FIA - WS 2014/15

1 Introduction

1.1 Internet

- loosely hierarchical: tier 1 to tier 3
- communications infrastructure enables distributed applications
- hyper... ⇒ concentration of content (Amazon, Google, ...)

1.2 Protocol Layering

- ISO/OSI
- TCP, UDP, ICMP, UDP Like, SCTP, ...
- DHCP, NAT, ...

1.3 Analogue digital conversion

- Fourier transformation
- sampling theorem + Nyquist
- PCM-transmission
- small amplitudes are being encoded more detailed than large amplitudes

1.4 Color Coding

- monochrome, RGB, YCbCr
- RGB:
  - red, green, blue values between [0, {7,31,...,65535}]
  - nonlinear encoding of intensities
  - computer graphics
- YCbCr
  - TV and digital video
  - Y more important ⇒ encode with more details
  - more efficient than RGB
  - can handle downsampling better
  - YUV is based color model used in analog color TV
    * YCbCr is scaled and offset version
    * YPbPr is the analog version of YCbCr

2 Digital coding of audio and video

2.1 Rate-Distortion Theory

- lossless compression algorithms
  - allow perfect reconstruction
  - low compression ratios
  - frequently encountered data is encoded more efficiently
- lossy compression algorithms
  - result is only a close approximation of original data
  - trade-off: distortion vs. required rate
  - much higher compression rate than lossless compression

  //TODO: bild einfuegen
- rate and distortion as measures for efficiency of compression and difference between reconstructed and original data
  - goal is to minimize distortion and rate
  - basic problem:
    * minimum expected distortion at a given rate?
    * minimum rate to achieve given distortion?
- distortion measures
  - mean square error $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$
  - signal to noise ratio $SNR = 10log_{10}\frac{\sigma_x^2}{\sigma_d^2}$
  - peak signal to noise ratio $PSNR = 10log_{10}\frac{x_{peak}^2}{x_d^2}$
- in order to maximize efficient communication maximize mutual information between x and y
- rate distortion function

2.2 digital image

- conversion between RGB and YUV
- downsampling J:a:b
  - 4:4:4 $\widehat{=}$ no downsampling
  - 4:2:2 $\widehat{=}$
  - 4:2:0 $\widehat{=}$
- statistical image modeling
  - all natural images occupy a tiny and unknown sloped space of all images
  - pixel intensities are dependent and correlated to the corresponding image $\Rightarrow$ pixel within images are highly correlated
  - correlation of pixels has high impact on image compression
- image transformations
  - negative transformations

  - log transformations
  - power-law transformations

- intensity:
  - change in brightness ⇒ shift of histogram
  - change in contrast ⇒ stretch/? of histogram

- filters based on convolution of neighbor pixels ⇒ enhancement, smoothing, edge, detection, ...

- the hoar transform
  - simplest useful energy compression
  - lossless retransformation is possible
  - transform original image into 4 parts (in 20)
    1. top-left: a+b+c+d, 4 point average, very important
    2. top-right: a-b+c-d, average horizontal gradient, less important
    3. bottom-left: a+b-c-d, average vertical gradient, less important
    4. bottom-right: a-b-c+d, diagonal curvature, less important
  - 1 is more expensive while 2-4 is cheaper to encode
  - reordering required to provide "typical" representation

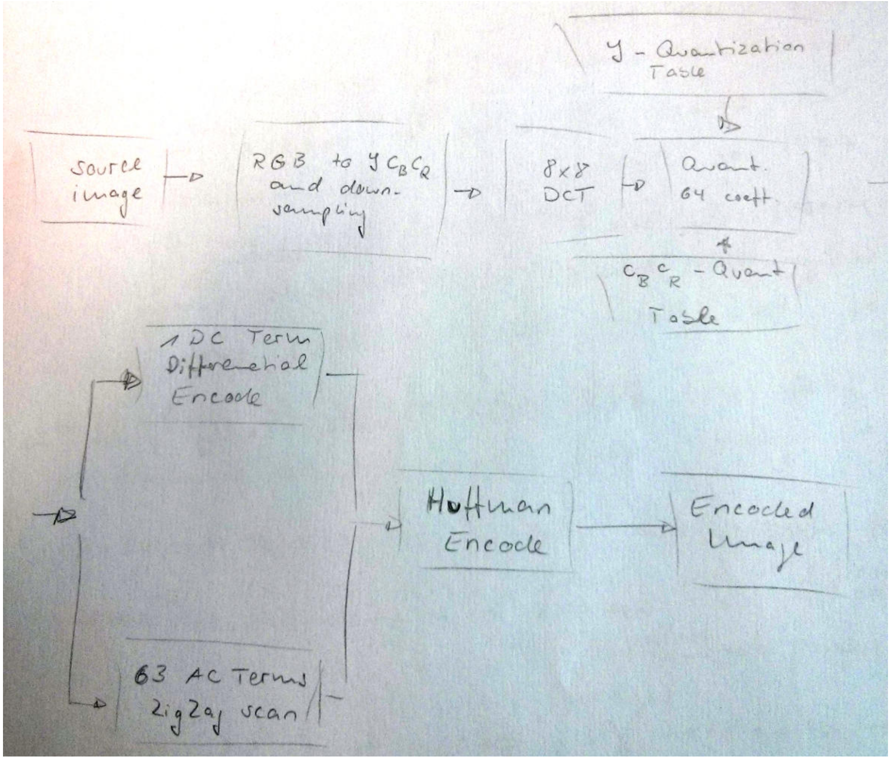
- entropy – quantifying the required bitrate
  - entropy H_x represents the mean number of bits per pixel that are required to encode the image
  - $H_x = \sum_{i=0}^{M-1} p_i - log_2\left(\frac{1}{p_i}\right)$ where $p_i = \frac{\text{pixel in bin i}}{N}$ where N is the number of pixels
  - estimated number of bits needed is $H_x \cdot N$

- the Karkunen-Loeve transform (KLT)
  - is optimal for minimizing bitrate
  - not or seldom used in practice ⇒ slow and complex

- the discrete cosine transformation (DCT)
  - seems to put most energy on a small number of values
  - linear transformation
  - each real world image can be represented by a combination of the DCT bases
  - DCT bases fit better to typical images than FFT or DST
  - first coefficient of DCT is mean of values being transformed and represents average tone of the block. subsequent blocks add further detail
  - JPEG: 8×8 DCT (empirically found to be good)

- need for wavelets
  - signals are available in time-domain but processing frequency information is much more easy
  - transformations (e.g. FFT) translate time-domain signals into frequency domains
    * useful for stationary signals where all frequencies are present at all times
    * not useful for instationary signals: both information required
    * solution: short term transformation—transformation of small fixed timewindows (similar to blocks of images)
  - wavelet analysis uses different sized windows
    * high frequency parts ⇒ small windows ⇒ good time resolution
    * low frequency parts ⇒ big windows ⇒ good frequency resolution

2.3 JPEG compression

- joint photographic experts groups
- 

2.4 digital video

- high corellation between successive frames
- interframe differential pulse code modulation (⇒ vgl. skript skizze)
- frame replenishment ⇒ pixels are only transmitted if a specified threshold is exceeded, otherwise nothing is transmitted
- change detection
  - pixel wise change detector
  - block wise change detector
  - comparision between current and previous frame
    * substract previous frame from current one
    * convert to absolute value
    * generate 3×3 averages
    * check for threshold
- motion compensated coding