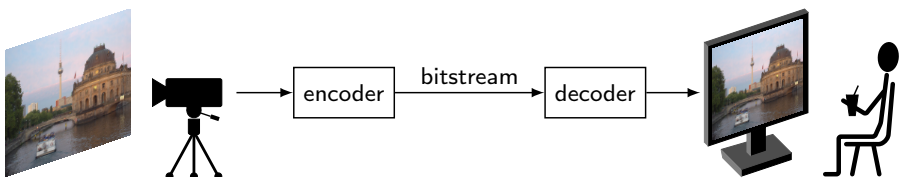
Image and Video Coding: Exam Preparation



What Type of Exam? - When and Where?

General

- Online exam via WebEx (or any other system)
- Possible dates: Starting at August, 10

➡ Write e-mail to heiko.schwarz@hhi.fraunhofer.de with 2–3 suggestions for an exam date

Choose Your Type of Exam

1 Conventional Oral Exam

- Oral exam of 30 minutes
- Explain fundamental concepts of image and video coding
- Answer questions to topics discussed in lectures

2 Alternative: Implementation Project

- Implement an improvement for our image codec or our video codec
- Present your implementation and simulation results
- Answer questions about implemented coding tool, actual implementation, and closely related topics

Example Questions: Human Visual Perception and Raw Data Formats

- How are images and videos represented in raw data formats?
- Calculate the raw data rate (in bits per second) of a full HD (1920×1080) video with 50 pictures per second and a bit depth of 8 bits per sample?
- Why do we need at least 3 color components for representing color images?
- What is the XYZ color format and how can it be converted into an RGB format?
- What are linear color spaces?
- Why do we always need a color conversion between image capture and display?
- What does the chromaticity diagram show?
- There are multiple RGB formats. How do they differ?
- What is the "color gamut" of a display or a representation format?
- Is it possible to represent all colors (perceivable by human beings) with an RGB display?
- What is "gamma encoding" and why is it used?
- Why are images and videos typically coded in the Y'CbCr format (instead of an RGB format)?
- What is the YCbCr 4:2:0 chroma sampling format and why is it used?

Example Questions: Transform Coding of Sample Blocks

- What are the components of a transform encoder and decoder?
- Why is the quantization done in the transform domain instead of the original sample space?
- What is the main goal of the transform?
- What are orthogonal transforms and why are non-orthogonal transforms not used in lossy coding?
- What is the Karhunen Loève transform (KLT) and why is it often considered as optimal transform?
- Why do we mainly use separable transforms in image and video coding?
- Why is the DCT-II used in most image and video codecs?
- Why do modern codecs also support alternative transforms as, for example, the DST-VII?
- Could the usage of non-separable transforms improve the coding efficiency?
- What signal characteristics determine the effectiveness of transform coding?
- What is the purpose of quantization?
- What is a uniform reconstruction quantizer (URQ) and why is it used in most lossy codecs?
- What is the quantization step size (for URQs) and how does it impact the bit rate and reconstruction quality?

Example Questions: Entropy Coding

- What is the goal of entropy coding?
- What influences the efficiency of entropy coding?
- Explain the Huffman algorithm for a simple given probability mass function.
- What are advantages of conditional entropy coding?
- What are advantages and disadvantages of block entropy coding?
- What is entropy, conditional entropy, block entropy, entropy rate?
- Why are neither scalar Huffman coding, conditional Huffman coding, nor block Huffman coding used for coding the quantization indexes in image and video codecs?
- Explain the concept of run-level coding (as used in JPEG or MPEG-2 Video) and its advantages.
- Why is the zig-zag scan (or a similar diagonal scan) used for transform coefficient levels?
- Explain the basic idea of arithmetic coding? What are the advantages?
- Why do most practical codecs use binary arithmetic coding?
- Explain how arithmetic coding can be efficiently combined with probability estimation.
- What design aspects impact the efficiency of arithmetic coding for quantization indexes?

Example Questions: Advanced Intra-Picture Coding

- Why is it important to utilize statistical dependencies between transform blocks?
- Explain the effect of the prediction of the DC quantization index used in JPEG.
- What are the advantages of intra prediction in the sample domain?
- Explain the idea of directional intra prediction modes.
- Why do modern video coding standards support a large variety of intra prediction modes?
- How can the intra prediction mode be selected in an encoder?
- How does the block size impact the effectiveness of intra prediction and transform coding?
- Why do modern codecs support the concept of variable block sizes?
- Why are only simple partitioning concepts supported in modern codecs?
- Explain the idea of quadtree partitioning.
- How can an encoder determine a suitable quadtree partitioning for a given block.
- What are the main design aspects in modern video coding standards that lead to a substantial coding efficiency improvement relative to JPEG for images?

Example Questions: Encoder Control

- What parameters have to chosen in a video encoder?
- Explain the difficulty in selecting coding parameters in a video encoder (based on examples).
- Explain why encoders that are based on distortion minimization are not very efficient.
- Explain the basic concept of Lagrangian optimization.
- Explain the basic approach of a Lagrangian encoder control.
- How can the Lagrangian concept be used for optimizing the selection of
  - ➡ coding modes or intra prediction modes
  - ➡ motion vectors and reference indexes
  - ➡ a block partitioning
  - ➡ quantization indexes
- Explain why using a "rounding offset" in encoder quantization represents a suitable low-complexity alternative to rate-distortion optimized quantization.
- Explain one idea for implementing a "fast" motion estimation algorithm.
- Explain why the encoding algorithm used is crucial for achieving a good coding efficiency with a given standard? Why is it somewhat difficult to fairly compare different video coding standards?

Example Questions: Motion-Compensated Prediction / Hybrid Video Coding

- What is the main difference between image and video coding?
- Explain the main idea of motion-compensated prediction. What is motion estimation?
- Draw a block diagram of a hybrid video encoder and explain the main components.
- Why do modern video coding standards use sub-sample accurate motion vectors?
- Why do we need interpolation for sub-sample accurate motion vectors?
- Why can longer interpolation filters provide better coding efficiency than bi-linear interpolation?
- Why are motion vectors coded predictively?
- Explain one example for motion vector prediction.
- Explain one example for a coding mode, in which the motion data are not explicitly coded, but derived based on already coded data.
- Is it possible to use non-translational motion (e.g., rotations) in motion-compensated prediction? Are there advantages? What are the disadvantages?
- Why do all video coding standards support intra coding modes in predictively coded pictures?
- Why are video encoders much more complex than video decoders?

Example Questions: Advanced Concepts in Video Coding

- Explain the concept of "multiple reference pictures".
- How can we select the reference picture for a block in a video encoder?
- Explain the concept of bi-prediction and why it can improve coding efficiency.
- Why provide video consing standards the possibility to select between conventional uni-prediction and bi-prediction on a block by block basis?
- What are B pictures and why do they typically improve coding efficiency?
- Why is it typically advantageous to code video pictures in an order that differs from the acquisition and display order? Explain an example for such an improved coding structure.
- What are "random access points" and why do we need them in broadcast and streaming?
- Explain the difference between random access points with "closed GOP" and "open GOP".
- Why are "advanced coding structures" (i.e., coding orders that differ from the display order) not used in video conferencing? Can we still use B pictures?
- What is the goal and the basic concept of a deblocking filter?
- Why are deblocking filters applied inside the motion compensation loop, and not as post filters?

On Implementation Projects (Alternative Exam)

Possible Implementation Projects

- Any potential improvement for our image codec or our video codec
- You can basically select a topic of your choice (should require a certain amount of work)
  - Some example topics are listed on the following slides
  - Topic has to be agreed with lecturer in advance (during last tutorial or by e-mail)

Alternative Exam

- Provide your code in a git branch (at least two days before exam)
- Provide simulation results for at least one complete video sequence (or four pictures)
  - Use our usual set of QP values (8, 12, 16, 20, 24)
  - You can select the video sequence (or the 4 pictures) for which your addition works best
  - Plot PSNR-rate curves for the base version (master branch) and your version into the same diagram
  - ! *Grade does not depend on coding efficiency gain (but correctness of implementation is important)*
  - ! *Could also be interesting to compare different variants of your implementation*
- Be prepared to explain your implementation and answer questions on related topics

Possible Improvements for our Image Codec I

1 Simple Version with Variable Block Sizes

- Two partitioning options for each $B \times B$ block (indicated by flag in the bitstream):
  - ➡ code $B \times B$ block as $B \times B$ block (as currently implemented)
  - ➡ code $B \times B$ block as four $(B/2) \times (B/2)$ blocks
- Use Lagrangian decision for selecting between these two partitioning options
- Compare new version ($B = 16$) with both fixed block sizes (16×16 and 8×8)

2 Coding of Color Images

- Extend our codec for PGM images (gray) to a codec for PPM images (given in RGB)
- Compare two different versions:
  - ➡ Independent coding of R, G, and B color components
  - ➡ Color transform to YCbCr (or YCoCg), coding of these components (with or without chroma subsampling), reconstruction of R, G, and B after decoding and output as PPM image
- Compare both variants with JPEG (data available in git)

Possible Improvements for our Image Codec II

3 Improved Entropy Coding of Quantization Indexes

- Implement and test ideas for improving the entropy coding of quantization indexes of a block
- Suggestion: Use some ideas from modern video coding standards, for example,
  - ➡ code last position using x and y coordinates with different probability models for x and y
  - ➡ optimized number and assignment of diagonal classes based on coding experiments
  - ➡ select probability models based in sum of absolute values in local neighbourhood (in addition to diagonal classes, see VVC)
- You can disable the mode dependent scanning (i.e., use diagonal scan for all modes)
- Test multiple ideas step by step and also present intermediate results

4 Improved Intra-Picture Prediction

- Add multiple directional intra prediction modes (see AVC, HEVC, ...)
- Modify entropy coding of intra prediction modes accordingly
- Decide mode using Lagrangian mode decision
- Suggestion: Test with rather small block sizes (e.g., 8×8 blocks)

Possible Improvements for our Video Codec I

1 Motion-Compensated Prediction with Quarter-Sample Accurate Motion Vectors

- Increase motion vector resolution to one quarter-sample
- Test different interpolation filters (e.g., bi-linear, AVC, and HEVC filters)
- Suggestion: Use "BQSquare" as test sequence

2 Motion-Compensated Prediction with Multiple Reference Pictures

- Store reconstructed reference pictures in a sliding window buffer
- Coding of reference index on block basis with block-adaptive selection of reference picture
- Run tests with 1, 2, and 4 reference pictures

3 Simple Variant of Motion-Compensated Prediction with Variable Block Sizes

- Add a block coding mode, in which the block is partitioned into 4 square blocks for motion compensation, motion estimation, and motion vector coding
- Use same block size for transform coding as for motion compensation (need second transform size)
- Include the new mode in the Lagrangian mode decision

Possible Improvements for our Video Codec II

4 Block-Adaptive Bi-Prediction

- Add another coding mode (for inter pictures) that uses bi-prediction
  - ➡ At the encoder: Find second motion vector using a conditional motion search (see lecture)
  - ➡ Add suitable concept for entropy coding of second motion vector
  - ➡ Modify entropy coding of mode information in a suitable way
- Include new coding mode in Lagrangian mode decision
- Test with the bi-linear interpolation filter and the HEVC filter

5 Implement MPEG-2 style "IBBP" coding (suitable for joint work of 2 students)

- Modify the coding order of pictures according to the MPEG-2 "IBBP" concept
- Introduce B pictures with 4 coding modes (selected on block basis)
  - ➡ Forward prediction using preceding I/P picture
  - ➡ Backward prediction using succeeding I/P picture
  - ➡ Bi-prediction using preceding and succeeding I/P picture (have to code 2 motion vectors)
  - ➡ DC prediction (i.e., intra coding)
- Coding mode for a block should be selected via Lagrangian mode decision
- Suggestion: Increase QP for B pictures (e.g., $\mathrm{QP}(B) = \mathrm{QP}(P) + 2$)

Possible Improvements for our Video Codec III

6 Investigate Different Cost Measures for Motion Estimation

- Investigate the following distortion measures for motion estimation:
  - ➡ Sum of absolute differences (SAD) with $\lambda_m = \sqrt{\lambda}$ (current version)
  - ➡ Sum of square differences (SSD) with $\lambda_m = \lambda$
  - ➡ Sum of absolute differences in transform domain (SATD) with $\lambda_m = \sqrt{\lambda}$ (use DCT-II as transform)
- Investigate coding efficiency and encoder complexity
- Is it useful to use different measures for the integer motion search and the sub-sample refinement ?

7 Extent Our Video Codec to the YCbCr 4:2:0 format

- Color versions of the test sequences (YCbCr 4:2:0) can be found on web page and in Whiteboard
- Raw data for a picture: $W \times H$ luma samples, $(W/2) \times (H/2)$ Cb samples, $(W/2) \times (H/2)$ Cr samples
- Suggestions:
  - ➡ Conduct motion estimation with luma component only
  - ➡ Re-use motion vectors for chroma (note: chroma has only half the resolution in both directions)
  - ➡ Choose joint coding mode for all color components of a block (use sum of distortions for mode decision)
- Compare coding efficiency with gray-only version (and MPEG-2 in YCbCr 4:2:0 format)