Techniques for Video Compression and Analysis (5LSE0), Module 01 - A

Introduction to Coding and Information Measures

Peter H.N. de With
(p.h.n.de.with@tue.nl)

slides version 1.0

Lecture Information

- Lecturers

Prof.dr.ir. Peter H.N. de With and Joy (Yue) Sun MSc.
Faculty Electrical Engineering, University Technology Eindhoven
SPS – Video Coding & Architectures (VCA)
Address: Flux 5.092, Tel: 040-247.25.40,
Email: P.H.N.de.With@tue.nl & Y.Sun1@tue.nl

- Information: **http://vca.ele.tue.nl/courses/5LSE0 (version 2022-23)**
  - Manuscript
  - Lecture slides
  - Software Practical Exercises
- **Examination: Oral Examn (70%) + Practical exercises (30%)**

Purpose of the course 5LSE0

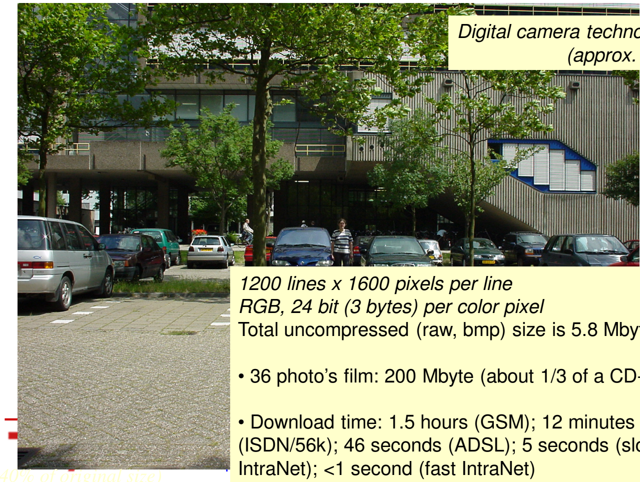
- **Learn the power of signal transformations**
  - For video compression, like as in transform coding
  - For video analysis, like as in the Gabor wavelet transform
  - For combining them with Deep Learning
- **Exploit the concept of “condensing information”**
  - This is useful for coding and compression
  - But also applicable to video analysis: what is the key informative feature
  - Use specific transformation for that, like the PCA,
- **Therefore, this course is divided in two parts**
  - A. Video compression using Discrete Cosine Transformation
  - Learn the concept of entropy as information measure both for compression and video analysis
  - B. Analysis techniques, e.g. wavelet transform, for finding feature
  - Expand the tool set of computer vision and for Deep Learning

Mod 01 - A / Part 1: Introduction / Motivation for Compression

What is Digital Signal Coding?

* Digital signals**: audio, image(s), video (multimedia)**
  – Will not discuss *document* compression
  – Focus on image compression (video and images are signals)
* Coding/Compression **= compact representation:**
  – Less space required on storage media (hard disks, CD, DVD, BD, embedded memory)
  – Smaller transmission bandwidth required (Internet, terrestrial broadcast, satellite links, mobile phones)
  – Allows for faster uploading/downloading of multimedia files (Internet, authoring environment)

Why Digital Signal Coding? – (1)



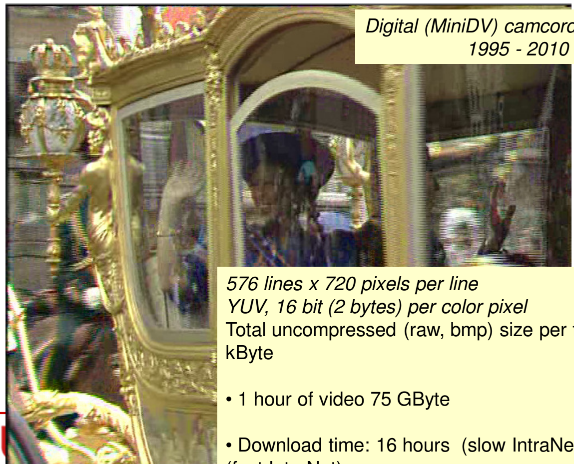
Digital camera technology 2002 (approx. 400 Euro)

1200 lines x 1600 pixels per line
RGB, 24 bit (3 bytes) per color pixel
Total uncompressed (raw, bmp) size is 5.8 Mbyte

• 36 photo's film: 200 Mbyte (about 1/3 of a CD-ROM)

• Download time: 1.5 hours (GSM); 12 minutes (ISDN/56k); 46 seconds (ADSL); 5 seconds (slow IntraNet); <1 second (fast IntraNet)

Why Digital Signal Coding? – (2)



Digital (MiniDV) camcorder technology 1995 - 2010 (1000 Euro)

576 lines x 720 pixels per line
YUV, 16 bit (2 bytes) per color pixel
Total uncompressed (raw, bmp) size per frame is 830 kByte

• 1 hour of video 75 GByte

• Download time: 16 hours (slow IntraNet); 2 hours (fast IntraNet)

Usage of Digital Signal Coding

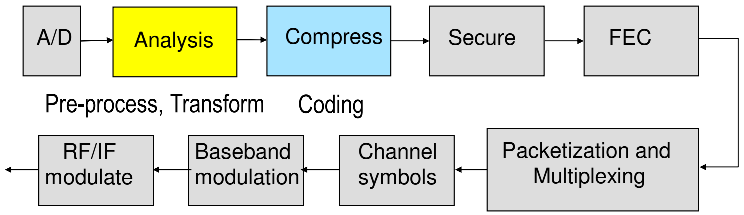
* Coding/Compression **=> compact representation:**
  – Less space required on storage media
  – Smaller transmission bandwidth required
  – Allows for faster uploading/downloading of multimedia files / communication
* **Examples of "products"/systems using compression:**
  – Digital audio: MP3
  – Digital image and video cameras: JPEG, MPEG, MPEG-4 AVC
  – Internet movie downloading/streaming DivX, Qtime, Real players
  – Image compression for mobile phones: Whatsapp
  – Digital Cinema: JPEG-2000

Objectives of this course

* **Objectives of this course:**
  - Fundamental video/image compression, basics Information Theory
  - Key Techniques by which compression can be achieved
  - What is inside the MPEG standard, DVB, DV, JPEG, etc.
  - Introduce new signal transformations: DCT (compression, Wavelet transform (compression and analysis), PCA (analysis)
  - Introduce metrics that benefit video analysis
  - Discuss video analysis cases using signal transforms
  - Show video analysis where models and metrics are used
  - Practical experiments with compression/quality and transform video analysis, software package

Compression is now an embedded system

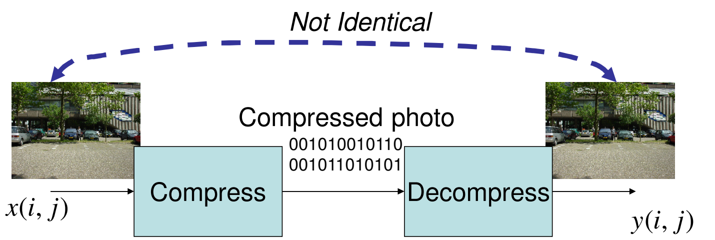
Signal to be coded ***x(t), x(t,s)***



Pre-process, Transform   Coding

* "Coding" includes Compression, Security, and Error Control techniques
* This course focuses on Compression and Transform Analysis only

Lossless vs. Lossy Compression



Lossy or Non-reversible Compression (if identical: reversible)

$y(i,j) \neq x(i,j)$

Examples: JPEG, MPEG, MP3, DivX
Can trade compression for amount of difference (quality)

Some Ballpark Figures

* **JPEG Images**
  » Digital camera: 1600x1200 pixels at 24 bits/color pixel ==> 46 Mbit/file = 5.8 MByte/file
  » "Good quality" JPEG: 600 kByte
    » Bit rate: 2.5 bit / color pixel (bpp), Compr. Factor: 9.6

* **MPEG Video**
  » Digital Broadcast TV: 720x576 at 16 bit/color pixel at 25 pictures/second = 166 Mbit/sec = 20 MByte/sec
  » "Commercial tv station" MPEG: 4.5 MBit/sec
    » Bit rate: 0.43 bit / pixel (bpp), Compr factor: 36.8
  » Full HD video: 1920 pixels x 1080 lines at 16 bit/color pixels, 30Hz frames/s (fps), compressed to 8-16 Mbit/s

Compression is not "for free" – (1)

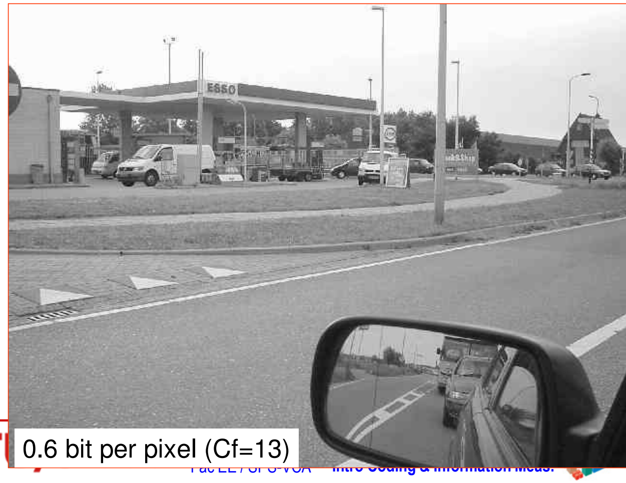
13

"lossy"

13

Compression is not "for free" – (2)

14



0.6 bit per pixel (Cf=13)

14

Compression is not "for free" – (3)

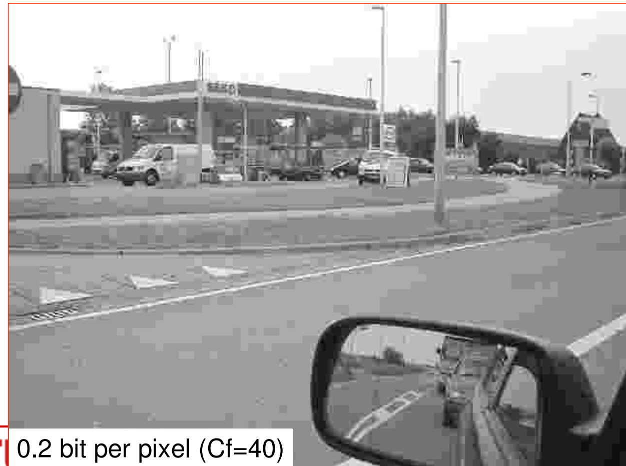
15

0.3 bit per pixel (Cf=26)

15

Compression is not "for free" – (4)

16

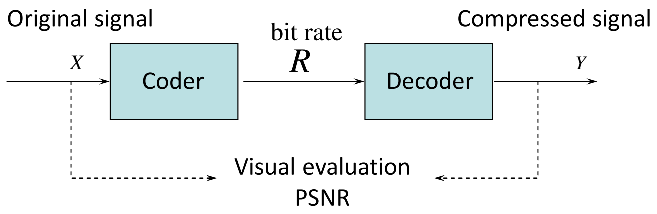


0.2 bit per pixel (Cf=40)

16

Quality of Compressed Signals – (1)

* **Visual evaluation (always important)**
* **Numerical evaluation: PSNR**



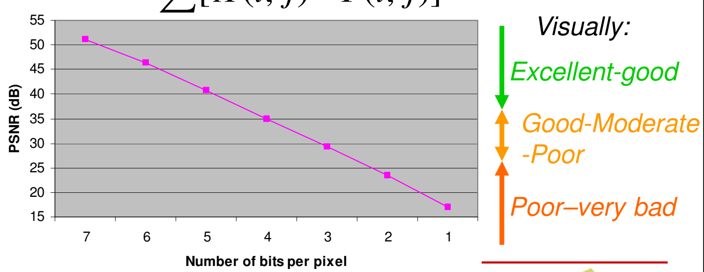
Quality of Compressed Signals – (2)

Number Score	Quality Scale	Impairment Scale
5	Excellent	Imperceptible
4	Good	(Just) perceptible but not annoying
3	Fair	(Perceptible and) slightly annoying
2	Poor	Annoying (but not objectionable)
1	Unsatisfactory (Bad)	Very annoying (Objectionable)

Quality of Compressed Signals

$$PSNR = \text{Peak Signal - to - Noise Ratio} =$$

$$10\log_{10}\frac{255^2}{\sum[X(i,j)-Y(i,j)]^2} \quad (\text{dB} = \text{d}$$



Obvious Compression of Video/Images

Two obvious ways to compress signals

1. Reduce the sampling rate (subsampling)
   - Fewer audio samples per second
   - Fewer pixels per picture line and fewer lines (images)
   - Fewer pictures per second (video)
2. Reduce the No. of amplitude levels per signal sample:
   - Pulse coded modulation *or* PCM
   - Obviously we will develop far more advanced techniques in this course

Case 2: Reduce the #Amplitudes – (1)

R(ed)G(reen)B(lue) samples, 24 bits (16,777,200 diff. colors)

Reduce the #Amplitudes – (4)

But real coding is always better…!

Mod 01 - A Part 2: Introduction / Relation to statistics of images

25

Questions to be Answered

Three questions will play a central role:

1. **Why can signals be compressed?**
2. How much can signals be compressed?
3. Which signal processing/information theory algorithms are most efficient in reaching maximum compression?

Answer: ***Because signal amplitudes are statistically redundant***

26

Why can Signals be Compressed? – (1)

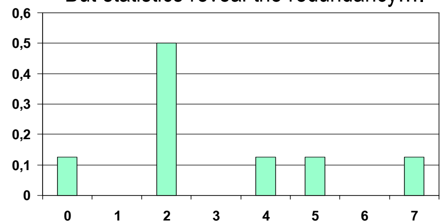
Example: Signal amplitudes being statistically redundant

Example**: What is required no. bits/sample for transmission?**

– Signal has integer amplitudes in range [0,7]
– 3 bit per signal sample (3 bps or bpp)

Signal	Code
0	*000*
1	*001*
2	*010*
3	*011*
4	*100*
5	*101*
6	*110*
7	*111*

But statistics reveal the redundancy….



27

Why can Signals be Compressed? – (2)

Because signal amplitudes are statistically redundant

Entropy Coding or Variable Length Coding

Example**:**

– Signal has integer amplitudes in range [0,7]
– **2 bits** per signal sample

Signal value	Probability	Codeword
0	*0.125*	*100*
1	*0*	
2	*0.5*	*0*
3	*0*	
4	*0.125*	*101*
5	*0.125*	*110*
6	*0*	
7	*0.125*	*111*

Question: What is the **shortest** average codelengh?

28

Why can Signals be Compressed? – (3)

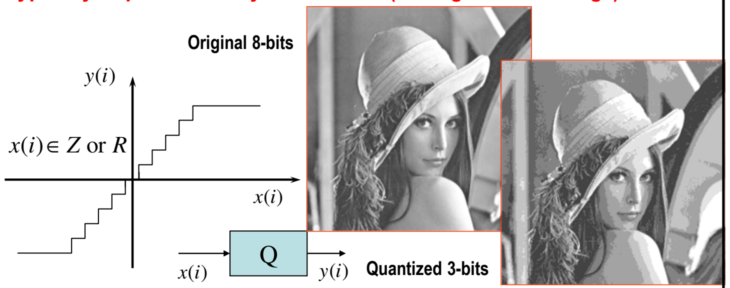
Because infinite accuracy of signal amplitudes is (perceptually) irrelevant : example color/detail irrelevancy

The difference between these two pictures is ***perceptually irrelevant***
because our visual system is insensitive for such small differences in color and fine details
But the files differ a factor of 3 in size!

Why can Signals be Compressed? – (4)

Because infinite accuracy of signal amplitudes is (perceptually) irrelevant

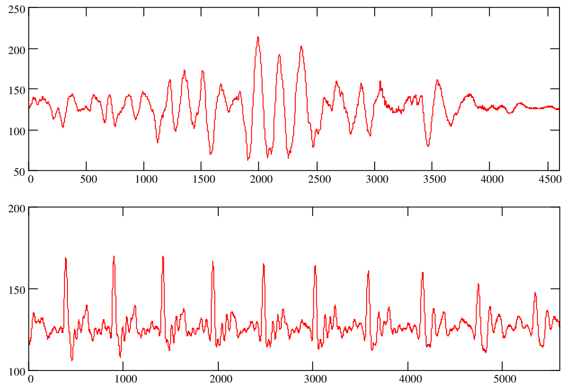
Typically implemented by a Quantizer (intelligent "rounding")



29

Why can Signals be Compressed? – (5)

Because signal amplitudes are mutually dependent



Signals are stochastic, change over time!

"S"

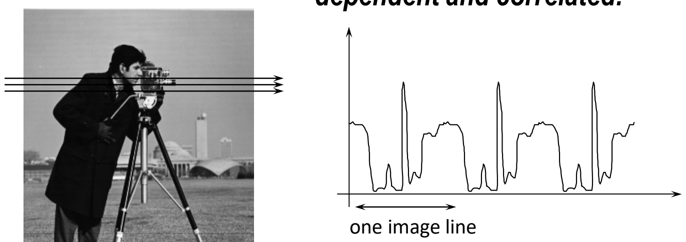
"O"

30

Why can Signals be Compressed? – (6)

Because signal amplitudes are mutually dependent

Example of video signal: change over image, but locally dependent and correlated!

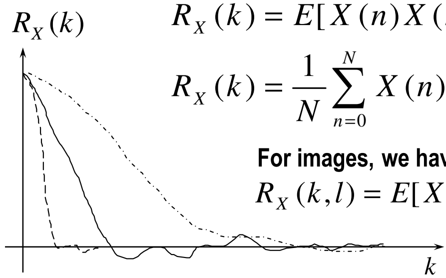


31

Why can Signals be Compressed? – (7)

Because signal amplitudes are mutually dependent

Dependency is commonly expressed by the *autocorrelation function*:



$$R_X(k) = E[X(n)X(n-k)]$$

$$R_X(k) = \frac{1}{N}\sum_{n=0}^{N} X(n)X(n-k)$$

For images, we have a 2D correlation

$$R_X(k,l) = E[X(n,m)X(n-k,m-l)]$$

32

Summarizing …