(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 673 545**

(51) Int. Cl.:

G06T 13/80 (2011.01)
G06T 15/20 (2011.01)
H04N 5/272 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **01.10.2012 PCT/JP2012/075416**

(87) Fecha y número de publicación internacional: **23.01.2014 WO14013629**

(96) Fecha de presentación y número de la solicitud europea: **01.10.2012 E 12881482 (9)**

(97) Fecha y número de publicación de la concesión europea: **09.05.2018 EP 2775455**

(54) Título: **Dispositivo de procesamiento de imágenes en movimiento, método de procesamiento de imágenes en movimiento, y medio de registro de información**

(30) Prioridad:

20.07.2012 JP 2012161924
07.09.2012 WO PCT/JP2012/072988
07.09.2012 WO PCT/JP2012/072989

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
22.06.2018

(73) Titular/es:

RAKUTEN, INC. (100.0%)
1-14-1, Tamagawa, Setagaya-ku
Tokyo 158-0094, JP

(72) Inventor/es:

ONAI, RIKIO;
OKABE, MAKOTO;
KURIHAR, TATSUYA;
SANJO, MASAHIRO y
HIRANO, HIROMI

(74) Agente/Representante:

CAMPELLO ESTEBARANZ, Reyes

ES 2 673 545 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

Dispositivo de procesamiento de imágenes en movimiento, método de procesamiento de imágenes en movimiento, y medio de registro de información

Campo técnico

La presente descripción se refiere a un dispositivo de procesamiento de vídeo, un método de procesamiento de vídeo y un medio de registro de información para determinar si los objetos interfieren entre sí cuando los objetos acompañados por la información de profundidad representada en múltiples vídeos se compilan en un solo vídeo.

Técnica antecedente

La tecnología para crear imágenes acompañadas de información de profundidad adquiriendo la distancia (profundidad o distancia z) entre un objeto fotografiado y una cámara se conoce desde antes. Dichas imágenes a menudo se denominan imágenes de 2,5 dimensiones (por ejemplo, véase la Bibliografía no perteneciente a patente 1).

Por otro lado, se ha propuesto la tecnología para separar regiones donde se representan fondos prácticamente estáticos y regiones donde se representan objetos en movimiento, a partir de vídeos que representan estados en los que cambian la posición, forma, color, y/o similares, de un objeto (por ejemplo, véase la Bibliografía de patente 1).

Combinando dichas tecnologías, es posible generar vídeos de 2,5 dimensiones en los que se distinguen la profundidad de fondo y la profundidad del objeto, después de separar el fondo y los objetos que se mueven delante de los mismos. En marcos que comprenden las imágenes de 2,5 dimensiones, se representa la superficie del objeto (superficie frontal del objeto) como se observa desde la cámara, y en píxeles incluidos en la región de representación donde se representa el objeto, se asocian el color de la parte de superficie del objeto correspondiente a esos píxeles, junto con la profundidad de esa parte de superficie (profundidad de superficie anterior).

Este tipo de vídeo de 2,5 dimensiones se puede obtener usando la detección mediante un sensor de profundidad simultáneamente fotografiando el mundo real mediante una cámara CCD, o fotografiando el mundo real con una cámara estéreo para estereopsis, encontrando la paralaje para cada píxel correspondiente y calculando la profundidad desde esa paralaje. Aquí, en un vídeo de 2,5 dimensiones realizado al filmar el mundo real, el estado de la superficie del objeto (superficie posterior del objeto) que no está representado en el fotograma no se fotografía, y se desconoce la información sobre la profundidad de la superficie posterior.

Además, después de crear un vídeo que expresa un estado que toma en consideración y causa el movimiento y/o similar de un objeto en un espacio virtual, añadiendo solamente la profundidad de superficie anterior en cada píxel del vídeo, es posible para obtener un vídeo de 2,5 dimensiones. Que este tipo de transformación se logre es conveniente por razones tales como la consistencia del formato, manteniéndose el vídeo de 2,5 dimensiones y siendo la edición de vídeos cada vez más fácil.

En el campo de los gráficos tridimensionales, la tecnología de determinación de conflictos se usa ampliamente para evitar que las regiones ocupadas por objetos se invadan entre sí de manera que los objetos que se mueven en el espacio virtual no interfieran, y para evitar situaciones en las que uno penetra en el otro (por ejemplo, véase la Bibliografía de patente 2). Conflicto como se denomina aquí naturalmente incluye contacto superficial entre objetos, y también incluye casos en los que las regiones ocupadas entre sí se superponen y casos en los que la distancia entre dos objetos es más cercana que un valor umbral dado aunque los dos estén realmente separados, teniendo de este modo un significado más amplio que la definición cotidiana.

La Bibliografía no perteneciente a patente 2 describe un método para detectar colisiones que pueden producirse en un único vídeo bidimensional. El método descrito comprende un algoritmo de prevención de colisiones para la simulación del comportamiento de multitudes. Dichos métodos no incorporan ni simulan información de profundidad y son diferentes de las simulaciones destinadas a la detección de colisiones en vídeos tridimensionales.

La Bibliografía no perteneciente a patente 3 describe un método para evitar colisiones en un único espacio 2,5D en el que se evita el cálculo tridimensional representando objetos usando un mapa de sombra 2,5D. El método descrito se refiere al cálculo de sombras para entornos grandes que incluyen miles de objetos dinámicos. Las colisiones se

pueden evitar deformando la trayectoria (y, por lo tanto, la sombra) de cada objeto.

Lista de citas

Bibliografía de patentes

Bibliografía de patente 1: Publicación Kokai de solicitud de patente japonesa sin examinar N.º 2009-080660.
Bibliografía de patente 2: Publicación de Solicitud de Patente de Estados Unidos N.º US2006/149546.

Bibliografía no perteneciente a patente

Bibliografía no perteneciente a patente 1: Artículo sobre "2,5D" en la versión inglesa del diccionario en línea gratuito Wikipedia, en http://en.wikipedia.org/wiki/2.5D, adquirido el 18 de julio de 2012.
Bibliografía no perteneciente a patente 2: Simulating the Collision Avoidance Behaviour of Pedestrians. Artículo de investigación de Franck Fuetey, adquirido el 9 de febrero de 2015.
Bibliografía no perteneciente a patente 3: Real-Time Shadows for animated crowds in virtual cities. Artículo de S. Kockara et al, adquirido el 10 de julio de 2015.

Resumen de la invención

Problema técnico

Aquí, cuando se compilan múltiples objetos representados en diferentes vídeos de 2,5 dimensiones en un solo vídeo, si se representan múltiples objetos superpuestos en la misma región, el objeto al que se debe dar prioridad en la representación se determina basándose en el tamaño de la profundidad.

Sin embargo, para empezar, los objetos que mutuamente invaden regiones entre sí deben ocupar espacio tridimensional, uno penetrando en el otro de manera que la interferencia que se produce entre los objetos no sea natural en un vídeo final.

En el campo de los gráficos tridimensionales, es posible determinar el resultado de la interferencia a través de la superposición de regiones tridimensionales que ocupan los objetos y la superposición de recuadros delimitadores y esferas delimitadoras que encierran el entorno de los objetos.

Sin embargo, como se ha descrito anteriormente, con vídeos de 2,5 dimensiones, aunque se entiende la forma de la superficie de un objeto que se puede ver desde la cámara (la superficie anterior del objeto), la forma de la superficie del objeto que no se puede ver desde la cámara (la superficie posterior del objeto) no se puede obtener, por lo que es imposible aplicar la tecnología de determinación de interferencias sin cambios en los gráficos tridimensionales.

Por consiguiente, es muy deseable la tecnología para determinar la interferencia entre objetos representados por vídeos de 2,5 dimensiones en condiciones en las que la información sobre la superficie posterior del objeto es inalcanzable.

En consideración de lo anterior, es un objetivo de la presente descripción proporcionar un dispositivo de procesamiento de vídeo, un método de procesamiento de vídeo y un medio de registro de información adecuado para determinar si los objetos interfieren entre sí cuando los objetos van acompañados de información de profundidad representada en múltiples vídeos se compilan en un vídeo.

Solución al problema

El dispositivo de procesamiento de vídeo de acuerdo con la presente descripción es un dispositivo de procesamiento de vídeo para determinar, cuando se forma un compuesto de un primer vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes y un segundo vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes para generar un tercer vídeo que comprende una pluralidad de fotogramas de imágenes, interferencia entre un primer objeto representado en el primer vídeo 2,5D y un segundo objeto representado en el segundo vídeo 2,5D, en el que el primer vídeo 2,5D va acompañado de una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie anterior del primer objeto representado en el primer vídeo 2,5D, y el segundo vídeo 2,5D va acompañado de una segunda distancia de profundidad, siendo la segunda distancia de profundidad una distancia a la superficie anterior del segundo el objeto representado en el segundo vídeo 2,5D,

comprendiendo el dispositivo de procesamiento de vídeo:

un adquiriente de profundidad de superficie posterior para adquirir una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie posterior del primer objeto, y una segunda distancia de profundidad, estando la segunda distancia de profundidad a una distancia con respecto a la superficie posterior del segundo objeto; y
un determinador de interferencia para determinar, si una primera área de dibujo donde el primer objeto debe representarse en un fotograma dado del tercer vídeo y una segunda área de dibujo donde el segundo objeto debe representarse en el fotograma dado de la tercera superposición de vídeo, si el primer objeto y el segundo objeto interfieren, basándose en la superposición de una primera zona de ocupación desde la primera profundidad de la superficie anterior hasta la primera profundidad de la superficie posterior del primer objeto en una área de superposición donde se superponen la primera área de dibujo y la segunda área de dibujo, y una segunda zona de ocupación desde la segunda profundidad de superficie anterior hasta la segunda profundidad de superficie posterior del segundo objeto en el área de superposición donde la primera área de dibujo y la segunda área de dibujo se superponen.

Además, el dispositivo de procesamiento de vídeo de la presente descripción puede comprenderse de tal forma que el adquiriente de profundidad de superficie posterior adquiere la primera profundidad de superficie posterior y la segunda profundidad de superficie posterior en base a los atributos que un usuario ajusta para el primer objeto y el segundo objeto.

Además, el dispositivo de procesamiento de vídeo de la presente descripción puede comprenderse de tal forma que el adquiriente de profundidad de superficie posterior adquiere la primera profundidad de superficie posterior del primer objeto en un primer fotograma del primer vídeo 2,5D estimando la longitud de la primera zona de ocupación desde un primer tamaño promedio de la primera área de dibujo en el primer vídeo 2,5D y un tamaño de área de la primera área de dibujo en el primer fotograma del primer vídeo 2,5D.

Además, el dispositivo de procesamiento de vídeo de la presente descripción puede comprenderse de tal forma que la diferencia entre la primera profundidad de superficie posterior y la primera profundidad de superficie anterior, y la diferencia entre la segunda profundidad de superficie posterior y la segunda profundidad de superficie anterior, son constantes no negativas.

Además, el dispositivo de procesamiento de vídeo de la presente descripción puede comprenderse de tal forma que el adquiriente adquiere una primera longitud de profundidad asociada por adelantado con los atributos del primer objeto, de la base de datos, y adquiere una segunda longitud de profundidad asociada por adelantado con los atributos del segundo objeto, de una base de datos, y adquiere la primera profundidad de superficie posterior de la primera profundidad de superficie anterior y la primera longitud de profundidad, y adquiere la segunda profundidad de superficie posterior de la segunda profundidad de superficie anterior y la segunda longitud de profundidad.

El método de procesamiento de vídeo de acuerdo con la presente descripción es un método de procesamiento de vídeo para determinar, cuando se forma un compuesto de un primer vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes y un segundo vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes para generar un tercer vídeo que comprende una pluralidad de fotogramas de imágenes, interferencia entre un primer objeto representado en el primer vídeo 2,5D y un segundo objeto representado en el segundo vídeo 2,5D, en el que el primer vídeo 2,5D va acompañado de una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie anterior del primer objeto representado en el primer vídeo 2,5D, y el segundo vídeo 2,5D va acompañado de una segunda distancia de profundidad, siendo la segunda distancia de profundidad una distancia a la superficie anterior del segundo el objeto representado en el segundo vídeo 2,5D, comprendiendo el método de procesamiento de vídeo:

un procedimiento de adquisición de profundidad de superficie posterior para adquirir una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie posterior del primer objeto, y una segunda distancia de profundidad, estando la segunda distancia de profundidad a una distancia con respecto a la superficie posterior del segundo objeto; y
un procedimiento de determinación de interferencia para determinar, si una primera área de dibujo donde el primer objeto debe representarse en un fotograma dado del tercer vídeo y una segunda área de dibujo donde el segundo objeto debe representarse en el fotograma dado de la tercera superposición de vídeo, si el primer objeto y el segundo objeto interfieren, basándose en la superposición de una primera zona de ocupación desde la primera profundidad de la superficie anterior hasta la primera profundidad de la

superficie posterior del primer objeto en una área de superposición donde se superponen la primera área de dibujo y la segunda área de dibujo, y una segunda zona de ocupación desde la segunda profundidad de superficie anterior hasta la segunda profundidad de superficie posterior del segundo objeto en el área de superposición donde la primera área de dibujo y la segunda área de dibujo se superponen.

El medio de registro de información legible por ordenador no transitoria de acuerdo con la presente descripción comprende un medio de registro de información legible por ordenador no transitorio en el que se registra un programa para determinar, cuando se forma un compuesto de un primer vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes y un segundo vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes para generar un tercer vídeo que comprende una pluralidad de fotogramas de imágenes, interferencia entre un primer objeto representado en el primer vídeo 2,5D y un segundo objeto representado en el segundo vídeo 2,5D, en el que el primer vídeo 2,5D va acompañado de una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie anterior del primer objeto representado en el primer vídeo 2,5D, y el segundo vídeo 2,5D va acompañado de una segunda distancia de profundidad, siendo la segunda distancia de profundidad una distancia a la superficie anterior del segundo el objeto representado en el segundo vídeo 2,5D, haciendo el programa que un ordenador funcione como:

un adquiriente de profundidad de superficie posterior para adquirir una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la parte posterior del primer objeto, y una segunda distancia de profundidad, estando la segunda distancia de profundidad a una distancia con respecto a la superficie posterior del segundo objeto; y
un determinador de interferencia para determinar, si una primera área de dibujo donde el primer objeto debe representarse en un fotograma dado del tercer vídeo y una segunda área de dibujo donde el segundo objeto debe representarse en el fotograma dado de la tercera superposición de vídeo, si el primer objeto y el segundo objeto interfieren, basándose en la superposición de una primera zona de ocupación desde la primera profundidad de la superficie anterior hasta la primera profundidad de la superficie posterior del primer objeto en una área de superposición donde se superponen la primera área de dibujo y la segunda área de dibujo, y una segunda zona de ocupación desde la segunda profundidad de superficie anterior hasta la segunda profundidad de superficie posterior del segundo objeto en el área de superposición donde la primera área de dibujo y la segunda área de dibujo se superponen.

El programa descrito anteriormente puede grabarse en un medio de registro de información no transitorio legible por ordenador tal como un disco compacto, disco flexible, disco duro, disco magneto-óptico, disco de vídeo digital, cinta magnética, memoria de semiconductor, y/o similares. El medio de registro de información se puede distribuir y vender independientemente de un ordenador.

Además, el programa descrito anteriormente puede distribuirse y venderse a través de un medio de transmisión tal como una red de comunicación informática y/o similares, independientemente de un ordenador en el que se ejecuta el programa.

Efectos ventajosos de la invención

Con la presente descripción, es posible proporcionar un dispositivo de procesamiento de vídeo, un método de procesamiento de vídeo, y un medio de registro de información adecuado para determinar si los objetos interfieren entre sí cuando los objetos acompañados por la información de profundidad representada en múltiples vídeos se compilan en un vídeo.

Breve descripción de los dibujos

La Figura 1 es un dibujo explicativo que muestra la relación entre un objeto representado en un vídeo de 2,5 dimensiones y los elementos del mismo;
la Figura 2A es una vista en sección transversal de un objeto;
la Figura 2B es una vista en sección transversal que muestra la superficie anterior del objeto;
la Figura 2C es una vista en sección transversal de un objeto cuya forma de superficie posterior se estima haciendo que el espesor sea constante;
la Figura 2D es una vista en sección transversal de un objeto cuya forma de superficie posterior se estima al encontrar una profundidad de superficie anterior representativa del objeto y ajustando el valor encontrado añadiendo espesor a éste como la profundidad de superficie posterior;
la Figura 2E es una vista en sección transversal de un objeto cuya forma de superficie posterior se estima

que es una superficie esférica;
la Figura 3 es un dibujo explicativo que muestra una composición resumida de los elementos que acompañan a la determinación de conflictos, en un dispositivo de procesamiento de vídeo;
la Figura 4 es un diagrama de flujo que muestra el flujo del proceso de determinación de conflictos;
la Figura 5A es un dibujo explicativo que muestra el estado de la pantalla antes del inicio del arrastre;
la Figura 5B es un dibujo explicativo que muestra el estado de la pantalla en el momento en que se inicia el arrastre;
la Figura 5C es un dibujo explicativo que muestra el estado de la pantalla en un punto en el tiempo durante el arrastre;
la Figura 5D es un dibujo explicativo que muestra el estado de la pantalla en un punto en el tiempo cuando la imagen se ha soltado;
la Figura 6 es un dibujo explicativo que muestra la composición resumida de la parte de acuerdo con un proceso de arrastrar y soltar del dispositivo de procesamiento de vídeo;
la Figura 7 es un diagrama de flujo que muestra el flujo de control del proceso de arrastrar y soltar;
la Figura 8 es un diagrama de flujo que muestra el flujo de control de un proceso de ajuste automático;
la Figura 9A es un dibujo explicativo que muestra el estado de la pantalla antes del inicio del arrastre;
la Figura 9B es un dibujo explicativo que muestra el estado de la pantalla en el momento en que se inicia el arrastre;
la Figura 9C es un dibujo explicativo que muestra el estado de la pantalla en un punto en el tiempo durante el arrastre;
la Figura 9D es un dibujo explicativo que muestra el estado de la pantalla en un punto en el tiempo cuando la imagen se ha soltado;
la Figura 9E es un dibujo explicativo que muestra el estado en el que se realizó una corrección automática a la posición de caída después de la caída;
la Figura 10 es un diagrama de flujo que muestra el flujo de control de un proceso de visualización de resaltado en la región desprendible;
la Figura 11A es un dibujo explicativo que muestra el estado de una trayectoria pos[X,i](t) y una trayectoria pos[Y,j](t);
la Figura 11B es un dibujo explicativo que muestra el estado de la trayectoria pos[X,i](t) y la trayectoria pos[W,j](t);
la Figura 12A es un dibujo explicativo que muestra el estado de una trayectoria que se deforma gradualmente a través de la repetición de un proceso;
la Figura 12B es un dibujo explicativo que muestra el estado de una trayectoria que se deforma gradualmente a través de la repetición de un proceso;
la Figura 12C es un dibujo explicativo que muestra el estado de una trayectoria que se deforma gradualmente a través de la repetición de un proceso;
la Figura 12D es un dibujo explicativo que muestra el estado de una trayectoria que se deforma gradualmente a través de la repetición de un proceso;
la Figura 13 es un dibujo explicativo que muestra el estado de ajuste de un vector normal primario en una parte específica de una trayectoria para evitar un conflicto.

Descripción de las realizaciones

A continuación, se describen las realizaciones preferidas de la presente descripción. Las realizaciones preferidas son para propósitos explicativos y no limitan el alcance de la presente descripción. Por consiguiente, un experto en la técnica podría aplicar una realización en la que diversos elementos o todos los elementos en el presente documento están sustituidos por equivalentes de los mismos, pero dichas realizaciones también se incluyen dentro del alcance de la presente descripción.

El dispositivo de procesamiento de vídeo de acuerdo con las realizaciones preferidas descritas a continuación se realiza mediante un ordenador que ejecuta un programa. El programa se lee en un dispositivo de memoria temporal tal como una RAM (memoria de acceso aleatorio) y/o similares, desde un dispositivo de memoria no transitorio tal como una ROM (memoria de solo lectura), un disco duro y/o similares. Una CPU (unidad central de procesamiento) interpreta sucesivamente los programas leídos en la RAM, lee y escribe datos entre diversos dispositivos de memoria y lleva a cabo procesos de cálculo en los datos. Además, la CPU controla la interacción con los dispositivos de entrada y salida, tales como el teclado, el ratón, el panel táctil, la pantalla y la interfaz de red, y/o similares, bajo el control del programa.

Sin embargo, el dispositivo de procesamiento de vídeo de acuerdo con la presente descripción no está limitado a ser

realizado por un ordenador que ejecuta un programa. Por ejemplo, es posible realizar las mismas funciones utilizando un circuito electrónico especial para el procesamiento de imágenes, una FPGA (matriz de puertas programables de campo), DSP (procesador de señales digitales) y/o similares.

En el campo de los gráficos tridimensionales, la interferencia entre objetos se denomina conflicto. En el lenguaje cotidiano, un conflicto significa objetos que colisionan violentamente, haciendo contacto las superficies mutuas y en algunos casos deformándose las formas de los objetos, pero cuando se habla de interferencia de objetos o conflicto de objetos, el significado incluirá cuando las superficies hacen contacto, casos donde las regiones dentro del espacio tridimensional ocupado por un objeto son invadidas y penetradas, y casos en los que incluso a través de los objetos que están separados, la distancia es suficientemente cercana, y/o similares. En las realizaciones a continuación, la frase "conflicto" se usa en lugar del término "interferencia" para facilitar la comprensión.

(Relación entre un vídeo de 2,5 dimensiones y objetos)

A continuación, la notación para expresar un objeto representado en un vídeo de 2,5 dimensiones se organiza para facilitar la comprensión. La Figura 1 es un dibujo explicativo que muestra la relación entre un objeto representado en un vídeo de 2,5 dimensiones y los elementos del mismo. La explicación a continuación hace referencia a este dibujo.

En la explicación a continuación, la frase "vídeo de 2,5 dimensiones" se abrevia apropiadamente como "vídeo" para facilitar la comprensión.

Una trama reproducida en un tiempo transcurrido t desde el punto de inicio de reproducción (clásicamente, esto se expresa por número de fotogramas) entre los fotogramas incluidos en un vídeo dado X se indica como X(t).

En el sistema de coordenadas de píxeles fijado para cada fotograma, en el fotograma X(t), el píxel situado en el valor de coordenadas x en la dirección horizontal y el valor de coordenadas y en la dirección vertical se indica como X(t)[x,y].

Que el valor de coordenadas x en la dirección horizontal y el valor de coordenadas y en la dirección vertical sean valores eficaces en cada fotograma del vídeo X se indica como $[x,y] \in X$. Es decir, [x,y] significa la posición de un píxel.

La Figura 1 ilustra un estado en el que la condición de un objeto posicionado en el espacio virtual se representa en un fotograma del vídeo se ve directamente desde el lateral, usando tecnología de gráficos tridimensionales, pero es posible concebir condiciones similares existentes incluso cuando el mundo real es fotografiado con una cámara.

Como se muestra en el dibujo, cuando desde un punto de filmación 11 en el espacio una media línea 13 hacia el píxel X(t)[x,y] en el fotograma X(t) posicionado en una superficie de proyección 12 se encuentra un punto de conflicto 15 de la superficie de un objeto 14, el estado del punto de conflicto encontrado 15 se representa en el píxel (X)t[x,y]. Esta es una proyección de perspectiva típica.

Como se ha descrito anteriormente, este dibujo es el estado de filmación visto directamente desde el lateral. Por consiguiente, en este dibujo, la superficie de proyección 12 se expresa por una línea recta.

El punto de filmación 11 corresponde a la posición de la cámara o la posición de observación, y la distancia entre el punto de filmación 11 y la superficie de proyección 12 se determina por la longitud focal de la lente y las unidades de longitud de píxel. Además, la dirección de filmación 16 es la dirección de una línea perpendicular desde el punto de filmación 11 hacia la superficie de proyección 12.

A continuación, el color del píxel X(t)[x,y] y la profundidad de superficie anterior se indican como X(t)[x,y].color y X(t)[x,y].fore.

X(t)[x,y].color se determina basándose en el color y el estado de la escritura en el punto de conflicto 15, y la distancia entre el punto de filmación 11 y el punto de conflicto 15, y/o similares. X(t)[x,y].color expresa un valor de coordenadas en cada espacio de color, tal como RGB, CYMK, HSV y/o similares, y se formula la conversión del valor de coordenadas entre espacios de color. Además, también sería correcto añadir un valor α que indique transparencia a X(t)[x,y].color.

X(t)[x,y].fore expresa la distancia de profundidad entre la parte correspondiente al punto de conflicto 15 del objeto 14

representado en el píxel X(t)[x,y], y el punto de filmación 11. Como la distancia de profundidad, es típico utilizar una longitud 17 de un componente en la dirección de filmación 16 de un vector desde el punto de filmación 11 al punto de conflicto 15 (correspondiente a la denominada "distancia Z"), pero como una aproximación del mismo, sería correcto utilizar la distancia desde el punto de filmación 11 al punto de conflicto 15.

La distancia de profundidad se puede expresar en diversos sistemas de unidades. Cuando un píxel comprende un cuadrado, si la longitud del lado del píxel se toma como la unidad, es posible normalizar la profundidad en múltiples vídeos. A continuación, la explicación supone que la profundidad se ha normalizado.

En general, en el vídeo X se representa un fondo además de los objetos en movimiento. Se puede pensar que el fondo es un objeto sustancialmente inmóvil.

Si se analiza cada fotograma del vídeo usando la tecnología descrita en la Bibliografía de patente 1 y/o similares, es posible identificar si se representa un fondo o se representa un objeto en movimiento en el píxel X(t)[x,y].

A continuación, se considerará un caso en el que se han asignado números de identificación a objetos identificados para facilitar la comprensión. El número de identificación del objeto que expresa el fondo es 0, y los números de identificación de otros objetos se asignan como 1, 2, 3....

Además, el número de identificación de un objeto del que se representa una parte en el píxel X(t)[x,y] se indica como X(t)[x,y].id, y el valor máximo de los números de identificación de objetos que aparecen en el vídeo X se indica como X.maxid.

Si los números de identificación de objetos que aparecen en el vídeo X son 0, 1, 2, 3, ..., N, después X.maxid = N.

En el vídeo X, si solo se extraen los píxeles X(t)[x,y] que satisfacen la condición X(t)[x,y].id ≠ 0, se puede obtener información solamente sobre los objetos (información que excluye el fondo). Además, para obtener un vídeo de solo el i-ésimo objeto, sería correcto extraer solo los píxeles X(t)[x,y] que satisfagan la condición X(t)[x,y].id = i. Estos corresponden a vídeos para la composición de clave cromática usando fondo azul en tecnología de vídeo convencional.

Como resultado de llevar a cabo la extracción descrita anteriormente, en píxeles que han dejado de ser objetivos de procesamiento, se supondrá que se ha dado un valor negativo como número de identificación para facilitar la comprensión. Es decir, si X(t)[x,y].id <0, el píxel X(t)[x,y] se interpretará como un píxel transparente. Aquí, "transparente" significa que corresponde al fondo azul en la composición de clave cromática.

Basándose en esta interpretación, es posible pensar en un valor negativo que se ajusta como X(t)[x,y].id cuando [x,y] $\in$ X no está establecido. A continuación, la explicación continuará para el caso en que los valores negativos se ajustan como .id para píxeles que no son objetivos de procesamiento y "píxeles" en posiciones fuera del fotograma del vídeo, para facilitar la comprensión.

(Diversos tipos de conversiones para vídeos de 2,5 dimensiones)

A continuación, se describen métodos de conversión para crear un vídeo diferente a partir de un vídeo dado. Entre este tipo de método de conversión se encuentran los siguientes:

(1) Traslación arriba, abajo, izquierda o derecha a lo largo de un fotograma del vídeo;
(2) Escalamiento del vídeo;
(3) Traslación en la dirección de profundidad del vídeo;
(4) Rotación del vídeo en torno al eje horizontal, el eje vertical o el eje de profundidad;
(5) Desplazamiento del vídeo en la dirección de tiempo;
(6) Selección de solo objetos específicos del vídeo;
(7) Formación de un compuesto superponiendo un vídeo diferente en un vídeo dado.

Estos se explican en orden a continuación.

En primer lugar, se considera la traslación hacia arriba, abajo, izquierda o derecha a lo largo de un fotograma del vídeo. Si se obtiene un vídeo Y al trasladar el vídeo X por p en la dirección horizontal y q en la dirección vertical dentro del fotograma, entonces:

$$Y(t)[x+p,y+q] = X(t)[x,y],$$

y

$$Y(t)[x,y] = X(t)[x-p, y-q].$$

Aquí, para las igualdades descritas anteriormente, los valores asignados a cada píxel tal como .color y .fore y/o similares, todos tienen un significado equivalente.

A continuación, el vídeo obtenido al trasladar el vídeo X por p en la dirección horizontal y por q en la dirección vertical dentro del fotograma se indica como move(p,q,X).

A continuación, se considera el escalamiento del vídeo. En una proyección transparente, la representación como c veces más grande significa que la profundidad se ha convertido en 1/c veces y los puntos de fondo se han vuelto más cercanos. Por consiguiente, si el vídeo Y es la ampliación del vídeo x c veces dentro del fotograma, para otro que no sea el .fore asignado a cada píxel,

$$Y(t)[c\times x, c\times y] = X(t)[x,y],$$

es decir,

$$Y(t)[x,y] = X(t)[x/c, y/c].$$

Además, por lo tanto, para .fore asignado a cada píxel,

$$Y(t)[c \times x, c\times y].fore = X(t)[x,y].fore/c,$$

es decir,

$$Y(t)[x,y].fore = X(t)[x/c, y/c].fore/c.$$

A continuación, el vídeo obtenido al ampliar el vídeo X c veces dentro del fotograma se indica como scale(c,X).

A continuación, se considera la traslación del vídeo en la dirección de profundidad. Cuando el vídeo Y se obtiene por objetos tomados en el vídeo X que se hace más distante por r en la dirección de profundidad, la distancia de profundidad de X(t)[x,y] cambia de X(t)[x,y].fore a X(t)[x,y].fore + r. Es decir, la distancia de profundidad se convierte en (X(t)[x,y].fore + r)/X(t)[x,y].fore = k veces.

En la proyección transparente, si la distancia a un punto de filmación se convierte en k veces, el tamaño representado en la superficie de proyección se amplía o se reduce a 1/k veces. Por consiguiente, para otro que no sea el .fore asignado a cada píxel,

$$Y(t)[x/k, y/k] = X(t)[x,y],$$

y para .fore,

$$Y(t)[x/k, y/k].fore = X(t)[x,y].fore + r.$$

A continuación, el vídeo obtenido trasladando el vídeo X por r en la dirección de profundidad se indica como push(r,X).

Además, se considera la rotación del vídeo en torno al eje horizontal, el eje vertical o el eje de profundidad. Si el vídeo Y se obtiene rotando el vídeo X en θ en torno al eje horizontal, los elementos diferentes de .fore se convierten en:

$$Y(t)[x,y] = X(t)[x, X(t)[x,y].fore \times \cos\theta - y \times \mathrm{sen}\,\theta],$$

y para .fore,

$$Y(t)[x,y].fore = X(t)[x,y].fore \times \mathrm{sen}\,\theta + y \times \cos\theta$$

se establece.

De forma similar, si el vídeo Y se obtiene rotando el vídeo X en φ en torno al eje vertical, los elementos distintos de .fore se convierten en:

$$Y(t)[x,y] = X(t)[X(t)[x,y].fore \times \cos\phi - x \times \text{sen}\,\phi, y],$$

y para .fore,

$$Y(t)[x,y].fore = X(t)[x,y].fore \times \text{sen}\,\phi + x \times \cos\phi$$

se establece.

Además, cuando el vídeo Y se obtiene rotando el vídeo X en ψ en torno al eje de profundidad,

$$Y(t)[x,y] = X(t)[x \times \cos\psi - y \times \text{sen}\,\psi, x \times \text{sen}\,\psi + y \times \cos\psi]$$

se establece.

A continuación, los vídeos obtenidos por las respectivas rotaciones se indican como rothor(θ,X), rotver(φ,X) y rotdep(ψ, X).

Cuando se realizan estas transformaciones, los valores de coordenadas de los píxeles no son valores enteros y, a veces, se produce un agujero en los valores de coordenadas. En dichos casos, los valores de .color y .fore y/o similares asignados a cada píxel se interpolan y se realiza anti-solapamiento, y para .id, es posible adoptar un método tal como utilizar un valor en el punto de cuadrícula más cercano a esa coordenada, o encontrar en qué posición el límite donde .id cambia se coloca a través de interpolación y determinar .id basándose en ese límite.

Los píxeles que están fuera del alcance del fotograma del vídeo debido a una traslación y/o similar pueden considerarse píxeles transparentes, por lo que sería correcto establecer .id en un valor negativo como anteriormente.

A continuación, se consideran las traslaciones del vídeo en la dirección del tiempo. Si el vídeo Y se encuentra desplazando el vídeo X por d en el tiempo,

$$Y(t)[x,y] = X(t-d)[x,y]$$

se establece.

A continuación, el vídeo en el que el tiempo del vídeo X se desplaza por d se indica como shift(d,X).

Además, se considera un proceso para seleccionar únicamente un objeto específico de un vídeo. Si un vídeo y representa solo un objeto con el i-ésimo número de identificación en el vídeo X, para elementos distintos de .id,

$$Y(t)[x,y] = X(t)[x,y], \text{ if } X(t)[x,y].id = i$$

se establece.
Además, para .id, si

$$Y(t)[x,y].id = 1, \text{ if } X(t)[x,y].id = i \;\; \text{e} \;\; i > 0;$$

$$Y(t)[x,y].id = 0, \text{ if } X(t)[x,y].id = i \;\; \text{e} \;\; i = 0;$$

$$Y(t)[x,y].id = -1,$$

de otro modo,
entonces es posible minimizar en la medida de lo posible el valor máximo Y.maxid de los números de identificación en el vídeo Y.

A continuación, un vídeo que representa solo el objeto que tiene el número de identificación i representado en el vídeo X se indica como select(i,X).

Finalmente, se considera el proceso de superponer un vídeo diferente en un vídeo dado. Si no hay conflicto entre los objetos representados en los vídeos X e Y, superponiendo un objeto distinto del fondo del vídeo Y en el vídeo X (sería correcto incluir el fondo, y también sería correcto incluir solo los objetos distintos del fondo), se obtiene un vídeo Z en el que los objetos representados en el vídeo X y los objetos representados en el vídeo Y se compilan en uno.

En el vídeo Z, para otro distinto de .id,

$$Z(t)[x,y] = X(t)[x,y], \text{ if } Y(t)[x,y].id \leq 0;$$

$$Z(t)[x,y] = X(t)[x,y], \text{ if } X(t)[x,y].fore < Y(t)[x,y].fore;$$

$$Z(t)[x,y] = Y(t)[x,y],$$

de otro modo
se establece.
Además, para .id, si

$$Z(t)[x,y].id = X(t)[x,y].id, \text{ if } Y(t)[x,y].id \leq 0;$$

$$Z(t)[x,y].id = X(t)[x,y].id, \text{ if } X(t)[x,y].fore < Y(t)[x,y].fore;$$

$$Z(t)[x,y].id = Y(t)[x,y].id + X.maxid,$$

de otro modo
entonces es posible evitar la superposición de números de identificación de objetos.

Además, de los números enteros 1, 2, ..., Z.maxid, si hay uno que no aparece como un valor de Z(t)[x,y].id, es posible minimizar Z.maxid desplazando el .id de los objetos posteriores.

A continuación, el vídeo obtenido al superponer objetos distintos del fondo del vídeo Y en el vídeo X se indica como superimpose(Y,X).

Cuando los procesos descritos anteriormente move, scale, push, rothor, rotver, rotdep, select, superimpose, y/o similares, se realizan por ordenador, es posible usar una función de calculadora poseída por una CPU (unidad central de procesamiento), y el cálculo de alta velocidad también es posible usando un procesador para el procesamiento de gráficos, y/o similares.

En scale y push, el escalamiento y la traslación en la dirección de profundidad se consideran centradas en el origen, y en rothor, rotver y rotdep, la suposición es que el eje de rotación pasa a través del origen.

Por consiguiente, cuando el deseo es lograr el escalamiento o la traslación en la dirección de profundidad centrada en una posición arbitraria [p,q] dentro del vídeo X, sería correcto trasladar primero [p,q] arriba, abajo, izquierda o derecha al origen, luego realizar el escalamiento o la traslación en la dirección de profundidad y luego realizar la traslación inversa arriba, abajo, izquierda o derecha. Es decir, sería correcto usar:

move(p,q, scale(c, move(-p, -q, X)));
move(p,q, push(r, move(-p, -q, X))).

Cuando la posición del eje de rotación se ha de desplazar hacia arriba, abajo, izquierda o derecha o en la dirección de profundidad, el eje de rotación se traslada primero en la dirección arriba, abajo, izquierda o derecha o en la dirección de profundidad para pasar por el origen, la rotación se lleva a cabo, y después de esto se lleva a cabo una traslación inversa en la dirección arriba, abajo, izquierda o derecha o en la dirección de profundidad.

Además de esto, en una transformación tal como una escala o empuje y/o similar cuando se traduce la profundidad, sería correcto corregir el brillo o la intensidad y/o similares de .color de acuerdo con el grado de los mismos.

La normalización del vídeo también se puede llevar a cabo basándose en los ajustes y/o similares en el momento de la filmación, pero también es posible establecer esto de acuerdo con los deseos del usuario o se puede lograr automáticamente.

En el caso de las configuraciones por parte del usuario, primero la altura o el ancho del vídeo simplemente se escala en base a las instrucciones del usuario y/o similares, y al ajustar la resolución, el tamaño de los objetos representados en la imagen se convierte en el deseado. Con "scale", la distancia de profundidad de los píxeles correspondientes se multiplica por un coeficiente correspondiente a la relación de ampliación, pero cuando se normaliza, la distancia de profundidad de los píxeles correspondientes se utiliza sin cambios.

A continuación, el usuario selecciona un coeficiente y la distancia de profundidad de cada píxel se multiplica por este

coeficiente.

Cuando se intenta generar un vídeo Z en el que el objeto i y el objeto j se representan juntos bajo la situación en la que se representa un objeto i en el vídeo X y se representa un objeto j en el vídeo Y, a menudo es preferible que el tamaño del objeto i y el tamaño del objeto j sean sustancialmente equivalentes, y que la profundidad del objeto i y la profundidad del objeto j sean sustancialmente equivalentes. Cuando se establecen dichos requisitos previos, es posible la normalización automática.

Es decir, el vídeo Y simplemente se escala para que el área promedio y la altura promedio y/o similares de la región en la que se representa el objeto j en el vídeo Y coincida con los valores correspondientes del objeto i en el vídeo X, de manera que se obtenga el vídeo Z. Es decir, para todos los valores, incluido .fore,

$$Z(t)[x,y] = Y(t)[x/c, y/c].$$

Además, la profundidad en el vídeo Z se multiplica por un coeficiente para que la profundidad promedio del objeto j en el vídeo Z y la profundidad promedio del objeto i en el vídeo X coincidan, obteniendo de ese modo un vídeo W. Es decir, para otro distinto de .fore,

$$W(t)[x,y] = Z(t)[x,y],$$

y para .fore,

$$W(t)[x,y].fore = k \times Z(t)[x,y].fore.$$

Si c y k se ajustan automáticamente de manera que se satisfagan las condiciones anteriores, el vídeo W se obtiene normalizando el vídeo Y de modo que el objeto j corresponda con el objeto i. Naturalmente, como se ha descrito anteriormente, sería correcto lograr la normalización configurando c y k de acuerdo con los deseos del usuario.

Anteriormente, se describieron la notación de diversas cantidades características de los vídeos y diversos tipos de transformaciones, pero estas anotaciones y transformaciones están destinadas a facilitar la comprensión. En particular, es posible expresar .color, .fore, .id y/o similares a través de otros formatos. Incluso cuando se utilizan métodos o transformaciones equivalentes a esto en diversos formatos utilizados en el procesamiento de vídeo, naturalmente esto se incluye dentro del alcance técnico de la presente descripción.

Primera realización preferida

Esta realización preferida se usa para determinar si un objeto i que tiene un número de identificación i en el vídeo X y un objeto j que tiene un número de identificación j en el vídeo Y entran en conflicto cuando el vídeo Y se superpone al vídeo X.

Se consideran las siguientes condiciones como un caso en el que un conflicto entre el objeto i y el objeto j se puede determinar con relativa facilidad.

Primero, hay un caso en el que hay un instante durante el cual la superficie anterior del objeto i (la superficie representada en el vídeo X) y la superficie anterior del objeto j (la superficie representada en el vídeo Y) ocupan la misma posición tridimensional. Es decir, para cualquier t, x e y, si

$$X(t)[x,y].id = i;$$

y

$$Y(t)[x,y].id = j;$$

la posición [x,y] está contenida duplicada en el área donde debería representarse el objeto i y el área donde debería representarse el objeto j.

En esta posición [x,y], si

$$X(t)[x,y].fore = Y(t)[x,y].fore$$

se establece, se determina que el objeto i y el objeto j entran en conflicto.

Con este método, la suposición es que los objetos i y j son objetos en forma de lámina sin nada en la superficie posterior, que no poseen profundidad y que comprenden solo la forma de la superficie anterior.

Para que las superficies anteriores de los objetos i y j tengan la misma profundidad en una posición proyectada sobre un píxel dado X(t)[x,y] significa que las superficies anteriores de los objetos i y j se superponen en esa posición. Por lo tanto, en este caso, es posible determinar que los dos entran en conflicto.

También es posible extender este estándar de determinación en la dirección del tiempo. A continuación, la suposición es para un caso en el que el tiempo transcurrido t se expresa mediante el número de fotogramas, es decir, que el fotograma en el tiempo t + 1 sigue al fotograma en el tiempo t.

En cualquier t, x e y, si

$$X(t)[x,y].id = X(t+1)[x,y].id = i;$$

$$Y(t)[x,y].id = Y(t+1)[x,y].id = j;$$

$$X(t)[x,y].fore < Y(t)[x,y].fore;$$

y

$$X(t+1)[x,y].fore > Y(t+1)[x,y].fore$$

se establecen, se puede pensar que desde el tiempo t hasta el tiempo t+1, el objeto j ha venido desde detrás del objeto i al frente. Además, si

$$X(t)[x,y].id = X(t+1)[x,y].id = i;$$

$$Y(t)[x,y].id = Y(t+1)[x,y].id = j;$$

$$X(t)[x,y].fore > Y(t)[x,y].fore;$$

y

$$X(t+1)[x,y].fore < Y(t+1)[x,y].fore$$

se establecen, se puede pensar que desde el tiempo t hasta el tiempo t+1, el objeto i ha venido desde detrás del objeto j al frente.

Por lo tanto, cuando se produce este tipo de "deslizamiento", se determina que se ha producido un conflicto de objetos.

Con una determinación de conflicto para los objetos i y j en estas condiciones, la suposición es que los objetos i y j están compuestos por las formas de solo las superficies anteriores de los mismos. Cuando los objetos son objetos delgados en forma de placa, se puede pensar que el espesor de los objetos es cero y que coinciden la superficie anterior del objeto y la superficie posterior del objeto. Por consiguiente, mediante la determinación de conflicto lograda a través de este tipo de método, es posible una determinación fácil y rápida.

La Figura 2A es una vista en sección transversal del objeto 14 mostrado en la Figura 1, y la Figura 2B es una vista en sección transversal que muestra la superficie anterior del objeto 14. La explicación a continuación hace referencia a estos dibujos.

Como se muestra en la Figura 2A, el objeto 14 tiene profundidad y la sección transversal de la misma tiene una extensión.

Sin embargo, cuando el objeto 14 se expresa a través de la información contenida en un vídeo de 2,5 dimensiones, como se muestra en la Figura 2B, la forma del mismo es un objeto plano que comprende solo la superficie anterior del objeto 14. Como este dibujo es una vista en sección transversal, un objeto plano se expresa mediante una línea curva.

Por consiguiente, para lograr una determinación de conflicto entre objetos con mayor precisión, es necesario establecer o estimar la forma de la superficie posterior del objeto 14 a partir de la información sobre el objeto plano

como se muestra en la Figura 2B.

Aquí, la profundidad con respecto a la superficie anterior del objeto 14 se obtiene a partir de información incidente al vídeo de 2,5 dimensiones, por lo que para conocer la profundidad del objeto 14, sería correcto obtener la profundidad con respecto a la superficie posterior del objeto 14.

A continuación, la profundidad con respecto a la superficie posterior de un objeto representado en el píxel X(t)[x,y] en el tiempo t en el vídeo X se indica como X(t)[x,y].back, y se describe un método para obtener esta profundidad de superficie posterior.

En un primer método, el usuario establece por adelantado la distancia entre la superficie anterior y la superficie posterior para cada objeto.

Es decir, si la distancia (espesor) entre la superficie anterior y la superficie posterior se establece como una constante thick(i) para un objeto i en el vídeo X, entonces el resultado es:

$$X(t)[x,y].back = X(t)[x,y].fore + thick(X(t)[x,y].id).$$

La Figura 2C es una vista en sección transversal de un objeto para el cual se estima la forma de la superficie posterior tomando el espesor como una constante. Como se muestra en el dibujo, la superficie posterior del objeto 14 es la forma de la superficie anterior del objeto 14 que ha sufrido una traslación en la dirección de filmación 16.

Además, existe un método para determinar una profundidad representativa de superficie anterior para el objeto i y tomar como profundidad de superficie posterior un valor encontrado al añadir a esto el valor de thick(i).

Lo siguiente puede concebirse como una profundidad de superficie anterior representativa repfore(X,t,i) del objeto i en un tiempo dado t:

(1) Utilizando el valor máximo de la profundidad de superficie anterior del objeto i:

$$repfore(X,t,i) - \max_{x,y;X(t)[x,y].id=i} X(t)[x,y].fore$$

(2) Utilizando el valor promedio de la profundidad de superficie anterior del objeto i:

$$repfore(X,t,i) - \mathrm{avg}_{x,y;X(t)[x,y].id=i} X(t)[x,y].fore$$

Aquí, max y avg significan el valor máximo y el valor promedio de los valores de la parte principal cuando, en la parte del subíndice, las variables dispuestas delante del punto y coma cambian dentro del intervalo que satisface las condiciones dispuestas después del punto y coma. Cuando se usa "constantemente satisfecho" como condición, solo las variables se indican como el subíndice.

Es decir, esto significa enumerar combinaciones de x, y que satisfacen X(t)[x,y].id = i en X, t e i dados en los puntos (1) y (2) descritos anteriormente, encontrando el valor de X(t)[x,y].fore para cada una de esas combinaciones y encontrando el valor máximo o el valor promedio de las mismas.

Si se usa repfore(X,t,i) establecido de este modo, la profundidad de superficie posterior del píxel X(t)[x,y] cuando se representa el objeto i en el píxel X(t)[x,y] (i=X(t)[x,y].id) puede establecerse de tal forma que:

$$X(t)[x,y].back = repfore(X,t,i) + thick(i).$$

o,

$$X(t)[x,y].back = \max[repfore(X,t,i) + thick(i), X(t)[x,y].fore].$$

Aquí, max[...] significa el valor máximo de los valores dispuestos entre corchetes.

La Figura 2D es una vista en sección transversal de un objeto para el cual se estima la forma de la superficie posterior del objeto encontrando una profundidad de superficie anterior representativa del objeto y añadiendo espesor para encontrar la profundidad de superficie posterior. Como se muestra en el dibujo, la superficie posterior del objeto 14 es un plano perpendicular a la dirección de filmación 16, y el objeto 14 se asemeja a una forma

columnar que se extiende en la dirección de filmación 16.

Además, existe un método para aproximar la superficie posterior estableciendo un punto representativo del objeto i en el tiempo t, y creando una superficie esférica centrada en ese punto representativo.

Primero, el número de píxeles area(X,t,i) de la región donde se representa el objeto i en el tiempo t en el vídeo X se encuentra de la siguiente manera:

$$\text{area}(X,t,i) = \Sigma_{x,y;X(t)[x,y].id=i}\ 1$$

Aquí, el subíndice de Σ tiene el mismo significado que con max y avg.

Una coordenada horizontal xc(X,t,i) y una coordenada vertical yc(X,t,i) del punto representativo del objeto i en el tiempo t se establecen de la siguiente manera:

$$xc(X,t,i) = \Sigma_{x,y;X(t)[x,y].id=i}\ x/\text{area}(X,t,i);$$

$$yc(X,t,i) = \Sigma_{x,y;X(t)[x,y].id=i}\ y/\text{area}(X,t,i).$$

Además, una anchura w(X,t,i) y una altura h(X,t,i) de la región en la que se representa el objeto i en el tiempo t en el vídeo X se establecen de la siguiente manera:

$$w(X,t,i) = \max_{x,y;X(t)[x,y].id=i}\ x - \min_{x,y;X(t)[x,y].id=i}\ x;$$

$$h(X,t,i) = \max_{x,y;X(t)[x,y].id=i}\ y - \min_{x,y;X(t)[x,y].id=i}\ y.$$

Existen varios métodos para establecer el diámetro D(X,t,i) de la superficie esférica, tales como los que se muestran a continuación:

$$D(X,t,i) = \max\ [w(X,t,i), h(X,t,i)];$$

$$D(X,t,i) = (w(X,t,i)^2 + h(X,t,i)^2)^{1/2};$$

$$D(X,t,i) = \text{area}(X,t,i)^{1/2};$$

$$D(X,t,i) = \max_{x,y;X(t)[x,y].id=i}\ ((x-xc(X,t,i))^2 + (y-yc(X,t,i))^2)^{1/2};$$

$$D(X,t,i) = \text{avg}_t\ \text{area}(X,t,i)^{3/2}/\text{area}(X,t,i)$$

Aquí, avg_t area(X,t,i) significa el promedio de tiempo para la región en la que se representa el objeto i en el vídeo X. Por consiguiente, avg_t $\text{area}(X,t,i)^{3/2}$ corresponde al valor aproximado del volumen que ocupa el objeto i, y cuando se divide por el area(X,t,i), se obtiene un valor aproximado de la longitud de la profundidad.

Ahora, una vez que se establece el diámetro D(X,t,i), se puede establecer la coordenada de profundidad zc(X,t,i) del punto representativo de la siguiente manera:

$$zc(X,t,i) = X(t)[xc(X,t,i), yc(X,t,i)].\text{fore} + D(X,t,i)/2$$

Además, sería correcto simplemente establecer esto de la siguiente manera:

$$zc(X,t,i) = X(t)[xc(X,t,i), yc(X,t,i)].\text{fore}.$$

Con lo cual, si la profundidad de superficie posterior del píxel X(t)[x,y] cuando el objeto i se representa en el píxel X(t)[x,y] (i=X(t)[x,y].id) se establece por:

$$X(t)[x,y].back = \max [zc(X,t,i) + [\max [(D(X,t,i)/^{2}) - (x-xc(X,t,i))^{2} - (y-yc(X,t,i))^{2}],0]^{1/2}, X(t)[x,y].fore],$$

la superficie esférica se puede aproximar mejor, y si

$$X(t)[x,y].back = \max [zc(X,t,i) + D, X(t)[x,y].fore]$$

se establece, un cilindro puede aproximarse mejor.

La Figura 2E es un dibujo en sección transversal de un objeto con una forma cuya superficie posterior se supone que es una superficie esférica. En el ejemplo mostrado en este dibujo, la superficie posterior del objeto 14 es una superficie esférica centrada en un punto representativo 21, pero desde la superficie anterior del objeto 14 con respecto a la superficie esférica, la forma se asemeja más a la forma de un cilindro que se extiende en la dirección de filmación 16.

Además, como método para estimar la profundidad de superficie posterior del objeto, es posible usar una base de datos que almacena un atributo que indica el tipo y/o similares del objeto y la longitud de profundidad con respecto a ese atributo, vinculados entre sí.

En esta base de datos, el enlace es tal que la longitud de profundidad es "OO cm" con respecto al atributo "persona", y la longitud de profundidad es "OO m" con respecto al atributo "aeronave". Además, la longitud de profundidad del objeto se adquiere a partir de un atributo estimado mediante la identificación de imagen desde la apariencia externa del objeto representado en el vídeo, y un atributo seleccionado y establecido por el usuario independientemente para el objeto.

Cuando el objeto es un producto que se vende en un distrito comercial de productos electrónicos, tal como un armario o lavadora y/o similares, es posible adquirir la longitud de profundidad del objeto con respecto al atributo de la información sobre el tamaño del producto indicado en la especificación del producto.

Una vez que se obtiene la longitud de profundidad del objeto, es posible estimar la profundidad de superficie posterior del objeto combinando esto con los diversos métodos descritos anteriormente.

Hasta este punto, se han descrito métodos para establecer y estimar la profundidad de superficie posterior de un objeto en movimiento, pero sería correcto pensar en la profundidad de superficie posterior como ∞ para objetos correspondientes al fondo. Es decir, cuando

$$X(t)[x,y].id = 0$$

se establece, entonces

$$X(t)[x,y].back = \infty.$$

La zona de ocupación en la dirección de profundidad ocupada en el píxel X(t)[x,y] por el objeto X(t)[x,y].id representado en el píxel X(t)[x,y] en el vídeo X en el tiempo t se extiende desde la profundidad de superficie anterior X(t)[x,y].fore a la profundidad de superficie posterior X(t)[x,y].back.

Si hay superposición en la zona de ocupación desde la profundidad de superficie anterior a la profundidad de superficie posterior de dos objetos representados en la misma posición de píxel, esos dos objetos están en conflicto.

Por consiguiente, para cualquier t, x, y:

$$X(t)[x,y].id = i;$$

$$Y(t)[x,y].id = j,$$

y si se da alguna de las cuatro relaciones:

$$(1)\quad X(t)[x,y].fore \leq Y(t)[x,y].fore \leq X(t)[x,y].back;$$

$$(2)\quad X(t)[x,y].fore \leq Y(t)[x,y].back \leq X(t)[x,y].back;$$

$$(3)\quad Y(t)[x,y].fore \leq X(t)[x,y].fore \leq Y(t)[x,y].back;$$

$$(4)\quad Y(t)[x,y].fore \leq X(t)[x,y].back \leq Y(t)[x,y].back.$$

se establece, cuando el vídeo X y el vídeo Y se superponen, es posible determinar que el objeto i representado en el vídeo X y el objeto j representado en el vídeo Y entran en conflicto.

A continuación, la cantidad de superposición en las dos zonas de ocupación se indica como overlap(X,Y,t,x,y). Si no se cumple ninguno de los puntos (1) a (4) anteriores, overlap(X,Y,t,x,y) = 0, y si overlap(X,Y,t,x,y) > 0, los objetos i y j entran en conflicto.

Específicamente, en el caso del punto (1) anterior,

$$overlap(X,Y,t,x,y) = \min [X(t)[x,y].back, Y(t)[x,y].back] - Y(t)[x,y].fore,$$

en el caso del punto (2) anterior

$$overlap(X,Y,t,x,y) = Y(t)[x,y].back - \max [X(t)[x,y].fore, Y(t)[x,y].fore],$$

en el caso del punto (3) anterior,

$$overlap(X,Y,t,x,y) = \min [X(t)[x,y].back, Y(t)[x,y].back] - X(t)[x,y].fore,$$

y en el caso del punto (4) anterior,

$$overlap(X,Y,t,x,y) = X(t)[x,y].back - \max [X(t)[x,y].fore, Y(t)[x,y].fore].$$

Aquí, min [...] es lo opuesto de max [...] y devuelve el valor mínimo de los valores entre corchetes.

Cuando se establece cualquiera de los puntos (1) a (4) anteriores y overlap(X,Y,t,x,y) = 0, es posible pensar que los objetos i y j no están separados sino que están en contacto entre sí. Por consiguiente, es posible distinguir entre un conflicto en el que los dos se invaden mutuamente entre sí y un conflicto en el que las superficies se tocan.

Además, sería correcto en la determinación de conflictos no usar la propia zona de ocupación descrita anteriormente, sino averiguar si existe una superposición ampliando ligeramente la zona de ocupación añadiendo un margen al límite superior y el límite inferior de la misma. En este caso, al hacer que se añada un margen a la determinación de conflictos de los objetos i y j, se determina que los objetos i y j están en conflicto (o en peligro) si no se separan por una cierta distancia.

Por otro lado, si el margen descrito anteriormente es negativo, es posible permitir conflictos con una ligera invasión. Esto es adecuado cuando el objeto es elástico y se asumen indentaciones, y/o similares.

La Figura 3 es un dibujo explicativo que muestra la composición resumida de elementos para llevar a cabo una determinación de conflictos, en el dispositivo de procesamiento de vídeo de acuerdo con esta realización preferida.

Como se muestra en este dibujo, los elementos para llevar a cabo una determinación de conflictos en el dispositivo de procesamiento de vídeo 101 comprenden un adquirente de profundidad de superficie posterior 102 y un determinador de conflictos 103.

El dispositivo de procesamiento de vídeo 101 toma el primer vídeo X y el segundo vídeo Y como objetivos de procesamiento. Por ejemplo,

(a) como primer vídeo X, se utiliza una imagen de 2,5 dimensiones filmando a un bailarín que baila en el mundo real,

(b1) como el segundo vídeo Y, se utiliza un vídeo que expresa un personaje que baila sin fondo, y
(b2) como el segundo vídeo Y, se utiliza un vídeo filmado para expresar otro usuario bailando y luego el otro usuario ha realizado la edición y eliminado toda la información de fondo para que el otro usuario esté bailando sin fondo.

El dispositivo de procesamiento de vídeo 101 en esta realización preferida finalmente compila un tercer vídeo en el que un bailarín baila junto con un personaje u otro usuario.

En consecuencia, en primer lugar, el dispositivo de procesamiento de vídeo 101 determina un conflicto entre el primer objeto i representado en el primer vídeo X y el segundo objeto j representado en el segundo vídeo Y que debería superponerse en el primer vídeo X. El dispositivo de procesamiento de vídeo 101 se realiza típicamente ejecutando un programa en un ordenador.

Aquí, el primer vídeo X va acompañado de una primera profundidad de superficie anterior del primer objeto i en el lado representado en el primer vídeo X, y el segundo vídeo Y va acompañado de una segunda profundidad de superficie anterior del segundo objeto j en el lado representado por el segundo vídeo Y.

Usando la notación anterior, si el primer objeto i se representa en un píxel en la posición [x,y] en el fotograma X(t) en el tiempo t contenido en el primer vídeo X, entonces X(t)[x,y].id = i y la primera profundidad de superficie anterior es X(t)[x,y].fore.

El primer vídeo X puede comprender un objeto en movimiento solo o puede incluir fondo. En el primer caso, un objeto con i >0 se utiliza como el primer objeto, y en el último caso, un objeto con i ≥0 se utiliza como el primer objeto.

Además, si el segundo objeto j se representa en un píxel en la posición [x,y] en un fotograma Y(t) en el tiempo t contenido en el segundo vídeo Y, entonces Y(t)[x,y].id = j y la segunda profundidad de superficie anterior es Y(t)[x,y].fore.

Los objetos que son el objetivo de la determinación de conflictos en el segundo vídeo Y no incluyen el fondo. Es decir, como el segundo objeto, se utilizan objetos con j >0.

En primer lugar, el adquiriente de profundidad de superficie posterior 102 adquiere la primera profundidad de superficie posterior del primer objeto i en el lado no representado en el primer vídeo X, y la segunda profundidad de superficie posterior del segundo objeto j en el lado no representado en el segundo vídeo Y.

Si se utiliza la notación descrita anteriormente, el adquirente de superficie posterior 102 encuentra la primera profundidad de superficie posterior X(t)[x,y].back a través del método descrito anteriormente para el tiempo t y la posición [x,y] que satisface X(t)[x,y].id = i. Además, el adquirente de superficie posterior 102 encuentra la segunda profundidad de superficie posterior Y(t)[x,y].back a través del método descrito anteriormente para el tiempo t y la posición [x,y] que satisface Y(t)[x,y].id = j.

Por otro lado, el determinador de conflictos 103 funciona como un determinante de interferencias para determinar la interferencia entre objetos. Además, el determinador de conflictos 103 determina que el primer objeto y el segundo objeto entran en conflicto si la primera área de dibujo donde debería representarse el primer objeto i y la segunda área de dibujo donde debería representarse el segundo objeto j se superponen como resultado de la superposición del primer vídeo X y el segundo vídeo Y, y la primera zona de ocupación desde la primera profundidad de superficie anterior hasta la primera profundidad de superficie posterior del primer objeto i y la segunda zona de ocupación desde la segunda profundidad de superficie anterior hasta la segunda profundidad de superficie posterior del segundo objeto j que se superponen en un área de superposición donde la primera área de dibujo y la segunda área de dibujo se superponen.

Si se usa la notación descrita anteriormente, la primera área de dibujo en un tiempo dado t es el conjunto de posiciones [x,y] que satisface X(t)[x,y].id = i, y la segunda área de dibujo en un tiempo dado t es el conjunto de posiciones [x,y] que satisface Y(t)[x,y].id = j. Si la parte común de estos conjuntos no es el conjunto vacío, entonces la primera área de dibujo y la segunda área de dibujo se superponen en el tiempo t, y la parte común de estos conjuntos corresponde al área de superposición en el tiempo t.

Para una posición [x,y] que es un elemento del área de superposición en el tiempo t, claramente:

$$X(t)[x,y].id = i;$$

$$Y(t)[x,y].id = j$$

se establecen.

En esta posición [x,y], el primer objeto i ocupa la primera zona de ocupación de X(t)[x,y].fore a X(t)[x,y].back, como la profundidad. Además, el primer objeto j ocupa la segunda zona de ocupación de Y(t)[x,y].fore a Y(t)[x,y].back.

Cuando la primera zona de ocupación y la segunda zona de ocupación se superponen, es decir, cuando la segunda profundidad de superficie anterior o la segunda profundidad de superficie posterior se incluye en la primera zona de ocupación o la primera profundidad de superficie anterior o la primera profundidad de superficie posterior se incluye en la segunda zona de ocupación, se determina que el primer objeto i y el segundo objeto j entran en conflicto en el tiempo t y la posición [x,y].

En la explicación descrita anteriormente, si la primera área de dibujo en la que se representa el primer objeto i y la segunda área de dibujo en la que se representa el segundo objeto j se solapan en el fotograma en el tiempo t, se logra una determinación de conflictos para el primer el objeto i y el segundo objeto j en el fotograma en el tiempo t usando información sobre la profundidad de superficie posterior.

En consecuencia, para confirmar que no existe un conflicto en un vídeo en su conjunto, sería bueno realizar la determinación de conflictos descrita anteriormente estableciendo tiempos sucesivos para todos los fotogramas en el primer vídeo X y el segundo vídeo Y.

La Figura 4 es un diagrama de flujo que muestra el flujo del proceso de determinación de conflictos de acuerdo con esta realización preferida. A continuación, el proceso de determinación de conflictos ejecutado por el dispositivo de procesamiento de vídeo 101 de acuerdo con esta realización preferida se describe con referencia a este dibujo.

En primer lugar, el dispositivo de procesamiento de vídeo 101 recibe el primer vídeo X y el segundo vídeo Y que son objetivos de determinación de conflictos (etapas S151). Los objetos se representan en el primer vídeo X y el segundo vídeo Y, y se adjunta la información de profundidad del objeto.

A continuación, se repite un proceso de escaneo de fotogramas del primer vídeo X y el segundo vídeo Y en orden desde el inicio (etapa S152).

Es decir, en el fotograma que se está escaneando actualmente, se hace una determinación de si un área en la que se representa un objeto en el primer vídeo X y un área en la que se representa un objeto en el segundo vídeo Y se superponen (etapa S153). Si no hay superposición (etapa S153: No), el proceso avanza a la etapa S157 y el proceso se repite.

Si hay superposición (etapa S153: Sí), se adquiere información de profundidad sobre el objeto en el primer vídeo X para el que hay superposición en esa región de dibujo (correspondiente al "primer objeto" descrito anteriormente), e información de profundidad sobre el objeto en el segundo vídeo Y (correspondiente al "segundo objeto" descrito anteriormente) (etapa S154).

Además, se realiza una determinación de si existe una superposición en las zonas de ocupación de los objetos obtenidos a partir de la información de profundidad (etapa S155). Si no hay solapamiento, el proceso avanza a la etapa S157 y el proceso se repite.

Si hay una superposición (etapa S155: Sí), se produce un resultado de determinación al efecto de que el objeto representado en el primer vídeo X y el objeto representado en el segundo vídeo Y entran en conflicto (etapa S156), y el proceso luego concluye.

Cuando concluye la repetición de las etapas S 152 a S 157, se produce un resultado de determinación al efecto de que el objeto representado en el primer vídeo X y el objeto representado en el segundo vídeo Y no entran en conflicto (etapa S158), y el proceso luego concluye.

En la descripción anterior, para facilitar el entendimiento de que la suposición era solo un primer objeto se

representa en el primer vídeo X y solo un segundo objeto se representa en el segundo vídeo Y, pero cuando hay múltiples objetos, sería correcto repetir el mismo proceso para cada objeto.

Además, en la descripción anterior, se realiza una determinación de superposición de zonas de ocupación basándose en información de profundidad y superposición de áreas de dibujo para cada fotograma, pero cuando hay numerosos fotogramas, es posible realizar un muestreo adecuado.

Por ejemplo, si se utiliza un método en el que inicialmente es posible determinar si la primera área de dibujo y la segunda área de dibujo se superponen seleccionando fotogramas mientras se salta con un espacio equivalente, y cuando hay una superposición, para detener el muestreo de los fotogramas antes y después, es posible ahorrar el tiempo de cálculo necesario para hacer una determinación de conflictos para los vídeos en su conjunto.

En general, cuando no se determina que "en cualquier momento y en cualquier posición, cualquiera de los objetos representados en el primer vídeo X y cualquiera de los objetos representados en el segundo vídeo Y entran en conflicto", es decir, cuando no se generan conflictos en absoluto, no surge contradicción en la relación frontal y posterior de los objetos, incluso si el segundo vídeo Y se superpone al primer vídeo X. Por consiguiente, es posible encontrar el tercer vídeo superimpose(Y, X) con el segundo vídeo Y superpuesto en el primer vídeo X.

Por otro lado, cuando se determina que "en cualquier momento y en cualquier posición, cualquiera de los objetos representados en el primer vídeo X y cualquiera de los objetos representados en el segundo vídeo Y entran en conflicto", no es posible superponer el primer vídeo X y el segundo vídeo Y. Por consiguiente, es necesario un esquema tal como el que se describe a continuación.

Segunda realización preferida

Esta realización preferida diseña una interfaz de usuario cuando se superpone desplazando el segundo vídeo Y al primer vídeo X a través de una operación de arrastrar y soltar y/o similar usando un ratón.

Con esta realización preferida, para facilitar la comprensión se asumirá un caso en el que a través de la operación de arrastrar y soltar del usuario, se especifica una cantidad de traslación hacia arriba, abajo, izquierda o derecha para desplazar y superponer el segundo vídeo Y en el primer vídeo X. Sin embargo, el sistema de operación de esta realización preferida no está necesariamente limitado a arrastrar y soltar, ya que también es posible reemplazar la función arrastrar y soltar a través de la utilización de diversas operaciones de entrada especificadas, tales como una acción de deslizamiento en una pantalla táctil, o dos veces repitiendo una operación de clic, y/o similares.

La Figura 5A es un dibujo explicativo que muestra el estado de la pantalla antes de iniciarse el arrastre, la Figura 5B es un dibujo explicativo que muestra el estado de la pantalla en un punto en el tiempo en que se inicia el arrastre, la Figura 5C es un dibujo explicativo que muestra el estado de la pantalla en un punto en el tiempo durante el arrastre, y la Figura 5D es un dibujo explicativo que muestra el estado de la pantalla en un punto en el momento en que se realiza la caída.

Además, la Figura 6 es un dibujo explicativo que muestra la composición resumida de partes del dispositivo de procesamiento de vídeo 101 relacionado con el proceso de arrastrar y soltar. La Figura 7 es un diagrama de flujo que muestra el flujo de control del proceso de arrastrar y soltar ejecutado por el dispositivo de procesamiento de vídeo 101 de acuerdo con esta realización preferida. La explicación a continuación hace referencia a estos dibujos.

En el proceso de arrastrar y soltar de acuerdo con esta realización preferida, en el dispositivo de procesamiento de vídeo 101, un ajuste de cantidad de deslizamiento 104 y un generador de vídeo 105 actúan basándose en la operación de arrastrar y soltar del usuario.

El dispositivo de procesamiento de vídeo 101, al deslizar y superponer el segundo vídeo Y sobre el primer vídeo X, está dotado de un determinante de interferencias para determinar si las condiciones de posición se satisfacen, de manera que la relación posicional de los objetos satisfaga las condiciones de posición prescritas.

Sin embargo, en esta realización preferida, para utilizar un conflicto de objetos que no se produce como una condición de posición, el determinador de conflictos descrito anteriormente 103 se utiliza como el determinante de interferencia, y se usa el resultado de la determinación del mismo. En este dibujo, el adquirente de profundidad de superficie posterior 102 frente al determinador de conflictos 103 se omite del dibujo.

Como se describe a continuación, como condición de posición, además de los objetos que no interfieran, es decir, que no se satisfagan las condiciones de interferencia, es posible establecer otras condiciones, por ejemplo, que los objetos prescritos estén inmediatamente adyacentes, y/o similares. Para los resultados de la determinación del determinador de conflictos descrito anteriormente 103, es posible usar también una determinación de las condiciones de posición que haga referencia a relaciones de posición distintas de la interferencia de objetos.

Además, un corrector 106 que se muestra en el dibujo es un elemento para corregir la cantidad de deslizamiento por la cual el segundo vídeo Y se desliza desde "sin deslizamiento" de manera que no se produzca interferencia entre los objetos.

En primer lugar, el corrector 106 muestra en la pantalla un fotograma representativo del primer vídeo X y un fotograma representativo del segundo vídeo Y (etapa S181). Por ejemplo, como se muestra en la Figura 5A, una ventana 411 del primer vídeo X, una ventana 412 del segundo vídeo Y, y un cursor 413 se muestran en una pantalla 401.

Por ejemplo, utilizar el primer fotograma de cada vídeo como los fotogramas representativos es el método más fácil. Además de esto, cuando el primer vídeo X y el segundo vídeo Y se desplazan en el tiempo, sería correcto que el usuario seleccione los fotogramas para la sincronización y que estos sean los fotogramas representativos. En este caso, la diferencia en los tiempos transcurridos en los fotogramas sincronizados corresponde a la cantidad de desplazamiento para el desplazamiento en el tiempo del segundo vídeo Y.

En este dibujo, el primer vídeo X es un vídeo con un fondo y el borde de la ventana 411 se representa mediante una línea continua. Por otro lado, el segundo vídeo Y es un vídeo sin fondo. La periferia de la ventana 412 se representa con un color transparente y el borde de la misma se representa por una línea de puntos. El borde de la ventana 412 no necesita necesariamente mostrarse en la pantalla 401. En este caso, el usuario ve una visualización en la pantalla 401 como si el objeto sin fondo solo fuera representado en el segundo vídeo Y.

A continuación, el dispositivo de procesamiento de vídeo 101 recibe una acción de clic (inicio de arrastre) en la ventana 412 del segundo vídeo Y del usuario (etapa S182). Aquí, como se muestra en la Figura 5B, el usuario ha usado el ratón para hacer que el cursor 413 se mueva y hace clic en una posición [s,t] en el segundo vídeo Y.

Además, el dispositivo de procesamiento de vídeo mueve la posición de visualización de la ventana 412 del segundo vídeo Y de acuerdo con esto (etapa S184) mientras que el arrastre por parte del usuario continúa (etapa S183: Sí). Por ejemplo, cuando el usuario hace que el cursor 413 se mueva mientras se mantiene el clic, como se muestra en la Figura 5C, se hace que la ventana 412 en su conjunto se mueva de tal manera que el cursor 413 indique la posición [s,t] en el segundo vídeo Y.

Cuando el usuario ha hecho que el cursor 413 se mueva hasta una posición [u,v] en el primer vídeo X, como se muestra en la Figura 5D, y detiene el clic, la ventana 412 se suelta en la ventana 411.

Después de lo cual, la posición [u,v] en el primer vídeo X y la posición [s,t] en el segundo vídeo Y se superponen. Por consiguiente, se puede pensar que el usuario ha superpuesto el segundo vídeo Y sobre el primer vídeo X realizando una transformación que hace que el segundo vídeo Y se mueva por u-s en la dirección horizontal y por v-t en la dirección vertical.

De esta manera, cuando el arrastre por parte del usuario finaliza (etapa S183: No), el graduador de cantidad de desplazamiento 104 del dispositivo de procesamiento de vídeo 101 establece la dirección horizontal u-s y la dirección vertical v-t como la cantidad de desplazamiento (etapa S185). Es decir, el dispositivo de procesamiento de vídeo 101 interpreta el cambio a medida que el usuario superpone el vídeo move(u-s, v-t, Y) obtenido moviendo el segundo vídeo hacia arriba, abajo, izquierda o derecha en paralelo con el fotograma, sobre el primer vídeo X.

Por lo tanto, el dispositivo de procesamiento de vídeo 101, usando el determinador de conflictos 103, determina si el objeto representado en el primer vídeo X y el objeto representado en el vídeo intermedio move(u-s, v-t, Y) entran en conflicto al desplazar el segundo vídeo Y en cualquier tiempo de fotograma o cualquier posición en el fotograma, o si no se produce un conflicto (etapa S186).

Cuando el resultado de la operación de soltar mostrada en la Figura 5C es que no se produce un conflicto (etapa S186: No), el generador de vídeo 105 del dispositivo de procesamiento de vídeo 101 produce el tercer vídeo superimpose(move(u-s, v-t, Y), X) a través de la composición, y el proceso concluye.

A través de este tipo de composición, es posible que el usuario obtenga un tercer vídeo en el que los objetos representados en el vídeo Y se colocan de manera que no entren en conflicto con los objetos del vídeo X. En el tercer vídeo, dado que no se producen conflictos con objetos representados en el vídeo Y, es posible evitar representaciones antinaturales tales como objetos que se ingieren entre sí.

Por otro lado, cuando se produce un conflicto incluso cuando se hace la operación de soltar mostrada en la Figura 5C (etapa S186: Sí), el dispositivo de procesamiento de vídeo 101 produce una advertencia tal como audio, vibración, una visualización en pantalla y similares (etapa S188), el proceso concluye y no se produce un nuevo vídeo. Es decir, si los comandos del usuario son apropiados, se genera el tercer vídeo, pero si los comandos son inapropiados, solo se genera una advertencia y el tercer vídeo no se genera.

Las diversas transformaciones descritas anteriormente se llevan a cabo en unidades de fotograma. Por consiguiente, cuando se lleva a cabo una transformación en un vídeo determinado para obtener un vídeo diferente, después de los procesos de transformación realizados tales como move y/o similares en cada fotograma, sería correcto ejecutar un proceso de cálculo diferente tal como una determinación de conflictos y/o similares, y sería correcto utilizar una evaluación de extensión impulsada por la demanda para llevar a cabo los procesos de transformación.

Si los procesos de transformación se llevan a cabo impulsados por la demanda, incluidos los ejemplos de aquí en adelante, es necesario realizar una determinación de superposición y/o similares para un fotograma dado, e inicialmente la transformación se logra para ese fotograma. Por consiguiente, para obtener una conclusión de que no se producen conflictos, es necesaria una transformación para todos los fotogramas, pero cuando se identifica un conflicto en un fotograma en el vídeo, el proceso de transformación no se ejecuta para fotogramas posteriores.

Además, con el método para especificar la cantidad de desplazamiento a través de arrastrar y soltar, la cantidad de desplazamiento se expresa según cuánto se ha desplazado la posición de la ventana caída 412 del segundo vídeo Y respecto a la posición de la ventana 411 del primer vídeo X. Sin embargo, debido a que la cantidad de desplazamiento es una cantidad aplicada al segundo vídeo Y, al ingresar la cantidad de desplazamiento, no se requiere una comparación con el primer vídeo X. Por consiguiente, sería correcto que el usuario use el teclado e ingrese directamente un valor numérico, o use un método para ingresar visualmente la cantidad de desplazamiento usando un ratón y/o similares.

A continuación, se describe un método para realizar el ajuste automático en lugar de simplemente emitir una advertencia cuando las instrucciones del usuario son inapropiadas.

Tercera realización preferida

Con esta realización preferida, cuando un objeto del primer vídeo X y un objeto del segundo vídeo Y entran en conflicto debido a la cantidad de traslación del segundo vídeo Y especificado por el usuario, ese conflicto se evita y el ajuste automático se lleva a cabo de manera que el objeto del segundo vídeo Y se posicione para tocar el suelo o el techo del primer vídeo X.

Con este ajuste automático, se repite una prueba de si se produce un conflicto al trasladar el segundo vídeo Y hacia arriba, abajo, izquierda o derecha al menos una vez, y en muchas condiciones se ejecuta repetidamente varias veces. Este proceso se realiza por el corrector 106 del dispositivo de procesamiento de vídeo 101. La Figura 8 es un diagrama de flujo que muestra el flujo de control del proceso de ajuste automático de acuerdo con esta realización preferida. La explicación a continuación hace referencia a este dibujo.

La cantidad de traslación hacia arriba, abajo, izquierda o derecha en una prueba se indica como Δp en la dirección horizontal y Δq en la dirección vertical. Cuando se usan constantes para Δp y Δq, el segundo vídeo Y se ajusta mientras se desplaza en una dirección preestablecida.

Para facilitar la comprensión, de forma similar a las Figuras 5A a 5D, el caso se considerará cuando un usuario haga clic en una posición [s,t] en el segundo vídeo Y, arrastre a una posición [u,v] en el primer vídeo X mientras se mantiene el clic y luego detiene el clic y suelta.

En este caso, el dispositivo de procesamiento de vídeo 101 de la realización preferida realiza el siguiente proceso.

Primero, el graduador de cantidad de desplazamiento 104 del dispositivo de procesamiento de vídeo 101 establece respectivamente el valor inicial de la variable p como u y el valor inicial de la variable q como v (etapa S191). A través de esto, el valor inicial de la cantidad de desplazamiento en la dirección horizontal se establece como p-s = u-s y el valor inicial de la cantidad de desplazamiento en la dirección vertical se establece como q-t = v-t, y el graduador de cantidad de desplazamiento 104 funciona como un graduador de valor inicial.

A continuación, el determinador de conflictos 103 del dispositivo de procesamiento de vídeo 101 determina si se establece la condición de posición "no se produce un conflicto de objetos entre el primer vídeo X y el vídeo intermedio move(p-s, q-t, Y) que es el segundo el vídeo Y desplazado en una cantidad de desplazamiento establecida" (etapa S192).

Si esto se establece (etapa S192: Sí), el generador de vídeo 105 produce el tercer vídeo superimpose(move(p-s, q-t,Y),X) a través de la composición (etapa S913), y este proceso concluye.

Por otro lado, si no se establece esta condición, el corrector 106 aumenta el valor de la variable p por Δp y aumenta el valor de la variable q por Δq (etapa S194). A través de esto, la cantidad de desplazamiento se actualiza ligeramente.

Además, cuando se satisface una condición de abandono, tal como cuando la cantidad de traslación (p-s, q-t) del vídeo Y llega a ser demasiado grande, o cuando el número de repeticiones excede un valor umbral prescrito, y/o similar (etapa S195: Sí), se determina que no se puede evitar un conflicto, y se emite una advertencia (etapa S196) y este proceso concluye.

Por otro lado, cuando la condición de abandono no se satisface y la repetición continúa (etapa S 195: No), el control del dispositivo de procesamiento de vídeo 101 vuelve a la etapa S192.

En los sistemas de coordenadas de numerosos vídeos, la dirección de aumento del valor de coordenadas en la dirección vertical corresponde a la dirección descendente en la pantalla. En el posicionamiento moviendo el objeto representado en el segundo vídeo Y en la dirección descendente de la pantalla, sería correcto establecer (Δp,Δq) = (0,1).

Además, como resultado de que (Δp,Δq) = (0,1) se establezca inicialmente y se ejecute el proceso de ajuste automático descrito anteriormente, cuando la determinación es que un conflicto es inevitable, sería correcto restablecer (Δp,Δq) = (0,-1) y ejecutar de nuevo el mismo proceso de ajuste automático.

Esto corresponde a la primera búsqueda de una cantidad de traslación en la que pueda evitarse un conflicto mientras se mueve de arriba a abajo, y si no se puede encontrar, entonces a buscar una cantidad de traslación en la que pueda evitarse un conflicto mientras se mueve de abajo a arriba.

En la explicación descrita anteriormente, el determinador de conflictos 103 se utilizó simplemente como el determinador de interferencia. Esto corresponde a la imposición de la condición de posición de los conflictos de objeto que no se producen para la relación de posición entre los objetos. Sin embargo, si se usan resultados de determinación del determinador de conflictos 103, es posible utilizar otras condiciones como condiciones de posición en el determinador de interferencia.

Por ejemplo, es posible cambiar la condición de posición a ""no se producen conflictos de objetos entre el primer vídeo X y el vídeo move(p-s, q-t, Y)" y "se produce un conflicto de objetos entre el primer vídeo X y el vídeo move(p-s, q+1-t, Y)"". Al hacer este cambio, el objeto representado en el segundo vídeo Y se coloca de manera que virtualmente toque el suelo al superponer el objeto representado en el segundo vídeo Y sobre el primer vídeo X.

Además, también es posible cambiar la condición de posición a ""no se producen conflictos de objetos entre el primer vídeo X y el vídeo move(p-s, q-t, Y)" y "se produce un conflicto de objetos entre el primer vídeo X y el vídeo move(p+Δp-s, q+Δq-t, Y)"".

En este caso, cuando se encuentra una cantidad de traslación en una dirección dada, el ajuste se realiza para avanzar en la medida de lo posible y se repite la prueba. Por ejemplo, cuando se mueve de arriba hacia abajo, el objeto del segundo vídeo Y se mueve hasta tocar el suelo del primer vídeo X, y cuando se mueve de abajo a arriba, el objeto del segundo vídeo Y se mueve hasta tocar el techo del primer vídeo X.

En los casos en que el suelo se representa en el primer vídeo X, se representa un objeto que se mueve en un plano horizontal en el segundo vídeo Y y la orientación de la cámara en el primer vídeo X con respecto al suelo y la orientación de la cámara en el segundo vídeo Y con respecto a la horizontal coinciden sustancialmente y/o similares, si se lleva a cabo el proceso descrito anteriormente, cuando el objeto representado en el segundo vídeo Y se superpone en el primer vídeo X, el objeto representado en el segundo vídeo Y toca sustancialmente el suelo o está posicionado para flotar separado del suelo.

Hay casos en los que un objeto representado en el primer vídeo X está parado, tal como el suelo, el techo, la pared y/o similares. Por ejemplo, hay casos en los que un usuario ha especificado un objeto en un fotograma determinado de antemano, casos en los que se entiende el historial de edición anterior por el usuario, y casos en que cada fotograma del primer vídeo X se comprende a través del contraste.

En este caso, establecer las condiciones de posición como se ha descrito anteriormente significa que el objeto compuesto establece el rango móvil que se puede ocupar en el espacio tridimensional. Es decir, es posible lograr la composición de vídeo en la que se imponen las condiciones para el rango móvil del objeto en el espacio tridimensional, tal como "el objeto puede flotar en el espacio si no hay conflicto", "el objeto toca el piso o el suelo", "el objeto toca el techo" y/o similares.

Como se ha descrito anteriormente, $(\Delta p,\Delta q)$ es un vector que expresa el desplazamiento de la posición de dibujo del objeto representado en el segundo vídeo Y, y ajustando la orientación del mismo, es posible cambiar apropiadamente la dirección del movimiento y la cantidad de movimiento.

En la situación descrita anteriormente, moviendo el segundo vídeo Y a lo largo de una dirección establecida de antemano y superponiendo la imagen en el primer vídeo X, se genera un nuevo vídeo de tal forma que los objetos representados en el primer vídeo X y el segundo vídeo Y no entren en conflicto, y se asume un movimiento paralelo solo en la dirección ascendente o descendente del vídeo Y.

Sin embargo, después de que se haya realizado el movimiento descrito anteriormente, sería correcto corregir la cantidad de desplazamiento para minimizar la diferencia entre la cantidad de desplazamiento y el valor inicial establecido por el usuario, al mismo tiempo que se satisfacen las condiciones de posición. En este caso, se usa la cantidad de desplazamiento más cercana a la designación del usuario.

Además, dado que la dirección de $(\Delta p,\Delta q)$, es posible utilizar la diferencia vectorial entre un punto representativo del primer objeto y un punto representativo del segundo objeto. En este caso, se realiza una traslación de manera que el segundo objeto se mueva en una dirección que se aproxima al primer objeto o en una dirección que se aleja del primer objeto.

Procediendo adicionalmente con estos métodos, sería correcto utilizar un método que establezca apropiadamente una función de evaluación y evitar conflictos de objetos al minimizar esta función. Esto se puede concebir como la minimización de la "cantidad de conflicto" cuando los objetos entran en conflicto.

Cuando se considera la magnitud de conflicto del vídeo X y el vídeo Y como la suma de las áreas de zonas de conflicto donde se representan los objetos en conflicto en el vídeo X y el vídeo Y, la cantidad de conflicto se puede calcular a través de:

$$\mathrm{collision}(Y,X) = \Sigma_{t,x,y;X(t)[x,y].id\geq 0,\ Y(t)[x,y].id>0,\ overlap(X,Y,t,x,y)>0}\ 1$$

Además, cuando la cantidad estimada de la suma de volúmenes donde los objetos en conflicto se superponen en el espacio tridimensional se toma como la cantidad de conflicto, es posible utilizar:

$$\mathrm{collision}(Y,X) = \Sigma_{t,x,t;X(t)[x,y].id\geq 0,\ Y(t)[x,y].id>0,\ overlap(X,Y,t,x,y)>0}\ \mathrm{overlap}(X,Y,t,x,y).$$

La cantidad de conflicto calculada de esta manera se toma como la función de evaluación.

Para minimizar la cantidad de conflicto, sería correcto actualizar repetidamente el valor de la variable que indica el grado de desplazamiento del vídeo Y, tal como p y q y/o similares ajustando la dirección de actualización $(\Delta p,\Delta q)$ de la variable desajustada a través de un camino aleatorio, el método de descenso más pronunciado o el método de gradiente conjugado, preparando (1,0), (0,10), (-1,0), (0,-1) y/o similares de antemano como candidatos para $(\Delta p,\Delta q)$, encontrando el cambio en la cantidad de conflicto al utilizar estos y seleccionando aquello para lo cual la

cantidad del conflicto desciende en mayor medida.

Por ejemplo, cuando se lleva a cabo la traslación, es posible encontrar las cantidades de traslación p-s y q-t para evitar conflictos entre objetos si s, t, u y v se establecen como se ha descrito anteriormente, los valores iniciales de las variables p y q se toman como u y v, se encuentra que p y q minimizan collision(move(p-s, q-t, Y),X) y la collision(move(p-s, q-t, Y),X)=0 se establece en el punto final al que ha llegado.

Adicionalmente, además de las traslaciones arriba, abajo, izquierda y derecha, sería correcto combinar arbitrariamente la traslación en la dirección de profundidad, rotación, escala y desfase de tiempo.

Por ejemplo, cuando se combinan las traslaciones arriba, abajo, izquierda o derecha y en la dirección de profundidad, s, t, u y v se configuran como se ha descrito anteriormente, u, v y 0 se toman como valores iniciales de p, q y r, y se encuentra que p, q y r minimizan la collision(push(r, move(p-s, q-t, Y)),X).

Cuando se combinan las rotaciones en torno al eje horizontal y el eje vertical y las traslaciones, s, t, u y v se configuran como se ha descrito anteriormente, u, v, 0 y 0 se toman como los valores iniciales de p, q, θ y φ, y se encuentra que p, q, θ y φ minimizan collision(rotver(φ, rothor(θ, move(p-s, q-t, Y))), X).

Cuando se producen traslaciones o rotaciones extremas, se hace necesaria una representación convencional de la superficie posterior del objeto. Sin embargo, si el vídeo Y es un vídeo de 2,5 dimensiones, no hay información sobre la superficie posterior del objeto. Por lo tanto, los límites superiores y los límites inferiores se establecen por adelantado para las cantidades de traslación p-s, q-t y r y las cantidades de rotación θ, φ, ψ y/o similares, y a través de esto es posible evitar que el movimiento de un objeto en el vídeo Y se vea de forma no natural.

Además, para reducir la cantidad de desplazamiento en la medida de lo posible, también hay un método para tomar como función de evaluación el resultado de añadir la cantidad de conflicto descrita anteriormente a un valor encontrado multiplicando una constante entera prescrita por el valor cuadrado o el valor absoluto de un parámetro de transformación de las cantidades de traslación p-s, q-t o r o las cantidades de rotación θ, φ, y/o similares, y continuar actualizando el parámetro de transformación hasta que la cantidad de conflicto se convierte en 0. Cuando el valor inicial del parámetro de transformación se ajusta a un valor especificado por el usuario, si el desplazamiento en el parámetro de transformación se combina con la función de evaluación, es posible encontrar la cantidad de desplazamiento más cercana al valor especificado por el usuario.

Además, a continuación se considera una situación en la que la cantidad de traslación en la dirección de profundidad se ajusta por orden del usuario.

En primer lugar, con la cantidad de traslación en la dirección de profundidad tomada como 0, en el método descrito anteriormente, p y q se seleccionan de manera que el objeto toque el suelo. Como resultado, se obtuvo un vídeo con el vídeo move(p-s, q-t, Y) superpuesto en el primer vídeo X.

A continuación, el usuario especifica una cantidad de traslación r en la dirección de profundidad a través del manejo de la rueda del ratón y/o similares.

Después, la actualización de la variable (p,q) se repite usando (Δp,Δq) igual que en la situación descrita anteriormente hasta que se satisface la condición "objetos que no entran en conflicto entre el primer vídeo X y el vídeo push(r, move(p-s, q-t, Y))" y "objetos que entran en conflicto entre el primer vídeo X y el vídeo push(r, move(p-s, q+1-t, Y))".

Al utilizar estas situaciones, es posible generar un vídeo en el cual el movimiento de profundidad es solo la cantidad de traslación especificada por el usuario y el objeto continúa hasta tocar el suelo.

Cuando se ajusta automáticamente, es posible utilizar primero una situación en la que se genera el tercer vídeo Z formando un compuesto del primer vídeo X y el segundo vídeo Y (o un vídeo en el cual el segundo vídeo Y se desplaza de acuerdo con los comandos de usuario). En esta situación, cuando se produce la composición, si cada píxel en el tercer vídeo Z se deriva de uno de los objetos y se representa un objeto dado ocultando un objeto diferente, el valor de píxel en la zona oculta del objeto oculto y la información de profundidad se almacenan por separado.

El tercer vídeo Z se corrige mientras se hace referencia a esta información almacenada para que no se produzcan

conflictos de objetos. Al hacer esta corrección, hay momentos en los que cambia la dirección de profundidad del objeto. En este caso, sería correcto hacer referencia a la información de píxeles en el segundo vídeo original Y a fin de preservar en la medida de lo posible la resolución de la apariencia externa del objeto porque el escalamiento se lleva a cabo cuando el objeto se representa dentro del fotograma.

Estas situaciones, cuando se comparan con los métodos descritos anteriormente, simplemente sustituyen el orden del proceso de cálculo como una situación en la que el método de almacenamiento de información difiere, y el principio de acción del mismo puede verse como sustancialmente idéntico.

Además, se considera que la cantidad de conflicto muestra la cantidad de penetración entre los objetos, por lo que es posible lograr la composición de un vídeo al imponer la condición de que "un cierto grado de penetración entre los objetos es bueno".

Incluso si se utilizan los diversos métodos descritos anteriormente, hay ocasiones en que es inevitable un conflicto de objetos. En tales casos, son necesarios esquemas tales como los explicados a continuación.

Cuarta realización preferida

Con las realizaciones preferidas descritas anteriormente, cuando el usuario especificó la cantidad de traslación hacia arriba, abajo, izquierda o derecha para un objeto en el vídeo Y y trató de superponerlo en el vídeo X, si surgía un conflicto de objetos, la explicación fue para una situación para producir una advertencia (segunda realización preferida) y una situación para ajustar automáticamente la cantidad de traslación (tercera realización preferida).

Esta realización preferida se puede aplicar a ambas realizaciones preferidas descritas anteriormente, y muestra un área donde el usuario puede soltar un objeto, es decir, un área que comprende una posición de caída donde se genera un vídeo en el que los objetos no entran en conflicto a través de una operación de soltar por parte del usuario, de una manera fácil de entender antes de la operación de soltar por parte del usuario.

Un área desprendible mostrada al usuario está compuesta por información correspondiente al rango en el que se encuentra un parámetro de transformación que indica la cantidad de desplazamiento en la transformación de desplazamiento del vídeo Y. A través de esto, en el espacio representado en el primer vídeo X, se expresa un área donde el segundo objeto no interfiere con el primer objeto. El rango posible se establece por el corrector 106 que funciona como un graduador de rango.

La Figura 9A es un dibujo explicativo que muestra la condición de la pantalla antes de que se inicie una caída, la Figura 9B es un dibujo explicativo que muestra la condición de la pantalla en el punto en el tiempo en que se inicia una caída, la Figura 9C es un dibujo explicativo que muestra la condición de la pantalla en un punto en el tiempo durante una caída, y la Figura 9D es un dibujo explicativo que muestra la condición de la pantalla en el momento en tiempo en que se realiza la caída. La explicación a continuación hace referencia a estos dibujos.

La Figura 10 es un diagrama de flujo que muestra el flujo de control de un proceso de visualización resaltado para un área desprendible. La explicación a continuación hace referencia a este dibujo.

Primero, una disposición F que tiene un índice candidato para cantidades de desplazamiento arbitrarias utilizables cuando el desplazamiento del segundo vídeo Y en la dirección horizontal y la dirección vertical se prepara en la RAM (etapa S501). A continuación, se supone que [ps, qt] se prepara como el índice, y que ps y qt son valores negativos. Es decir, ps puede ser un número entero entre el valor encontrado al anular el signo de la anchura del segundo vídeo Y y la suma de la anchura del primer vídeo X y la anchura del segundo vídeo Y, y qt puede ser un número entero entre el valor encontrado al negar el signo de la altura del segundo vídeo Y y la suma de la altura del primer vídeo X y la altura del segundo vídeo Y.

Además, para los candidatos [ps, qt] para la cantidad de desplazamiento utilizable para el segundo vídeo Y (etapa S501), se logra una determinación de conflictos para un objeto entre el primer vídeo X y el vídeo intermedio move(ps, qt, Y) que es el segundo vídeo Y desplazado (etapa S502), y se repite un proceso de sustitución de resultados de la determinación de conflictos en los elementos F[ps,qt] de la matriz F (etapa S503) (etapa S504).

A continuación, de forma similar a la realización preferida descrita anteriormente, la ventana 411 del primer vídeo X y la ventana 412 del segundo vídeo Y se muestran en la pantalla (etapa S506) y se recibe el comienzo de un arrastre (etapa S507).

Aquí, el arrastre se inicia cuando el usuario hace clic en una posición [s,t] dentro del segundo vídeo Y usando un ratón.

Suponiendo que la posición [s,t] en el segundo vídeo Y se deja soltar en una posición [p,q] dentro del primer vídeo X, el vídeo intermedio encontrado al desplazar el segundo vídeo Y se convierte en move(p-s, q-t, Y).

Por consiguiente, sería correcto referirse al elemento F[p-s, q-t] de la matriz para determinar si los objetos entran en conflicto entre el primer vídeo X y el vídeo intermedio move[p-s, q-t, Y].

El dispositivo de procesamiento de vídeo 101, para cada posición [p,q] dentro del primer vídeo X (etapa S508), repite (etapa S510) un proceso para hacer que el color (cualquiera de la saturación del color, brillo o matiz, o una combinación de estos) de los píxeles en cada posición [p, q] dentro del primer vídeo X cambie (etapa S509) de acuerdo con el valor de F[p-s, q-t].

A través de esto, el usuario puede distinguir entre las áreas donde las caídas son posibles y las áreas donde las caídas no son posibles, y una de las áreas se muestra resaltada.

Los métodos para llevar a cabo una visualización resaltada incluyen "hacer que el color de los píxeles que se van a resaltar cambie y hacer que el color de otros píxeles no cambie" o "no hacer que el color de los píxeles a resaltar cambie de color y hacer que el color de otros píxeles cambie".

Siguiendo esto en las realizaciones preferidas descritas anteriormente, sería correcto avanzar a la etapa S183.

La explicación a continuación cita un ejemplo. Antes de iniciar el arrastre, el primer vídeo X se representa sin cambios en la ventana 411, como se muestra en la Figura 9A.

En el instante en que se inicia un arrastre, el dispositivo de procesamiento de vídeo calcula el área que se puede soltar. Además, el área desprendible dentro de la ventana 411 se muestra resaltada, como se muestra en la Figura 9B. En este dibujo, el resaltado se ilustra con sombreado.

Si el usuario arrastra el cursor 413 hasta dentro del área desprendible, como se muestra en la Figura 9C, y luego lleva a cabo un proceso de soltar como se muestra en la Figura 9D, se obtiene un tercer vídeo en el que los objetos no entran en conflicto.

Con el método descrito anteriormente, es posible hacer que se mejore la velocidad de experiencia de la visualización resaltada después de un clic, y es posible realizar un redibujado a alta velocidad cuando se vuelve a hacer clic (no se muestra).

Sin embargo, cuando se puede realizar una velocidad de cálculo suficiente, cuando el usuario inicie una operación de arrastre, sería correcto realizar una determinación de conflictos y realizar el resaltado para una posición arbitraria [p,q] dentro del primer vídeo X después haber establecido [s,t], basándose en la cantidad de cambio [p-s, q-t].

De esta manera, cuando el usuario comienza a arrastrar el segundo vídeo Y, es fácil entender dónde se puede realizar una operación de soltar en el primer vídeo X para obtener un nuevo vídeo.

El ajuste automático de la cantidad de traslación en la tercera realización preferida corresponde a corregir la posición de caída dentro del área desprendible cuando se hace una caída fuera del área desprendible.

En esta realización preferida, cuando esta realización preferida se utiliza después de determinar si son posibles caídas para todas las posiciones dentro del fotograma del primer vídeo X, cuando se hace una caída fuera del área desprendible, simplemente corrigiendo la posición de caída en "un punto más cercano a esa posición de caída dentro del área desprendible" o "el punto central del área desprendible" y/o similares, es posible realizar un ajuste automático de la cantidad de traslación hacia arriba, abajo, izquierda o derecha.

Corregir la posición de caída con respecto al límite que delinea el área desprendible y distinta del área desprendible significa que hay un instante en el que un objeto en el primer vídeo y un objeto en el segundo vídeo se tocan mutuamente.

Por consiguiente, es concebible que para hacer que el objeto toque el suelo en la medida de lo posible, sería correcto corregir la posición de caída hasta el límite inferior del área desprendible.

La Figura 9E es un dibujo explicativo que muestra la condición cuando se lleva a cabo la corrección automática de la posición de caída después de una caída. Como se muestra en este dibujo, la posición [s,t] dentro de la ventana 412 se corrige de modo que se superpone en el límite inferior del área desprendible, debajo de la posición [u,v] dentro de la ventana 411.

Con esta realización preferida, es fácil para el usuario entender las posiciones desprendibles, e incluso cuando se realiza una caída en una posición no desprendible, la posición de caída puede corregirse fácilmente usando la información ya calculada.

Con la realización preferida descrita anteriormente, a través de una interfaz de usuario conocida como arrastrar y soltar, el usuario especificó la cantidad de desplazamiento en la dirección horizontal y la dirección vertical, y al mostrar el área desprendible dentro de la pantalla, se mostró un rango donde es posible la cantidad de desplazamiento. A continuación, se explica un método del usuario que especifica cantidades de desplazamiento en la dirección horizontal, la dirección vertical y la dirección de la profundidad.

Con este método, se utiliza una interfaz de usuario de tal forma que, después de que el usuario haya especificado una posición [s,t] dentro de la ventana 412 haciendo clic y ha soltado el botón del ratón (correspondiente a la Figura 9B), después de mover el ratón con el botón del ratón liberado (correspondiente a la Figura 9C), el usuario vuelve a hacer clic en la posición [u,v] dentro de la ventana 411 (correspondiente a la Figura 9D).

Además, si el usuario hace que la rueda del ratón gire durante el intervalo entre hacer clic en la posición [s,t] y hacer clic en la posición [u,v], la cantidad de desplazamiento en la dirección de profundidad cambia.

Con el método descrito anteriormente de soltar y arrastrar, la cantidad de desplazamiento en la dirección de profundidad se establece en 0, pero con el método actual, el valor inicial de la cantidad de desplazamiento en la dirección de profundidad se establece en 0 y la cantidad de cambio aumenta o disminuye según la cantidad de rotación de la rueda.

Si la cantidad de desplazamiento en la dirección de profundidad cambia, el vídeo intermedio que es el segundo vídeo X desplazado en la dirección de profundidad se usa en lugar del segundo vídeo X. Por consiguiente, el tamaño de la ventana 412 y el tamaño del objeto 413 que se muestra en la misma cambia. Además, debido a que la posición del segundo objeto en la dirección de la profundidad cambia, el "área desprendible (área donde está contenida la posición especificada [u,v]" también cambia.

En consecuencia, cuando la rueda del ratón se gira en circunstancias correspondientes a la Figura 9C, la forma del área indicada por las líneas inclinadas (el área en que se encuentran las posiciones [u,v] que se pueden especificar) también cambia.

Cuando se utiliza este tipo de condición, cuando se determina un conflicto entre fotogramas correspondientes al primer vídeo y al segundo vídeo, es imposible un proceso de cálculo de alta velocidad. Esto puede resolverse reduciendo apropiadamente las determinaciones de conflicto.

El método más sencillo de reducción es un método para realizar determinaciones de conflicto solamente entre un fotograma representativo del primer vídeo X presentado en la ventana 411 y un fotograma representativo del segundo vídeo Y presentado en la ventana 412.

Aquí se proporciona un ejemplo de un método en el que, para el primer vídeo Y, el área ocupada por el primer objeto se suma para todos los fotogramas o cada fotograma de los fotogramas reducidos, y para la suma (unión) de las áreas totales a calcular de antemano como el espacio de ocupación del primer objeto. Este ejemplo no está dentro del alcance de la invención reivindicada, y se proporciona solo con fines ilustrativos. En este caso, si el espacio de ocupación se encuentra combinando recuadros delimitadores que comprenden sólidos y esferas rectangulares que indican la forma aproximada de cada primer objeto, es posible hacer el cálculo aún más simple.

Con este método, la determinación del conflictos se realiza de la misma manera descrita anteriormente entre el espacio de ocupación calculado de antemano y el segundo objeto representado en un fotograma representativo del segundo vídeo Y visualizado en la ventana 412. Como resultado, cuando se produce un conflicto, se entiende que la

cantidad de desplazamiento no está contenida en el rango posible.

Por consiguiente, después de la finalización de la especificación de posición dentro de la ventana 412, cuando la cantidad de desplazamiento en la dirección de profundidad cambia como resultado de la operación de la rueda del ratón, en primer lugar si se logra la determinación sencilla descrita anteriormente, es posible indicar al usuario las cantidades de desplazamiento que no se pueden especificar.

Después de esto, se realiza una determinación de conflictos entre la zona de ocupación calculada de antemano a través del procesamiento paralelo y el segundo objeto representado en el fotograma representativo del segundo vídeo Y mostrado en la ventana 412, y se determinan cantidades de desplazamiento que no se pueden especificar. En esta condición, el posible rango representado por líneas inclinadas en la ventana 411 se estrecha gradualmente con el paso del tiempo.

Este método encuentra el espacio de ocupación por adelantado para un vídeo complementario compuesto y realiza una determinación de conflictos con cada fotograma de un vídeo en el lado corregido por la cantidad de desplazamiento, pero sería correcto sustituirlo.

Es decir, hay un método en el que el espacio de ocupación de los objetos representados en el vídeo Y se encuentra por adelantado y se realiza una determinación de conflictos entre este espacio de ocupación y cada fotograma del vídeo X, y a través de esto se encuentra el rango de posibles cantidades de desplazamiento del vídeo Y. Se puede pensar que este método cambia las funciones del primer vídeo y el segundo vídeo.

El método de reducir las determinaciones de conflictos se puede aplicar al método descrito anteriormente arrastrando y soltando también.

Con el método que utiliza el espacio de ocupación, es posible establecer el rango posible de cantidades de desplazamiento de manera que los dos objetos se toquen en el espacio de ocupación. Este método es particularmente útil cuando el primer objeto es un objeto prácticamente inmóvil.

Por ejemplo, es posible utilizar el suelo, la superficie del suelo, la carretera, la vía férrea y/o similares como el primer objeto y llevar a una persona en movimiento sobre la superficie del suelo para que sea el segundo objeto, o tomar un automóvil en movimiento sobre la carretera como el segundo objeto, o un tren en movimiento sobre la vía férrea como el segundo objeto.

Como se ha descrito anteriormente, es posible adquirir los atributos del primer objeto (suelo, superficie del suelo, carretera, vía férrea y/o similares) y los atributos del segundo objeto (persona, automóvil, tren y/o similares) a través de la designación del usuario o a través del reconocimiento de imágenes.

Además, si las condiciones de colocación para el segundo objeto, tal como "tocar el espacio de ocupación" y/o similares, se almacenan en una base de datos como condiciones de colocación para un atributo, es posible establecer automáticamente las condiciones de colocación para el espacio de ocupación simplemente estableciendo los atributos de los objetos en cada vídeo, y es posible establecer las condiciones en las que los objetos se mueven naturalmente.

Quinta realización preferida

Las coordenadas de un punto representativo en el espacio tridimensional para el objeto i representado en el vídeo X pueden indicarse como (xc(X,t,i), yc(X,t,i), zc(X,t,i)), como se ha descrito anteriormente. Esto se puede ver como un vector de posición tridimensional pos[X,i](t) que cambia con el tiempo.

En las realizaciones preferidas descritas anteriormente, cuando el objeto j representado en el vídeo Y se superpone al vídeo X, cuando el objeto j entra en conflicto con un objeto representado en el vídeo X, se obtuvo un vídeo W por traslación del vídeo S en las direcciones arriba, abajo, izquierda o derecha o en la dirección de profundidad, o girando el vídeo Y, o haciendo un cambio de hora, y este vídeo W se superpuso entonces en el vídeo X.

Debido a que la traslación y la rotación y/o similares, son transformaciones congruentes, cuando solo se usan tales transformaciones, el vector de posición pos[Y,j](t) y el vector de posición pos[W,j](t) que son las trayectorias del objeto j, pueden tener diferentes posiciones y orientaciones, pero la forma es la misma. Es decir, se aplica una transformación no deformante que no deforma la trayectoria, por lo que la trayectoria pos[Y,j](t) y la trayectoria

pos[W,j](t) son congruentes.

La Figura 11A es un dibujo explicativo que muestra la condición de la trayectoria pos[X,i](t) y la trayectoria pos[Y,j](t).

La Figura 11B es un dibujo explicativo que muestra la condición de la trayectoria pos[X,i](t) y la trayectoria pos[W,j](t).

En estos dibujos, la posición de cada trayectoria en la posición horizontal se muestra en el eje horizontal y el paso del tiempo se muestra en el eje vertical. Además, a lo largo de cada trayectoria, se forman rectángulos estrechos que indican la extensión de ese objeto en la dirección horizontal.

La trayectoria pos[W,j](t) en la Figura 11B es la trayectoria pos[Y,j](t) en la Figura 11A movida en la dirección horizontal, y las formas de los dos son congruentes.

Si el objeto i en el vídeo X y el objeto j en el vídeo Y entran en conflicto en el tiempo T, como se muestra en la Figura 11A, la trayectoria pos[X,i](t) y la trayectoria pos[Y,j](t) se tocan en el tiempo T, y los rectángulos que representan la extensión de los objetos en la dirección horizontal se superponen.

Por otro lado, no se produce ningún conflicto entre el objeto i del vídeo X y el objeto j del vídeo W. Por consiguiente, como se muestra en la Figura 11B, los rectángulos que representan la extensión de los objetos en la dirección horizontal no se superponen.

Sin embargo, con su tipo de transformación congruente, hay casos en los que los conflictos son inevitables. Por lo tanto, con esta realización preferida, el corrector 106 evita conflictos entre objetos al aplicar sobre el vídeo una transformación de deformación que deforma la trayectoria del objeto.

A continuación, se explica un método en el que la trayectoria pos[X,i](t) del objeto i representado en el vídeo X que tiene un fondo no se deforma, pero la trayectoria pos[Y,j](t) del objeto j representado en el vídeo Y superpuesto en el vídeo X se deforma.

Primero, el vector de posición pos[Y,j](t) que indica que la trayectoria se deforma para convertirse en el vector de posición morph(t).

A continuación, los diversos valores de coordenadas del vector de posición se indican adjuntando .x, .y y .z.

Mediante la deformación de la trayectoria, la profundidad se convierte en morph(t).z/pos[Y,j](t).z veces. Esto significa que en el tiempo t, el objeto j se amplía a pos[Y,j](t).z/morph(t).z veces dentro del fotograma del vídeo Y.

Además, en el tiempo t el objeto j se mueve en la dirección horizontal y la dirección vertical. Las cantidades de estos movimientos son respectivamente (morph(t).x - pos[Y,j](t).x) x pos[Y,j](t).z/morph(t).z, (morph(t).y - pos[Y,j](t).y) x pos[Y,j](t).z/morph(t).z.

Por consiguiente, cuando la trayectoria se deforma, combinando la selección del objeto, la traslación y el escalamiento, es posible obtener un vídeo en el que la trayectoria de un objeto dado se deforme.

Si la deformación de una trayectoria se lleva a cabo para evitar un conflicto entre objetos, al superponer el vídeo con la trayectoria deformada, es posible generar un vídeo en el que los objetos representados en vídeos múltiples se representan en un estado en movimiento sin entrar en conflicto.

A continuación, se describen diversos métodos de deformación de trayectorias.

En un primer método, en un tiempo t en el que se establece overlap(X,Y,t,x,y) >0, un proceso de deformación de la trayectoria del objeto j de manera que la posición pos[Y,j](t) de un punto representativo del objeto j = Y(t)[x,y].id contenido en el vídeo Y se traslada en una dirección alejada de la posición pos[X,i](t) de un punto representativo del objeto i = X(t)[x,y].id contenido en el vídeo X, o en una dirección del vector normal primario a la trayectoria pos[Y,j](t), se repite para que el cambio en el grado de deformación antes y después del tiempo se vuelva más pequeño, y para que overlap(X,Y,t,x,y) = 0 se establezca en todo momento.

Las Figuras 12A a 12D son dibujos explicativos que muestran el estado en el que la trayectoria se deforma

gradualmente repitiendo el proceso. La explicación a continuación hace referencia a estos dibujos.

En la Figura 12A, similar a la Figura 11A, la trayectoria pos[X,i](t) y la trayectoria pos[Y,j](t) se tocan en el tiempo T.

En la Figura 12B, la posición del objeto j en el tiempo T se corrige para que no se produzca ningún conflicto.

Debido a que esta cantidad de corrección era grande, en la Figura 12C, las posiciones del objeto j en los tiempos T-1 y T+1 antes y después del tiempo T se corrigen en una cantidad menor que la cantidad de corrección en la Figura 12B.

Debido a que esta cantidad de corrección era grande, en la Figura 12D, las posiciones del objeto j en los tiempos T-2 y T+2 antes y después de la corrección previa se corrigen adicionalmente en una cantidad menor que la cantidad de corrección en la Figura 12C.

Con esto, la corrección es suficiente y la trayectoria mostrada en la Figura 12D corresponde a la trayectoria morph(t) finalmente obtenida.

La cantidad de corrección propagada a fotogramas adyacentes se reduce elevando la cantidad de corrección para el propio fotograma a una potencia constante no inferior a 0 pero inferior a 1, y si la cantidad de corrección es inferior a un valor umbral prescrito, sería correcto detener la propagación.

Al deformar la trayectoria, sería correcto no propagar las correcciones antes y después del tiempo como se ha descrito anteriormente, sino lograr una corrección uniforme mediante la utilización de interpolación de splines.

Además, con este método, si aún se produce un conflicto después de realizar el tipo de corrección y propagación de la cantidad de corrección que se ha descrito anteriormente en el momento en que la cantidad de conflicto es mayor, es posible hacer la cantidad de corrección y la frecuencia de corrección lo más pequeña posible repitiendo el mismo proceso.

Un segundo método es un método que, al deformar la trayectoria pos[Y,j](t) en el tiempo t, causa solo el tamaño del vector normal primario (un vector ortogonal al vector de velocidad y correspondiente a una dirección que dobla la orientación del vector de velocidad) sin cambiar el tamaño del vector tangente con respecto a la trayectoria (correspondiente al vector de velocidad) y minimiza a 0 collision(Y,X) mientras que también minimiza la suma de los cambios en el tamaño del vector normal primario (típicamente la suma de cuadrados).

La Figura 8 es un dibujo explicativo que muestra el estado en el que se ajusta el vector normal primario en partes específicas de la trayectoria. Este dibujo muestra las formas de la trayectoria sin cambios.

En este dibujo, la trayectoria pox[X,i](t) entra en conflicto con la trayectoria pos[Y,j](t) en el tiempo T. Por lo tanto, en la zona ajustada T-a a T+a antes y después del tiempo T, la curva de la trayectoria pos[Y,j](t) se ajusta de manera que no se produzca un conflicto, y se obtiene una trayectoria postcorrección morph(t).

Las formas de la trayectoria pos[Y,j](t) y la trayectoria morph(t) después del tiempo T+a son congruentes.

En el primer método, el cálculo es sencillo pero la divergencia entre la trayectoria y el movimiento del objeto j que se expresa por el vídeo se vuelve grande, y en este tiempo, hay momentos en los que parece que el objeto j está haciendo repentinamente un "Moonwalk".

Puede determinarse si el objeto j realiza movimientos no naturales según si el tamaño y la orientación del vector tangente con respecto a la trayectoria del objeto j, el tamaño del vector normal primario (este es un valor correspondiente a la curvatura de la trayectoria) y el cambio en la orientación exceden los valores umbral prescritos.

Cuando se determina que el movimiento del objeto no es natural debido a la deformación de la trayectoria, si se emite una advertencia a tal efecto, la comodidad del usuario es alta.

En la repetición, del proceso de deformación de la trayectoria, sería correcto establecer de antemano un límite superior a la cantidad de deformación de la trayectoria. En este caso, el movimiento del objeto no se vuelve antinatural, pero existe la posibilidad de que no sea posible evitar los conflictos.

Con el segundo método, los cálculos son más complejos, pero la dirección de movimiento del objeto j cambia solo ligeramente y el tamaño de la velocidad de movimiento no cambia, por lo que no se produce un "moonwalk".

Cuando se representa un fondo en el vídeo X junto con el objeto i, cuando se provoca que la trayectoria del objeto i cambie, hay casos en los que se producen huecos sin información de píxeles entre el fondo y el objeto i. Por lo tanto, con esta realización preferida, solamente se deforma la trayectoria del objeto j representado en el vídeo Y.

Sin embargo, cuando el fondo no está contenido dentro del vídeo X, cuando el objeto i y el objeto j entran en conflicto, sería correcto deformar ambas trayectorias. Con el primer método, sería correcto hacer que el objeto i y el objeto j se muevan en direcciones que se alejan mutuamente, y con el segundo método, sería correcto minimizar tanto la cantidad de cambio en el vector normal primario de la trayectoria del objeto i como la cantidad de cambio en el vector normal primario de la trayectoria del objeto j.

En estos métodos, es posible evitar el movimiento antinatural del objeto y aumentar la comodidad del usuario estableciendo un límite superior en la cantidad de cambio en el vector normal primario, o producir una advertencia cuando la cantidad de cambio en el vector normal primario excede un valor umbral prescrito.

En esta configuración, incluso cuando un conflicto es inevitable con una simple traslación, rotación, escalamiento y similares de los objetos representados en los vídeos, es posible evitar un conflicto deformando la trayectoria.

En esta realización preferida y las realizaciones preferidas descritas anteriormente, se describen métodos para evitar conflictos de objetos. Sin embargo, también es posible imponer la condición limitante de que el contacto entre las superficies de los objetos se produzca en un punto deseado en el tiempo.

Es decir, al imponer la restricción de que en el punto de tiempo deseado hay superposición en la zona de ocupación desde la profundidad de superficie anterior hasta la profundidad de superficie posterior de los dos objetos en una posición donde los dos objetos se representan superponiéndose, y la superposición es 0, el movimiento del objeto y la deformación de las trayectorias se logra de manera que la restricción se cumple.

Además, sería correcto causar la deformación de ambas trayectorias de los objetos i y j representados en ambos vídeos X e Y. En este caso, sería correcto lograr la deformación para minimizar la cantidad de conflicto compilando y tratando los parámetros de deformación para el objeto i y los parámetros de deformación para el objeto j.

Sexta realización preferida

Cuando el objeto j representado en el vídeo Y es un objeto que baila en una superficie plana y se incluye un suelo en el fondo en el vídeo X, es deseable que el objeto j se superponga para bailar sobre el suelo del vídeo X. Esta realización preferida se da cuenta de esto.

Con la realización preferida descrita anteriormente, la explicación fue para la condición en la que coinciden la inclinación de una cámara con respecto al suelo en el vídeo X y la inclinación de la cámara con respecto a un plano en el vídeo Y, pero en esta realización preferida, este tipo de coincidencia no se requiere necesariamente.

En primer lugar, como un punto representativo del objeto j representado en el vídeo Y, se selecciona un punto que se muestra más abajo en la pantalla. Cuando el objeto j es un personaje que baila, el punto representativo seleccionado de esta manera corresponde a los pies del personaje.

La posición de píxel [xb(Y,t,j), yb(Y,t,j)] del punto representativo puede calcularse, por ejemplo, de la siguiente manera:

$$yb(Y,t,j) = \min_{x,y;Y(t)[x,y].id=j} y$$

$$xb(Y,t,j) = \min_{x,y;Y(t)[x,yb(Y,t,j)].id=j} x.$$

Aquí, si el punto más bajo del objeto j hace contacto con el suelo del vídeo X, entonces:

$$X(t)[xb(Y,t,j), xb(Y,t,j)].id = 0;$$

$$Y(t)[xb(Y,t,j), xb(Y,t,j)].id = j;$$

$$X(t)[xb(Y,t,j), xb(Y,t,j)].fore = Y(t)[xb(Y,t,j), xb(Y,t,j)].fore$$

debería establecerse.

Por consiguiente, si la cantidad de traslación en las direcciones arriba, abajo, izquierda y derecha y también la dirección de profundidad del vídeo Y y la cantidad de rotación en torno a los diversos ejes se determina de manera que, además de minimizar collision(Y,X) a 0, la suma de las diferencias (típicamente, suma de cuadrados) de X(t)[xb(Y,t,j), xb(Y,t,j)].fore e Y(t)[xb(Y,t,j), xb(Y,t,j)].fore también se minimiza, el objeto j representado en el vídeo Y toca la superficie del piso representada en el vídeo X.

Cuando el objeto j salta del suelo en el medio de un baile, es imposible que X(t)[xb(Y,t,j), xb(Y,t,j)].fore e Y(t)[xb(Y,t,j), xb(Y,t,j)].fore coincidan perfectamente, pero al lograr la minimización, es posible evitar la condición de separarse del piso en instantes en los que no salta.

Además, a través de la minimización, es posible hacer que el comportamiento del objeto j se superponga al vídeo X natural en la medida de lo posible haciendo que el suelo sobre el que está el objeto j en el vídeo Y y el suelo en el vídeo X coincidan en la medida de lo posible.

Séptima realización preferida

Con la realización preferida descrita anteriormente, en las diversas transformaciones del vídeo de 2,5 dimensiones, se realizó un proceso de cálculo independiente para cada posición y cada fragmento de información de profundidad dentro de un único objeto, y se hicieron cálculos para reflejar esto en la medida de lo posible. En esta realización preferida, la carga de cálculo se reduce adicionalmente al realizar aproximaciones usando puntos representativos de objetos.

Es decir, en esta realización preferida, se calculan las coordenadas (xc(X,t,i), yc(X,t,i)) en el vídeo X del punto representativo del objeto i en el tiempo t en el vídeo X y la coordenada de profundidad zc(X,t,i) del punto representativo.

El objeto i se mueve como un grupo en su totalidad, por lo que las coordenadas de profundidad de las diversas partes del objeto i se pueden aproximar mediante zc(X,t,i). Por consiguiente, cuando la profundidad del punto representativo cambia de zc(X,t,i) a zc(X,t,i)+r, la profundidad en cada posición se aproxima a [zc(X,t,i)+r / zc(X,t,i)] veces. Este tipo de aproximación es eficaz para las traslaciones en el método de profundidad, y las rotaciones en torno al eje horizontal o al eje vertical, del vídeo.

Específicamente, el destino del movimiento de solo el punto representativo del objeto i que es el objetivo de traslación o rotación en torno al eje horizontal o al eje vertical se encuentra a través de los mismos cálculos que en las realizaciones preferidas descritas anteriormente. Además, escalando y posicionando el vídeo de acuerdo con los cambios en la profundidad del punto representativo centrado en este destino de movimiento, el cálculo descrito anteriormente se simplifica.

Con la transformación en esta realización preferida, se supone que la forma del objeto i es un objeto en el que se representa una imagen en la superficie de una lámina perpendicular a la dirección de profundidad, e incluso cuando se realiza una traslación en el método de profundidad o una rotación en torno al eje horizontal o el eje vertical, esto corresponde al movimiento de la lámina perpendicularmente a la dirección de la profundidad.

Con una traslación en la dirección de profundidad, la lámina permanece perpendicular a la dirección de profundidad. En esta realización preferida, al establecer un límite superior al ángulo de rotación en torno al eje horizontal o al eje vertical, se hace que la posición de la lámina cambie a través de la rotación con un ángulo o una rotación de hasta un cierto número de grados, pero se piensa que la orientación de la lámina permanece perpendicular a la dirección de la profundidad.

A continuación, se describe el cálculo de transformación específico.

El vídeo obtenido moviendo el objeto i representado en el vídeo X por r en la dirección de profundidad se obtiene realizando los siguientes procesos.

(a) Solo el objeto i se selecciona del vídeo X,
(b) se realiza una traslación a lo largo del fotograma con (xc(X,t,i), yc(X,t,i)) como el origen,
(c) el vídeo se escala a zc(X,t,i) / [zc(X,t,i)+r] veces, y
(d) se hace una traslación a lo largo del fotograma de manera que el origen se convierta en (xc(X,t,i), yc(X,t,i)).

Por consiguiente, cuando la transformación de mover el objeto i representado en el vídeo X por r en la dirección de profundidad se indica como push'(i,r,X), el resultado es:

$$\text{push}'(i,r,X) = \text{move}(xc(X,t,i), yc(X,t,i), \text{scale}\ (zc(X,t,i)/\ [zc(X,t,i)+r],$$

$$\text{move}(-xc(X,t,i), -yc(X,t,i), \text{select}(i,X)))).$$

Además, cuando el objeto i representado en el vídeo X gira en θ en torno al eje horizontal, las coordenadas en el vídeo del punto representativo del objeto se mueven desde (xc(X,t,i), yc(X,t,i)) a (xc(X,t,i), xc(X,t,i) x cos θ - yc(X,t,i) x sen θ), y la profundidad se mueve desde zc(X,t,i) a zc(X,t,i) x sen θ + yc(X,t,i) x cos θ. Es decir, sería correcto realizar los siguientes procesos.

(a) Seleccionar solo el objeto i del vídeo X,
(b) mover (xc(X,t,i), yc(X,t,i)) a (xc(X,t,i), xc(X,t,i) x cos θ - yc(X,t,i) x sen θ) a lo largo del fotograma, y
(c) escalar el vídeo a zc(X,t,i)/ [zc(X,t,i) x en θ + yc(X,t,i) x cos θ].

Por consiguiente, cuando la transformación de la rotación del objeto i representado en el vídeo X por θ en torno al eje horizontal se indica como rothor'(i, θ, X), entonces:

$$\text{rothor}'(i, \theta, X) = \text{scale}(zc(X,t,i)/\ [zc(X,t,i) \times \sin\theta + yc(X,t,i) \times \cos\theta], \text{move}\ (0,$$

$$zc(X,t,i) \times \cos\theta - yc(X,t,i) \times \sin\theta - yc(X,t,i), \text{select}\ (i,X))).$$

Una transformación rotver'(i, θ, X) para girar el objeto i representado en el vídeo X por φ en torno al eje vertical se define de manera similar. Es decir,

$$\text{rothor}'(i, \phi, X) = \text{scale}(zc(X,t,i)/\ [zc(X,t,i) \times \sin\theta + xc(X,t,i) \times \cos\theta], \text{move}$$

$$(zc(X,t,i) \times \cos\theta - xc(X,t,i) \times \sin\theta - xc(X,t,i), 0\ \text{select}\ (i,X))).$$

Con esta realización preferida, una transformación de un único objeto representado en el vídeo se expresa a través de una combinación de los cálculos de movimiento, selección y escalamiento que tienen una carga de cálculo ligera, por lo que es posible acelerar el procesamiento.

Octava realización preferida

Con la realización preferida descrita anteriormente, el tiempo del vídeo X experimentó la transformación de desplazamiento shift(d,X) por d, pero es posible concebir una transformación en la cual el vídeo X se repite lentamente o se avanza rápido.

Si el vídeo Y es el vídeo X en avanzado rápido en a veces, entonces:

$$Y(t)[x,y] = X(t/a)[x,y]$$

se establece. A continuación, esta transformación se indica como fastforward(a,X).

Debido a que el objeto i representado en el vídeo X y el objeto j representado en el vídeo Y repiten ambos una acción dada, cuando el deseo es componer un vídeo Z de tal forma que el objeto i y el objeto j aparecen alternativamente en una ubicación designada dentro del vídeo, es posible usar una combinación de desplazamiento y avance rápido.

Por ejemplo, se considera el caso en el que el deseo es formar un compuesto del vídeo X, en el que una persona

(objeto i) que está haciendo la acción de moler torta de arroz en un mortero, y el vídeo Y, en el que una persona (objeto j) está haciendo la acción de amasar a mano torta de arroz en el mismo mortero, y generar un vídeo que expresa el estado de la persona A y la persona B que hacen conjuntamente arroz.

Cuando dentro y alrededor del mortero, y la posición y orientación de la cámara que realizó la filmación, son comunes en el vídeo X y el vídeo Y, sería correcto encontrar parámetros para el desplazamiento y el avance rápido para minimizar la función de evaluación que expresa la cantidad de conflictos.

Además, cuando se desea determinar a partir del vídeo si es posible colocar muebles y productos electrónicos dentro de una casa, es posible aplicar este tipo de transformación.

Por ejemplo, cuando se desea determinar a partir del vídeo si es posible colocar una lavadora cuya puerta se abre y se cierra dentro de un cuarto de lavado de una casa, sería correcto realizar el siguiente tipo de proceso.

Es decir, un fabricante de productos eléctricos o un proveedor de productos eléctricos proporciona en la web un vídeo X en el que la situación de la puerta de la lavadora que se abre y se cierra periódicamente se filma desde la parte delantera.

Por otro lado, un usuario que considera la compra de la lavadora prepara un vídeo Y en el que se filma desde el frente la situación de una puerta con respecto a la habitación en la que se instalaría la lavadora abriendo y cerrando.

Después de lograr la normalización de manera que la distancia desde la cámara a la lavadora en el vídeo X y la distancia desde la cámara a la ubicación candidata para la instalación de la lavadora en el vídeo Y coincidan sustancialmente, el usuario arrastra y suelta el objeto de la lavadora en el vídeo X a la ubicación candidata de instalación en el vídeo Y.

Además, se encuentra la cantidad de conflicto de los objetos en los vídeos X e Y. Si no hay conflicto, es posible estimar que la lavadora se puede instalar en la ubicación deseada. Además, incluso cuando hay un conflicto, si es posible encontrar parámetros para el desplazamiento y el avance rápido a fin de minimizar la cantidad de conflictos y hacer que no haya conflictos, naturalmente es posible estimar que la lavadora se puede instalar en la ubicación deseada.

Cuando el usuario filma el interior de la habitación, hay momentos en los que la dirección de filmación es limitada. En tal caso, sería correcto proporcionar vídeos en los que la lavadora cuya puerta se abre y se cierra periódicamente se filma desde varias direcciones, tales como situaciones en las que la filmación es desde el frente y también situaciones en las que la filmación es desde el lado izquierdo y situaciones en las que la filmación es desde el lado derecho, y para que el usuario seleccione la filmación de vídeo desde la misma dirección en la que el usuario filmó la habitación.

Novena realización preferida

En las realizaciones preferidas descritas anteriormente, es posible componer un vídeo de modo que no se produzcan conflictos en los que los objetos penetren mutuamente entre sí, y también es posible formar un compuesto de vídeos de tal forma que las superficies de los objetos se toquen en un instante deseado o la parte inferior de un objeto toque el suelo. A continuación, se describen aplicaciones específicas de las realizaciones preferidas descritas anteriormente.

El primer ejemplo de aplicación es permitir a los expositores preparar fácilmente vídeos que muestren la situación de los productos en mercados electrónicos tales como subastas de Internet y/o similares.

Es decir, el operador de un mercado electrónico proporciona materiales de vídeo para que los usuarios puedan usarlos libremente, además de proporcionar herramientas de edición para diversos tipos de vídeos. Este material de vídeo puede prepararse por expositores o postores en el mercado electrónico, o un sistema de distribución de vídeo puede prepararse por separado para que los vídeos públicamente utilizados sean cargados y puedan ser examinados por otros usuarios, y los vídeos acumulados en este sistema de distribución de vídeos pueden ser reutilizados.

Sería correcto para los expositores filmar con anticipación vídeos que expresen fondos, tal como en las habitaciones de una casa y/o similares, y al hacerlo, preparar una mesa decorada o una pantalla plegable y/o similares para que

el producto se pueda presentar de forma atractiva, y también sería correcto que los productos en la habitación se mantengan ordenados y en orden, y se limpien, para que no se pueda interpretar información privada. Además, el vídeo de fondo puede seleccionarse del material de vídeo.

A continuación, cuando el expositor desea exhibir un nuevo producto, es posible realizar la filmación en una ubicación preferida dentro de la habitación sin tener que ordenar y organizar los productos o realizar la limpieza.

A continuación, el expositor arrastra y suelta el vídeo del nuevo producto que se filmó en la toma de vídeo de fondo por adelantado. Con lo cual, es posible componer un vídeo que muestra una situación en la cual el producto no entra en conflicto con los productos (la pantalla plegable y/o similares) de la imagen de fondo y toca el suelo (la superficie de la mesa decorada y/o similares) de la imagen de fondo, a través de evitar conflictos descritos anteriormente y el ajuste automático de posición.

La composición del vídeo del producto en el vídeo de fondo puede ser hecha tanto por postores como por expositores. Por ejemplo, si se exhibe un vestido de fiesta, se puede proporcionar un vídeo del producto que muestre una situación en la que la modelo del expositor y/o similar, esté usando ese vestido, caminando y dando la vuelta.

El postor podría filmar por su cuenta un vídeo de fondo filmando el interior de un salón de actos donde se hay una fiesta a la que desean asistir con el vestido, o podría seleccionarse de entre los materiales de vídeo. Además, el postor forma un compuesto de un vídeo de producto del vestido usado por la modelo, en el vídeo de fondo del salón de fiestas.

Al aplicar la realización preferida descrita anteriormente, la modelo se mueve sobre el suelo, pero el vídeo se compone para que no haya conflictos con otras personas, muebles o instalaciones dentro del salón.

Al examinar este vídeo, el postor puede verificar con anticipación si el vestido que se exhibe como producto coincide con la atmósfera del salón de fiestas, antes de pujar.

Un segundo ejemplo de aplicación utiliza además los materiales de vídeo.

Por ejemplo, si se proporciona un vídeo de fondo que muestra una situación en la que un artista famoso está bailando música como material de vídeo, un usuario puede componer un vídeo que muestra una situación en la que el usuario está bailando junto con el cantante famoso, filmando y componiendo un vídeo del usuario bailando.

Además, cuando se proporciona un vídeo de fondo que muestra una situación en la que un carrusel se mueve con asientos vacíos como material de vídeo, es posible componer un vídeo que muestra una situación con el usuario que monta en el carrusel, filmando un vídeo del usuario que monta una bicicleta, un vagón y/o similar, u otro vehículo que se mueve sustancialmente a lo largo de la misma trayectoria que el tiovivo y realiza la composición del vídeo mientras impone la condición restrictiva de que "el usuario está tocando continuamente el tiovivo".

Además, cuando un usuario forma un compuesto de un vídeo balanceando su brazo o una raqueta con fuerza hacia abajo con un vídeo de fondo en el que una pelota golpea ferozmente el suelo, si se impone la condición restrictiva de que la mano del usuario o la raqueta y la pelota hacen contacto en el instante en que el usuario balancea su brazo hacia abajo, es posible generar un vídeo que muestre al usuario golpeando una pelota de voleibol o golpeando una pelota de tenis.

Además, se pueden componer múltiples materiales de vídeo filmando a una sola persona practicando esgrima o kendo, y si se impone la condición de que los objetos entren en contacto con el tiempo en el instante en que el jugador que se filmó realiza un movimiento decisivo, se puede componer un vídeo que parece como si un partido estuviera en curso.

Aplicabilidad industrial

Con la presente descripción, es posible proporcionar un dispositivo de procesamiento de vídeo, un método de procesamiento de vídeo, y un medio de registro de información adecuados para determinar si los objetos entran en conflicto cuando se compilan objetos acompañados de información de profundidad representada en múltiples vídeos en un solo vídeo.

Lista de signos de referencia

11	Punto de filmación
12	Superficie de proyección
13	Radio
14	Objeto
15	Punto de conflicto
17	Distancia Z
21	Punto representativo
101	Dispositivo de procesamiento de vídeo
102	Adquirente de profundidad de superficie posterior
103	Determinador de conflictos
104	Graduador de cantidad de cambio
105	Generador de vídeo
106	Corrector
401	Pantalla
411	Ventana
412	Ventana
413	Cursor

REIVINDICACIONES

1. Un dispositivo de procesamiento de vídeo (101) para determinar, cuando se forma un compuesto de un primer vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes y un segundo vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes para generar un tercer vídeo que comprende una pluralidad de fotogramas de imágenes, interferencia entre un primer objeto (14) representado en el primer vídeo 2,5D y un segundo objeto (14) representado en el segundo vídeo 2,5D, en el que el primer vídeo 2,5D va acompañado de una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie anterior del primer objeto (14) representado en el primer vídeo 2,5D, y el segundo vídeo 2,5D va acompañado de una segunda distancia de profundidad, siendo la segunda distancia de profundidad una distancia a la superficie anterior del segundo el objeto (14) representado en el segundo vídeo 2,5D, comprendiendo el dispositivo de procesamiento de vídeo (101):

un adquiriente de profundidad de superficie posterior (102) para adquirir una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie posterior del primer objeto (14), y una segunda distancia de profundidad, estando la segunda distancia de profundidad a una distancia con respecto a la superficie posterior del segundo objeto (14); y
un determinador de interferencia para determinar, si una primera área de dibujo donde el primer objeto (14) debe representarse en un fotograma dado del tercer vídeo y una segunda área de dibujo donde el segundo objeto (14) debe representarse en el fotograma dado de la tercera superposición de vídeo, si el primer objeto y el segundo objeto (14) interfieren, basándose en la superposición de una primera zona de ocupación desde la primera profundidad de la superficie anterior hasta la primera profundidad de la superficie posterior del primer objeto (14) en una área de superposición donde se superponen la primera área de dibujo y la segunda área de dibujo, y una segunda zona de ocupación desde la segunda profundidad de superficie anterior hasta la segunda profundidad de superficie posterior del segundo objeto (14) en el área de superposición donde la primera área de dibujo y la segunda área de dibujo se superponen.

2. El dispositivo de procesamiento de vídeo (101) de acuerdo con la reivindicación 1, en el que el adquiriente de profundidad de superficie posterior (102) adquiere la primera profundidad de superficie posterior y la segunda profundidad de superficie posterior en base a los atributos que un usuario ajusta para el primer objeto (14) y el segundo objeto (14).

3. El dispositivo de procesamiento de vídeo (101) de acuerdo con la reivindicación 1, en el que el adquiriente de profundidad de superficie posterior (102) adquiere la primera profundidad de superficie posterior del primer objeto (14) en un primer fotograma del primer vídeo 2,5D estimando la longitud de la primera zona de ocupación desde un primer tamaño promedio de la primera área de dibujo en el primer vídeo 2,5D y un tamaño de área de la primera área de dibujo en el primer fotograma del primer vídeo 2,5D.

4. El dispositivo de procesamiento de vídeo (101) de acuerdo con la reivindicación 1, en el que la diferencia entre la primera profundidad de superficie posterior y la primera profundidad de superficie anterior, y la diferencia entre la segunda profundidad de superficie posterior y la segunda profundidad de superficie anterior, son constantes no negativas.

5. El dispositivo de procesamiento de vídeo (101) de acuerdo con la reivindicación 1, en el que el adquiriente adquiere una primera longitud de profundidad asociada por adelantado con un atributo del primer objeto (14), de una base de datos, y adquiere una segunda longitud de profundidad asociada por adelantado con un atributo del segundo objeto (14), de la base de datos, y adquiere la primera profundidad de superficie posterior desde la primera profundidad de superficie anterior y la primera longitud de profundidad, y adquiere la segunda profundidad de superficie posterior desde la segunda profundidad de superficie anterior y la segunda longitud de profundidad.

6. Un método de procesamiento de vídeo para determinar, cuando se forma un compuesto de un primer vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes y un segundo vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes para generar un tercer vídeo que comprende una pluralidad de fotogramas de imágenes, interferencia entre un primer objeto (14) representado en el primer vídeo 2,5D y un segundo objeto (14) representado en el segundo vídeo 2,5D, en el que el primer vídeo 2,5D va acompañado de una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie anterior del primer objeto (14) representado en el primer vídeo 2,5D, y el segundo vídeo 2,5D va acompañado de una segunda distancia de

profundidad, siendo la segunda distancia de profundidad una distancia a la superficie anterior del segundo el objeto (14) representado en el segundo vídeo 2,5D, comprendiendo el método de procesamiento de vídeo:

un procedimiento de adquisición de profundidad de superficie posterior para adquirir una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie posterior del primer objeto (14), y una segunda distancia de profundidad, estando la segunda distancia de profundidad a una distancia con respecto a la superficie posterior del segundo objeto (14); y
un procedimiento de determinación de interferencia para determinar, si una primera área de dibujo donde el primer objeto (14) debe representarse en un fotograma dado del tercer vídeo y una segunda área de dibujo donde el segundo objeto (14) debe representarse en el fotograma dado de la tercera superposición de vídeo, si el primer objeto (14) y el segundo objeto (14) interfieren, basándose en la superposición de una primera zona de ocupación desde la primera profundidad de la superficie anterior hasta la primera profundidad de la superficie posterior del primer objeto (14) en una área de superposición donde se superponen la primera área de dibujo y la segunda área de dibujo, y una segunda zona de ocupación desde la segunda profundidad de superficie anterior hasta la segunda profundidad de superficie posterior del segundo objeto (14) en el área de superposición donde la primera área de dibujo y la segunda área de dibujo se superponen.

7. Un medio de registro de información legible por ordenador no transitorio en el que se registra un programa para determinar, cuando se forma un compuesto de un primer vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes y un segundo vídeo 2,5D que comprende una pluralidad de fotogramas de imágenes para generar un tercer vídeo que comprende una pluralidad de fotogramas de imágenes, interferencia entre un primer objeto (14) representado en el primer vídeo 2,5D y un segundo objeto (14) representado en el segundo vídeo 2,5D, en el que el primer vídeo 2,5D va acompañado de una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie anterior del primer objeto (14) representado en el primer vídeo 2,5D, y el segundo vídeo 2,5D va acompañado de una segunda distancia de profundidad, siendo la segunda distancia de profundidad una distancia a la superficie anterior del segundo el objeto (14) representado en el segundo vídeo 2,5D, haciendo el programa que un ordenador funcione como:

un adquiriente de profundidad de superficie posterior (102) para adquirir una primera distancia de profundidad, siendo la primera distancia de profundidad una distancia a la superficie posterior del primer objeto (14), y una segunda distancia de profundidad, estando la segunda distancia de profundidad a una distancia con respecto a la superficie posterior del segundo objeto (14); y
un determinador de interferencia para determinar, si una primera área de dibujo donde el primer objeto (14) debe representarse en un fotograma dado del tercer vídeo y una segunda área de dibujo donde el segundo objeto (14) debe representarse en el fotograma dado de la tercera superposición de vídeo, si el primer objeto (14) y el segundo objeto (14) interfieren, basándose en la superposición de una primera zona de ocupación desde la primera profundidad de la superficie anterior hasta la primera profundidad de la superficie posterior del primer objeto (14) en una área de superposición donde se superponen la primera área de dibujo y la segunda área de dibujo, y una segunda zona de ocupación desde la segunda profundidad de superficie anterior hasta la segunda profundidad de superficie posterior del segundo objeto (14) en el área de superposición donde la primera área de dibujo y la segunda área de dibujo se superponen.

FIG. 1

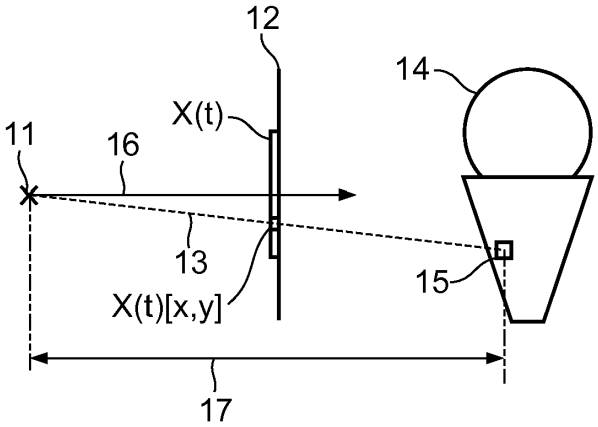

FIG. 2A

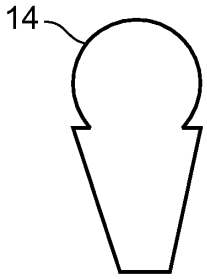

FIG. 2B

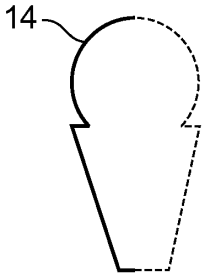

FIG. 2C

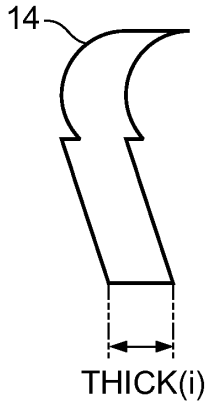

FIG. 2D

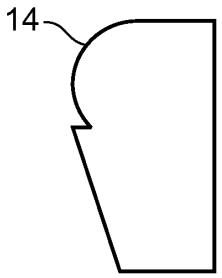

FIG. 2E

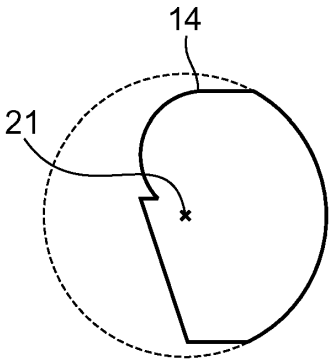

FIG. 3

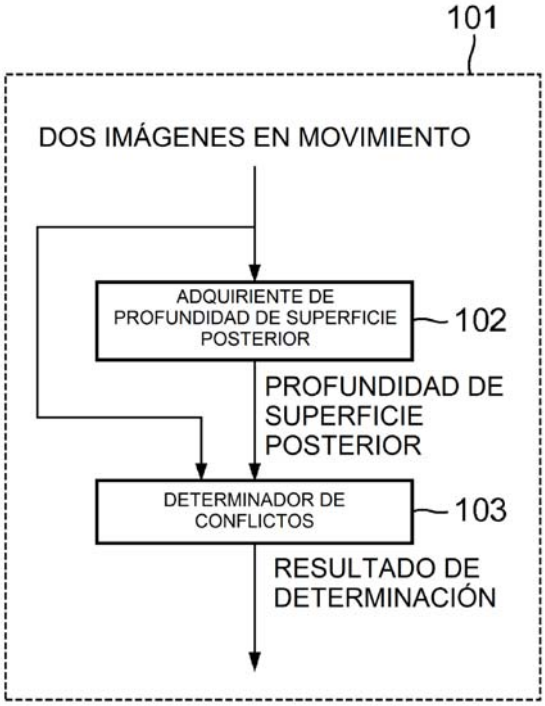
101
DOS IMÁGENES EN MOVIMIENTO
ADQUIRIENTE DE PROFUNDIDAD DE SUPERFICIE POSTERIOR
102
PROFUNDIDAD DE SUPERFICIE POSTERIOR
DETERMINADOR DE CONFLICTOS
103
RESULTADO DE DETERMINACIÓN

FIG. 4

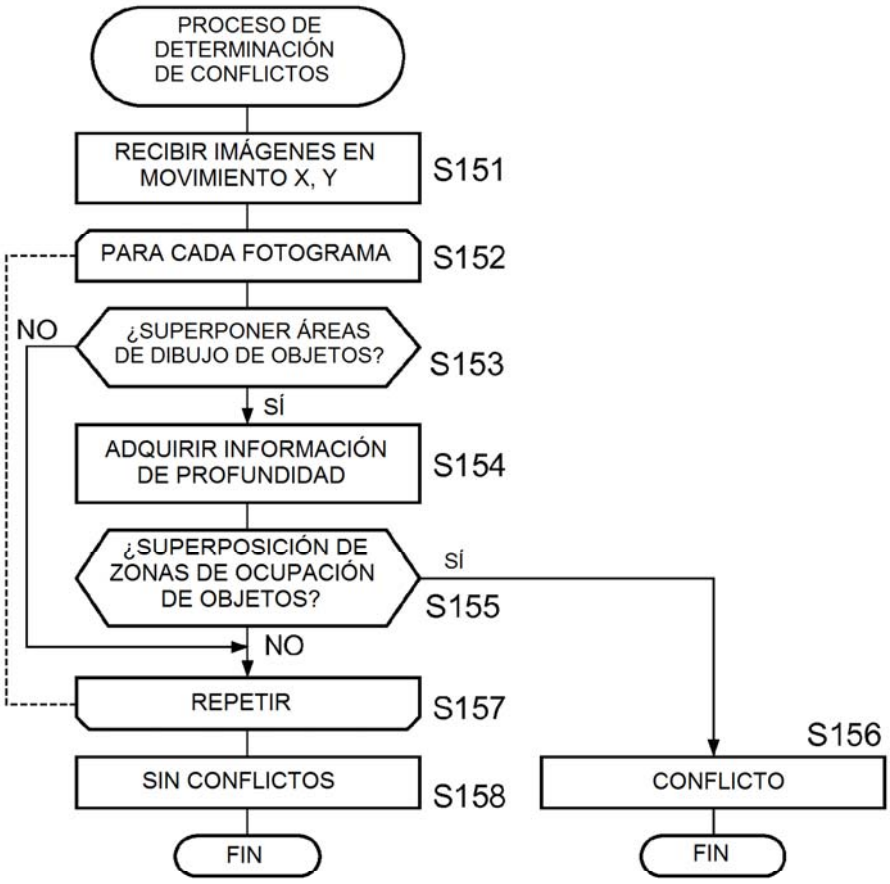
PROCESO DE DETERMINACIÓN DE CONFLICTOS
RECIBIR IMÁGENES EN MOVIMIENTO X, Y
S151
PARA CADA FOTOGRAMA
S152
NO
¿SUPERPONER ÁREAS DE DIBUJO DE OBJETOS?
S153
SÍ
ADQUIRIR INFORMACIÓN DE PROFUNDIDAD
S154
¿SUPERPOSICIÓN DE ZONAS DE OCUPACIÓN DE OBJETOS?
SÍ
S155
NO
REPETIR
S157
S156
SIN CONFLICTOS
S158
CONFLICTO
FIN
FIN

FIG. 5A

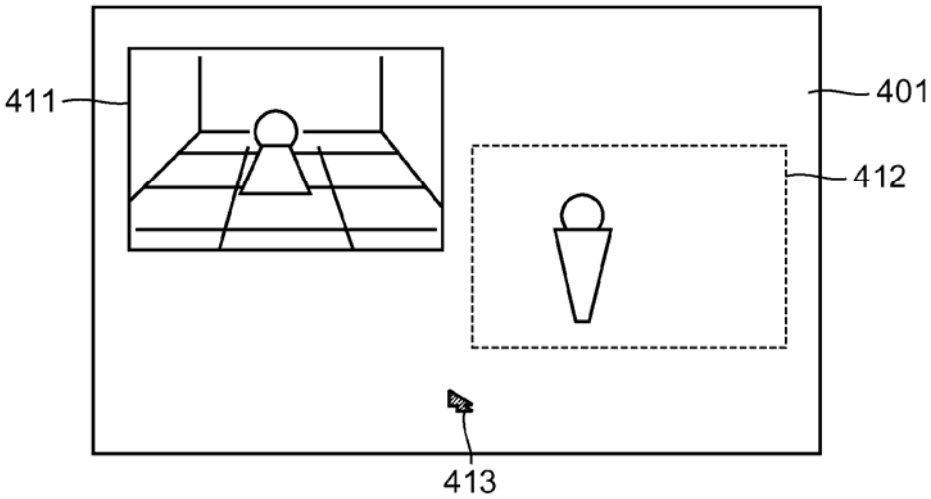
411
401
412
413

FIG. 5B

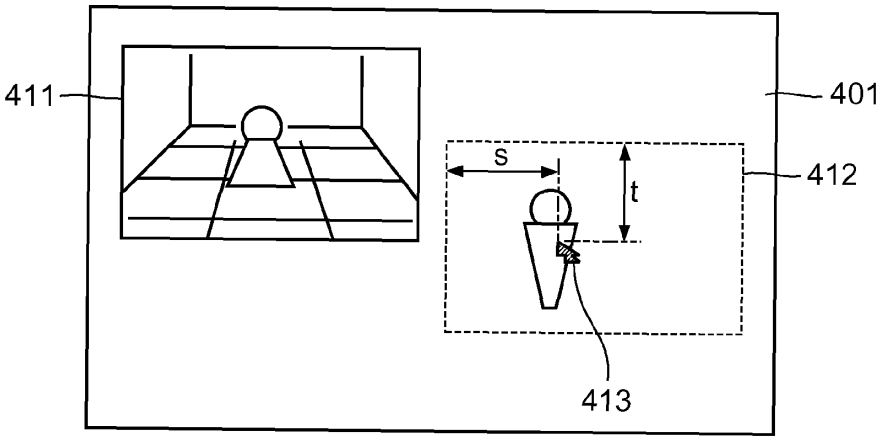

FIG. 5C

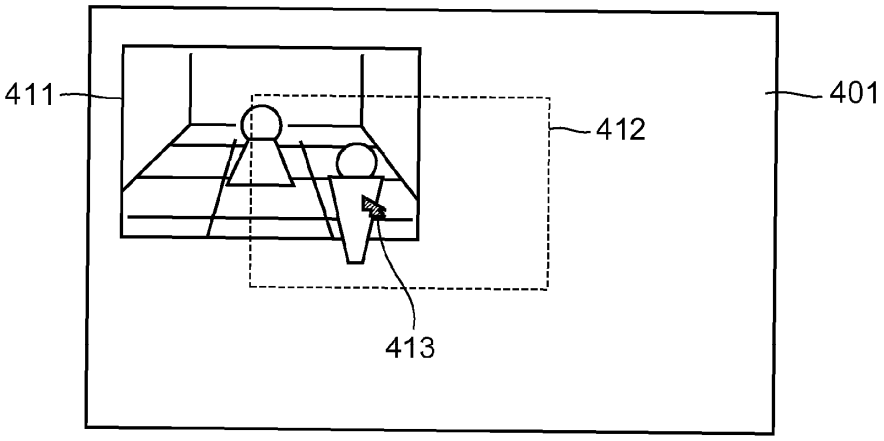

FIG. 5D

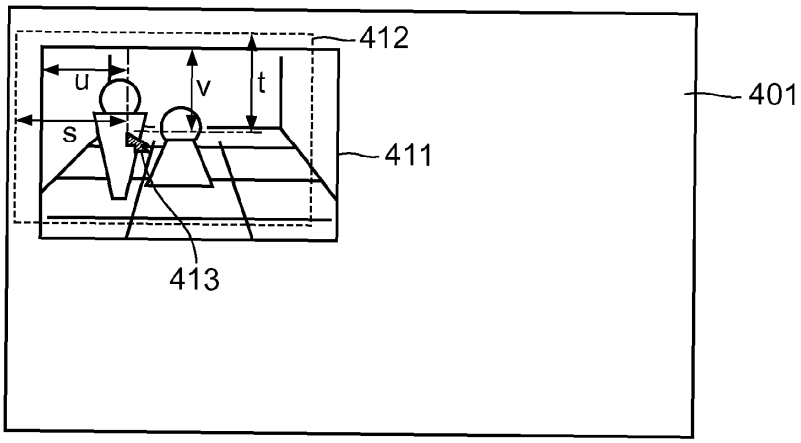

FIG. 6

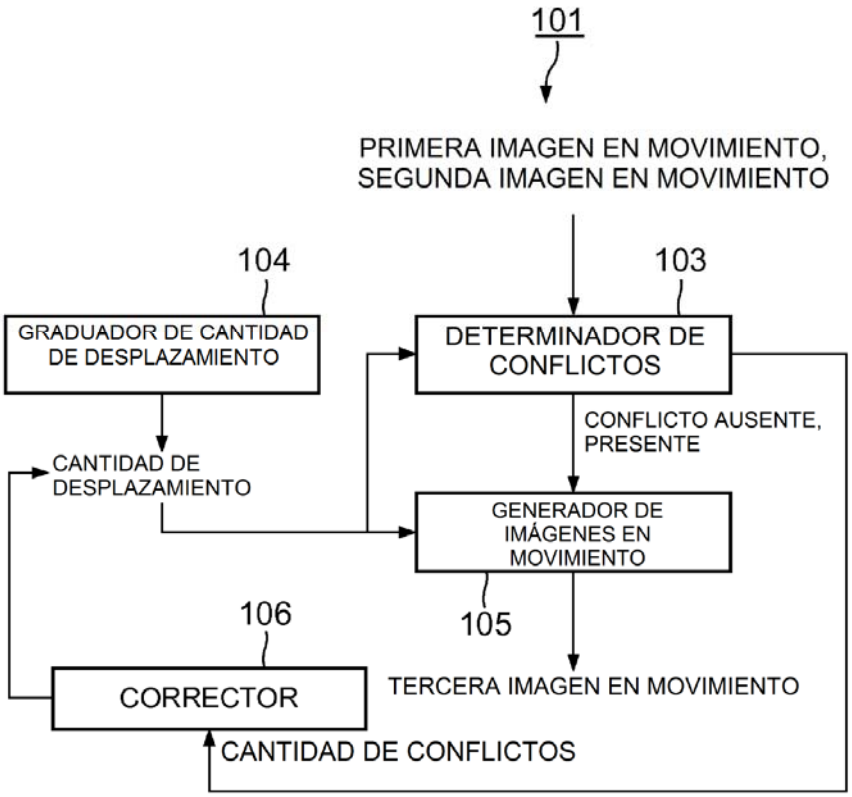
101
PRIMERA IMAGEN EN MOVIMIENTO,
SEGUNDA IMAGEN EN MOVIMIENTO
104
GRADUADOR DE CANTIDAD DE DESPLAZAMIENTO
103
DETERMINADOR DE CONFLICTOS
CONFLICTO AUSENTE, PRESENTE
CANTIDAD DE DESPLAZAMIENTO
GENERADOR DE IMÁGENES EN MOVIMIENTO
106
105
CORRECTOR
TERCERA IMAGEN EN MOVIMIENTO
CANTIDAD DE CONFLICTOS

FIG. 7

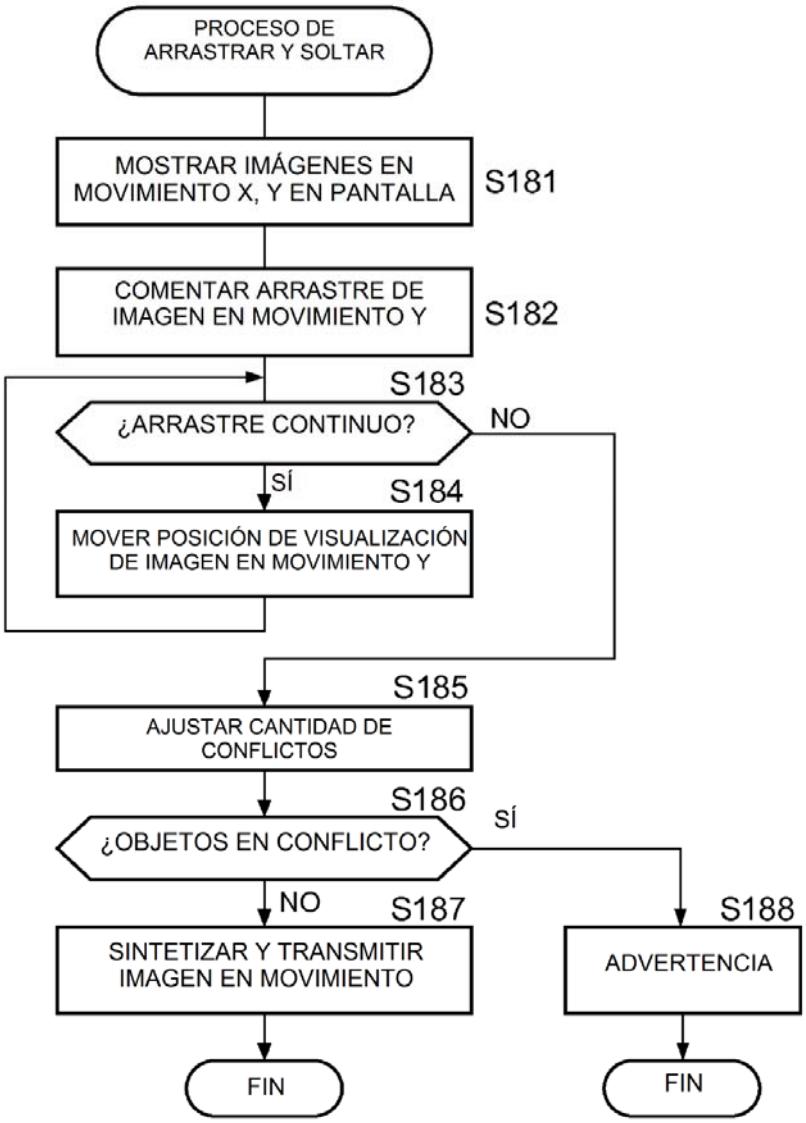
PROCESO DE ARRASTRAR Y SOLTAR
MOSTRAR IMÁGENES EN MOVIMIENTO X, Y EN PANTALLA
S181
COMENTAR ARRASTRE DE IMAGEN EN MOVIMIENTO Y
S182
S183
¿ARRASTRE CONTINUO?
NO
SÍ
S184
MOVER POSICIÓN DE VISUALIZACIÓN DE IMAGEN EN MOVIMIENTO Y
S185
AJUSTAR CANTIDAD DE CONFLICTOS
S186
¿OBJETOS EN CONFLICTO?
SÍ
NO
S187
SINTETIZAR Y TRANSMITIR IMAGEN EN MOVIMIENTO
S188
ADVERTENCIA
FIN
FIN

FIG. 8

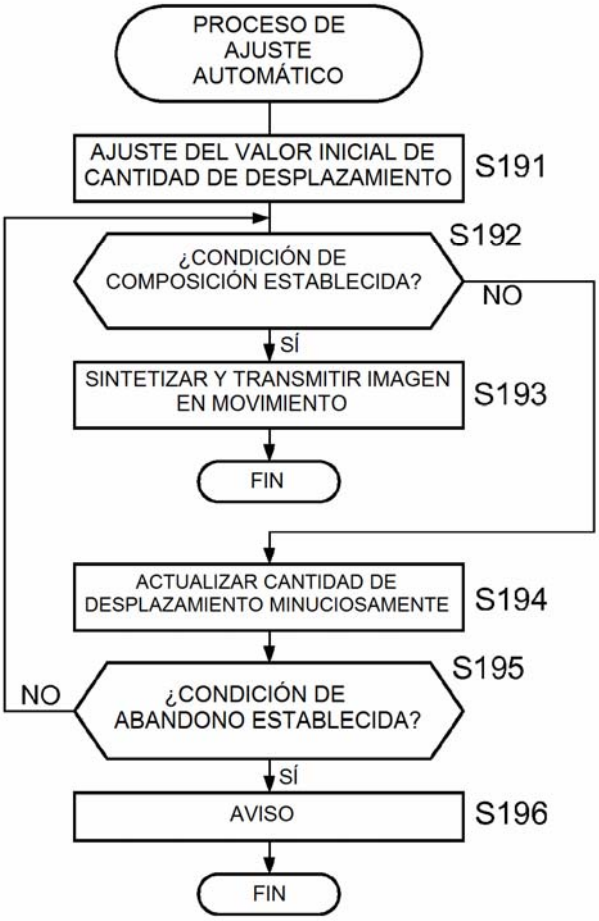
PROCESO DE AJUSTE AUTOMÁTICO
AJUSTE DEL VALOR INICIAL DE CANTIDAD DE DESPLAZAMIENTO
S191
S192
¿CONDICIÓN DE COMPOSICIÓN ESTABLECIDA?
NO
SÍ
SINTETIZAR Y TRANSMITIR IMAGEN EN MOVIMIENTO
S193
FIN
ACTUALIZAR CANTIDAD DE DESPLAZAMIENTO MINUCIOSAMENTE
S194
S195
NO
¿CONDICIÓN DE ABANDONO ESTABLECIDA?
SÍ
AVISO
S196
FIN

FIG. 9A

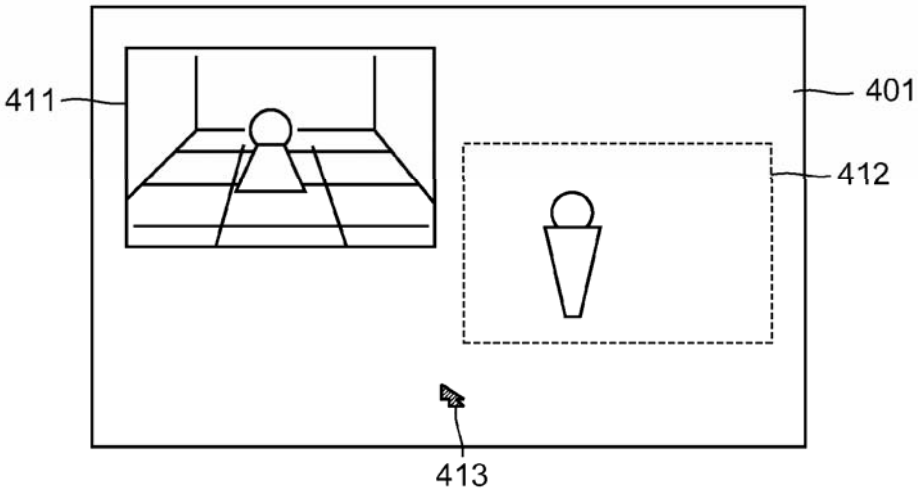
411
401
412
413

FIG. 9B

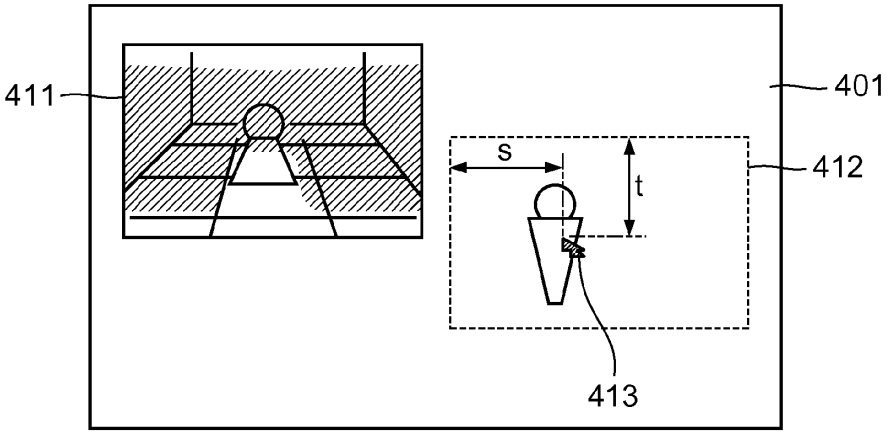

FIG. 9C

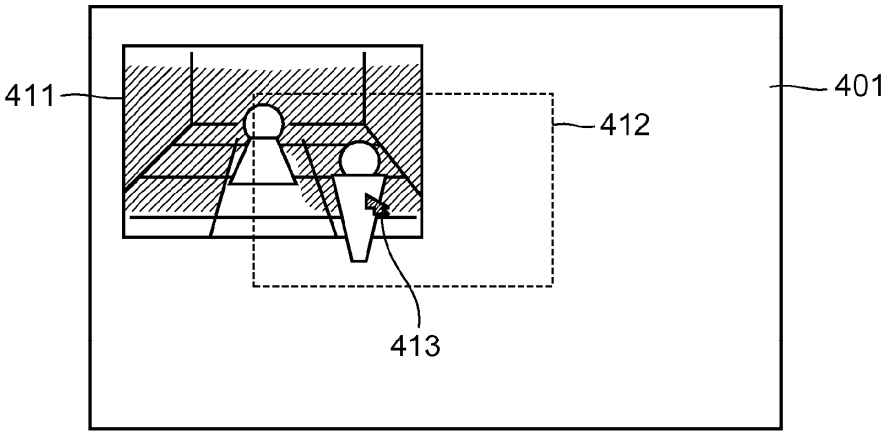

FIG. 9D

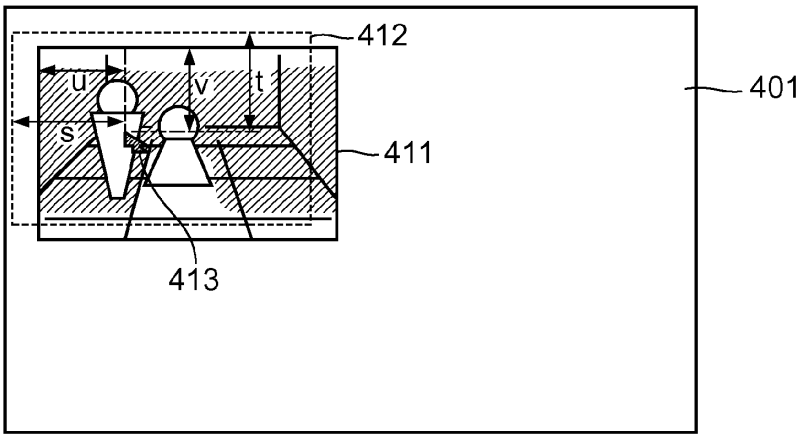

FIG. 9E

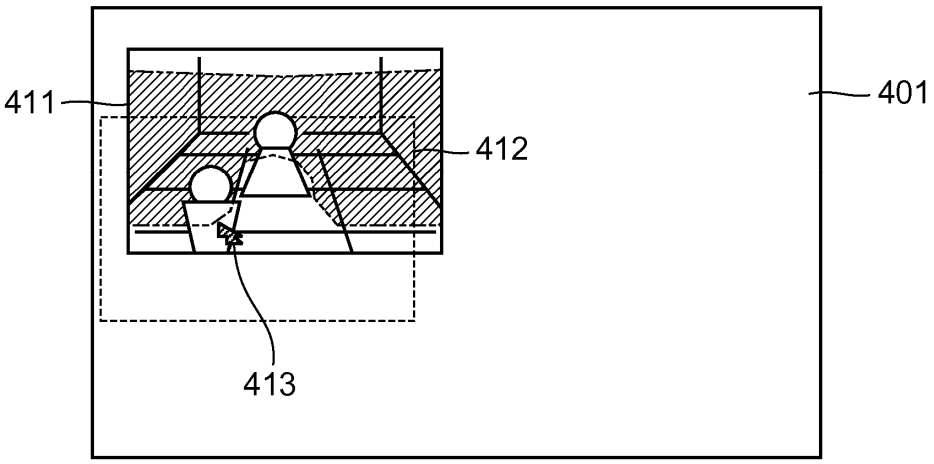

FIG. 10

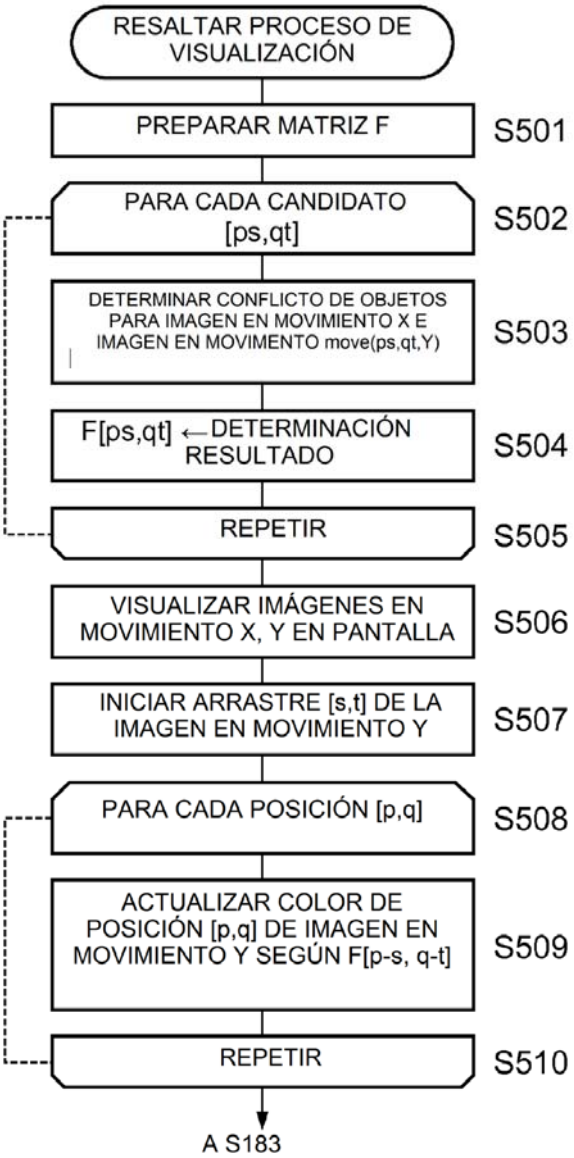
RESALTAR PROCESO DE VISUALIZACIÓN
PREPARAR MATRIZ F
S501
PARA CADA CANDIDATO [ps,qt]
S502
DETERMINAR CONFLICTO DE OBJETOS PARA IMAGEN EN MOVIMIENTO X E IMAGEN EN MOVIMENTO move(ps,qt,Y)
S503
F[ps,qt] ← DETERMINACIÓN RESULTADO
S504
REPETIR
S505
VISUALIZAR IMÁGENES EN MOVIMIENTO X, Y EN PANTALLA
S506
INICIAR ARRASTRE [s,t] DE LA IMAGEN EN MOVIMIENTO Y
S507
PARA CADA POSICIÓN [p,q]
S508
ACTUALIZAR COLOR DE POSICIÓN [p,q] DE IMAGEN EN MOVIMIENTO Y SEGÚN F[p-s, q-t]
S509
REPETIR
S510
A S183

FIG. 11A

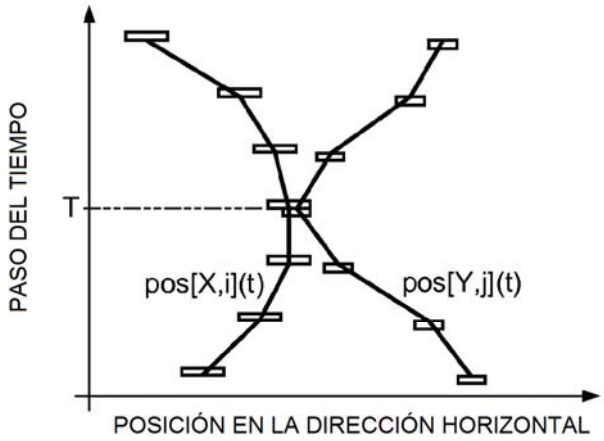

FIG. 11B

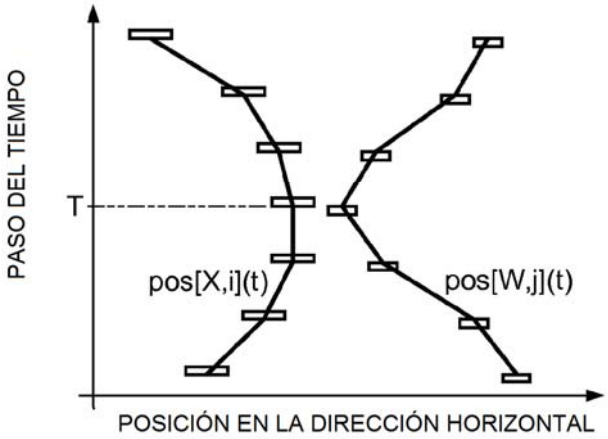

FIG. 12A

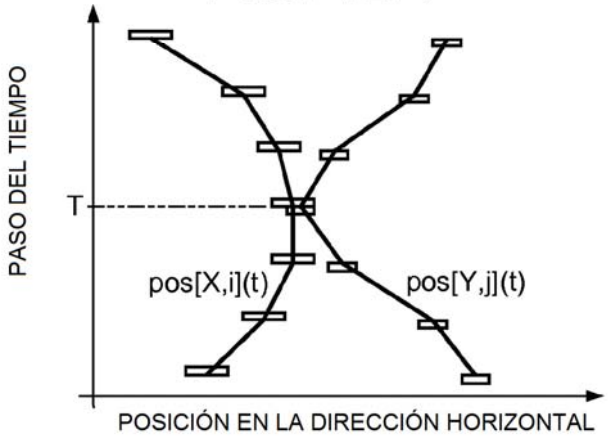

FIG. 12B

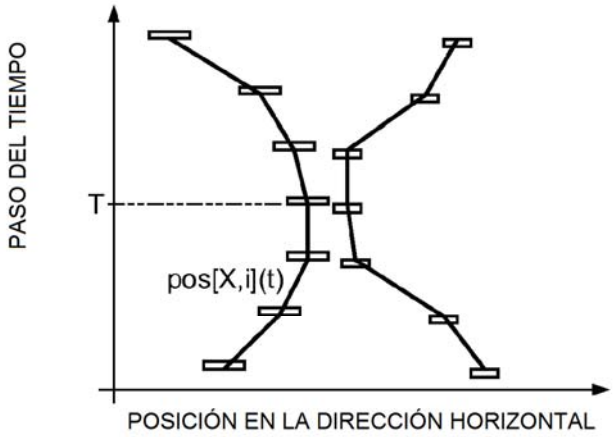

FIG. 12C

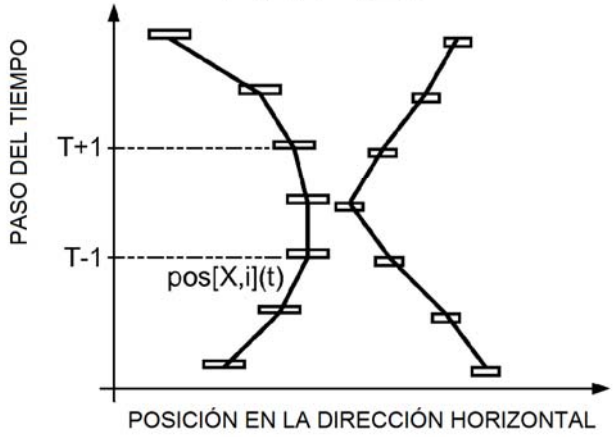

FIG. 12D

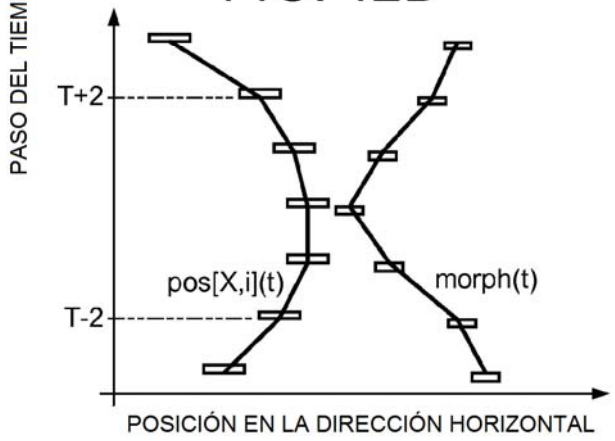
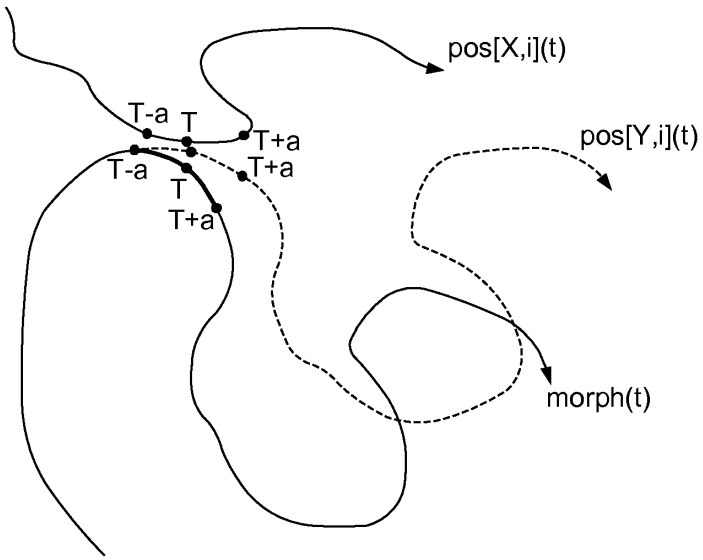
FIG. 13