
Audio y vídeo digital

Jorge Juan Chico <jjchico@dte.us.es>, Julián Viejo Cortés <julian@dte.us.es> 2011-2020
Departamento de Tecnología Electrónica
Universidad de Sevilla

Usted es libre de copiar, distribuir y comunicar públicamente la obra y de hacer obras derivadas siempre que se cite la fuente y se respeten las condiciones de la licencia Attribution-Share alike de Creative Commons. Puede consultar el texto completo de la licencia en <http://creativecommons.org/licenses/by-sa/3.0/>

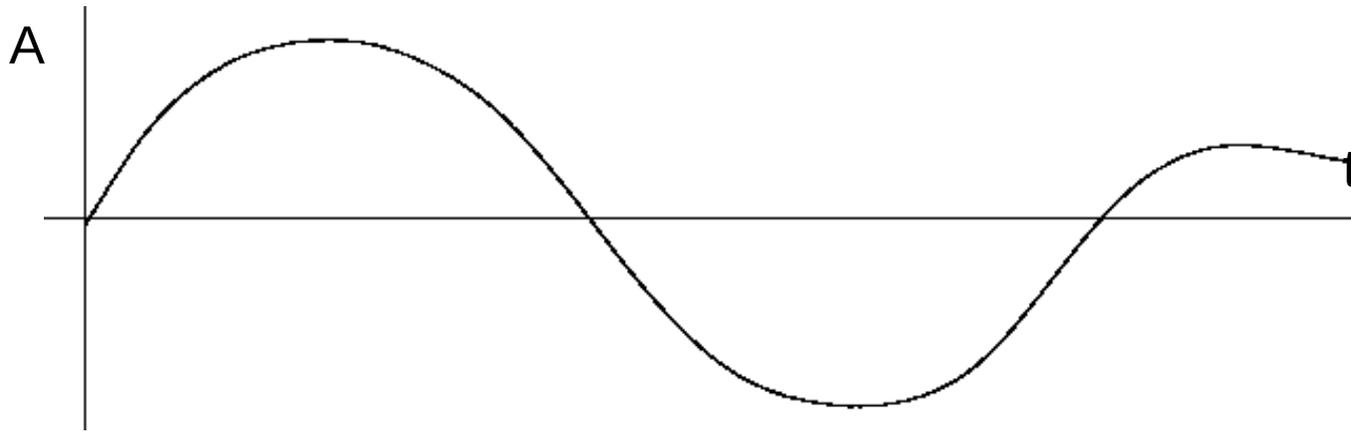
Objetivos

- Comprender los fundamentos de la representación digital del audio y el vídeo.
- Conocer los principales formatos de codificación de datos multimedia.
- Comprender la importancia de los algoritmos de compresión en los formatos multimedia.
- Saber reproducir, transformar y realizar ediciones básicas sobre datos multimedia mediante las aplicaciones adecuadas.

Contenidos

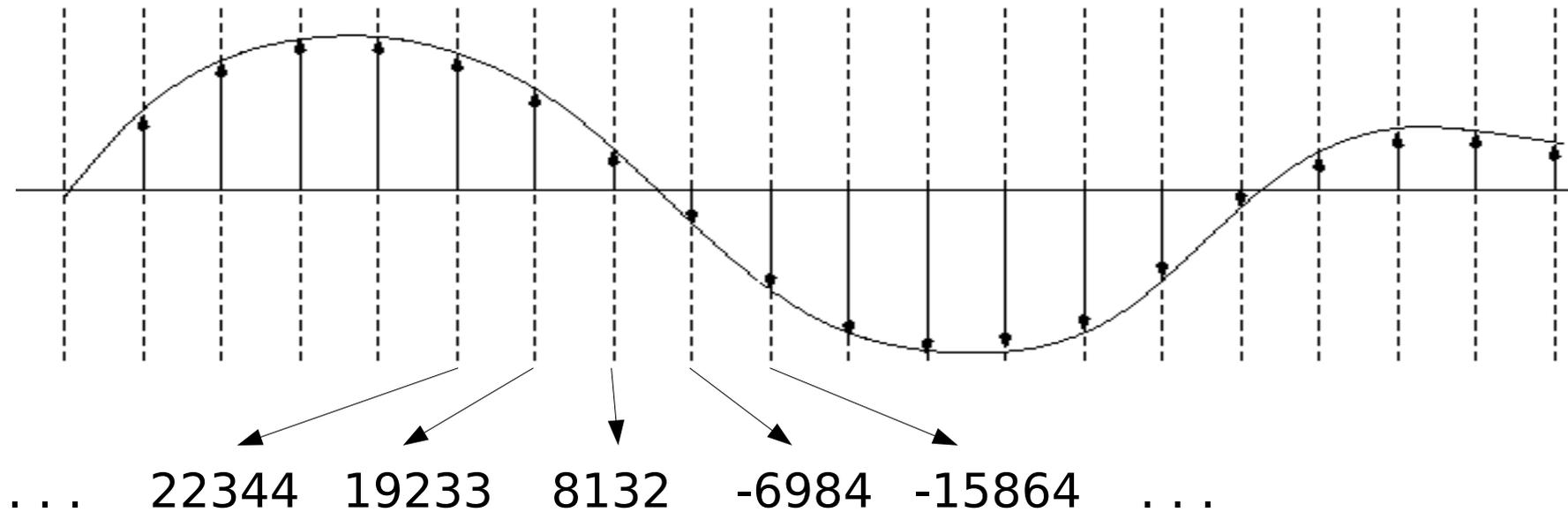
- Audio
 - Fundamentos de la codificación digital del audio.
 - Formatos de audio.
 - Parámetros y cálculos sobre audio codificado.
 - Aplicaciones típicas.
- Vídeo
 - Fundamentos de la codificación digital del vídeo.
 - Formatos de vídeo
 - Contenedores multimedia
 - Aplicaciones típicas.

Audio. Codificación PCM



- El sonido se representa por una curva de amplitud (presión, señal eléctrica, etc.) frente al tiempo.
- Codificación PCM (Pulse Code Modulation):
 - Muestreo: se toma el valor de la señal a intervalos regulares. Frecuencias típicas de muestreo (Hz): 11025, 22050, 44100
 - Cuantización: la amplitud se representa con un número determinado de bits. Ej. 8, 16 bits.

Audio. Codificación PCM



- La calidad será mayor cuanto mayor sea la frecuencia de muestreo y el número de bits por muestra.
- Ejemplo: calidad CD: 44100Hz, 16 bits/muestra, 2 canales.

Audio PCM. Cálculos

- Parámetros
 - Frecuencia de muestreo (f) [Hz] [muestra/segundo]
 - Periodo de muestreo (T) [s]
 - Tamaño de la muestra (m) [bits]
 - Número de canales (n)
 - Duración (t) [segundos]
 - Tamaño de datos (L)
 - Tasa de datos (R)
 - Cantidad de datos par codificar una unidad de tiempo.

$$L = f \times m \times c \times t$$

$$L = R \times t$$

$$R = f \times m \times c$$

Audio. Formatos

- Formato básico (PCM):
 - Se almacenan directamente las muestras de la codificación PCM.
 - Se almacenan los parámetros de la codificación: Frecuencia de muestreo, bits por muestra, número de canales,
 - Puede dar lugar a archivos de gran tamaño.
- Formatos comprimidos
 - Compresión sin pérdidas (ratio típico 2:1)
 - Emplean algoritmos similares a los usado para comprimir otros tipos de datos, pero aprovechando la correlación temporal de los datos.
 - Compresión con pérdidas (ratio típico 10:1)
 - Emplean algoritmos que reducen o eliminan las componentes del audio que se consideran poco perceptibles para la audiencia, reduciendo así el volumen de datos necesario para representar la onda de audio.
 - El factor de compresión depende de la calidad solicitada al algoritmo: mayor calidad → mayor tamaño de datos.

Audio. Formatos

- Codec: implementación de un algoritmos de codificación/decodificación multimedia (audio o vídeo).
 - Incluidos en programas o como bibliotecas de programación

Formato	Tipo	Comentario
MP3	Con pérdidas	Libre (recientemente)
Vorbis	Con pérdidas	Libre
Opus	Con pérdidas	Libre, sucesor de Vorbis
FLAC	Sin pérdidas	Libre
AAC	Con pérdidas	No libre. MPEG, multicanal, con vídeo
AC3	Con pérdidas	Libre, Dolby, multicanal, con vídeo

[Comparison of audio coding formats \(wikipedia\)](#)

Audio. Contenedores

- Los datos codificados se guardan en un archivos con una cabecera que contiene datos sobre el formato utilizado (metadatos):
 - Parámetros usados en la codificación.
 - Título de la pista
 - Autor
 - Fecha
 - ...
- Un contenedor puede contener audio en diferentes formatos pero frecuentemente cada contenedor tiene un formato o formatos asociados.

Audio comprimido. Cálculos

- Tasa de datos (tasa de bits/bit rate)
 - Cantidad de datos empleados para codificar una unidad de tiempo.
- La calidad del audio depende de:
 - La tasa de datos
 - La eficiencia del algoritmos de compresión
- Tipos de tasa de datos
 - Tasa de datos constante
 - Uniforme a lo largo de toda la muestra
 - Calidad variable a lo largo de la muestra
 - Tasa de datos variable
 - Variable a lo largo de la muestra
 - Objetivo: calidad uniforme a lo largo de la muestra
 - Menor tamaño para una misma calidad percibida
 - Se define una tasa de datos promedio para toda la muestra

Tasa vs. calidad.

Ejemplo MP3, Ogg Vorbis

Tasa (kb/s)	Calidad
32	Voz
96	Voz o música de baja calidad
128–160	Música calidad mínima/media
192	Música calidad media
256	Música alta calidad
320	Mayor calidad soportada por MP3
500	Mayor calidad soportada por Ogg Vorbis
1411,2	Tasa de datos PCM CD-DA (referencia)

Bit rate (Wikipedia)

Operaciones típicas y aplicaciones

- Grabar/codificar/reproducir
 - Grabadores/reproductores de audio digital
- Cambios de formato y reducción de tamaño/calidad
 - Convertidores de formato
 - Ej: codificar música para escuchar en el coche.
- Transformaciones del audio
 - Editores de audio
 - Aumentar/bajar volumen, normalizar
 - Filtrado: reducir ruido, eliminar frecuencias, etc.
 - Efectos: acelerar, enlentecer, desvanecimientos, fusiones, etc.
- Mezclar audio
 - Mezcladores multi-pista
 - Post-producción de música/audio de cine, etc.

Contenidos

- Audio
 - Fundamentos de la codificación digital del audio.
 - Formatos de audio.
 - Parámetros y cálculos sobre audio codificado.
 - Aplicaciones típicas.
- **Vídeo**
 - **Fundamentos de la codificación digital del vídeo.**
 - **Formatos de vídeo**
 - **Contenedores multimedia**
 - **Aplicaciones típicas.**

Vídeo. Codificación

- El vídeo digital consiste en una sucesión de imágenes (fotogramas) digitales a un determinado ratio (ej: 25 fps - frames per second-)
- El “tasa de datos” sin compresión es muy elevada
 - Ej: DVD PAL 720x576 pixel, 25fps: 29,7 MB/s = 106920 MB/h
- Todos los formatos prácticos emplean compresión.

Formato	Tasa (MB/s)	Tasa (GB/h)
DVD PAL 720x576, 25fps	31,10 MB/s	112 GB/h
Vídeo FHD 1920x1080, 30fps	186,62 MB/s	672 GB/h

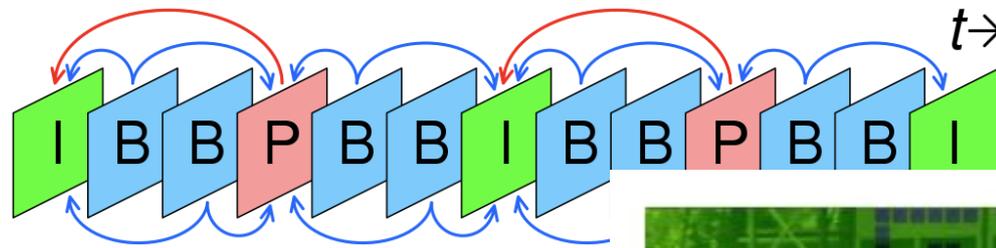
Blu-ray estándar: 25/50 GB

Vídeo. Compresión

- Fotograma intra-codificados
 - Redundancia espacial
 - Compresión de cada fotograma similar a JPEG
- Fotogramas inter-codificados (predictivos)
 - Redundancia temporal
 - Fotogramas adyacentes suelen ser similares
 - Se codifican sólo las diferencias entre fotogramas, teniendo en cuenta posibles transformaciones:
 - cambios de posición (desplazamientos)
 - cambios de orientación (rotaciones)
 - cambios de brillo
 - etc.
- Ratios de compresión típicos: x100-200

Vídeo compresión

Fotogramas inter-codificados



I frame



P frame



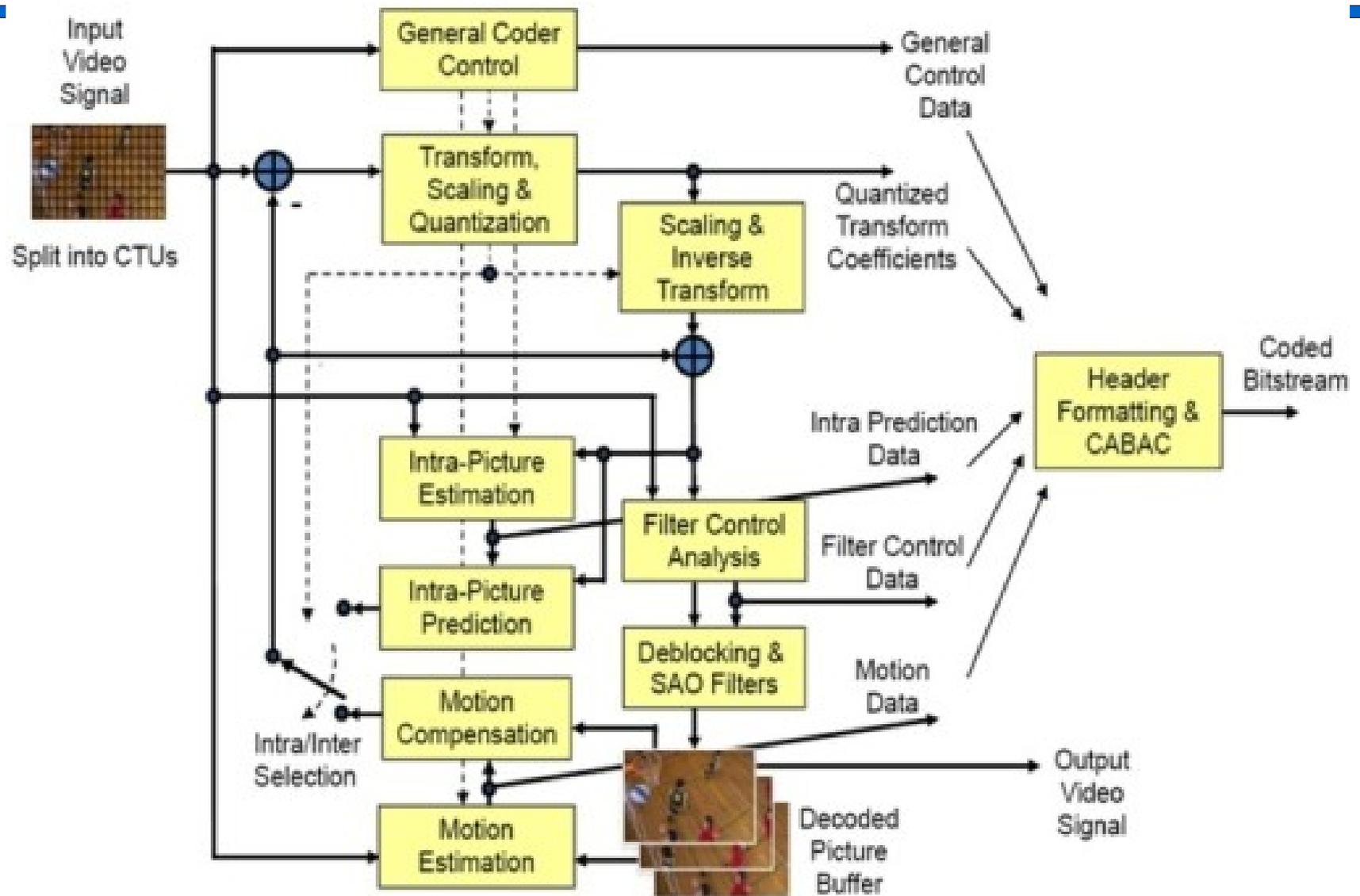
B frame

- Intra
- Inter prediction
- Biprediction
- Direct mode

- Abundance of Direct macroblock in B frames

I: Intra-coded frame
P: Predictive-coded frame
B: Bidirectionally-predictive-coded frame

HEVC - h265 - MPEG H Part 2



Vídeo

Formatos de codificación

- La mayoría de los formatos en uso han sido promovidos por el Moving Pictures Expert Group (MPEG) en forma de estándares
- Los formatos MPEG están protegidos por patentes
 - Pago de licencias por cada codificador (incluso software)
 - Bloqueo a la producción de codificadores libres
- Diversos formatos libres alternativos a los de MPEG han surgido para evitar los problemas de patentes y pago de derechos

Vídeo

Formatos MPEG

Formato	Usado	Comentario
MPEG-1 part 2	VCD, DVD	Calidad similar a VHS Define audio MP3
MPEG-2 part 2 (H.262)	DVD, SVCD, DVB-T (TDT), HDTV	Formato estándar DVD y primera generación TDT
MPEG-4 part 2	HD-DVD, Blu-ray	Usado en ordenadores y reproductores domésticos
MPEG-4 part 10 (H.264, AVC)	DVB-T2 (TDT2)	Formato más extendido en 2019-20
MPEG-H part 2 (H.265, HEVC)	Blu-ray	Sustituto de H.264

Vídeo

Formatos libres

Formato	Similar	Comentario
Theora	MPEG-4 part 2	Obsoleto
Dirac (VC-2)	H.264	Producido por la BBC. Poco usado
VP8	H.264	Producido por Google
VP9	H.265	Sucesor de VP8

Codecs

- Implementación particular de un algoritmo para codificar o decodificar vídeo (o audio).
- Ejemplos:
 - MPEG4 part 2: divx, xvid
 - h264: x264, h264_nvenc
 - h265: x265, hevc_nvenc
 - vp8: libvpx
 - vp9: libvpx-vp9
 - Theora: libtheora
- ffmpeg: programa universal de codificación/decodificación de audio y vídeo. Incorpora multitud de codecs.

```
$ ffmpeg -codecs  
...
```

Formatos de vídeo

Perfiles y niveles

- Perfiles
 - Selección (subconjunto) de características de un formato de vídeo.
 - Un reproductor o codificador de vídeo puede soportar sólo ciertos perfiles: básico, medio, alto, etc.
 - Usar un perfil básico permite la reproducción en más dispositivos.
- Niveles
 - Restricciones de formato: resolución, tasa de bits, FPS, etc.
 - Un reproductor puede soportar sólo ciertos niveles (ej. reproductores portátiles)
 - Si un reproductor soporta un nivel tiene que soportar los inferiores.
- Sólo los perfiles más completos y los niveles más altos permiten la codificación con la máxima calidad y rendimiento (nivel de compresión).

Formatos de vídeo

Contenedores multimedia

- Formato para almacenar juntos vídeo y audio codificados.
- Cada contenedor soporta un conjunto de formatos de audio y vídeo.
 - Genéricos: soportan múltiples formatos.
 - Específicos: soportan combinaciones concretas de formatos.
- Combinan varios flujos de información (streams)
 - Vídeo
 - Audio
 - Subtítulos
- Realizan el “multiplexado” de los flujos.
- Almacenan los metadatos de los flujos (formato, parámetros, etc.)

Formatos contenedores (Wikipedia)

Formatos de vídeo

Contenedores multimedia

Contenedor	Vídeo (ej.)	Audio (ej.)	Comentario
AVI	MPEG4-2, ...	AC3, MP3, ...	Microsoft, en desuso
MP4	H.264, H.265, ...	AAC, AC3, ...	MPEG, muy extendido
Matroska (MKV)	(muchos)	(muchos)	Abierto, universal
WebM	VP8, VP9	Vorbis, Opus	Google, abierto, Web

Formatos contenedores (Wikipedia)

Cálculos

- Tasa de datos (“bit rate” cuando es bits por segundo)
 - Número de datos / tiempo de reproducción.
 - Determina el tamaño de la reproducción y/o archivo contenedor.
 - Cada “stream” tiene una tasa de datos propia.
 - La tasa de datos total es la suma de las tasas de datos de cada “stream”.
 - La tasa de datos es una medida de la calidad del “stream”: más datos → menos compresión → más calidad.
- Tasa de datos constante
 - Misma tasa para cada fragmento.
 - Fácil de decodificar y codificar (1 pasada)
- Tasa de datos variable (tasa promedio)
 - Tasa variable según complejidad de datos
 - Mejor calidad (adaptativo)

Operaciones típicas y aplicaciones

- Reproducción multimedia
 - Reproductores, centros multimedia, etc.
- Captura de audio/vídeo
 - Cámaras, grabadores de escritorio, etc.
- Cambios de formato: calidad, resolución, etc.
 - Recodificadores
- Edición de vídeo: unir tomas, combinar audio, etc.
 - Editores no lineales
- Emisión en red
 - Servidores de streaming (webcast)

Algunas aplicaciones

- Audacity
 - Editor de audio
 - Modificación de pistas, combinación de pistas (limitado), cambios de formato (limitado).
- Handbrake
 - Recodificador de vídeo
 - Cambios de formato, de calidad, adaptación a dispositivos, etc.
- Avidemux
 - Recodificador de vídeo y edición básica
 - Cambios de formato, recorte sin pérdidas de calidad, etc.
- Shotcut
 - Editor multimedia no lineal
 - Combinación de tomas, mezcla de audio, mezcla de fotos, efectos, transiciones, títulos, etc.