

Estimación y Compensación de Movimientos en Video Codificado

*Claudia Cecilia Russo*¹ - *Hugo Ramón*²

LIDI - Laboratorio de Investigación y Desarrollo en Informática
Facultad de Informática
*Universidad Nacional de La Plata*³

Resumen

Los sistemas distribuidos multimedia requieren transferencia de datos continua sobre periodos de tiempo relativamente altos, sincronización en el manejo de los diferentes tipos de datos (ejemplo: voz, sonido y video), espacios de almacenamiento extremadamente grandes, manejo de tiempo real y técnicas especiales de indexación y recuperación de los datos de tipo multimedia, además de otros problemas que surgen a partir de éstos.

En los últimos años han ido apareciendo nuevas aplicaciones en el mundo de las telecomunicaciones, que requieren una continua mejora de los equipos terminales y de los canales de comunicación. El desarrollo y la explotación de nuevos sistemas de comunicación como la videoconferencia o la videotelefonía, deben tratar de compaginarse con la necesidad de aprovechar canales de transmisión de baja capacidad, que en un principio fueron diseñados para transmitir voz o texto. La necesidad anterior ha impulsado el diseño de técnicas de codificación que permitan trabajar con unas tasas de transmisión restringidas, manteniendo una *calidad aceptable*[1], para ello se hace hincapié en el estudio de técnicas de segmentación como también en la implementación de métodos de estimación y compensación de movimiento.

Esa creciente necesidad de incrementar la interconexión de las cada vez más poderosas estaciones de trabajo multimedia da como resultado una evolución de las comunicaciones en búsqueda de las redes (sus características) que soporten la transmisión de este tipo de información multimedia.

Palabras Claves: Compensación de movimiento, estimación de movimiento, segmentación, video, MPEG, Tiempo Real.

¹ Magister en Informática - Profesor Adjunto Dedicación Exclusiva, Facultad de Informática, U.N.L.P. E-mail: crusso@lidi.info.unlp.edu.ar

² Magister en Informática - Profesor Adjunto Dedicación Exclusiva, Facultad de Informática, U.N.L.P. E-mail: hramon@lidi.info.unlp.edu.ar

³ Calle 50 y 115 Primer Piso, La Plata (1900), Argentina, Pcia. de Bs. As. Teléfono 54-21-227707 lidi@ada.info.unlp.edu.ar

Introducción

Las técnicas de compresión de datos juegan un papel fundamental en el desarrollo de las aplicaciones multimedia debido a los grandes requerimientos de espacio que exige el audio, las imágenes, la animación y el video como a los limitados anchos de banda de las redes que no permiten la transmisión de altos volúmenes de datos multimedia (como el video) en tiempo real.

Los métodos de compresión, aprovecha la redundancia espacial de una imagen (áreas uniformes), la correlación entre puntos cercanos y la menor sensibilidad del ojo a los detalles finos de las imágenes fijas (JPEG) y, para imágenes animadas (MPEG), se saca provecho también de la redundancia temporal entre imágenes sucesivas[1] [2] [3].

Se pueden obtener grandes factores de compresión teniendo en cuenta la redundancia entre imágenes sucesivas. Esto involucra al eje del tiempo. Este proceso se denomina codificación inter o codificación temporal.

La codificación temporal permite altos factores de compresión, pero con la desventaja de que una imagen individual existe en términos de la diferencia entre imágenes previas. Si una imagen previa es quitada en la edición, entonces los datos de diferencia pueden ser insuficientes para recrear la siguiente imagen. El estándar ISO MPEG (Motion Pictures Experts Group) utiliza esta técnica.

Compresión de Video en el estándar MPEG [4]

En el año de 1990, la ISO, preocupada por la necesidad de almacenar y reproducir imágenes de video digitales y su sonido estereofónico correspondiente, creó un grupo de expertos que llamó MPEG (Moving Pictures Expert Group) procedentes de aquellas áreas implicadas en el problema (telecomunicaciones, informática, electrónica, radio difusión, etc).

El primer trabajo de este grupo se conoció como la norma ISO/IEC 11172, mucho más conocida como MPEG-1, en el año 1992. La idea inicial era la de permitir el almacenamiento y reproducción en soporte CD-ROM con un flujo de transmisión de datos del orden de 1,5 Mbits/s, transportando tanto imagen como sonido.

El estándar MPEG además de aprovechar la redundancia espacial intrínseca de una imagen fija utilizada en la codificación JPEG, aprovecha la redundancia temporal que aparece en la codificación de imágenes animadas, permitiendo encontrar similitudes entre las imágenes sucesivas de video. [5]

Debido a que la calidad en la compresión de video en el estándar MPEG-1 era de baja calidad y no servía para otras aplicaciones, se creó la norma ISO/IEC 13818, mucho más conocida con el nombre de MPEG-2. Esta norma permite un flujo de transmisión hasta el orden de los 20 Mbits/s, transportando tanto imagen como sonido. Norma que se utilizaría en la televisión de alta definición.

En la actualidad, se está trabajando en una norma llamada MPEG-4 y está encaminada a la transmisión de datos del orden de los 8 a 32 Kbits/s, norma que será utilizada en las aplicaciones de video conferencia o video teléfono.

La compresión de video utiliza los mismos principios que JPEG con pérdidas, a la que se le añaden nuevas técnicas que, juntas, forman el MPEG-1, que permiten reducir considerablemente la cantidad de información necesaria para la transmisión de imágenes sucesivas muy correlacionadas temporalmente. [7]

Estas técnicas, llamadas de "predicción con compensación de movimiento", consisten en reducir, con un mínimo de información adicional, la mayoría de las imágenes precedentes (incluso las que le siguen).

Esto requiere un dispositivo de estimación de movimiento en el decodificador, que es la parte más compleja.

Tratándose de imágenes en movimiento o animadas, la descompresión deberá poder hacerse en "tiempo real" durante la reproducción. Por otro lado, la necesidad de un tiempo de sincronización y de una respuesta de acceso aleatorio a una secuencia no demasiado largos (0.5 segundos máximo) limita el número de imágenes que pueden depender de la misma primera imagen a diez o doce para un sistema de 25 imágenes por segundo. [6]

Es de nuestro interés poner énfasis en el estudio de la optimización de algunos de los procesos involucrados en la compresión de video como la segmentación que será la base para aplicar los métodos de compensación y estimación de movimiento.

Compensación y Estimación de Movimiento

La compensación de movimientos [7] es un proceso mediante el cual se mide eficazmente el movimiento de los objetos de una imagen a otra. De este modo se consigue medir qué tipos de movimientos redundan entre imágenes.

Las técnicas, llamadas de "predicción con compensación de movimiento", consisten en reducir, con un mínimo de información adicional, la mayoría de las imágenes precedentes (incluso las que le siguen).

El objetivo de la compensación de movimiento es la reconstrucción del cuadro en proceso, empleando el cuadro de referencia y la información de movimiento.

Esto requiere un dispositivo de estimación de movimiento en el decodificador, que es la parte más compleja.

Tratándose de imágenes en movimiento o animadas, la descompresión deberá poder hacerse en "tiempo real" durante la reproducción. Por otro lado, la necesidad de un tiempo de sincronización y de una respuesta de acceso aleatorio a una secuencia no demasiado largos (0.5 segundos máximo) limita el número de imágenes que pueden depender de la misma primera imagen a diez o doce para un sistema de 25 imágenes por segundo.

La estimación de movimiento [8][9] es uno de los problemas fundamentales en el tratamiento de vídeo digital.

Su objetivo es calcular el campo de vectores que describe el movimiento aparente entre dos cuadros de una secuencia.

Es importante hablar de movimiento aparente, porque no hay que olvidar que las señales de imagen en movimiento son la proyección en el plano, en instantes discretos de tiempo, de escenas tridimensionales. Esto supone una pérdida de información que hace necesario distinguir entre el movimiento real que se proyecta sobre el plano y el movimiento aparente que, junto con su velocidad, es lo que el observador percibe mediante la variación del contenido visual de las imágenes de la secuencia. Una consecuencia de lo anterior es que, en situaciones en las que hay cambios en la iluminación o existe una ausencia del suficiente gradiente espacial en la imagen, el movimiento real y el flujo óptico (velocidad aparente) no tienen por qué coincidir.

Temas de investigación

Nuestro estudio se centra en:

- las técnicas de estimación y compensación
- el análisis de la segmentación de las imágenes para lograr un mayor nivel de estimación y compensación, para esto se estudia la segmentación por bloques fijos, por bloques variables.
- la segmentación basada en regiones tomando los objetos dentro de la imagen.
- estudio de la complejidad de los algoritmos de búsqueda y matching.
- eficiencia y tiempo de respuesta sobre canales reales.

Bibliografía

[1] Introduction to Data Compression. Khalid Sayood, Morgan Kaufmann 1996.

[2] Digital Image Processing, Bernd Jahne, Springer 1997.

[3] Digital Image Processing Algorithms, Ioannis Pitas, Prentice Hall 1995.

[4] Digital Video Processing, A. Murat Tekalp, Prentice Hall Signal Processing Series 1995.

[5] Digital Compression of Still Images and Video, R. J. Clake, Academic Press 1995.

[6] Image and Video Compression, Antonio Ortega and Kannan Ramchandran, IEEE Signal Processing Nov. 1998.

[7] Video Compression, Gary J. Sullivan and Thomas Wiegand, IEEE Signal Processing Nov. 1998.

[8] Estimating Motion in Image Sequences, Christoph Stiller and Januz Konrad, IEEE Signal Processing, July. 1999.

[9] Correspondence Estimation in Image Pairs, Andre Redert, Emile Hendriks and Jan Biemond, IEEE Signal Processing, May 1999.

[10] Colección IEEE