

Licenciatura em Engenharia Informática

Sistemas Multimédia

Compressão de Dados com Perdas (Som, Imagem e Vídeo)

Telmo Reis Cunha

Departamento de Eletrónica, Telecomunicações e Informática

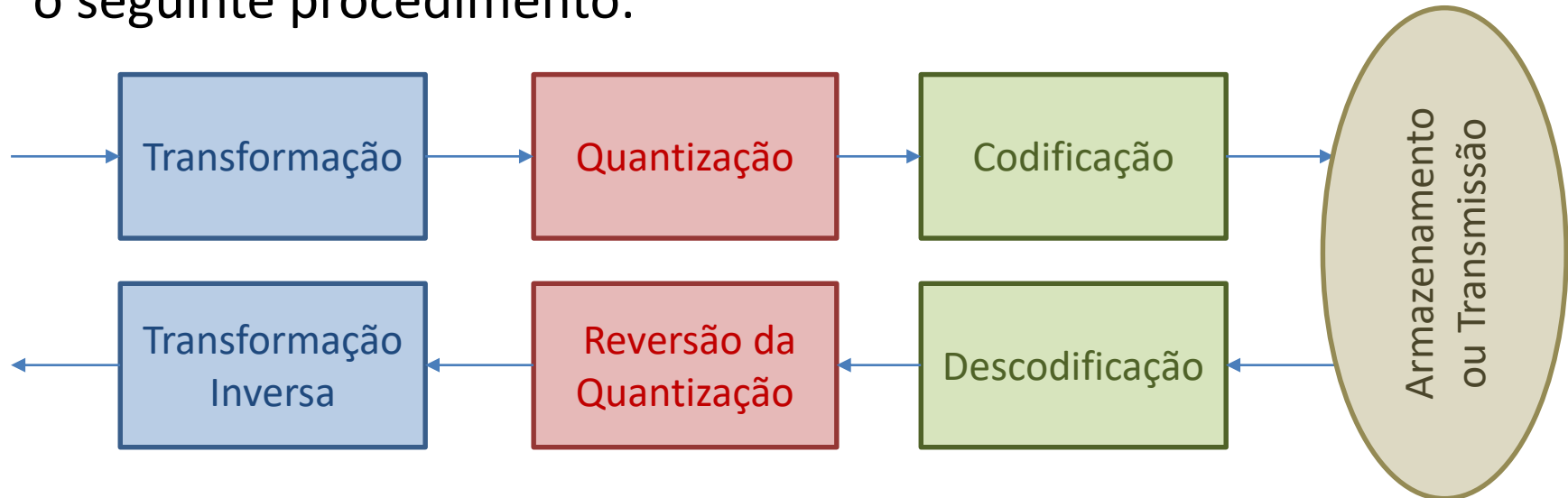
Universidade de Aveiro – 2018/2019

1. Compressão de Dados com Perdas

- Na transmissão e armazenamento de som, imagem e vídeo torna-se fundamental comprimir substancialmente a informação.
- Tal implica o recurso a perdas de informação, relativamente aos dados originais.
- As técnicas de compressão com perdas consideram informação específica sobre a fonte geradora dos sinais, e também sobre o recetor (destinatário da informação).
- Por exemplo, o conhecimento sobre o funcionamento do ouvido humano pode ser usado para remover informação de um sinal de áudio (informação que o ouvido não iria interpretar).

1. Compressão de Dados com Perdas

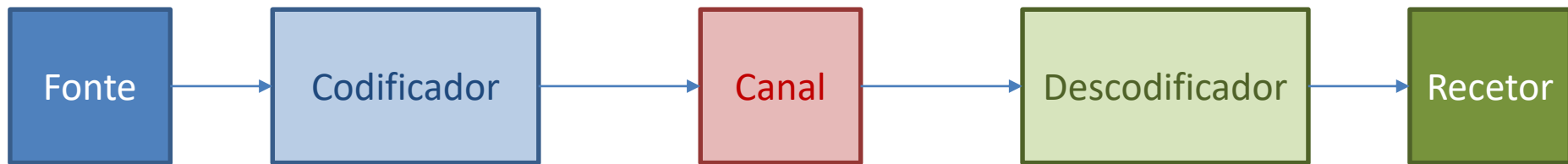
- Muitos processos de compressão de dados com perdas seguem o seguinte procedimento:



- Transformação:** Operação no domínio da frequência (por exemplo, filtrar determinadas bandas).
- Quantização:** Alocar diferentes números de bits a determinadas bandas de frequência (usado no áudio, por exemplo).
- Codificação:** Aplicação de códigos de Huffman, ou aritmético.

2. Compressão de Áudio

- Na compressão de sinais de áudio, a serem transmitidos entre dois interlocutores, são usualmente tidas em conta duas características: **Redundância** e **Irrelevância**.



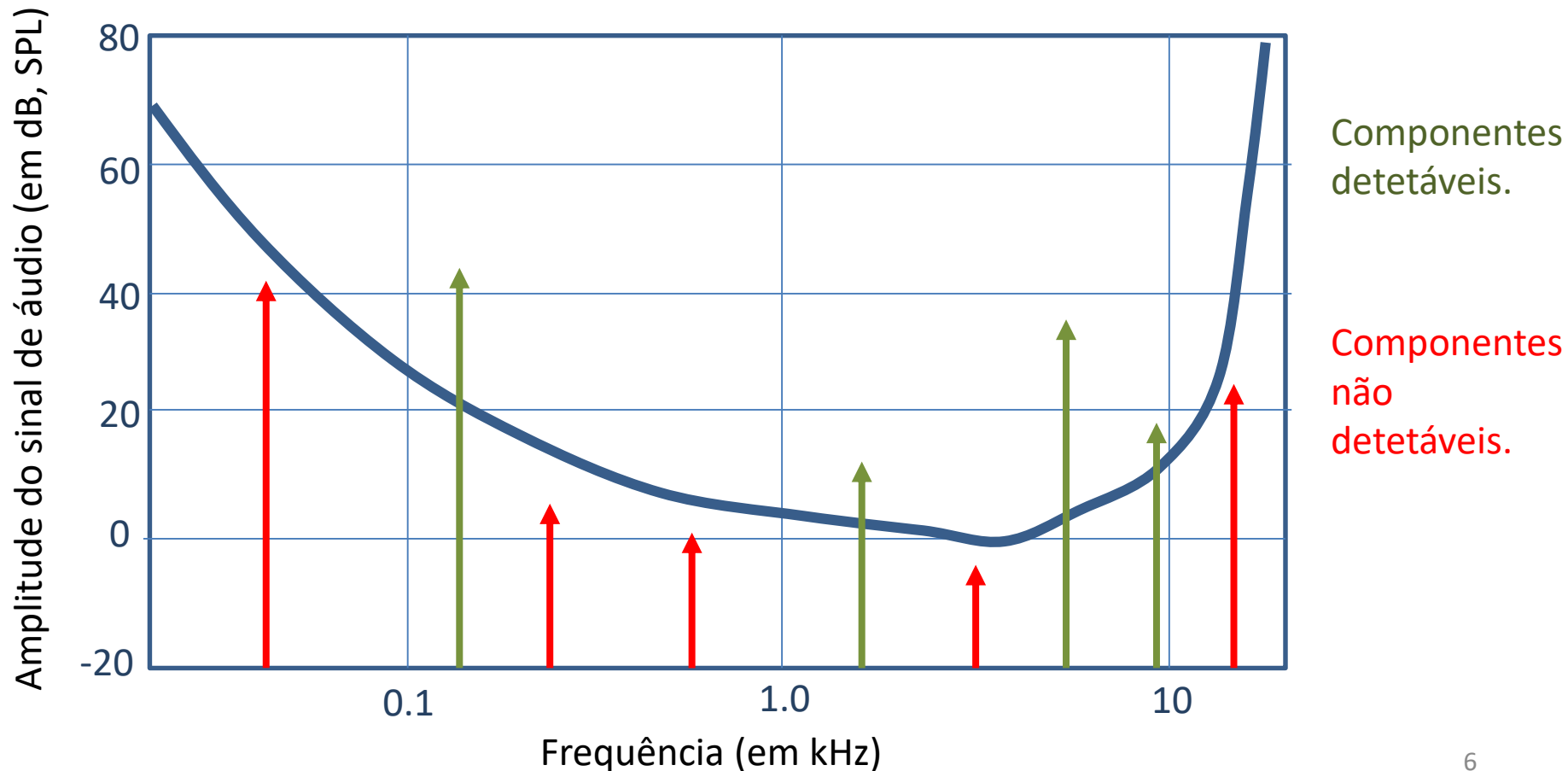
- A eliminação de componentes redundantes no sinal implica conhecimento sobre as características de geração desse sinal.
- A eliminação de informação irrelevante pressupõe conhecimento sobre a forma como o recetor interpreta a informação.

2. Compressão de Áudio

- A codificação de música de elevada qualidade considera unicamente a **irrelevância** associada às limitações do ouvido humano.
- A **Psicoacústica** é a disciplina que estuda a relação entre a fisiologia auditiva e a acústica, conjugando as características dos sinais sonoros com a sensação auditiva que eles provocam.
- Estes estudos permitem identificar modelos que são usados na compressão de sinais de áudio.
- Normalmente, são três as características principais consideradas:
 - Limiar Auditivo;
 - Mascaramento na Frequência;
 - Mascaramento no Tempo.

2. Compressão de Áudio

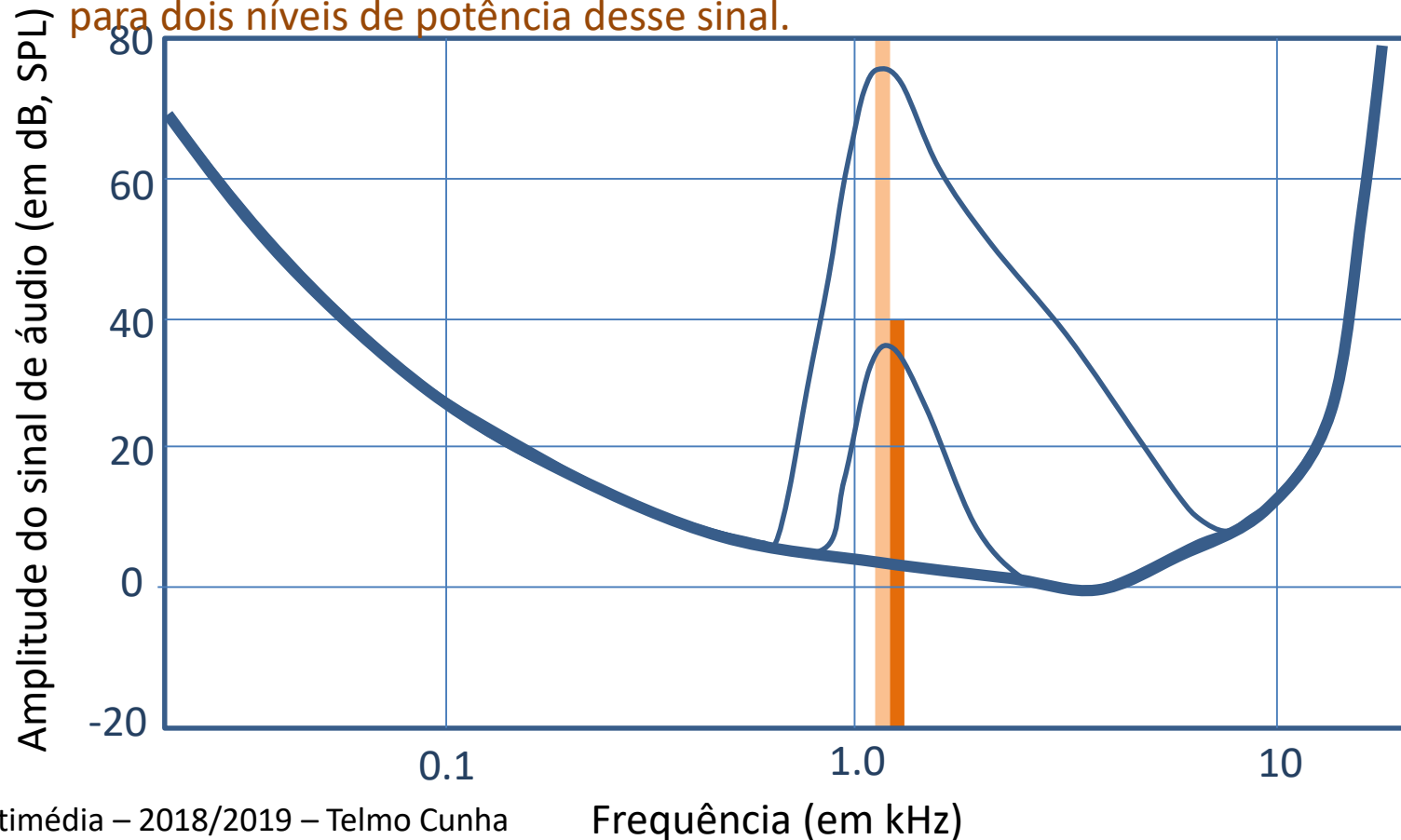
Limiar Auditivo: nível mínimo de pressão acústica eficaz necessária para provocar uma sensação auditiva, num ambiente silencioso.



2. Compressão de Áudio

Mascaramento na Frequência: A presença de sinais sonoros em determinadas bandas de frequência diminui a capacidade do ouvido para detetar sinais a outras frequências.

Variação do limiar auditivo na presença de um sinal de áudio em torno de 1.2 kHz, para dois níveis de potência desse sinal.



2. Compressão de Áudio

Mascaramento no Tempo:

- A presença de um sinal de amplitude elevada diminui a capacidade de deteção de sons mesmo algum tempo após deixar de ser aplicado o sinal forte.
- Também a sensibilidade à deteção de um determinado sinal depende, também, da sua duração.
- Por exemplo, sinais de duração inferior a 50 ms são mais dificilmente detetados do que os de duração entre 50 e 400 ms.
- Estas características inerentes ao ouvido humano são usadas para reduzir adequadamente a informação de sinais áudio recolhidos, sem que o recetor o sinta.
- Estes algoritmos entram em conta com o **espectrograma** do sinal.

2. Compressão de Áudio

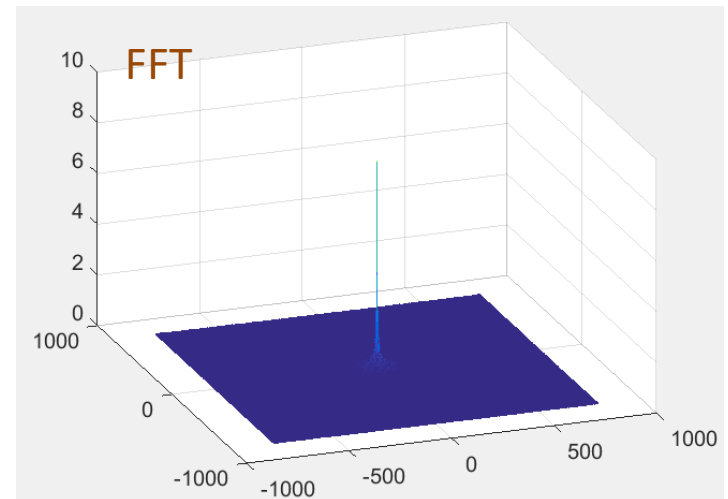
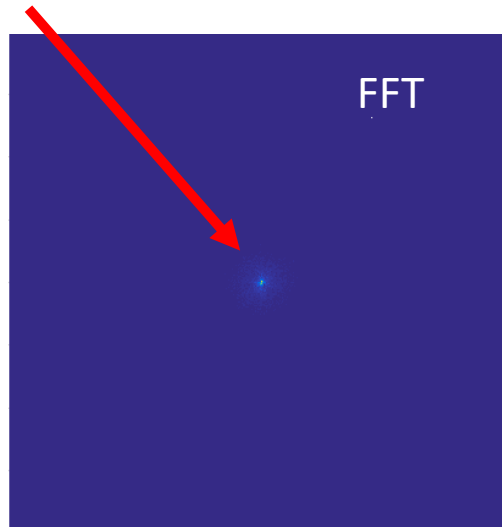
MPEG-I:

Modelo psico-acústico

Camada (Layer)	Máscara no Tempo	Máscara na Frequência	kbit/s	Notas
1		x	30 - 448	Philips DCC
2	x	x	64 - 256	DVD, DAB
3 (MP3)	Modelo mais complexo		64 (mono)	Maior taxa de compressão

3. Compressão de Imagem

- São várias, e variadas, as técnicas de compressão de imagem, envolvendo, em diversos casos, a perda de alguma informação.
- Uma abordagem comum consiste numa quantificação reduzida das componentes de elevada frequência, por se verificar que, na maior parte dos casos, predominam as componentes de frequências mais baixas.



(A componente dc foi removida.)

3. Compressão de Imagem

- Também as características do sensor ótico humano são consideradas para se proceder a uma maior compressão de imagens, sem que se note uma diminuição da qualidade da imagem.
- Por exemplo, o olho humano não é igualmente sensível às cores R, G e B, pelo que se tira melhor proveito se se usar uma outra forma de representar as cores, em vez do RGB.
- O olho humano é mais sensível à intensidade de luz de cada pixel (luminância) do que à sua cor (crominância).
- Assim, é usual transforma-se a representação RGB para a YC_bC_r (uma matriz de luminância (Y) e duas de crominância (C_b e C_r)).

3. Compressão de Imagem

R



G



B



Y



C_b

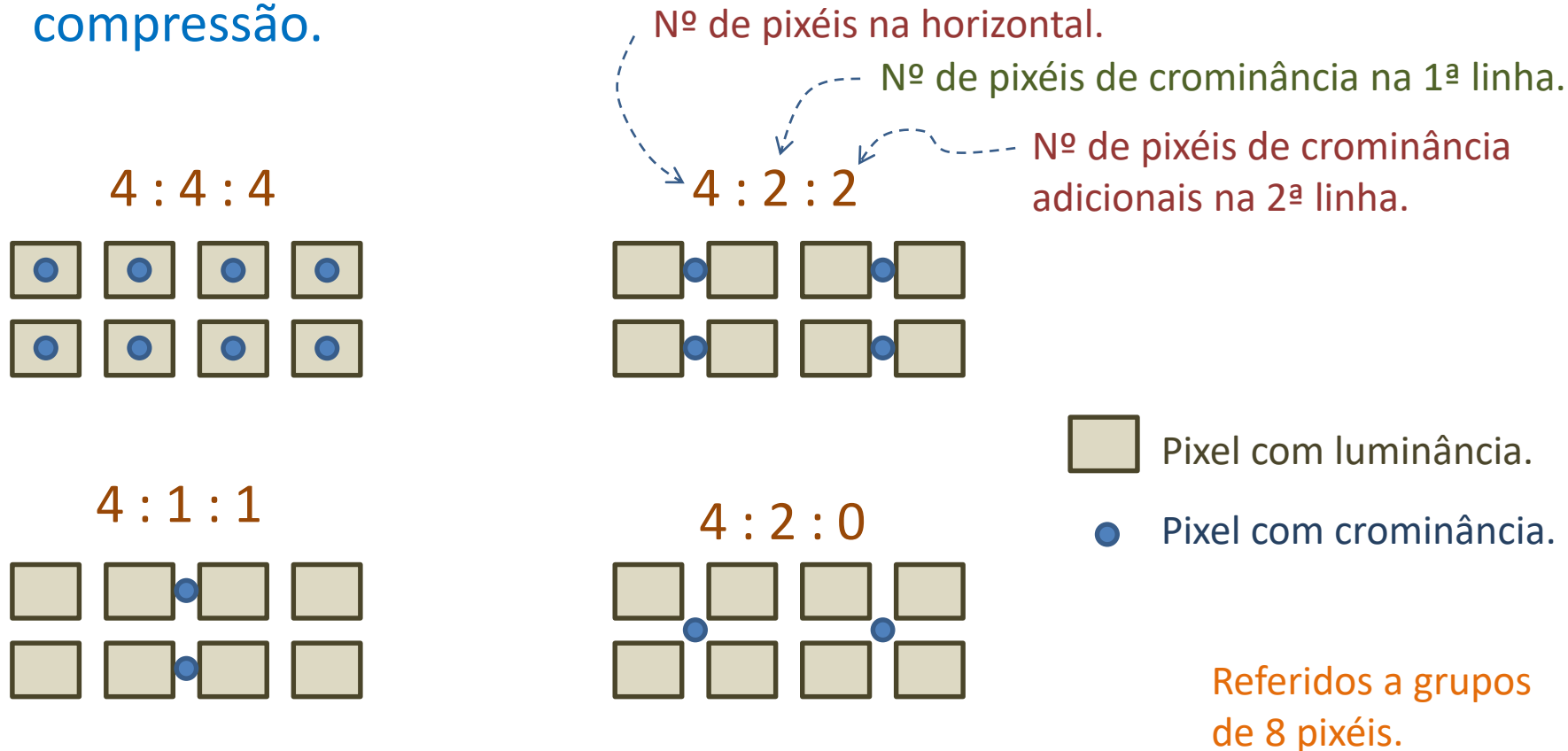


C_r



3. Compressão de Imagem

- Aproveitando a menor variabilidade da crominância, e o facto de o olho humano não ser muito sensível a variações de cor, a informação da crominância é muitas vezes subamostrada na compressão.



3. Compressão de Imagem

Alguns Formatos Standard:

- *JPEG – Joint Photographic Experts Group*
Efetua compressão com perda de informação.
Obtém imagens com elevada taxa de compressão.
- *JPEG2000*
Pode efetuar compressão com ou sem perda de informação.
- *TIFF – Tagged Image File Format*
Pode efetuar compressão com ou sem perda de informação.

4. Compressão de Vídeo

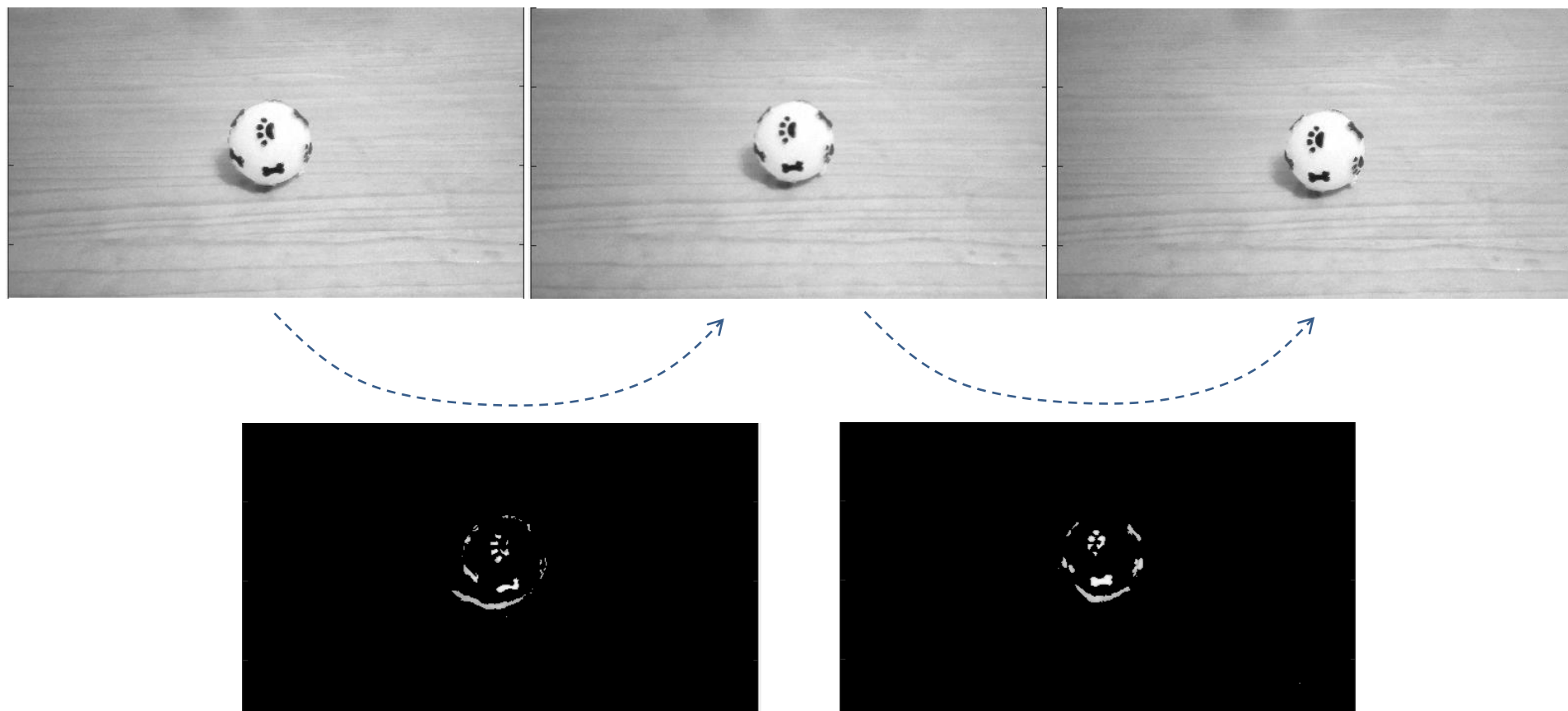
- Se uma imagem envolve um volume de dados considerável, o vídeo (com 25/30 *frames* por segundo) requer certamente procedimentos dedicados que conduzam a um elevado nível de compressão.
- Por exemplo, considere-se o caso de vídeo sem compressão, com:
 - 720x576 pixéis por imagem, RGB (formato PAL)
 - Cada *frame* requer, então, 1244160 bytes (1.24 MB).
 - Cada segundo de vídeo (25 *frames* em PAL) requer 31.1MB.
 - Uma hora de vídeo seriam, assim, quase 112 GB de dados!
- Comprimir vídeo é, portanto, obrigatório.

4. Compressão de Vídeo

- Naturalmente, as técnicas de compressão de imagem (como a redução da **resolução espacial** da cromaticidade) são usadas em compressão de vídeo.
- Mas o vídeo pode ainda tirar proveito da **Redundância Temporal**: duas *frames* consecutivas apresentam muita informação comum (pelo que não é necessário repeti-la em cada *frame*).
- Para tal, vários algoritmos são usados, tais como:
 - **Codificadores preditivos** (modelam o comportamento no tempo de componentes das imagens)
 - **Detetores de movimento** (identificam as diferenças entre imagens consecutivas e modelam a evolução das diferenças).

4. Compressão de Vídeo

Exemplo (bola a girar):



4. Compressão de Vídeo

- É usual um vídeo ser composto por diferentes tipos de *frame*, por exemplo:
- **Frame de referência** – imagem em bruto, sem ser o resultado do processamento da informação de *frames* anteriores.
- **Frame preditiva** – resultado de um modelo preditivo (usualmente complexo) que opera com base na *frame* de referência anterior, e nas diferenças entre *frames*.
- Operando sobre a evolução das diferenças, nomeadamente nas componentes de luminância e crominância separadamente, conseguem-se taxas de compressão muito significativas.
- Os modelos preditivos são muitas vezes suportados por algoritmos de identificação de entidades nas imagens diferença.

4. Compressão de Vídeo

Alguns formatos standard:

- **Motion JPEG**

Cada *frame* é comprimida em JPEG.

Não considera a redundância temporal.

É, por exemplo, usada em algumas máquinas fotográficas.

- **MPEG-2**

Considera desde 352x288@25fps até 1920x1250@60fps (HD).

Usada em transmissão de TV e HDTV.

- **H.216, H.263 e H.263+**

Vídeo-conferência sobre rede telefónica fixa (ISDN) ou móvel.

- **H.264 (MPEG-4 AVC)**

Atualmente muito usado em transmissão de vídeo.

4. Compressão de Vídeo

- Algumas aplicações concretas tiram partido de determinadas especificidades para reduzir ainda mais o volume de dados.
- Muitas dessas técnicas consideram estudos sobre a forma como o cérebro humano discrimina a informação visual que recebe.
- Por exemplo, num jogo de futebol, o cérebro está mais atento à posição e movimentação da bola e dos jogadores, não prestando tanta atenção aos pormenores da relva do campo.
- Tal pode ser aproveitado para reduzir muito significativamente a resolução espacial das zonas do relvado vazias, sem que o espetador se aperceba de tal.