

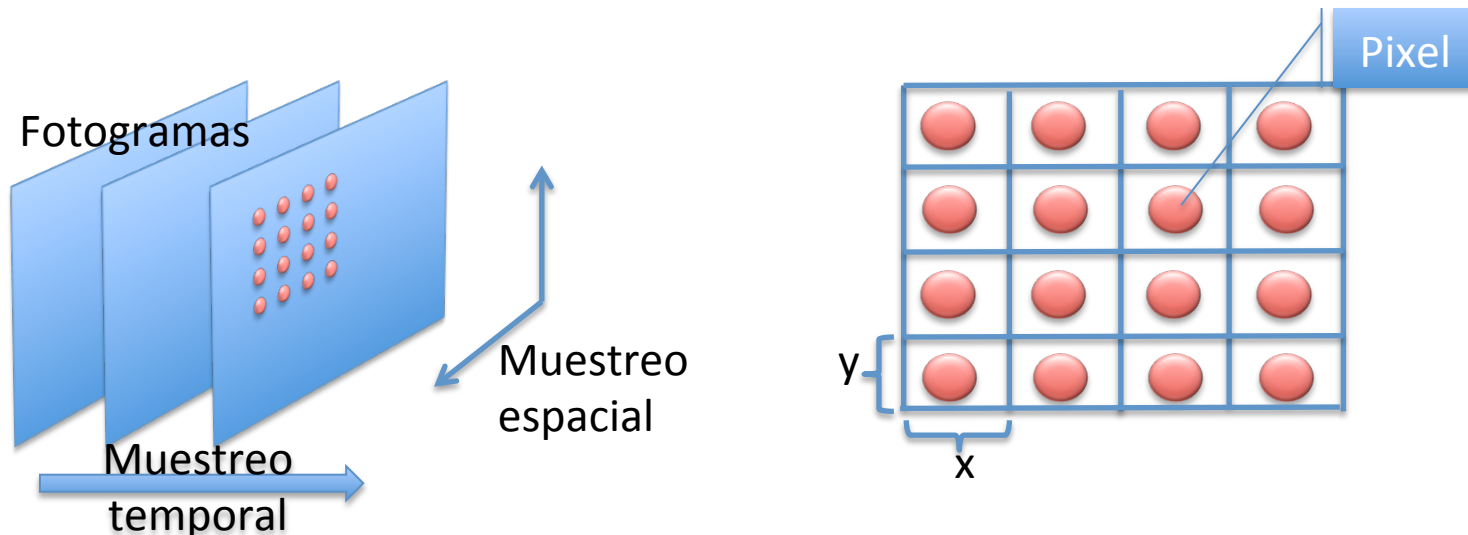
# **Codificación multimedia**

Departamento de Sistemas Telemáticos y Computación  
(GSyC)

**Curso 2012/2013**

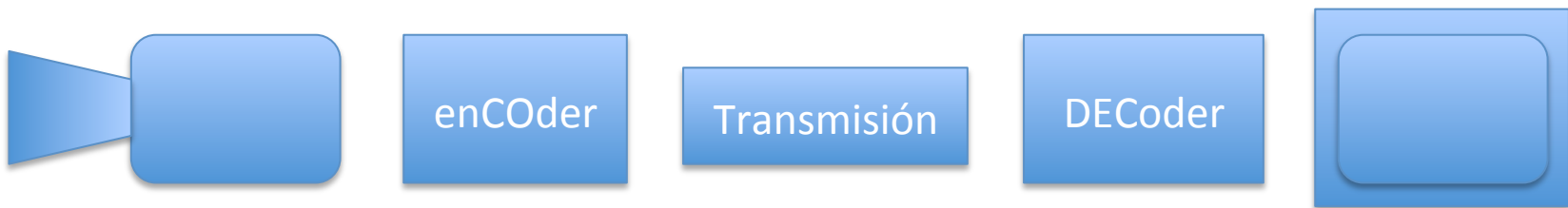
# Captura de vídeo

- La captura de vídeo requiere un muestreo en el tiempo y en el espacio para su representación digital
- Cada muestra temporal se denomina fotograma (**frame**) y es una imagen estática del momento de la captura
- Para cada *frame* se capturan un conjunto de muestras según una rejilla regular (normalmente rectangular). El área de influencia de cada muestra se denomina **pixel** (*picture element*). Cada pixel viene definido por sus coordenadas en el plano del fotograma y por su valor.
- Se denomina relación de aspecto del pixel al cociente entre la dimensión horizontal y vertical



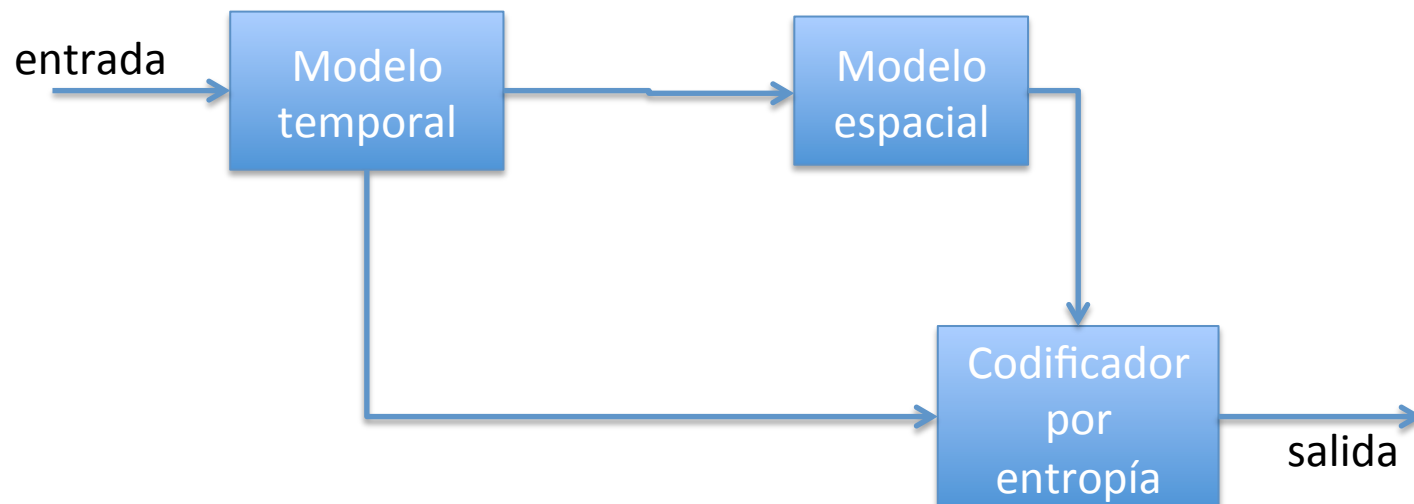
# Compresión de vídeo

- ¿Por qué es necesario comprimir el vídeo?
  - Una señal PAL de televisión requiere 216 Mbps
  - Un DVD puede almacenar 21 segundos de señal PAL
- El proceso de compresión de vídeo compacta la señal digital reduciendo el número de bits, permitiendo de esta forma su almacenamiento y transmisión.
- La compresión consta de dos sistemas complementarios, el codificador (compresor) y el decodificador (decompresor) El par codificador/decodificador es habitualmente nombrado como el *CODEC* (enCOder/ DECoder).



# Compresión de vídeo

- La compresión de datos se obtiene eliminando redundancia. Existen dos modalidades:
  - **Compresión sin pérdidas:** Es un proceso de compresión reversible que permite obtener la imagen original a través de la decodificación. Permite obtener relaciones de compresión de un orden de magnitud
  - **Compresión con pérdidas:** La señal codificada no es idéntica a la original. Permite el aumento del ratio de compresión a costa de una pérdida de calidad en la imagen
- La compresión elimina redundancia temporal (entre fotogramas) y espacial (entre píxeles)
- El nivel de redundancia de un vídeo se mide a través de la correlación temporal y espacial entre muestras. Una mayor correlación implica una mayor similitud



# Encapsulado de vídeo

- El resultado de la compresión de media se encapsula por fotogramas. Cada CODEC define sus propios formatos, pero en general se puede esperar encontrar la siguiente información



Extracto de la estructura de encapsulado de fotogramas usada por h.263

# Encapsulado de vídeo

- ¿Qué información podemos esperar en cada PDU de media?
  - Código de inicio y fin de imagen: Un patrón de bits que señalan el principio de una imagen
  - Referencia Temporal: TS de muestreo, que permite establecer el momento de la presentación
  - Formato de vídeo: Tamaño de pantalla
  - Tipo de fotograma: I, P o B
  - Mecanismo de predicción
  - Vector de movimiento
  - Cuantizador
  - Perfil
  - Vídeo codificado
  - ...

# Codecs de vídeo más comunes

- **MPEG2:** Publicado en 1996, se usa de forma extensiva en televisión digital (TDT)
- **H263:** Publicado en 1995, resulta muy fácil de codificar, por lo que se ha usado muy extensamente en sistemas móviles
- **Spark Sorenson:** Es una variación de H263. Ha sido durante mucho tiempo el Codec de referencia de Flash Video (FLV)
- **MPEG4:** Es una evolución de MPEG2 que ofrece mejor calidad y funcionalidades.
- **H264:** Publicado en 2003, ofrece una alta calidad y ratio de compresión, aunque requiere muchos recursos de procesado. Se está convirtiendo en el Codec de referencia para Internet y redes móviles
- **Theora:** Codec sin patentes ni royalties. Ofrece una calidad algo inferior a MPEG4
- **VP8:** Es un codec propiedad de Google, pero libre de patentes. Ofrece una calidad algo inferior a H264

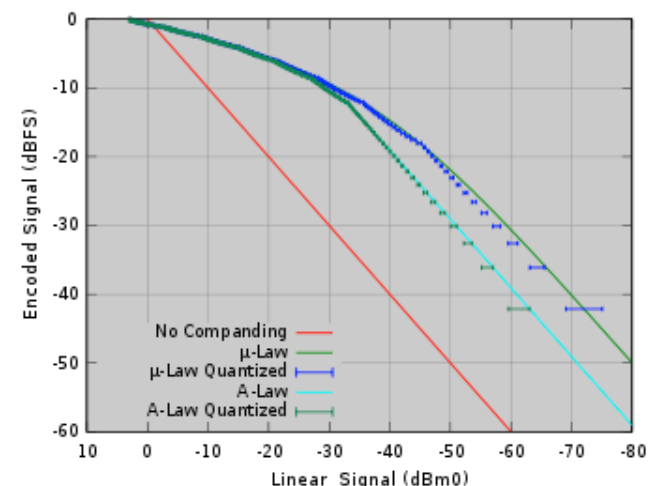
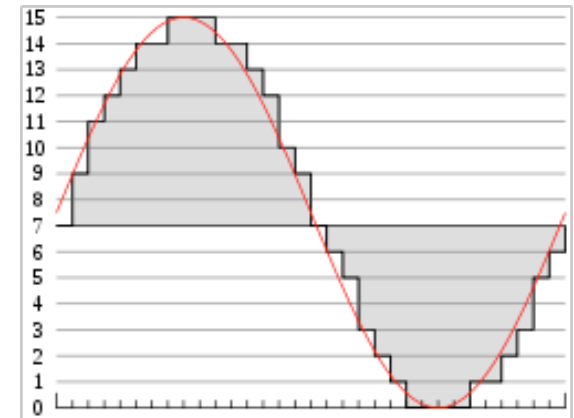
# Captura de audio

- La captura de audio requiere un muestreo en el tiempo de la onda de presión
  - ¿Requiere también un muestreo espacial?
- El audio digitalizado se caracteriza por
  - Muestras por segundo
  - Bits por muestra
  - ¿Se usa el bit rate como una característica del audio?



# PCMA – PCMU

- **PCM:** Pulse Code Modulation
  - Es el primer formato digital de audio
  - Asociado a telefonía
  - 8000 muestras por segundo
  - 8 bits por muestra
  - 64 kbps
- Se establece un paso de cuantización variable (truco equivalente al uso de YUV en vídeo)
  - Ley A: Europa
  - Ley U: EEUU
- Se incrementa el SRN con el mismo consumo de recursos



# LPCM

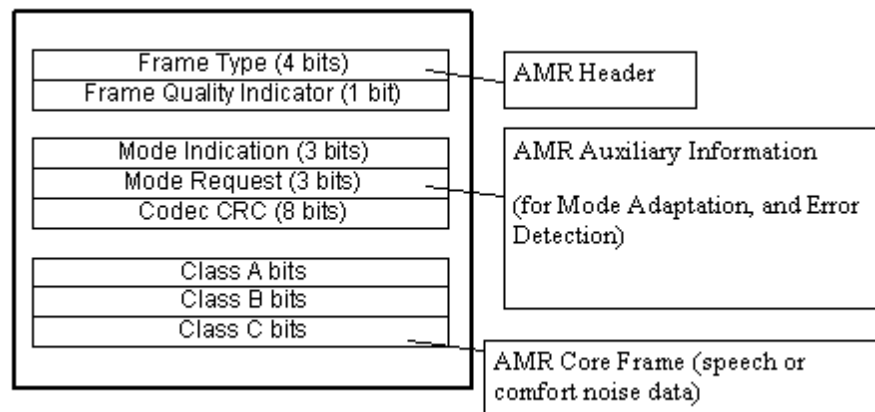
- PCM es el primer formato de audio elegido para distribución multimedia
  - ¿Dónde se usa por primera vez?
- Se mejoran las características
  - Frecuencias de muestreo superior: 11.025, 16.000, 22.050 ... 44,100, 48000, 192k
  - Bits por muestra: 16, 24, 32, 48
- CD usa 44,1k a 16 bits
- DVD usa 192k a 24 bits

# Encapsulado de audio

- Las muestras de audio codificado se agrupan en frames, igual que el video
- Cada frame contiene un fragmento de audio con una longitud fija
- El frame debe ser corto para evitar latencia
  - **PCM**: No existe
  - **AMR**: 20ms con 160 muestras a 8KHz
  - **SPEEX**: 20ms con 160 muestras a 8KHz
- ¿Qué información podemos esperar en un frame de audio?

# Formato de frame AMR

- Código de tipo de frame
- Indicador de calidad
- CRC
- Audio comprimido
- ¿Dónde está el TS?



# Codecs de audio más comunes

- **PCM:** Diseñado originalmente para telefonía. En la actualidad se usa en distribución de audio de alta calidad. No tiene pérdidas
- **AMR:** Codec de audio de bit rate variable usado en telefonía (UMTS)
- **SPEEX:** Codec de audio sin royalties ni patentes
- **AAC:** Sucesor del MP3. Diseñado para proporcionar alta calidad. Se usa en la actualidad junto con H263 en la mayor parte de servicios de streaming.

# Contenedores multimedia

- Los contenedores multimedia permiten encapsular y organizar flujos de media
- ¿Para que sirven?
  - Permiten agrupar canales y multiplexarlos como en una unidad coherente
- Multiplexación de media: Aglutina vídeo audio y subtítulos
  - Audio y vídeo
  - Audio estéreo
  - Video con audio 5.1 con subtítulos
- Organización. Permiten establecer divisiones o capítulos y generar menús de visualización
- Identificación. Proporciona información detallada sobre el CODEC de vídeo y los parámetros de codificación. FourCC
- Indexación. El contenedor marca el principio, el final y la duración del media, permitiendo realizar de forma rápida saltos dentro del flujo (SEEK)
- Incorporan meta-información
  - Características del media
  - Información genérica
    - Autor
    - Título
    - Fecha de captura
    - Etc.
- Cada contenedor soporta un número limitado de formatos

# FourCC

## **FourCC = Four Character Code**

- Es una cadena de 4 caracteres que identifica un CODEC
- Se añade a la cabecera de los contenedores para identificar el formato del media encapsulado
- Se usan exclusivamente caracteres de tipo *Printable ASCII* para que sean legibles
  - DIVX, XVID, H264, DX50
  - <http://www.fourcc.org/codecs.php>
- Los códigos FourCC solo aplican a CODECs de vídeo. Para audio se definen TAGs, un número hexadecimal de 32 bits
- Originalmente pensado para el contenedor AVI, en la actualidad se usa en multitud de contenedores

# Contenedores más comunes

- **AVI:** Creado por Microsoft en 1992. Presenta limitaciones importantes, como la imposibilidad de definir la relación de aspecto o permitir al CODEC usar B-frames
- **MP4:** MPEG-4 Part 14. Es una evolución del contenedor QuickTime (mov). Presenta una modalidad específica para audio con extensión m4a
- **VOB:** MPEG2-PS. Es el contenedor de almacenamiento usado en CDs y DVDs. Permite definir menús de gran complejidad.
- **3GP:** Es un contenedor definido por el 3GPP y específico para almacenar contenidos móviles. La aparición de los nuevos smartphones de Android y Apple ha hecho que se reduzca su uso
- **FLV:** Flash Video. Es un contenedor propietario de Adobe Flash. El único soportado por el plugin Flash. Presenta importantes limitaciones técnicas y un conjunto de CODECs soportado muy reducido
- **MKV:** Es un contenedor basado en estándares abiertos sin patentes ni royalties. Está diseñado para soportar cualquier tipo de CODEC. Es el contenedor con más uso en la distribución de vídeo de alta definición
- **WebM:** Es un contenedor de media abierto que soporta pistas de vídeo codificadas con VP8 y audio en formato Vorbis. Su diseño se basa en el formato Matroskay está especialmente diseñado para streaming.



# Audio Video Interface (AVI)

- Defines bloques dentro del contenedor.

```
RIFF ('AVI '
  LIST ('hdr1'
    'avih'(<Main AVI Header>)
    LIST ('strl' 'strh'(<Stream header>)
      'strf'(<Stream format>)
      [ 'strd'(<Additional header data>) ]
      [ 'strn'(<Stream name>) ] ... ) ... )
  LIST ('movi'
    {SubChunk | LIST ('rec '
      SubChunk1
      SubChunk2
      ...
    )
    ...
  }
  ... )
  [ 'idx1' (<AVI Index>) ]
)
```

# Matroska (MKV)

- Aceleración de búsquedas dentro del fichero
- Soporte para capítulos
- Soporte completo de meta-información
- Selección de pistas de audio, video y subtítulos
- Modular, permite generar extensiones de forma sencilla
- Resistente a fallos. Permite recuperar la reproducción incluso con el flujo dañado
- Capacidad para transmisión a través de internet.
- *Soporte para menús (igual que el formato DVD)*

# Bibliografía

- H.264 and MPEG-4 Video Compression Iain E. G. Richardson. The Robert Gordon University, Aberdeen, UK. Wiley
- Recommendation ITU-R BT.601-5, Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios, ITU-T, 1995.
- <http://www.fourcc.org/>
- <http://www.speex.org/docs/>
- 3GP
  - <http://www.3gpp.org/>
- AVI
  - <http://msdn.microsoft.com/en-us/library/ms779636%28VS.85%29.aspx>
  - <http://tools.ietf.org/html/rfc2361>
- MKV
  - <http://www.matroska.org/technical/specs/index.html>
  - <http://www.matroska.org/technical/diagram/index.html>