cámaras, que habrán sido previamente dispuestas sobre el modelo tridimensional en la etapa de configuración.

Para visualizar el espacio virtual tridimensional, se establecerá en el espacio virtual tridimensional un punto de visualización (X_o, Y_o, Z_o) y una ventana de visualización 3, conocida también como viewport, determinando un tronco 4 de visualización en el espacio virtual tridimensional, conocido como frustrum, del modo ilustrado en la Fig. 2. De esta manera se formarán los fotogramas de salida a partir de las proyecciones de los fotogramas de las secuencias de entrada proyectadas en dicha ventana de visualización 3 en el mismo instante temporal, convenientemente codificadas. En caso que el número de fotogramas por segundo de las señales de entrada proyectadas en la ventana de visualización 3 sea diferente al número de fotogramas por segundo de la señal de salida, se deberán descartar algunos fotogramas de entrada si estos son más rápidos que el número de fotogramas por segundo de la señal de salida o bien mantener por más de un instante los fotogramas de entrada si son más rápidos que el número de fotogramas por segundo de la señal de salida, de modo conocido.

Ventajosamente, el procedimiento permite que solamente se deban decodificar para formar la señal de video de salida las señales de entrada que proporcionen fotogramas cuyas proyecciones $S_1'[n], S_2'[n]$ sean visibles en la ventana de proyección. Además, si existen combinaciones de fotogramas de entrada cuyas proyecciones queden al menos parcialmente superpuestas en la ventana de visualización 3, es posible que no se tengan que procesar, es decir, decodificar, todas las señales de entrada que de los fotogramas superpuestos para recodificar la señal de salida.

Cuando existan combinaciones de fotogramas de entrada, por ejemplo en la ventana de visualización 3 ilustrada en la Fig. 3, se deberá primeramente calcular una densidad de píxel umbral $d(S_o[n])$ asociada para los píxeles en la ventana de visualización 3, a partir de la resolución de salida R_o y las dimensiones espaciales asociadas de dicha ventana de visualización 3. Esta densidad de píxel umbral, será en las Figs. 3, 4a y 4b la relación entre el número de píxeles que vendrá dada por la resolución de salida y la superficie de la ventana de visualización 3, es decir, dividiendo el número de píxeles totales equivalentes a la resolución de salida, entre la superficie que ocupa la ventana de visualización 3 en el espacio virtual tridimensional. De esta manera se obtendrá una densidad de píxel umbral $d(S_o[n])$ a partir de la cual, si solamente se puede obtener una densidad de píxel mayor de las proyecciones, se tendría que interpolar los píxeles, con lo que la señal de salida quedaría degradada. Por tanto, es muy importante que la densidad de píxel del fotograma proyectado $S_1'[n], S_2'[n]$ que siempre se utilice sea aquel que permita obtener la resolución de salida deseada. Naturalmente, en otras variantes de realización se puede combinar ese fotograma proyectado que siempre se utilice con el resto de fotogramas en la zona solapada, por ejemplo $S_1'[n], S_2'[n]$ para hacer una transición progresiva entre fotogramas de diferente resolución en la zona que se solapan en la ventana de visualización 3 o en los extremos de la zona solapada, para evitar que el usuario pueda observar transiciones bruscas, aunque de esta manera se tenga que sacrificar parte de la calidad del fotograma de salida en esta zona, ya que se prevé que los usuarios toleren más una pérdida de calidad para realizar una transición progresiva que un salto de calidad que podría ser brusco.

No es conveniente para esta zona de la ventana de visualización 3 en la que exista combinaciones de fotogramas de entrada tomar siempre el fotograma de mayor resolución, puesto que, aunque se aseguraría que la calidad de los fotogramas de la señal de salida es siempre la mejor que se puede obtener, conllevaría un coste de procesado que puede no ser necesario si existe otra señal de entrada cuya densidad de píxel del fotograma proyectado sea más adecuada.

Por tanto, primeramente será necesario determinar las combinaciones de fotogramas de entrada cuyas proyecciones quedan al menos parcialmente superpuestas en la ventana de visualización 3, como se ilustra en la Fig. 3, para cada combinación de fotogramas, seleccionar el fotograma de cada combinación de fotogramas que proyectado sobre la ventana de visualización 3 tiene una densidad de píxel $d(S_{i1}'[n])$, $d(S_2'[n])$, expresada en píxeles por unidad cuadrada del espacio virtual tridimensional 1, en toda la parte superpuesta mayor o igual a la densidad de píxel umbral $d(S_0'[n])$ o, en su defecto, seleccionar el fotograma que proyectado sobre la ventana de visualización 3 tiene una mayor densidad de píxel, para formar la parte de fotograma de salida que queda superpuesta.

De los fotogramas superpuestos de la Fig. 3, las Figs. 4a y 4b muestran la comparación entre la densidad de píxel umbral $d(S_0'[n])$ y las densidades de píxeles de las proyecciones de cada uno de los fotogramas $d(S_{i1}'[n])$, $d(S_2'[n])$, determinándose que para esta ventana de visualización 3, la densidad de píxel proporcionada por la proyección del fotograma de la primera señal, $d(S_{i1}'[n])$ es menor que la densidad de píxel umbral $d(S_0'[n])$, tal y como se puede observar visualmente en la Fig. 4a y la densidad de píxel proporcionada por la proyección del fotograma de la segunda señal, $d(S_2'[n])$ es mayor que la densidad de píxel umbral $d(S_0'[n])$, tal y como se puede observar visualmente en la Fig. 4b.

Por tanto, será más conveniente utilizar la proyección del fotograma de la segunda señal de entrada $S_{i2}'[n]$ para formar la porción en la que las proyecciones quedan parcialmente superpuestas $S_{i2}'[n]$ en el fotograma de salida $S_0[n]$, ya que la densidad de píxel de esta segunda señal proyectada será la única mayor al umbral de las disponibles, mientras que para el resto de la ventana de visualización 3, al tener solamente información de la primera señal de entrada S_1 , se deberán formar los píxeles del correspondiente fotograma de la secuencia de fotogramas de salida a partir de la primera señal de entrada S_{i1} , $S_{i1}''[n]$ tal y como se indica en la Fig. 5.

La Fig. 6 muestra otra visualización del mismo espacio virtual tridimensional en el que de la proyección en la ventana de visualización 3, mostrada en la Fig. 7, se desprende que la densidad de píxel de ninguna de las señales de entrada, comparadas por separado en las Figs. 8a y 8b será mayor o igual a la densidad de píxel umbral $d(S_0'[n])$, ya que tal y como muestra la Fig. 8a la densidad de píxel proporcionada por la proyección del fotograma de la primera señal, $d(S_{i1}'[n])$, todavía menor a la densidad de píxel umbral $d(S_0'[n])$, y tal y como muestra la Fig 8b la densidad de píxel proporcionada por la proyección del fotograma de la segunda señal, $d(S_{i2}'[n])$, será también menor a la densidad de píxel umbral $d(S_0'[n])$.

Por tanto, en este caso, para la combinación de fotogramas, se seleccionará el fotograma de entre la combinación de fotogramas proyectados de mayor densidad de píxel, en este caso el de la segunda señal de entrada S_{i2}'' .

La densidad de píxel de cada fotograma proyectado sobre la ventana de visualización 3 se puede calcular tanto píxel a píxel como a partir de la proyección de agrupaciones de píxeles 4 del fotograma. Naturalmente, cuando se calcula densidad de píxel a píxel el coste computacional es mayor, pero se determina mejor el punto en el que debe utilizarse un fotograma con densidad de píxel mayor o menor. Se ha observado que utilizando agrupaciones de píxeles 5 en forma de divisiones triangulares se consigue una buena estimación de la densidad de píxel, por ejemplo como se muestra en el espacio virtual de la Fig. 10 con fotogramas de una tercera y cuarta señal de entrada S_{i3} , S_{i4} , en el que el contenido proyectado en la ventana de visualización 3 se ilustra en la Fig. 11. Naturalmente, para este espacio virtual tridimensional 1 deberá tenerse en cuenta la perspectiva que presentará la proyección de agrupaciones 5' de píxeles. Se observa que en ejemplo ilustrado en la Fig. 10 la resolución de salida R_0 mostrado en la ventana de visualización 3 en la Fig. 11 es muy baja, comportando que la densidad de píxel umbral sea igualmente baja, que será inferior a las de las proyecciones $S_{i3}'[n]$ y $S_{i4}'[n]$ por lo que al evaluar progresivamente los fotogramas, empezando primero por el fotograma de menor resolución de entrada y acabando por el de mayor resolución de entrada, se determinará que el fotograma de la combinación de fotogramas S_{i3}, S_{i4} el fotograma S_{i3} proyectado sobre la ventana de visualización 3 tiene una densidad de píxel $d(S_{i3}'[n])$ en toda la parte superpuesta mayor o igual a la densidad de píxel umbral $d(S_0'[n])$. Al ser la densidad de píxel $d(S_{i4}'[n])$ del otro fotograma todavía superior, se tomará solamente el fotograma que proyectado sobre la ventana de visualización tenga una densidad de píxel más próxima a la densidad de píxel umbral, es decir, $S_{i3}'[n]$

Por tanto, los píxeles del fotograma de la señal de salida $S_0[n]$ se formarán solamente a partir de la proyección de fotogramas $S_{i3}'[n]$ recodificada a la resolución de salida R_0 , $S_{i3}''[n]$, como se ilustra en la Fig. 12.

Ventajosamente, el procedimiento de la presente invención se puede utilizar para combinar señales de video de entrada generadas por una o varias cámaras de vídeo, en las que se represente al menos parte del mismo contenido, por ejemplo una parte de una escena, a diferentes resoluciones y que queden combinadas formando un video compuesto, también conocido como video stitching en el espacio virtual tridimensional. Es también posible que señales

de video representen el mismo contenido a diferentes resoluciones y que durante la etapa de configuración se dispongan en el espacio virtual tridimensional superpuestas.

- 5 Se contempla además que al menos una de las señales de video puede estar adaptada para formar fotogramas cuyo contenido sea una imagen de síntesis, es decir, formada mediante computador a partir de vectores, datos o de un modelo tridimensional de modo que tenga su densidad de píxel sea infinita a la práctica.

REIVINDICACIONES

- 1.- Procedimiento para componer una señal de video de salida (S_o), que comprende una secuencia de fotogramas de salida ($(S_o[n])$) con un número de píxeles prestablecidos para cada fotograma de salida asociado a una resolución de salida (R_o), a partir de una pluralidad de señales de video de entrada ($S_{i1}, S_{i2}, S_{i3} \dots S_{im}$), que comprenden respectivas secuencias de fotogramas de entrada ($S_{i1}[n], S_{i2}[n], S_{i3}[n] \dots S_{im}[n]$) con un número de píxeles para cada fotograma de entrada asociado a una resolución de entrada ($R_{i1}, R_{i2}, R_{i3} \dots R_{im}$), que comprende:
- 5 una etapa de configuración en la que se determina una superficie espacial sobre un espacio virtual tridimensional (1) formado por un conjunto de puntos con coordenadas tridimensionales (X,Y,Z) asociada a cada píxel (p) de cada fotograma de cada secuencia de fotogramas de entrada ($S_{i1}[n], S_{i2}[n], S_{i3}[n] \dots S_{im}[n]$); establecer en el espacio virtual tridimensional (1) un punto de visualización (X_0, Y_0, Z_0) y una ventana de visualización (3), que determinan un tronco (4) de visualización en el espacio virtual tridimensional (1), y formar fotogramas de salida ($S_o[n]$) a partir de las proyecciones de los fotogramas de entrada ($S'_{i1}[n], S'_{i2}[n], S'_{i3}[n] \dots S'_{im}[n]$) de las secuencias de entrada proyectadas en dicha ventana de visualización; comprendiendo el método, además:
- 10 calcular una densidad de píxel umbral ($d(S'_o[n])$) asociada a los píxeles en la ventana de visualización (3), a partir de la resolución de salida (R_o) y las dimensiones espaciales asociadas de dicha ventana de visualización (3); determinar combinaciones de fotogramas de entrada cuyas proyecciones ($S'_{i1}[n], S'_{i2}[n], S'_{i3}[n], S'_{i4}[n]$) quedan al menos parcialmente superpuestas en la ventana de visualización (3) y, para cada combinación de fotogramas, seleccionar el fotograma de cada combinación de fotogramas de entrada que proyectado sobre la ventana de visualización tiene una densidad de píxel ($d(S'_{i1}[n]), d(S'_{i2}[n]), d(S'_{i3}[n]), d(S'_{i4}[n])$) en toda la parte superpuesta mayor o igual a la densidad de píxel umbral ($d(S'_o[n])$) o, en su defecto, seleccionar el fotograma de entrada que proyectado sobre la ventana de visualización tiene la densidad de píxel más próxima a la densidad de píxel umbral, para formar la parte de fotograma de salida ($S_o[n]$) que queda superpuesta.
- 20 2.- Procedimiento según la reivindicación anterior, **caracterizado por que** si en alguna de las combinaciones de fotogramas de entrada existe una pluralidad de fotogramas de entrada que proyectados ($S'_{i1}[n], S'_{i2}[n], S'_{i3}[n], S'_{i4}[n]$) sobre la ventana de visualización (3) tienen una densidad de píxel ($d(S'_{i1}[n]), d(S'_{i2}[n]), d(S'_{i3}[n]), d(S'_{i4}[n])$) en toda la parte superpuesta mayor o igual a la densidad de píxel umbral ($d(S'_o[n])$), se seleccionan de esa combinación los fotogramas de entrada que proyectados sobre la ventana de visualización tienen una densidad de píxel ($d(S'_{i1}[n]), d(S'_{i2}[n]), d(S'_{i3}[n]), d(S'_{i4}[n])$) más próxima a la densidad de píxel umbral ($d(S'_o[n])$).
- 3.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** para formar la parte de fotograma de salida ($S_o[n]$) que queda superpuesta se realiza una mezcla entre el fotograma de entrada seleccionado de cada combinación de fotogramas y al menos otro fotograma de entrada de la misma combinación de fotogramas de entrada.
- 4.- Procedimiento según una cualquiera de las reivindicaciones 1 a 2, **caracterizado por que** para formar la parte de fotograma de salida ($S_o[n]$) que queda superpuesta se utiliza solamente el fotograma de entrada seleccionado de cada combinación de fotogramas de entrada.
- 5.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** la densidad de píxel ($d(S'_{i1}[n]), d(S'_{i2}[n]), d(S'_{i3}[n]), d(S'_{i4}[n])$) de cada fotograma de entrada proyectado sobre la ventana de visualización (3) se calcula a partir de la proyección de agrupaciones de píxeles (5) del fotograma.
- 6.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** el paso de seleccionar el fotograma de entrada de cada combinación de fotogramas de entrada se realiza evaluando progresivamente los fotogramas de entrada, empezando primero por el fotograma de entrada de menor resolución y acabando por el fotograma de entrada de mayor resolución.
- 7.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** las superficies espaciales asociadas a cada píxel de cada fotograma de cada secuencia de fotogramas de entrada en el espacio virtual tridimensional (1) conforman una superficie equivalente al interior de una esfera.
- 8.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** al menos una de las señales de video de entrada ($S_{i1}, S_{i2}, S_{i3} \dots S_{im}$) es una señal de video generada por una cámara de vídeo.
- 9.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** al menos dos de las señales de video de entrada ($S_{i1}, S_{i2}, S_{i3} \dots S_{im}$) están adaptadas para formar secuencias de fotogramas con al menos parte del mismo contenido a diferentes resoluciones.

65

10.- Procedimiento según la reivindicación anterior, **caracterizado por que** al menos dos de las señales de video están adaptadas para formar secuencias de fotogramas con el mismo contenido a diferentes resoluciones.

5 11.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** al menos dos señales de video de entrada (S_{i1} , S_{i2} , S_{i3} ... S_{im}) provienen de una misma cámara de vídeo.

10 12.- Procedimiento según una cualquiera de las reivindicaciones anteriores, **caracterizado por que** al menos una de las señales de video de entrada (S_{i1} , S_{i2} , S_{i3} ... S_{im}) está adaptada para formar fotogramas cuyo contenido es una imagen de síntesis de al menos parte del contenido del fotograma de otra de las señales de video de entrada.

13.- Programa de ordenador adaptado para realizar el procedimiento según una cualquiera de las reivindicaciones 1 a 12.

15 14.- Aparato para procesar datos que comprende medios para realizar el procedimiento según una cualquiera de las reivindicaciones 1 a 12.

15.- Instalación que comprende un aparato según la reivindicación anterior y al menos una cámara de vídeo conectada al aparato para proporcionar una señal de vídeo de entrada (S_{i1} , S_{i2} , S_{i3} ... S_{im}).

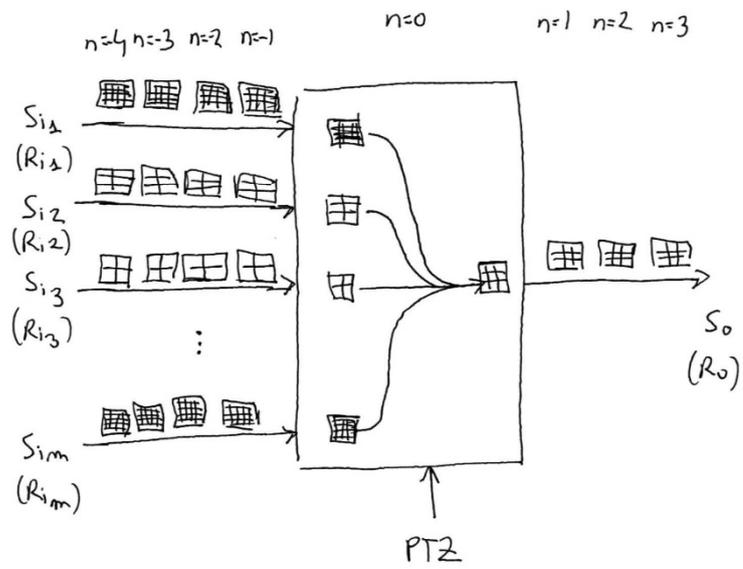


Fig 1

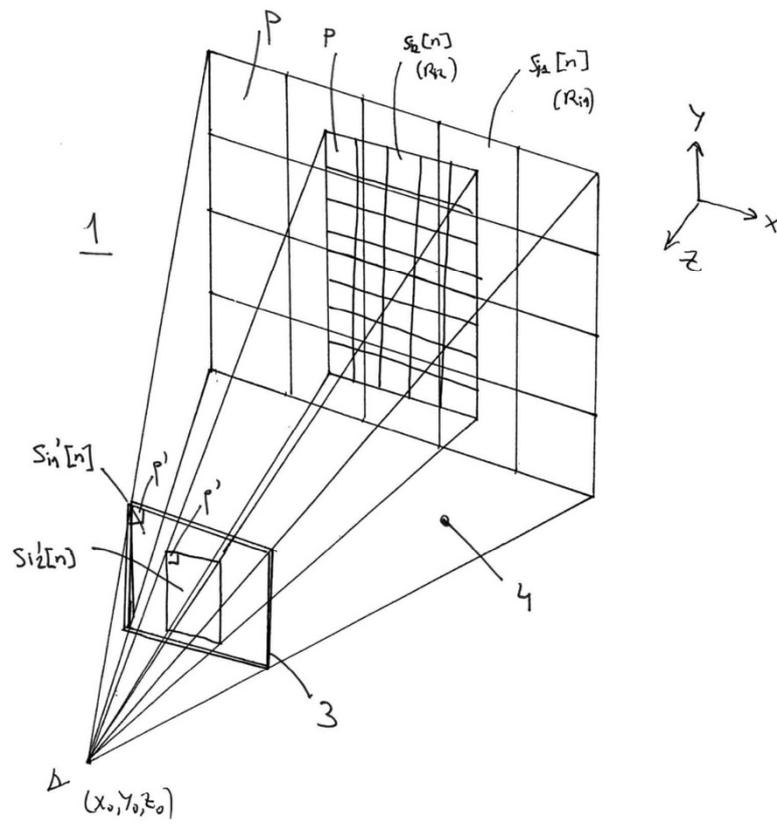


Fig 2

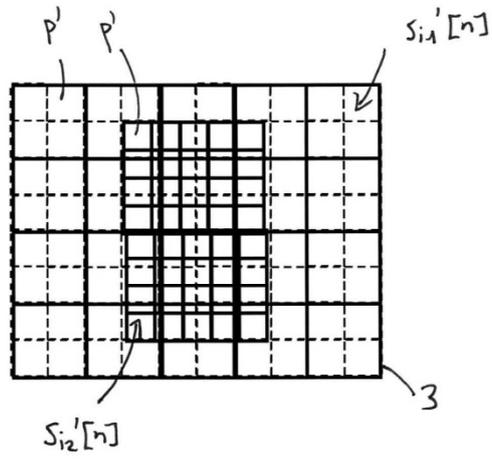


Fig 3

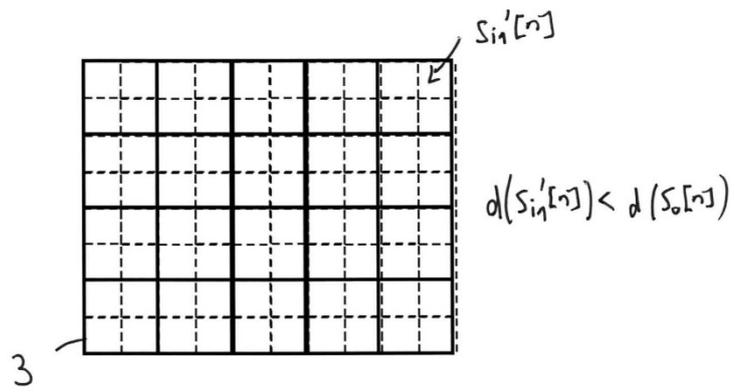


Fig 4a

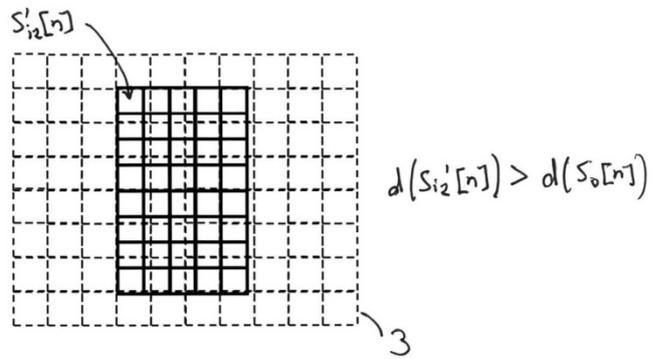


Fig 4b

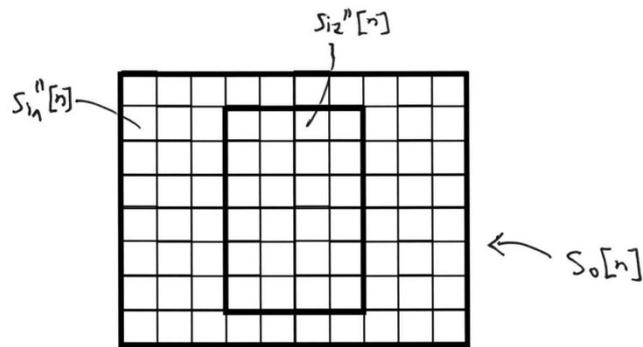


Fig 5

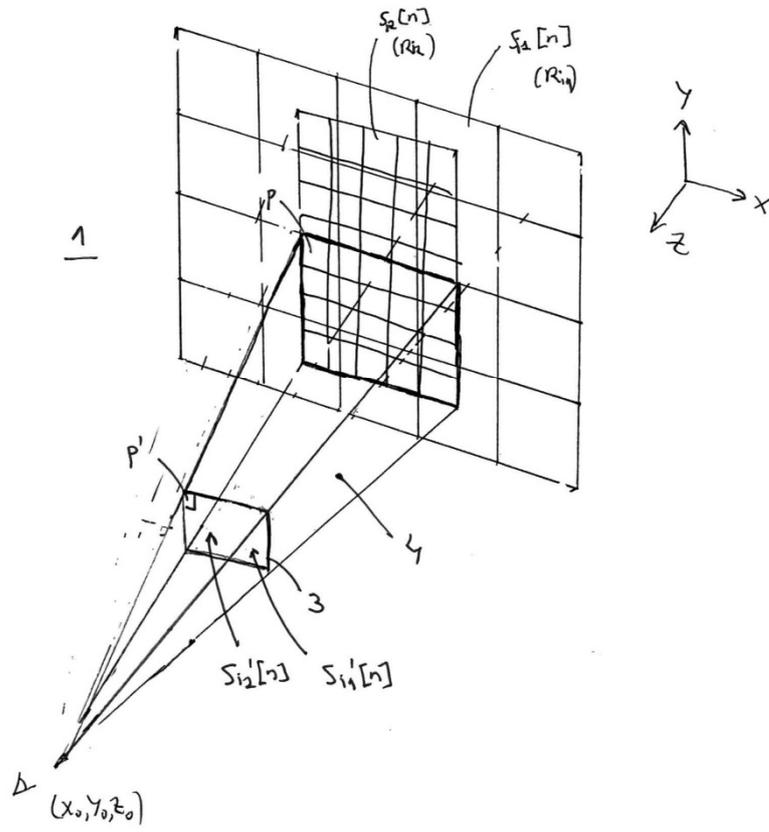


Fig 6

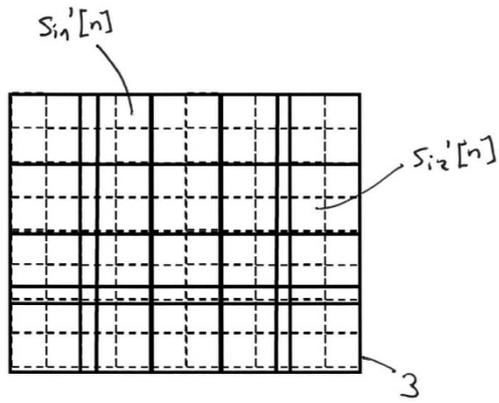


Fig 7

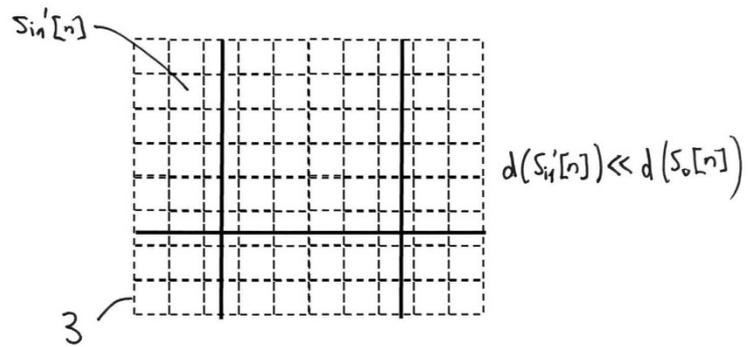


Fig 8a

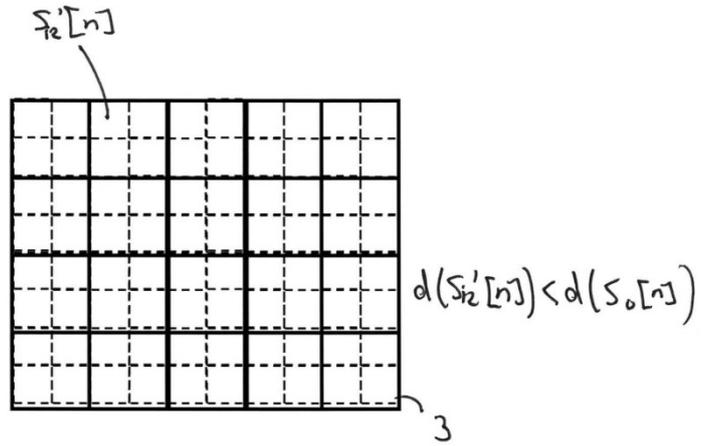


Fig 8b

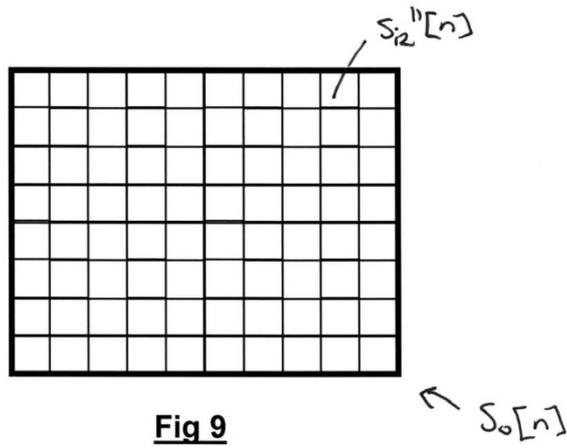


Fig 9

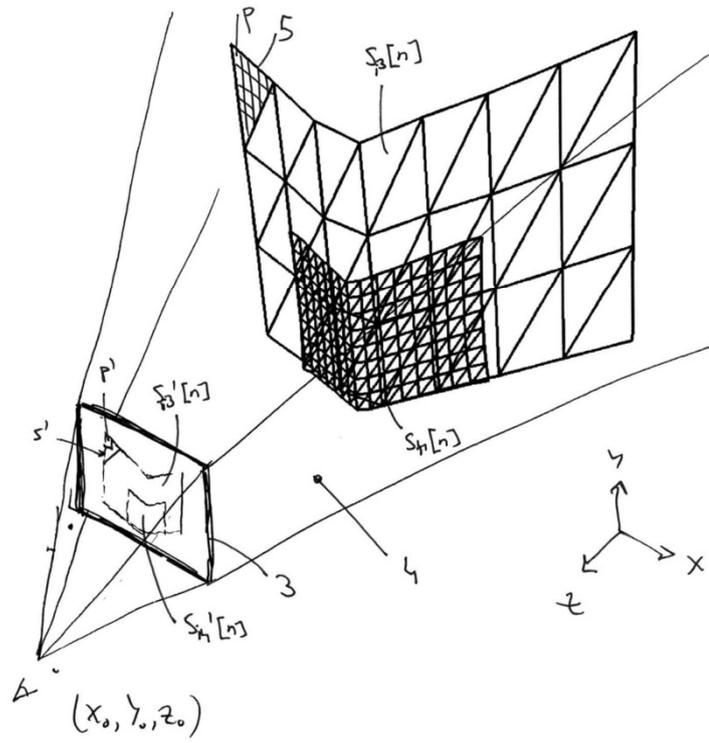


Fig 10

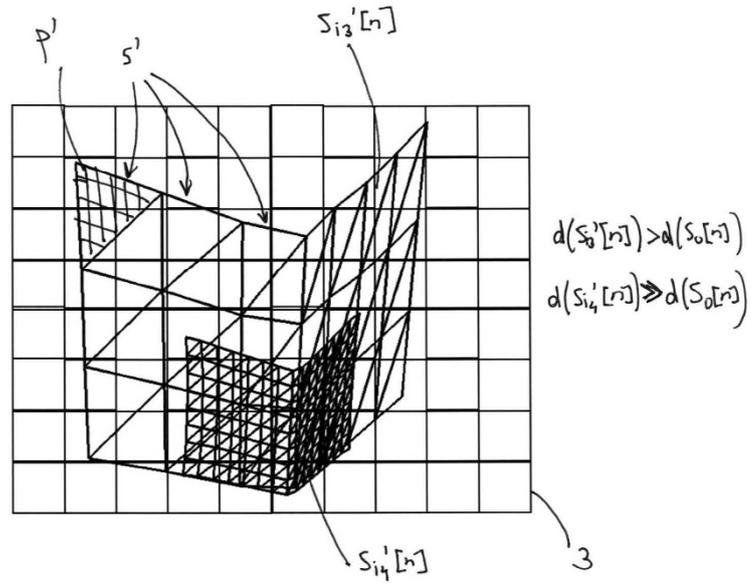


Fig 11

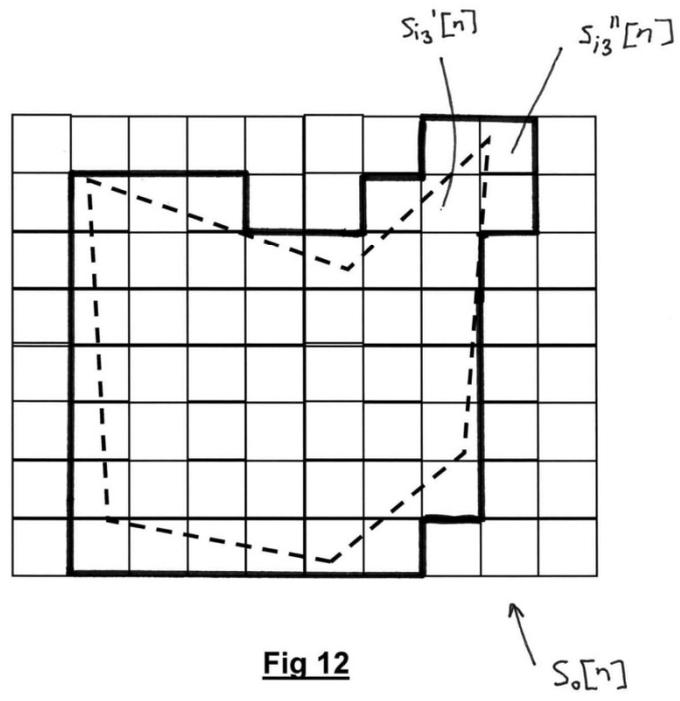


Fig 12