



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 291 451**

(51) Int. Cl.:
H04N 7/24 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **02711963 .5**

(86) Fecha de presentación : **15.01.2002**

(87) Número de publicación de la solicitud: **1354479**

(87) Fecha de publicación de la solicitud: **22.10.2003**

(54) Título: **Procedimiento y equipamiento para la gestión de interacciones dentro de la norma MPEG-4.**

(30) Prioridad: **15.01.2001 FR 01 00486**
07.02.2001 FR 01 01648

(45) Fecha de publicación de la mención BOPI:
01.03.2008

(45) Fecha de la publicación del folleto de la patente:
01.03.2008

(73) Titular/es:
Groupe des Ecoles des Telecommunications
46, rue Barrault
75634 Paris Cédex 13, FR
FRANCE TELECOM

(72) Inventor/es: **Dufourd, Jean-Claude;**
Concolato, Cyril;
Preteux, Françoise y
Preda, Marius

(74) Agente: **Elzaburu Márquez, Alberto**

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

Procedimiento y equipamiento para la gestión de interacciones dentro de la norma MPEG-4.

5 El presente invento se refiere a la gestión de las interacciones multimedia efectuadas por uno o varios usuarios a partir de terminales multimedia. Las interacciones pueden ser tanto textuales, vocales o de gestos. Las interacciones son adquiridas a partir de no importa qué dispositivo de entrada clásico tal como el ratón, el joystick, el teclado o no clásico tal como sistemas de reconocimiento y síntesis de voz o enlaces de mando por la mirada y/o el gesto. Estas interacciones multimedia son tratadas en el contexto de la norma internacional MPEG-4.

10 Contexto del invento

La norma MPEG-4 (ISO/IEC 14496) especifica un sistema de comunicación para las escenas audiovisuales interactivas. Conforme a la presente memoria, los usuarios tienen la posibilidad de interactuar con la escena multimedia. 15 Estas interacciones de usuarios pueden ser tratadas localmente por el lado del cliente o retransmitidas al servidor para tratamiento.

La norma ISO/IEC 14496-1 (MPEG-4 Systems) define el formato binario de descripción de escena (BIFS: Formato Binario para Escena) que se refiere a la organización de objetos audiovisuales en una escena. Los comportamientos de los objetos y sus respuestas a las interacciones efectuadas por los usuarios pueden ser representadas en el formato BIFS por medio de fuentes y de objetivos (rutas) de eventos así como con la ayuda de captadores (nodos especiales capaces de provocar los eventos). Las interacciones por el lado del cliente consisten en la modificación de los atributos de los objetos de la escena conforme a las acciones especificadas por los usuarios. Sin embargo, MPEG-4 Systems no define un enlace de usuario particular o un mecanismo que asocia la interacción del usuario con eventos BIFS. 25

BIFS-Command es el subconjunto de la descripción BIFS que permite modificar propiedades del gráfico de la escena, de sus nodos o de sus comportamientos. BIFS-Command es por tanto utilizado para modificar un conjunto de propiedades de la escena en un instante dado. Las órdenes son reagrupadas en CommandFrames con el fin de permitir el envío de varias órdenes en una sola unidad de acceso (Unidad de Acceso). Las cuatro órdenes o comandos de base 30 son los siguientes: reemplazamiento de una escena entera, e inserción, borrado o reemplazamiento de las estructuras de nodo, entrada de eventos (eventin), exposedField, valor indexado en un MFField o ruta. La identificación de un nodo en la escena es proporcionada por un nodeID (Identificador de nodo). La identificación de los campos de un nodo es proporcionada por el INid del campo.

BIFS-Anim es el subconjunto de la descripción BIFS que se refiere a la puesta en marcha continua de ciertos campos de nodos en el gráfico de la escena. BIFS-Anim es utilizado para integrar diferentes tipos de animación, incluida la animación de modelos de cara, de cuerpo humano, de mallas, así como diversos tipos de atributos tales como posiciones bidimensionales y tridimensionales, rotaciones, factores de escala o información colorimétrica. El BIFS-Anim especifica un flujo así como los procedimientos de codificación y decodificación para cómo animar 40 ciertos nodos de la escena que comprenden campos dinámicos particulares. El inconveniente mayor del BIFS-Anim en el marco del presente invento es el siguiente: El BIFS-Anim no especifica cómo animar todos los campos susceptibles de ser actualizados de todos los nodos de una escena. Además, el BIFS-Anim utiliza una máscara de animación que forma parte de las informaciones de configuración del decodificador. La máscara de animación no es modificable por una interacción directa de un usuario. Por consiguiente, el BIFS-Anim no está adaptado para la interacción del 45 usuario que necesita un nivel elevado de flexibilidad y la posibilidad de hacer evolucionar dinámicamente los nodos de la escena a modificar.

El MPEG-J es un sistema de programación que especifica los enlaces para garantizar la interoperatividad de un difusor de medios MPEG-4 con código Java. El código Java llega al nivel del terminal MPEG-4 bajo forma de un flujo elemental distinto. A continuación, es dirigido hacia el entorno de ejecución MPEG-J que comprende una máquina virtual Java a partir de la cual el programa MPEG-J tendrá acceso a los diversos componentes del difusor de medio 50 MPEG-4. El enlace de programación SceneGraph proporciona un mecanismo por el cual las aplicaciones MPEG-J acceden a la escena utilizada para la composición por el difusor de medio BIFS y la manipulan. Se trata de una enlace de bajo nivel, que permite a la aplicación MPEG-J controlar los eventos de la escena y modificar la arborescencia de la escena por programa. Puede igualmente crearse y manipularse nodos, pero sólo los campos de los nodos para los cuales se ha definido una identificación de nodo son accesibles a la aplicación MPEG-J. Además, una puesta en práctica del MPEG-J requiere recursos demasiado importantes para numerosas aplicaciones, particularmente cuando se trata de dispositivos portátiles de pequeño tamaño y de decodificadores. Así, el MPEG-J no es apropiado en la definición de procedimientos de interacción de usuarios disponibles sobre los terminales que presentan capacidades 60 limitadas.

El análisis del estado de la técnica presentada más arriba describe y examina brevemente los principales procedimientos que podrían ser utilizados para generar las interacciones de usuarios multimedia. Queda por completar por los aspectos relativos a las arquitecturas de gestión de interacción actuales. Hasta ahora, no existen más que dos maneras de abordar la interacción. En primer lugar, en el marco MPEG-4 y únicamente para interacciones de tipo puntero, el 65 dispositivo de composición es encargado de transcodificar los eventos que provienen de usuarios en acción de modificación de la escena. En segundo lugar, fuera del contexto normativo MPEG-4, todas las interacciones que no sean las de tipo puntero deben ser puestas en práctica en una aplicación específica. En consecuencia, la interoperatividad se ha

perdido. Las dos opciones precedentemente descritas están demasiado limitadas para aprender en su generalidad y su tipo genérico el concepto de interactividad multiusuario que resulta el reto principal de los sistemas de comunicación.

El presente documento propone procedimientos y un sistema genéricos para gestionar las interacciones multimedia efectuadas por uno o varios usuarios desde un terminal multimedia. El sistema propuesto es una extensión de las especificaciones de la parte "Sistemas MPEG-4". Especifica como asociar las interacciones mono o multiusuarios con eventos BIFS, reutilizando la arquitectura de Sistemas MPEG-4. El sistema unido al invento es genérico puesto que permite tratar todos los tipos de interacciones mono o multiusuarios a partir de dispositivos de entrada tanto primitivos (ratones, teclados) como complejos (que necesitan tener en cuenta 6 grados de libertad o que ponen en práctica sistemas de reconocimiento de voz). Por la simple reutilización de herramientas existentes, este sistema es utilizable en todas las situaciones, incluidas las que no pueden soportar más que un nivel muy débil de complejidad.

En el estado de la técnica, se conoce el documento WO 0000898. Se refiere a una interacción multiusuario para una comunicación multimedia que consiste en gestionar un mensaje en un ordenador de usuario local, conteniendo el mensaje datos de medio orientados a objeto (por ejemplo un flujo de datos audio numéricos o un flujo de datos video numéricos o los dos) y en transmitir el mensaje a un ordenador de usuario a distancia. El ordenador de usuario local presenta una escena que comprende los datos del medio orientados a objeto y compartidos entre el ordenador de usuario local y el ordenador de usuario a distancia. Este último construye el mensaje con la ayuda de una clase de gestor de mensajes. La interacción multiusuario para la comunicación multimedia es una extensión de la Versión 1 MPEG-4.

Otro documento PCT W09939272 se refiere a un sistema de comunicación interactiva fundado sobre MPEG-4, en el que los descriptores de mando o control son utilizados con nodos de encaminamiento de orden o vías de encaminamiento de servidor en la descripción de la escena con el fin de aportar un soporte a la interactividad específica para la aplicación. Se puede ayudar a la selección del contenido indicando la presentación en parámetros de mando, indicando el identificador de mando que la orden es una orden de selección de contenido. Se puede crear una escena inicial que comprende varias imágenes y un texto que describe una presentación asociada a una imagen. Un descriptor de selección de contenido está asociado a cada imagen y al texto correspondiente. Cuando el usuario hace clic sobre una imagen, el cliente transmite la orden que contiene la presentación seleccionada y el servidor dispara una nueva presentación. Se puede poner en aplicación esta técnica en cualquier contexto de aplicación, de la misma manera que se puede utilizar HTTP y CGI para poner en práctica toda la funcionalidad de aplicación fundada sobre un servidor.

En el marco de la nueva aproximación propuesta de la interacción multimedia mono o multiusuario, los datos de interacción generados por un dispositivo de entrada cualquiera son considerados como flujos elementales MPEG-4. Se deduce de ello que operaciones similares a las aplicadas, a no importa qué flujo de datos elementales, pueden ser entonces puestas en práctica, utilizando directamente la cadena de descodificación estándar.

En consecuencia, el presente invento propone utilizar un modelo similar al presentado en MPEG-4 con el fin de tratar los datos de la interacción.

A este efecto, el invento se refiere según su acepción más general a un procedimiento para la gestión de las interacciones entre periféricos de mando y aplicaciones multimedia que explotan la norma MPEG-4, entregando dichos periféricos de mando señales numéricas función de acciones de uno o varios usuarios, procedimiento caracterizado porque comprende una etapa de construcción de una secuencia numérica que presenta la forma de un nodo BIFS [Formato Binario para Escena con el sentido de la norma MPEG-4], comprendiendo este nodo uno o varios campos que definen el tipo y el número de datos de interacción a aplicar a objetos de la escena.

Según un modo de puesta en práctica preferido, este nodo comprende un indicador cuyo estado permite o prohíbe que una interacción sea tenida en cuenta por la escena.

Según una variante, dicho nodo comprende una etapa de señalización de la actividad del dispositivo asociado.

Ventajosamente, el procedimiento comprende una etapa de designación de la naturaleza de la acción o acciones a ejercer sobre uno o varios de los objetos de la escena por medio del o de los campos del nodo.

Según un modo de puesta en práctica preferido, el procedimiento comprende una etapa de construcción, a partir del o de los campos del nodo, de otra secuencia numérica compuesta de al menos una acción a ejercer sobre la escena, y al menos un parámetro de dicha acción, cuyo valor corresponde a una variable entregada por dicho periférico.

Según un modo de puesta en práctica preferido, el procedimiento comprende una etapa de transferencia de dicha secuencia numérica en la memoria de composición.

Según un modo de puesta en práctica preferido, la transferencia de dicha secuencia numérica utiliza la cadena de descodificación de los sistemas MPEG-4 para introducir las informaciones de interacción en el dispositivo de composición.

ES 2 291 451 T3

Según un modo de puesta en práctica particular, la etapa de transferencia de la secuencia numérica es realizada bajo control de un flujo que comprende al menos un descriptor de flujo, que transporta a su vez las informaciones necesarias para la configuración de la cadena de descodificación con el descodificador apropiado.

5 Según una variante, la etapa de construcción de dicha secuencia es realizada en un descodificador dotado del mismo enlace hacia el dispositivo de composición que un descodificador BIFS ordinario para ejecutar los BIFS-Commands descodificados sobre la escena sin pasar por un tampón de composición.

10 Según una variante, el nodo BIFS que realiza la primera etapa de construcción comprende un número de campos variable, dependiendo del tipo de periférico de mando utilizado, dichos campos están conectados a los campos de los nodos a modificar por las rutas. El descodificador de interacción transfiere entonces los valores producidos por el periférico en los campos de este nodo BIFS, encargándose el mecanismo de las rutas de propagar estos valores a los campos objetivo.

15 Según un modo de puesta en práctica particular, el flujo de datos de interacción mono o multiusuario pasa por un cliente DMIF asociado al dispositivo que genera las unidades de acceso a situar en la memoria tampón de descodificación unida al descodificador correspondiente.

20 Según un ejemplo específico, el flujo de interacción mono o multiusuario entra en el descodificador correspondiente, bien directamente, o bien a través de la memoria tampón de descodificación asociada, acortando así el trayecto seguido por el flujo de interacción del usuario.

25 El invento se refiere igualmente a un equipo informático que comprende una calculadora para la ejecución de una aplicación multimedia que explota la norma MPEG-4 y al menos un periférico para la representación de una escena multimedia, así como al menos un periférico para mandar llamado programa caracterizado porque comprende además un circuito de enlace que comprende un circuito de entrada para recibir las señales de un medio de mando y un circuito de salida para entregar una secuencia numérica, y un medio para la construcción de una secuencia de salida en función de las señales suministradas por el periférico de entrada, conforme al procedimiento seguido.

30 El invento será mejor comprendido con la lectura de la descripción que sigue, relativo a un ejemplo no limitativo de realización, que se refiere a los dibujos adjuntos, donde:

La fig. 1 representa el organigrama del modelo de descodificador del sistema;

35 La fig. 2 representa el flujo de datos de interacción del usuario.

La fig. 1 describe el modelo estándar. La fig. 2 describe el modelo en el que aparecen dos conceptos principales: el descodificador de interacción que produce unidades de composición (CU) y el flujo de interacción de usuario. Los datos pueden, bien provenir de la memoria tampón de descodificación situada en una unidad de acceso (AU), si el acceso al gestor de dispositivo de entrada es efectuado por medio de DMIF (Estructura de Integración Multimedia de Entrega) de la norma MPEG-4, o bien pasar directamente del dispositivo de entrada al propio descodificador, si la puesta en práctica es tal que el descodificador y el gestor de dispositivo de entrada están situados en el mismo componente. En este último caso, la memoria tampón de descodificación es inútil.

45 Para gestionar la interacción del usuario, los elementos siguientes son necesarios:0

- un nuevo tipo de flujo, teniendo en cuenta los datos de interacción del usuario (UI);
- 50 - un nuevo nodo único BIFS para especificar la asociación entre el flujo de interacciones del usuario y los elementos de la escena e, igualmente, para autorizar o prohibir esta interacción; y,
- un nuevo tipo de descodificador, para interpretar los datos que provienen del dispositivo de entrada, o quizás de la memoria tampón de descodificación, y transformarlos en modificaciones de escena. Estas modificaciones tienen el mismo formato que las órdenes o comandos BIFS. Dicho de otro modo, la salida del descodificador de interacción es exactamente equivalente a la salida de un descodificador BIFS.
- 55

El nuevo tipo de flujo, llamado flujo de interacción de usuario (flujo UI, véase tabla 1) está definido aquí. Está compuesto de unidades de acceso (AU) que provienen de un dispositivo de entrada (por ejemplo, un ratón, un teclado, un guante instrumentalizado,...). Con el fin de ser más genérico, la sintaxis de una unidad de acceso no está definida aquí. Podría ser, sin estar limitada a ello, idéntica a otra unidad de acceso que proviene de otro flujo elemental, si el acceso se hace por DMIF. El tipo de flujo especificado aquí comprende igualmente el caso de un dispositivo local de creación de medio utilizado como dispositivo de interacción. Así, un dispositivo local que produce no importa qué tipo de objeto definido por la indicación de tipo de objeto (Indicación de Tipo de Objeto) de MPEG-4, tal como un objeto visual o audio, es generado por la presente mejora del modelo.

65

ES 2 291 451 T3

La sintaxis del nuevo nodo BIFS, llamada InputSensor, es la siguiente:

5	InputSensor			
	{			
	ExposedFiel	SFBool	Enabled	TRUE
10	d			E
	ExposedFiel	SFCommandBuffe	InteractionBuffe	[]
	d	r	r	
15	Field	SFUrl	url	" "
	EventOut	SFBool	lsActive	
	}			
20				

El campo "habilitado" ("enabled") permite controlar si el usuario desea o no autorizar la interacción que proviene de los flujos de interacción de usuario referenciados en el campo "url". Este último especifica el flujo elemental a utilizar, como se ha descrito en la plataforma de descripción de objeto de la norma MPEG-4.

El campo "InteractionBuffer" es un SFCommandBuffer que describe lo que el descodificador debe hacer del flujo de interacción especificado en "url". La sintaxis no es obligatoria pero la semántica de la memoria tampón es descrita por el ejemplo siguiente:

30	InputSensor{	
	Enabled	TRUE
35	InteractionBuffe	["REEMPLAZAR N1.size", "REEMPLAZAR N2.size"
	r	, "REEMPLAZAR N3.size"
	url	"4"
40	}	

Este captador recupera al menos tres parámetros que provienen del dispositivo de entrada asociado al descriptor de objeto 4 y reemplaza respectivamente, el campo "tamaño" ("size") de los nodos N1, N2 Y N3 por los parámetros recibidos.

La función del descodificador de interacción de usuario es transformar las unidades de acceso recibidas, bien provenientes de la memoria tampón de descodificación, o bien directamente del dispositivo de entrada. Las transforma en unidades de composición (CU) y las coloca en la Memoria de composición (CM), como es especificado por la norma MPEG-4. Las unidades de composición generadas por el descodificador del flujo de interacciones de usuario son BIFS-Updates, más específicamente las órdenes o comandos REEMPLAZAR, como se ha especificado por los sistemas MPEG-4. La sintaxis es estrictamente idéntica a la definida por la norma MPEG-4 y deducida de la memoria tampón de interacción.

Por ejemplo, si el dispositivo de entrada ha generado el entero 3 y si la memoria tampón de interacción contiene "REEMPLAZAR N1.size", entonces la unidad de composición será la BIFS-Update descodificada equivalente a "REEMPLAZAR N1.size por 3".

Una de las variantes mencionadas consiste en reemplazar el campo InteractionBuffer del nodo InputSensor por un número de campos variable, dependiendo del tipo de periférico de mando utilizado, del tipo EventOut. La misión del descodificador de interacción de usuario es entonces modificar los valores de estos campos, encarga al autor de la presentación multimedia crear rutas que conectan los campos del nodo InputSensor a los campos objetivo en el árbol de escena.

REIVINDICACIONES

- 5 1. Un procedimiento para la gestión de interacciones entre al menos un periférico de comando y al menos una aplicación multimedia que explota la norma MPEG-4, entregando dicho periférico de mando señales de función de acción de uno o varios usuarios, **caracterizado** porque comprende una etapa de construcción de una secuencia numérica que presenta la forma de un nodo BIFS [Formato Binario para Escena en el sentido de la norma MPEG-4], comprendiendo dicho nodo al menos un campo que define el tipo y el número de datos de interacción a aplicar en los objetos de la escena.
- 10 2. Un procedimiento para la gestión de interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según la reivindicación 1ª **caracterizado** porque la transferencia de dicha secuencia numérica utiliza la cadena de descodificación de los sistemas MPEG-4 para introducir informaciones de interacción en el dispositivo de composición.
- 15 3. Un procedimiento para la gestión, de las interacciones entre al menos un periférico de mando y al menos una aplicación multimedia que explota la norma MPEG-4, según la reivindicación 1ª o 2ª **caracterizado** porque comprende una etapa de designación de la naturaleza de la acción o de las acciones a ejercer sobre uno o varios objetos de la escena por medio del o de los campos del nodo.
- 20 4. Un procedimiento para la gestión de interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según la reivindicación 1ª o 2ª **caracterizado** porque el nodo BIFS comprende un número de campos variable, dependiendo del tipo de periférico de mando utilizado, y porque la transferencia de datos de interacción de los campos de este nodo hacia los campos objetivo es hecha por medio de rutas.
- 25 5. Un procedimiento para la gestión de interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según la reivindicación 1ª o 2ª **caracterizado** porque comprende una etapa de señalización de la actividad del dispositivo asociado.
- 30 6. Un procedimiento para la gestión de interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según una cualquiera de las reivindicaciones precedentes **caracterizado** porque la etapa de transferencia de la secuencia es realizada en forma de un flujo señalado para un descriptor que contiene las informaciones necesarias para la configuración de la cadena de descodificación con el descodificador apropiado.
- 35 7. Un procedimiento para la gestión de interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según una cualquiera de las reivindicaciones precedentes **caracterizado** porque la etapa de construcción de la secuencia de datos de interacción es realizada en una memoria tampón de descodificación del terminal de ejecución de la aplicación multimedia.
- 40 8. Un procedimiento para la gestión de interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según una cualquiera de las reivindicaciones 1ª a 7ª **caracterizado** porque la etapa de traducción de la secuencia de datos de interacción es realizada en un descodificador dotado de un enlace hacia el dispositivo de composición similar al de un descodificador BIFS ordinario para ejecutar los BIFS-Commands descodificados en la escena.
- 45 9. Un procedimiento para la gestión de interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según una cualquiera de las reivindicaciones precedentes **caracterizado** porque el flujo de interacciones del usuario pasa por un cliente DMIF asociado al dispositivo que genera las unidades de acceso a colocar en la memoria tampón de descodificación unida al descodificador correspondiente.
- 50 10. Un procedimiento para la gestión de las interacciones entre un periférico de mando y una aplicación informática que explota la norma MPEG-4 según una cualquiera de las reivindicaciones precedentes **caracterizado** porque el flujo de interacciones del usuario entra en el descodificador correspondiente, bien directamente, o bien a través de la memoria tampón de descodificación asociada, acortando así el trayecto seguido por el flujo de interacciones del usuario.
- 55 11. Un equipamiento informático que comprende una calculadora para la ejecución de una aplicación multimedia que explota la norma MPEG-4 y al menos un periférico para la representación de una escena multimedia, así como al menos un periférico para mando llamado programa **caracterizado** porque comprende además un circuito de enlace que comprende un circuito de entrada para recibir las señales de un medio de mando y un circuito de salida para entregar una secuencia de tipo BIFS, y un medio para la construcción de una secuencia de salida en función de las señales suministradas por el periférico de entrada, conforme al procedimiento según la reivindicación 1ª.
- 60
- 65

Fig. 1

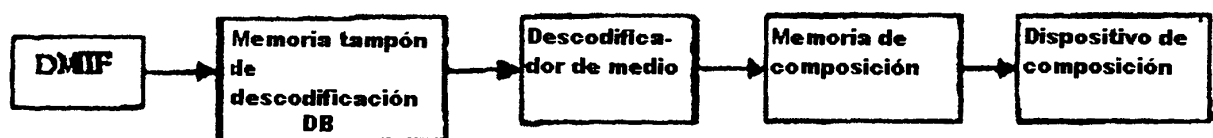


Fig. 2

