

## "Estimación y Compensación de Movimientos en Video Codificado"

*Magister Claudia Cecilia Russo (UNLP)<sup>1</sup>*

*Magister Hugo Ramón (UNLP)<sup>2</sup>*

### Resumen

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.

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

### Introducción

La compresión de video surge de la necesidad de transmitir imágenes a través de un canal que contenga un ancho de banda aceptable. Los métodos de compresión, recurren a los procedimientos generales de compresión de datos, aprovechando además 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].

Cuando las imágenes individuales son comprimidas sin referencia a las demás, el eje del tiempo no entra en el proceso de compresión, esto por lo tanto se denomina codificación intra (intra=dentro) o codificación espacial. A medida que la codificación espacial trata cada imagen independientemente, esta puede emplear ciertas técnicas de compresión desarrolladas para las imágenes fijas. El estándar de compresión ISO (International Standards Organization) JPEG (Joint Photographic Experts Group), está en esta categoría. Donde una sucesión de imágenes codificadas en JPEG también se usan para la televisión, esto es llamado "JPEG en movimiento"[4].

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 (inter=entre) 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

---

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

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

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).

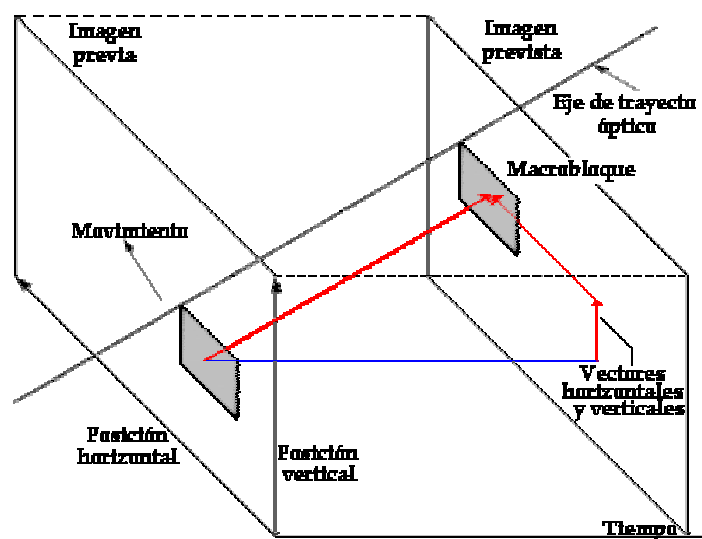
Esto requiere un dispositivo de estimación de movimiento en el decodificador, que es la parte más compleja.

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 de movimiento [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. La siguiente muestra que la imagen en movimiento puede ser expresada en un espacio de tres dimensiones que resulta de mover a lo largo del eje del tiempo dos imágenes consecutivas.



**Objetos viajando en un espacio de tres dimensiones**

En el caso de un objeto que permanezca estático, el movimiento de este solo se ve en el eje del tiempo. Sin embargo, cuando un objeto está en movimiento, este se mueve en el eje de trayecto óptico (eje horizontal y vertical en el tiempo) que no es paralelo al eje del tiempo.

El eje de trayecto óptico une los puntos de movimiento de un objeto a medida que este se mueve a través de varias imágenes. Este trayecto puede ser nulo en valores que representen un objeto en movimiento que solo cambia con respecto al eje del tiempo. De igual manera, al mirar un objeto en movimiento que cambia su apariencia; uno de estos movimientos es el de rotación.

Para un simple movimiento de traslación los datos que representan un objeto son altamente redundantes con respecto al eje de trayecto óptico. Así, que de este modo, el eje de trayecto óptico puede ser hallado, generando un código de ganancia cada vez que se observa la presencia de un mismo objeto en movimiento.

Un codificador de compensación de movimiento trabaja de la siguiente forma. Una imagen I es enviada, pero esta es almacenada de tal modo que pueda ser comparada con la siguiente imagen de entrada para encontrar así varios vectores de movimiento, los cuales pueden ser utilizados en diferentes áreas de la imagen. Luego la imagen I es combinada de acuerdo a

estos vectores o cancelada a una codificación espacial debido a su no conveniencia. La imagen prevista resultante es comparada con la imagen actual para producir una predicción de error también llamada residual. La predicción de error es transmitida con los vectores de movimiento. En el receptor la imagen I original es también retenida en la memoria, esta es cambiada de acuerdo con los vectores de movimiento transmitidos para crear la imagen prevista y luego la predicción de error es adicionada recreando la imagen original. Cuando una imagen es codificada de esta manera, es llamada imagen P en MPEG.

### **Métodos de estimación [8] [9]**

Existen varios criterios para clasificar los diferentes métodos de estimación.

Podemos distinguir entre métodos que emplean la imagen en proceso y los que no la emplean si estamos trabajando con información la cual se emplea para el cálculo de vectores.

Entre los métodos que no emplean la imagen en proceso se encuentran las técnicas recursivas, basadas de forma directa en la ecuación de flujo óptico y que poseen la ventaja de que no necesitan transmitir ninguna información de movimiento; el precio que pagan dichas técnicas, es una predicción con un error mayor que el que se obtiene con el otro grupo de técnicas. Por otro lado, las técnicas que emplean para realizar la estimación información sobre el propio cuadro en proceso, necesitan transmitir información de movimiento al decodificador, pero obtienen una predicción mucho mejor que las anteriores, por lo que el análisis se ajustará a éstas.

Otra posible clasificación de los métodos de estimación, se realiza de acuerdo a si llevan a cabo una estimación del movimiento del plano de la imagen (movimiento 2D) o bien una estimación del movimiento de objetos (movimiento 3D, donde el tiempo es la tercera dimensión).

### **Sentido de la estimación de movimiento. [9]**

La estimación de movimiento trata de obtener los vectores que describen el movimiento aparente entre dos cuadros de una secuencia. Al igual que en la compensación, existen dos posibilidades para su aplicación:

La estimación hacia adelante consiste en la selección de puntos en la imagen de referencia. Dicha imagen esta disponible tanto en el codificador como en el decodificador, por lo que si se emplea en ambos extremos el mismo mecanismo de obtención de esos puntos, sólo resulta necesaria la transmisión de los vectores de movimiento.

La estimación hacia atrás toma los puntos en la imagen actual, obligando a transmitir al decodificador tanto los vectores como las coordenadas de dichos puntos, porque dicha imagen no existe aún en el decodificador.

### **Temas actuales de estudio e investigación**

Nuestro estudio se centra tanto en las técnicas de estimación y compensación como en 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 quadtree y la segmentación basada en regiones tomando los objetos dentro de la imagen.

### **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 Hendrinks and Jan Biemond, IEEE Signal Processing, May 1999.