

Compresión de video

Sergio Fukushima

Technical Manager South America
de Axis Communications
sergio.fukushima@axis.com



Las técnicas de compresión de video consisten en reducir y eliminar datos redundantes del video para que el archivo de video digital pueda enviarse a través de la red y almacenar en discos informáticos. Este capítulo trata sobre los conceptos básicos de la compresión y proporciona una descripción de los estándares más difundidos.



El material técnico que se publica en este informe fue proporcionado por Axis Communications a Revista Negocios de Seguridad®. Prohibida su reproducción (parcial o total) sin el expreso consentimiento del autor o este medio.

■ Índice

Capítulo 1.

Video en red (RNDS N° 45)

Capítulo 2.

Cámaras de red /Cámaras IP (RNDS N° 46)

Capítulo 3.

Elementos de la cámara (RNDS N° 47)

Capítulo 4.

Protección de la cámara y carcacas (RNDS N° 48)

Capítulo 5.

Codificadores de video (RNDS N° 49)

Capítulo 6.

Resoluciones (RNDS N° 51)

Capítulo 7

Compresión de video

7.1. Conceptos básicos

7.1.1. Códec de video

7.1.2. Compresión de imagen vs. compresión de video

7.2. Formatos de compresión

7.2.1. Motion JPEG

7.2.2. MPEG-4

7.2.3. H.264 o MPEG-4 Part 10/AAVC

7.3. Frecuencia de bits variable y constante

Capítulo 8.

Audio

Capítulo 9.

Tecnologías de red

Capítulo 10.

Tecnología inalámbrica

Capítulo 11.

Sistemas de gestión de video

Capítulo 12.

Consideraciones sobre ancho de banda y almacenamiento

Con técnicas de compresión eficaces se puede reducir considerablemente el tamaño del fichero sin que ello afecte la calidad de la imagen. Sin embargo, la calidad del video puede verse afectada si se reduce en exceso el tamaño del fichero, aumentando el nivel de compresión de la técnica que se utilice.

Existen diferentes técnicas de compresión, tanto patentadas como estándar. En la actualidad la mayoría de los proveedores de video en red utilizan técnicas de compresión estándar.

Los estándares son importantes para asegurar la compatibilidad y la interoperabilidad y tienen un papel especialmente relevante en la compresión de video, puesto que éste se puede utilizar para varias finalidades y, en algunas aplicaciones de videovigilancia, debe poder visualizarse varios años después de su grabación. Gracias al desarrollo de estándares, los usuarios finales tienen la opción de escoger entre diferentes proveedores, en lugar de optar no solo para un sistema de videovigilancia.

Los más difundidos estándares de compresión de video son Motion JPEG, MPEG-4 Parte 2 (MPEG-4) y H.264, siendo este último el estándar más actual y eficaz. Este capítulo trata sobre los conceptos básicos de la compresión y proporciona una descripción de cada uno de los estándares mencionados.

7.1 .Conceptos básicos

7.1.1. Códec de video

En el proceso de compresión se aplica un algoritmo al video original para crear un archivo comprimido y ya listo para ser transmitido o guardado. Para reproducir el archivo comprimido, se aplica el algoritmo inverso y se crea un video que incluye prácticamente el mismo contenido que el video original. El tiempo que se tarda en comprimir, enviar, descomprimir y mostrar un archivo es lo que se denomina "latencia". Cuanto más avanzado sea el algoritmo de compresión, mayor será la latencia.

El par de algoritmos que funcionan conjuntamente se denomina códec de video (codificador/decodificador). Los códecs de video de estándares diferentes no suelen ser compatibles entre sí, es decir que el contenido de video comprimido con un estándar no se puede descomprimir con otro diferente. Por ejemplo, un decodificador MPEG-4 no funcionará con un codificador H.264. Esto ocurre simplemente porque un algoritmo no puede descodificar correctamente los datos de salida del otro algoritmo, pero es posible usar muchos algoritmos diferentes en el mismo software o hardware, que permitirían la coexistencia de varios formatos.

Continúa en página 156

Viene de página 152

7.1.2. Compresión de imagen vs. compresión de video

Los diferentes estándares de compresión utilizan métodos distintos para reducir los datos y, en consecuencia, los resultados en cuanto a frecuencia de bits y latencia son diferentes. Existen dos tipos de algoritmos de compresión: compresión de imágenes y compresión de video.

La compresión de imagen utiliza la tecnología de codificación intrafotograma. Los datos se reducen a un fotograma de imagen con el fin de eliminar la información innecesaria que puede ser imperceptible para el ojo humano. Motion JPEG es un ejemplo de este tipo de estándar de compresión. En una secuencia Motion JPEG, las imágenes se codifican o comprimen como imágenes JPEG individuales.



Con el formato Motion JPEG, las tres imágenes de la secuencia se codifican y se envían como imágenes únicas separadas (fotogramas), sin que dependan unas de otras.

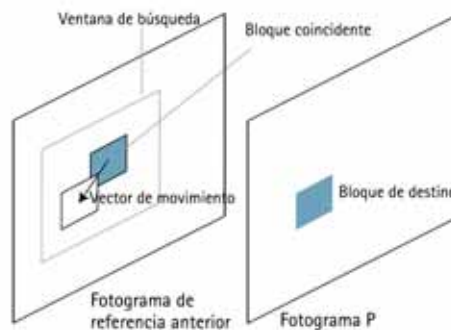
Los algoritmos de compresión de video -4 como el MPEG-4 y el H.264 utilizan la predicción interfotograma para reducir los datos de video entre las series de fotogramas. Esta consiste en técnicas como la codificación diferencial, en la que un fotograma se compara con un fotograma de referencia y sólo se codifican los píxeles que han cambiado con respecto al fotograma de referencia. De esta forma, se reduce el número de valores de píxeles codificados y enviados. Cuando se visualiza una secuencia codificada de este modo, las imágenes aparecen como en la secuencia de video original.



Con la codificación diferencial sólo la primera imagen se codifica en su totalidad. En las dos imágenes siguientes existen referencias a la primera imagen en lo que se refiere a elementos estáticos, como la casa. Sólo se codifican las partes en movimiento (el hombre que corre) mediante vectores de movimiento, reduciendo así la cantidad de información que se envía y almacena.

Para reducir aún más los datos, se pueden aplicar otras técnicas como la compensación de movimiento basada en bloques, que tiene en cuenta que gran parte de un fotograma nuevo está ya incluido en el fotograma anterior, aunque quizás en un lugar diferente del mismo. Esta técnica divide un fotograma en una serie de macrobloques (bloques de píxeles). Se puede componer o "predecir" un nuevo fotograma bloque a bloque, buscando un bloque que coincida en un fotograma de referencia. Si se encuentra una coincidencia, se codifica la posición en la que se debe encontrar el bloque coincidente en el fotograma de referencia. La codificación del vec-

tor de movimiento, como se denomina, precisa de menos bits que si hubiera de codificarse el contenido real de un bloque.

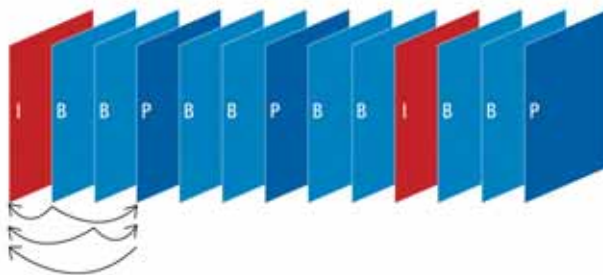


Con la predicción interfotograma, cada fotograma de una secuencia de imágenes se clasifica como un tipo de fotograma concreto, como un fotograma I, P o B.

Un fotograma I, o intrafotograma, es una imagen autónoma que se puede codificar de forma independiente sin hacer referencia a otras imágenes. La primera imagen de una secuencia de video es siempre un fotograma I. Los fotogramas I sirven como puntos de inicio en nuevas visualizaciones o como puntos de resincronización si la transmisión de bits resulta dañada. Los fotogramas I se pueden utilizar para implementar funciones de avance o retroceso rápido o de acceso aleatorio. Un codificador insertará automáticamente fotogramas I a intervalos regulares o a petición de nuevos clientes que puedan incorporarse a la visualización de una transmisión. La desventaja de este tipo de fotogramas es que consumen muchos más bits, pero por otro lado no generan demasiados efectos provocados por los datos que faltan.

Un fotograma P (de interfotograma Predictivo), hace referencia a partes de fotogramas I o P anteriores para codificar el fotograma. Los fotogramas P suelen requerir menos bits que los fotogramas I, pero con la desventaja de ser muy sensibles a la transmisión de errores, debido a la compleja dependencia con fotogramas P o I anteriores.

Un fotograma B, o interfotograma Bipredictivo, es un fotograma que hace referencia tanto a fotogramas anteriores como posteriores. El uso de fotogramas B aumenta la latencia.



Secuencia típica con fotogramas I, B y P. Un fotograma P sólo puede hacer referencia a fotogramas I o P anteriores, mientras que un fotograma B puede hacerlo a fotogramas I o P tanto anteriores como posteriores.

Cuando un decodificador de video restaura un video descodificando la transmisión de bits fotograma a fotograma, la descodificación debe comenzar siempre por un fotograma I. Los fotogramas P y B, en caso de usarse, deben descodificarse junto a los fotogramas de referencia.

Continúa en página 160

Viene de página 156

Los productos de algunas marcas permiten a los usuarios configurar la longitud de GOV (grupo de video), la cual determina la cantidad de fotogramas P que deberían enviarse antes de realizar el envío de otro fotograma I. La frecuencia de bits se puede disminuir mediante la reducción de la frecuencia de fotogramas (GOV más largo). Para reducir la latencia no se utilizan fotogramas B.

Además de la codificación diferencial y la compensación de movimiento, se pueden emplear otros métodos avanzados para reducir aún más los datos y mejorar la calidad de video. El H.264, por ejemplo, admite técnicas avanzadas como los esquemas de predicción para codificar fotogramas I, la compensación de movimiento mejorada con una precisión inferior a un píxel y el filtro de eliminación de bloques en bucle para suavizar los bordes de los bloques (defectos).

7.2. Formatos de compresión

7.2.1. Motion JPEG

Motion JPEG o M-JPEG es una secuencia de video digital compuesta por una serie de imágenes JPEG individuales. (JPEG son las siglas de Joint Photographic Experts Group - Grupo de Expertos Fotográficos Unidos). Cuando se visualizan 16 o más imágenes por segundo, el ojo humano lo percibe como un video en movimiento. Un video en completo movimiento se percibe a 30 (NTSC) o 25 (PAL) imágenes por segundo.

Una de las ventajas de Motion JPEG es que cada imagen de una secuencia de video puede conservar la misma calidad garantizada que se determina mediante el nivel de compresión elegido para la cámara de red o codificador de video. Cuanto más alto es el nivel de compresión, menor es el tamaño del archivo y la calidad de imagen. En algunas situaciones, como cuando hay poca luz o la escena es compleja, el tamaño del archivo puede ser bastante grande y, por lo tanto, usar más ancho de banda y espacio de almacenamiento. Para evitar que esto ocurra, los productos de video en red permiten al usuario establecer un tamaño máximo para un fotograma de imagen.

Al no haber dependencia alguna entre los fotogramas de Motion JPEG, un video Motion JPEG es resistente, lo que significa que si falla un fotograma durante la transmisión, el resto del video no se verá afectado.

Motion JPEG es un estándar que no requiere licencia. Tiene una amplia compatibilidad y su uso es muy habitual en aplicaciones donde se requieren fotogramas individuales en una secuencia de video -por ejemplo, para el análisis- y donde se utiliza una frecuencia de imagen de 5 fotogramas por segundo o inferior. Motion JPEG también puede ser útil para aplicaciones que requieren integración con sistemas solo compatibles con Motion JPEG.

Sin embargo, el principal inconveniente de Motion JPEG es que no utiliza ninguna técnica de compresión de video para reducir datos, ya que consiste en una serie de imágenes fijas y completas. El resultado es una frecuencia de bits relativamente alta o una relación de compresión baja para la calidad proporcionada, en comparación con estándares como JPEG-4 y H.264.

7.2.2. MPEG-4

Cuando se menciona MPEG-4 en las aplicaciones de videovigilancia, normalmente nos referimos a MPEG-4 Parte 2, también conocido como MPEG-4 Visual. Como todos los estándares MPEG (Moving Picture Experts Group), requiere una licencia, es decir, los usuarios deben pagar una tasa de licencia por cada estación de supervisión. MPEG-4 es compatible con aplicaciones de ancho de banda reducido y aplicaciones que requieren imágenes de alta calidad, sin limitaciones de frecuencia de imagen y con un ancho de banda virtualmente limitado.

7.2.3. H.264 o MPEG-4 Part 10/AAVC

El H.264, también conocido como MPEG-4 Parte 10/AAVC para Codificación de Video Avanzada, es el estándar MPEG más actual para la codificación de video. Se espera que el H.264 se convierta en la alternativa de estándar en los próximos años. Ello se debe a que, sin comprometer la calidad de la imagen, un codificador H.264 puede reducir el tamaño de un archivo de video digital en más de un 80% si se compara con el formato Motion JPEG, y hasta un 50% más en comparación con el estándar MPEG-4. Esto significa que se requiere menos ancho de banda y espacio de almacenamiento para los archivos de video. O, visto de otra manera, se puede lograr mayor calidad de imagen de video para una frecuencia de bits determinada.

El H.264 ha sido definido conjuntamente por organizaciones de normalización del sector de las telecomunicaciones (ITU-T's Video Coding Experts Group) y de las tecnologías de la información (ISO/IEC Moving Picture Experts Group), y se espera que tenga una mayor adopción que los estándares anteriores. En el sector de la videovigilancia, H.264 encontrará su mayor utilidad en aplicaciones donde se necesiten velocidades y resoluciones altas, como en la vigilancia de autopistas, aeropuertos y casinos, lugares donde por regla general se usa una velocidad de 30/225 (NTSC/PAL) imágenes por segundo. Es aquí donde las ventajas económicas de un ancho de banda y un almacenamiento reducidos se harán sentir de forma más clara.

Se espera que H.264 acelere también la adopción de cámaras megapíxel, ya que con esta eficiente tecnología de compresión se pueden reducir los archivos de gran tamaño y las frecuencias de bits sin que la calidad de la imagen se vea afectada. En cualquier caso, tiene sus exigencias: aunque H.264 permite ahorrar en costes de ancho de banda y almacenamiento, también necesita cámaras de red y estaciones de control de mejor rendimiento.

Los codificadores H.264 utilizan el perfil base, lo que supone que sólo se usan los fotogramas I y P. Este perfil es el ideal para cámaras de red y codificadores de video, ya que la latencia se reduce gracias a la ausencia de fotogramas B. La latencia baja es esencial en aplicaciones de videovigilancia donde se realice supervisión en directo, sobre todo si se emplean cámaras PTZ o domos PTZ.

7.3. Frecuencia de bits variable y constante

Con el MPEG-4 y el H.264, los usuarios pueden determinar que una transmisión de video codificado tenga una frecuencia de bits variable o constante. La selección óptima dependerá de la aplicación de la infraestructura de red.

Con la VBR (frecuencia de bits variable), se puede mantener un nivel predefinido de calidad de imagen independientemente del movimiento o falta de movimiento en una escena. Esto significa que el uso de ancho de banda aumentará cuando haya mucha actividad en una escena y disminuirá cuando no haya movimiento. A menudo esta opción es ideal para las aplicaciones de videovigilancia que requieren una alta calidad, especialmente si hay movimiento en una escena. Debido a que la frecuencia de bits puede variar, incluso aunque se haya definido una frecuencia de bits media de destino, la infraestructura de red (ancho de banda disponible) debe poder adaptarse a grandes caudales de datos.

Con un ancho de banda limitado se recomienda utilizar el modo CBR (frecuencia de bits constante), ya que este modo genera una frecuencia de bits que el usuario puede predefinir. La desventaja que tiene la CBR es que si, por ejemplo, hay mucha actividad en una escena que da como resultado una frecuencia de bits mayor que la velocidad de destino, la restricción para mantener una frecuencia de bits constante conlleva una calidad y frecuencia de imagen inferiores. ■