

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 646 446**

51 Int. Cl.:

H04N 5/272 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.07.2012** **E 12176388 (2)**

97 Fecha y número de publicación de la concesión europea: **30.08.2017** **EP 2574042**

54 Título: **Reposicionamiento automático de elementos de video**

30 Prioridad:

23.09.2011 US 201113242167

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

13.12.2017

73 Titular/es:

DISNEY ENTERPRISES, INC. (100.0%)
500 South Buena Vista Street
Burbank, CA 91521, US

72 Inventor/es:

DESHPANDE, NIKHIL y
GEFEN, SMADAR

74 Agente/Representante:

ROEB DÍAZ-ÁLVAREZ, María

ES 2 646 446 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

Reposicionamiento automático de elementos de video

5 **Campo de la invención**

Las realizaciones ilustrativas se refieren a sistemas y métodos que reposicionan automáticamente elementos virtuales y físicos en una escena, de tal forma que su imagen en el video no es conflictiva. Específicamente, cuando una región de inserción del elemento virtual, como puede determinarse en directo durante un evento, ocluye uno o más elementos, ya sean elementos virtuales o físicos, tendrá lugar automáticamente un reposicionamiento de los elementos conflictivos, además, además del reposicionamiento, puede realizarse modificación estructural de elementos para acomodar espacialmente elementos cercanos.

15 **Información de antecedentes**

Debido al empleo de tecnologías de realidad aumentada llevado a cabo por sistemas de inserción de video la experiencia de ver muchos tipos de programas, incluyendo eventos deportivos, se ha mejorado mediante la capacidad de insertar mejoras virtuales (también denominadas como inserciones, gráficos, logotipos o signos) en una ubicación particular de la imagen de video que el espectador está viendo en la televisión. Por ejemplo, en fútbol americano, se inserta una Línea de Primer Intento (FDL) en la transmisión en tiempo real de un partido para indicar el punto del campo hacia el que el equipo en la actualidad en ataque debe dirigirse para conseguir cuatro intentos más. En otro ejemplo, se inserta una flecha de Intento y Distancia (DnD) indicando el número de jugada y la distancia que falta para alcanzar el FDL. Mientras estas posiciones y apariciones de elementos virtuales se determinan en directo en base a la progresión del juego, otros elementos virtuales pueden no estar relacionados con los eventos del partido, tales como impresiones publicitarias insertadas en diversas áreas del campo de juego o en diversas estructuras del estadio como una pared del estadio.

Un sistema de inserción es un sistema y método para insertar gráficos (elementos virtuales) en transmisiones de video en directo de una manera realista en tiempo real. En general, la perspectiva de la cámara se estima continuamente de modo que elementos gráficos, ya sean en 2D o 3D, pueden proyectarse en la imagen de video desde la perspectiva de la cámara actual como si estos elementos gráficos se ubicaran en una posición y orientación predeterminadas en la escena.

Los sistemas de inserción de Video de Transmisión en Directo (VIS) se desarrollaron y se usan comercialmente para el propósito de insertar publicidad y otros signos en secuencias de video, incluyendo transmisiones en directo de eventos deportivos. Un ejemplo de un VIS de transmisión en directo de este tipo se usa comercialmente con el nombre comercial de L-VIS®. En ejemplos adicionales, en las Patentes de Estados Unidos N.º 5.264.933, 5.543.856, 6.750.919 a Rosser et al. y Patente de Estados Unidos N.º 5.491.517 a Kreitman et al se describen VIS de transmisión en directo. Estos VIS, con diversos grados de éxito, incorporan fluidamente y realísticamente signos en el video original en tiempo real. El realismo se mantiene incluso a medida que la cámara cambia su perspectiva a través de la cobertura del evento y teniendo en cuenta elementos en movimiento en la escena que pueden ocluir los signos insertados.

La Figura 1 muestra un diagrama de bloques de nivel superior de un típico VIS 100. El componente 120 de cálculo de VIS principal recibe una señal de video 110 en directo de una cámara y a continuación emite, posiblemente con alguna latencia, un video mejorado 115. Además, el sistema incluye un componente GUI 150 con el que un operador controla el sistema antes y durante un evento y una unidad de signos 170 en la que se almacenan representaciones de los elementos virtuales insertados.

En la esencia de cada sistema de inserción está la capacidad de asociar un punto 226 en la escena a su proyección en el espacio de imagen de video 246, como se ilustra en la Figura 2. En general, se conoce el modelo de la escena. Por ejemplo, las dimensiones de un campo de fútbol americano se definen dentro del espacio de coordenadas 3D 210 y su modelo de la escena incluye la ubicación 3D de cada marca de referencia distintiva (por ejemplo, líneas 225, puntos de unión 226, etc.) en el campo. El plano X-Y 210 del campo muestra una región de inserción indicada mediante las coordenadas 3D – P_1 , P_2 , P_3 y P_4 . Esta región de inserción se asocia con un elemento virtual (por ejemplo, gráfico 2D) que se insertará 240 en la imagen de video actual 230 desde la perspectiva de la cámara actual. Por lo tanto, una cámara proyecta la escena en su espacio de imagen 230, con una proyección dictada por los parámetros de la cámara (por ejemplo, longitud focal, posición, orientación, etc.). Una vez que se conocen estos parámetros, cualquier región 220 dentro del espacio de mundo real 210 puede proyectarse en el espacio de imagen 240 de la cámara. La estimación de los parámetros de la cámara, a su vez, requiere conocimientos de puntos de referencia (marcas de referencia en la escena, por ejemplo 225 y 226, y sus correspondientes puntos en la imagen, por ejemplo 245 y 246). A continuación, se describe en detalle la forma en la que un VIS típico estima, de forma continua y en tiempo real, los parámetros de la cámara actual (denominada en este documento como el modelo de la cámara) y usa la misma para insertar virtualmente signos.

65

Encontrar pares de puntos correspondientes, en los que marcas de referencia en el campo se hacen coincidir con sus proyecciones en el fotograma de video actual, comienza con el proceso de reconocimiento como se realiza mediante el módulo de reconocimiento y seguimiento 125. A través del procesamiento de la imagen de video actual, se detectan características únicas tales como líneas, cónicas, uniones, esquinas, etc. En base a su estructura geométrica, aspecto o cualquier otro atributo se determinan sus correspondencias con marcas de referencia en el modelo de escena. Este proceso de reconocimiento puede llevarse a cabo cada varios fotogramas. Para los fotogramas que suceden entre el proceso de reconocimiento, el rastreo de las características detectadas mediante el módulo de reconocimiento y seguimiento 125 puede mantener sus correspondencias con las marcas de referencia de la escena. A continuación, en base a los pares correspondientes encontrados, el modelo de la cámara actual puede estimarse usando el módulo estimador 130 de modelo de cámara. Como se ha mencionado anteriormente, un modelo de cámara es un operador matemático (matriz) que convierte un punto 3D del espacio de escena 210 a su correspondiente punto en el espacio 230 de imagen de video. El modelo de la cámara se compone de parámetros intrínsecos, tales como longitud focal, y parámetros extrínsecos, tales como la posición y orientación de la cámara (barrido, inclinación y rotación).

Teniendo la estimación de modelo de la cámara actual, la unidad de deformación 135 deforma (proyecta) un elemento virtual dado en una postura 3D dada en el espacio 230 de imagen de video actual. Por ejemplo, un elemento virtual puede ser un logotipo. Este logotipo puede representarse en la base de datos 185 de signos mediante su imagen (por ejemplo, formato BMP o GIF) y su ubicación deseada (región de inserción) dentro del espacio 3D de la escena: P_1 , P_2 , P_3 y P_4 . La unidad de deformación 135 a continuación deformará esta imagen del logotipo, usando el modelo de la cámara, en una nueva imagen de signo dentro del espacio de imagen de video actual: C_1 , C_2 , C_3 y C_4 ; esta nueva imagen de signo está a continuación lista para convertirse en la imagen de video mediante el mezclador 145. Obsérvese que, un elemento virtual no se limita a un gráfico 2D, sino que puede ser cualquier estructura 3D. En este caso, unos datos representativos de elemento virtual 3D en la base de datos 185 de signos puede ser su modelo 3D (malla poligonal o representación basada en puntos), textura y posición deseada, orientación y escala en la escena. De manera similar, el conocimiento del modelo de la cámara actual puede usarse para convertir este elemento 3D desde esta perspectiva de cámara.

A continuación, el generador 140 de máscara de oclusión genera una función de transparencia o clave de máscara, que se aplica a continuación al proceso de inserción en el mezclador 145 para dar cuenta apropiadamente de cualquier obstáculo que pueda estar presente en la región de inserción. Realizando un procesamiento de oclusión antes de inserción, VIS 100 garantiza que la verisimilitud del logotipo insertado en la imagen de video se conserva cuando un elemento físico como un jugador entra en la región de inserción. En vez de ocluir al jugador con el logotipo insertado, la función de transparencia o clave de máscara garantiza que en cada ubicación de píxel en el que se produce un solapamiento entre el jugador y el logotipo, el píxel que corresponde al logotipo se suprime a favor del píxel de la imagen del jugador. Por lo tanto, en el mezclador 145 las imágenes de signos deformadas se superponen con la imagen de video basada en la máscara de oclusión.

Un operador, a través de un componente GUI 150, controla el sistema de procesamiento de inserción 120. Antes del partido, el operador establece y entrena el sistema, preparando el mismo para el evento en directo. Típicamente, el operador introduce datos relativos a la escena normalmente a través de interfaz gráfica. El operador define las coordenadas 3D de marcas de referencia en la escena dentro de un sistema de coordenadas 3D de la unidad 155 de modelado de escena. Por ejemplo, en un evento deportivo se introducirá la estructura de campo 210. El operador también puede entrenar al sistema para reconocer características de color de los primeros planos dinámicos (jugadores) y características de color del fondo estático (campo) utilizando unidad 160 de modelado de color. Estos datos se utilizarán posteriormente para la generación de máscara de oclusión. Otra información que el operador habitualmente introduce en el sistema es la ubicación y orientación 3D de inserción deseadas usando la unidad 165 de posicionamiento de signos de cada elemento virtual almacenado en la base de datos 185 de signos. Como se explicará a continuación, dependiendo del tipo de signos, esta información puede introducirse durante el establecimiento previo al evento o durante el partido.

En los sistemas VIS descritos anteriormente, particularmente, aunque no exclusivamente, en sistemas VIS capaces de insertar un elemento virtual en una región de inserción dinámicamente determinada, surge un problema cuando este elemento se inserta en un área de la imagen de video que no debería ocluirse. Por ejemplo, en un partido de fútbol americano, un logotipo virtual no debería ocluir el nombre del equipo en la zona de anotación o en un partido de baloncesto el logotipo no debería cubrir un anuncio real (en lugar de virtual) en la pared del estadio. En sistemas anteriores, la responsabilidad para realizar esto recaía en el operador manual que tenía que reposicionar el logotipo insertado para no interferir con ninguna porción de la imagen que debiera permanecer visible durante la transmisión. Tal reposicionamiento manual tiene el desafortunado efecto secundario de retrasar la inserción del logotipo en su posición deseada hasta tal punto que el espectador se da cuenta del aspecto repentino del logotipo en la imagen como si saliera de la nada. Un retraso visualmente perceptible de este tipo destruye la fluidez y realismo que son las señas de identidad de los VIS. El documento US 2003/0169366 analiza un sistema de subtítulos opcionales preparado remotamente mediante un subtítulador para reposicionar por alguien distinto al subtítulador.

Breve descripción de los dibujos

La Figura 1 muestra un sistema de inserción de video de transmisión en directo convencional.

5 La Figura 2 muestra un modelo de mundo real ilustrativo de una escena y una proyección de imagen de la misma.

La Figura 3 muestra un sistema de inserción de video de transmisión en directo con reposicionamiento automático de mejoras de video de acuerdo con una realización ilustrativa.

10 La Figura 4 ilustra reposicionamiento de un elemento de flecha virtual con relación a elementos físicos en el plano del campo de acuerdo con una realización ilustrativa.

La Figura 5 muestra un diagrama de flujo ilustrativo para reposicionar automáticamente elementos virtuales y físicos de acuerdo con restricciones predeterminadas y de modo que sus regiones abarcadas no se solapen.

15 **Descripción detallada**

Un método y sistema se definen en las reivindicaciones independientes 1 y 8, respectivamente, a las que el lector debería ahora referirse. Realizaciones específicas se definen en las reivindicaciones dependientes.

20 Las realizaciones ilustrativas pueden entenderse adicionalmente con referencia a la siguiente descripción de las realizaciones ejemplares y los dibujos adjuntos relacionados, en los que elementos similares están provistos de los mismos números de referencia. Las realizaciones ilustrativas se relacionan con sistemas y métodos para resolver la visualización de signos conflictivos en señales de video en directo o pregrabadas en las que signos insertados (por ejemplo, representados mediante elementos virtuales) y signos existentes (por ejemplo, representados mediante elementos físicos) están interfiriendo espacialmente entre sí y por lo tanto se reposicionan automáticamente dentro de la imagen de video. En esta descripción, el término "representado" cuando se usa para referirse a los elementos del fotograma de video (por ejemplo, ya sean los elementos virtuales o los elementos físicos) significa que el fotograma de video, cuando se visualiza, se concibe para mostrar estos elementos. Sin embargo, esto no significa que las realizaciones ilustrativas descritas en el presente documento requieran que el fotograma de video se visualice. Es decir, el procesamiento del fotograma de video y los elementos pueden llevarse a cabo sin ninguna visualización real del fotograma de video.

35 En la descripción anterior, se observó que existen muchos ejemplos de sistemas y métodos para la inserción de signos en secuencias de video y lo anterior proporcionó varios ejemplos de tales sistemas y métodos. Sin embargo, se observa que las realizaciones ilustrativas no se limitan a tales sistemas y métodos. Es decir, como se describirá en más detalle a continuación, los sistemas y métodos ilustrativos pueden aplicarse a cualquier tipo de sistema de inserción. También se observa que en la descripción anterior y en la siguiente descripción, los sitios de eventos ilustrativos se describen como sitios de eventos deportivos. Sin embargo, las realizaciones ilustrativas no se limitan a tales sitios de eventos remotos. Puede ser que las realizaciones ilustrativas se implementen para un espectáculo de estudio, por ejemplo.

45 La Figura 3 muestra un VIS 300 equipado con funcionalidad para garantizar que unos signos se insertan en una ubicación en la imagen de video que evitan ocluir cualquier otro elemento virtual o físico que son parte de la escena. Durante toda esta descripción, la expresión "elemento virtual" se usará para referirse a una representación por ordenador 3D de un gráfico generado por ordenador. Ejemplos de elementos virtuales incluyen la FDL y DnD descritos anteriormente, pero también puede incluir anuncios y otros tipos de gráficos. También, durante toda esta descripción, la expresión "elemento físico" se usará para referirse a una representación por ordenador 3D de un objeto actual. Ejemplos de objetos representados mediante elementos físicos incluyen signos que se pintan en el campo de juego (por ejemplo, nombre del equipo en la zona de anotación, logotipo del equipo en el medio del campo, etc.) y signos que se publican en los carteles laterales de la cancha. Por lo tanto, cuando se usa el término "elemento", este puede referirse a cualquiera o ambos de elementos físicos y elementos virtuales.

55 El VIS 300 puede ubicarse en cualquier ubicación adecuada en el sistema de producción de programa para crear un programa de transmisión. Por ejemplo, el VIS 300 puede ubicarse en un camión de producción en el lugar que se filma o remotamente en un estudio de producción. Además, el VIS 300 no se limita a señales de video de eventos en directo, sino que puede aplicarse a programas prealmacenados en cualquier medio adecuado, tales como cinta magnética, disco óptico o almacenamiento de memoria flash.

60 De acuerdo con una realización ilustrativa el reposicionamiento automático de mejoras de video reside esencialmente en la unidad de signos 370, incluyendo una unidad 375 de reconstrucción de signos, una unidad de reposicionamiento automático 380 y una base de datos 385 de signos. Como se ilustra en la Figura 4, una situación típica en una producción en la que se emplea realidad aumentada tiene una pluralidad de elementos dentro de la escena 400. Algunos de estos elementos pueden ser gráficos que se publican físicamente en el plano del campo, tales como elementos físicos 405, 410 y 415. Algunos de estos elementos pueden ser gráficos que se insertan virtualmente en el plano del campo, posiblemente en una ubicación dictada por los eventos de un partido, tales como

la flecha de DnD 420 y otros elementos 425 de mejoras (elementos virtuales). Por ejemplo, durante el partido el operador introduce a través de GUI 350 la ubicación deseada de la flecha de DnD 420 que se procesa mediante la unidad 365 de posicionamiento de signos, y, como se muestra en escena 400, el gráfico de DnD 420 puede ocultar a otro elemento 415 en el campo. Para resolver este problema, un método ilustrativo automáticamente reposiciona los elementos en el campo de acuerdo con reglas predefinidas como se explicará en mayor detalle a continuación.

Antes del evento, se modelan elementos físicos para facilitar su posible "eliminación" de la escena y re inserción en una nueva ubicación en la escena. Un remodelado de elemento físico tiene lugar en la unidad 375 de reconstrucción de signos. Por ejemplo, antes del partido, un elemento físico que corresponde al logotipo del equipo se genera capturando la imagen del logotipo del equipo y mapeando el mismo en el espacio 3D de la escena. Además, se genera un elemento de fondo (un segundo elemento físico) que representa el patrón y marcas del campo que se cubren por la región abarcada de elemento físico. Por lo tanto, en la base de datos 385 de signos, todo elemento físico almacenado tiene un elemento de fondo asociado al mismo. Durante el partido, cuando existe una necesidad de "mover" un elemento físico, el sistema de inserción se dirigirá para insertar el elemento de fondo asociado a la misma ubicación del elemento físico y a continuación insertar el elemento físico a la nueva ubicación deseada.

La Figura 5 muestra un diagrama de bloques de un método de reposicionamiento automático 500 ilustrativo que puede residir esencialmente en la unidad de reposicionamiento automático 380 del VIS 300. Para cada fotograma de video, antes de emplear la inserción, el método de reposicionamiento automático 500 recibe una lista de elementos y sus datos asociados desde la base de datos 385 de signos. Como se ha mencionado anteriormente, elementos en la lista pueden ser o bien físicos o virtuales y pueden ser o bien de dos dimensiones o tres dimensiones. Unos datos asociados del elemento incluyen su ubicación, dimensión y orientación dentro del espacio 3D de la escena y su modelo de aspecto - ya sea textura plana para un elemento 2D o una maya poligonal y plantillas de texturas para un elemento 3D. Para simplicidad de la descripción, el método se ilustra en la Figura 4 usando elementos 2D.

En la etapa 510, se reciben todos los elementos - representados mediante $R_{1,..,N}$, con sus prioridades asignadas - representadas mediante $P_{1,..,N}$. Una prioridad del elemento puede establecerse mediante un operador durante la configuración de sistema, por ejemplo. En la etapa 515, los elementos se asignan o bien un estado estático, R_i^s , o un estado dinámico, R_i^d , en el que elementos estáticos estarán en su lugar mientras elementos dinámicos pueden reposicionarse de acuerdo con este método. Asignar estado estático o dinámico puede hacerse como sigue. Las regiones abarcadas de los elementos y su relativa distancia espacial se analizan para detectar un conflicto. Un "conflicto" puede ser cualquier situación en la que es deseable mover o posicionar un elemento con respecto a la posición de otro elemento, incluyendo situaciones en las que se solapan elementos, en las que elementos están demasiado cerca entre sí o cualquier otra situación definida por prioridades o una regla predefinida. Si se encuentra que un grupo de elementos se solapan entre sí, se asignará un estado estático al elemento con la prioridad más alta y se asignará un estado dinámico a todos los demás elementos. Por ejemplo, los elementos 405, 410, 415, 420 y 425 pueden tener designadas prioridades 5, 4, 3, 2, 1, respectivamente. En este caso, se asigna un estado estático a los elementos 405, 410 y 415 y se asigna un estado dinámico a los elementos 420 y 425. Significando que en esta situación, primero, el elemento 420 se reposicionará y, segundo, el elemento 425 se reposicionará, ya que el elemento 420 tiene una prioridad más alta con respecto al elemento 425. Se observa que en este ejemplo se ha asignado una prioridad más alta a los elementos físicos 405, 410, 415 que a los elementos virtuales 420 y 425. Sin embargo, esto es únicamente ilustrativo y puede haber casos en los que los se asignan prioridades más altas a elementos virtuales que a elementos físicos. La asignación de prioridades para elementos puede basarse en cualquier número de reglas o prácticas acogidas por el operador. Por ejemplo, podría ser que los elementos físicos reciban prioridades más altas. En otro ejemplo, ciertos anunciantes pueden pagar para que sus elementos virtuales tengan una prioridad más alta que otros anunciantes o incluso que elementos físicos. Como puede observarse a partir de estos ejemplos, la asignación de prioridades a elementos puede ser específica para el operador del sistema.

El reposicionamiento de los elementos dinámicos se efectúa en las etapas 520 y 525. La realización ilustrativa describe un mecanismo específico para reposicionar un elemento. Sin embargo, los expertos en la materia pueden contemplar otros mecanismos de reposicionamiento equivalentes. En la etapa 520 se calcula un mapa de inserción en base a los elementos estáticos actuales. Como se ilustra en escena 430 un mapa de inserción se define dentro de una cierta región de interés ("ROI") 435. Se asigna un valor a los puntos de red 450-454 representando una distancia al límite más cercano de cualquiera de un elemento estático o el límite ROI (por ejemplo, el punto de red 450 está más cerca del elemento estático 440, el punto de red 452 está más cerca del elemento estático 445 y el punto de red 454 está más cerca del límite de la ROI 435). El valor de los puntos de red ubicados dentro de una región 456 del elemento estático o fuera de la ROI es cero. La resolución de red puede depender del tamaño de etapa de reposicionamiento deseado. Tal tamaño de etapa de reposicionamiento deseado puede ser ajustable por un usuario.

A continuación, en la etapa 525, se selecciona el R_i^d de elemento con la prioridad más alta 420 de la lista actual de elementos dinámicos para reposicionar como sigue. El área del mapa de inserción que se solapa con la región de R_i^d se sumará a un valor representado como A_i^d . Se puede observar que A_i^d se vuelve mayor a medida que R_i^d 420 se aleja de los elementos estáticos 405-415 y de los límites de ROI 435. Por lo tanto, reposicionar R_i^d puede llevarse a cabo a través de un proceso de maximización de A_i^d . Este proceso se ilustra en la escena 460 en la que la flecha

de DnD se mueve desde su posición actual 490 a su nueva posición 492. Encontrar la nueva posición de R_i^d para la que se maximiza A_i^d puede realizarse a través de una búsqueda exhaustiva o a través de métodos de optimización conocidos en la técnica tales como descenso de gradiente, recocido simulado, algoritmo genético, etc. La ROI 435 puede usarse para restringir el reposicionamiento. Una característica de restricción de este tipo puede ser útil donde se necesita reposicionar más de un elemento, como se explicará a continuación. Expertos en la materia entenderán que mientras la ROI 435 es útil en los elementos de reposicionamiento, no es necesaria. En un caso de este tipo, el elemento puede posicionarse en cualquier ubicación dentro de la escena.

En la etapa 530 se determina si después del reposicionamiento de R_i^d esta región del elemento todavía se solapa con la de regiones abarcadas de los otros elementos estáticos. Si esto es afirmativo, en la etapa 535, el tamaño o forma del elemento puede modificarse. Por ejemplo, la flecha de DnD puede disminuirse en tamaño para acomodar los elementos circundantes. Otra posibilidad es permitir algún grado de transparencia cuando se convierten las regiones solapantes de R_i^d . En la etapa 540, la nueva posición, y, posiblemente, nuevo tamaño y estructura del R_i^d del elemento dinámico actual se almacenan en la base de datos 385 de signos y el estado del elemento cambia a estático. En este punto, en la etapa 545, si hay más elementos dinámicos, el método se repite desde la etapa 520 en la que el mapa de inserción se actualiza en base a la lista actual de elementos estáticos. Como se ha descrito anteriormente, el elemento dinámico 420 anteriormente reposicionado ahora se asigna un estado estático. En este segundo paso, por ejemplo, el reposicionamiento del elemento dinámico 425 tendrá lugar en relación con los ahora elementos estáticos 405, 410, 415 y 420.

Como se ha mencionado anteriormente, la ROI 435 que circunscribe el mapa de inserción se define para servir como una restricción para limitar el alcance del reposicionamiento de un elemento dinámico. Por ejemplo, en el caso de una flecha de DnD, únicamente se permite movimiento en la dirección Y. En otro ejemplo, si hay dos elementos dinámicos, tales como los elementos 420 y 425, un posible esquema de reposicionamiento puede ser definir para cada uno de ellos una ROI separada en la que el movimiento se limitará dentro de la ROI de cada uno.

En una realización ilustrativa puede definirse un seudo-elemento e introducirse (junto con otros elementos físicos o virtuales) en el método descrito en la Figura 5. Este seudo-elemento, cuando se asigna un estado estático, por ejemplo, puede servir como un "ocupante de espacio" que evita que otros elementos ocupen la misma región ocupada por este elemento. Evitar la inserción de elementos virtuales o reposicionar elementos físicos en o de una cierta área en la cancha/campo, respectivamente, puede ser instrumental para asegurar que los elementos no están interfiriendo con objetos de primer plano tales como jugadores. Uno o más seudo-elementos, por ejemplo, pueden introducirse en el sistema durante un segmento de tiempo limitado ligado opcionalmente a un evento específico.

Los expertos en la materia entenderán que las realizaciones ilustrativas descritas anteriormente pueden implementarse en cualquier configuración adecuada de software o hardware o una combinación de los mismos. Una plataforma de hardware ilustrativa para implementar las realizaciones ilustrativas puede incluir, por ejemplo, una plataforma basada en Intel x86 con sistema operativo compatible, una plataforma Mac y MAC OS, etc. La plataforma de hardware también puede ser una plataforma de procesamiento de video estándar y su hardware y componentes de software asociados. En un ejemplo adicional, las realizaciones ilustrativas de los sistemas y métodos de inserción pueden ser un programa que contiene líneas de código almacenado en un medio de almacenamiento legible por ordenador no transitorio que, cuando se compila, puede ejecutarse en un procesador.

Será evidente para expertos en la materia que pueden hacerse diversas modificaciones en la presente invención. Se concibe que la presente invención cubre modificaciones y variaciones de esta invención siempre que estén dentro del alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un método, que comprende:

5 recibir un fotograma de vídeo de una escena;
 recibir datos, incluyendo datos de posición, describiendo una pluralidad de elementos mapeados en un espacio
 3D de la escena, en el que los datos comprenden un elemento físico y un elemento virtual;
 analizar los datos que describen la pluralidad de elementos para determinar una existencia de un conflicto entre
 el elemento físico y el elemento virtual de la pluralidad de elementos;
 10 en el que los elementos se asignan o bien un estado estático o un estado dinámico;
 cuando existe un conflicto, resolver el conflicto reposicionando automáticamente o modificando estructuralmente
 al menos un elemento de la pluralidad de elementos que tiene un estado dinámico;
 en el que el reposicionamiento automático o modificación estructural se realiza maximizando una suma de
 valores en un área de un mapa de inserción que se solapa con el al menos uno de la pluralidad de elementos;
 15 en el que el mapa de inserción se define dentro de una región de interés (ROI), en el que el límite de la región de
 interés circunscribe el mapa de inserción y sirve como una restricción para limitar el alcance del
 reposicionamiento del elemento dinámico;
 en el que los valores del mapa de inserción incluyen valores representativos de una distancia al límite más
 cercano de cualquiera de la pluralidad de elementos estáticos o al límite de la región de interés; e
 20 insertar al menos un elemento reposicionado o modificado estructuralmente de la pluralidad de elementos en el
 fotograma de video para crear un fotograma de video modificado.

2. El método de la reivindicación 1,

25 en el que el elemento físico es una representación por ordenador de un objeto real en la escena representada en
 el fotograma de vídeo, comprendiendo el método además:
 modelar el objeto real para crear el elemento físico;
 crear un elemento de fondo que corresponde al elemento físico, en el que el elemento de fondo es una
 30 representación por ordenador de un fondo que está tapado por el objeto real; e
 insertar el elemento de fondo en el fotograma de video cuando se reposiciona el elemento físico.

3. El método de la reivindicación 1,

35 en el que el elemento virtual es un gráfico generado por ordenador.

4. El método de la reivindicación 1,

40 en el que los datos que describen la pluralidad de elementos adicionalmente incluyen uno o más de datos de
 tamaño, aspecto y transparencia.

5. El método de la reivindicación 1,

45 en el que la etapa de reposicionar automáticamente o modificar estructuralmente la pluralidad de elementos se
 realiza en un orden definido mediante prioridades asignadas a los elementos.

6. El método de la reivindicación 5,

50 en el que un anuncio asociado con un elemento virtual de la pluralidad de elementos se prioriza en base a las
 prioridades asignadas.

7. El método de la reivindicación 1,

55 en el que el conflicto incluye dos o más elementos conflictivos y adicionalmente en el que los elementos se
 solapan entre sí cuando se insertan en el fotograma de video y en el que la etapa de resolver el conflicto implica
 reposicionar automáticamente o modificar estructuralmente un elemento físico y/o un elemento virtual en la
 escena.

8. Un sistema (300), que comprende:

60 una memoria que almacena datos, incluyendo datos de posición, describiendo una pluralidad de elementos
 mapeados en un espacio 3D de una escena, en el que la escena se representa en un fotograma de vídeo, en el
 que los datos comprenden un elemento físico y un elemento virtual; y
 un procesador (300) que comprende una unidad de reposicionamiento automático (380) configurada para
 65 analizar al menos uno de la pluralidad de elementos para determinar una existencia de un conflicto entre el
 elemento físico y el elemento virtual de la pluralidad de elementos;

en el que elementos se asignan o bien un estado estático o un estado dinámico; y cuando existe un conflicto, resolver el conflicto reposicionando automáticamente o modificando estructuralmente al menos un elemento de la pluralidad de elementos que tiene un estado dinámico; en el que el reposicionamiento automático o modificación estructural se realiza maximizando una suma de valores en un área de un mapa de inserción que se solapa con el al menos uno de la pluralidad de elementos; en el que el mapa de inserción se define dentro de una región de interés (ROI), en el que la región de interés circunscribe el mapa de inserción y sirve como una restricción para limitar el alcance del reposicionamiento del elemento dinámico; en el que los valores del mapa de inserción incluyen valores representativos de una distancia al límite más cercano de cualquiera de la pluralidad de elementos estáticos o al límite de la región de interés; e insertar al menos un elemento reposicionado o modificado estructuralmente de la pluralidad de elementos en el fotograma de video para crear un fotograma de video modificado.

9. El sistema (300) de la reivindicación 8,

en el que el elemento físico es una representación por ordenador de un objeto real dentro de una escena representada mediante el fotograma de video, en el que la unidad de reposicionamiento automático (380) modela un objeto real para crear el elemento físico, crea un elemento de fondo que corresponde al elemento físico, en el que el elemento de fondo es una representación por ordenador de un fondo que está tapado por el objeto real e inserta el elemento de fondo en el fotograma de video modificado cuando se reposiciona el elemento físico.

10. El sistema de la reivindicación 8,

en el que los datos que describen la pluralidad de elementos adicionalmente incluyen uno o más de datos de tamaño, aspecto y transparencia.

11. El sistema de la reivindicación 8,

en el que la unidad de reposicionamiento automático (380) se configura para realizar la etapa de reposicionar automáticamente o modificar estructuralmente la pluralidad de elementos en un orden definido mediante prioridades asignadas a los elementos.

12. El sistema de la reivindicación 11,

en el que la unidad de reposicionamiento automático (380) se configura adicionalmente para:

definir al menos un elemento de la pluralidad de elementos como un elemento estático y en el que el elemento estático no se reposiciona o modifica; definir los otros elementos de la pluralidad de elementos como elementos dinámicos, en el que los elementos dinámicos se reposicionan o modifican con respecto al elemento estático; y después de reposicionar o modificar el elemento dinámico de los elementos dinámicos, definir el elemento dinámico particular como un elemento estático.

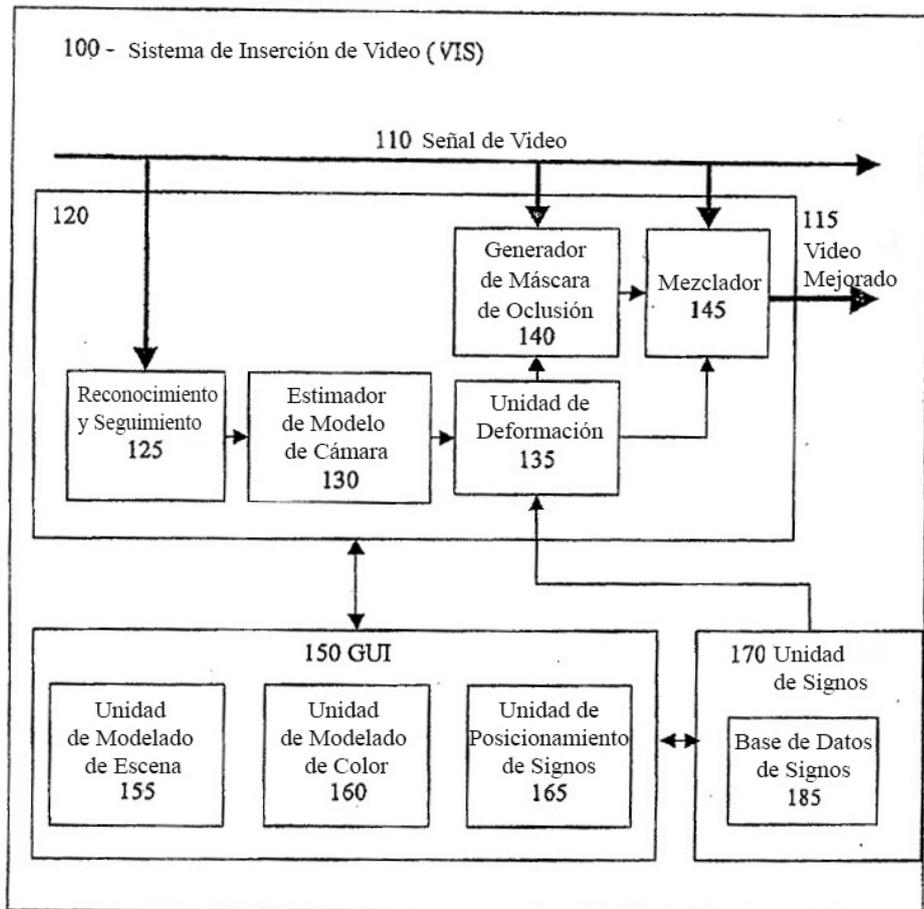


FIG. 1

TÉCNICA ANTERIOR

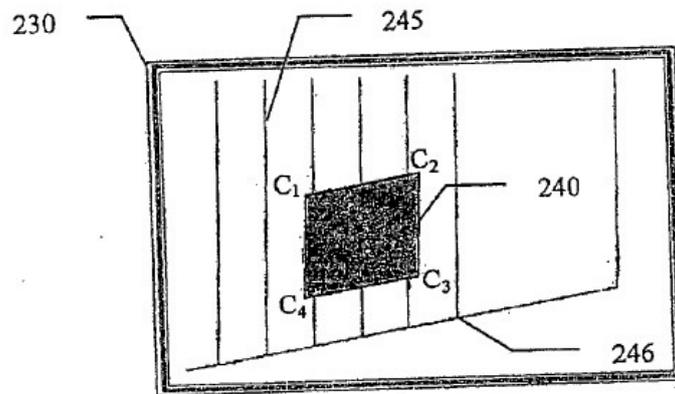
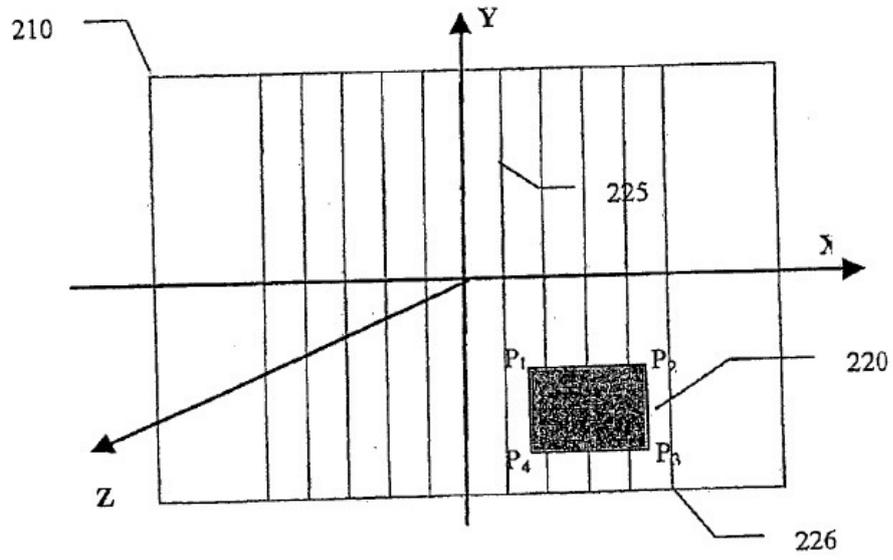


FIG. 2

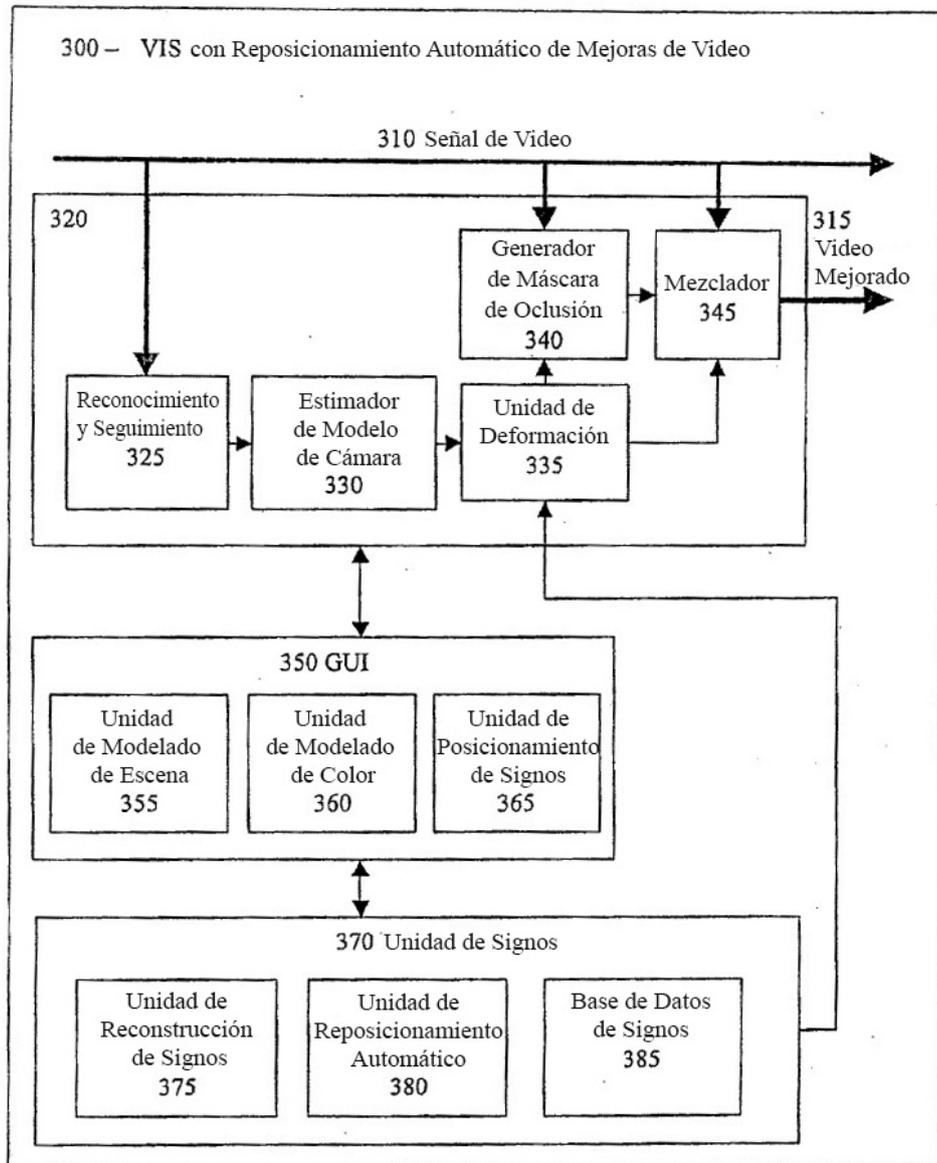


FIG. 3

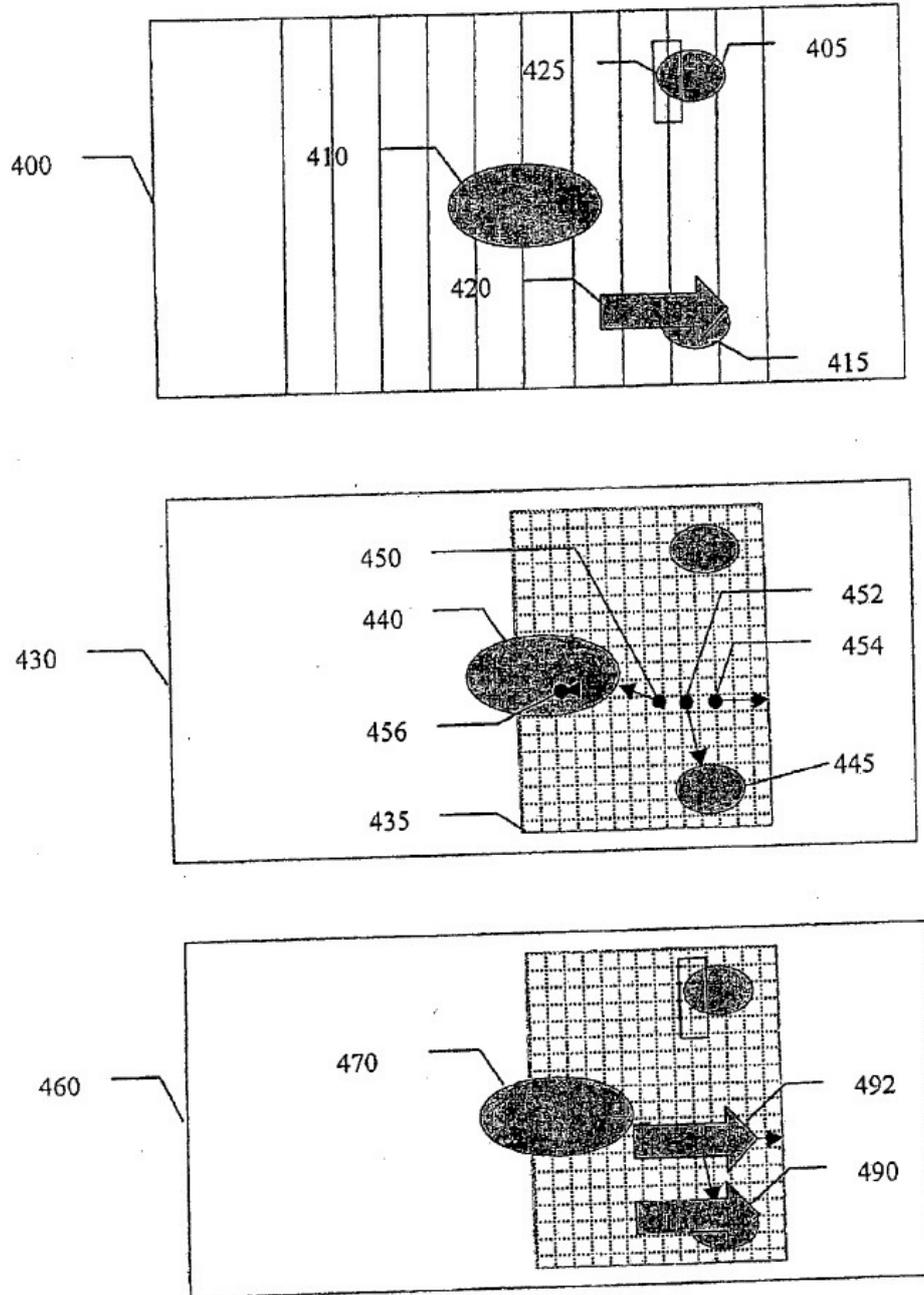


FIG. 4

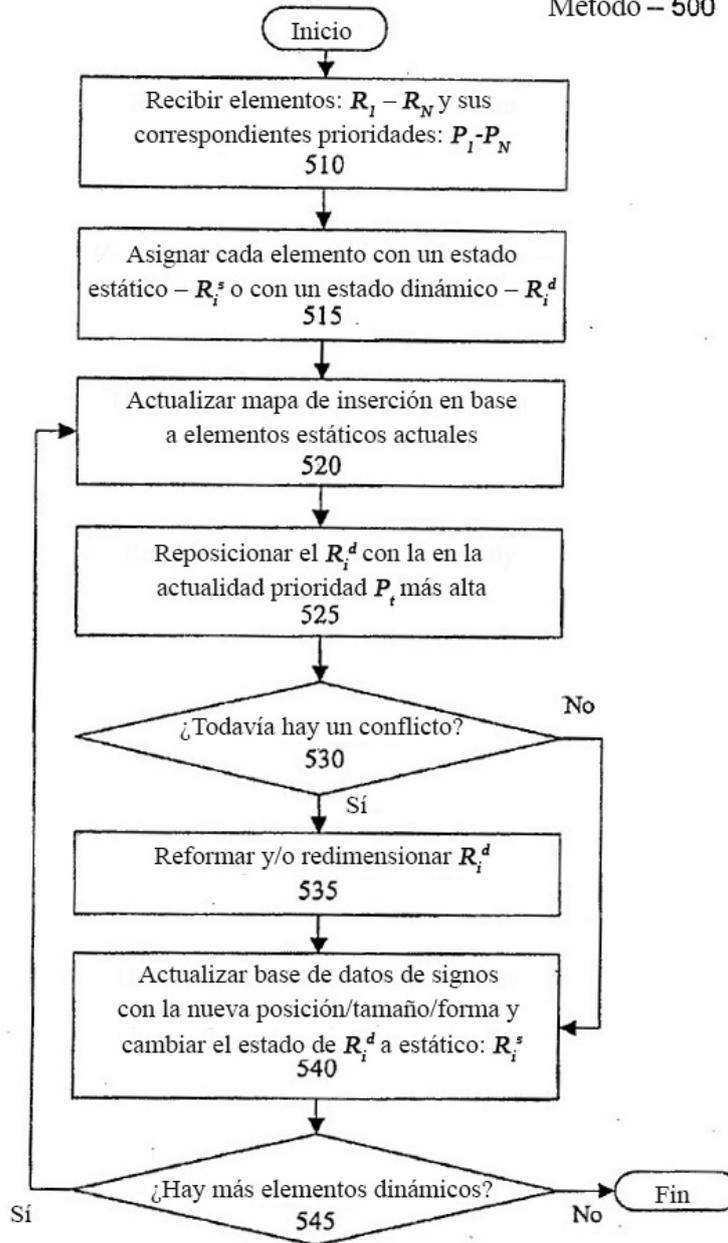


FIG. 5