

Sample Questions - II

CS4551 Multimedia Software Systems

- How does transform coding contribute to the compression?
- What does JPEG stand for?
- JPEG – JPEG compression, the image components are each broken into 8x8 blocks and each block is transformed to the frequency domain giving a DC and many AC coefficients.
 - Given a RGB image, JPEG performs two transformations for compression. Name these transformations and explain how each contribute for the compression.
 - What are AC and DC Components and what do they represent?
 - Why are the DC coefficients encoded separately from the AC coefficients? Can you think of an application scenario where this separation can prove to be helpful?
 - What is reason for zigzag coding of AC Coefficients?
 - Why does JPEG consider 8x8 blocks instead of whole image?
 - The entry a_{ij} of a JPEG quantization table specifies the number of quantization levels for the $(i,j)^{th}$ DCT coefficient. Explains why a_{ij} typically decreases with increasing i or j . How does this fact influence the compression?
 - Let's assume that DCT coefficients in a particular 8x8 image block is as shown below:

48	12	0	0	0	0	0	0
-10	8	0	0	0	0	0	0
2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

SIZE	AMPLITUDE
1	-1,1
2	-3,-2,2,3
3	-7,-4,4,7
4	-15,-8,8,15
5	-31,-16,16,31
6	-63,-32,32,63
7	-127,-64,64,127
8	-255,-128,128,255
9	-511,-256,256,511
10	-1023,-512,512,1023

For entropy coding, we need to convert DC coefficient and AC coefficients to the intermediate representations i.e. (SIZE)(AMPLITUDE) for DC, (RUNLENGTH,SIZE)(AMPLITUDE) for ACs. Represent the given 8x8 block DCT coefficients in the intermediate format. Use the table (on the right side) for computing SIZE.

Ans: DC – (6)(48), ACs – (0,4)(12), (0,4)(-10), (0,2)(2), (0,4)(8), (0,0)

- JPEG supports three modes for progressive encoding: spectral selection, SNR scalability, and pyramid mode. Which one is best indicated if we want to use the same compressed file structure to accommodate different sizes of display? Why?
- Suppose that we have the following 4x4 image block.

10	11	12	13
11	12	13	14
12	13	14	14
13	14	14	15

Encode this block using the lossless mode of JPEG. For prediction, use A or B or C or A+B-C.

- What are the advantages of using a different color space than RGB (e.g., YCrCb) for the sampling and compression of color images and video?
- Consider a video format with 325 lines/frame, 490 pixels/line, 30 frames/s, color subsampling scheme 4:2:2, image aspect ratio: 4:3. Compute the bit-rate of the system (assuming each luminance and chrominance sample is quantized with 8 bits).

Ans: 76.44 Mbits

- Using the following video format table, compute bit-rate for NTSC and two HDTV videos.

Format name	Lines/ frame	Pixels/ line	Frames/s	Interlaced	SS scheme	Aspect ratio
CIF	288	352		N	4:2:0	4:3
QCIF	144	176		N	4:2:0	4:3
SQCIF	128	96		N	4:2:0	4:3
SIF-525	240	352	30	N	4:2:0	4:3
SIF-625	288	352	25	N	4:2:0	4:3
NTSC	486	720	29.97	Y	4:2:2	4:3
PAL	576	720	25	Y	4:2:2	4:3
HDTV	720	1280	59.94	N	4:2:0	16:9
HDTV	1080	1920	29.97	Y	4:2:0	16:9

- Suppose a camera has 450 lines per frame, 520 pixels per line, and 25 Hz frame rate. The color-subsampling scheme is 4:2:0. The camera uses interlaced scanning, and each sample of Y, Cr, Cb is quantized with 8 bits.
 - What is the bit-rate produced by the camera?
 - Suppose we want to store the video signal on a hard disk, and, in order to save space, re-quantize each chrominance (Cr, Cb) signals with only 6 bits per sample. What is the minimum size of the hard disk required to store 10 minutes of video?
 - Repeat the exercise (both questions) assuming color subsampling scheme 4:2:2.
- What is gamma correction?
- What is motion compensation and why is it used in video compression?
- In motion compensation video coding, one motion vector per macroblock is transmitted, together with the encoded difference between the color values in the macroblock and the motion compensated reference frame. What are the criteria to bear in mind when choosing the size of the macroblocks? Why is it not a good idea to encode the motion vectors with a lossy mechanism?
- Consider motion compensation video coding: the displacement vector (dx,dy) is transmitted, together with the encoded residual. What happens if, instead of a closed-loop motion-based encoder and decoder we use an open-loop system (i.e., if the residual is computed with respect to the “original” previous frame instead than with respect the reconstructed frame)?
- Motion compensation is a technique that usually allows us to save bits when encoding video. However, motion compensation should not be used when there is a scene change (i.e., between the last frame of a scene and the first frame of a new scene). Explain why.
- Name at least three block matching criteria and explain how the best matching block is selected.
- What is the advantage of Integral Projection method used of the block matching?
- Computing a motion vector for each macroblock is time consuming. Write at least two methods which can reduce search time for finding the best matching block.
- Which movie coding standard is most indicated for teleconferences over POTS with modems up to 33.6 Kb/s? And which standard is indicated for broadcasting digital video and HDTV?
- What is the compression standard for video in digital television (DTV)? What is an appropriate video compression standard for videotelephony over ISDN lines? What are the main differences between HDTV and NTSC video formats? What is the bit-rate generated by a MPEG-1 video coder?
- What are possible bit-rates of video compressed according to the ITU H.261? Does H.263 yield higher or lower bit-rate than H.261? What was the intended application for MPEG-1? Which audio and video compression systems are used for Digital TV and HDTV?
- Suppose we are using H.263 to encode video, and we monitor the bit-rate at the output of the encoder. Why do we usually see a peak (burst) in the bit-rate in correspondence of a change of scene? What happens if the first frame of the next scene is encoded in predictive mode instead than in intraframe mode?

- What are possible bit-rates of video compressed according to the ITU H.261? Does H.263 have higher compression ratio with respect to H.261?
- Write at least three improvements that H.263 has made compared to H.261.
- Explain why, in motion-compensated compressed MPEG movies, I-frames typically require more bits than P- or B-frames?
- Write at least three differences (improvements) that MPEG-1 has compared to H.261.
- Write at least three differences (improvements) that MPEG-2 has compared to MPEG-1.
- What is the bit-rate of MPEG-2 movie?
- Suppose we encode a sequence of 23 frames using MPEG. Each frame is encoded in I (intra frame), P (forward prediction) or B (bi-directional prediction) mode according to the following order: I P P P B B P I P P P I P P B B B B P I P I P
 - a) Derive the correct transmission and decoding order.
 - b) What is the minimum delay (in terms of frame period) due to bi-directional encoding in this case?
- Write five scalabilities that MPEG-2 provides. Pick two and explain them.
- What are the advantages of using scalable MPEG-2 encoding when users have connections to the network at different bit-rates and have different processing capabilities? What are the main scalable modes of MPEG-2?
- Explain why with bidirectional motion compensation of videos the camera and display order of frames is different from the transmission and decoding order. How many displacement vectors do we need to transmit for each 16x16 macroblock of a B-frame in MPEG-2?
- Consider a hypothetical video standard with progressive (non-interlaced) scanning, $N=486$ lines per frame and frame rate $F=60$ Hz. We would like to broadcast video with such a standard, but we expect that most of the receiver will be able to decode and display NTSC video only ($N=486$ lines, $F=30$ Hz, interlaced scanning). We decide to use a temporal scalable encoder, so that all users will be able to decode at least the base layer. Describe how the base layer and the enhancement layer are encoded.
- What is DPCM? What does the encoder side quantize and transmit to the decoder side?
- Supposed that you have the following sampled values: 10.3, 10.6, 10.9, 11.2, 11.7, 11.3, 10.9. Assume that you use the previous value for the prediction and quantize values using round function. Show the encoded values by a closed loop DPCM. Also, show the decoded values in the decoder side.
- Suppose we are sampling and quantizing a signal using DPCM lossy coding. Assume that the variance of the difference $d(n)$ is 100 times less than the variance of the signal. What is the quantization SNR if we quantize $d(n)$ with 16 bits per sample? How many bits per sample are needed to ensure that the quantization SNR stays above 70 dB? (Recall how to compute SNR and its measurement unit. Also recall the relationship between SNR and the quantization bit.)
- Assume that you quantize 16-bit samples into k -bits ($k < 16$) using DPCM. Describe one inherent problem of DPCM and propose one method that can improve DPCM.
- What is DM? When is this coding method useful?
- What is ADPCM? How does it improve PCM, DM, or DPCM?
- What is the ITU audio compression standard that mainly uses PCM? Name one ITU standard that heavily relies on ADPCM.
- What are μ -law and A-law? Why do we need them?

- Masking as defined by the American Standards Association (ASA) is the amount (or the process) by which the threshold of audibility for one sound is raised by the presence of another (masking) sound. There are three types of masking. What are frequency masking and temporal masking?
- What is the main idea of MPEG-1 audio compression method?
- Write at least two significant differences between MPEG-1 Layer-1 and Layer-2.
- Write at least two significant differences between MPEG-1 Layer-2 and Layer-3.
- Read JPEG-2000 slides for some of EC questions.

SAMPLE