

# H.264

## PARA EL RESTO DE NOSOTROS

POR  
KUSH AMERASINGHE

TRADUCCIÓN AL ESPAÑOL POR  
ALVARO CALANDRA

Una guía independiente acerca de la compresión de video y F4V, una explicación específica de las preferencias y configuración de F4V en Adobe Media Encoder y Adobe Flash Media Encoding Server, y un ejemplo de estimación de los límites de tasa de bits aceptables en cualquier video.



*“La compresión del video es como empacar las valijas”*

## Introducción

H.264 ha penetrado el mundo del video en la web como una tormenta. Esta tecnología relativamente nueva tiene un poco de misterio e intriga. Siendo una persona sin la paciencia necesaria para comprender todas las particularidades sobre cómo funciona esta maravilla, y aun así pudiendo obtener lo que la misma tenía para ofrecer, tuve un rara ocurrencia aparte de experiencia de aprendizaje (lo que me hizo sentir un adolescente otra vez).

Ahora viendo lo difícil que es resumir este tema (a la hora de escribir este trabajo) pensé que podría haber otras personas que se pudiesen beneficiar del mismo, si podía compartir lo que he aprendido hasta ahora.

Los elementos básicos de este texto incluyen:

- Una guía independiente sobre compresión y F4V.
- Una explicación sobre la configuración específica de F4V y sobre el Adobe Media Encoder y el Adobe Flash Media Encoder Server.
- Un caso de estudio sobre lo que considero sería un tasa de bits aceptable para cualquier video.

## La Versatilidad de H.264

H.264 es muy sofisticado y aun así es una tecnología muy versátil creada para servir a una variedad de aplicaciones que van desde una teleconferencia, con un tamaño de imagen pequeño y súper comprimida, hasta formatos cinematográficos enormes y de gran calidad. Usar tecnología con un rango posible de opciones tan grande para el relativamente estrecho campo del video para web (que está siendo constantemente modificado y redefinido al aumentar los anchos de banda y la potencia de la tecnología disponible, y cambiar el espectro de calidad) puede ser un gran desafío. El grado de control que H.264 ofrece unido a la complejidad de los métodos de compresión en sí mismos puede ser abrumador, tanto para aquellas personas nuevas en el tema de video online, así como para los más avezados productores de video para la web.

## La metáfora de la valija

Me gusta pensar que la compresión de video es como empacar las valijas para un viaje. Espero que esto sea algo que todos puedan asociar. Aun si uno no es un viajero experimentado en algún momento se debe haber enfrentado al desafío de poner la mayora cantidad de ropa posible dentro de una valija determinada, preservando las condiciones de la ropa lo mejor posible, para cuando la vayamos a usar en el destino final. Así que volveremos varias veces a esta analogía para describir los aspectos de la compresión de video que sean aplicables.

# ¿H-punto-qué?

## Códecs

Un códec (codificador-decodificador) es un método o tecnología particular que se usa para comprimir una señal electrónica como una grabación de video o audio. Los variados códecs son a veces formas distintas de hacer lo mismo, sin embargo puede suceder que un códec no tenga nada que ver con otro. Cada códec tiene sus ventajas y desventajas. Casi siempre los códecs más nuevos proveen una mejor calidad y eficiencia que los más antiguos, que a su vez tiene mayor base instalada. Los códecs más usados para comprimir video para Adobe Flash son H.264, On2 VP6 y Sorenson Spark. H.264 fue introducido en Adobe Flash Player en la versión 9.0.r115. On2 VP6 era el códec anterior introducido en Flash Player 8. Sorenson Spark es el más viejo de los códecs soportados por Flash Player introducido en la versión 6.

## Compatibilidad

Al momento de escribir esto alguna versión del Flash Player está instalada aproximadamente en el 98% de las computadoras conectadas a Internet en el mundo. De esta cantidad la proporción de usuarios que aun usan Flash Player 6 o anterior es nula en todo el mundo. La adopción de las nuevas versiones del Flash Player que pueden reproducir los últimos códecs de H.264 es de cerca del 90% de esta base de usuarios. Esto convierte a al video codificado usando H.264 para Flash en la opción más atractiva para las aplicaciones de distribución de video a nivel global.

## Seleccionando un códec

Hay algunos otros factores a tener en cuenta al elegir un códec. Si bien algunos usuarios tienen la capacidad de reproducir video H.264 a nivel de software, su capacidad de procesamiento (potencia de la computadora) y su ancho de banda disponible, unido a su configuración particular (o rango de configuración en el caso de codificación multi-bit) usados actualmente para entregar video, podría afectar seriamente la calidad de la experiencia que el usuario recibe. Si su franja de audiencia es un grupo altamente especializado (por ejemplo, de un país, una franja etaria, o un grupo de características comunes) la proporción de tales usuarios en este grupo puede (o puede no ser) ser diferente que la audiencia en general.

Por ejemplo, la mayor relación de calidad/tasa de transferencia de H.264 generalmente tiene una necesidad de mayor poder de procesamiento; ese poder de procesamiento puede no ser accesible para la audiencia para la cual se generó el video.

También como los códecs son simplemente diferentes formas de comprimir una señal, más que una serie lineal e incremental de mejoras en las especificaciones, algunos métodos pueden ser más efectivos que otros según las necesidades. Por ejemplo, un archivo FLV codificado usando On2 VP6 puede contener Alpha Channels (Canal Alfa o sea transparencia) que se podría utilizar para crear video con transparencia para usarse con interactividad en una aplicación Flash, mientras que H.264 en Flash no soporta hoy Alpha Channels.

Aun así creo que H.264 es el ganador en términos de ventajas y en la capacidad de distribuir eficientemente video de gran tamaño y alta calidad (aun comparado con códecs competitivos que no están disponibles para Flash Player). Por esta razón es que me enfocaré en el uso de ésta tecnología específicamente aplicada a Flash Player.

*“La gran mayoría de los consumidores de video en Internet ya han elegido las nuevas versiones de Flash Player capaces de reproducir video H.264”*

Entonces ¿Dónde calzaría el ejemplo de la valija en el universo de los códecs y sus aspectos de compresión?

Al principio estaba inclinado a pensar en esto como en la marca de valija a utilizar. Luego me di cuenta de la variedad de factores que participan en la elección del códec es casi tan diversa como el proceso de elección de cómo comenzar el viaje – en avión, tren, barco, etc. Cada uno tiene su costo, tiempo, esfuerzo, y factores de accesibilidad que dependen mucho del destino particular. Así como algunos destinos tienen muchas opciones sensatas para llegar ahí, otras no serán alcanzadas con determinadas formas de viaje.

Así como algunas formas de viajar inventadas más tarde en la historia hicieron más práctico ir a determinados lugares, algunas aplicaciones de video son posibles a partir de la invención de ciertos códecs.

*Es importante entender lo que los algoritmos de compresión tienen que soportar (aparte del compañero de cuarto molesto)*

## Contenedores

Es importante recordar que los datos de video en cualquier códec se pueden entregar en varios tipos de formatos “contenedores” junto a otras informaciones en otros formatos (por ejemplo, la pista de audio y su códec, metadata y otros). Entonces cuando consideramos la habilidad de determinado dispositivo de interpretar cierto tipo de video hay que asegurarse que soporte los códecs específicos + la combinación con el contenedor para que sea totalmente compatible.

## ¡La Compresión no es mezquina, solamente incomprendida!

Mucha gente piensa que la compresión es un proceso desagradable que degrada completamente sus preciadas creaciones audiovisuales en algo mucho peor que lo que era el original. No se los puede culpar porque es cierto. Aun si nos tomamos el tiempo de ver el mundo desde la perspectiva de un compresor, no solo vamos a apreciar las maravillas y la practicidad que la compresión va a agregar al proceso de publicar sus videos, sino que nos haremos amigos de la compresión y nos llevaremos muy bien! H.264 es uno de los códecs más complicados en uso global en el mundo hoy y puede ser un gran desafío producir con él. Sin embargo lo que puede lograr, el esfuerzo extra en lograr una compresión básica de la filosofía que tiene detrás y de su comportamiento nos puede convertir a todos en mejores “compresores” (el que hace la compresión).

Antes de comenzar a pensar sobre cómo obtener los mejores resultados de una combinación particular de configuración de compresión, hay que considerar los factores que hacen más o menos trabajosa la tarea para el método de compresión en sí mismo.

Algunos son obvios. Por ejemplo, a mayor duración del video, mayor es la cantidad de datos que se necesitan transmitir al espectador (así como hay que cargar más ropa para unas vacaciones más largas). Análogamente un cuadro de mayor tamaño o de mayor resolución en el video, requerirá una cantidad mayor de datos que se necesitan transportar o almacenar (el guardarropa de un adulto para la misma estadia ocupa más espacio que el de un niño). Estas dimensiones temporales y espaciales de un video son aspectos claves en lo que está focalizado la compresión de H.264 para poder reproducir el video lo más cercano posible al original, con la menor cantidad de datos posibles (esto se discute más adelante).

Otros factores también dictaminan que tan amigable es la compresión del video, basado en el contenido. La naturaleza del contenido puede hacer que el video sea más fácil de comprimir mientras mantiene una calidad razonable, o muy difícil de comprimir adecuadamente. (Piense como algunos artículos de vestimenta puede doblarse fácilmente en capas, mientras que otros son imposibles de doblar o apretar sin destruirlos, por ejemplo camisetitas vs sombreros).

Otro factor crucial que afecta la eficiencia de la compresión es cuan dinámico o estático es el contenido del video, en términos de movimiento del sujeto filmado y/o del movimiento de la cámara. Esto es de alguna forma independiente de la tasa de cuadros (cuantos cuadros se reproducen por segundo). 4

Como H.264 tiene formas sofisticadas de detectar y empaquetar eficientemente cuadros repetidos o redundantes, aun cuando reducir la tasa de cuadros pueda reducir la cantidad de información que se necesita empaquetar. (Unas vacaciones con una agenda complicada que incluye diferentes actividades necesitando varios cambios de ropa al día demanda más ropa que una en la que sólo necesito una muda de ropa al día).

## El acto de balancear

Pienso que hay tres ejes fundamentales que hay que tener en cuenta para lograr la mejor solución de codificación para uno, y la mejor experiencia de reproducción para la audiencia



**Figura 1: La consideración tridimensional de la compresión:** No hay bien o mal, sino una elección sobre que tomar y que dejar al decidir en qué cuadrante de la gráfica de compresión se quiere estar.

## Calidad versus velocidad de datos

Lo primero es el balance entre calidad y velocidad de datos. Aquellos que sean viejos como para recordar aquellos días en que Internet comenzó a utilizar imágenes estáticas por primera vez (para aquellos que no lo son, si hubo un tiempo en que Internet era solo texto) también recuerdan que el gran dilema era comprimir las imágenes lo justo para que cargaran razonablemente rápido, sin destruir su aspecto completamente. Esto es aun hoy considerado “buenas prácticas” a pesar de la adopción general de la banda ancha. Al día de hoy tanto los desarrolladores de páginas web así como los usuarios finales tienen una comprensión implícita acerca de donde está el balance. Como el video de larga duración en la web es relativamente nuevo, no es sorprendente que los viejos desarrolladores de páginas web tengan una sensación de “deja-vú”.

Similarmente al mundo de las imágenes, en el mundo del video, incrementar la velocidad de datos durante la fase de compresión permite empaquetar una mayor (humanamente percibida) calidad visual en cualquier video.

Sin embargo el incremento en la velocidad de datos se puede degradar o a veces impedir la reproducción, debido a limitaciones en el ancho de banda del público objetivo. Mientras tanto reducir la velocidad de datos limita la calidad del video; pasando de cierto límite degrada el video al punto de hacerlo incomprensible, a pesar de la reproducción mejorada y/o velocidad de carga – de la misma forma que apretar demasiada ropa en una valija pequeña puede estropearla (especialmente si se toman medidas drásticas como las que

**Figura 2: El "Artifacting" de la Compresión:** Cuando se comprime demasiado los efectos de la compresión comienzan a hacerse obvios



se discutirán más adelante) mientras que una valija de mayor tamaño cuesta más pero a veces puede ser rechazada por la aerolínea si no cumple con las medidas de peso y tamaño. (El hosting de archivos en diferentes tasas de bits permite entregar el mismo video en un rango de opciones transparentes para el usuario, lo que podría resolver este problema en el caso de que el público objetivo tenga un rango de anchos de banda más amplio). Hay que tener en cuenta que la velocidad de datos de lo codificado impacta en la cantidad de datos entregados desde un sitio web, lo que a su vez afecta los costos del hosting.

## Eficiencia versus rendimiento de la reproducción

Los métodos de compresión complejos como H.264 pueden consumir un poder de procesamiento significativo en comparación con métodos más sencillos al decodificar durante la reproducción. Aun así utilizar métodos de compresión complejos permite producir con mejor calidad, manteniendo la misma velocidad de datos.

Así que el segundo eje de la ecuación es rendimiento versus eficiencia, lo que significa que uno puede ganar en una mayor eficiencia, obteniendo más calidad y velocidades de datos menores, si se está dispuesto a invertir en más poder de procesamiento y viceversa (de forma desigual en las dos caras de la ecuación de codificación y decodificación). Este es un punto importante a recordar ya que H.264 tiene muchos parámetros que manejan este aspecto, y comprender esta filosofía de eficiencia versus mayor poder de cómputo es crucial para entender la potencialidad que tienen estos parámetros.

## Eficiencia en la compresión versus velocidad de codificación

Mientras existe una cierta correlación entre una mayor complejidad en la decodificación y la demanda de mayor poder de cómputo en la codificación, en general se utilizará más tiempo en la codificación que en la decodificación. Esto tiene sentido en la mayoría de

los casos, donde la reproducción debe ser en tiempo real mientras que la codificación se maneja con otros tiempos (salvo en la transmisión en vivo). No todas las medidas que incrementan el tiempo de codificación a cambio de una mejor relación de tasa de bits/calidad, resultan en un video difícil de decodificar. Si se tienen los recursos se debería invertir el tiempo que lleve obtener los mejores ratios de eficiencia de compresión durante la codificación. De todas formas puede haber situaciones en que algunas funciones generen un beneficio imperceptible en cuanto a la calidad para un tasa de bits determinado pero a su vez incrementen los tiempos de codificación significativamente. En esos casos uno debería evaluar cuidadosamente el balance entre practicidad y el resultado final.

## Comprendiendo el balance de rendimiento

Balancear esos dos ejes es similar a si usamos más tiempo doblando y organizando la ropa, uno podría poner más cosas en determinada valija, mientras un empaque más rápido y descuidado podría ahorrar tiempo de empaque a costo de menos eficiencia en el uso del espacio. Más aun si tomásemos más pasos para hacer que la ropa se empacara mejor, digamos planchar al vapor una por una las prendas, se podría obtener un mejor resultado al costo de perder más tiempo. Podríamos ir un paso más allá al usar algún sistema de envasado al vacío para lograr un paquete de ropa más compacto. Esto consumiría más tiempo y dinero. Sin embargo planchar la ropa incrementaría el tiempo que lleva empacar, pero no implicaría ninguna demora para desempacar, una vez llegado a destino; mientras que empacar al vacío la ropa en plástico tomaría más tiempo al empacar y al desempacar. Esto es también cierto en el mundo de H.264, donde algunas configuración hacen que un video demore más en codificar, pero no tiene ningún impacto en la decodificación, mientras hay otros configuración que impactan gravemente la descompresión y la reproducción debido al mayor proceso de compresión.

## Las pérdidas y ganancias de compresión con pérdida

Hay dos grandes grupos de compresión de datos en general: Compresión sin Pérdida (Lossless) y Compresión con Pérdida (Lossy). Compresión sin pérdida es cuando luego del proceso de compresión y descompresión el archivo es idéntico al original. En otras palabras los datos son reacomodados y procesados de forma tal que lleva menos espacio y aun así toda la información está disponible cuando es reconstruido. (Un ejemplo de compresión sin pérdida es un archivo ZIP, donde el archivo es igual cuando se hace unzip). En compresión con pérdida los datos son reducidos de forma tal que ocupa un espacio significativamente menor que en el caso de compresión sin pérdida, simplemente por eliminar una parte de la información del archivo original.

La idea detrás de la compresión con pérdida es hacerlo de forma tal que la información eliminada pase desapercibida para la audiencia (especialmente si no tienen con que compararla) y así ahorrar un gran volumen de datos a la hora de transmitirlos. Cuando esto se exagera, los efectos de esta forma tan destructiva de reducir se hacen notorios, hasta a los ojos de los profanos. Estos efectos secundarios debidos a la sobre compresión se conocen como Artefactos de la Compresión o Compression Artifacts en inglés.

Para hacerlo más interesante, los algoritmos de compresión por si mismos no tienen la conciencia humana de cuando estos artefactos se hacen notorios, al punto que la continua reducción de datos causa más daño que beneficio. Análogamente el algoritmo no tiene conciencia si agrega valor al incrementar la calidad, al costo de generar más datos a transferir, ya que el ojo humano no registra el detalle agregado al original. Aun peor, este nivel de calidad aceptable o excelente es altamente subjetivo, no solo de aplicación a aplicación (por ejemplo video conferencia versus largometraje), pero también entre públicos objetivos y aun individuos con diferentes niveles de expectativas. Este es el punto donde la compresión de video se acerca más a ser un arte que a una ciencia.



**Figura 3: la Compresión con pérdida es destructiva.** Empacar su ropa con compresión con pérdida significa que cuando llegue a destino no tendrá la misma ropa.

Este también es un punto en que la comparación con la valija se hace un poco bizarra. Aun para mantener la consistencia del argumento, imagínese que decide remover algunos de los abultados botones que están obstaculizando la forma prolija en que esta doblando la ropa. Para ser más drástico podría decidir cortar las mangas largas de las camisas con una tijera para alivianar el peso, obteniendo una camisa que es aun utilizable (por lo menos en teoría). En ambos casos cabe señalar como la ropa al llegar a destino sería diferente que la que empaco; y el cambio es permanente, así como no hay vuelta atrás con el video con compresión con pérdida con respecto al original. A diferencia del empacador desesperado que aun tiene un poco de sentido común en su mente sobre la usabilidad de la ropa una vez desempacada, los métodos computacionales completamente inconscientes de compresión no sabrían donde parar y podrían destejer una prenda, para convertirla en un ovillo de lana, que podría en teoría ocupar menos espacio ¡si bien no tendría uso alguno al llegar a destino!

Aun los cambios más pequeños que puede hacer la compresión con pérdida a un video pueden resultar en problemas luego cuando se vuelva a comprimir. Es por esta razón que siempre se debe usar la máxima calidad disponible hasta el final del proceso. (Las herramientas de Adobe hacen muy práctico trabajar en su formato nativo.) Aun un video que parece de alta calidad y parece igual al original tendrá información faltante y patrones ocultos inesperados y propiedades no tan obvias al ojo inexperto al comprimir.

Por lo tanto si se tiene la alternativa, siempre se debe optar por el archivo menos procesado, aun si la versión comprimida “parece” que no tuviese problemas de calidad, ya que los problemas que permanecen ocultos al ojo humano se puede ver amplificados en el proceso de recompresión.

## Trastorno obsesivo de compresión

Los esquemas de compresión sofisticados como H.264 brillan en su actividad por lo obsesionados que son por naturaleza por obtener la calidad visual máxima, con la mínima cantidad de datos posibles. Otro aspecto notable del esquema de compresión de H.264 es como usa referencias a otras áreas del video clip para definir una nueva área sin tener que redefinirla desde cero.

Por ejemplo, la cantidad de datos que se requieren para definir la imagen de un automóvil que se mueve a través de la pantalla de un cuadro al otro, con un set de datos nuevo para cada cuadro, es mucho mayor que en comparación con definirlo una vez y reutilizar esa información para decir “Ok, ponga el mismo automóvil un poco más a la derecha en este cuadro”. Ahora ¿qué pasa con el área que antes estaba cubierta por el auto cuando sea visible porque el auto se movió? Ningún problema: “fíjese como se veía antes que el auto entrara en cuadro para comenzar” y así sucesivamente.

Este nivel de obsesión por ahorrar espacio al punto que solo podría ilustrarse con un emparador tan desesperado en ahorrar espacio al punto que simplemente omitiría ciertos prendas de ropa al poner una nota de cómo recrear un conjunto como por ejemplo “compre otro par de pantalones de este estilo, pero del mismo color que la chaqueta” o empaclar una sola media con una nota que diga “trate de encontrar una parecida a esta”.

Este tipo de referencia puede ocurrir tanto a nivel del aspecto espacial de comprimir el cuadro individual así como el aspecto temporal de comprimir movimiento sobre varios cuadros.



**Figura 4:** La compresión de H.264 llega a los extremos para reducir lo que necesita llevar en la valija. En lugar de empaclar las dos medias en un par, ¿porqué no cargar una sola con una nota que diga que la llegar al destino se haga una copia?

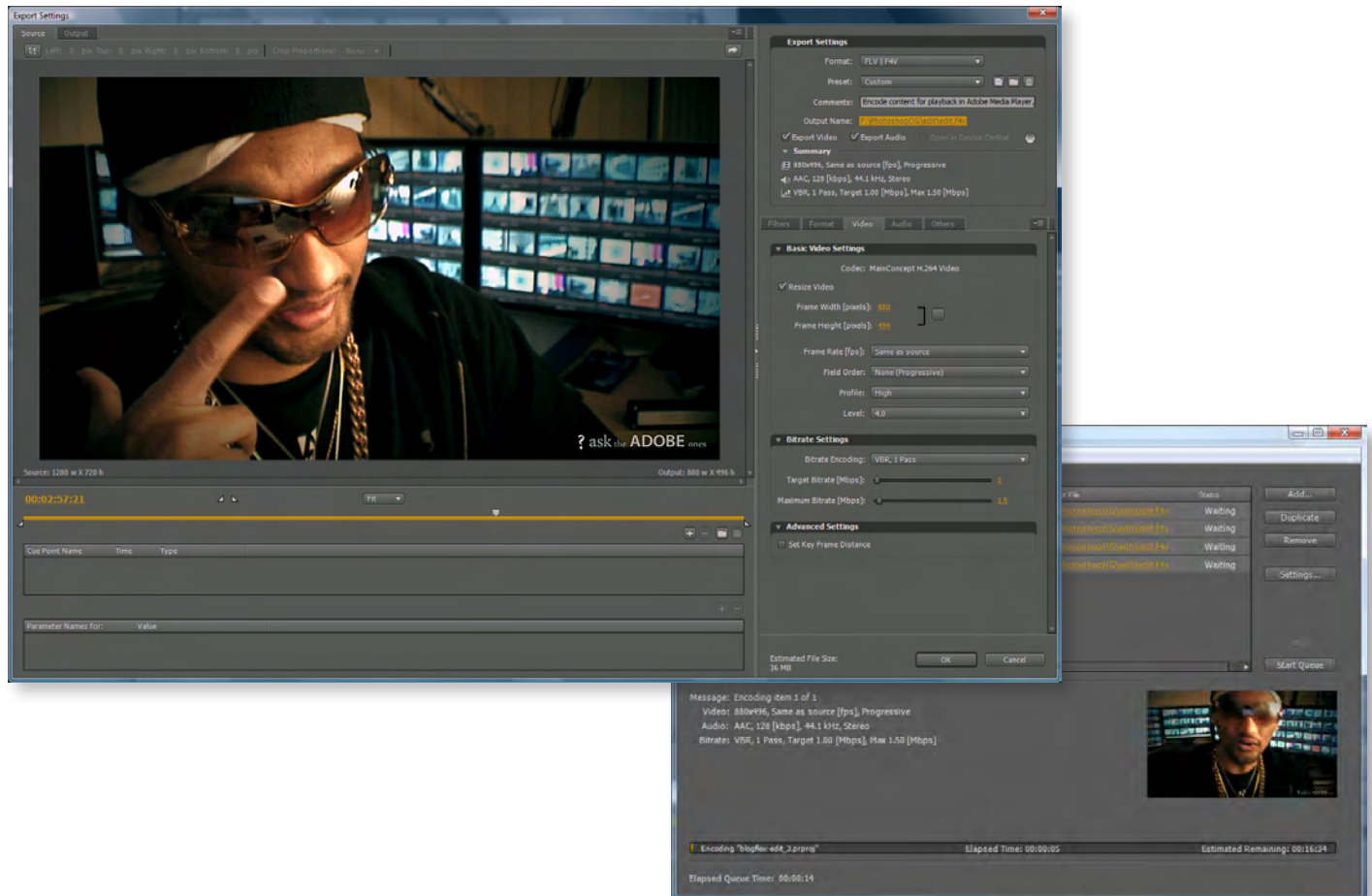
## Domando los controles

Se dice que el poder no sirve sin control – ¡Ignorar que es lo que permiten hacer los controles en codificación H.264 puede dejarlo impotente y perdido!

**Figura 5: Conocer lo que los controles hacen ayuda a Extra-H a verse bien en la web.**

AME esta por comprimir un video donde estoy yo y mi alter ego “extra-h” en “Ask the Adobe Ones”.

Ahora que tiene una idea básica de lo que sucede en el misterioso pero encantador mundo de la compresión H.264, lo llevare a un paseo por los controles y configuración de los más utilizados codificadores ofrecidos por Adobe.



## Adobe® Media Encoder

Adobe Media Encoder (AME) viene incluido como parte del paquete de herramientas dentro de Adobe Creative Suite 4, se inicia automáticamente cuando manda renderear un video desde Adobe Premiere Pro CS4. Si bien Adobe Media Encoder puede renderear una variedad de formatos de video, ahora solo nos focalizaremos en la configuración específica de H.264.

Cuando uno exporta media desde Adobe Premiere Pro CS4, se abre una amplia selección de formatos disponibles (como Quicktime, AVI, FLV y otros). Esta lista de formatos es la misma serie de opciones que se obtienen al agregar un archivo raw directo al AME y se asigna el formato usando el menú emergente en la columna de formato. (Uno puede especificar estas configuraciones para cada archivo en la cola de AME en forma independiente). Estas configuraciones determinan el formato del archivo o el contenedor de la media codificada.

## AME configuración de exportación

### Formato

Usted puede notar que existen varias opciones de formato dentro de H.264. Si está exportando a un formato estándar como Blu-ray o iPod, seleccione de una plantilla desde configuración del formato y luego seleccione del subgrupo apropiado de ese formato, desde el menú emergente. Para video para web la configuración puede necesitar algunos ajustes dependiendo de su caso particular.

Mientras Adobe Flash Player soporta una amplia variedad de estándares, ciertos datos adjuntos que el contenedor de F4V puede almacenar no se interpretan correctamente en iTunes y QuickTime. Por lo tanto hay que hacer una diferencia entre el H.264 genérico y el especialmente renderado para Flash Player que utilizan la extensión F4V.

### Presets

Los Presets son combinaciones predefinidas de configuraciones de codificación preparadas para el formato que usted seleccionó. En el caso de la categoría del formato FLV/F4V, hay algunos presets que usan el formato FLV, en cuyo caso se puede optar por los códecs On2 VP6 o Sorenson Spark, y en los diseñados para F4V se puede usar el códec Main Concept H.264. Estos presets se pueden usar como están o como base para crear su propio set de configuración de compresión optimizada. Aparte de la configuración del códec en sí mismo, hay otras opciones relacionadas al códec que están disponibles para cada tipo de archivo (FLV y F4V).

En el diálogo de exportación de Premiere Pro, se puede ajustar la configuración de un determinado preset luego de haberlo seleccionado. En el caso de la cola de AME, usted puede seleccionar el preset utilizando el menú emergente en la columna Preset; luego si se hace clic sobre el preset subrayado, un cuadro de diálogo aparece similar al de Premiere Pro Media Export, en donde se podrán cambiar las opciones del preset que usted seleccionó. (Además de modificar la configuración, usted puede también seleccionar un formato diferente y/o un preset, y aun luego de eso puede modificarlo y puede grabarlo como un preset personal).

Independientemente del preset seleccionado, se puede alternar entre subformatos de FLV o F4V en la pestaña de Formato antes de modificar la configuración de video y audio en sus respectivas pestañas.

## Configuración de Video de AME

### Tamaño

Si selecciona la opción “Resize Video” (redimensionar video), se puede definir el ancho y el alto del cuadro de video en pixeles. Si esta opción no está seleccionada el video se codificará en su tamaño nativo. Mientras codificar en su tamaño nativo retendrá la máxima calidad, mientras se provea la tasa de bits necesario para que pueda ser interpretado sin artefactos aparentes, en la práctica usted probablemente va a querer reducir el tamaño del cuadro para que sea más práctico para el compresor H.264 de empaquetar en un flujo de datos o en un archivo lo suficientemente pequeño, para ser descargado por una conexión de Internet promedio. Generalmente uno debería de abstenerse de agrandar los videos más allá de su tamaño original, ya que no tiene sentido agrandar la carga de trabajo mientras degrada la calidad inicial de la imagen en la precompresión.

El número de píxeles en un cuadro de video es un múltiplo del ancho por el alto del cuadro. Por eso incrementar cualquiera de ambos tendría un efecto dramático en el “área” o el número total de píxeles. Incrementar ambos tendría un impacto mayor aun. Por ejemplo, uno pensaría que 640x480 es “el doble de grande” comparado con 320x240. De hecho duplicar tanto ancho como largo, se cuadruplica la cuenta de píxeles. Este es un punto importante para recordar cuando se hacen las cuentas para obtener tasa de bits que funcionen bien con determinados tamaños de video, ya que la cantidad total de píxeles es un factor más sensible a tener en cuenta que las meras medidas del cuadro de video. Sin embargo, mientras que en general cuantos más píxeles en el video, más carga de procesamiento para la codificación en cualquier método, en compresión en H.264 se utilizan complejos algoritmos que en algunos casos hace que esto no sea así. Por ejemplo un cuadro con un color solido en el fondo y un logo pequeño se puede comprimir de la misma forma que el mismo logo pegado en un fondo mucho más grande – siempre y cuando el logo no sea proporcionalmente más grande- a pesar de la mayor cantidad de píxeles en este último.

Más allá del efecto relativo que puede tener el tamaño del cuadro de video en la tasa de bits, el valor de los píxeles en sí mismos, puede jugar un rol preponderante. Esto tiene que ver con cómo se hace la compresión a nivel de bloques o conjuntos de píxeles más que de la imagen entera. Por ejemplo tamaños de cuadro que son divisibles entre 16 son preferibles a los que no lo son (divisibles por 8 es la siguiente mejor opción).

Al modificar el tamaño es importante mantener el “Aspect Ratio” o sea la relación de aspecto (la proporción de ancho sobre alto) que sea igual al video original para prevenir que se degrade estéticamente (aplastando o estirando) la imagen. El ojo humano es sensible a semejantes cambios en las proporciones, especialmente cuando los objetos deben ser proporcionales (como las ruedas de un auto) sufren una modificación desproporcionada. Las relaciones de aspecto más comunes en video son 4:3 y 16:9. Otras relaciones de aspecto existen en cine, por ejemplo, y aun así cuando se “hacen para televisión” estas imágenes son modificadas para que coincidan con una de las relaciones de aspecto básicas.

Otra consideración acerca del tamaño es que si la fuente del video tiene píxeles con relaciones de aspecto que no son cuadradas ( $no=1$ ) el tamaño del video codificado debe ser traducido a lo que sería si los píxeles fuesen cuadrados.

## ¿Los píxeles también tienen relación de aspecto?

Mientras que la relación de aspecto de la imagen hace referencia a la proporción relativa del ancho y alto de la imagen (como en 16:9) la relación de aspecto del pixel se refiere a la escala proporcional de cada pixel que compone la imagen. Ciertos formatos tradicionales de televisión usan lo que se llama píxeles “non-square” (no cuadrados). Por ejemplo, con HDV (que es distinto que HDTV) la imagen consiste de 1440x1080 píxeles. Esta imagen es estirada desproporcionalmente para que complete un área de 1920x 1080 cuando es reproducida en un televisor compatible con HD.

Cuando se selecciona específicamente el formato F4V, no hay configuración para definir la relación de aspecto del pixel (que si está presente en el caso de la configuración de H.264 genérico). Esto es así porque Flash Player reproduce mejor cuando el video contiene píxeles cuadrados o una relación de aspecto de pixel de 1. Los videos que no tiene esta relación de aspecto en sus píxeles, deben tomar pasos adicionales al reproducir y se resiente la relación calidad/rendimiento/tasa de bits, lo que no vale la pena. Como la relación de aspecto de los píxeles se va a definir como 1 al seleccionar F4V, se debería reflejar ese cambio en el video original si no tuviera esa relación de aspecto de píxeles. Por ejemplo, 1440x1080 HDV debería ser 1920x1080, porque  $1080 \times (16/9) = 1920$ ; y 720 x 576 PAL DV pantalla apaisada corresponde a  $1024 \times 576$  en píxeles cuadrados:  $720 \times (16/9)=1024$ .

Esta es una excepción a la regla sobre no escalar el tamaño del video hacia arriba, ya que la otra opción sería mantener el ancho real de los píxeles y reducir la altura para reflejar la correcta relación de aspecto, resultando en pérdida de detalle vertical.

Lo mejor es hacerlos cambios al cuadro de video justo al final de la etapa de codificación y mantener el tamaño nativo durante todo el proceso de la edición del video, usando así las capacidades nativas de las herramientas de video de Adobe.

## Tasa de cuadros (velocidad de cuadros o fotogramas)

La tasa de cuadros es cuántas imágenes fijas se reproducen en el video para crear la ilusión de imágenes en movimiento. El ojo humano puede ser engañado para que perciba este movimiento de imágenes alrededor de 24 cuadros por segundo en adelante (este es la tasa de cuadros del cine). PAL (común en Europa parte de Asia y Latinoamérica) usa 25 fps (frames per second o sea cuadros por segundo), mientras que NTSC estándar (en uso en Estados Unidos y Japón por ejemplo) usa 29,97 fps. A medida que la tasa de cuadros disminuye, el movimiento parece entrecortado, especialmente si el sujeto filmado cambia de ubicación en el cuadro rápidamente. De todas formas los primeros videos que venían en CD-ROM usaban tasa de cuadros bajos de hasta 15 o 10 fps para reproducir video, y algunos temas como la captura de pantalla de la computadora en tutoriales se pueden experimentar razonablemente con tasa de cuadros tan bajos como 5 fps. En contraposición los monitores para jugar con la computadora pueden manejar tasa de cuadros altos como 120fps. Finalmente hay una tendencia en crecimiento entre los creadores de video de simular el “efecto cine” usando 24 fps.

A no ser que el video haya sido procesado con una técnica especializada de creación de cuadros intermedios como el efecto Time Warp de Premiere Pro, no se debería nunca exceder la tasa de cuadros del video original. De hecho los mejores resultados se obtienen si se conserva la tasa de cuadros original de la fuente original. Como esto depende en gran manera de la cámara utilizada, es importante tener esta información antes de codificar el video y tomar la decisión de que tasa de cuadros usar. Así como en el tamaño del cuadro en la velocidad de los cuadros lo mejor siempre es mantener la original durante todo el proceso de la edición.

Si bien en un mundo ideal, los tasa de cuadros originales dan el mejor resultado posible, en la práctica la tasa de cuadros es una carga pesada para la tasa de bits. Una vez más, la complejidad de los algoritmos de compresión temporal de H.264 hacen difícil hacer una relación lineal entre la tasa de cuadros y su impacto en la tasa de bits. La naturaleza del contenido del video afecta dramáticamente la forma en que dicho video es comprimido sin presentar artefactos. Por ejemplo una cámara fija con una persona hablándole a la pantalla (talking head video) es mucho más fácil de comprimir que una secuencia de acción con tomas que cambian continuamente y que tienen mucho movimiento. En el mundo de H.264 la tasa de cuadros debería ser más bien visto como un techo. Los tasa de cuadros más altos permiten mayor suavidad en los movimientos al precio de un ratio menor de calidad/tasa de bits, mientras que reducir la tasa de cuadros reduce la carga, permitiendo más “espacio para respirar” para poder empacar mayor calidad visual (comprometiendo la calidad temporal).

Si usted va a reducir la velocidad de cuadros, trate de hacerlo en cifras pares de forma tal de no tener problemas de sincronización de audio y video que surgen al utilizar números impares en las conversiones.

## Field Order (orden de campos)

El orden de los campos es otra reliquia del pasado que sigue complicando el mundo del video digital. A diferencia de la película donde cada cuadro se proyecta a la pantalla, en una sucesión de imágenes tan rápida que no notamos el parpadeo de una a la siguiente (creando la ilusión de movimiento continuo), la tecnología tradicional de la televisión creaba la imagen en movimiento por intermedio de haz de electrones que escaneaban continuamente las filas de arriba hacia abajo con la velocidad suficiente para que se percibiera por el ojo

como una imagen completa. Para poder empacar más información, en cada cuadro cada fila de “píxeles” saltea una (por ejemplo escanear la primera, tercera, quinta, etc, hasta completar el primer campo de la imagen, al terminar volver al inicio o sea arriba y escanear la segunda, cuarta, sexta, etc, hasta que el segundo campo estaba completo). Estos dos campos eran luego “interlaceados” para formar un cuadro de video. ¡Si la imagen cambiaba entre estos dos escaneos durante el proceso de grabación, las imágenes en las filas pares e impares serían diferentes dentro del mismo cuadro! Para complicar más las cosas tanto PAL como NTSC escanean los campos en diferente orden (impar/par también llamado upper/lower o arriba/abajo). Cuando estos estándares pasaron a ser digitales (DV) la convención en el orden de los campos volvió a cambiar nuevamente. Por esta razón se debe seleccionar la configuración de los campos durante el proceso de edición utilizando los presets de Premiere Pro para la norma en la que uno trabaja.

Sin embargo, al llegar al momento de codificar para la web, el video final debería ser cuadros completo o progresivos, sin ninguna división dentro del cuadro. Si su video original contiene estos campos en lugar de ser progresivo, asegúrese de desinterlacear su video antes de codificarlo. No se debe desinterlacear video que ya fue desinterlaceado o que es progresivo para empezar, hacerlo estropearía el video. Una vez que su video es progresivo, en la configuración de campos se debe usar “ninguna” o progresivo en el codificación final.

## Perfiles

Los Perfiles son sets de características que apuntan a determinadas aplicaciones diferentes. Si bien hay varios perfiles dentro del estándar H.264, las más comunes son Baseline (línea de base) Main (principal) y High (alta). Ayuda a comprender los temas de eficiencia versus rendimiento cuando seleccionamos un perfil. Los perfiles con más capacidades tienden a obtener mejor calidad para un tasa de bits determinado, pero consumen más recursos al implementarse.

Siendo que estos perfiles son simplemente sets de características sobre las capacidades más que opciones sobre una configuración lineal (no confundir con calidad baja/media/alta), estos no son comparables en una escala lineal. A continuación las descripciones de cada perfil proveerán una comparación relativa, ya que la complejidad de los factores que las hacen más eficientes (y más pesados de manejar), o menos eficientes (y menos pesados de manejar) – unido a la forma en que cada capacidad es utilizada en los diferentes escenarios posibles en cualquier video- pueden variar en forma impredecible. Qué perfil utilizar dependerá en gran medida en la aplicación final, y ahí no hay ganadores absolutos.

**Baseline** (línea de base): Este perfil es usado cuando el video es liviano como en el caso de video conferencia o reproducción en teléfonos celulares con limitaciones en poder de cómputo. Provee la compresión menos eficiente entre las tres opciones y el consumo más bajo de CPU en la decodificación.

**Main** (principal): Este perfil es más capaz que Baseline, lo que generalmente se traduce en mayor eficiencia, si bien trae unido un mayor demanda de poder de procesamiento (menos que el perfil High). Este perfil se usa en los videos de calidad media para las aplicaciones de video en la web.

**High** (alta): Este es el perfil más eficiente de los tres. Tiene las mayores capacidades, lo que logra empacar con la mejor calidad en un tasa de bits determinado, y es el más pesado de procesar debido a las operaciones adicionales. Si bien originalmente se pensó solo para las aplicaciones que manejaban alta definición como Blu-ray, este perfil esta poniéndose de moda para video en la web a medida que los usuarios finales promedio acceden a mayor poder de cómputo y más ancho de banda.

## Nivel

La configuración de Nivel es una restricción a la cantidad de trozos que el proceso de decodificación podría encontrar. En la práctica esto se traduce en restricciones en las combinaciones de la velocidad de cuadros y el tamaño de los cuadros. Cuanto más alto el nivel, mayor la restricción. Debido al amplio espectro de aplicaciones que utilizan H.264 (desde pequeños video para teléfonos a videos enormes para cine), se hace necesario a veces filtrar el rango de posibilidades de determinados contenidos para una plataforma o aparato específico, dependiendo de sus capacidades de proceso. Por ejemplo, un dispositivo pequeño con capacidad de procesamiento limitado debería declinar la posibilidad de reproducir un largometraje enorme que involucra cantidades enormes de datos y requiere altas prestaciones de procesamiento, independientemente que el dispositivo eventualmente pudiera interpretar la información recibida (a un ritmo totalmente impráctico) simplemente porque reconoce el formato del video.

Por ejemplo una configuración de nivel 1 restringe la cantidad de “paquetes” de pixeles (macroblocks) de forma tal de permitir solo un video de unos 200 pixeles de ancho, corriendo a 15 fps, mientras que un nivel 3 podría acomodar un video de 720x480 corriendo a 30fps.

Como el nivel define solo un límite y no una combinación de configuraciones absoluta, es posible usar un nivel que permite más que lo requerido. Sin embargo, hacerlo puede restringir la reproducción de su video en un dispositivo que puede hacerlo, si lo confunde con un video de características de mayor volumen de datos. Por lo tanto, incrementa el nivel escalón por escalón hasta que sea capaz de lograr la combinación de tamaño del cuadro y la tasa de bits deseado sin hacerlo innecesariamente alto.

Si bien las configuraciones de los niveles son cruciales para los dispositivos, la reproducción de video en la computadora de escritorio normalmente ignora el nivel que trae el archivo al momento de ser reproducido.

## Modos de tasa de bits de Video: VBR versus CBR

VBR o codificación de Tasa de bits Variable, permite en la compresión definir un valor o tasa de flujo de destino en conjunto con un valor máximo. La idea es que para utilizar eficientemente la compresión para mantener una calidad alta permitiendo picos ocasionales de debido a partes del video con más dificultad para comprimir. Generalmente VBR es más eficiente comparativamente con codificación CBR o de Tasa de bits Constante, cuando hay que empacar un archivo con la máxima calidad para una cantidad de almacenamiento determinada. Sin embargo, permitir que estos picos impredecibles de datos se mantengan a una tasa constante de calidad puede interrumpir la reproducción si los picos se hacen demasiado frecuentes o el límite máximo es muy alto. Por lo tanto, VBR es comúnmente utilizado para obtener las descargas progresivas y basadas en archivos de vídeo en la web. Sin embargo, con la actual oferta de servicios de banda ancha capaces de lograr picos de transferencia mucho más altos que lo que pueden mantener en un nivel constante, VBR se puede convertir en una opción también para hacer streaming.

CBR se usa para los casos de streaming y otras aplicaciones donde un flujo predecible y constante de datos es esencial. Esta predictibilidad tiene un costo, que es no permitir a H.264 utilizar sus capacidades de compresión adaptables para entregar la máxima calidad posible. CBR en alguna manera cambia consistencia por calidad para ganar predictibilidad y una reproducción pareja sin interrupciones o pausas.

Si bien la regla general es que VBR es para descargas progresivas y CBR para streaming, experimentar con ambas puede arrojar resultados contradictorios en determinados casos. Por lo tanto es importante que experimente con su contenido específico en su ambiente específico.

## Número de Pasadas

Esta configuración determina si el video se codificará en una sola pasada (ejecución de compresión) o si el codificador vuelve a revisar el video desde el principio al final, una segunda vez para ver si se puede mejorar la compresión de alguna parte. Esto se puede aplicar a CBR tanto como a VBR. El número de pasadas es uno de esos factores que puede resultar en un mejor “empaque” mientras no tiene ningún impacto en la dificultad de desempacar el video. Generalmente dos pasadas demoran el doble en codificar, pero generan una mejor relación entre calidad y tasa de bits. De todas formas la mayoría de las veces la doble codificación no genera el doble de calidad. Por lo tanto, elija dos pasadas si desea la mejor calidad posible a cualquier precio, ya que esta mayor inversión es solo en el lado de la codificación y no agrega más procesamiento a la hora de la reproducción (no como otros parámetros que impactan tanto la codificación como la decodificación). Sin embargo, si usted cree que la pequeña mejora en la economía de tasa de bits no vale el aumento en el tiempo de codificación, elija una sola pasada. Una vez más, experimente para poder establecer que funciona mejor en su caso particular.

## Configurar la distancia entre key frames (fotogramas o cuadros clave)

Los key frames son cuadros completos provenientes del la fuente original sin ninguna referencia a otros cuadros dentro del video. La distancia entre key frames o que tan seguido aparecen en el video afecta que tan cercano esta el video codificado al original sin comprimir. La frecuencia de estos key frames también afecta que tan bien se puede “navegar” el video (hacer scrubbing). Seleccionar esta opción permite ajustar la configuración manualmente. Generalmente la distancia óptima depende de la cantidad de movimiento en el video y de la velocidad de cuadros. Usualmente se configura entre uno y tres segundos, traducido a cuadros utilizando la tasa de cuadros (por ejemplo para un video en 30fps, un segundo son 30 cuadros).

*AME está dirigido a los individuos y grupos la producción de video, que trabajan en forma permanente (como en un empresa de post-producción), mientras que FMES es una solución de nivel corporativo de gran escala para transcodificación masiva (como un Canal de Televisión que publica en la web).*

## Servidor de Adobe® Flash® Media Encoding Server

Adobe Flash Media Encoder Server (FMES) es una plataforma de codificación de nivel corporativo especialmente diseñada para formatos que son soportados por Adobe Flash Player. FMES también provee una mayor variedad de controles específicos de H.264 en comparación con Adobe Media Encoder, lo que implica mayor flexibilidad y poder con una óptima combinación de configuraciones a medida para su situación específica.

## Configuración FMES

La mayoría de las configuraciones básicas son similares a las de Adobe Media Encoder (disponibles en la sección anterior sobre la configuración de AME). Aquí hay algunas configuraciones relacionadas a H.264 que no están disponibles en AME:

### Stream Básico

**Tipo de Stream:** El tipo de Stream puede ser F4V o MPEG-2. F4V es más adecuado para contenidos del tipo Flash Player.

**Usar modo Streaming:** Esto se puede poner como verdadero o falso. Seleccionando esto como verdadero trae al frente del archivo un set de datos que contiene información sobre cómo interpretar el archivo, lo que es favorable para el streaming.

### Video básico y Tasa de Bits

La configuración de Video Básico y Tasa de Bits es similar a la utilizada en AME, aun cuando la nomenclatura sea un poco diferente. La “Tasa de Bits Máxima” solo se puede definir en el modo VBR, porque CBR solo tiene una tasa fija (llamada simplemente Tasa de Bits de Video aquí).

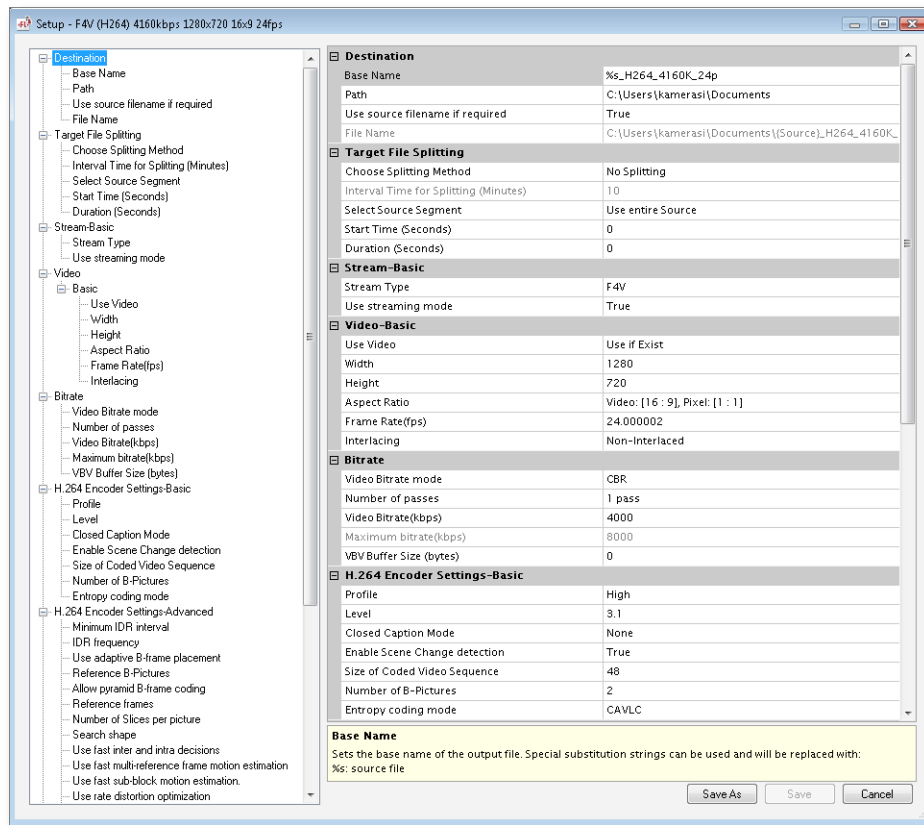


Figura 6: FMES configuración básica

## Configuración Básica del Codificador H.264

Bajo esta categoría se pueden seleccionar los ya conocidos Perfiles y Niveles (ver la explicación en la sección de configuración de AME). De la misma forma que la “Tasa de Bits Máxima”, ciertas configuraciones en esta categoría pueden no estar disponibles, dependiendo de la configuración que se haya seleccionado anteriormente. Por ejemplo, seleccionar Baseline inhabilita la opción CABAC Entropy Coding Mode, porque el perfil Baseline no provee esta funcionalidad.

**Habilitar detección de cambios de Escena:** Configurando esto como verdadero, fuerza al codificador a insertar un cuadro nuevo que define un nuevo tipo de información visual sin ningún parecido con el set de cuadros anteriores. Esto es posible porque H.264 puede definir grupos de cuadros interrelacionados de longitud variada (a diferencia de los viejos compresores que usaban grupos basados en ciclos en el proceso de compresión temporal). Habilitar detección de cambios de escena puede reducir las reducciones repentinas de calidad cuando el video tiene cambios visuales impredecibles; y aun haciéndolo reducirá la eficiencia de tasa de bits y puede a su vez aumentar el tiempo de codificado.

**Tamaño de la secuencia de vídeo codificado:** Esto hace referencia a la longitud de los grupos de imágenes usadas en los métodos de compresión referencial. Los videos que tienen menos cambios pueden permitir valores más altos para esta configuración, a la vez de aumentar la eficiencia de tasa de bits. Poniendo este valor en 1 detiene completamente los métodos de compresión temporal, resultando un archivo con mayor fidelidad con el original y una menor eficiencia en tasa de bits. En la mayoría de los casos un valor de 10 veces la tasa de cuadros es adecuado.

**Cantidad de imágenes B** (No disponible en perfiles Baseline): Esto hace referencia al número de imágenes referenciales insertadas entre imágenes derivadas del video original. Aumentar este número aumenta la eficiencia de tasa de bits con la desventaja de reducir la posibilidad de búsqueda y con un aumento en el poder de cómputo necesario para la reproducción.

**El modo de codificación Entropy:** Este modo puede ser configurado como CAVLC (Basado en el contexto de adaptación de longitud variable de codificación “Context-based Adaptive Variable Length Coding”) o como CABAC (Codificación aritmética adaptable binaria basada en contexto “Context-based Adaptive Binary Arithmetic Coding”). CABAC no está disponible cuando se selecciona el perfil Baseline. Cuando se elige entre los dos modos y se usan los perfiles Main y High, es importante recordar que CABAC es más avanzado, y por eso brinda una mejor economía de tasa de bits/calidad, con la contraprestación de requerir más poder de cómputo comparado con CAVLC. En general, para aplicaciones que requieren mayor calidad como video para web de gran tamaño, pruebe CABAC y vea si se reproduce sin cortes en el ambiente del cliente objetivo. Si el incremento en calidad trae aparejado una reducción en el rendimiento de la reproducción, entonces use CAVLC. (El incremento en poder de cómputo para usar CABAC por lo general vale la pena para la sutil mejoría de calidad para una tasa de bits específica).

## Configuración Avanzada del Codificador H.264

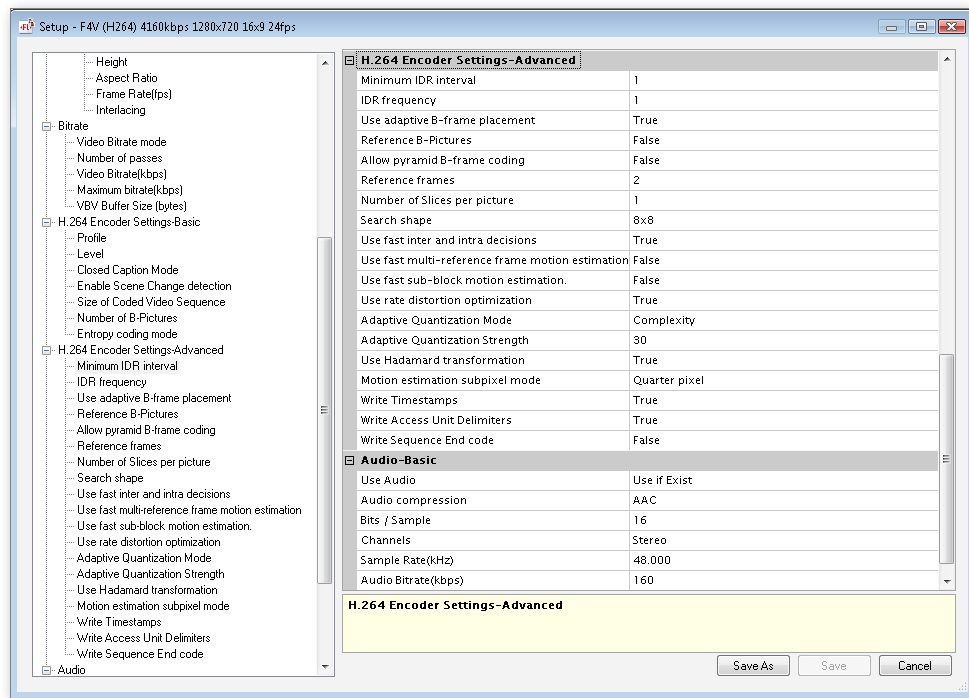


Figura 7: FMES configuración avanzada

Permítame comenzar esta tenebrosa sección diciendo que la mejor regla de pulgar es no tocar esta configuración si realmente no la entiende, ya que solo le agregarían variables a su experimentación. El valor del preset básico que escogió seguramente la servirá. Sin embargo si usted entiende cómo funcionan estas opciones, trataré de alinear sus conocimientos generales y como se aplican a FMES.

Un factor clave a tener en cuenta con estas configuraciones es cuáles afectan solo la codificación y cuáles afectan la codificación y la decodificación o sea la reproducción (análogamente a como algunas formas de empacar resultan en más tiempo empacando con ningún impacto en el desempaque y otras formas complican ambos procesos). En general usted debería aprovechar todo el poder de procesamiento del que disponga para que sus videos se compriman de la forma más eficiente sin sobrecargar el lado de la descompresión y la reproducción. Aun así en algunos casos esto podría resultar poco práctico debido a los largos tiempos de codificación, que finalmente no le reditarán debido a la pequeña o nula mejora en calidad.

Finalmente recuerde que la mayoría de estos algoritmos y técnicas son de alguna forma independientes entre sí, siendo que todas trabajan en diferentes etapas del proceso para lograr un objetivo común de obtener una buena economía de tasa de bits. (Es como usar las notas de “compra otro como este” Y planchar al vapor Y envolver al vacío, todo sobre la misma ropa para lograr la valija mejor empacada de la historia, y a su vez tratar de no destruir su ropa en ninguna de las etapas del proceso a la misma vez.)

**Mínimo Intervalo IDR:** En la vieja escuela había un solo tipo de key frame o cuadro clave (los “I” frames). Ahora en H.264 tenemos dos clases de key frames: I frames “normales” y los nuevos I frames IDR (Instantaneous Decoding Refresh o Decodificación de actualización instantánea). El factor que hace la diferencia entre cuadros IDR y no-IDR o comunes es que, como su nombre lo sugiere, fuerza al decodificador a actualizarse (en otras palabras olvidar que paso en el cuadro anterior) y de esa forma no hacer ninguna referencia a los cuadros anteriores y de ahí para adelante.

Como los IDR lideran cada grupo de imágenes, el intervalo entre un IDR y el siguiente es el tamaño del grupo de imágenes. A diferencia de los códecs tradicionales que usaban grupos

cíclicos de tamaño consistente, H.264 puede tener grupos con diferente cantidad de imágenes. “Mínimo intervalo IDR” define el mínimo número de imágenes en un grupo (o sea que puede fluctuar hacia arriba pero nunca ser menor a este número) Bajar este valor puede incrementar la exactitud del escaneo del video y la resiliencia de errores con la costo en eficiencia de tasa de bits. En algunos casos valores menores puede simplificar la compresión, como resultado se obtiene mejor codificación y mejor rendimiento en la reproducción. Generalmente es mejor dejar el intervalo en 1 y dejar que el codificador decida el menor intervalo en diferentes partes del video. Si por cualquier motivo usted quisiera que a H.264 le fuese restringida la inserción de IDRs muy cercanos, entonces puede aumentar este valor.

**Frecuencia de IDR:** Esta es la parte alta del rango de inserción automático de IDR. Reducir la frecuencia de IDR resulta en una mayor designación de cuadros I como cuadros IDR, lo que mejora los resultados en búsquedas aleatorias (escanear o saltar partes del video) y resiliencia a los errores con la contrapartida de menos economía de tasa de bits. Los IDR se pueden colocar aproximadamente cada 10 cuadros en la mayoría de los casos. Si se requiere una mayor respuesta al escanear, se puede usar un número menor y hasta 1. (Al usar 1 todos los cuadros serán IDR. Sin embargo si se pone cero, solo el primer cuadro de todo el archivo será un cuadro IDR.)

**Uso de cuadros B de colocación adaptable:** Como H.264 es capaz de usar grupos irregulares de imágenes, habilitar esta opción hace un uso completo de esta opción al ubicar los cuadros B basado en que está sucediendo en los cuadros particulares del video que se está codificando. Utilizar esto puede incrementar los tiempos de proceso pero puede mejorar la economía de tasa de bits. Referencia de Imágenes B: Habilitar esta opción permite al codificador usar imágenes B como cuadros de referencia. Cabe señalar que esto solo es posible si el número de imágenes B por secuencia es mayor que uno. Esto puede incrementar la carga de procesamiento como resultado de una mayor economía de tasa de bits.

**Permitir codificación piramidal de cuadros B:** Habilitar esto agrega un nivel más de complejidad al permitir que los cuadros B hagan referencia a otros cuadros B. Esto agrega más procesamiento en la codificación y en la reproducción para incrementar la economía de tasa de bits.

**Cuadros de Referencia:** A diferencia de configuraciones anteriores que afectan la frecuencia de determinado tipo de cuadro, esta configuración define el número de cuadros al que un cuadro resultante hace referencia de forma de construir un cuadro particular. Cuanto más alto el número, se logra mayor eficiencia en la compresión, si bien impacta negativamente el rendimiento. Este número puede ser entre 2 y 16.

**Número de partes por imagen:** Cada cuadro puede ser procesado como una entidad entera (una parte) o puede ser dividido en dos o cuatro partes. Una ventaja de dividir los cuadros es la habilidad de máquinas con múltiples procesadores (CPUs) de procesar una parte por procesador, reduciendo efectivamente el tiempo que demora procesar un cuadro, sin embargo esto puede reducir el ratio de calidad/tasa de bits.

**Buscar forma:** Esto hace referencia al tamaño del bloque usado para la estimación de movimiento. Las opciones son 8x8 ó 16x16. Bloques más pequeños proveen mayor exactitud en la búsqueda de movimiento con la contrapartida de mayor procesamiento durante la codificación, sin impacto significativo en el rendimiento de la reproducción.

**Uso rápido de de decisiones inter e intra:** Habilitar esta opción acelera la codificación con la contrapartida de menor economía de tasa de bits. En la mayoría de los casos el ahorro de tiempo puede ser más importante que la imperceptible pérdida de calidad y una pequeña baja en economía de tasa de bits. Deshabilitar esta opción solo en casos especiales cuando el tiempo no es primordial y se requiere el máximo ratio de calidad/ tasa de bits posible. Esta configuración no tiene ningún efecto sobre el rendimiento de la reproducción.

**Usar estimación de movimiento de cuadros multi-referencial:** Habilitar esta característica puede reducir aún más los tiempos de codificación con un costo pequeño en calidad y economía de tasa de bits. En la mayoría de los casos el beneficio supera las desventajas. Esto tampoco impacta en la reproducción.

**Usar estimación rápida de movimiento de sub-bloque:** Este algoritmo adicional usa un proceso de adaptación para utilizar estimación de movimiento de sub-bloque en forma selectiva, solamente cuando necesita incrementar la velocidad de codificación con mínimas desventajas en términos de pérdida de calidad.

**Utilice la optimización de distorsión de tasa:** Este método optimiza la tasa de bits basado en la diferencia entra la imagen comprimida resultante y la imagen original. Activar esta opción puede incrementar el tiempo de codificación y a la vez mejorar la economía en la tasa de bits. Puede no tener impacto en la reproducción.

**Modo de Cuantización adaptable:** Cuantización adaptable (AQ Adaptive quantization en inglés) es una forma inteligente de cuantización de áreas (bloques) de cuadros dependiendo de la naturaleza de su contenido. Por ejemplo, cuando el área es relativamente plana o continua (un cielo azul o una pared blanca), la mayoría de las áreas se pueden interpretar con menos datos mientras que áreas con detalles complejos obtienen más consideración; así que áreas que pueden producir artefactos son menos comprimidas, en forma selectiva. AQ tiene cuatro posibles opciones:

- Ninguno apaga AQ
- Modo de complejidad funciona tomando la cantidad de detalle en las zonas consideradas
- Brillo suaviza las zonas oscuras planas
- Contraste intenta dar menos importancia a las zonas menos contrastadas

Modo de complejidad funciona para la mayoría de los contenidos, de todas formas usted debería experimentar si su contenido es de un estilo particular.

AQ es una carga agregada a la codificación que mejora la economía en la tasa de bits como resultado, si bien no genera ningún impacto en el rendimiento de la reproducción.

**Potencia de la Cuantización adaptable:** Esta configuración controla hasta qué grado se aplica la Cuantización adaptable (0-100). Si siente que AQ esta sobresimplificando la imagen o ve artefactos significativos en áreas donde trabaja AQ , pruebe bajar el valor; o si siente que la imagen puede beneficiarse más de la función de AQ entonces incremente el valor.

**Utilice la transformación de Hadamard:** Este algoritmo aumenta más la eficiencia de la compresión en áreas suaves y es más efectivo en áreas donde la mayoría del área es suave o continua. Puede agregar un poco mas de proceso tanto en la codificación como en la reproducción.

**Modo de estimación de movimiento de sub pixel:** Esto determina el rigor y la exactitud de la estimación de movimiento. Pixel Completo es el modo más fácil y el menos efectivo, mientras que Cuarto de Pixel es el más difícil y a su vez el que requiere más poder de procesamiento para codificar. Medio pixel está en el medio. Búsquedas más afinadas dan una mejor relación calidad/tasa de bits. Esta configuración solo afecta la codificación y no la reproducción.

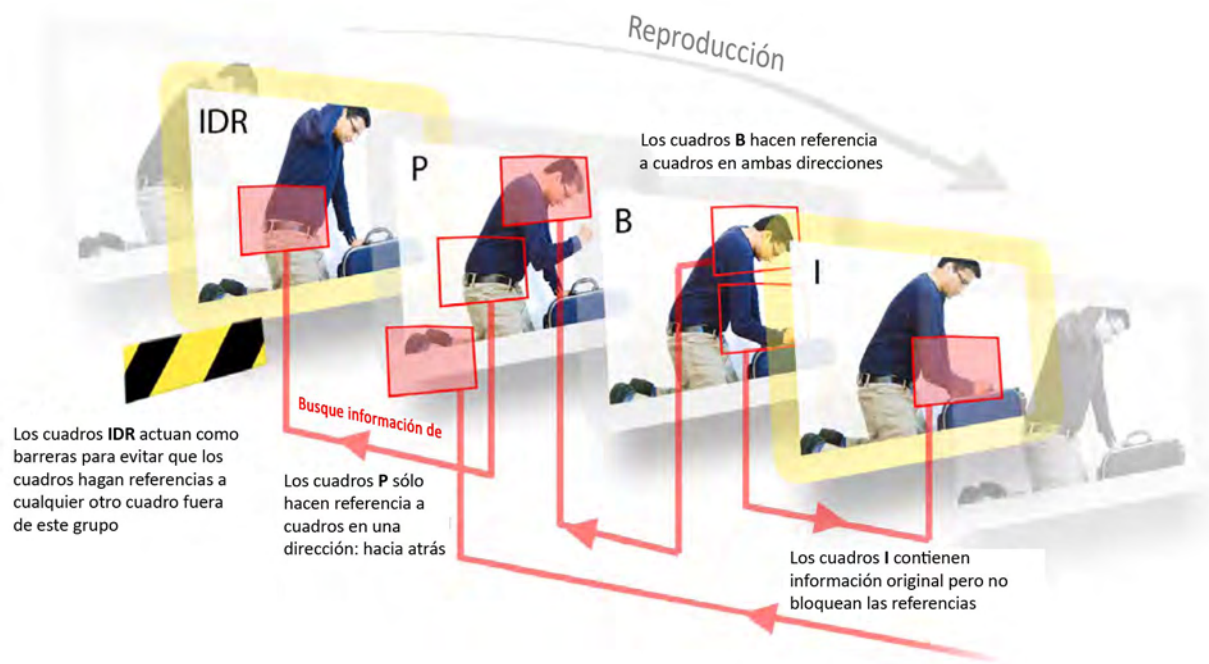
**Escribir Marcas de Tiempo:** No se debe confundir con una función de sobreimpreso de la imagen (en otras palabras esto no imprimirá el timecode en su video), esto es una pieza más de información que ayuda al decodificador a mantener la sincronía entre audio y video. Mientras teóricamente esto agrega datos, su utilidad es mayor que el problema que se puede generar si esta información está ausente. Así que póngalo como verdadero excepto en casos extremos, o en un caso especial en el que se quiera eliminar esta información por alguna razón particular.

**Delimitadores de Escritura de Acceso a la Unidad:** Los Delimitadores de Escritura de Acceso a la Unidad son una pieza de información agregada al archivo que pueden ayudar a cierto tipo de decodificadores a detectar los límites entre cuadros, manejar los cuadros que tienen campos y así sucesivamente. Esto puede ser requerido por algunos estándares de reproducción para mantener la compatibilidad.

**Escriba el código de Fin de Secuencia:** El código de Fin de Secuencia está relacionado a las Unidades de Acceso y puede ayudar a diagnosticar problemas de reproducción debido a incompatibilidades en ciertos ambientes. Esto no es requerido (ponerlo como falso) para la mayoría de las situaciones.

¿Aun confundido con todos estos tipos de cuadros? Aquí hay un diagrama que explica las diferentes funcionalidades y propósitos que tienen las imágenes dentro del grupo en un video comprimido.

Figura 8: Tipos de Cuadros



## El Medidor de Kush

### ¿Su economía de datos está por debajo o por encima del promedio?

Para ayudar a los novatos a tener una idea de qué combinación de configuraciones se podría considerar demasiado exigente en la tasa de datos, para un contenido de un tipo y tamaño, ya sea brillantemente ágil, haciendo quedar mal a los profesionales, o quizás demasiado avaro, o quizás casi perfecto, se me ocurrió un simple marco basado en un poco de experimentación y comparando apuntes con otras personas que también trabajan sobre los mismos temas.

Esto es solamente a modo de referencia y no debería tomarse como una regla o como medida o valor para cualquier uso, ya que su situación particular puede diferir notoriamente de la que yo usé para llegar a estos resultados. Aun así le puede ser útil comparar sus propias pruebas de compresión y los mejores resultados a los que haya llegado con estos; o podrían usarse como punto de partida o rangos hasta que llegue a descubrir sus propios números mágicos.

## Sopesando los factores

Comience considerando los factores que lo llevan a necesitar mayores tasas de bits para lograr un determinado nivel de calidad:

- Cantidad de píxeles en cada cuadro
- Cantidad de cuadros por segundo
- Cantidad de movimiento en la imagen (baja/media/alta)

Calcular la cantidad de píxeles por cuadro es fácil: simplemente multiplique el ancho por el alto. Por ejemplo un video de 640 x 360 tiene  $640 \times 360 = 230.400$  píxeles.

La velocidad de cuadros por segundo se sabe inmediatamente. En este ejemplo asuma que es 30fps. Esta debería ser velocidad de cuadros óptima para que un video fuera aceptable (por ejemplo una demostración en la computadora con captura de pantalla no necesita 30 fps si solo el ratón está en movimiento.)

Considere la cantidad de movimiento (lo llamaremos “rango de movimiento”). Como regla general trate de simplificar esto en tres rangos: Bajo, Medio y Alto. Para definir estos rangos en términos del mundo real:

- **Movimiento bajo:** es un video con mínimo movimiento. Por ejemplo una persona hablando delante de la cámara sin moverse demasiado mientras la cámara y el fondo están quietos.
- **Movimiento medio:** sería con algún grado de movimiento, pero con un orden predecible y ordenado, lo que implica que tanto el sujeto como la cámara se mueven lentamente, sin muchos cambios de escenario por cortes sorpresivos o movimientos repentinos de cámara ni acercamientos, donde la imagen cambia completamente en forma instantánea.
- **Movimiento alto:** sería como una sinopsis de una película de acción, donde no solo los movimientos son rápidos e impredecibles, sino que las escenas también cambian muy rápidamente.

Para convertir este factor tan subjetivo en un número cuantificable, tratemos de dar un factor de multiplicación a cada rango. Como estos rangos no son lineales, elegí los siguientes números para cada rango: Bajo=1, Medio=2, Alto=4. (En otras palabras un video con una cantidad de movimiento razonable es el doble de difícil de comprimir comparado con uno que tiene escaso o nulo movimiento. Un video con cambios rápidos e imprevisibles requerirá de cuatro veces más tiempo para comprimir manteniendo el mismo nivel de calidad).

Dado este multiplicador relativo basado en estos factores, he tratado de desarrollar una base numérica desde donde estos multiplicadores puedan producir estimaciones de tasas de bits del mundo real. Luego de numerosos experimentos, he notado cierto patrón que se podría considerar constante o como un valor de base (para los tamaños y tasas de cuadros más comunes). Al redondearlo ese valor es de 0.07 bps por pixel, por cuadro y por valor de rango de movimiento del cuadro.

En otras palabras, para estimar un valor de tasa de bits óptima para H.264 que resultaría en lo que se considera “buena calidad” resulta para cualquier video, usted puede multiplicar la cantidad de pixeles de destino por la tasa de cuadros ; luego multiplicar por el factor 1, 2 ó 4 dependiendo de rango de movimiento; ese resultado multiplicarlo por 0.07 para obtener una tasa de bits en bps (y dividir eso por 1.000 para tener un estimado de Kbps o por 1.000.000 para tener Mbps).

## Ejemplo práctico

1280 x 720 @ 24 fps, movimiento mediano (rango 2):

$$1280 \times 720 \times 24 \times 2 \times 0.07 = 3.096.576 = \sim 3.000 \text{ kbps}$$

Si el rango de movimiento es Alto (rango 4), queda alrededor de 6.000 kbps.

Por otro lado si el mismo clip se puede usar a 5 fps con movimiento bajo:

$$1280 \times 720 \times 5 \times 1 \times 0.07 = 32.256 = \sim 320 \text{ kbps}$$

Una reducción del tamaño del cuadro puede reducir dramáticamente la tasa de bits necesaria:

$$640 \times 360 \times 5 \times 1 \times 0.07 = 80.640 \text{ bps} = \sim 80 \text{ kbps.}$$

Este ejemplo muestra como estos factores generan diferencias enormes en las tasas de bits entre videos con el mismo tamaño de cuadro y con diferentes tasas de cuadros y rangos de movimiento.

En el caso de CBR, un valor cercano a este estimado se puede usar. En caso de VBR un valor aproximado al 75% del estimado puede ser un buen número y en un 150% si se va a usar el máximo de tasa de bits. Esta diferencia depende mayormente de la naturaleza del contenido y de la habilidad de absorber los picos de tasa de bits en el ambiente de reproducción.

## ¿Qué pasa con el Audio?

H.264 es estrictamente un códec de video. Como se empareja la información del audio con la corriente de video queda a cargo del contenedor (por ejemplo F4V). El códec AAC (Advanced Audio Coding o “codificación avanzada de audio”) es usado cuando el video es H.264. AAC tiene muchas ventajas sobre otros códecs más antiguos de compresión con pérdida de audio como MP3, incluyendo una alta eficiencia (la misma calidad con menores tasas de bits) y características adicionales (más canales para implementar sonido envolvente), una gama más amplia de opciones de tasa de muestreo, y más). AAC también tiene variantes: por ejemplo HE.AAC para mayor eficiencia, con tasas de bits más bajas para hacer streaming.

Igual que con el video, usted debería tratar de tener audio sin comprimir o sin pérdida para comenzar y comprimir solo en la etapa final.

## ¿Mono o estéreo?

El verdadero estéreo requiere dos canales independientes de audio. Por esta razón, usted debería pensar en audio en términos de una pista o dos pistas, más que como una característica agregada. Cuando tenga poco presupuesto de tasa de bits, elija estéreo solamente si el contenido lo necesita verdaderamente. Por ejemplo si es un video musical, el sonido estéreo puede ser importante para una buena experiencia del usuario, pero un video con una persona hablando, donde el contenido se puede consumir en mono, aun si tiene música en la introducción. Si la fuente del material original es mono no tiene ningún sentido codificarlo como estéreo. (Es como grabar una foto blanco y negro en un formato color.)

## Bits por muestra

Así como las fotos digitales son divididas en pixeles con distintos niveles de intensidad (como 256 niveles de gris) el audio digital tiene diferentes escalones en donde la onda de sonido puede estar en un momento determinado. Usando más bits por muestra puede definir un flujo de audio como más cercano a la onda de sonido, ya que tiene más “escalones”. Generalmente, audio de 16 bits es considerado de alta calidad, y todas las formas de reducir la tasa de bits debería ser considerada antes de reducir el valor de bits por muestra.

## Frecuencia de muestreo

La frecuencia de muestreo es el equivalente del audio a la resolución de una imagen. Cuanto más alta la resolución de las imágenes más pixeles que pueden contener detalles más finos, mientras que resoluciones bajas pueden mostrar las grandes formas perdiendo detalles, frecuencias de muestreo más altas permiten más detalle o que las altas frecuencias de los sonidos se puedan guardar con mayor fidelidad.

Como el oído humano no es sensible a frecuencias de más de 20kHz, una “resolución” de audio capaz de poder reproducir una señal así (una onda consistente en una muestra hacia arriba y otra muestra hacia abajo anda por los 40kHz. Entonces una frecuencia de muestreo de 44.1 kHz es adecuada para la mayoría de las aplicaciones. La excepción es audio que solo contenga la voz que se puede codificar en frecuencias de muestreo más bajas como 22.05 kHz, porque la frecuencia útil más alta no llega ni cerca del límite humano de percepción de sonido. Aun así sorprendentemente el rango de frecuencias puede degradar otros sonidos sutiles como sonidos de la respiración y silbidos, lo que podría dar una sensación de pérdida

de calidad de audio. Generalmente las frecuencias de muestreo más altas que 48 kHz no se usan salvo en aplicaciones de audio de alta fidelidad.

## Tasa de bits de audio

Normalmente una corriente de audio consume una proporción menor en la tasa de bits en un archivo de audio y video. Aun en tasas de bits bajas se puede conseguir una calidad de audio razonable. Por ejemplo una pista de música estéreo se puede codificar entre 96 y 128 kbps con una pérdida de calidad entre mínima a imperceptible. En sonido mono tasas de bits tan bajas como 56-80 kbps puede aun ser aceptables, mientras que en un audio que solo contenga voz, usado en aplicaciones que solo pretenden que el audio sea comprensible y sin tener una estética placentera las tasas de bits pueden bajar drásticamente.

## Consideraciones sobre las fuentes de audio

Cuando estamos buscando una tasa de bits optima para un audio en una determinada situación, comience por decidir si la pista de audio será mono o estéreo (¡o sonido envolvente!) primero. Luego experimente con diferentes tasas de bits manteniendo el audio en base a 16-bit 44.1kHz. Si la calidad baja mucho cuando baja las tasas de bits a los niveles que necesita, recién entonces comience a hacer concesiones más importantes (aparte de la decisión de mono o estéreo) como reducir la frecuencia de muestreo o los bits por muestra. Para mejores resultados asegúrese que el audio “está limpio” antes de apretarlo en el compresor. Así como usted corregiría color o el ajustaría los niveles para que la visualización sea lo mejor posible dentro de los colores disponibles usando controles de niveles/brillo/contraste, asegúrese que su onda de sonido utiliza la mayoría de los valores de la amplitud de la gama disponible sin recorte “clipping”, (salirse del rango, resultando en una onda de sonido plana)

## Conclusiones

Comprimir video en H.264 es más un arte que una ciencia, pero tener una comprensión básica de cómo los diversos factores influyen el resultado final puede mejorar notablemente su habilidad para conseguir mejores resultados. Si bien no hay fórmulas mágicas o reglas rigurosas, empezar desde una base sólida e ir gradualmente refinando las configuraciones mediante experimentación controlada pueda dar los mejores resultados en cualquier situación.

Espero que usted haya encontrado útil este viaje por las turbulentas aguas del mundo de la codificación. ¿Es usted un mejor compresor ahora?

## Reconocimientos

Quiero agradecer especialmente a las siguientes personas que me ayudaron con este trabajo:

Benoit Ambry, Karl Soule, Mike Melanson, Desiree Motamedi, Laurel Reitman, Jessica Fewless, Elen Gales (fotografía), Alvaro Calandra (traducción al español), Antonio Lirio..

..y cualquier otra persona que me haya olvidado.



## Acerca del autor

Kush Amerasinghe trabaja como científico informático en Adobe Systems en California. Actualmente utiliza varios sombreros siendo productor de contenidos para Adobe TV ([tv.adobe.com](http://tv.adobe.com)) mientras es el anfitrión de su propio show “Ask the Adobe Ones (pregúntele a los de Adobe)” — un show interactivo y divertido en la web, donde la audiencia le pregunta a él y a otros expertos de Adobe. (Es uno de los programas con más suscriptores que Adobe TV emite en Adobe Media Player.)

Antes de unirse a Adobe, Kush trabajo en variados campos, incluyendo multimedia, Internet, televisión, efectos visuales y educación.

Se le puede contactar en: [ask@adobe.com](mailto:ask@adobe.com)

