

MPEG-4 sobre Redes Inalámbricas

María E ÁLVAREZ

Departamento de Comunicaciones, Escuela de Ingeniería Eléctrica
Universidad Central de Venezuela, Caracas, Venezuela

y

Luis J. FERNÁNDEZ

Departamento de Comunicaciones, Escuela de Ingeniería Eléctrica
Universidad Central de Venezuela, Caracas, Venezuela

RESUMEN

MPEG-4 es un protocolo que se basa en el manejo de objetos de imágenes tanto simples como compuestas y sus aplicaciones tienen especial importancia en las transmisiones inalámbricas, particularmente referidas a la tecnología de tercera generación "3G", pues, además de la compresión en el ancho de banda, proporciona un buen control de errores, punto que servirá de base para la realización de un modelo de simulación que proporcionará una pre-visualización del comportamiento y todas las características esenciales entre un usuario móvil y una radio base.

Palabras Claves: 3G, inalámbricas.

1. INTRODUCCIÓN

MPEG-4, es un estándar que permite la codificación por objetos, y estructuras jerárquicas, mejora el control de errores, dando paso a lo que se denomina multimedia móvil que proporciona la movilidad a través de redes multiservicio y de acceso múltiple para el usuario. Esto permitirá el inicio del desarrollo de un modelo de simulación entre un usuario móvil de 3G y una radio base, tal que, se puedan estudiar ciertos aspectos que generen inconvenientes, en el establecimiento, mantenimiento, y finalización de la comunicación. Para la realización de dicho modelo se estudian los aspectos que afectan las comunicaciones móviles, tales como: ancho de banda menor que en sistemas cableados, ruido e interferencia, asignación de frecuencias (Handoff) y seguridad.

Inicialmente se explicará el protocolo MPEG-4, y se analizará su comportamiento en los sistemas de tercera generación "3G" basándose en aspectos relacionados con mejoras de: ancho de banda asociado, características de compresión, algoritmos matemáticos, control de errores, susceptibilidad a ruido e interferencia, etc, tal que se compare con diferentes tipos de compresión de video y se obtenga unas conclusiones certeras.

Una vez obtenidos los resultados de primer análisis, en los sucesivos estudios se explicará la escogencia de parámetros determinados; tales que sean el fundamento

para la realización de la herramienta de modelaje. Tales parámetros poseerán todas las características específicas que definan el comportamiento del sistema el cual se desea modelar (usuario-radiobase).

2. MPEG-4

MPEG-4 (Moving Picture Expert Group 4) tal y como se define en ISO/IEC 14496-2, es el estándar usado para proporcionar a los usuarios un nuevo nivel de interacción con información de tipo visual, basándose en tecnología para visualizar, acceder y manipular objetos en lugar de píxeles, además de presentar una gran robustez frente a errores en un amplio rango de anchos de banda. El estándar proporciona herramientas para codificación de formas, estimación y compensación de movimiento, codificación de texturas, recuperación frente a errores, codificación de "sprites" y escalabilidad. La base para la interoperabilidad viene dada por diversos puntos de conformidad definidos en el estándar en forma de tipos de objetos, perfiles y niveles.

La codificación de formas puede llevarse a cabo en modo binario, donde la forma de cada objeto se describe mediante una máscara binaria, o en modo de escala de grises, donde cada figura se define de forma similar a un canal alfa, técnica que permite la gestión de valores de transparencia y efectos de "anti-aliasing".

La compensación de movimiento está basada en bloques, contando con las modificaciones necesarias para prever los límites de los objetos. El tamaño de bloque puede ser 16x16 u 8x8, con resolución de medio píxel. MPEG-4 también proporciona un modo para compensación de movimiento superpuesto. Para la codificación de texturas se utiliza DCT (*Transformada Coseno Discreta*) modificada para los límites de los objetos. Es posible implementar predicción de coeficientes para mejorar la eficiencia. Las texturas estáticas pueden ser codificadas utilizando una transformada *wavelet*.

La tolerancia a errores se consigue mediante marcas de resincronización, particionamiento de datos, extensiones en los códigos de las cabeceras y códigos de longitud variable reversibles.

La escalabilidad permite mejorar tanto la resolución espacial como temporal. La escalabilidad en MPEG-4 se basa en la concepción de objetos, con la restricción de que la forma del objeto debe ser rectangular. Los puntos de conformidad se definen mediante perfiles y vienen detallados en el *Perfil Simple* (Simple Profile), *Perfil Interno* (Core Profile) y en el *Perfil Principal* (Main Profile). El estándar visual MPEG-4 consiste en una serie de herramientas que permiten a las

aplicaciones distintos tipos de funcionalidades. El estándar MPEG-4 ha sido optimizado para tres anchos de banda:

1. Por debajo de 64 kbit/s
2. 64-384 kbit/s
3. 384 - 4 Mbit/s

En cualquier caso, tanto para anchos de banda que permiten más calidad como para anchos de banda bajos, el conjunto de herramientas que se utilizan son las mismas.

MPEG-4 proporciona soporte tanto para material progresivo como entrelazado. El formato de crominancia soportado es 4:2:0. Cada componente puede representarse por un número de bits entre 4 y 12.

El concepto central en MPEG-4, como se desprende de los apartados anteriores, es el objeto audiovisual, que sustenta la representación basada en objetos. Esta representación es bastante apropiada para aplicaciones interactivas y permite acceso directo a los contenidos de una escena.

Una escena audiovisual MPEG-4 está formada por uno o más objetos visuales. Cada objeto de vídeo se caracteriza por cierta información temporal y espacial sobre la forma, movimiento y textura del objeto.

Los niveles jerárquicos que describen la escena más directamente son:

Secuencia de objetos visuales (SOV). La escena MPEG-4 completa, que puede contener cualquier objeto natural o sintético en 2D ó 3D y sus correspondientes capas.

Objeto de vídeo (OV). Un objeto de vídeo se corresponde a un objeto 2D particular de la escena. Generalmente se trata de un frame.

Capa de Objeto de Vídeo (COV). Cada objeto de vídeo puede codificarse de una forma escalable (multi-capa) o no escalable (única capa), dependiendo de la aplicación, representada por la capa de objeto de vídeo (COV). La COV es la base para la codificación escalable. Un objeto de vídeo puede codificarse usando escalabilidad espacial o temporal. Dependiendo de parámetros como el ancho de banda disponible, la potencia de computación o las preferencias del usuario, la resolución deseada debe estar disponible para el decodificador.

Grupo de planos de un objeto de vídeo (GPOV). El GPOV agrupa planos de objeto de vídeo. Los GPOV pueden proporcionar puntos en el flujo donde los planos de objeto de vídeo sean codificados unos independientemente de otros, para así proporcionar puntos de acceso aleatorio o directo al flujo. Los GPOV son opcionales.

Plano de objeto de vídeo (POV). Un POV es una muestra en el tiempo de un objeto de vídeo. Los POV pueden codificarse independiente o dependientemente unos de otros usando compensación de movimiento. Un frame de vídeo convencional puede ser representado por un POV con forma rectangular

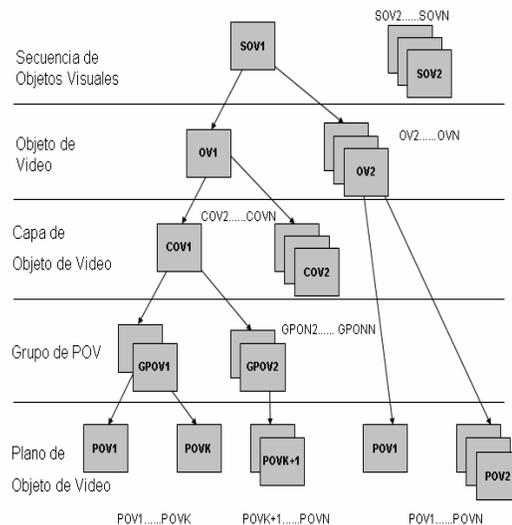


Figura I: Niveles jerárquicos de escena.

3. CARACTERÍSTICAS DE COMPRESIÓN DE MPEG-4.

El estándar visual MPEG-4 contempla dos tipos de información acerca de la forma como características inherentes de un objeto de vídeo: información binaria e información en escala de grises sobre forma. La información sobre forma en binario establece qué porción (píxeles) del soporte de un objeto pertenecen a un objeto de vídeo dado. La información binaria se representa generalmente mediante una matriz del tamaño del rectángulo que acota un POV. Cada elemento de la matriz tiene uno de los dos valores posibles, dependiendo de si el píxel pertenece o no al objeto de vídeo. La información en escala de grises es una generalización sobre el concepto de forma binaria y permite representar objetos transparentes.

Codificación binaria de forma

MPEG adoptó la técnica de compresión basada en mapa de bits para la información sobre forma. Se codifica mediante una técnica de compensación de movimiento basada en bloques que permite codificación tanto con pérdida como sin pérdida. En el algoritmo de compresión de vídeo de MPEG-4, la forma de cada POV es acotada por un marco rectangular de tamaño igual a un múltiplo de un rectángulo de 16x16 píxeles. La posición de este rectángulo de acotación puede elegirse de manera que contenga el mínimo número de bloques de 16x16 con píxeles transparentes. Los valores transparentes (0) son asignados a los píxeles que no pertenecen al POV y a los que definen su contorno. Sobre estos bloques de 16x16 en que se divide el marco de acotación se llevará a cabo (bloque a bloque) el proceso de codificación/decodificación.

La matriz binaria que representa la forma de un POV suele denominarse *máscara binaria*. En esta máscara, cada píxel que pertenece al POV tiene como valor 255 y el resto de los píxeles se ponen a 0 (transparentes). Después, esta matriz se divide en *bloques alfa binarios* (BAB) de tamaño 16x16, que son codificados por separado. Dado que estos BAB son marcos

cuadrados, es frecuente que todos los píxeles de un BAB tengan el mismo valor, 0 en cuyo caso se habla de un BAB o bloque transparente ó 255 que determina un bloque opaco. El algoritmo de compresión de formas prevé distintos modos para codificar un BAB. Las herramientas básicas para codificar un BAB son el algoritmo de Codificación Aritmética basada en el Contexto (CAE) y las técnicas de compensación de movimiento. Los vectores de movimiento pueden calcularse buscando la posición que más concuerde (aquella que minimice la suma de diferencias absolutas) y son codificados diferencialmente.

Codificación de forma con escala de grises

La información de forma en escala de grises tiene una estructura muy similar a la binaria, con la diferencia de que cada píxel (elemento de la matriz) puede tomar valor en un rango, generalmente 0-255, que representan el grado de transparencia de cada píxel. La escala de grises sobre forma se corresponde con la noción de plano alfa utilizada en informática gráfica, en la que el 0 corresponde a un píxel transparente y el 255 a un píxel completamente opaco.

La información en escala de grises se codifica mediante DCD basada en bloques con compensación de movimiento similar a la codificación de texturas, permitiendo solamente codificación con pérdida. La codificación de forma con escala de grises también utiliza codificación binaria para su soporte.

4. TOLERANCIA A FALLOS

MPEG-4 otorga suficientes mecanismos para permitir tolerancia a fallos con diferentes niveles de robustez y complejidad. Estos mecanismos se llevan a cabo a través de herramientas que proporcionan medios para la resincronización, detección de errores, recuperación de datos y ocultamiento de errores. La parte visual de MPEG-4 ofrece cuatro herramientas para la tolerancia a fallos. Éstas son las siguientes:

Resincronización. Consiste en insertar marcas en el flujo de manera que, en caso de error, el decodificador puede avanzar hasta la siguiente marca y saltar los bits intermedios. Situado sobre esta marca, el decodificador puede continuar con la decodificación. MPEG-4 permite insertar marcas de resincronización después de un número más o menos constante de bits codificados (paquetes de vídeo), a diferencia de MPEG-2 y H.263 que sólo permitían resincronización después de un número constante de macrobloques codificados (normalmente una fila de macrobloques). Distintos experimentos han mostrado que la resincronización propuesta en MPEG-4 es una forma más eficiente de recuperación para errores en la transmisión.

Particionamiento de datos. Este método separa los bits para la codificación de la información sobre movimiento de aquellos destinados a codificación de texturas. En caso de error, es posible soslayar de forma más eficiente un fallo ocurrido sólo en los bits de textura.

Extensión del código en cabeceras. Estos códigos binarios permiten incluir opcionalmente información de cabecera redundante, vital para la correcta decodificación de vídeo. De

esta forma, se reducirán las posibilidades de corrupción de la información de cabecera y la necesidad de saltar largas secuencias del flujo.

VLCs reversibles (RVLCs). Estos VLCs permiten reducir aún más la influencia de errores en los datos decodificados. Los RVLCs son códigos que pueden ser decodificados tanto hacia adelante como hacia atrás. En caso de que suceda un error y sea necesario saltar hasta la siguiente marca de resincronización, es posible todavía decodificar partes del flujo corrupto en orden inverso y así limitar la influencia del error.

5. PROYECCIÓN

Analizadas estas características del estándar MPEG-4, y basándonos en los niveles de capa 1 y 2 del modelo ISO/OSI, se prevé el soporte para que se pueda llevar a cabo el manejo de la información multimedia a través de la red inalámbrica. Los estudios que se están llevando actualmente aunando al análisis de diferentes programas de simulación permitirán determinar los parámetros necesarios para el desarrollo de la herramienta de modelaje.

6. REFERENCIAS

L. Chariglione, "MPEG and Multimedia Communications", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 7, No. 1, pp. 5-18, Feb. 1997.

MPEG-4 Video Group, "Generic Coding of Audio-Visual Objects: Part 2 - Visual," ISO/IEC JTC1/SC29/WG11 N1902, FDIS of ISO/IEC 14496-2, Atlantic City, Nov. 1998

C. Le Buan, F. Bossen, S. Bhattacharjee, F. Jordan, T. Ebrahimi, "Shape representation and coding of visual objects in multimedia applications - An Overview", (invited paper), in Annales des Telecommunications 53, No 5-6, pp. 164-178, May 1998.