Introduction to Media Informatics

Dr. Mathias Lux
Alpen-Adria Universität Klagenfurt

First Of All: Todays Plan

- Whoami?
- Whoareyou?
- Todays schedule:
  - 9:00-10:30 part I
  - 12:00-13:30 part II
- Tomorrow
  - ???

Motivation: A YouTube Video

Motivation: A YouTube Video

- FallOut Nuka Break, Wayside Creations
- http://youtu.be/_iq8swQ7kpg

Video Production: Storyboard

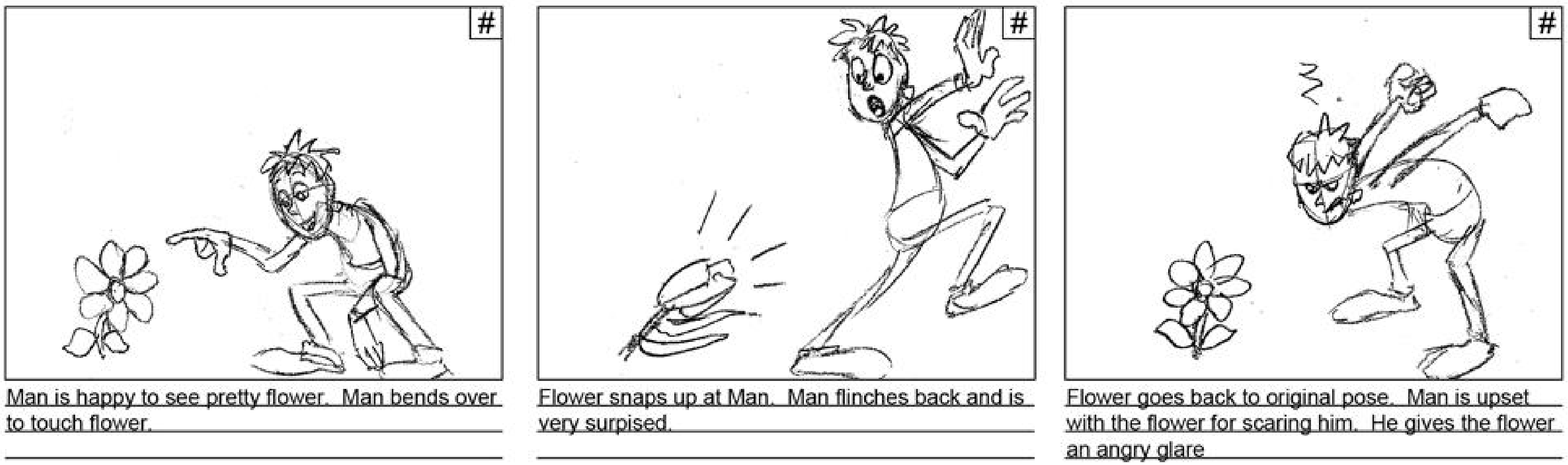
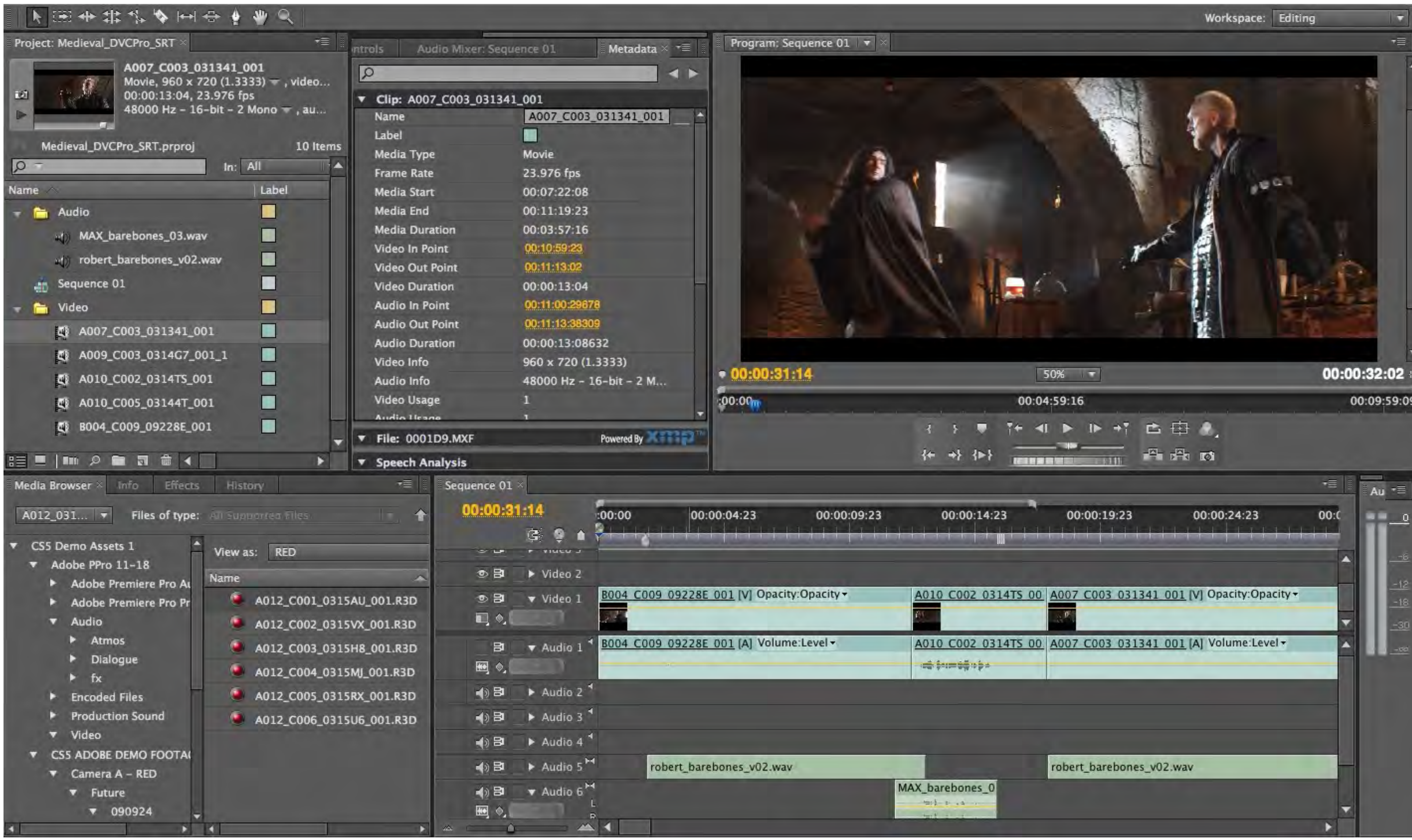


image src. http://software.intel.com/en-us/articles/3d-character-animation-part-1

Video Production: Recording

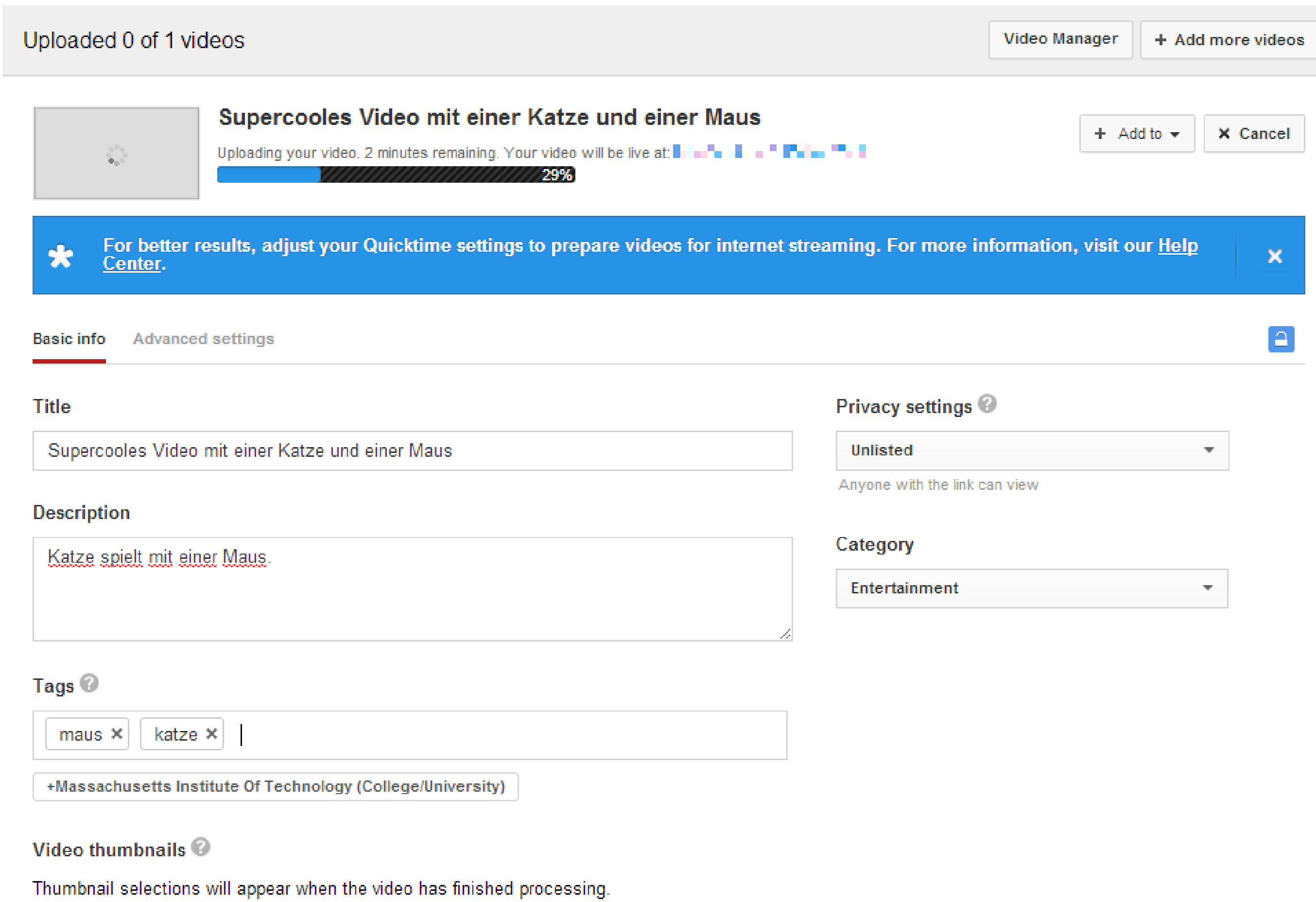
- Behind the Scenes, Wayside Creations
- http://youtu.be/dSjPaSWENw8?t=6m35s

Video Production: Cut, SFX & Editing



Video Production: Cut, SFX & Editing

Video Production: Publishing



Video Consumption: Viewing

- Web server, video server
- Client browser, client video player

Media Informatics – A Definition

- Informatik (de) == Computer Science (en)
- Computer Science and Media: Digital Media

- Digital Media is produced, consumed, shared, and edited by people
  - It`s not about data bases
  - It`s about people

Media – A Definition

- Storage
- Transmission Channels
- Systems & Services

Media – Storage

- Physical Media
  - analog: audio tapes, records, paper, …
  - digital: compact disc, DVD, hard drives
- Digital Formats
  - eBooks, HTML, …
  - MPEG-4, AAC, MP3 …

Media – Transmission Channel

- Distribution of Information
- Analog
  - Radio Channels, Newspaper
- Digital
  - DVB-T, Internet

Media – Systems and Services

- "In the Media"

Media Informatics - Challenges

- Transform traditional to digital media
- Authoring and Production
- New Media and Interaction

Challenges – Transform to Digital

- Digital Rights Management
  - How to borrow, annotate or preserve an eBook or an audio tape, ie. with iPad, Kindle, …
- Physical artifacts
  - Nothing to touch & possess, ie. games downloads
- Distribution
  - Costs and access, ie. mobile, distribution networ'
- Business models
  - Added value, profit, user content, …

Challenges – Authoring and Production

- Complex authoring tools
  - eg. motion capture, performance capture
  - http://youtu.be/ikrM5DrWj10 (D. Cage, Beyond)
- Automated methods
  - eg. multi camera, bullet time
  - http://youtu.be/rQmHFb4rd6U (Multicam system)

Challenges – Interaction

- Mobile devices & screens
  - Additional sensors, etc.
  - http://youtu.be/U0dOD3B8sg8 (B. Kaufmann)
- New devices
  - eg. Oculus Rift, …

Media – Interdisciplinary Approaches

- Cognitive Science
  - human factors and impact of media on people
- Social Science
  - impact and role of media on society
- Economic Science
  - economic aspects of media
- Computer Science
  - technical possibilities and limitations

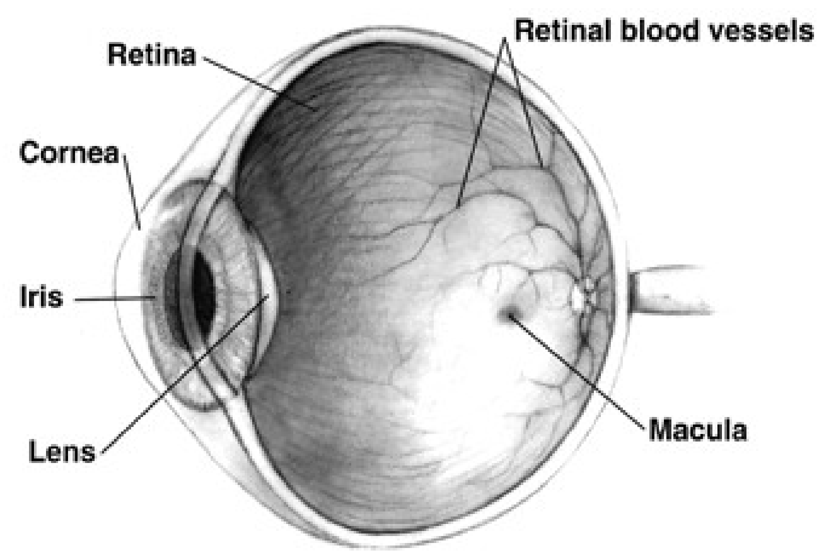
Human Perception vs. Reality

Human Perception

- How do people process information?
- Traditional senses
  - Sight, Hearing, Touch
  - Smell, Taste
- Other senses
  - Balance, acceleration, temperature, time, pain
  - Kinesthetic sense (Proprioception)

The Sense of Sight – Human Vision

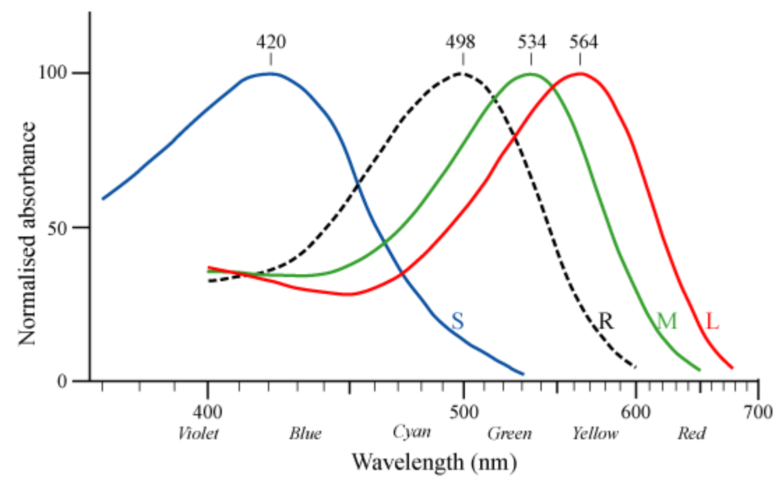
- The eye as instrument of perception
- Sensory capabilities
  - Cones (bright light): 6-7 Mio.
  - Rods (dim-light): 75-150 Mio.
  - Brain 'corrects' vision
    - e.g. blind spot



The human eye – Cones and Rods

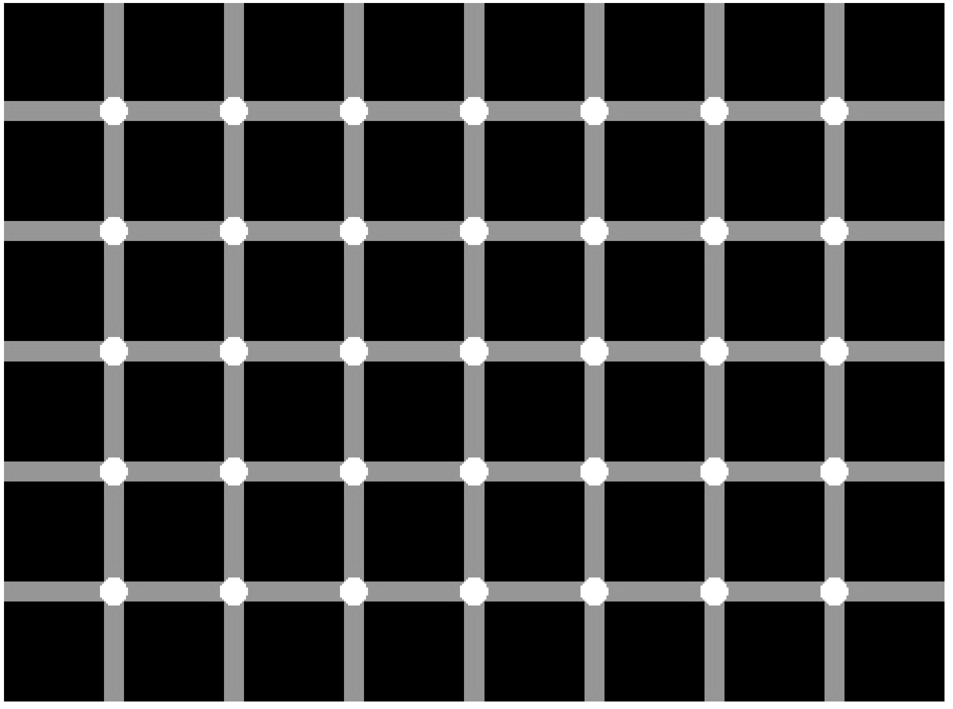
S-, M- and L-cones: Blue, green and red

- Responsiveness of cone types
- Number of cones / types
- etc.



The human eye …

- Count the black dots on the image:



The human eye …

- Rabbit or duck?

The human eye …

- Anamorphic illusions

See e.g. http://users.skynet.be/J.Beever/pave.htm

The human eye …

- Anamorphic illusions

See e.g. http://users.skynet.be/J.Beever/pave.htm

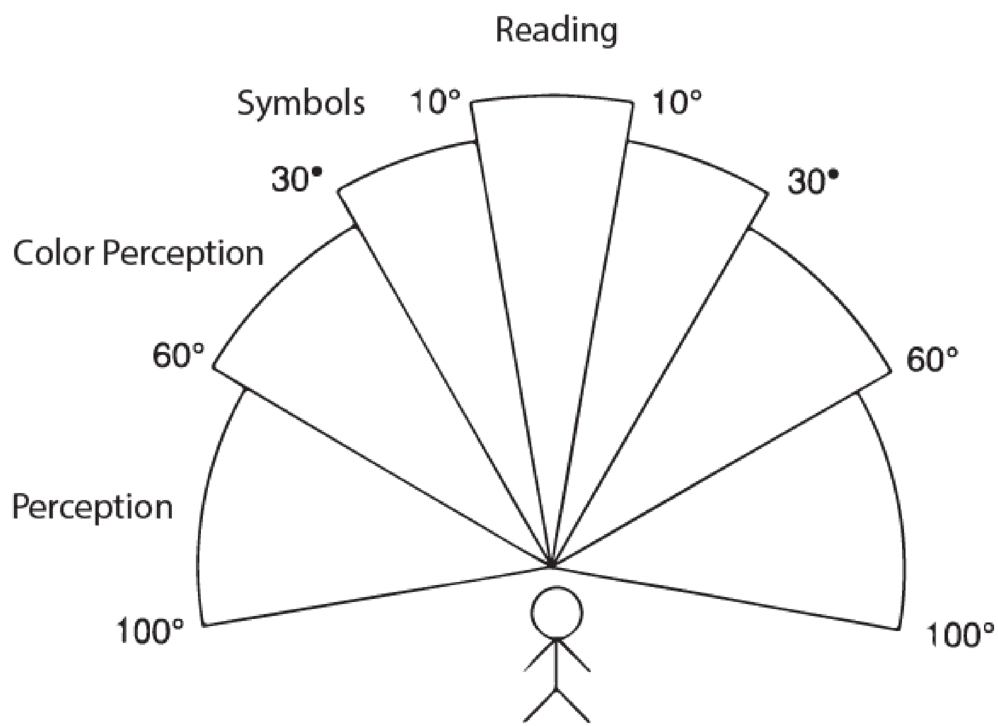
Human Perception – The Eye

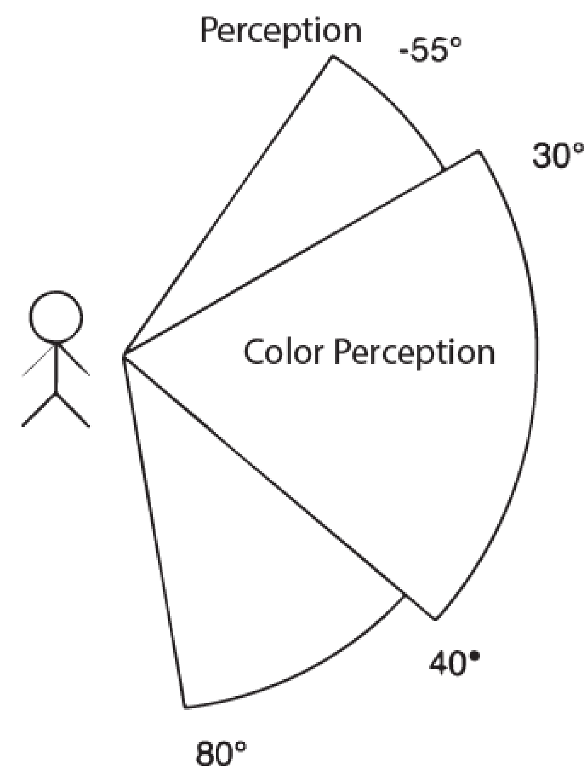
- Cones and rods are not evenly distributed
- Focusing is done by small muscles
- Color blindness affects 8% of the male, 1% of the female population
- 285M people are visually impaired worldwide (WHO, June 2012)

3D Vision

- Monocular cues
  - Relative & familiar size,
  - Motion Parallax, Occlusion
  - Depth from optical expansion, …
- Binocular cues
  - Stereoscopy, …

The human eye – Spatial Properties





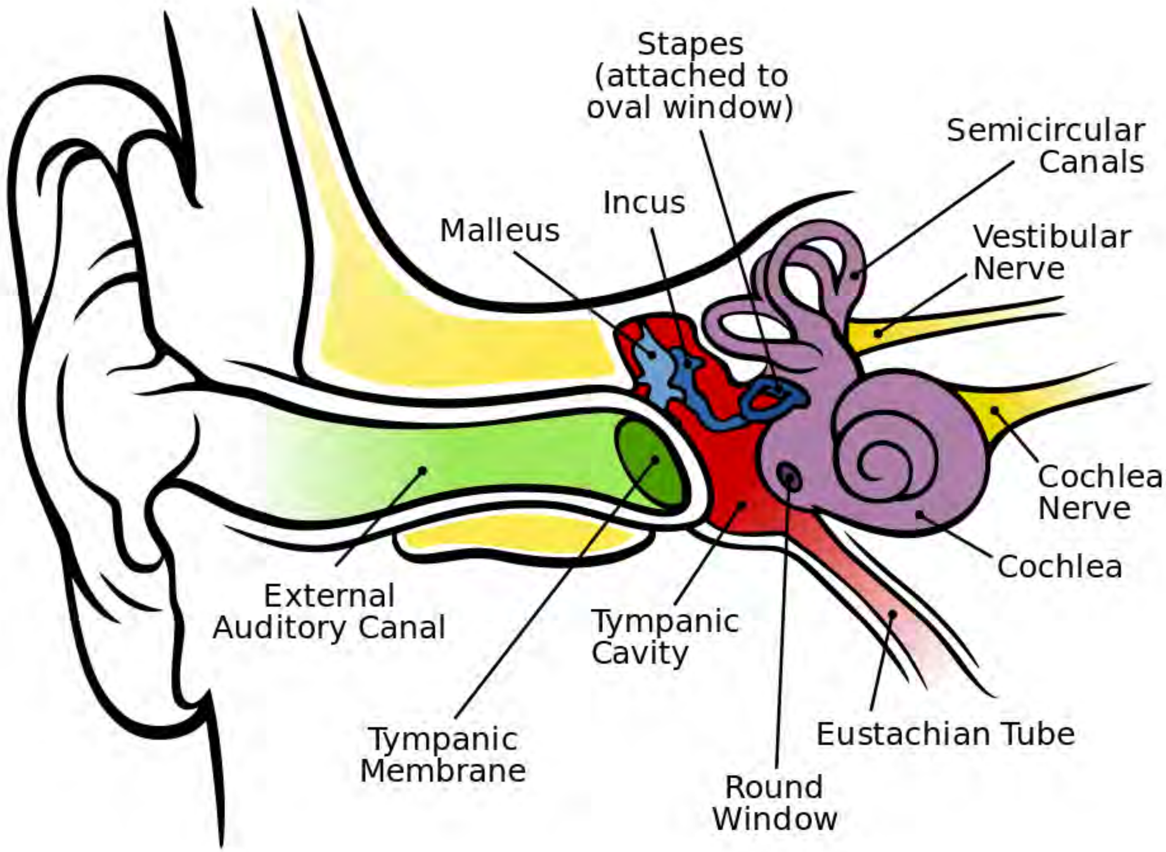
The human eye – Temporal Properties

- Visual information is perceived in 15-50ms
- Subsequent images are perceived as motion
  - ie. motion pictures, videos
- 3-5 images / second imply movement
- 20-25 images and more imply smooth movement

The human eye – Temporal Properties

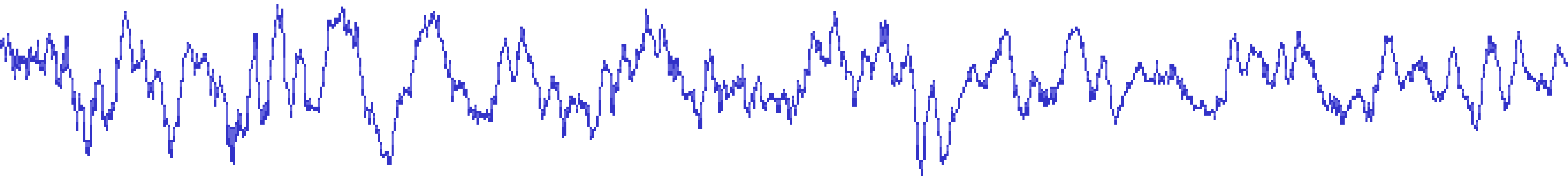
- Demo: Videos with 6, 12, 20, 24 frames.
  - The Simpsons Movie Trailer

The Sense of Hearing



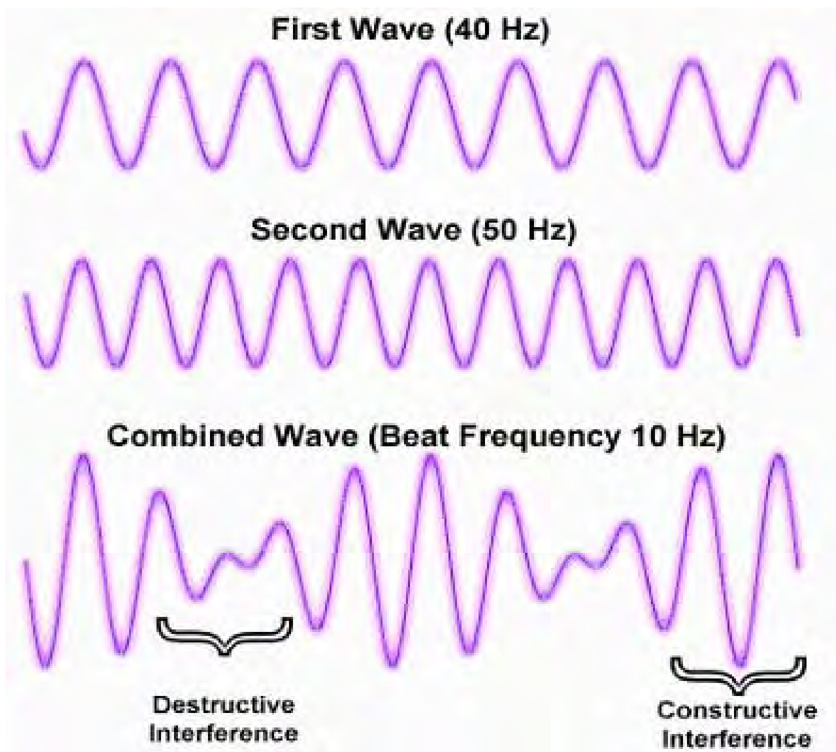
Src. http://en.wikipedia.org/wiki/File:Anatomy_of_the_Human_Ear.svg

What is sound?



What is sound?

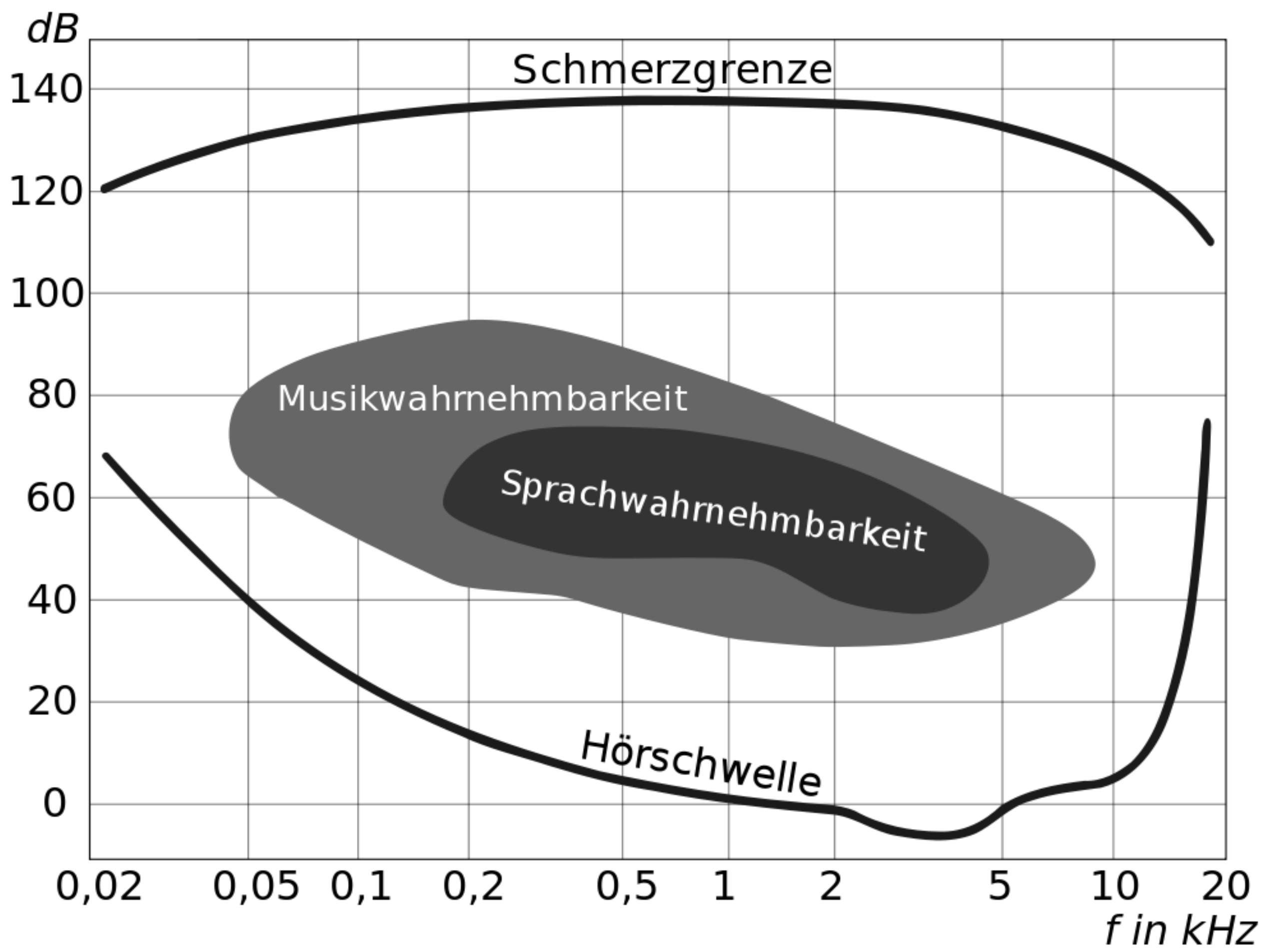
- Multiple sounds at the same time?



What is sound?

- Audacity Demo

The Sense of Hearing



Src. http://de.wikipedia.org/wiki/Datei:Hoerflaeche.svg

The Sense of Hearing

The Sense of Hearing

- Human ear is very sensitive to temporal changes (2-3ms)
- Spatial resolution is worse than with the eye
- Exponential frequency change is perceived linear
- Brain also "adds" information to everything heard

Digitizing Information

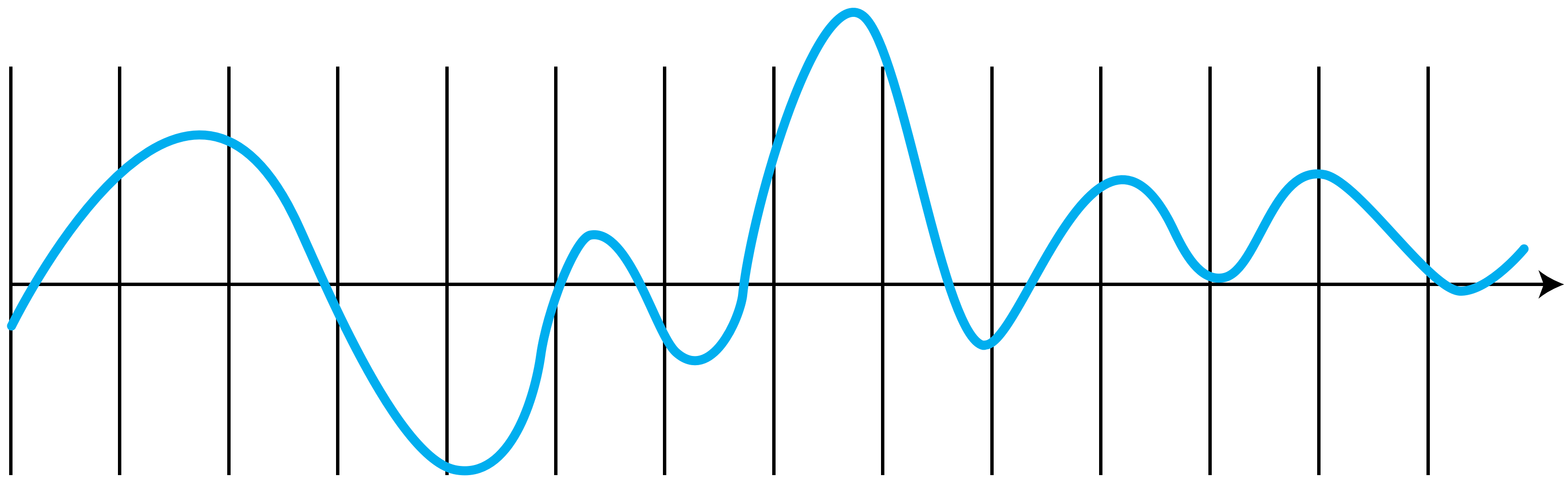
- Converting analog signals to digital ones
- Analog-to-digital converter

What is an analog Signal?

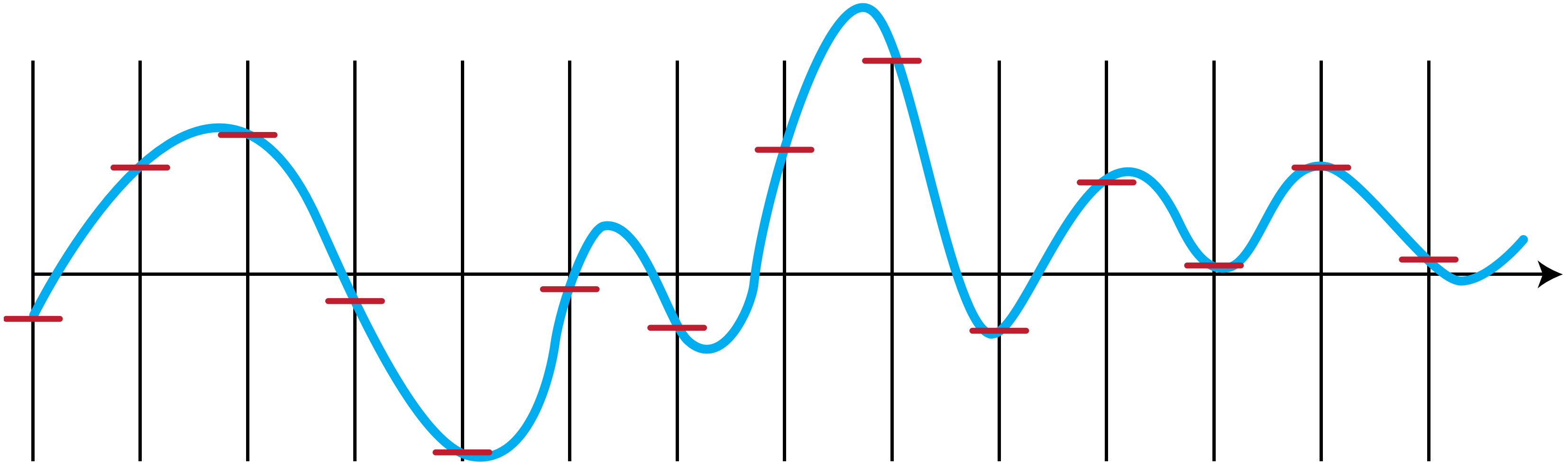
- Continuous signal
- Varying a physical feature
  - e.g. more light, less light
  - voltage, pressure, current, …

What is a digital Signal?

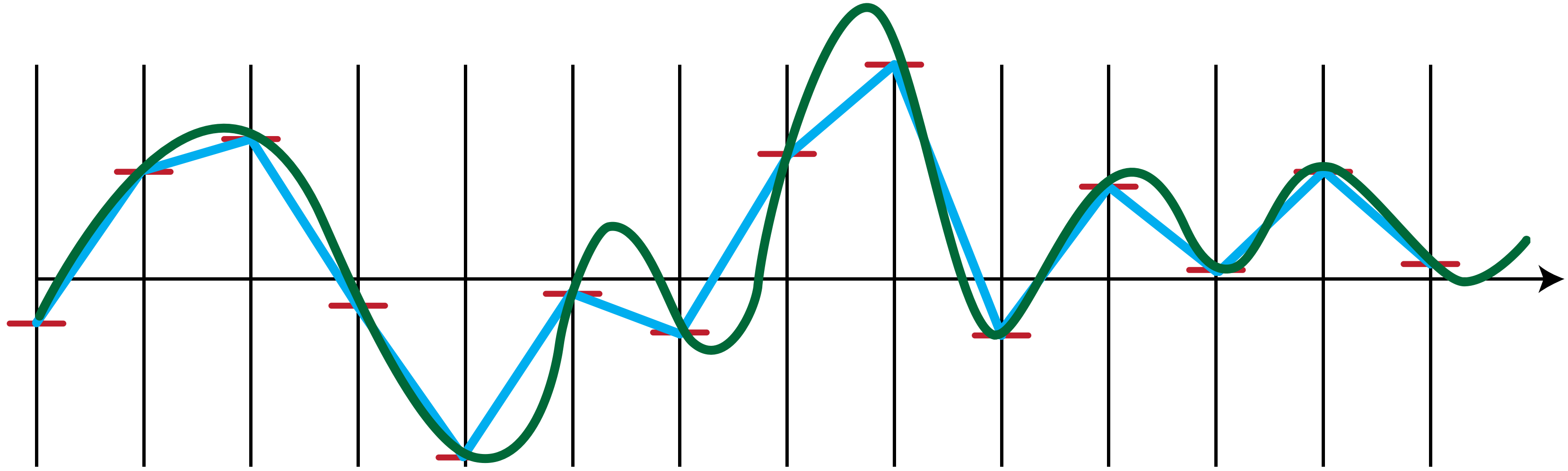
- Has a limited and discrete range for variables
- Is typically an approximation of a analog signal
- Allows for processing in a digital computer.



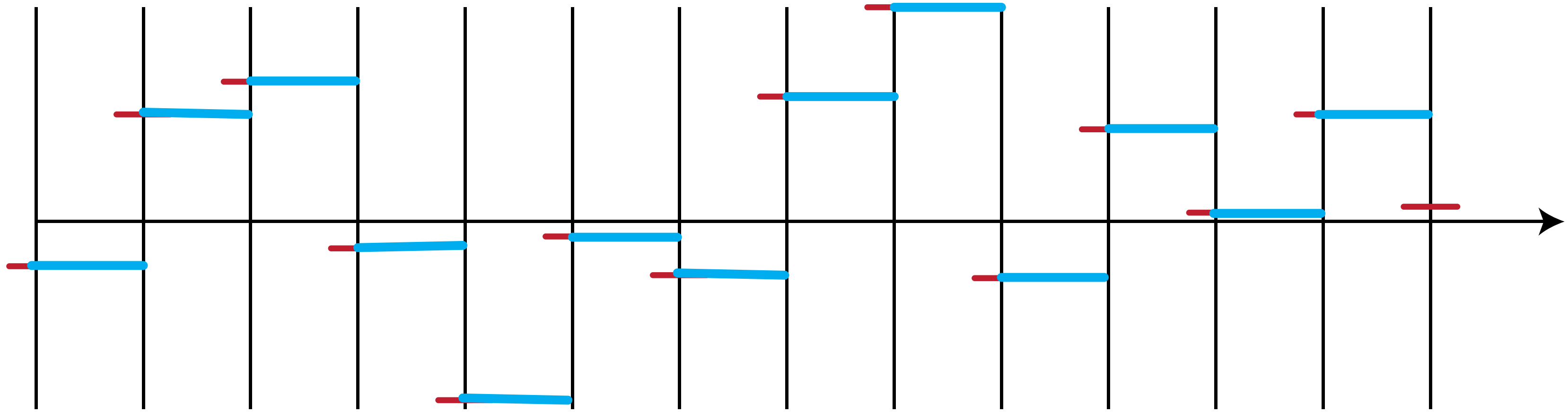
Digitization



Digitization

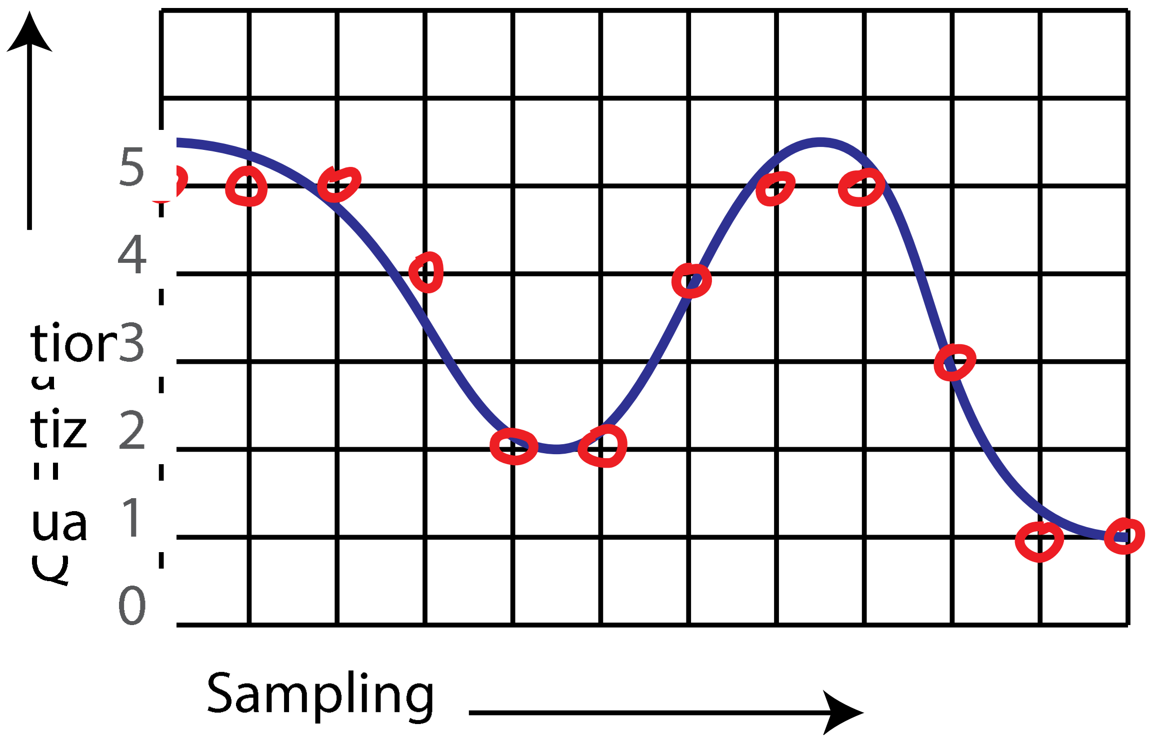


Digitization



Quantization & Sampling

- Digital signal: 5, 5, 5, 4, 2, 2, 4, 5, ...
- Binary: 101101101100010010100101...



Sampling Theory

- Nyquist–Shannon sampling theorem
  - Harry Nyquist and Claude Shannon (1949)

If a function $x(t)$ contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $1/(2B)$ seconds apart.

Additional Concepts

- Information Theory
- Entropy & Entropy Coding
- Digital Rights Management
- Licenses

Information Theory

- Quantification of information
- Limits for compression
- Reliable storing & communication

Information Theory

Which signal carries most information?

1. 010101010101010101010101
2. 00000000000000000
3. 01101001110101100111101010

Entropy

- Key measure in information theory
- Measures the uncertainty "for the next value"
  - Cp. How "chaotic" is your data?

- 00000000000000000
  - It's certain that the next value will be 0
  - Entropy is zero

Information Theory & Entropy

- The higher the entropy, the higher the quantified information
- The lower the entropy, the lower the quantified information
  - Information can be compressed
  - 0000000000000000 -> 16x0

Entropy

- Probability of occurrence of a symbol z_i is p_i
- If there is 50% for letter “a” =: z_1 to occur
  - $p_1 = 0.5$
- The entropy of a letter z_i is then

$$h_i = ld(\frac{1}{p_i})$$

Entropy

- Expected entropy H of a message
- Is based on each symbol of an alphabet A

$$H = \sum_{i=1}^{|A|} (p_i \cdot h_i) = \sum_{i=1}^{|A|} (p_i \cdot ld(\frac{1}{p_i}))$$

Entropy Example

z	p	1/p	h	p*h
a	0,50	2	1	0,5
b	0,25	4	2	0,5
c	0,25	4	2	0,5
				H=1,5

$$H = \sum_{i=1}^{|A|} (p_i \cdot h_i) = \sum_{i=1}^{|A|} (p_i \cdot ld(\frac{1}{p_i}))$$

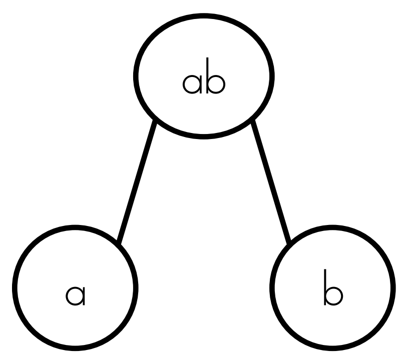
Huffman Coding

- Given an alphabet with probabilities for each symbols.
- We construct a code, that the entropy of a message is maximized.
- That is done with an algorithm …

Huffman Coding

- Find the two symbols with the lowest probabilities.
- Combine those two in a tree with to leaves.
- Combine them in the table.
- If there is more than one symbol left, start over.

Symbol	p
a	0,10
b	0,15
c	0,30
d	0,16
e	0,29



Symbol	p
ab	0,25
c	0,30
d	0,16
e	0,29

1. Lessig – Laws that choke Creativity

- Laws that choke creativity
  - http://www.ted.com/talks/larry_lessig_says_the_law_is_strangling_creativity.html
  - http://www.youtube.com/watch?v=7Q25-S7jzgs

Content Licenses

- Copyright
  - All rights reserved
  - Rights holder grants some rights to buyers
- Copyleft
  - GNU GPL: viral licenses
  - Can be used for media, interpretation of "source"
  - GNU Free Documentation License
  - See also http://www.gnu.org/licenses/license-list.html#FreeDocumentationLicenses

Creative Commons

- Video
  - http://creativecommons.org/about
- Demo
  - http://creativecommons.org/choose

Src. http://creativecommons.org/about

Introduction to Cryptography

- What is cryptography?
- What is cryptography used for?
- Basic principles & common methods

Readings: Menezes, Alfred J., Paul C. Van Oorschot, and Scott A. Vanstone. *Handbook of applied cryptography*. CRC press, 2010. http://cacr.uwaterloo.ca/hac/

What is Cryptography?

- Information as value
- Cryptography to secure this value
- For military, government& diplomatic services
  - ie. make sure "others" cannot access or manipulate information.

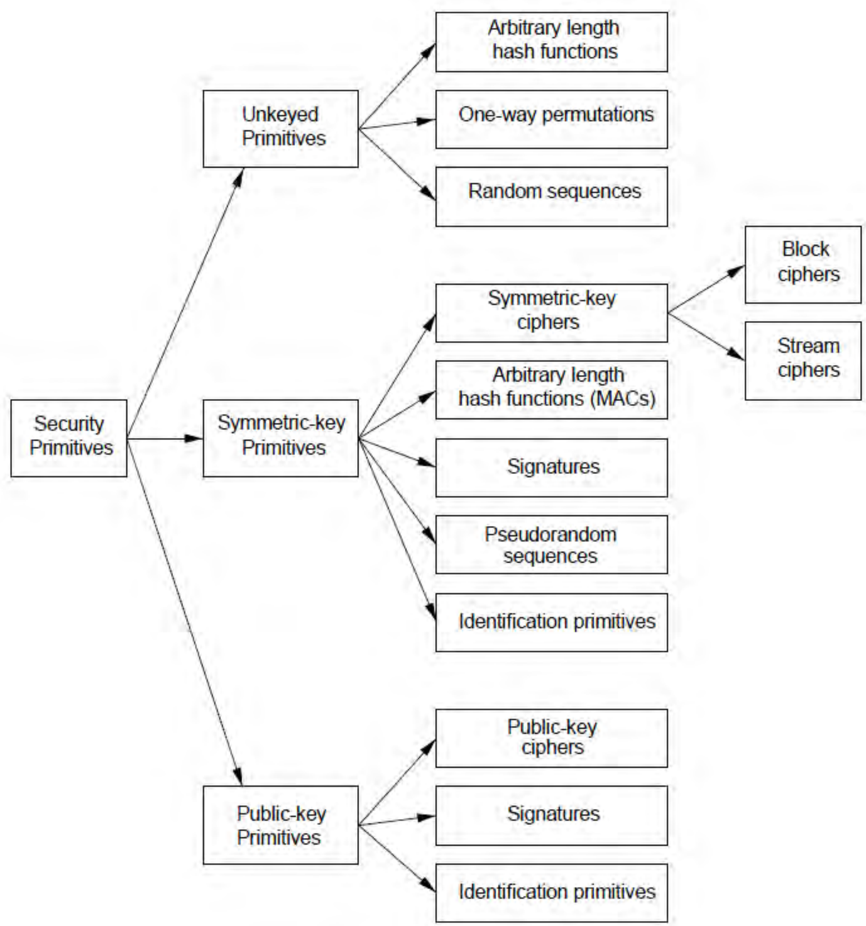
What is the use of Cryptography? (1)

privacy or confidentiality	keeping information secret from all but those who are authorized to see it.
data integrity	ensuring information has not been altered by unauthorized or unknown means.
entity authentication or identification	corroboration of the identity of an entity (e.g., a person, a computer terminal, a credit card, etc.).
message authentication	corroborating the source of information; also known as data origin authentication.
signature	a means to bind information to an entity.
authorization	conveyance, to another entity, of official sanction to do or be something.
validation	a means to provide timeliness of authorization to use or manipulate information or resources.
access control	restricting access to resources to privileged entities.

What is the use of Cryptography? (2)

certification	endorsement of information by a trusted entity.
timestamping	recording the time of creation or existence of information.
witnessing	verifying the creation or existence of information by an entity other than the creator.
receipt	acknowledgement that information has been received.
confirmation	acknowledgement that services have been provided.
ownership	a means to provide an entity with the legal right to use or transfer a resource to others.
anonymity	concealing the identity of an entity involved in some process.
non-repudiation	preventing the denial of previous commitments or actions.
revocation	retraction of certification or authorization.

Taxonomy of Crypto-Methods



from Menezes, Alfred J., Paul C. Van Oorschot, and Scott A. Vanstone. Handbook of applied cryptography. CRC press, 2010.

Trapdoor One-Way Functions

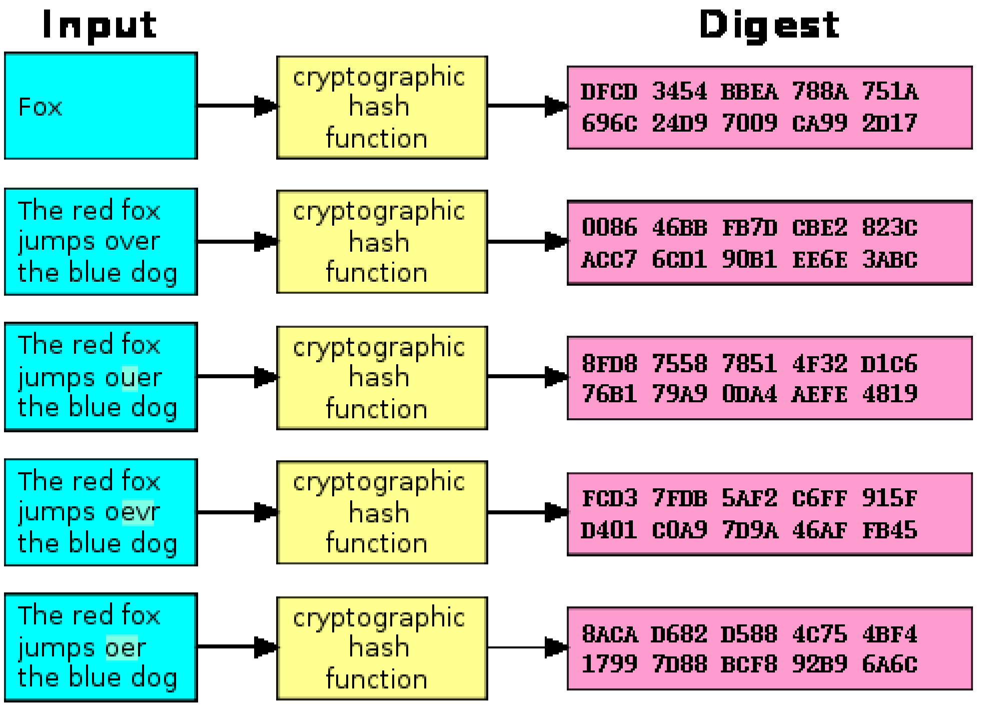
Definition. A function f from a set X to a set Y is called a one-way function if $f(x)$ is "easy" to compute for all x in X but for "essentially all" elements y in $Im(f)$ it is "computationally infeasible" to find any x in X such that $f(x) = y$.

Trapdoor One-Way Functions

- In other words: It's hard to reverse the function
  - Example: function y=2*x is easy to reverse
- Additional constraints
  - computationally inexpensive
  - few collisions
- Examples
  - document hashes
  - password hashes

Trapdoor One-Way Functions

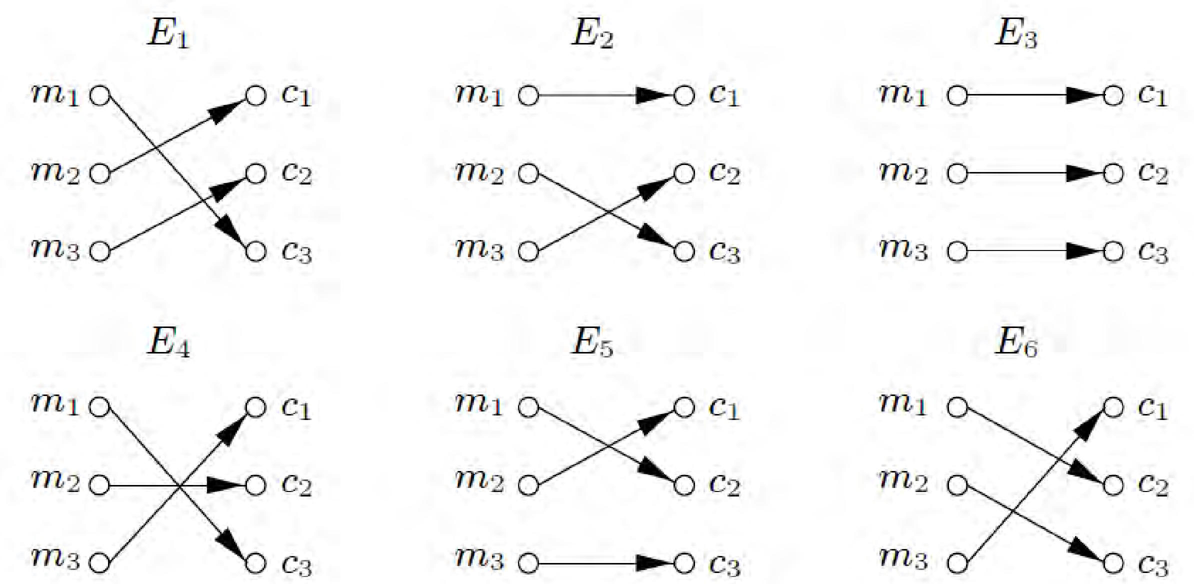
- Common functions are
  - MD5, SHA-0, SHA-1, SHA-2, RIPEMD, …



Src. http://en.wikipedia.org/wiki/Cryptographic_hash_function

Permutations

- Change symbols based on a rule

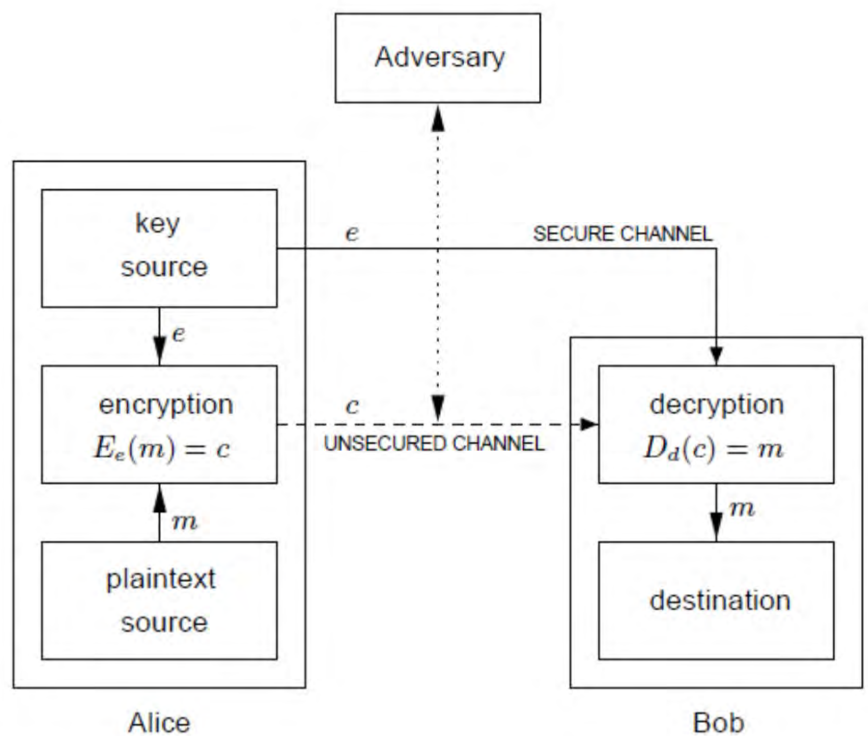


Permutations

- Historical Examples
  - Cesar Shift
  - Enigma: product of permutations
- Easy to break
  - Symbol probabilities lead to decryption table

Symmetric Key

- A key is applied to a message to encrypt it
- The same key decrypts the message

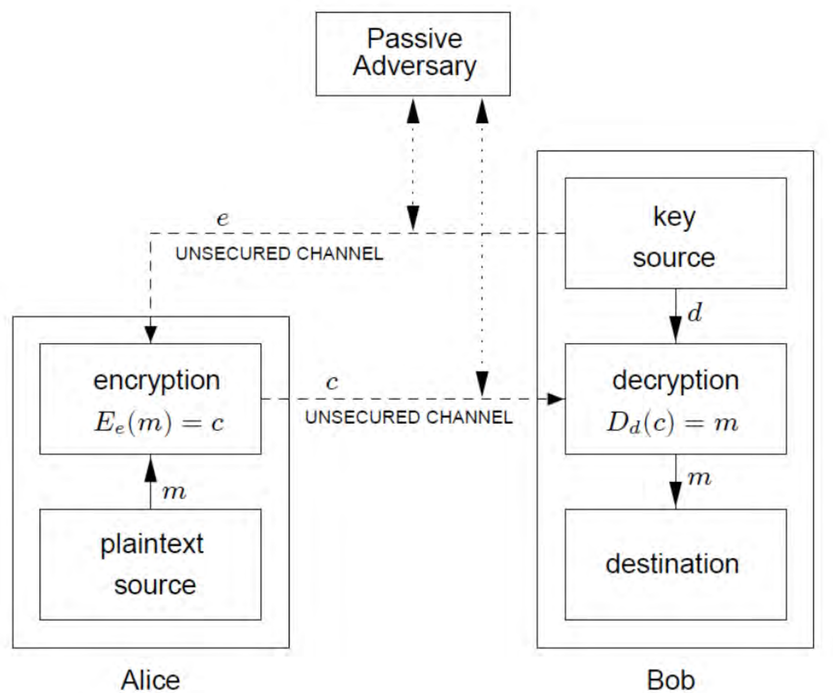


Symmetric Key

- The larger the key, the better the encryption
  - If the key is as large as the message and only used once this method is theoretically unbreakable
- This method is used for HTTPS, SSH, etc.
  - For on-the-fly encryption
- Common methods are
  - DES, Triple-DES, AES, Blowfish

Public Key Cryptography

- Two keys (a corresponding pair) are used
  - public key is known to anyone
  - private key is known only by one entity



Public Key Cryptography

- Most prominent example: RSA (1977)
  - Ron Rivest, Adi Shamir and Leonard Adleman
- Implementations are available
  - Pretty Good Privacy (PGP)
  - Gnu Privacy Guard (GPG)
- Public keys are stored on key server
  - Users can search for keys of others
- Based on two large prime numbers

Digital Rights Management

- Software & hardware framework that makes sure that content can only be consumed
  - in the right place
  - at the right time
  - by the right entity

Digital Rights Management

Examples

- Pay per view for a movie
- Download music for a specific ecosystem
- Disable storing of streams (audio, video)
- HDCP, Content Scramble System, etc.

Discuss turn in business models for digital music!

Exercise: Readings

- Read http://craphound.com/msftdrm.txt
- Prepare a discussion statement on
  - what you think about the things said
- Discussion will take place in the next lesson.

Introduction to Media Informatics: Text and Hypertext

Dr. Mathias Lux
Alpen-Adria Universität Klagenfurt

Agenda

- Text & Markup
- Hypertext
- WWW
  - URL, DNS, HTTP
- HTML
- CSS, JavaScript

Text Formats

- American Standard Code for Information Interchange (ASCII)
- Defines 128 character, 95 of them printable

USASCII code chart

				$b_7 b_6 b_5$ →	0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
b_4	b_3	b_2	b_1	Row ↓ Column →	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	|
1	1	0	1	13	CR	GS	-	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	SI	US	/	?	O	—	o	DEL

Src. Wikipedia, http://en.wikipedia.org/wiki/ASCII

ASCII

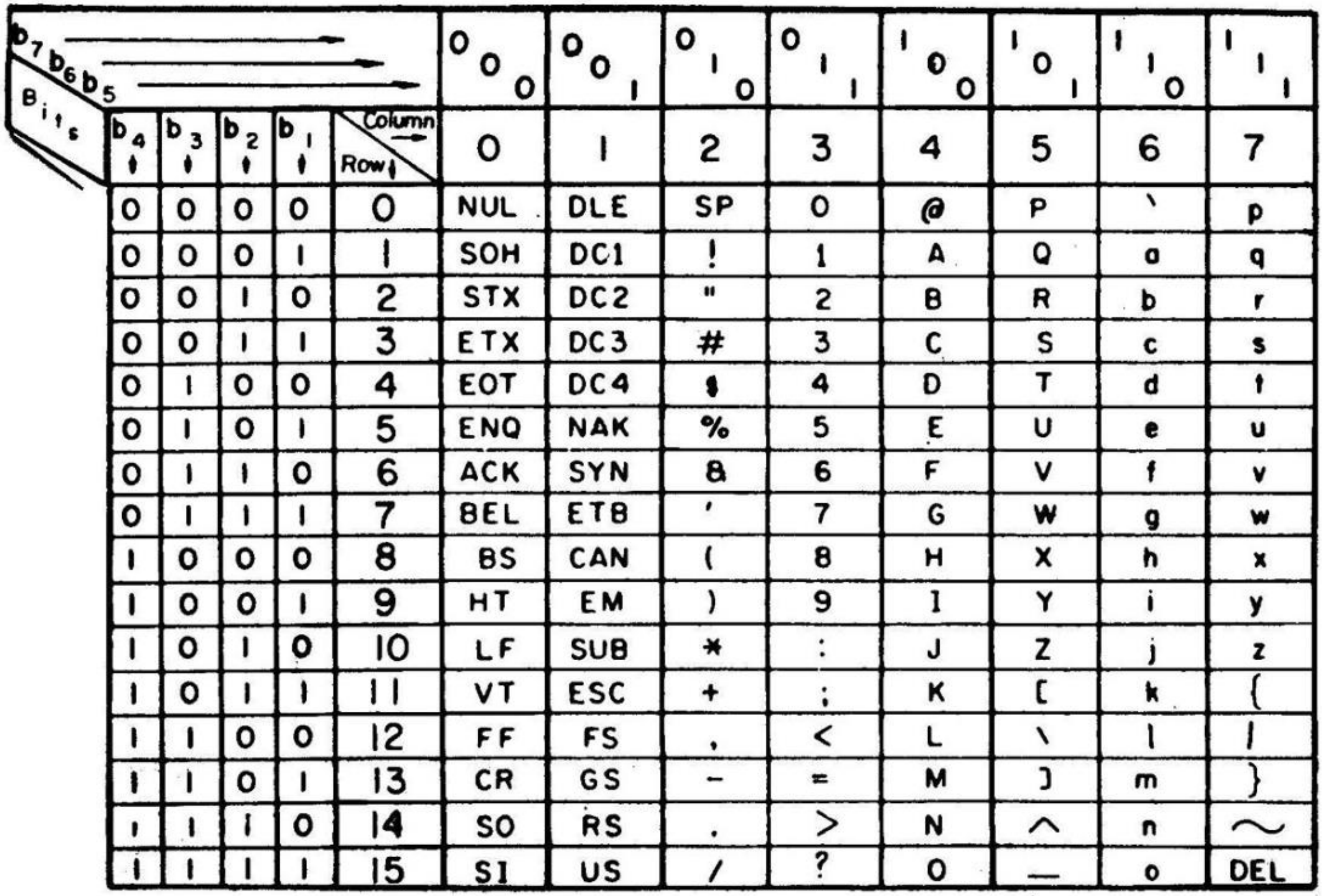
- Letter has a 7-bit code
  - A -> $1000001_2 = 64_{10}$
- Control Codes also have 7-bit codes

000 0111	007	7	07	BEL	BEL	^G	\a	Bell
000 1000	010	8	08	BS	BS	^H	\b	Backspace[d][e]
000 1001	011	9	09	HT	HT	^I	\t	Horizontal Tab[f]
000 1010	012	10	0A	LF	LF	^J	\n	Line feed
000 1011	013	11	0B	VT	VT	^K	\v	Vertical Tab
000 1100	014	12	0C	FF	FF	^L	\f	Form feed
000 1101	015	13	0D	CR	CR	^M	\r	Carriage return[g]

Src. Wikipedia, http://en.wikipedia.org/wiki/ASCII

Which Characters are Missing?

USASCII code chart



$b_7 b_6 b_5$					0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
b_4	b_3	b_2	b_1	Column → / Row ↓	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	|
1	1	0	1	13	CR	GS	-	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	SI	US	/	?	O	_	o	DEL

Src. Wikipedia, http://en.wikipedia.org/wiki/ASCII

Text Formats: ISO/IEC 8859-1

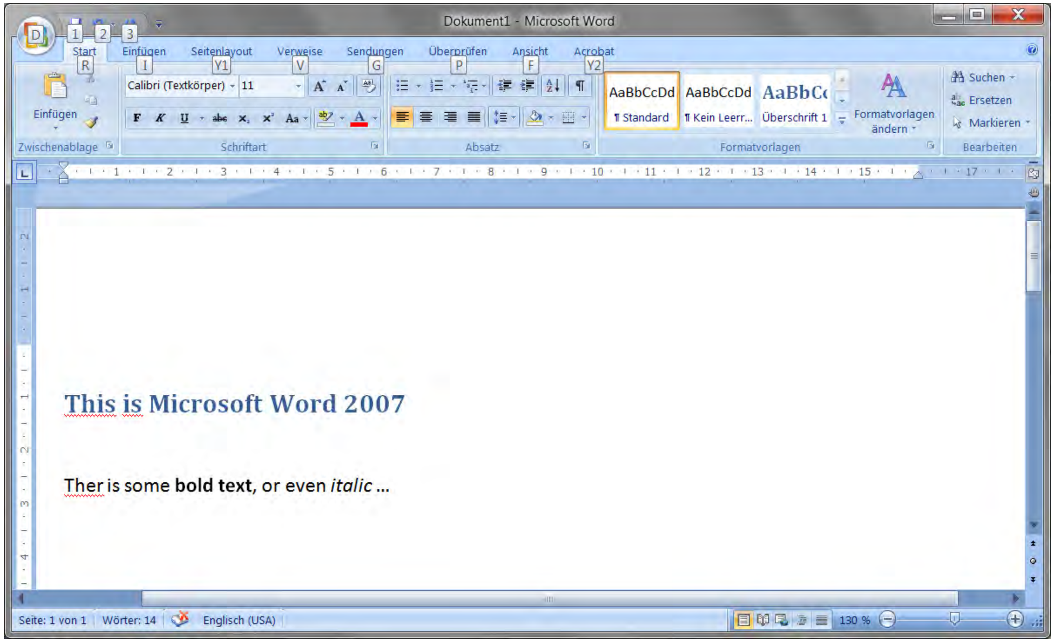
- Commonly referred to as "Latin-1"
- Each character has an 8-bit code
- 191 characters supported
- Sufficient for many Western European languages
  - But still some characters missing

Text Formats: Unicode Standard

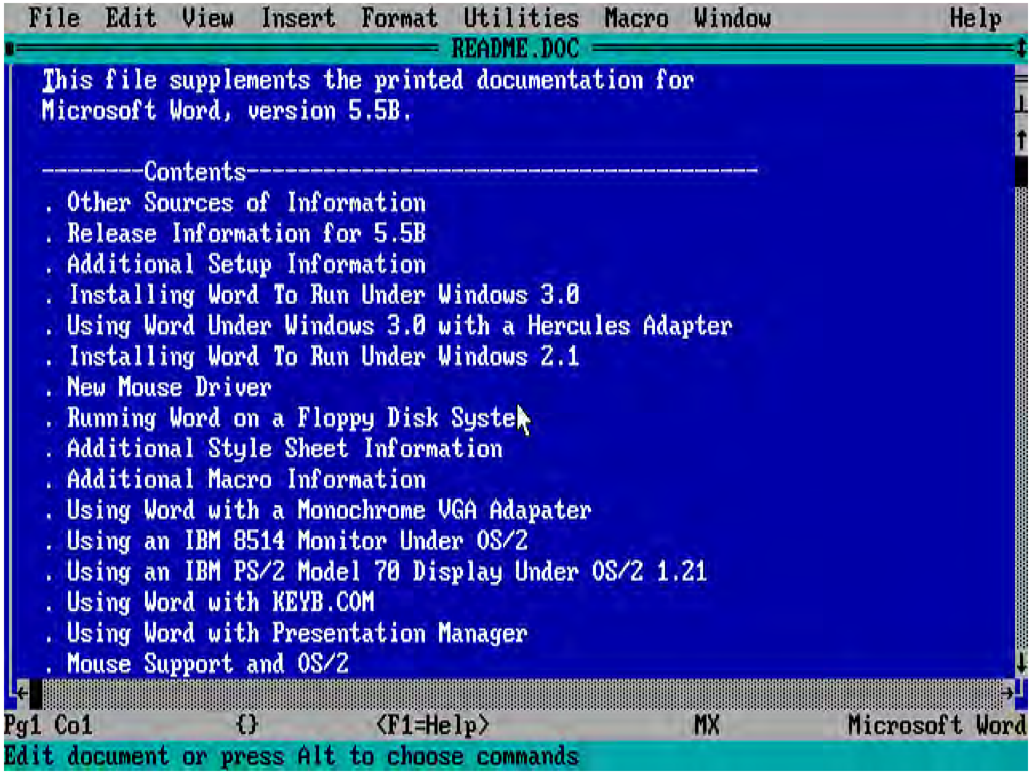
- More than 100,000 characters
- Supporting left-to-right and right-to-left scripts
- Defines character properties, rendering, ...
- Proposes different encodings like UTF-8, ..

WYSIWYG Text Editing

- Short for “What you see is what you get”
  - Means that you can actually see what it will look like when printing it …



WYSIWYG Text Editing



WYSIWYG: Related Concepts

- Type faces
  - were "bought with the printer"
  - expensive printers had nice postscript fonts
- Programs would send to a printer
  - a font set command and
  - the text to print
- Printers would even split documents in pages

WYSIWYG: True Type Fonts

- Developed 1980s by Apple and Microsoft
- Competitors of Adobe Type 1 (Postscript) fonts
  - Licensing, etc.
- Provide a more flexible way to
  - define outlines of fonts
  - control display at different font sizes
- First common font families
  - Times Roman, Courier, Helvetica

WYSIWYG: True Type Fonts

- Not the same look on every system
  - different outline -> pixel renderers
- Can now be loaded into printers as “soft fonts”
- The open source implementation of True Type is called “Free Type”
- OpenType is the successor created by Microsoft and Adobe

Fonts: Demo

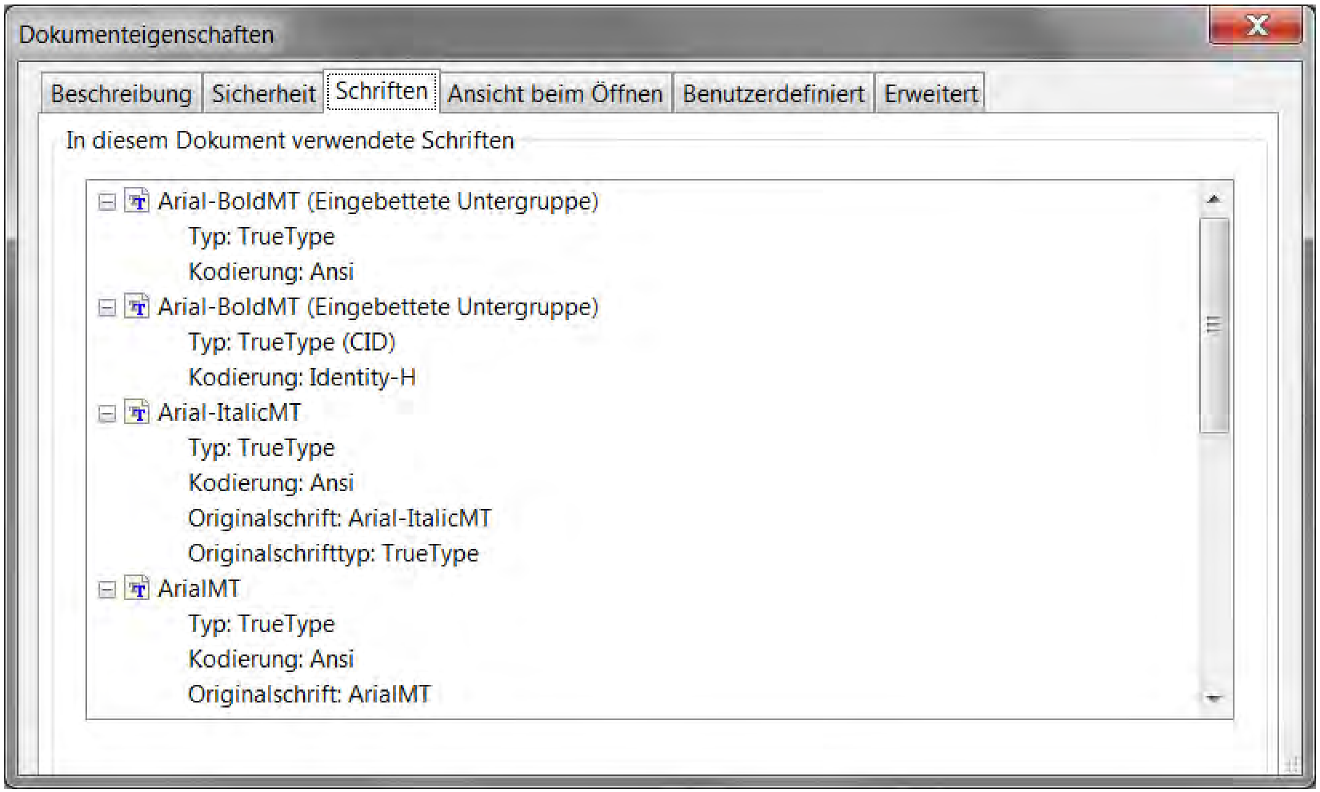
- Font Business
  - LinoType: http://www.linotype.com/
  - dafont.com: http://www.dafont.com/
- Font differences
  - PDF with bitmap fonts
  - PDF with Type 1 fonts
  - PDF with TrueType / OpenType

Some Font Facts ...

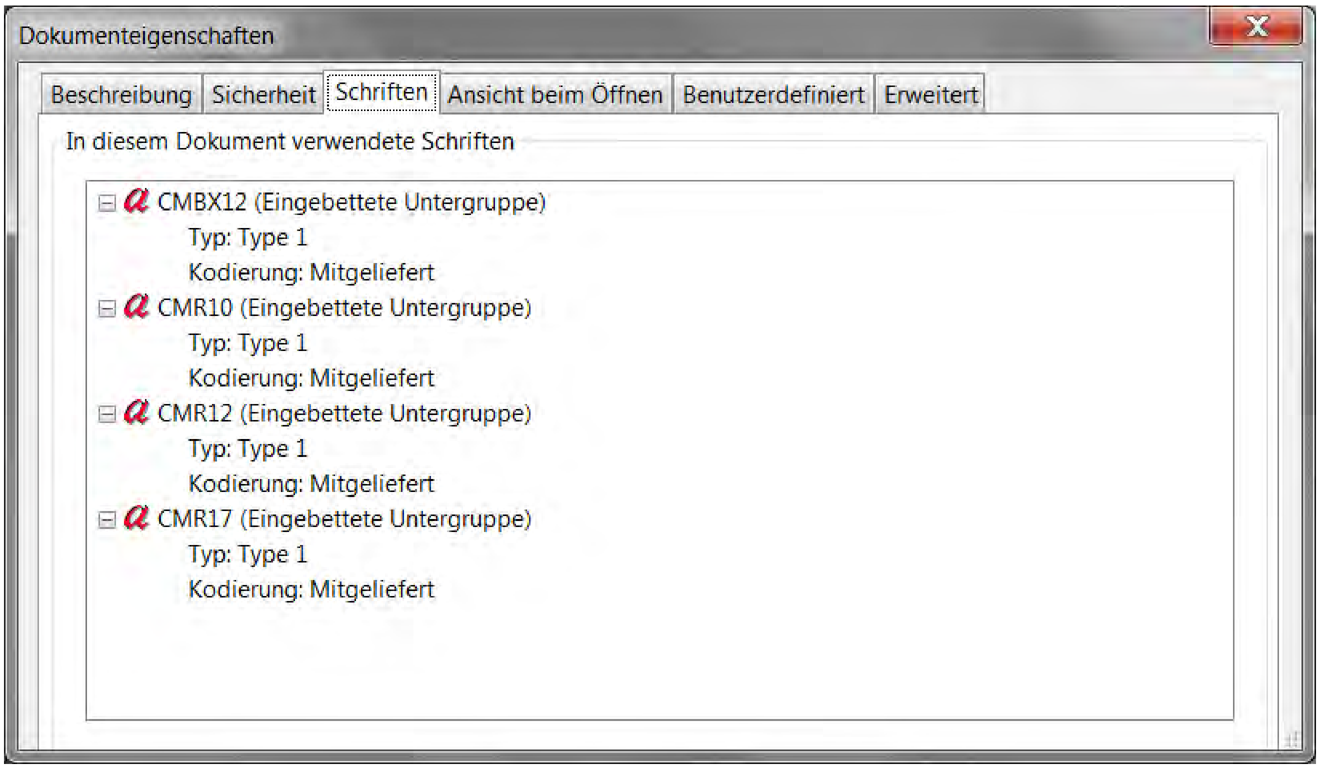
- A font can be
  - Serif, Sans Serif, Slab Serif
  - Monospaced

Times New Roman:	The quick brown fox jumps over the lazy dog.
Droid Sans:	The quick brown fox jumps over the lazy dog.
Courier New:	The quick brown fox jumps over the lazy dog.

Font Embedding vs. Reference

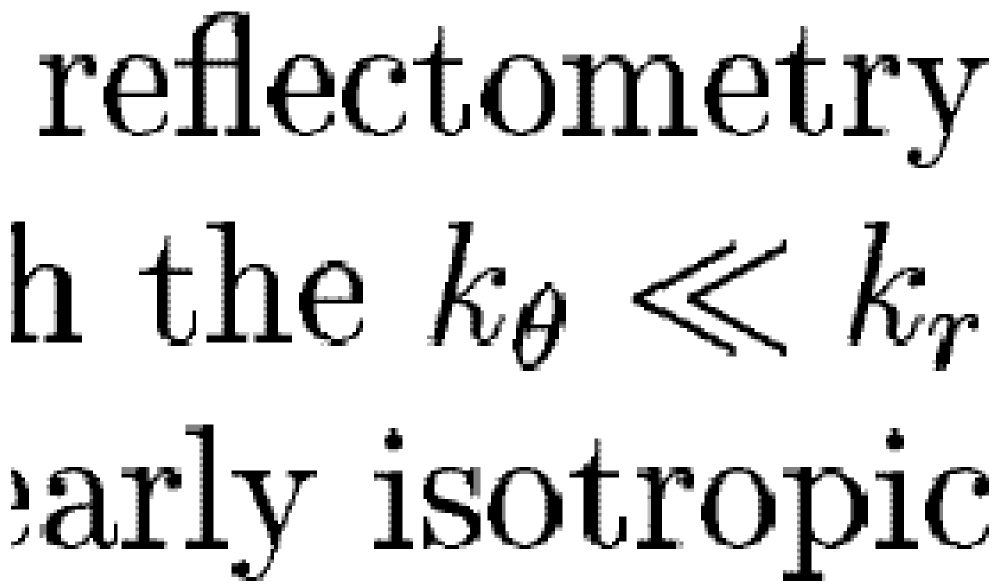


Font Embedding vs. Reference



Bitmap Fonts

- Pre-rendered fonts embedded in a document
- Lead to aliasing & artifacts when zoomed in



Typography

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec pellentesque odio in risus sodales aliquam. Aliquam orci tellus, elementum ultrices felis vitae, hendrerit gravida dui. Aenean sapien ligula, rhoncus et elementum quis, elementum eget augue. Aenean bibendum, elit sed interdum malesuada, diam turpis facilisis sem, vitae condimentum ipsum velit eget est. Sed in mauris tincidunt, tincidunt tellus sit amet, commodo ligula. Pellentesque porttitor, mauris sit amet malesuada fringilla, tortor justo dictum nibh, ac commodo lectus nisl at ante. Cras vitae neque fringilla, porttitor augue quis, euismod risus.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec pellentesque odio in risus sodales aliquam. Aliquam orci tellus, elementum ultrices felis vitae, hendrerit gravida dui. Aenean sapien ligula, rhoncus et elementum quis, elementum eget augue. Aenean bibendum, elit sed interdum male-suada, diam turpis facilisis sem, vitae condimentum ipsum velit eget est. Sed in mauris tincidunt, tincidunt tellus sit amet, commodo ligula. Pellen-tesque porttitor, mauris sit amet malesuada fringilla, tortor justo dictum nibh, ac commodo lectus nisl at ante. Cras vitae neque fringilla, porttitor augue quis, euismod risus.

LaTeX

- Demo ...

Markup Languages

- Text (content) is annotated by markup,
  - cp. “marking up” text with a red pencil
- Markups are syntactically distinguishable from the actual text (content)
- Different Types of markup languages
  - presentational … ie. WYSIWYG editors
  - procedural … processing instructions
  - descriptive … semantics of the text

SGML

- Direct ancestor of Scribe (1980, Brian Reid)
  - Scribe was based on a grammar
- Structural description of a document
  - instead of a presentation, ie. style is defined separately
- Long-term valid, machine-readable documents
- SGML ...
  - is a rooted acyclic directed graph.
  - supports document type declarations

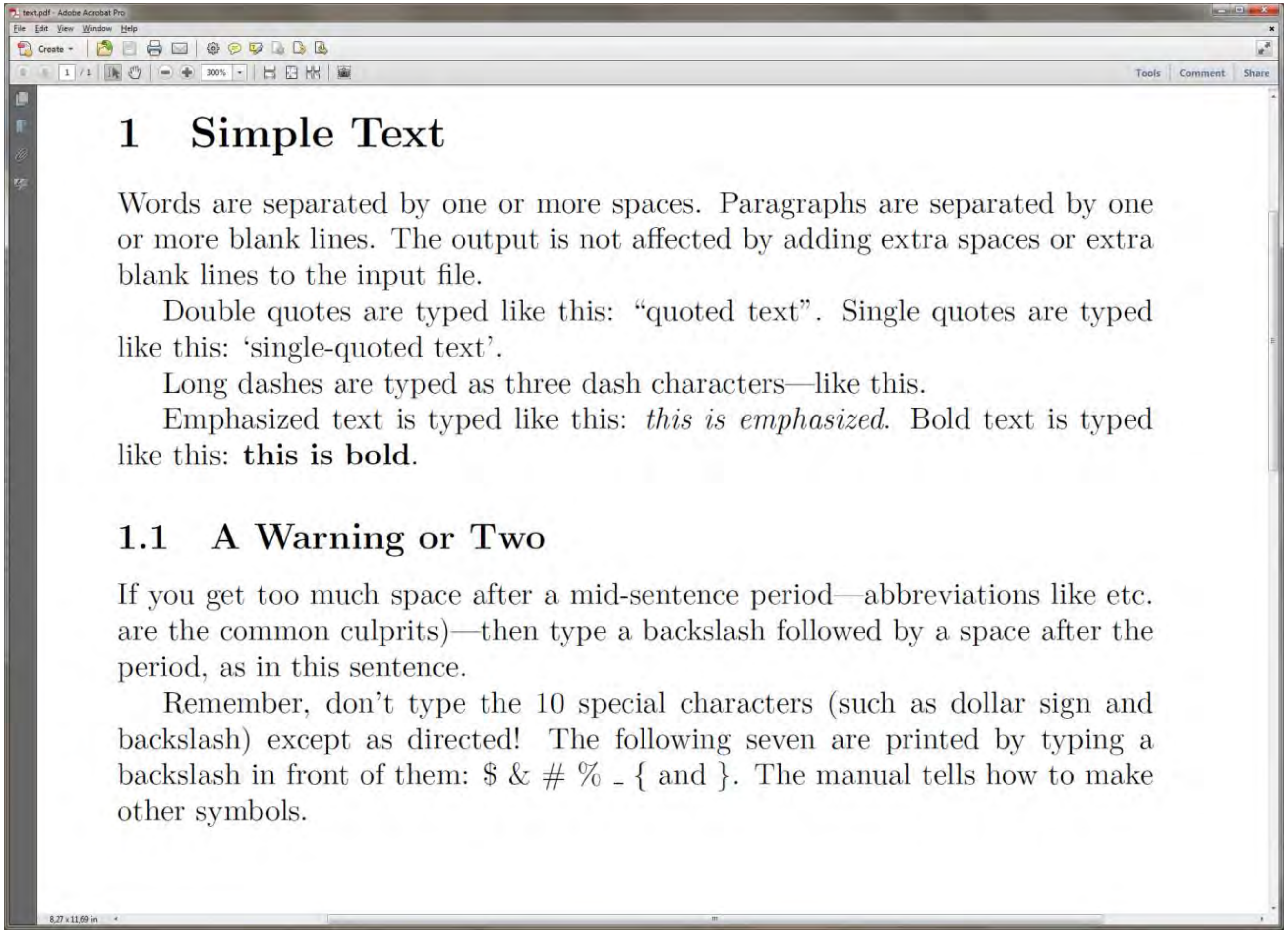
TeX

- Typesetting system widely spread in academia
  - Released in 1978 by Donald Knuth
- TeX is free software
  - available on Linux distributions, Mac and Win
  - LaTeX is a popular packaging for TeX
  - MikTeX is a popular collection of tools for Windows (http://miktex.org/)

Sample LaTeX Document

```latex
\documentclass{article}
\begin{document}
\section{Simple Text}

Words are separated by one or more spaces.  Paragraphs are separated by one or more blank lines.
The output is not affected by adding extra spaces or extra blank lines to the input file.

Double quotes are typed like this: ``quoted text''.
Single quotes are typed like this: `single-quoted text'.

Long dashes are typed as three dash characters---like this.

Emphasized text is typed like this: \emph{this is emphasized}.
Bold       text is typed like this: \textbf{this is bold}.

\subsection{A Warning or Two}

If you get too much space after a mid-sentence period---abbreviations like etc.\ are the common
culprits)---then type a backslash followed by a space after the period, as in this sentence.

Remember, don't type the 10 special characters (such as dollar sign and backslash) except as directed!
The following seven are printed by typing a backslash in front of them:  \$  \&  \#  \%  \_  \{  and
\}. The manual tells how to make other symbols.

\end{document}
```

Sample LaTeX Document



1 Simple Text

Words are separated by one or more spaces. Paragraphs are separated by one or more blank lines. The output is not affected by adding extra spaces or extra blank lines to the input file.

Double quotes are typed like this: “quoted text”. Single quotes are typed like this: ‘single-quoted text’.

Long dashes are typed as three dash characters—like this.

Emphasized text is typed like this: *this is emphasized*. Bold text is typed like this: **this is bold**.

1.1 A Warning or Two

If you get too much space after a mid-sentence period—abbreviations like etc. are the common culprits)—then type a backslash followed by a space after the period, as in this sentence.

Remember, don't type the 10 special characters (such as dollar sign and backslash) except as directed! The following seven are printed by typing a backslash in front of them: $ & # % _ { and }. The manual tells how to make other symbols.

XML

- Successor of SGML
  - less complex, easier to read (for humans), easier to parse (for machines)
- Standardized by W3C
  - Builds a basis for many standards
  - SVG, XHTML, SMIL, MPEG-7, ...

XML

- Based on a tree model & Unicode
  - Supports document type declarations, etc.
  - Documents can be strictly defined by XML Schema
- Main components
  - root element, child elements, both with attributes
  - text in between elements

XML

- Two promising parsing models
- Document Object Model – DOM
  - the XML document is treated as a tree data structure
- Simple API for XML - SAX
  - event based model for parsing
  - sample events: document begins, element begins, etc.
  - no need to store document in-memory

XML Benefits & Drawbacks

- XML is a tree structure
  - simple to maintain, but complicated to use for graph data
- XML cannot be streamed
  - only the complete document can be validated#
- XML brings serious overhead
  - markups, DOM data structure, etc.

XML Examples

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<俄语>данные</俄语>
```

```xml
<?xml version="1.0" encoding="ISO-8859-1"?>
<note>
<to>Tove</to>
<from>Jani</from>
<heading>Reminder</heading>
<body>Don't forget me this weekend!</body>
</note>
```

Lightweight Markup Languages

- Simple (non complex) markup language
- Easy to read and write for humans
  - Even with a simple text editor
- Used for
  - immediate editing, short texts
  - simple styled documents with clear guidelines

Bulletin Board Code - BBCode

- Used for many message boards
- Examples:
  - [b]bold text[/b]
  - [i]italicized text[/i]
  - [u]underlined text[/u]
  - [s]strikethrough text[/s]
  - [url]http://example.org[/url]
  - [url=http://example.com]Example[/url]
  - [img]http://www.cnn.com/test.png[/img]
  - [quote]quoted text[/quote]
  - [code]monospaced text[/code]

Wikitext

- Simple markup with bi-directional links
- Wiki Link: [[Another Site]]
  - link is resolved by wiki engine
  - target site is notified or created (implicitely)
- Wiki formatting
  - Different for different wiki engines, ie. Mediawiki, Dokuwiki, …
  - eg. http://en.wikipedia.org/wiki/Help:Cheatsheet

Creole

- Light weight markup language

```
//emphasized// (e.g., italics)

**strongly emphasized** (e.g., bold)

* Bullet list
* Second item
** Sub item

# Numbered list
# Second item
## Sub item
```

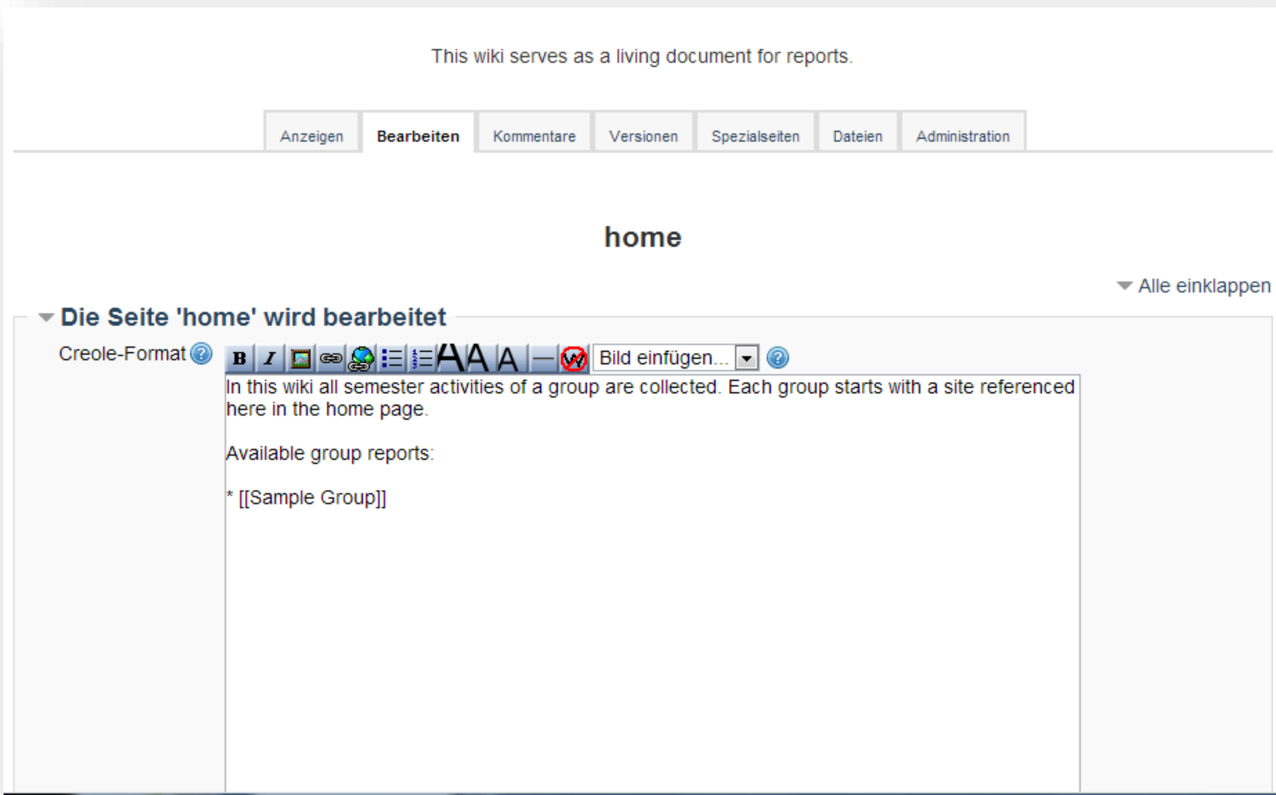
See http://en.wikipedia.org/wiki/Creole_(markup)

Creole

```
Link to [[wikipage]], [[link_address|link text]]

= Extra-large heading (closing optional)
== Large heading
=== Medium heading
==== Small heading

Force\\linebreak

---- (horizontal line)

{{Image.jpg|title}}

|= |= table |= header |
| a | table | row |
| b | table | row

{{{
This text will //not// be **formatted**.
}}} |
```

Creole - Demo

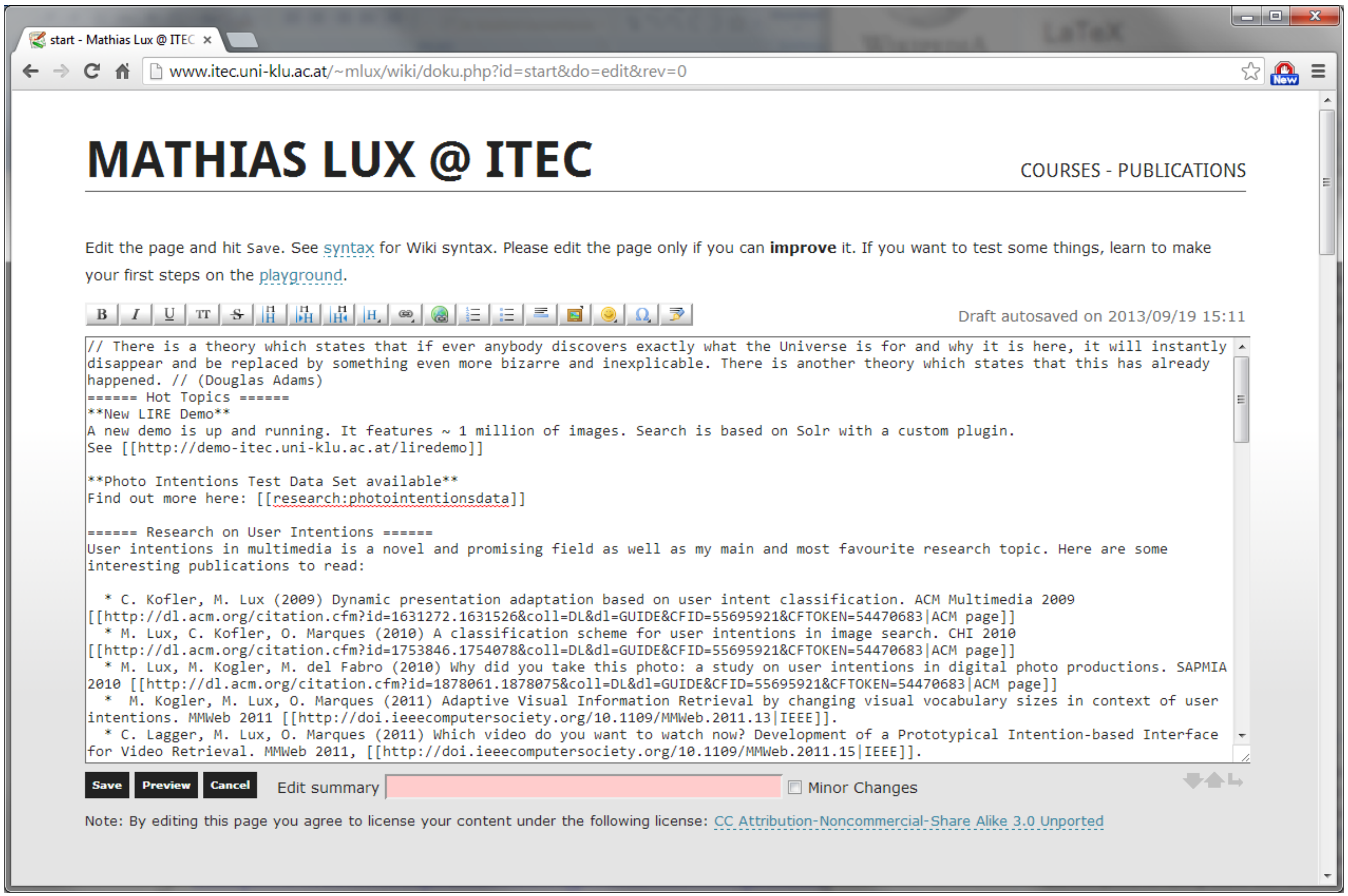
- Course Moodle ...



In-Course Exercise: Create your group page

- Create you group page in the moodle.
- Add an image showing the group members.
- Create a sub page.

Dokuwiki Demo



Markdown

- Very simple text format
- Can be rendered to HTML, etc.
  - text markup itself is much like ASCII formatting
- Renderers are available in several languages

Sample Markdown Document

```
LIRE Solr Integration Project
=============================
Includes a RequestHandler and some utility
classes for a fast start.

The request handler supports four different
types of queries

1. Get random images ...
2. Get images that are looking like the one with id ...
3. Get images looking like the one found at url ...
4. Get images with a feature vector like ...

Preliminaries
-------------

Supported values for feature field parameters,
e.g. lireq?field=cl_ha:

-  **cl_ha** .. ColorLayout
-  **ph_ha** .. PHOG
-  **oh_ha** .. OpponentHistogram
-  **eh_ha** .. EdgeHistogram
-  **jc_ha** .. JCD
```

LIRE Solr Integration Project

Includes a RequestHandler and some utility classes for a fast start.

The request handler supports four different types of queries

1. Get random images ...
2. Get images that are looking like the one with id ...
3. Get images looking like the one found at url ...
4. Get images with a feature vector like ...

Preliminaries

Supported values for feature field parameters, e.g. lireq?field=cl_ha:

- **cl_ha** .. ColorLayout
- **ph_ha** .. PHOG
- **oh_ha** .. OpponentHistogram
- **eh_ha** .. EdgeHistogram
- **jc_ha** .. JCD

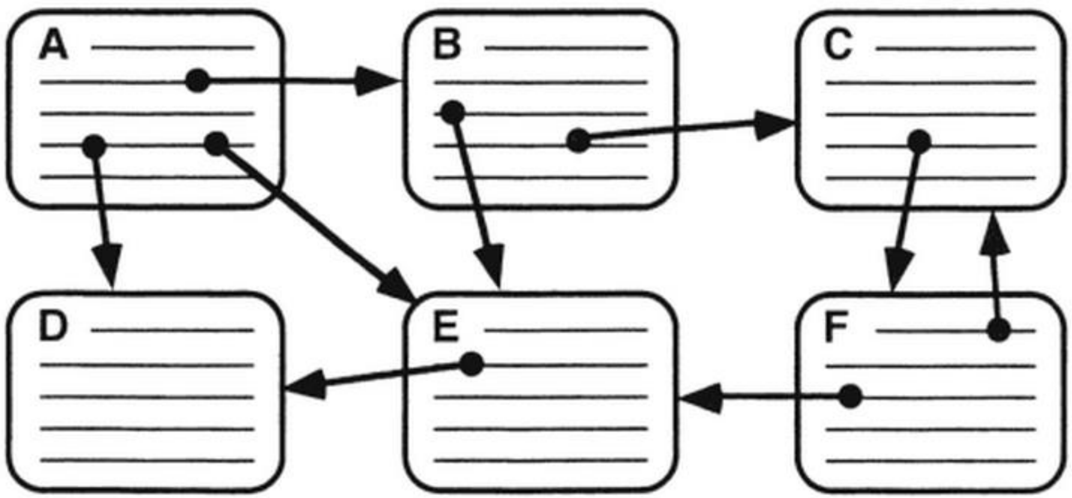
HTML & XHTML

- Prominent markup languages for the web
- HTML is based on SGML
- XHTML is based on XML

- Both are used for creating hypertext systems

Hypertext – A Definition

- Text is printed & consumed sequentially
- Hypertext is non-sequential



Src. Nielsen J (1995) Multimedia and Hypertext: The Internet and Beyond, Morgan & Kaufmann

Hypertext: A Definition

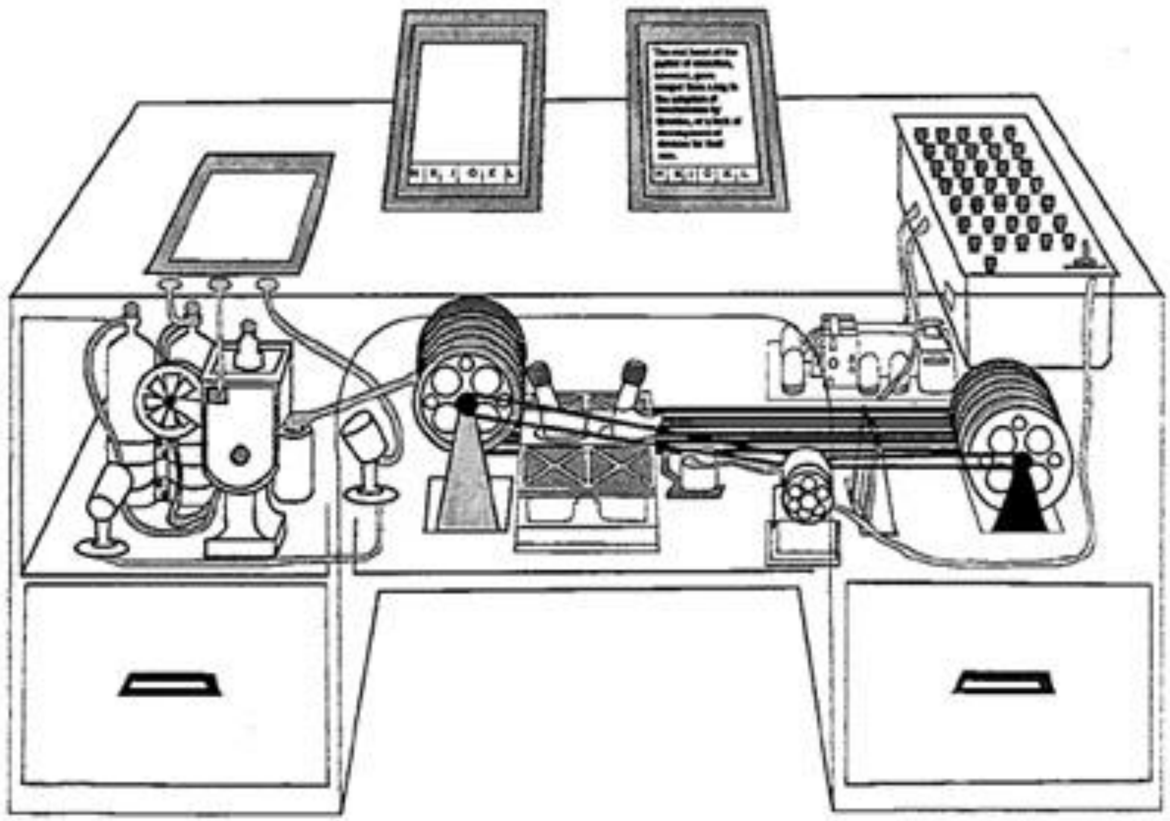
- Text has some “hypertext” elements
  - footnotes, indexes, (cross-)references, glossaries
- Hypertext is more general
  - hypertext nodes are connected by links
  - links can be bi- or unidirectional
  - links can be “typed”
    - footnote, related, include, …
- Readers of hypertext traverse links

MEMEX

Memory **Ex**tender – Vannevar Bush

- Published in 1945 (Atlantic Monthly)
- An electromechanical device for
  - viewing books and films
  - adding information and comments
  - interlinking information
  - browsing links
- MEMEX is an early hypertext system

Memex



Uniform Resource Locators (URLs)

A URL consists of

- the scheme name (protocol)
- a colon and two slashes
- a host (domain name or IP address)
- a port number (optional)
- full path of the resource

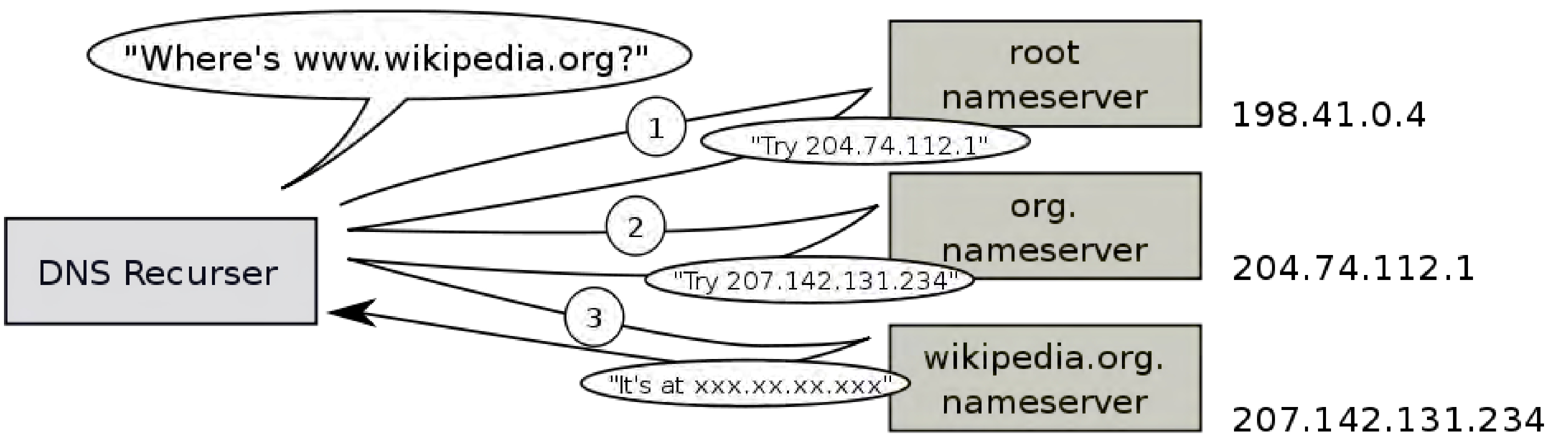
`https://code.google.com/p/lire/`

Domain Names

- A domain name is a string that defines a realm of authority
  – cnn.com, öamtc.at, uni-klu.ac.at, …
- A top level domain defines country and/or intent of the domain
  – .com, .ac.at, …
- A sub domain points to a specific IP address
  – www.aau.at, ftp.uni-klu.ac.at

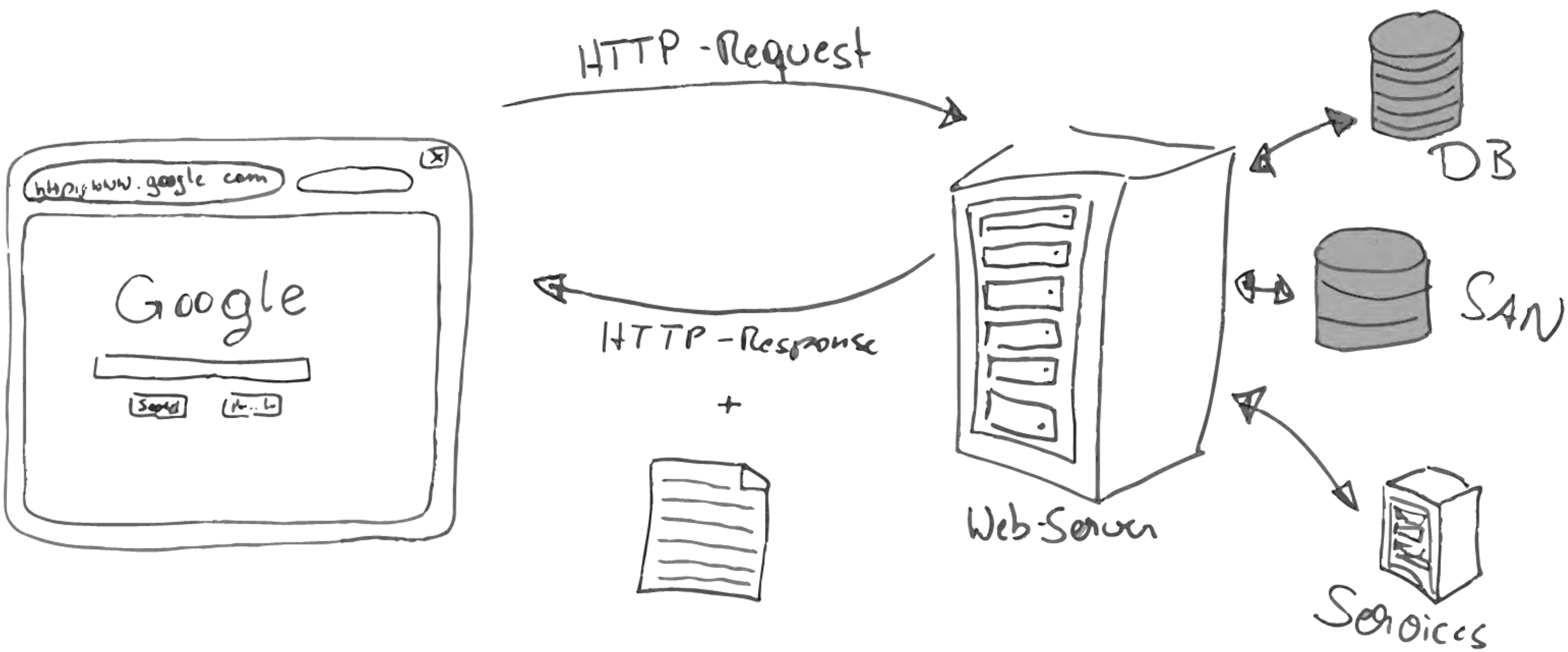
Domain Name System

- Hierarchical system where domain names are "registered"



Src. http://en.wikipedia.org/wiki/File:An_example_of_theoretical_DNS_recursion.svg

WWW



HTTP -Request
http://www.google.com
Google
HTTP -Response
+
Web-Server
DB
SAN
Services

WWW – Die Maus (WDR)

- http://www.youtube.com/watch?v=QKLz4ufCuKk

HTTP

Retrieve http://www.somehost.com/path/file.html

- Open socket to www.somehost.com:80
- Send something like this:

```
GET /path/file.html
HTTP/1.0
From: someuser@jmarshall.com
User-Agent: HTTPTool/1.0
[blank line here]
```

Src. http://www.jmarshall.com/easy/http/

HTTP Response

- 200 OK
  - The request succeeded, resource is returned in the message body.
- 404 Not Found
  - The resource doesn't exist.
- 302 Moved Temporarily
  - Used to redirect.
- 500 Server Error
  - An unexpected server error.

Sample Response

```
HTTP/1.0 200 OK
Date: Fri, 31 Dec 1999 23:59:59 GMT
Content-Type: text/html
Content-Length: 1354

<html>
  <body>
    <h1>Happy New Millennium!</h1>
    (more file contents) . . .
  </body>
</html>
```

Src. http://www.jmarshall.com/easy/http/

HTTP – Some more facts

- A web page typically is more than one file
- Browsers try to re-use HTTP connections
  - Ask for multiple files in one single connection
  - „keep alive"
- HTTP is based on TCP & defaults to port 80
  - Works well behind firewalls
- Other services try to use HTTP too
  - HTTP video streaming, etc.

HTML Document Structure

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML
  4.01 Transitional//EN"
  "http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
  <title>
    Beschreibung der Seite
  </title>
</head>
<body>
</body>
</html>
```

HTML Document Structure

- Document type Declaration
- HTML
  - Head
    - Title
  - Body
    - . . . .

Introduction to Media Informatics: Information Retrieval

Dr. Mathias Lux
Alpen-Adria Universität Klagenfurt

Information Retrieval History

Currently there are no museums for IR

IR is the process of **searching** through a **document collection** based on a particular **information need**.

IR Key Concepts

- Searching
  - Indexing, Ranking

- Document Collection
  - Textual, Visual, Auditive

- Particular Needs
  - Query, User based

A History of Libraries

Libraries are perfect examples for document collections.

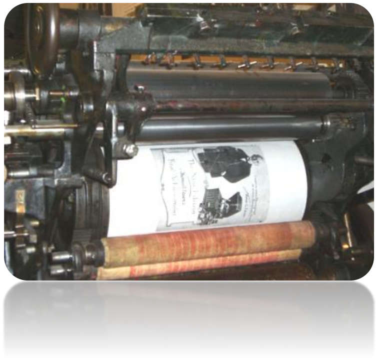
- Wall paintings in caves
  - e.g. Altamira, ~ 18,500 years old
- Writing in clay, stone, bones
  - e.g. Mesopotamian cuneiforms, ~ 4.000 BC
  - e.g. Chinese tortoise-shell carvings, ~ 6.000 BC
  - e.g. Hieroglyphic inscriptions, Narmer Palette ~ 3.200 BC

A History of Libraries (ctd.)

- Papyrus
  - Specific plant (subtropical)
  - Organized in rolls, e.g. in Alexandria
- Parchment
  - Independence from papyrus
  - Sewed together in books
- Paper
  - Invented in China (bones and bamboo too heavy, silk too expensive)
  - Invention spread -> in 1120 first paper mill in Europe

A History of Libraries (ctd.)

- Gutenberg's printing press (1454)
  - Inexpensive reproduction
  - e.g. "Gutenberg Bible"
- Organization & Storage
  - Dewey Decimal System (DDC, 1872)
  - Card Catalog (early 1900s)
  - Microfilm (1930s)
  - MARC (Machine Readable Cataloging, 1960s)
  - Digital computers (1940s+)



Library & Archives today

- Partially converted to electronic catalogues
  - From a certain time point on (1992 - …)
  - Often based on proprietary systems
  - Digitization happens slow
  - No full text search available
  - Problems with preservation
    - Storage devices & formats

History of Searching

- Browsing
  - Like "Finding information yourself"
- Catalogs
  - Organized in taxonomies, keywords, etc.
- Content Based Searching
  - `SELECT * FROM books WHERE title='%Search%'`
- Information Retrieval
  - Ranking, models, weighting
  - Link analysis, LSA, …

History of IR

- Starts with development of computers
- Term "Information Retrieval" coined by Mooers in 1950
  - Mooers, C. (March 1950). "The theory of digital handling of non-numerical information and its implications to machine economics". *Proceedings of the meeting of the Association for Computing Machinery at Rutgers University*.

- Two main periods (Spark Jones u. Willett)
  - 1955 – 1975: Academic research
    - Models and Basics
    - Main Topics: Search & Indexing
  - 1975 – … : Commercial applications
    - Improvement of basic methods

A Challenge: The World Wide Web

- First actual implementation of **Hypertext**
  - Interconnected documents
  - Linked and referenced
- World Wide Web (1989, T. Berners-Lee)
  - Unidirectional links (target is not aware)
  - Links are not typed
  - Simple document format & communication protocol (HTML & HTTP)
  - Distributed and not controlled

Some IR History Milestones

- Book "Automatic Information Organization and Retrieval", *Gerard Salton* (1968)
  - Vector Space Model
- Paper "A statistical interpretation of term specificity and its application in retrieval", Karen Sparck Jones (1972)
  - IDF weighting
  - http://www.soi.city.ac.uk/~ser/idf.html
- Book "Information Retrieval" of *C.J. Rijsbergen* (1975)
  - Probabilistic Model
  - http://www.dcs.gla.ac.uk/Keith/Preface.html

Some IR History Milestones

- Paper "Indexing by Latent Semantic Analysis", S. Deerwester, Susan Dumais, G. W. Furnas, T. K. Landauer, R. Harshman (1990).
  - Latent Semantic Indexing
- Paper "Some simple effective approximations to the 2-Poisson model for probabilistic weighted retrieval" Robertsen & Walker (1994)
  - BM25 weighting scheme
- Paper "The Anatomy of a Large-Scale Hypertextual Web Search Engine", *Sergey Brin & Larry Page* (1998)
  - World Wide Web Retrieval

Agenda

- Information Retrieval History
- **Information Retrieval & Data Retrieval**
- Searching & Browsing
- Information Retrieval Models
- Web Retrieval

Organizational: References

- in the Library
  - *Modern Information Retrieval*, Ricardo Baeza-Yates & Berthier Ribeiro-Neto, Addison Wesley
  - *Google's Pagerank and Beyond: The Science of Search Engine Rankings*, Amy N. Langville & Carl D. Meyer, University Presses of CA
  - *Distributed Multimedia Database Technologies supported by MPEG-7 and MPEG-21*, Harald Kosch, CRC Press
  - *Readings in Information Retrieval*, Karen Sparck Jones, Peter Willett, Morgan Kaufmann

Organizational: References

- WWW
  - *Skriptum Information Retrieval*, Norbert Fuhr, Lecture Notes on Information Retrieval - Univ. Dortmund, 1996. Updated in 2002
  - *Information Retrieval 2nd Edt.*, C.J. Rijsbergen, Butterworth, London 1979
- Through me:
  - *Lectures on Information Retrieval: Third European Summer-School, Essir 2000 Varenna, Italy, Revised Lectures*, Maristella Agosti, Fabio Crestani & Gabriela Pasi (eds.), Lecture Notes in Computer Science, Springer 2000

Information Retrieval & Data Retrieval

Information Retrieval

- Information Level
- Search Engine
- Bing / Google

Data Retrieval

- Data Level
- Data Base
- Oracle / MySQL

Information Retrieval & Data Retrieval

Information Retrieval	Data Retrieval
Content Based Search	Search for Patterns and String
Query ambigous	Query formal & unambigous
Results ranked by relevance	Results not ranked
Error tolerant	Not error tolerant
Multiple iterations	Clearly defined result set
Examples	*Examples*
Search for synonyms	Search for patterns
Bag of Words	SQL Statement

- Retrieval is nearly always a combination of both.

Agenda

- Information Retrieval History
- Information Retrieval & Data Retrieval
- **Searching & Browsing**
- Information Retrieval Models
- Web Retrieval

Information Retrieval Basics: Searching

A **user** has an **information need**, which needs to be **satisfied**.

- Two different approaches:
  - Browsing
  - Searching

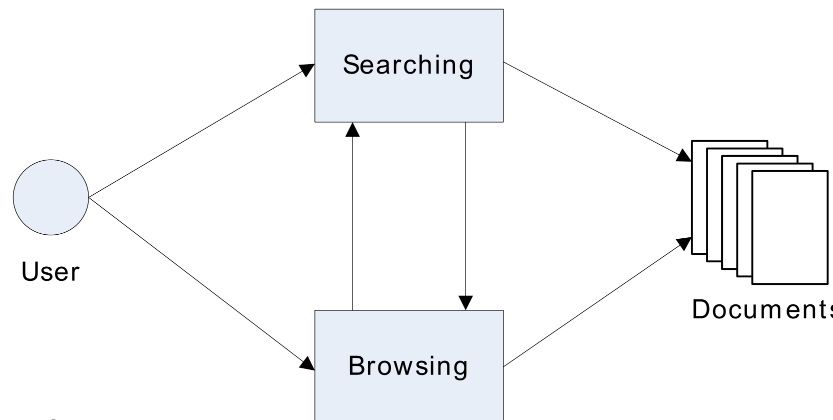
Searching & Browsing

Searching

- Explicit information need
- Definition through "query"
- Result lists
- e.g. Google

Browsing

- Not necessarily explicit need
- Navigation through repositories



Browsing

- Flat Browsing
  - User navigates through set of documents
  - No implied ordering, explicit ordering possible
  - Examples: One single directory, one single file
- Structure Guided Browsing
  - An explicit structure is available for navigation
  - Mostly hierarchical (file directories)
  - Can be generic digraph (WWW)
  - Examples: File systems, World Wide Web

Searching

- Query defines "Information Need"

- Ad Hoc Searching
  - Search when you need it
  - Query is created to fit the need
- Information Filtering
  - Make sets of documents smaller
  - Query is filter criterion
- Information Push
  - Same as filtering, delivery is different

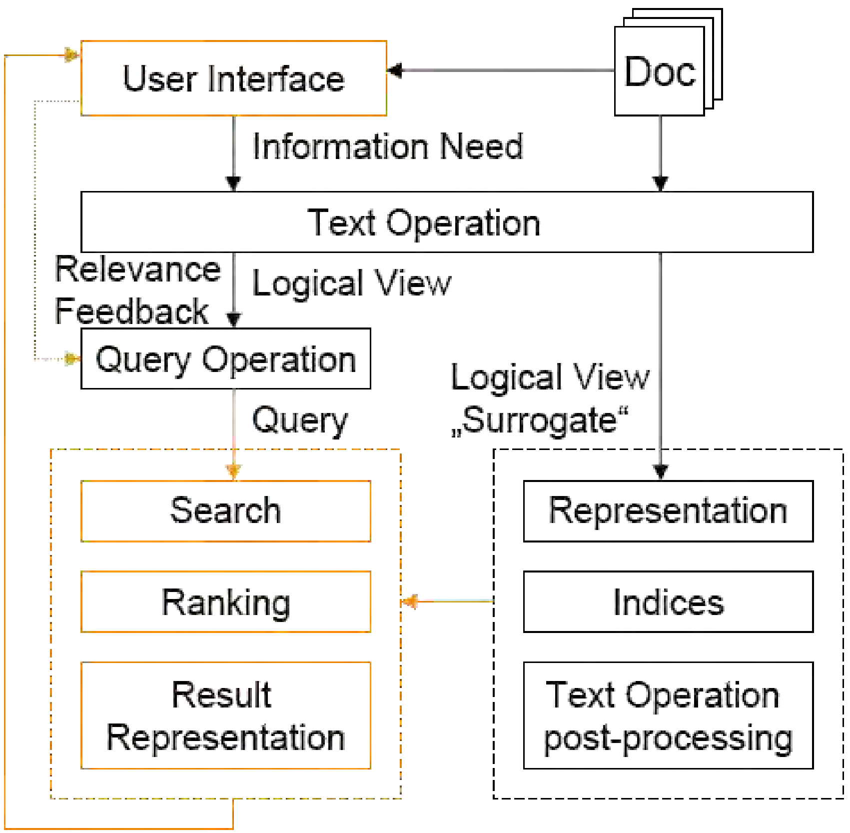
Agenda

- Information Retrieval History
- Information Retrieval & Data Retrieval
- Searching & Browsing
- **Information Retrieval Models**
- Web Retrieval

Information Retrieval System Architecture

Aspects

- Query & languages
- IR models
- Documents
- Internal representation
- Pre- and post-processing
- Relevance feedback
- HCI



Information Retrieval Models

- Boolean Model
  - Set theory & Boolean algebra
- Vector Model
  - Non binary weights on dimensions
  - Partial match
- Probabilistic Model
  - Modeling IR in a probabilistic framework

Formal Definition of Models

An information retrieval model is a quadruple $[D, Q, F, R(q_i, d_j)]$

- D is a set of logical views (or representations) for the **documents** in the collection.
- Q is a set of logical views (or representations) for the user needs or **queries**.
- F is a **framework** for modeling document representations, queries and their relationship.
- $R(q_i, d_j)$ is a **ranking function** which associates a real number with a query q_i of Q and a document d_j of D.

Definitions
in Context of Text Retrieval

- **index term** – word of a document expressing (part of) document semantics
- **weight $w_{i,j}$** – quantifies the importance of index term t_i for document d_j
- **index term vector** for document d_j (having t different terms in all documents):

$$\overrightarrow{d_j} = (w_{1,j}, w_{2,j}, \ldots, w_{t,j})$$

Boolean Model

- Based on set theory and Boolean algebra
  - Set of index terms
  - Query is Boolean expression
- Intuitive concept:
  - Wide usage in bibliographic system
  - Easy implementation and simple formalisms
- Drawbacks:
  - Binary decision components (true/false)
  - No relevance scale (relevant or not)

Boolean Model: Example

- Example queries
  - cat OR dog
  - cat AND dog
  - lecture AND (multimedia OR media AND informatics)

Boolean Model

- Advantages
  - Clean formalisms
  - Simplicity
- Disadvantages
  - Might lead to too few / many results
  - No notion of **partial match**
  - Sequential ordering of terms not taken into account.

Vector Model

- Integrates the notion of partial match
- Non-binary weights (terms & queries)
- Degree of similarity computed

$$\vec{d}_j = (w_{1,j}, w_{2,j}, \ldots, w_{t,j})$$

$$\vec{q} = (w_{1,q}, w_{2,q}, \ldots, w_{t,q})$$

Vector Model: Example

$$sim(d_j, q) = \frac{\vec{d}_j \bullet \vec{q}}{\left|\vec{d}_j\right| \times \left|\vec{q}\right|} = \frac{\sum_{i=1}^{t} w_{i,j} \cdot w_{i,q}}{\sqrt{\sum_{i=1}^{t} w_{i,j}^2} \cdot \sqrt{\sum_{i=1}^{t} w_{i,q}^2}}$$

$$\vec{d} = (0.3, 0.1, 0, 0.1, 1)$$

$$\vec{q} = (1, 0, 0, 0.5, 0)$$

$$sim(\vec{d}, \vec{q}) = \frac{1 \cdot 0.3 + 0.1 \cdot 0.5}{\sqrt{0.3^2 + 0.1^2 + 0.1^2 + 1^2} \cdot \sqrt{1 + 0.5^2}} \approx \frac{0.35}{2.24} \approx 0.17$$

Another Example

- Document & Query:
  - D = "The quick brown fox jumps over the lazy dog"
  - Q = "brown lazy fox"

$$sim(d_j, q) = \frac{\vec{d}_j \bullet \vec{q}}{\left|\vec{d}_j\right| \times \left|\vec{q}\right|} = \frac{\sum_{i=1}^{t} w_{i,j} \cdot w_{i,q}}{\sqrt{\sum_{i=1}^{t} w_{i,j}^2} \cdot \sqrt{\sum_{i=1}^{t} w_{i,q}^2}}$$

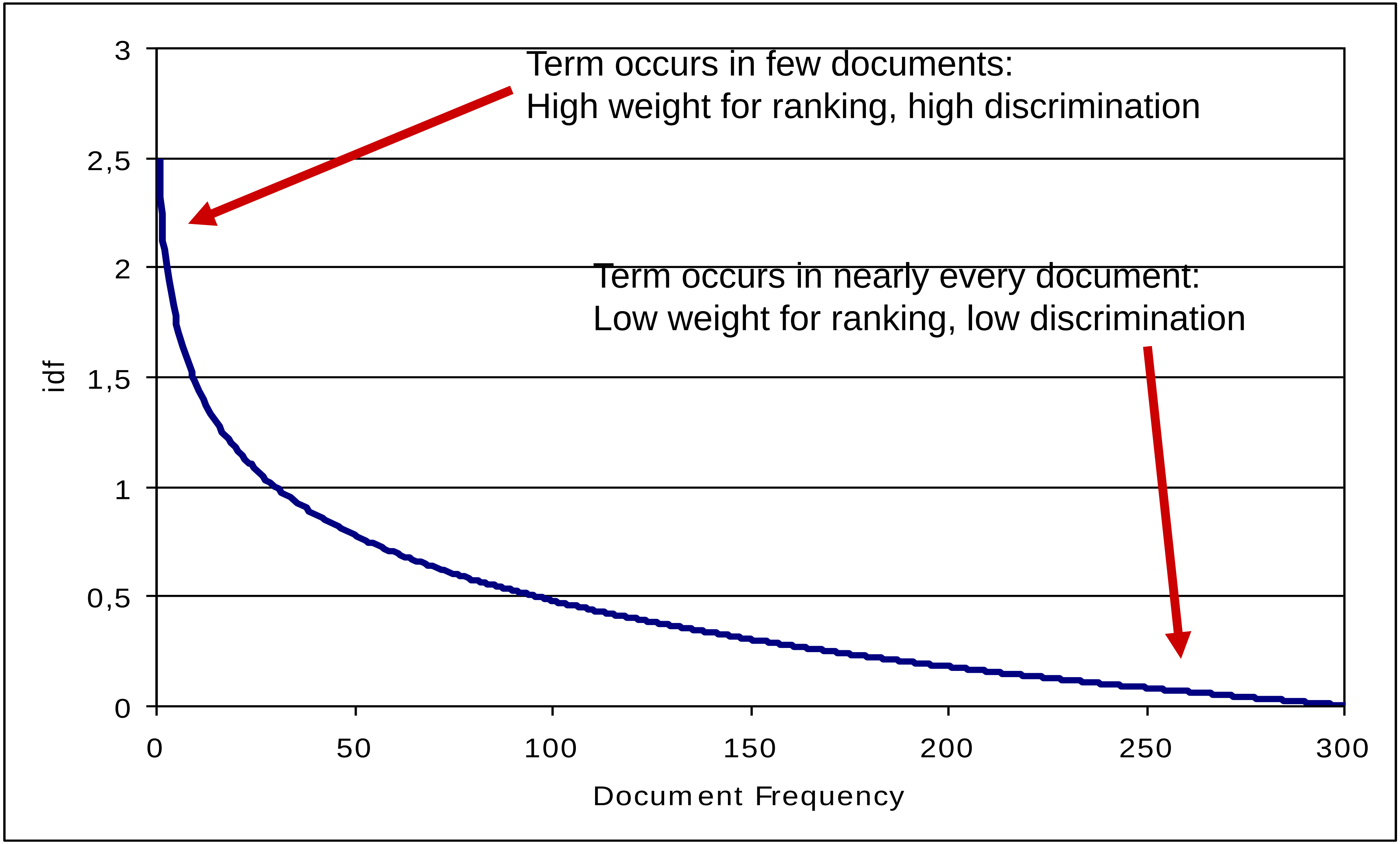
- Results:
  - $(1,1,1,1,1,1,1,2)^t * (1,1,1,0,0,0,0,0)^t = 3$
  - sqrt(11) * sqrt(3) = 5.04
  - Similarity = 3 / 6 = 0.595

Term weighting: TF*IDF

Term weighting increases retrieval performance

- Term frequency
  - How often does a term occur in a document?
  - Most intuitive approach
- Inverse Document Frequency
  - What is the information content of a term for a document collection?
  - Compare to *Information Theory* of Shannon

Example: IDF with 300 documents corpus



Definitions: Normalized Term Frequency

$$f_{i,j} = \frac{freq_{i,j}}{\max_l(freq_{l,j})} \text{ ... normalized term frequency}$$

$$freq_{i,j} \text{ ... raw term frequency of term } i \text{ in document } j$$

- Maximum is computed over all terms in a document
- Terms which are not present in a document have a raw frequency of 0

Definitions: Inverse Document Frequency

$$idf_i = \log \frac{N}{n_i}$$ ... inverse document frequency for term i

N ... number of documents in the corpus

n_i ... number of document in the corpus which contain term i

- Note that idf_i is independent from the document.
- Note that the whole corpus has to be taken into account.

TF*IDF

- TF*IDF is a very prominent weighting scheme
  - Works fine, much better than TF or Boolean
  - Quite easy to implement

$$w_{i,j} = f_{i,j} \cdot \log \frac{N}{n_i}$$

Vector Model

- Advantages
  - Weighting schemes improve **retrieval performance**
  - Partial matching allows retrieving documents that **approximate query** conditions
  - Cosine coefficient allows **ranked list** output
- Disadvantages
  - Term are assumed to be mutually independent

Simple example (I)

- Scenario
  - Given a **document corpus on birds**: nearly each document (say 99%) contains the word bird
  - someone is searching for a document about sparrow nest construction with a query "**sparrow bird nest construction**"
  - Exactly the document which would satisfy the user needs **does not have the word "bird"** in it.

Simple example (II)

- TF*IDF weighting
  - knows upon the low discrimative power of the term bird
  - The weight of this term is near to zero
  - This term has virtually no influence on the result list.

Agenda

- Information Retrieval History
- Information Retrieval & Data Retrieval
- Searching & Browsing
- Information Retrieval Models
- **Web Retrieval**

Retrieval in the WWW

- General Retrieval is based on content
  - Represented e.g. by terms, keywords …
- What is different with the WWW?
  - Structured text (markup)
  - Hypermedia (links)
  - Heterogeneous formats (gif, pdf, flv, …)
  - Distributed content (access over network)

Web based Retrieval: Challenges

- Working with an enormous amount of data
  - 10 billion pages a 500kB estimated in 01-2004
    - 2 pages / person on the globe
  - 1 trillion unique URLs indexed by Google in 2008
  - 109.5 million top-level domains operated in 2009
- Furthermore there is a **Deep Web**
  - Including the usenet, tor, torrent, non-indexed WWW, ftp, …

Web based Retrieval: Challenges

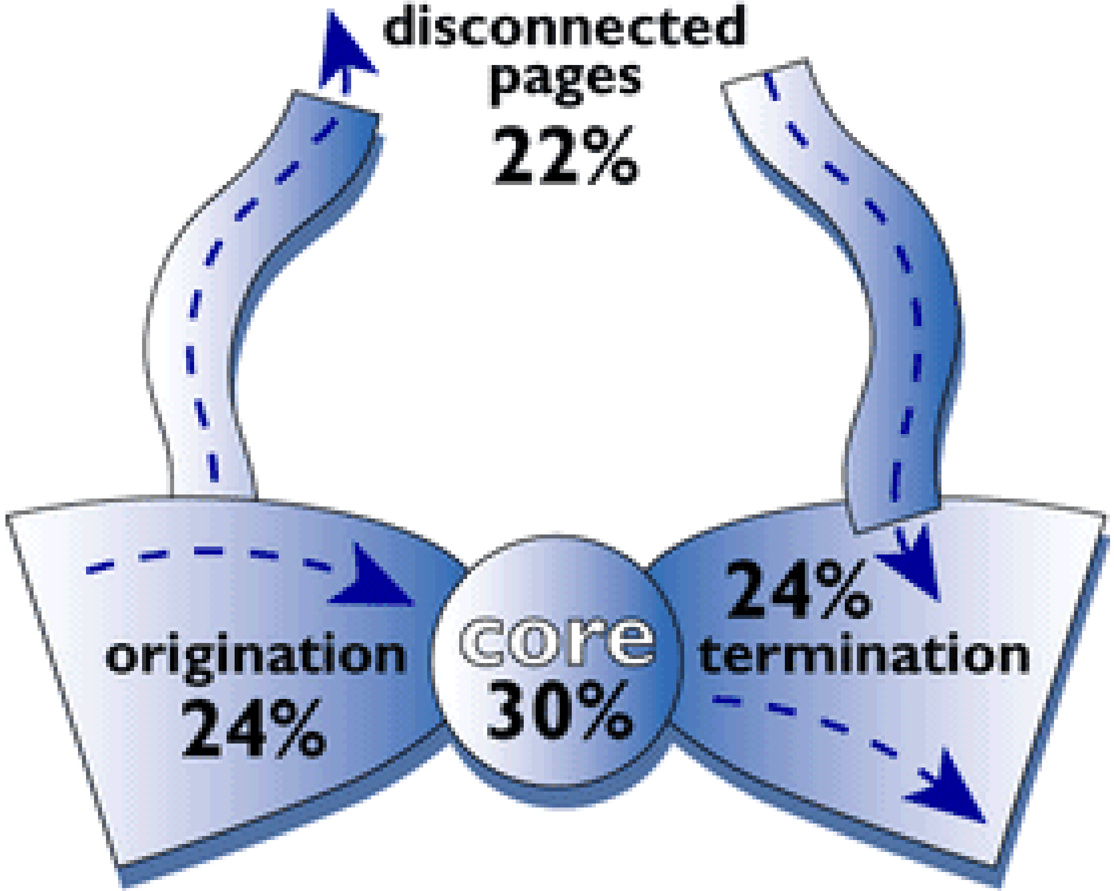
- Example for the amount of web pages:
  - Searching for 'Enterprise' yielded on Google ~ 435 millions of results
  - Users investigate up to 20 result list entries.
- What web page is the **most interesting**?
  - cp. Concept of relevance (IR)
- How to **index** this amount of pages?
  - eg. in an inverted list

Web based Retrieval: Challenges

The Web is self-organized

- No central authority / main index
  - For the WWW
- Everyone can add (or edit) pages
  - Cp. Personal homepages, blogs, wikis, ...
- Pages disappear on regular basis
  - US study claimed that in 2 investigated tech. journals 50% of the cited links were inaccessible after four years.
- Lots of errors and falsehood, no quality control

WWW – Bow Tie Structure



Ranking by Popularity

- Problem with amount of data
  - Queries on popular terms yield many results
- Idea for selecting the most relevant …
  - Combine content with **popularity** of page
  - More popular pages are "authorities"
- How to define popularity?
  - Only hypertext documents are given …

Popularity Ranking

- 2 Algorithms developed independently
  - PageRank, Brin & Page
  - Hypertext Induced Topic Search (HITS), Kleinberg
- Basic idea of popularity
  - Someone likes a page
  - Gives a recommendation (on another page)
  - Using a hyperlink

Popularity Ranking: Basic Idea

- There are different types of people:
  - Regarding their idea of recommendation
    - People giving a lot of recommendations (links)
    - People giving few recommendations (links)
  - Regarding their state of recommendation
    - Recommended by a lot of people
    - Recommended by few people
- Combinations are possible:
  - Having no recommendation, but recommending a lot, …

Popularity Ranking: Basic Idea

Think of ....

- people as pages
- recommendations as links

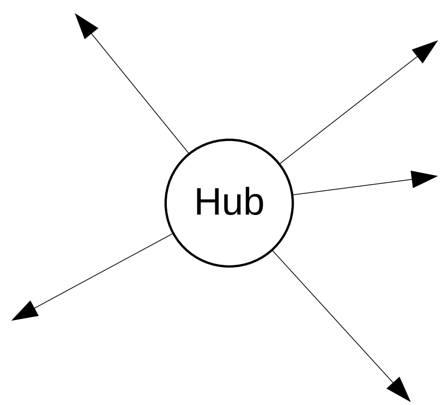
PageRank (Google)

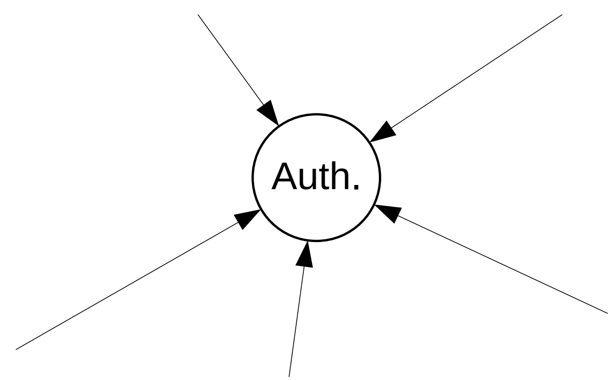
Therefore:

"Pages are popular, if popular pages link them"

Popularity Ranking: Basic Idea

- Additional assumptions:
  - **Hubs** are pages that refer a lot
  - **Authorities** are pages, which are referred a lot





HITS

PageRank: Original Summation Formula

- Original summation formula
  - PageRank of page P_i is given by the summation of all pages that link to P_i given by Set B_{P_i}

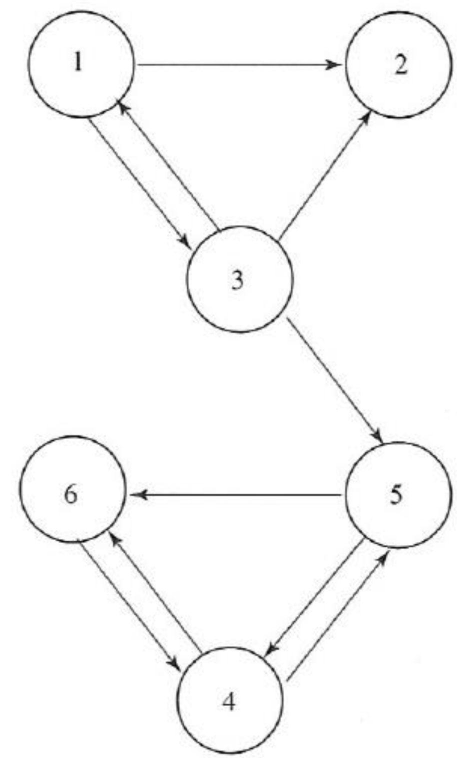
$$r(P_i) = \sum_{P_j \in B_{P_i}} \frac{r(P_j)}{|P_j|},$$

- Iterative formula, starting with rank 1/n for all n pages:

$$r_{k+1}(P_i) = \sum_{P_j \in B_{P_i}} \frac{r_k(P_j)}{|P_j|}$$

PageRank: Original Summation Formula

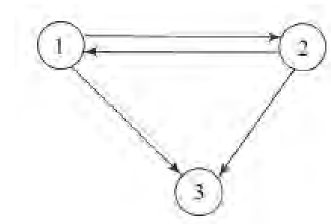
$$r_{k+1}(P_i) = \sum_{P_j \in B_{P_i}} \frac{r_k(P_j)}{|P_j|}$$

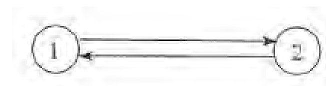


Iteration 0	Iteration 1	Iteration 2	Rank at Iter. 2
$r_0(P_1) = 1/6$	$r_1(P_1) = 1/18$	$r_2(P_1) = 1/36$	5
$r_0(P_2) = 1/6$	$r_1(P_2) = 5/36$	$r_2(P_2) = 1/18$	4
$r_0(P_3) = 1/6$	$r_1(P_3) = 1/12$	$r_2(P_3) = 1/36$	5
$r_0(P_4) = 1/6$	$r_1(P_4) = 1/4$	$r_2(P_4) = 17/72$	1
$r_0(P_5) = 1/6$	$r_1(P_5) = 5/36$	$r_2(P_5) = 11/72$	3
$r_0(P_6) = 1/6$	$r_1(P_6) = 1/6$	$r_2(P_6) = 14/72$	2

Initial Problems

- Rank sinks & cycles:
  - Some pages get all of the score, other pages none
  - Cycles just flip the rank
- How many iterations?
  - Will the process converge?
  - Will it converge to one single vector?





Approach of Brin & Page

- Notion of the random surfer
  - Someone navigates through the web using hyperlinks.
  - If there are 6 links, there is a probability of 1/6 that s/he takes a specific link
  - On dangling nodes (without out links) s/he can jump everywhere with equal chance
  - Furthermore s/he can leave the link path with a given probability every time

Features of PageRank

- Mathematical model
  - Created later on, based on Markov chains
- Can be handled in a distributed way
  - "Worlds biggest matrix multiplication"

HITS

- Every page i has a authority score x_i and a hub score y_i
- Successive refinement of scores:

$$x_i^{(k)} = \sum_{j:e_{ji} \in E} y_j^{(k-1)} \text{ and } y_i^{(k)} = \sum_{j:e_{ji} \in E} x_j^{(k)} \text{ for } k = 1, 2, 3, \ldots$$

Search Engine "Optimization"

- Business for "optimizing" rank in search listings (SEO)
- There are two ways:
  - Ethical: Good content and communication leads to extensive linking and a high content score as well as popularity
  - Unethical: Try to get a lot of links to the site of the customer or lay a *Google Bomb*.

Costs for Web Crawling

- How much does it cost to run a search engine?
  - Monthly amount of pages to crawl: 4 billion
  - 4.000.000.000 pages @ 200K = 80.000 GB per month.
- One connection:
  - 100mbs connection
    / 8 megabits per MB
    * 60 seconds in a minute
    * 60 minutes in an hour
    * 24 hours in a day
    * 30 days in a month=32.400 GB / month

Costs for Web Crawling

- Therefore at least 3 100 MBit connections are needed
  - Running at full capacity 24/7
  - Only with a simple calculation (w/o overhead)
- Also at least 3 servers are needed
- And a lot of storage
  - ~ 80.000 GB with caching

taken from http://www.mail-archive.com/nutch-user@lucene.apache.org/msg05577.html

Introduction to Media Informatics: Images

Dr. Mathias Lux
Alpen-Adria Universität Klagenfurt

Contents

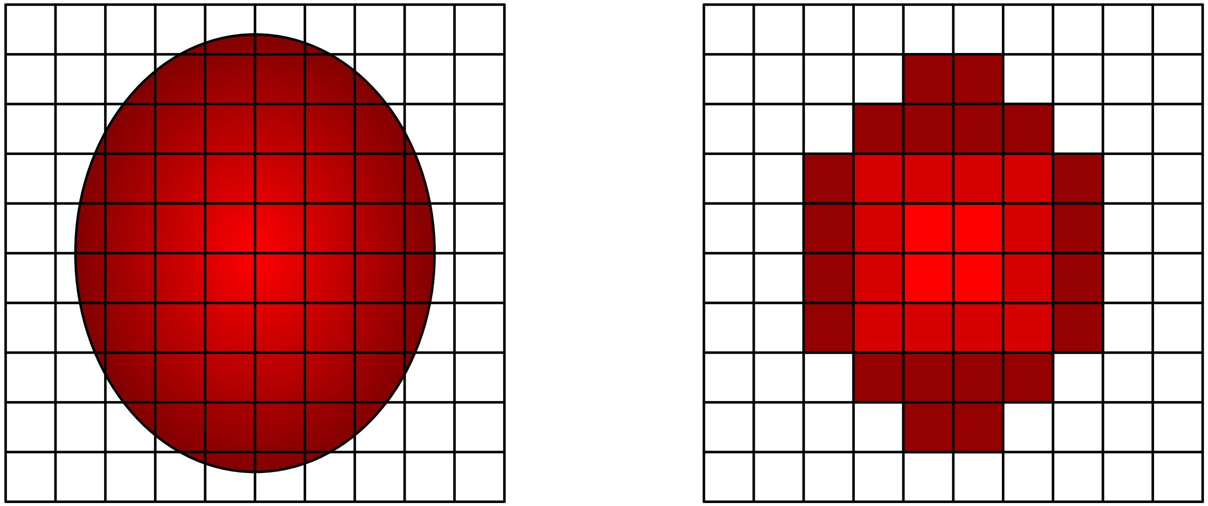
- Color Spaces
- Compression
- Formats
- Filtering

What are (digital) images?

- An Image is
  - Created by a set of photons
    - With different frequency
    - Moving from different sources
    - Along different vectors
  - A representation of sensor unit activation
    - Activated by the set of photons
- Storing an image
  - Based on the set of photons ???

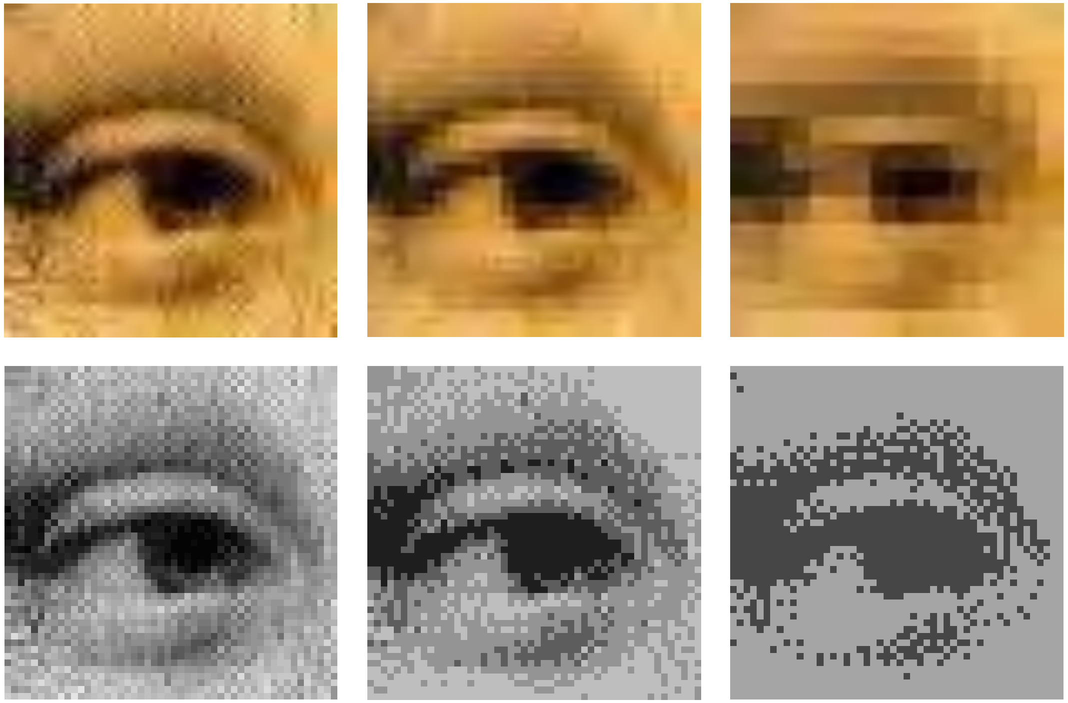
Sampling & Quantization

- Capturing continuous images on sensors
  - Sampling: Continuous to matrix
  - Quantization: Continuous color to value



Sampling & Quantization

- Size of a captured image:
  - # of samples (width*heigth) * # of colors

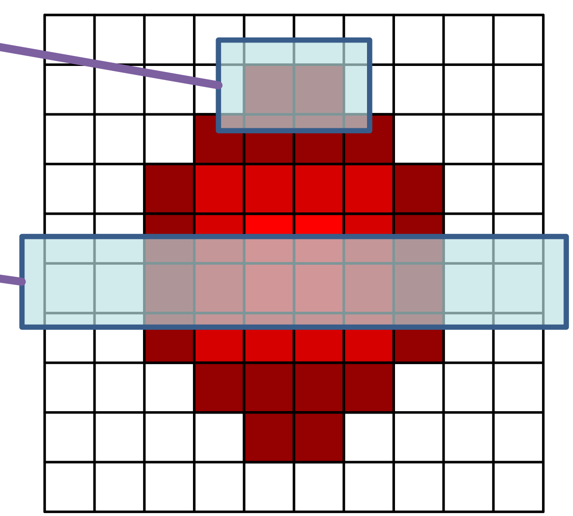


Storing digital images

- Defining color per cell
- Left to right & top to bottom

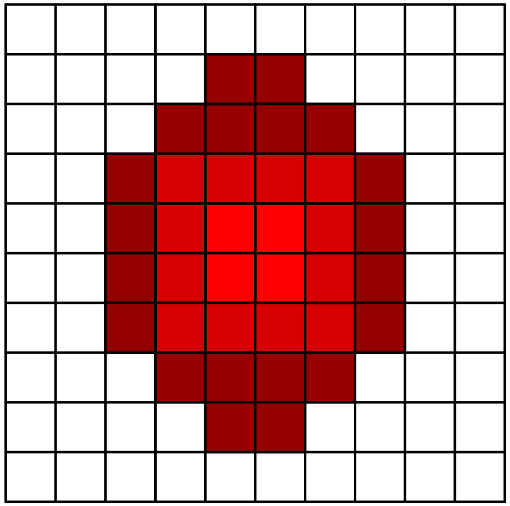
- 0, 0, 0, … 0, 1, 1, 0 …

- … 0, 0, 1, 2, 3, 3, 2, 1, 0, 0, …



Storing digital images

- Header: Additional information
  - Width and height: 10x10
  - DPI, etc.
  - Quantization
  - Color space or color table
    - 0 -> white
    - 1 -> dark red
    - 2 -> medium red
    - 3 -> light red



Color Models

- How to „specify“ color?
- Color models are an abstraction
  - mapping color <-> numbers
- Color palettes are an enumeration

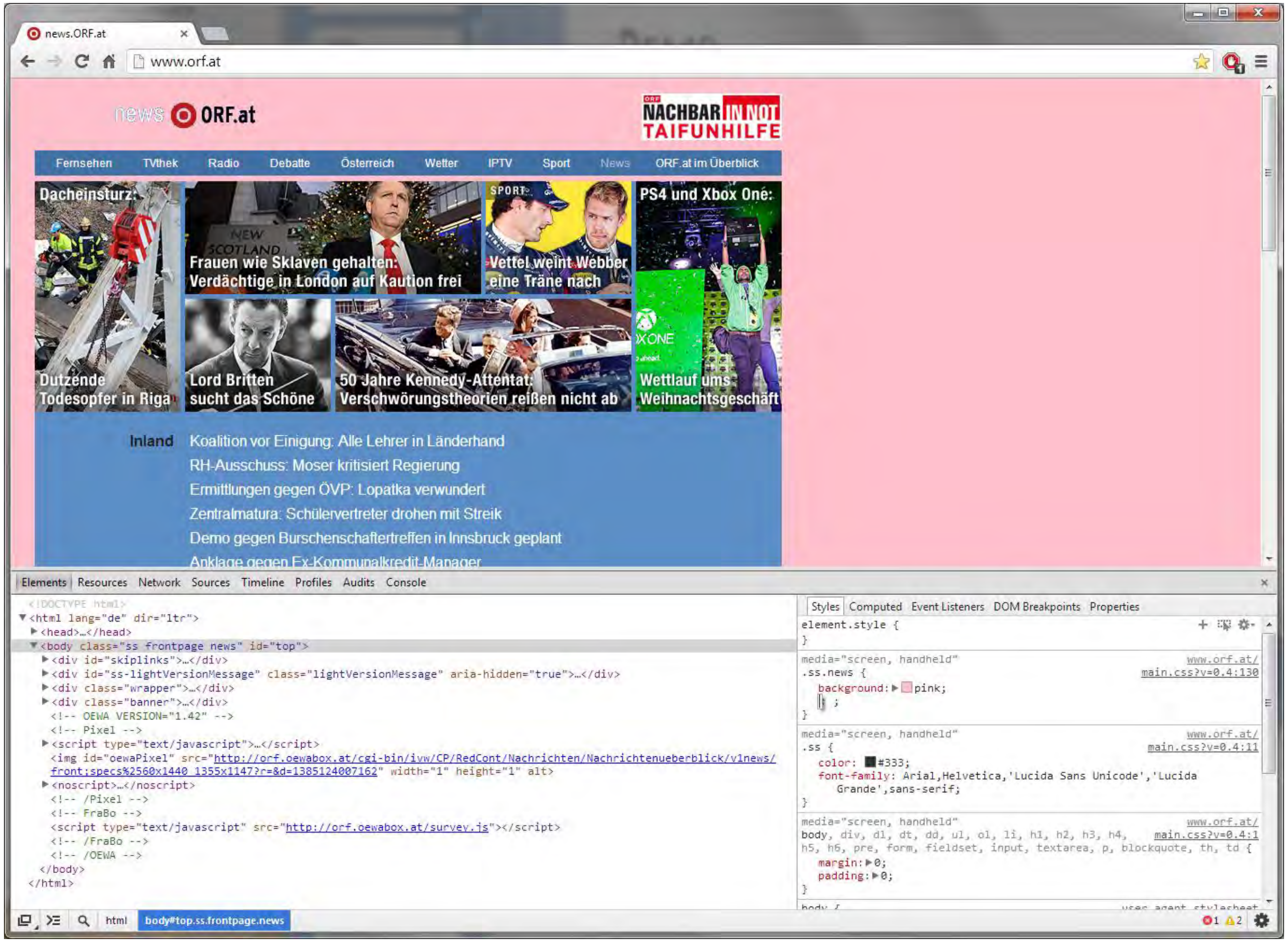
Color palettes
(HTML 4.01 spec.)

CSS1 / HTML3–4 / VGA color names

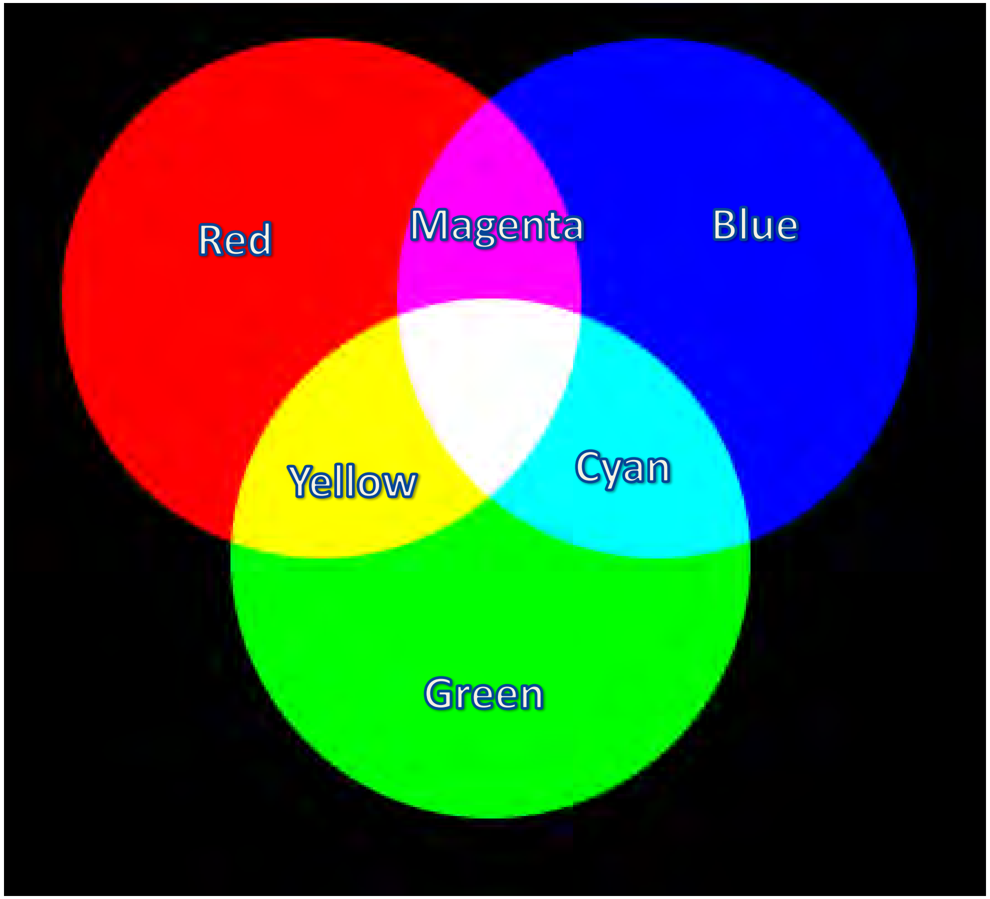
	Name	Hex triplet	Red	Green	Blue	Hue	Satur	Light	Satur	Value	
	White	#FFFFFF	100%	100%	100%	0°	0%	100%	0%	100%	
	Silver	#C0C0C0	75%	75%	75%	0°	0%	75%	0%	75%	
	Gray	#808080	50%	50%	50%	0°	0%	50%	0%	50%	
	Black	#000000	0%	0%	0%	0°	0%	0%	0%	0%	
	Red	#FF0000	100%	0%	0%	0°	100%	50%	100%	100%	
	Maroon	#800000	50%	0%	0%	0°	100%	25%	100%	50%	
	Yellow	#FFFF00	100%	100%	0%	60°	100%	50%	100%	100%	
	Olive	#808000	50%	50%	0%	60°	100%	25%	100%	50%	
	Lime	#00FF00	0%	100%	0%	120°	100%	50%	100%	100%	green
	Green	#008000	0%	50%	0%	120°	100%	25%	100%	50%	
	Aqua	#00FFFF	0%	100%	100%	180°	100%	50%	100%	100%	cyan
	Teal	#008080	0%	50%	50%	180°	100%	25%	100%	50%	
	Blue	#0000FF	0%	0%	100%	240°	100%	50%	100%	100%	
	Navy	#000080	0%	0%	50%	240°	100%	25%	100%	50%	
	Fuchsia	#FF00FF	100%	0%	100%	300°	100%	50%	100%	100%	magenta
	Purple	#800080	50%	0%	50%	300°	100%	25%	100%	50%	

Src. Wikipedia

Demo

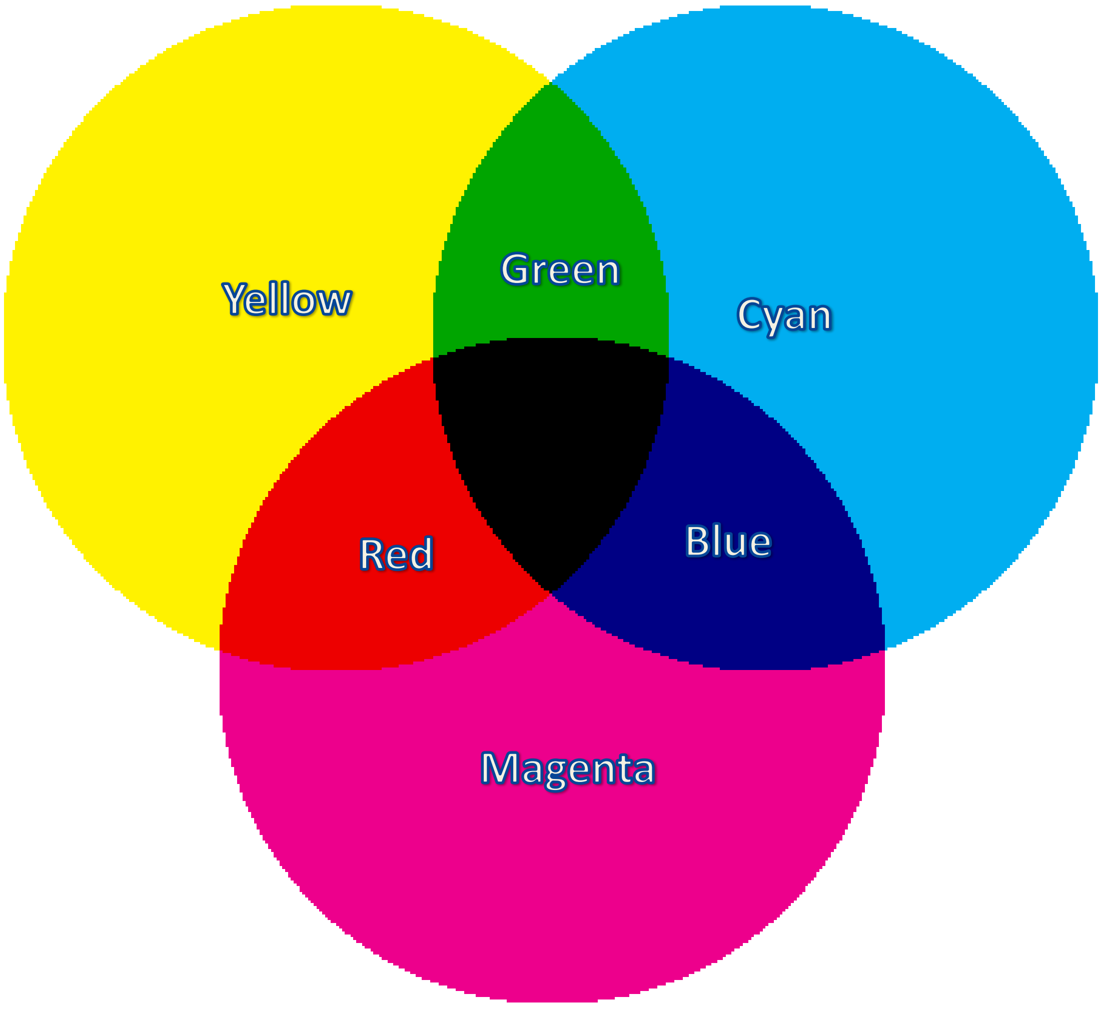

news.ORF.at
www.orf.at
NACHBAR IN NOT
TAIFUNHILFE
Fernsehen TVthek Radio Debatte Österreich Wetter IPTV Sport News ORF.at im Überblick
Dacheinsturz: Dutzende Todesopfer in Riga
Frauen wie Sklaven gehalten: Verdächtige in London auf Kaution frei
Vettel weint Webber eine Träne nach
PS4 und Xbox One: Wettlauf ums Weihnachtsgeschäft
Lord Britten sucht das Schöne
50 Jahre Kennedy-Attentat: Verschwörungstheorien reißen nicht ab
Inland
Koalition vor Einigung: Alle Lehrer in Länderhand
RH-Ausschuss: Moser kritisiert Regierung
Ermittlungen gegen ÖVP: Lopatka verwundert
Zentralmatura: Schülervertreter drohen mit Streik
Demo gegen Burschenschaftertreffen in Innsbruck geplant
Elements Resources Network Sources Timeline Profiles Audits Console
Styles Computed Event Listeners DOM Breakpoints Properties
.ss.news {
background: pink;
}
body#top.ss.frontpage.news

Color: Additive Scheme
(Mixtures of Light)



Color: Subtractive Scheme

(Mixtures of Pigments)

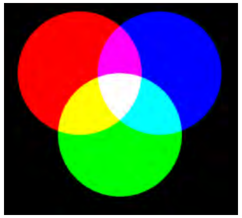


CIE 1931 color space

- First mathematically defined color space
- Tristimulus values: X, Y, Z
  - roughly red, green and blue
  - not observed as such
- Defined based on experiments in the 1920s

RGB Color Model

- Additive color model
- Based on red, green and blue
- Color is defined as mixture of intensities in all 3 channels
- Examples with channel intensities in [0,1]
  - (0, 0, 0) -> no intensities, darkest color
  - (1, 1, 1) -> full intensities, lightest color
  - (1, 0, 0) -> saturated red



RGB Color Model

- Intensities are typically quantized
  - E.g. 8 bit / channel (24 bit per pixel)
  - 16-bit Highcolor: 5 bits / channel (+1 for green)
- With 8 bits per channel
  - range from 0-255 (decimal) or 00-FF (hex)
- Web: 8 bits / channel, defined in hex
  - #000000 -> black
  - #FF0000 -> red

sRGB Color Space

- Standard defined by HP & Microsoft (1996)
- Used for monitors, printer & internet
- Uses 8 bits per channels
- Transformation to and from CIE XYZ defined by:

$$\begin{bmatrix} R_{\text{linear}} \\ G_{\text{linear}} \\ B_{\text{linear}} \end{bmatrix} = \begin{bmatrix} 3.2406 & -1.5372 & -0.4986 \\ -0.9689 & 1.8758 & 0.0415 \\ 0.0557 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Color Space: HSV

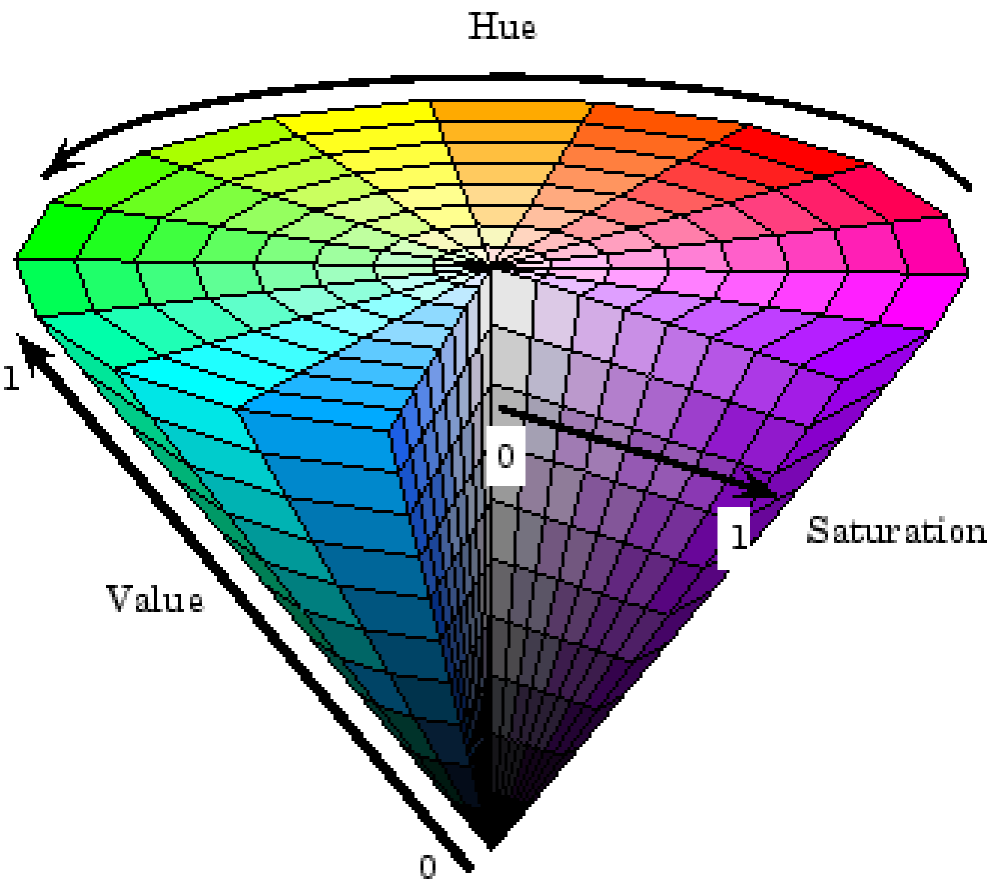
- Transformation of RGB to
  - Hue: The actual color (0°-360°)
  - Saturation: The brilliance of the color
  - Value: The lightness of the color
- Useful for changing hue (color) withouth changing lighness (value) or brilliance (saturation)

Color Space: HSV

Increase saturation

Set fixed hue

Color Space: HSV



Src. Mathworks.com

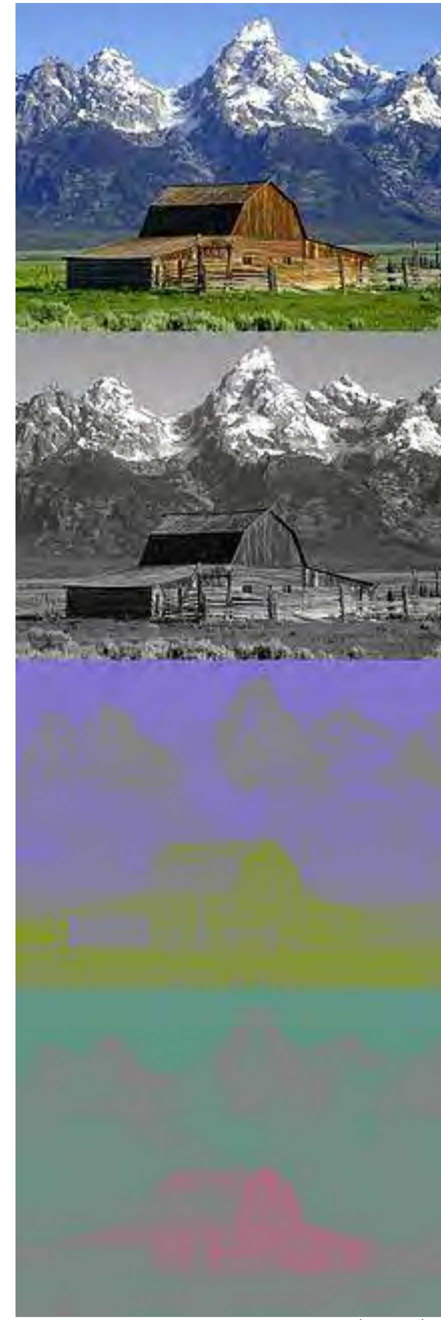
Opponent Color Space

Grayscale

- A pixel just has an intensity value
- Intensity channel := Y
- Conversion from RGB
  - $Y = 0.2126*R + 0.7152*G + 0.0722*B$

YCbCr

- Splits color into luma & chroma
  - Y: intensity
  - Cb: blue-difference
  - Cr: red-difference
- Used for chroma subsampling
  - Reducing "color" in images

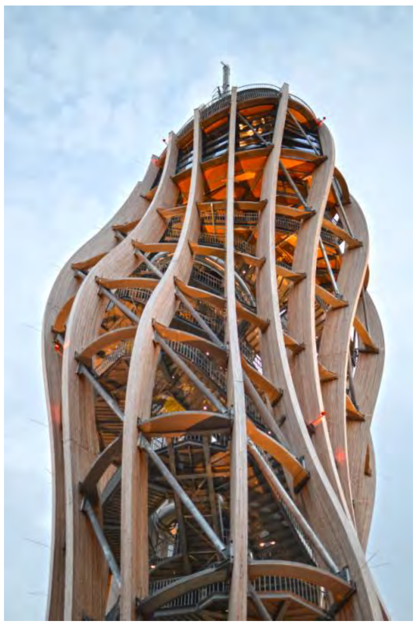


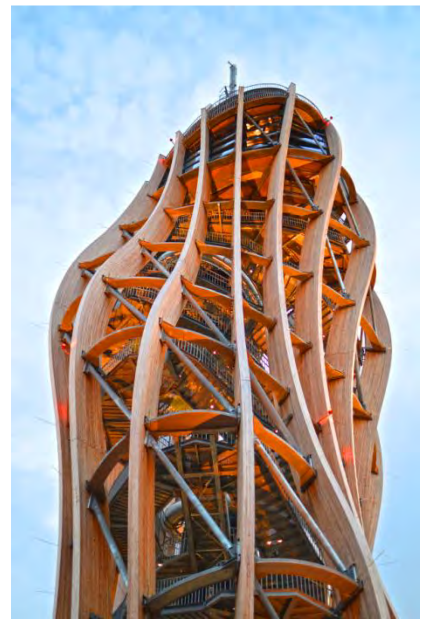
src. Wikipedia

EXERCISE: IMAGE COLORS

- Install GIMP
  - http://www.gimp.org/downloads/
- Use or take a photo
- Using GIMP create these versions
  - Fixed hue
  - High saturation
  - No saturation

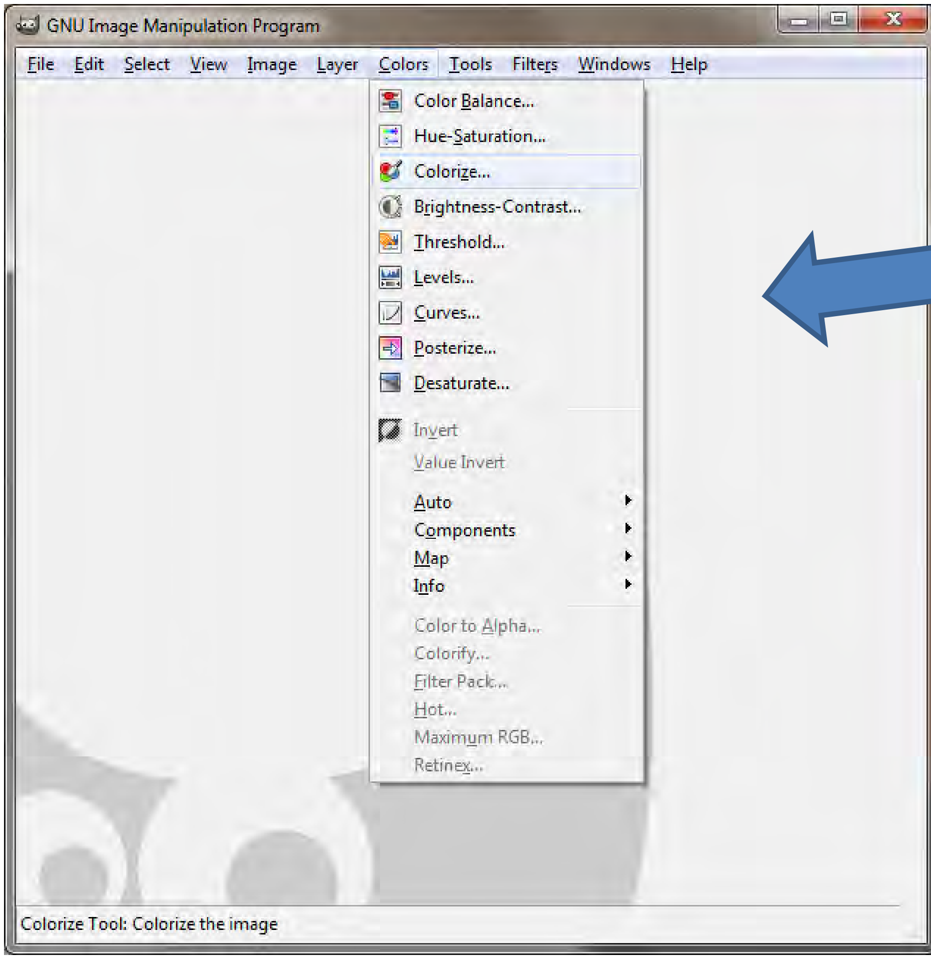
EXERCISE: IMAGE COLORS






It should look like these ...

EXERCISE: IMAGE COLORS



Use this menu in GIMP

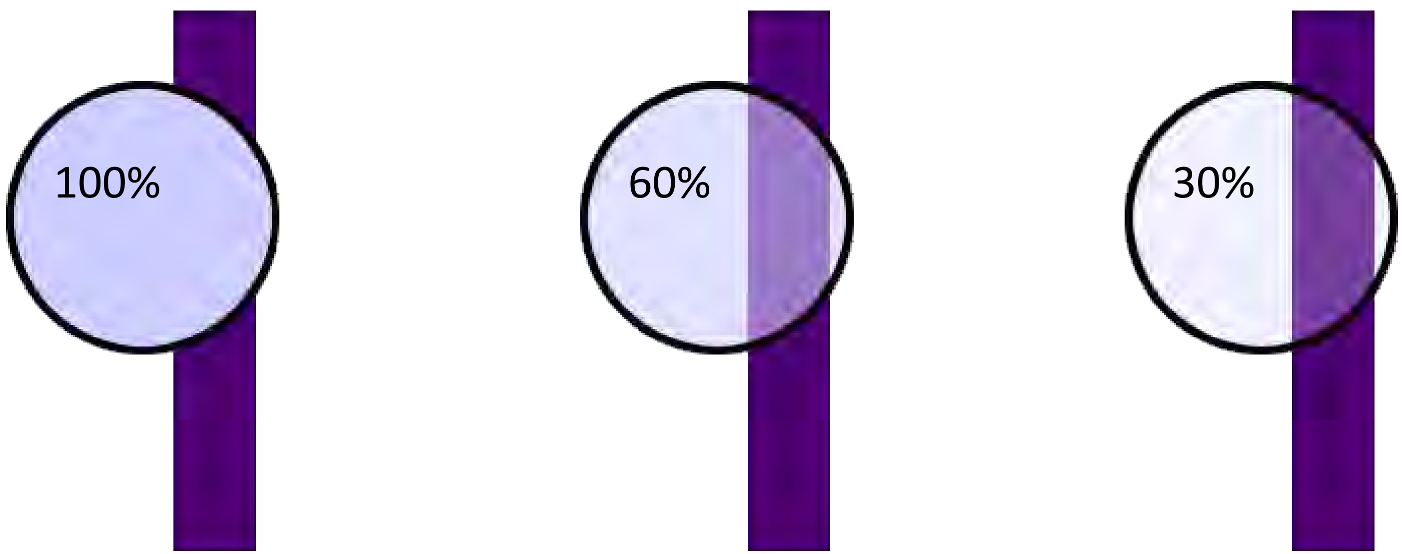
EXERCISE: Image Colors

- Put the images into your report.
- Add 500 words describing your experience.

Transparency & Alpha Channel

- How to define the opacity of a color / pixel?
- Alpha channel
  - 0% … fully transparent
  - 100% … fully opaque pixel
- Color spaces
  - RGBa – Red, Green, Blue & Alpha
  - ARGB – Alpha first, used in Flash, Silverlight

Transparency & Alpha Channel



Demo: Gimp

- Create a text layer on an image
- Change opacity of the layer

Size of digital images ...

- Example: Canon EOS M
  - 333 € (Nov 22, 2013)
  - 18 Mega Pixels

- What`s the size of an image?
  - 5,184 x 3,456 pixels
  - 24 bits per pixel for color = 3 bytes
  - Size: 5,184 x 3,456 x 3 bytes =~ <u>504.7 MB</u>

Compression: Lossless vs. Lossy

Compression is a tradeoff processing time vs. storage space

- Lossless
  - All information is retained
  - Image can be re-constructed without flaws
- Lossy
  - Unimportant information is dropped
  - Image reconstruction is similar to original image

Lossless vs. Lossy

463 KB -- 3.5 KB

Lossless compression
(examples)

- Run-length encoding
  - Crunching long sequences of same numbers
  - wwwwooowwoooowwww -> 4w3o2w4o4w
- Dictonary encoding (LZW, ZIP, …)
  - Re-occuring sequences get short codes
- Huffmann
  - Variable length, entropy encoding
  - Numbers with high frquency get short representations

Lossy Compression: DCT

- DCT == „discrete cosine transform"
- Expresses data points as sum of cosine functions
- Defined by following equation:

$$D(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right] \qquad 1$$

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0 \\ 1 & \text{if } u > 0 \end{cases} \qquad 2$$

Discrete Cosine Transform

Actual encoding process easier than the equation looks:

1. For each 8x8 pixel block from an image
2. Multiply with DCT matrix
3. Quantize with quantization matrix
4. Encode in zig-zag manner

DCT: Example

- Extract 8x8 pixels block
- Subtract 128 from each entry

$$Original = \begin{bmatrix} 154 & 123 & 123 & 123 & 123 & 123 & 123 & 136 \\ 192 & 180 & 136 & 154 & 154 & 154 & 136 & 110 \\ 254 & 198 & 154 & 154 & 180 & 154 & 123 & 123 \\ 239 & 180 & 136 & 180 & 180 & 166 & 123 & 123 \\ 180 & 154 & 136 & 167 & 166 & 149 & 136 & 136 \\ 128 & 136 & 123 & 136 & 154 & 180 & 198 & 154 \\ 123 & 105 & 110 & 149 & 136 & 136 & 180 & 166 \\ 110 & 136 & 123 & 123 & 123 & 136 & 154 & 136 \end{bmatrix}$$

$$M = \begin{bmatrix} 26 & -5 & -5 & -5 & -5 & -5 & -5 & 8 \\ 64 & 52 & 8 & 26 & 26 & 26 & 8 & -18 \\ 126 & 70 & 26 & 26 & 52 & 26 & -5 & -5 \\ 111 & 52 & 8 & 52 & 52 & 38 & -5 & -5 \\ 52 & 26 & 8 & 39 & 38 & 21 & 8 & 8 \\ 0 & 8 & -5 & 8 & 26 & 52 & 70 & 26 \\ -5 & -23 & -18 & 21 & 8 & 8 & 52 & 38 \\ -18 & 8 & -5 & -5 & -5 & 8 & 26 & 8 \end{bmatrix}$$

DCT: Example

- Apply DCT with $D = T*M*T^{-1}$

$$T = \begin{bmatrix} .3536 & .3536 & .3536 & .3536 & .3536 & .3536 & .3536 & .3536 \\ .4904 & .4157 & .2778 & .0975 & -.0975 & -.2778 & -.4157 & -.4904 \\ .4619 & .1913 & -.1913 & -.4619 & -.4619 & -.1913 & .1913 & .4619 \\ .4157 & -.0975 & -.4904 & -.2778 & .2778 & .4904 & .0975 & -.4157 \\ .3536 & -.3536 & -.3536 & .3536 & .3536 & -.3536 & -.3536 & .3536 \\ .2778 & -.4904 & .0975 & .4157 & -.4157 & -.0975 & .4904 & -.2778 \\ .1913 & -.4619 & .4619 & -.1913 & -.1913 & .4619 & -.4619 & .1913 \\ .0975 & -.2778 & .4157 & -.4904 & .4904 & -.4157 & .2778 & -.0975 \end{bmatrix}$$

$$M = \begin{bmatrix} 26 & -5 & -5 & -5 & -5 & -5 & -5 & 8 \\ 64 & 52 & 8 & 26 & 26 & 26 & 8 & -18 \\ 126 & 70 & 26 & 26 & 52 & 26 & -5 & -5 \\ 111 & 52 & 8 & 52 & 52 & 38 & -5 & -5 \\ 52 & 26 & 8 & 39 & 38 & 21 & 8 & 8 \\ 0 & 8 & -5 & 8 & 26 & 52 & 70 & 26 \\ -5 & -23 & -18 & 21 & 8 & 8 & 52 & 38 \\ -18 & 8 & -5 & -5 & -5 & 8 & 26 & 8 \end{bmatrix}$$

$$D = \begin{bmatrix} 162.3 & 40.6 & 20.0 & 72.3 & 30.3 & 12.5 & -19.7 & -11.5 \\ 30.5 & 108.4 & 10.5 & 32.3 & 27.7 & -15.5 & 18.4 & -2.0 \\ -94.1 & -60.1 & 12.3 & -43.4 & -31.3 & 6.1 & -3.3 & 7.1 \\ -38.6 & -83.4 & -5.4 & -22.2 & -13.5 & 15.5 & -1.3 & 3.5 \\ -31.3 & 17.9 & -5.5 & -12.4 & 14.3 & -6.0 & 11.5 & -6.0 \\ -0.9 & -11.8 & 12.8 & 0.2 & 28.1 & 12.6 & 8.4 & 2.9 \\ 4.6 & -2.4 & 12.2 & 6.6 & -18.7 & -12.8 & 7.7 & 12.0 \\ -10.0 & 11.2 & 7.8 & -16.3 & 21.5 & 0.0 & 5.9 & 10.7 \end{bmatrix}$$

DCT: Example

- Apply quantization matrix

$$D = \begin{bmatrix} 162.3 & 40.6 & 20.0 & 72.3 & 30.3 & 12.5 & -19.7 & -11.5 \\ 30.5 & 108.4 & 10.5 & 32.3 & 27.7 & -15.5 & 18.4 & -2.0 \\ -94.1 & -60.1 & 12.3 & -43.4 & -31.3 & 6.1 & -3.3 & 7.1 \\ -38.6 & -83.4 & -5.4 & -22.2 & -13.5 & 15.5 & -1.3 & 3.5 \\ -31.3 & 17.9 & -5.5 & -12.4 & 14.3 & -6.0 & 11.5 & -6.0 \\ -0.9 & -11.8 & 12.8 & 0.2 & 28.1 & 12.6 & 8.4 & 2.9 \\ 4.6 & -2.4 & 12.2 & 6.6 & -18.7 & -12.8 & 7.7 & 12.0 \\ -10.0 & 11.2 & 7.8 & -16.3 & 21.5 & 0.0 & 5.9 & 10.7 \end{bmatrix}$$

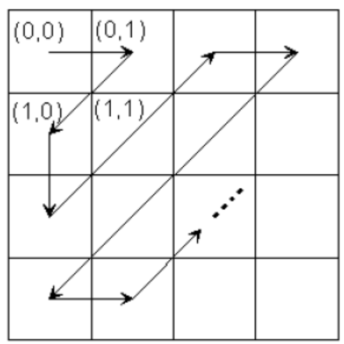
$$Q_{50} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

$$C_{i,j} = round\left(\frac{D_{i,j}}{Q_{i,j}}\right)$$

$$C = \begin{bmatrix} 10 & 4 & 2 & 5 & 1 & 0 & 0 & 0 \\ 3 & 9 & 1 & 2 & 1 & 0 & 0 & 0 \\ -7 & -5 & 1 & -2 & -1 & 0 & 0 & 0 \\ -3 & -5 & 0 & -1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

DCT: Example

- Encode in zig-zag manner



$$C = \begin{bmatrix} 10 & 4 & 2 & 5 & 1 & 0 & 0 & 0 \\ 3 & 9 & 1 & 2 & 1 & 0 & 0 & 0 \\ -7 & -5 & 1 & -2 & -1 & 0 & 0 & 0 \\ -3 & -5 & 0 & -1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- 10, 4, 3, -7, 9, 2, 5, 1, -5, -3, -2, ….

Note: There are a lot of zeros at the end!

DCT: Example

Inverse process:

- Multiply component-wise by the same quantization matrix & round
- Perform IDCT: $T^{-1}*M*T$
- Add 128 per component

DCT: Example

- DCT with Q_{50}
- Compressed version to the right

Image Containers & Formats

- Windows Bitmap File Format (BMP)
- Portable Network Graphics (PNG)
- Tagged Image File Format (TIFF)
- JPEG images and file format (JPEG)
- Scalable Vector Graphics (SVG)

Bitmap File Format

- Popular format for storing raster graphic
- Microsoft Windows and former OS/2 platforms
- Also known as Device Independent Bitmap (DIB) format
- Typical file extensions .bmp and .dib, MIME-Type image/x-ms-bmp

- Characteristics:
  - Stores pixel values typically uncompressed or in RLE
  - Color palette mode available
  - Supports RGB color model and greyscale bitmaps
  - Color depth of 1, 4, 8, 16, 24, or 32 bits per pixel
  - Simple format, easy handling
  - Typically used for icons and small pictures

Bitmap File Format

- File organization
  - Bitmap File Header (14 Bytes)
    - Magic Number (0x42 0x4D, ASCII Code for 'BM')
    - Total size of the bitmap file (4 Bytes)
    - Offset to bitmap data in file (4 Bytes)
  - Bitmap Information
    - Structure depends on file format version (e.g. V3, V4, V5)
    - Size ranging from 12 – 124 Bytes
    - Bitmap width and height in pixels (4 Bytes each)
    - Number of bits per pixel (color depth)
    - Optional compression method
    - Number of colors in palette (0 – 2^32)

Bitmap File Format

- File organization (cont'd)
  - Color Palette (optional)
    - Color descriptions using 24 Bit RGB values + 1 Byte Padding
    - Number of palette entries specified in Bitmap Information block
    - Colors referenced in Bitmap Data by index into Color Palette
  - Bitmap Data
    - Pixel values (either color or index into palette) stored row-by-row
    - Unless otherwise specified rows are stored in a bottom-up fashion, i.e. lower left corner to upper right corner
    - First byte of a row has to be word (32-bit) aligned
- Remarks
  - Actual storage depends on color-depth, compression etc.
  - Values are Little-Endian encoded (Intel x86)

Graphics Interchange Format (GIF)

- Developed by CompuServe in the late 80s
  - GIF87a – introduced 1987
  - GIF89a – extension in 1989 (animated GIFs)
- Storage of raster image
  - 8 Bit color palette
  - Each pixel consists of index into palette
- LZW-based compression of image data
  - Paper in 1984, Patent in 1985 in different countries,
    claims by Unisys in 1993,
    Patent expired in 2003 and 2004 (can be used freely)
- Supports
  - multiple images per file
  - transparent color
  - Progressive mode (interlaced image transmitted first)
- File extension .gif, MIME Type image/gif

Progressive mode ...

Src. http://msdn.microsoft.com/en-us/library/ee720036(VS.85).aspx

Portable Network Graphics (PNG)

- Design goals
  - Portable storage of raster-images
  - Lossless compression
  - Patent-free replacement for GIF format
- Development History
  - v1.0 in 1996, RFC 2083 and W3C Recommendation
  - v1.1 in 1998 and v1.2 in 1999 – minor modifications
  - official ISO standard since 2004 - ISO/IEC 15948:2004

Portable Network Graphics (PNG)

- Characteristics
  - indexed-color, greyscale and true-color raster images
  - Supports only RGB color model (no YCbCr or CMYK)
  - Progressive display mode
  - Streamability (files can be read or written serially)
  - Transparency and alpha channel support
  - Checksums for file corruption checking
- Metadata
  - Ancillary information (e.g. Textual descriptions)
  - Gamma correction and color calibration information

PNG – File Structure

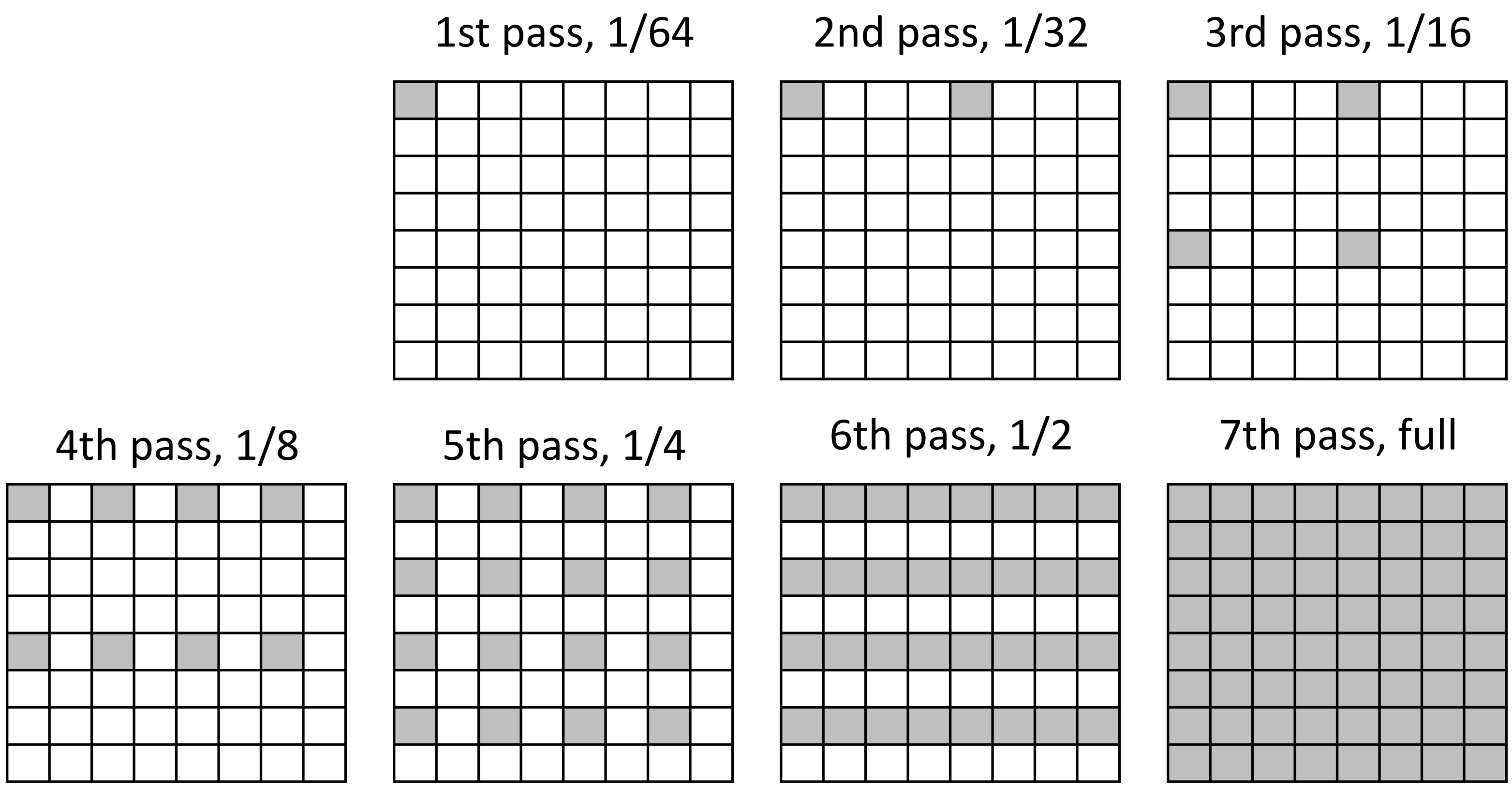
- PNG signature
  - File must start with fixed 8-Byte sequence
  - Hex: 0x89 0x50 0x4E 0x47 0x0D 0x0A 0x1A 0x0A
  - ASCII: \211 P N G \r \n \032 \n
- Sequence of chunks
  - First chunk: IHDR, last chunk: IEND
  - Each chunk consists of
    - Length (4 Byte): number of bytes in the chunk data (*n*)
    - Chunk Type (4 Byte): ASCII letters (a-z, A-Z)
    - Chunk Data (*n* Bytes): actual payload of chunk
    - CRC (4 Bytes): checksum over Type and Data field

PNG – Interlaced Mode

- For implementing progressive display of pictures/logos
- Use case
  - Slow internet connections
  - Low-res image is immediately visible
- Implementation – Adam7 interlace mode
  - Seven passes over the pixel of the image
  - In each pass a subset of pixels are encoded
  - Based on a 8x8 pattern replicated over the image data
  - Each pixel assigned to a specific pass

1	6	4	6	2	6	4	6
7	7	7	7	7	7	7	7
5	6	5	6	5	6	5	6
7	7	7	7	7	7	7	7
3	6	4	6	3	6	4	6
7	7	7	7	7	7	7	7
5	6	5	6	5	6	5	6
7	7	7	7	7	7	7	7

PNG – Interlaced Mode



TIFF – Tagged Image File Format

- Initially developed by Aldus (first version in 1986)
- Now part of Adobe Systems Inc.
- Latest revision 6.0 – June 1992,
  since than no further development
- Well-supported by many applications especially in the context of scanners and fax software
- Specification divided into two parts
  - Baseline TIFF
  - TIFF Extensions

TIFF – File Structure

- Each file consists of a fixed-size header and a set of (at least one) Image File Directory (IFD)

- Each file begins with a 8-byte image header
  - Byte 0-1: signals the byte order
    - 0x49 0x49 ("II")…Little-endian encoding
    - 0x4D 0x4D ("MM")…Big-endian encoding
  - Byte 2-3: magic number (0x2A or decimal 42)
  - Byte 4-7: offset (in bytes) of the first IFD

- Image File Directory
  - Represents an image in the file
  - One TIFF file can contain more than one image (e.g. Scanner)

TIFF – TIFF Extensions

- Support for multiple images per file
  - Storage of a scanned document in a single file
  - Storage of a received fax message
- Document Storage and Retrieval
  - Additional TIFF fields for annotations & metadata.
- Compression
  - Enhanced CCIT Bilevel Encodings
  - LZW
  - JPEG Compression support
- Color spaces
  - YCbCr (with chroma subsampling support)
  - CIE L*a*b*
  - RGB Image Colorimetry
  - CMYK (Cyan, Magenta, Yellow, Key) for printing

TIFF – Summary

- Supports a great variety of color spaces
  - RGB, CMYK, YCbCr, Grayscale, ...
- Multiple, arbitrary sized images per file
  - Scanned documents, Fax servers
  - Thumbnails
- Compression modes
  - Uncompressed, Lossless & Lossy
- Wide range of color depths
  - 1 Bit B/W up to 48 Bit RGB (e.g., for scientific imaging)
  - Support of Alpha Channels
- Fragmentation - Strips or Tiles can be decoded independently
- File extension .tif or .tiff, MIME-Type image/tiff

JPEG

- Standardized as CCITT Rec.81 in 1992
  - Specification of encoder
  - Specification of decoder
  - Definition of an interchange format (Annex B)
- Applications
  - Good compression of photographs and paintings
  - Not very suitable for drawings, icons or text (sharp edges)
  - Not recommended as intermediate format during image editing (generation loss)
- Coding tools
  - DCT-based lossly compression
  - Prediction-based lossless compression
  - Huffman or arithmetic entropy coding

JPEG – File Format

- JPEG Interchange Format (JIF)
  - Annex B defines the JPEG Interchange Format (JIF)
  - Some shortcomings (definition of color-space etc.)
  - Almost never used in applications nowadays
- JPEG File Interchange Format (JFIF) in 1992
  - Defines YCbCr as default color space
  - Allows for defining sample aspect ratios
  - Clarifies the location of chroma subsamples
  - Thumbnail support
- Wide support in applications as well as devices like digital still cameras, mobile phones etc.

JPEG

- File organization based on marker
  - Can be identified in the file without decoding the image
  - In contrast to chunks (PNG) the marker do not contain length information about the compressed image data
- Variety of compression modes
  - Lossless and DCT-based
  - Sequential, progressive and hierarchical storage
- 8 or 12 Bit sample precision per component
- JFIF supports inclusion of thumbnail
- File extension .jpg or .jpeg, MIME-Type image/jpeg

Image Format Selection …

Consider following use cases:

- Print: Color Model CMYK.
- Web: Storing a logo without compression artifacts.
- Email: Sending a photo to a friend.
- Image Processing: Store a preprocessed photo for more processing later on.

Spatial Filtering

- Methods for *enhancing* or *transforming* the image
- Typically a *kernel* or *filter* is used:
  - A matrix which is applied to the image
  - In a linear transformation

Spatial Filtering

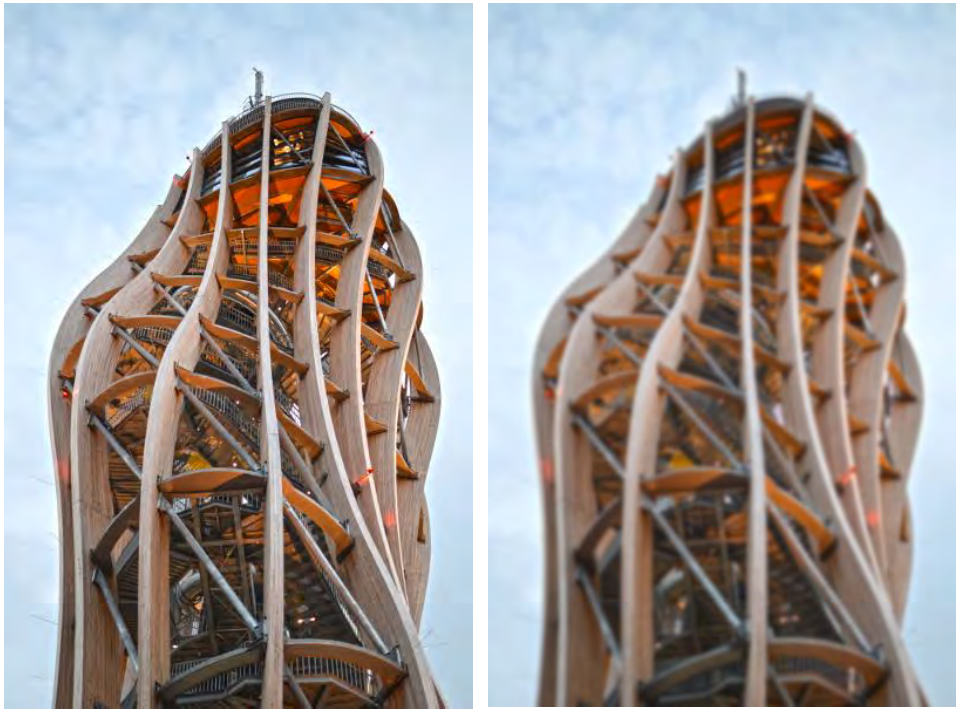
194	128	102	197	69
162	68	103	144	115
121	85	57	27	14
24	183	192	239	150
92	93	154	138	170

194	128	102	197	69
162	68	103	144	115
121	85	122	27	14
24	183	192	239	150
92	93	154	138	170

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Spatial Filtering

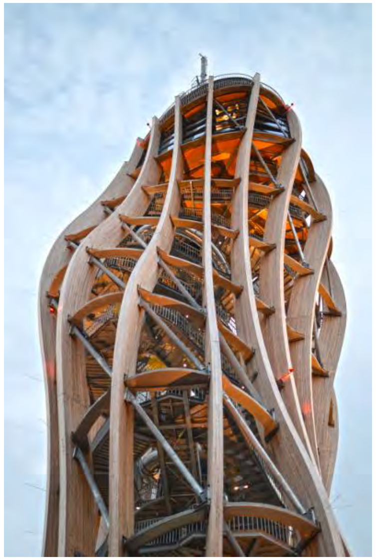
- This is a simple smoothing kernel
- Other operations
  - Sharpen
  - Gradient
  - ...

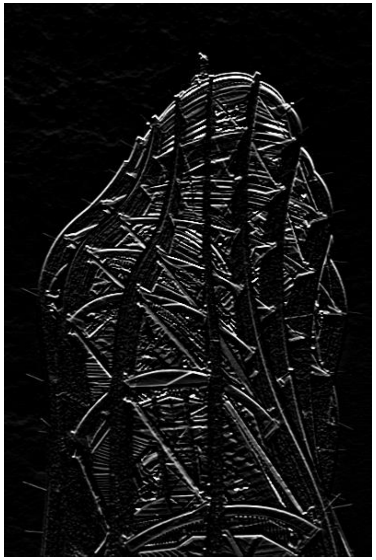


Sobel Filter

- 3x3 kernel

-1	-2	-1
0	0	0
1	2	1





Sharpening

- 3x3 kernel

-1/9	-1/9	-1/9
-1/9	9	-1/9
-1/9	-1/9	-1/9

Blur Filter

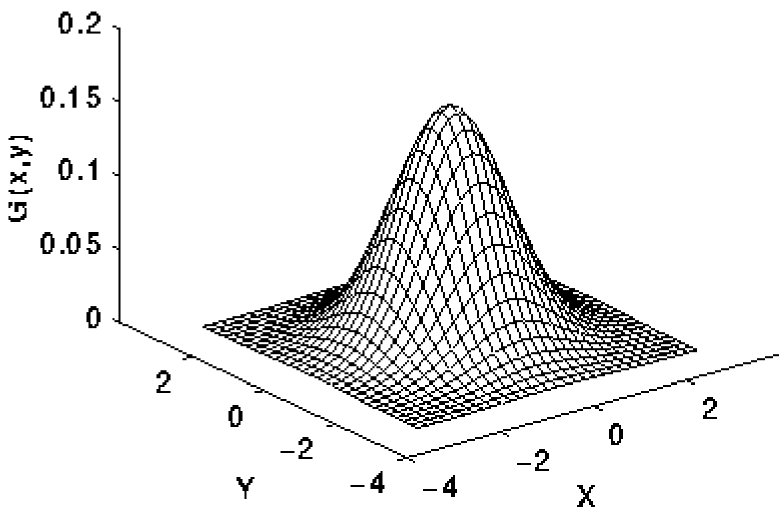
- 3x3 kernel

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Gaussian Blur Filter

- Kernel depends on Gaussian distribution

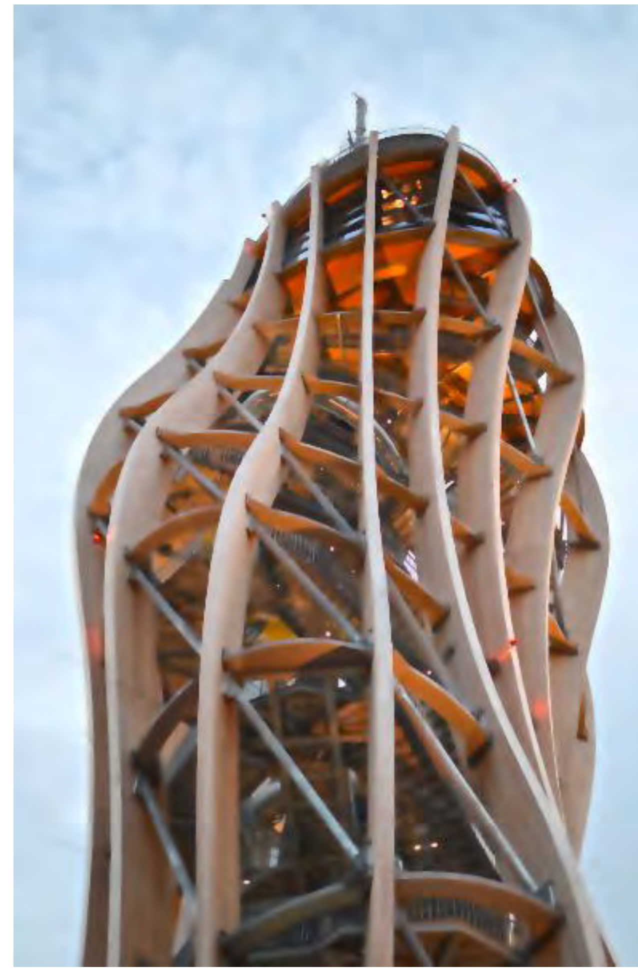
$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



Selective Blur (Egde Retaining)

- More complicated …

$$I^{\text{filtered}}(x) = \sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|)$$



Edge Detection

Edge Detection

- Apply two Kernels
  - with results L_x and L_y
- Then compute the gradient magnitude

$$|\nabla L| = \sqrt{L_x^2 + L_y^2}$$

- Edge direction: $\arctan(L_y/L_x)$

1	2	1
0	0	0
-1	-2	-1

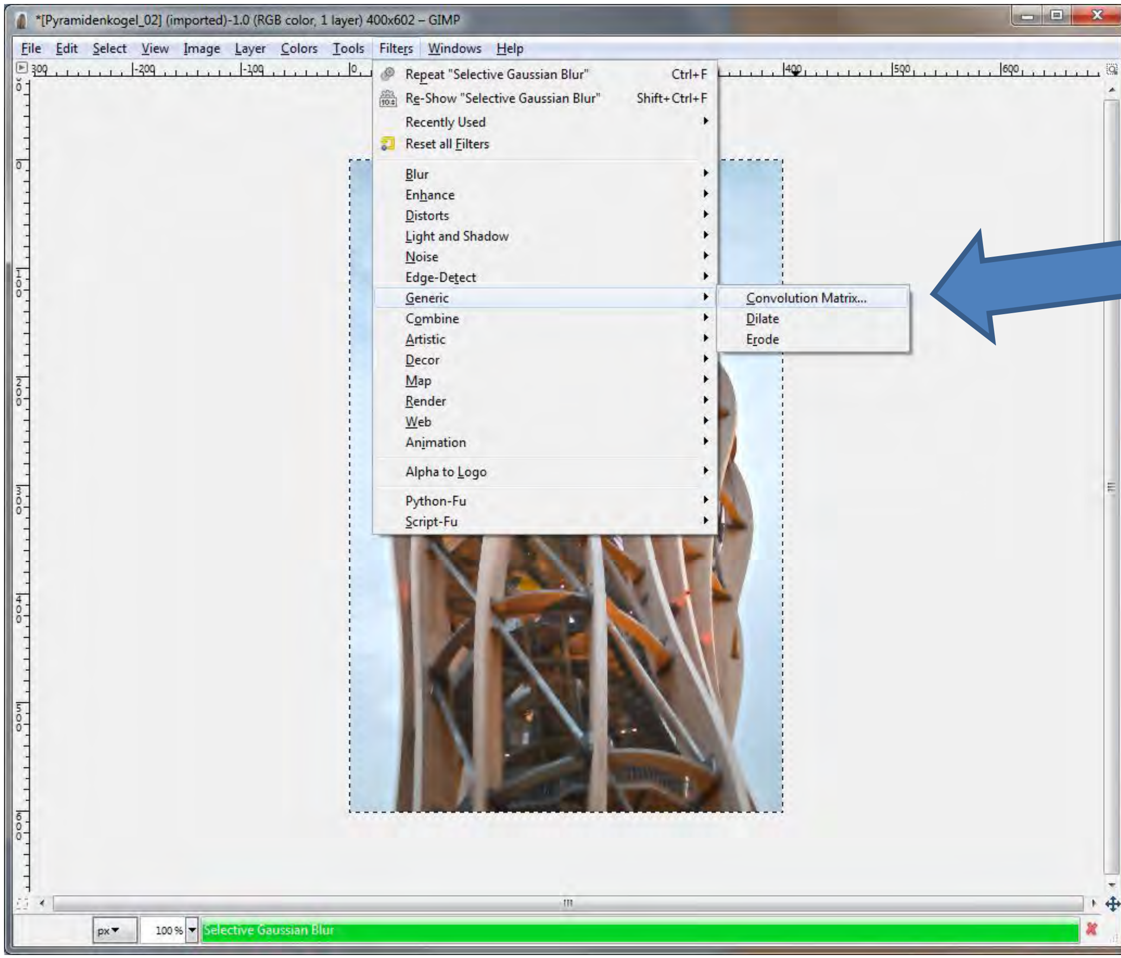
1	0	-1
2	0	-2
1	0	-1

Edge Detection

Exercise: Apply Filters in Gimp

- Install GIMP
  - http://www.gimp.org/downloads/
- Use or take a photo
- Using GIMP create these versions
  - Generic blur filter (from slides)
  - Generic Sobel filter (from slides)
  - Generic sharpen filter (from slides)
  - Use "Filter" -> "Selective Gaussian Blur …"

Exercise: Apply Filters in Gimp



Use this menu in GIMP

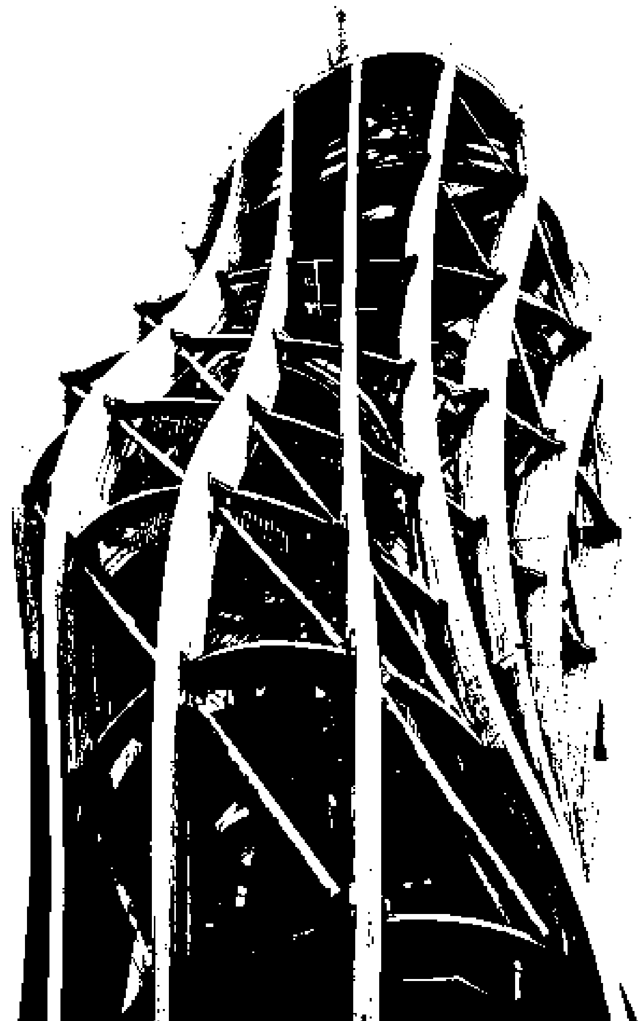
Exercise: Apply Filters in Gimp

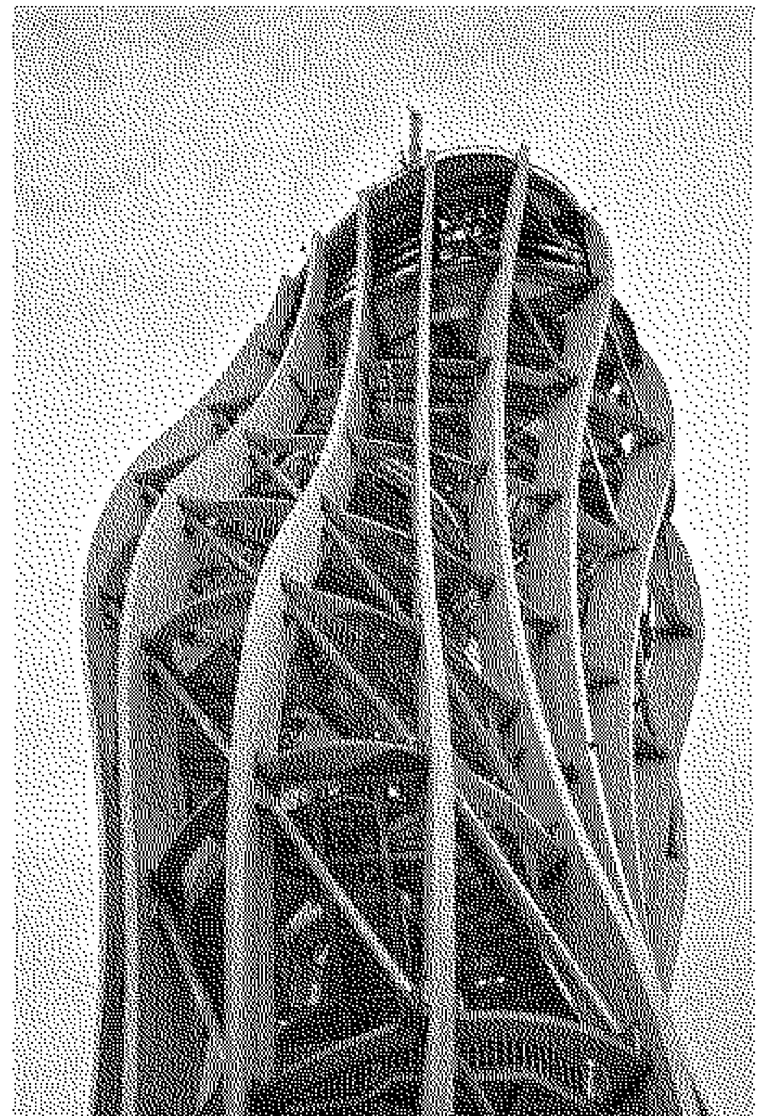
- Put the images into your report.
- Add 500 words describing your experience.

Dithering

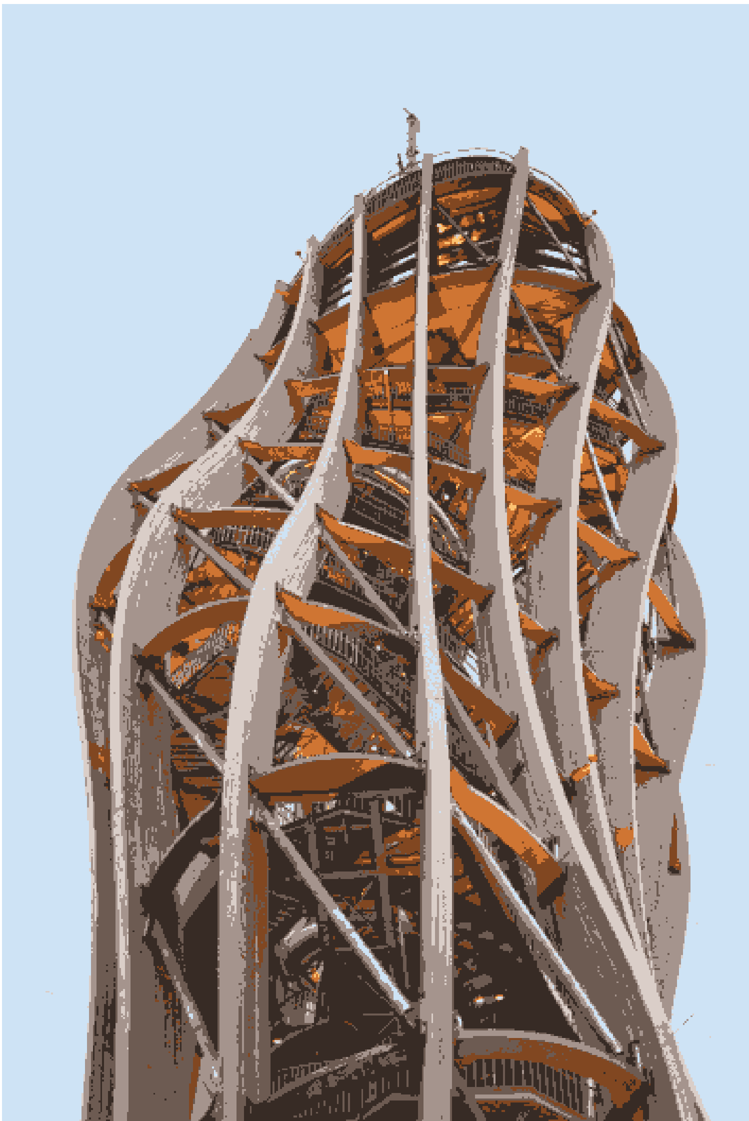
- Intentionally introduced noise
  - to prevent color banding.
  - to have a more appealing visual representation with fewer colors.
- Typically used in newspaper
  - ie. to print grayscale images in black & white

Dithering Example: B & W





Dithering Example: 8 Colors



Floyd-Steinberg Dithering

- Color quantization error is pushed to neighbor pixels

$$\begin{bmatrix} & & * & \frac{7}{16} & \ldots \\ \ldots & \frac{3}{16} & \frac{5}{16} & \frac{1}{16} & \ldots \end{bmatrix}$$

```
for each y from top to bottom
   for each x from left to right
      oldpixel  := pixel[x][y]
      newpixel  := find_closest_palette_color(oldpixel)
      pixel[x][y]  := newpixel
      quant_error  := oldpixel - newpixel
      pixel[x+1][y  ] := pixel[x+1][y  ] + 7/16 * quant_error
      pixel[x-1][y+1] := pixel[x-1][y+1] + 3/16 * quant_error
      pixel[x  ][y+1] := pixel[x  ][y+1] + 5/16 * quant_error
      pixel[x+1][y+1] := pixel[x+1][y+1] + 1/16 * quant_error
```

src. Wikipedia

Introduction to Media Informatics: Videos

Dr. Mathias Lux
Alpen-Adria Universität Klagenfurt

Contents

- Video Compression
- H.264
- Video Formats & Containers
- Video Retrieval
- Video Summaries

Video

- Video typically names the combination of
  - One or more audio streams and
  - One or more video streams

Video

h.264 compressed video stream

AAC encoded German audio stream/s

AAC encoded English audio stream/s

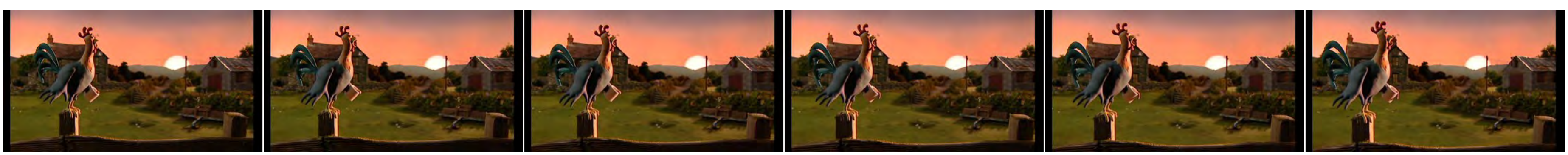
Video

We focus on the visual part:

- A video is a sequence of pictures
- Shown at a certain rate
  - More than 12 images / second needed
  - e.g. 25 (PAL) 30 (NTSC), 24 (24p)
- For equal length time slots
  - cp. jitter

Video Compression

- Basically there are two types of compression:
  - Intra frame compression: A single image is compressed (like a JPEG image).
  - Inter frame compression: Compression is based on surrounding frames.



Video Compression

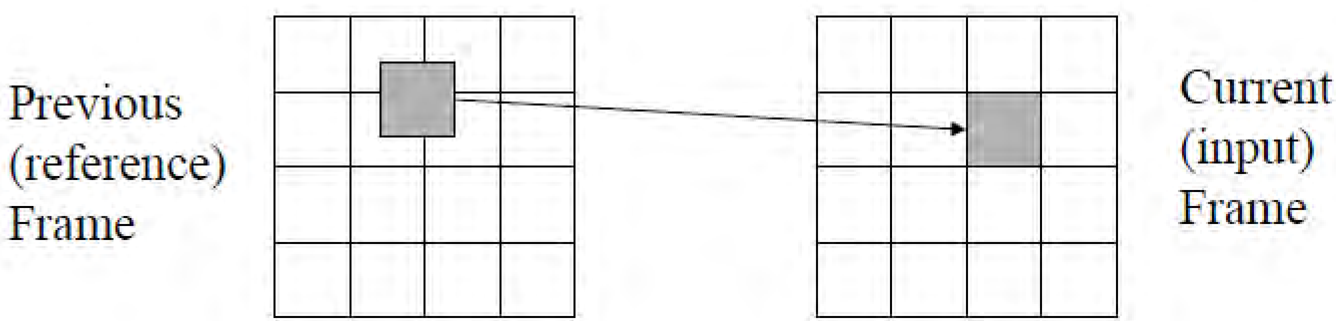
Difference between temporal neighbors

Idea: Compensate Motion

- Difference originates from motion
  - Foreground vs. background motion
- Encoder tries to estimate motion
  - E.g. rooster moves from left to right, 2 pixels
- Decoder "moves" last decoded picture
- Difference is encoded

Idea: Compensate Motion

- Image is split in blocks (e.g. 8x8 pixels)
- "Motion vectors" are computed
  - Where does the block move from the last picture?
- Difference is stored
  - Between predicted block and
  - Moved block in the previous picture



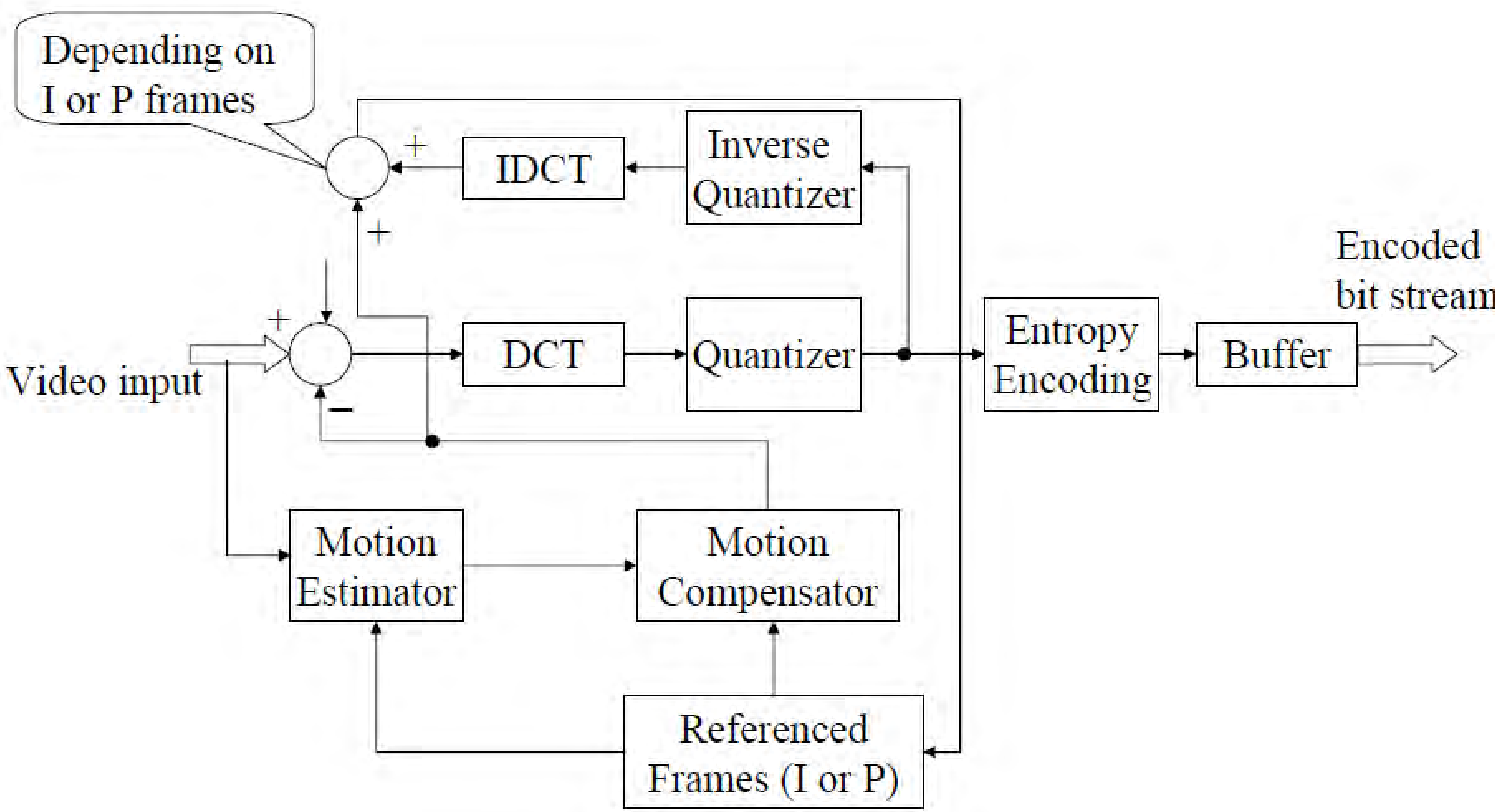
Inter Frame Compression

- I-Frame: Intra coded picture
- P-Frame: Predicted picture
- B-Frame: Bi-directional predicted picture

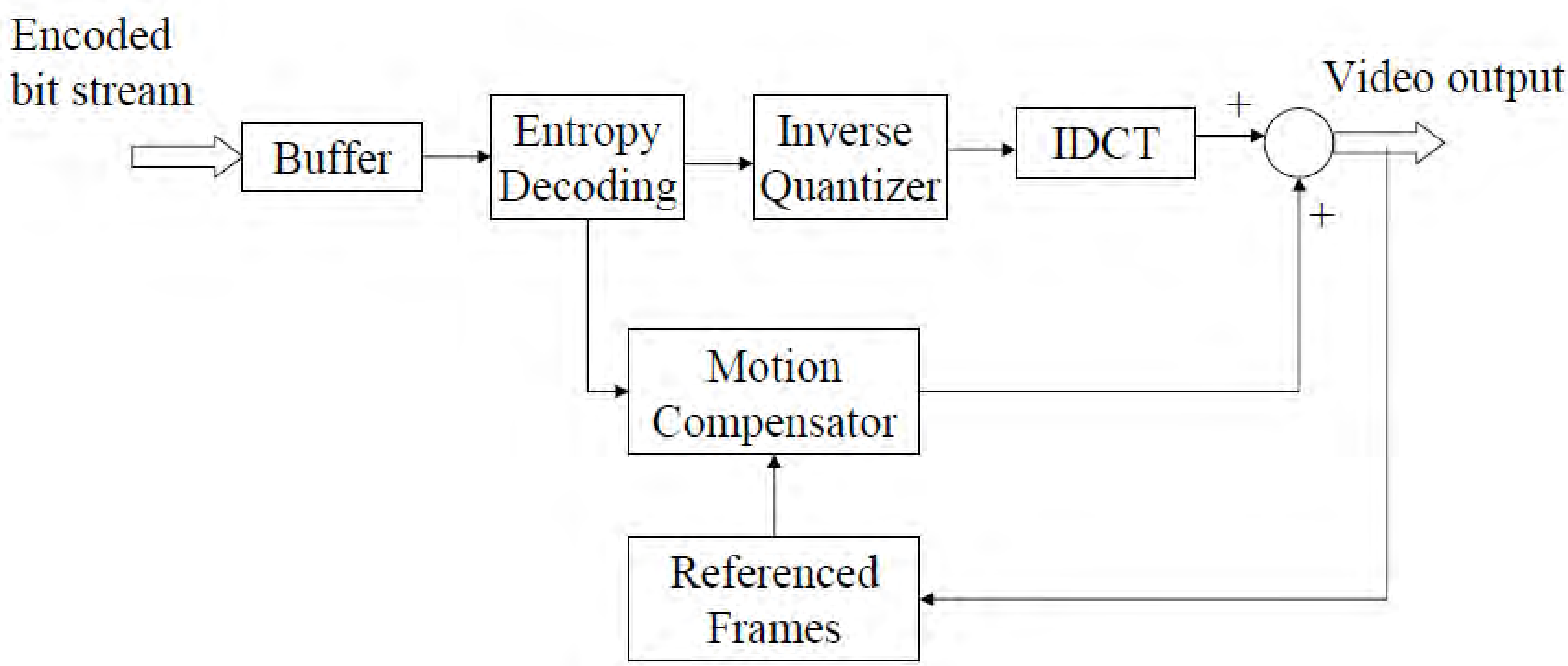
Inter Frame Compression

- Frames are encoded in GOPs (group of pictures)
- GOPs differ in size based on the
  - Compression standard and
  - Video content
- Examples
  - GOP size 5: I-P-P-P-P
  - GOP size 17: I-B-B-B-P-B-B-B-P-B-B-B-P-B-B-B-I

MPEG Video Encoding Scheme



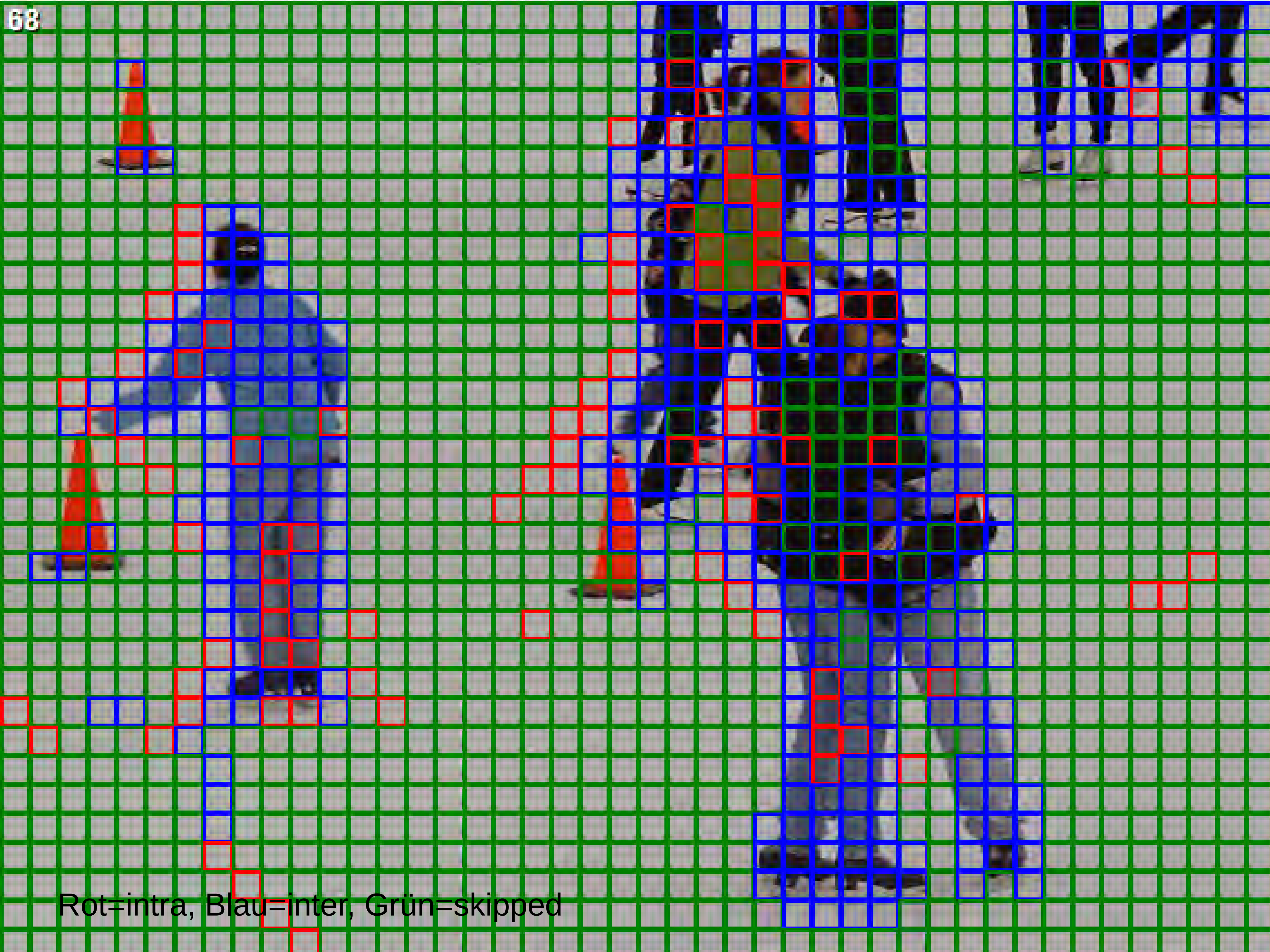
MPEG Video Decoding Scheme

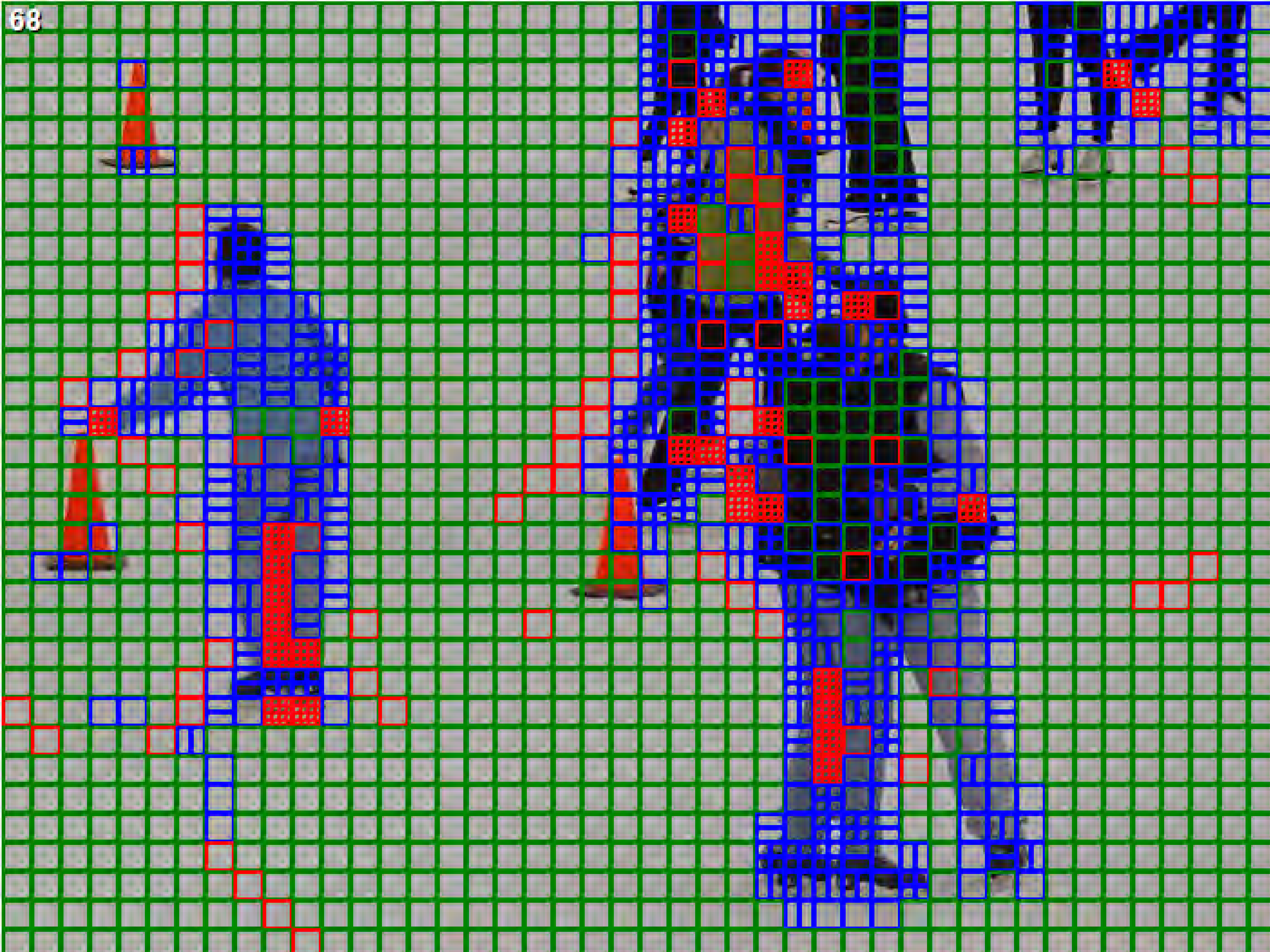


H.264 Internals: Example

- H.264 slices and macroblocks for prediction
- Skip, I- and P-macroblocks
- Motion vectors

Src. K. Schöffmann, ITEC, Uni Klu, with permission


68
Rot=intra, Blau=inter, Grün=skipped


68


68

Video Coding: Additional Thoughts

- Why does encoding take so long?
- What is multi-pass encoding?
- How can "fast forward" and "reverse" work?
- How do I jump to a certain time point in a video?

Video Formats & Containers: AVI

- AVI := Audio Video Interleave
- Widespread format for storing audio visual data
- Defined by Microsoft, very popular on Windows systems
- Based on the Resource Interchange File Format (RIFF)
  - File is organized in chunks and lists
  - Sequential and hierarchical composition of chunks and lists

Video Formats & Containers: ASF

- ASF := Advanced Systems Format
  - Meant as replacement for AVI
- Design goal:
  - Efficient playback from servers (e.g. HTTP server)
  - local storage
  - May be used for editing purposes (no explicit focus)
- Lessons learned
  - Support of incremental generation of content (live content)
  - 64 Bits size fields (to tackle 4 GB limitations)
- File extensions
  - .wmv (video data) and .asf (otherwise)

Video Formats & Containers: MP4 / MOV

- Very powerful and complex format
  - File organization based on boxes
  - Supports file sizes > 4 GB
  - Explicit timing information
  - Focus on MPEG-based audio and video formats
- Streaming support
  - Video streaming supported (e.g., used by iPhone)
  - Bit rate hint boxes available
  - File format concepts do not allow incremental generation and consumption
- MP4 is not the successor of MP3 ;)

Video Formats & Containers: Matroska

- Open & patent free container format
  - For streaming and storage
- Matroska is derived from matryoshka (**Матрёшка**) == „nesting doll"
- File extensions: .mkv .mka .mks

Video Containers & Formats: MPEG

- MPEG-1
  - Meant for Video CDs
- MPEG-2
  - Meant for high resolution materials
  - Used for DVDs & DVB, main cash-cow of MPEG
- MPEG-4 SP & ASP
  - First try on a better format
- AVC / H.264
  - Adapted from H.264 / ITU, used for e.g. Blue-Rays, etc.
- HEVC (High Efficiency Video Coding)
  - Supports 8K Ultra HD

Video Containers & Formats: WebM

- Royalty free online video format
  - Matroska container
  - VP9 video codec
  - Opus audio stream
- Native support in
  - Chrome, Firefox, Opera

Video Containers & Formats: Xvid & DivX

- Both are basically MPEG SP/ASP implementations
- DivX is proprietary
  - DivX ;) was hacked from a MS codec
  - Developer got hired by DivX, Inc.
  - Later DivX "learned" AVC, HEVC and uses MKV
- XVid is open source software
  - Typically packed into an AVI container

Other video formats …

- OGG (Vorbis & Theora)
  - Open source, patent free
  - Used in CC, Wikimedia, many game engines
- Bink (RAD Game Tools)
  - Proprietary, used in games, library & effects
- FLV – Flash Video
  - Adobe container for Sorenson, VP6 & AVC
- MXF
  - Professional container for B2B
- Real Media
  - Streaming server & player along with format

Demo: FFMPEG CMD-Line

- Show ffmpeg –i
- Show ffmpeg –i …. -> webm
- Show ffmpeg -i ED_HD.avi -s 640x360 -b 1024k ed.webm

Exercise: Transcode a video for the Web

- Get ffmpeg
  - http://ffmpeg.zeranoe.com/builds/ for windows
  - http://www.evermeet.cx/ffmpeg/ for Mac
- Create a video (mobile phone, etc.)
- Convert the video to WebM
  - using ffmpeg
- Create a web page featuring your video

Exercise: Transcode a video for the Web

- Consult slides of the course on HTML 5
- Search Google for "ffmpeg how convert webm"
- Add 500 characters to your report describing your experience.
- Deadline (for all exercises) Jan 19, 2014
- Test (multiple choice) Jan 21, 2014

Multimedia Networking

- Applications
  - Streaming stored data
  - Streaming live data
  - Real time interactive video

- Characteristics
  - Delay sensitive
  - Loss tolerant

Internet Multimedia: Download & Play

- Data stored in file
- File transferred via HTTP
  - Received at the client
  - Passed on to the player
- That's not streaming
  - Long start-up delays, etc.

Internet Multimedia: Progressive Download

- Video URL is handled by browser, plugin or video player
- Server is contacted via TCP/HTTP
- Streaming-like experience using HTTP-GET
- Player starts as soon as there is enough downloaded
- Stops or pauses if there is not enough data

Internet Multimedia: Streaming

- Non-HTTP protocol between server and media player
- Can use e.g. RTSP+RTP+UDP (see later)
  - Each frame is packaged in a UDP datagram
  - If one is lost the frame is lost
  - Client takes care of package order
  - Pause, forward, rewind are supported

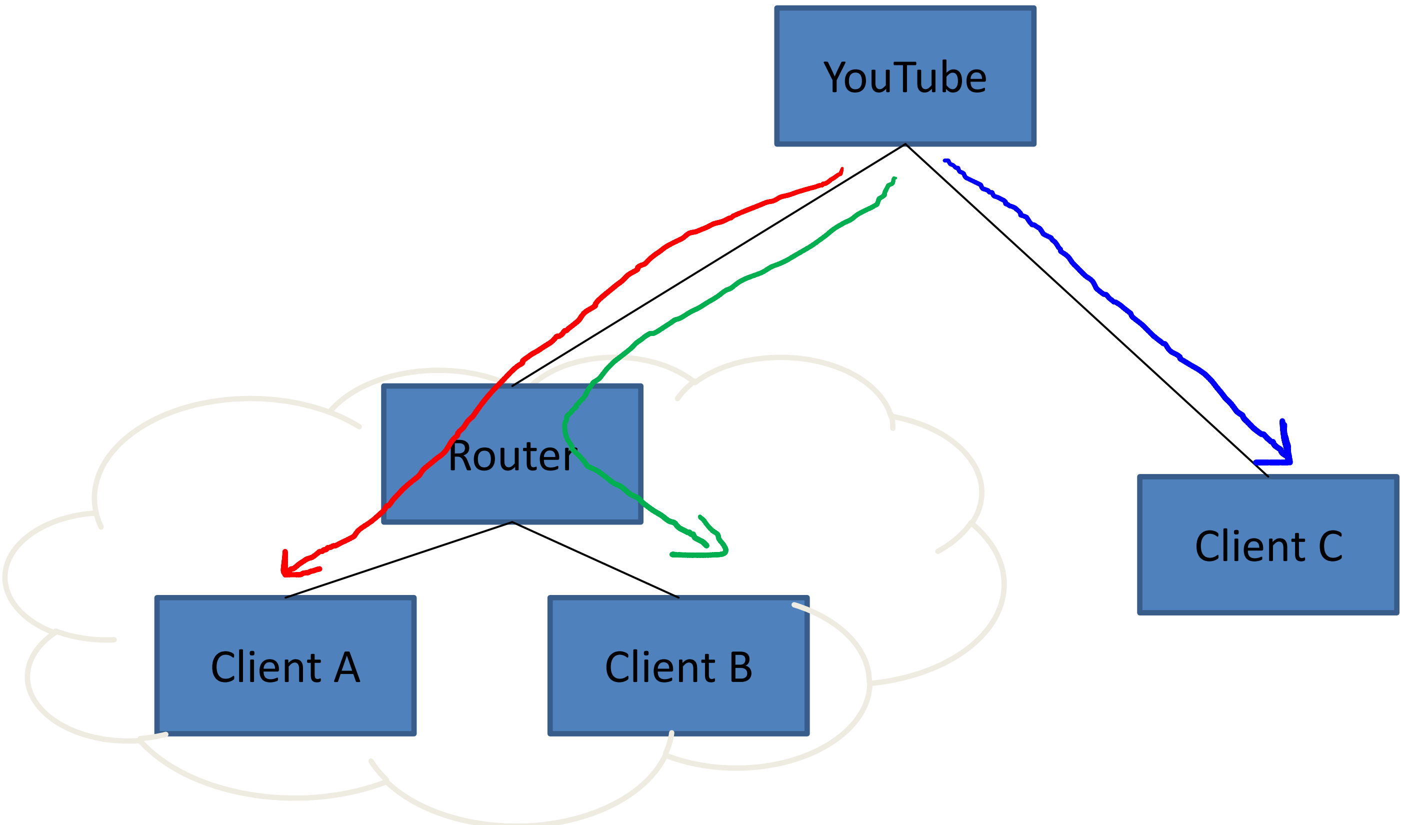
Internet Multimedia: Current Approaches

- HTTP pseudo-streaming
  - Like progressive download
  - But seeking is supported (jump 2 time)
- HTTP dynamic/live streaming
  - Server provides many small files
  - In different resolutions / bitrates
  - Client glues files together for full video
- RTMP Streaming
  - Proprietary Adobe protocol

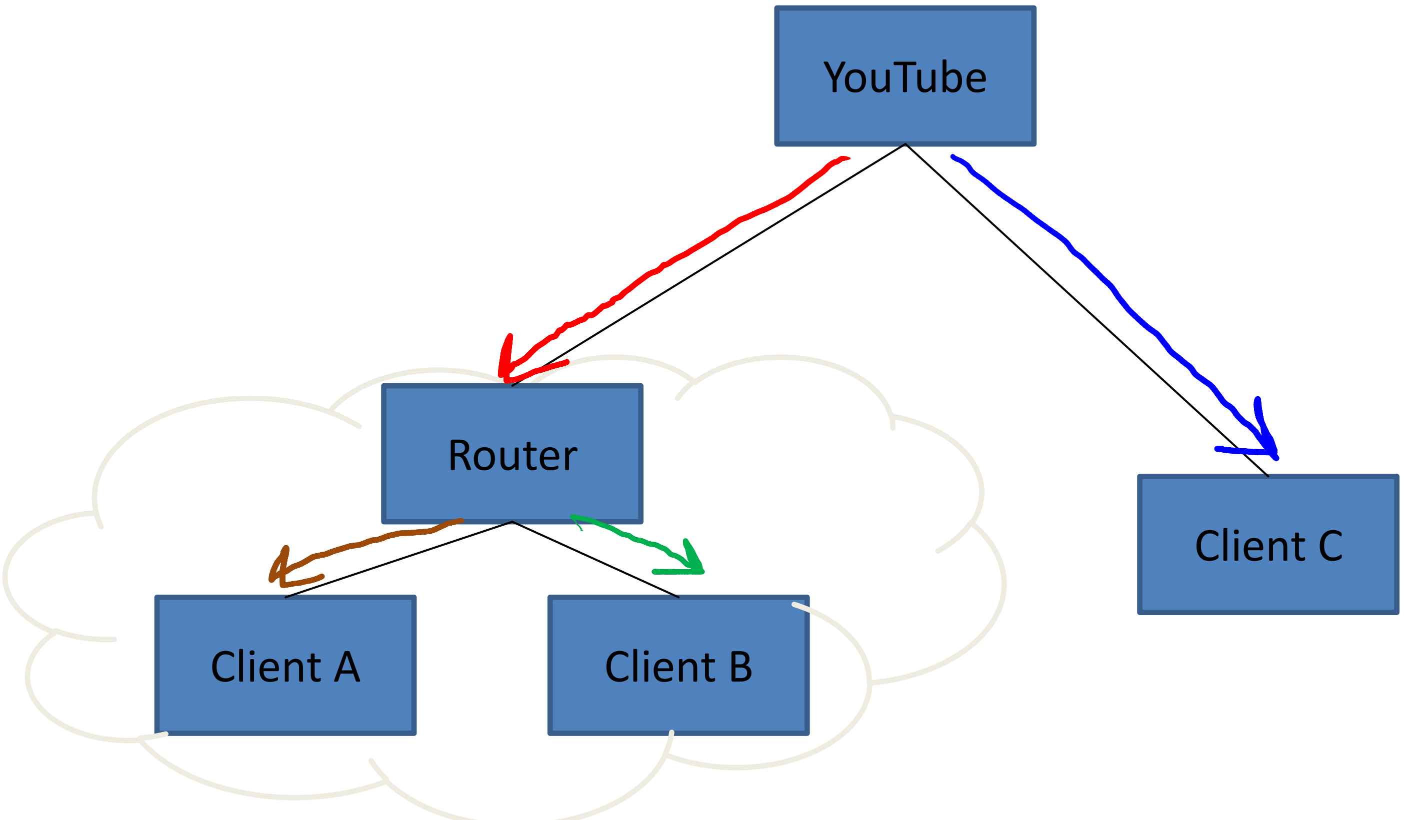
MPEG DASH

- Server has for each video
  - lots of small chunks, ie. 1 sec each
  - pre-transcoded in different sizes & bit rates
- Client gets Media Presentation Description
  - knows which chunks are available
  - requests what is appropriate
- Features:
  - stream & bitrate switching, HTTP transport, …

Multimedia Networking: Multicast vs. Unicast



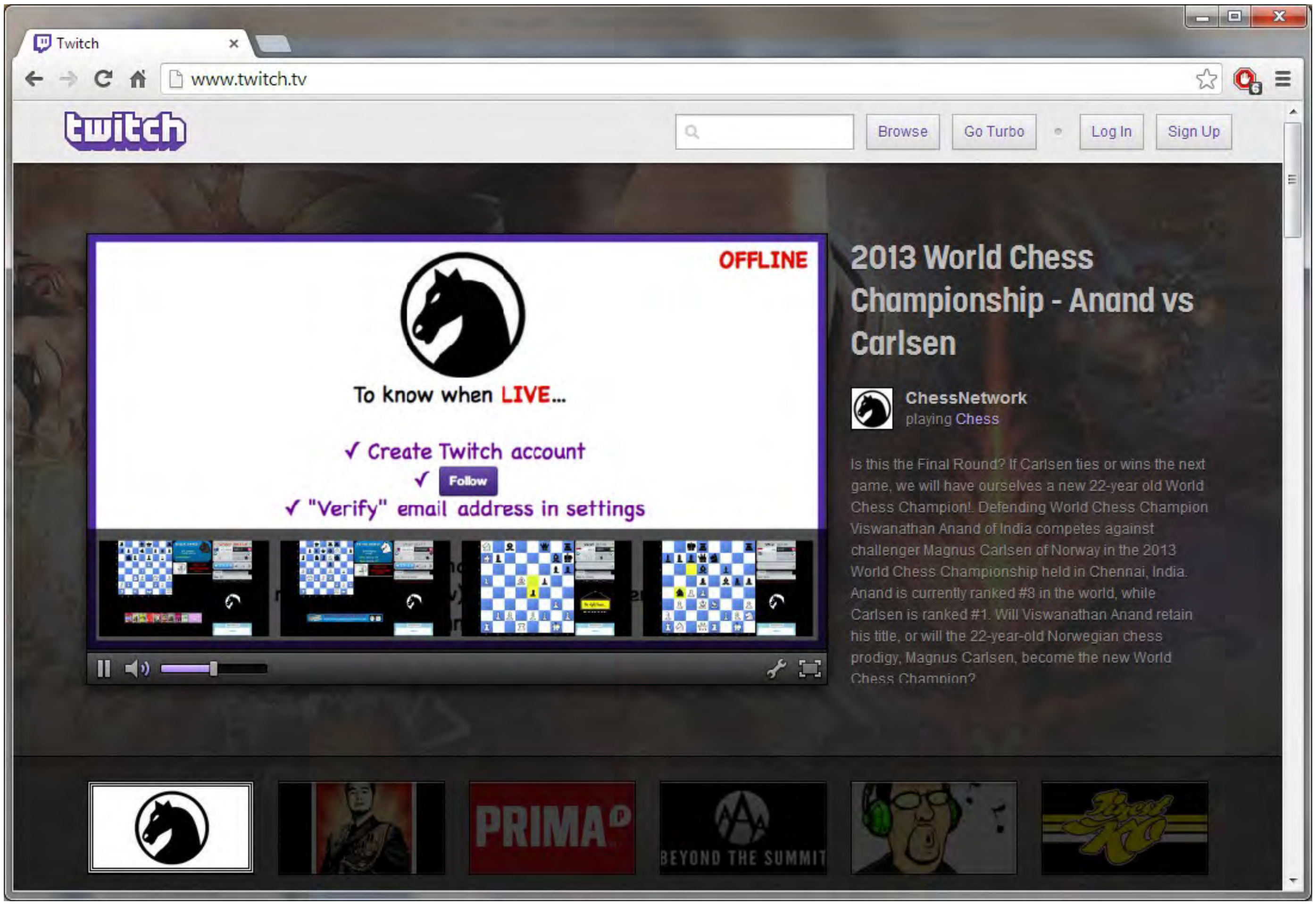
Multimedia Networking: Multicast vs. Unicast



Why Streaming?

- DRM & Security in …
  - IP-TV & Interactive TV: Aon TV, etc.
  - Movie & video rentals: Netflix, Xbox Video Store
- Real time services like
  - VoIP, video conferencing
- Live streaming
  - Sports event, etc.
- Multicast: Delivery to many clients at once

Demo: Twitch.tv



DVB

- DVB := Digital Video Broadcast
- European standardization and development effort
- Suite of open standards for digital television since 1990s
  - Based on existing standards

DVB

- DVB Project (www.dvb.org)
  - Industry-led consortium of 280 companies from 35 countries(broadcaster, network operator, manufacturer)
  - Currently more than 200 million DVB receivers deployed
- European Telecommunications Standards Institute (ETSI)
  - Founded in 1988, now 740 members
  - Independent, non-profit standardization organization

DVB - Transmission

- Common source coding and multiplexing
- DVB-S (1994)
  - DVB-S2, DVB-SH
  - 107 million receivers deployed
- DVB-T (1997)
  - DVB-T2, DVB-H
  - 81 million receivers deployed
- DVB-C (1998)
  - DVB-C2
  - 42 million receivers deployed

DVB-S

- Broadcast by geostationary satellites
- Coexistence with existing analog programmes
- Very popular in Europe
  - Inexpensive equipment (< 100 €)
  - No follow-up expenses
  - Coverage of sparse-populated areas
  - Variety of different channels (free to air, FTA)
  - Astra and Hotbird satellites

DVB-S

- Geostationary Communication Satellites
  - Appear at a fixed position for an earth-based observer
  - Moves with the same speed as earth rotation
  - Orbital position of about 36000 km above equator
  - Comparison:
    1/10 way to the moon
    ISS or Space Shuttle - orbit of 400 km
- Satellites can be positioned at different longitudes, e.g.
  - Astra at 19.2° East
  - Hotbird at 13° East

DLNA / UPNP

- UPnP is set of networking protocols
  - For residential networks
  - Without administrator
  - Devices discover compatible neighbors
  - Function services are established automatically
- UPnP AV
  - Extension to AV data
  - Definition of roles & functions

DLNA

- DLNA := Digital Living Network Alliance
- Non profit organization, > 250 members
- Common goal of
  - Using standard based technology to
  - Make it consumption of multimedia easier
  - Digital photos, videos and music
- > 8,000 devices certified
  - As of May 2010

DLNA – Home Network Devices

- Media Server
  - Storage and delivery of AV content
- Media Player
  - Access servers, playback and render
- Media Renderer
  - Render content, controlled by controller
- Media Controller
  - Access & find servers, control renderer
- Media Printer

DLNA Setup?

tversity

Introduction to Media Informatics: Audio

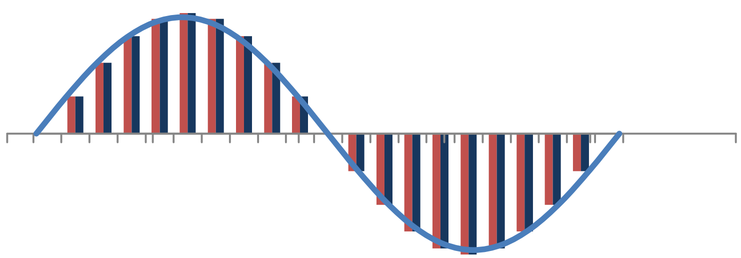
Dr. Mathias Lux
Alpen-Adria Universität Klagenfurt

Agenda

- Audio Basics
- Audio Visualization

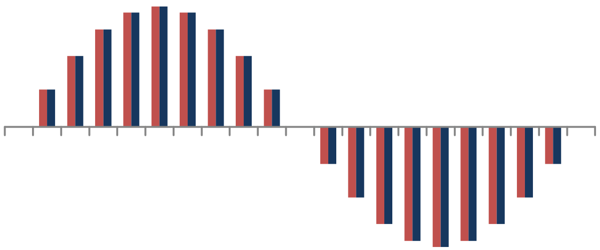
What is digital sound?

- A digitization of the wave.
  - Either a recipe for reconstruction
  - Or a discrete approximation



Sampled sound

- Wave gets sampled x times a second
  - E.g. 48.000 times -> 48 kHz sampling rate
- Obtained values are stored
  - E.g. 256, 240, 13, -7, -12, -44, ….
  - Quantization to e.g. 2^8 levels -> 8 Bit
- Possibly from different sensors
  - Stereo -> 2 channels



Sampled sound

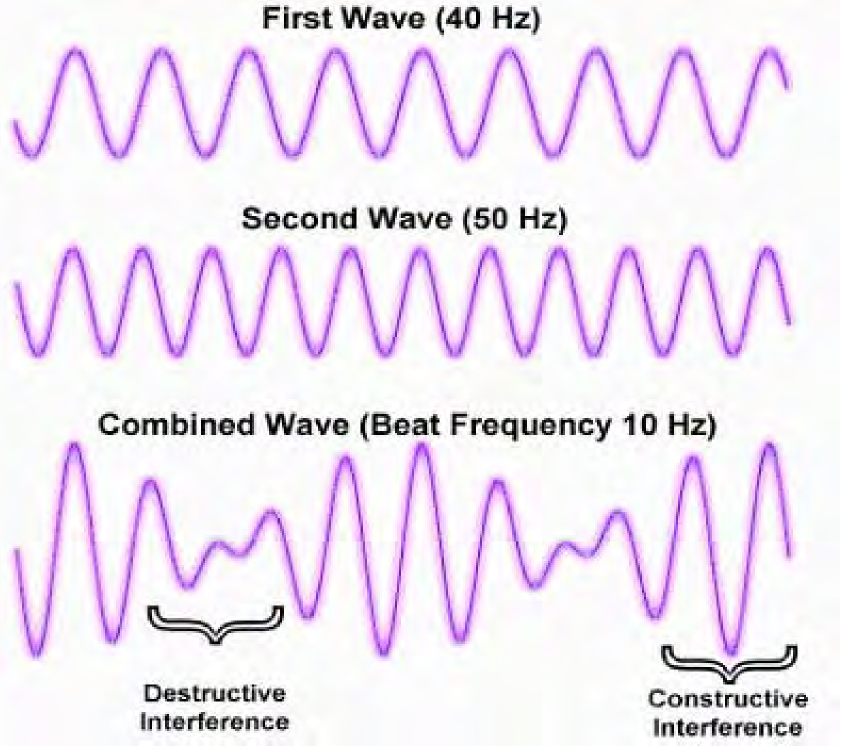
- Example: 8 kHz, 16 bit Stereo
  - Sound wave is sampled 8.000 times a second
  - Samples are stored in 16 bit numbers
- That's *Pulse Code Modulation* (PCM)
  - Often used in WAV files ...
  - Also as input from microphone or line in

What do we want to capture?

- Humans can hear
  - From around 16 - 21 Hz
  - To around 16 kHz - 19kHz
  - 16 bit is enough (CD), 32 bit even better

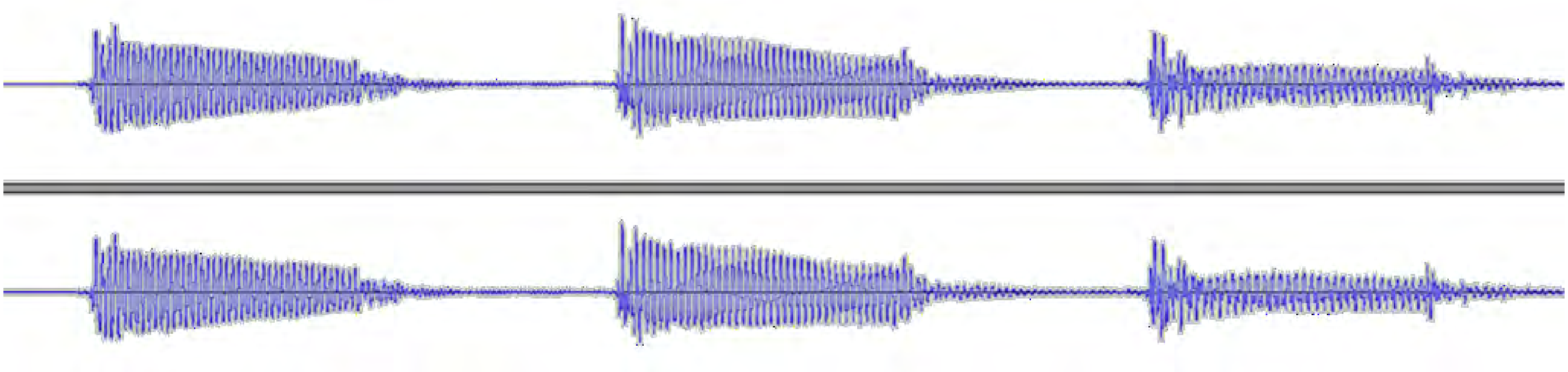
Sound Mixing

- Multiple sounds at the same time?



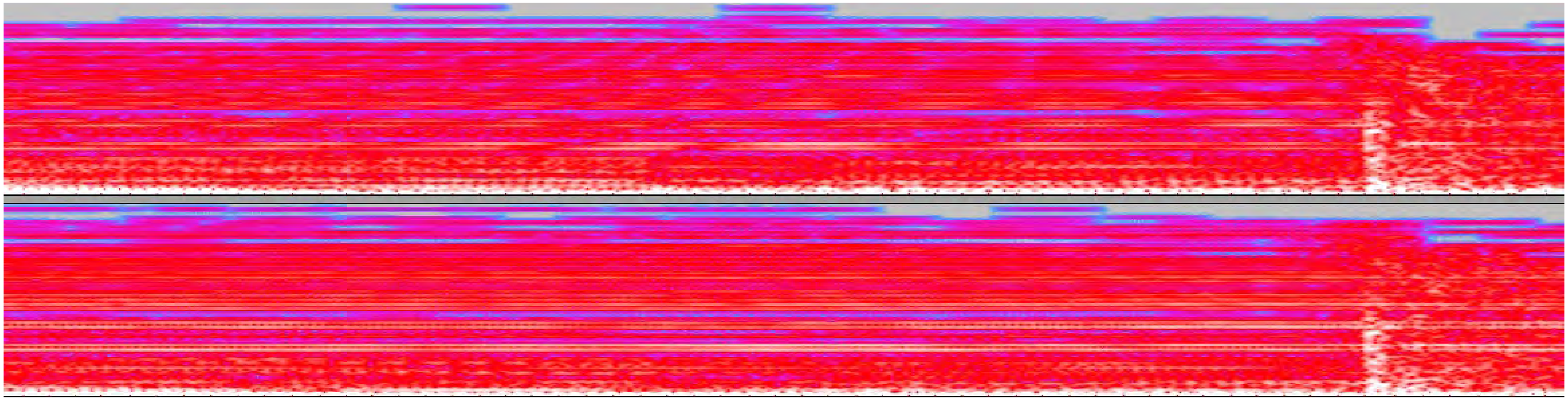
Audio Representation: WaveForm

- Gives the power of a signal of a given time
- Visualization of pulse code modulation



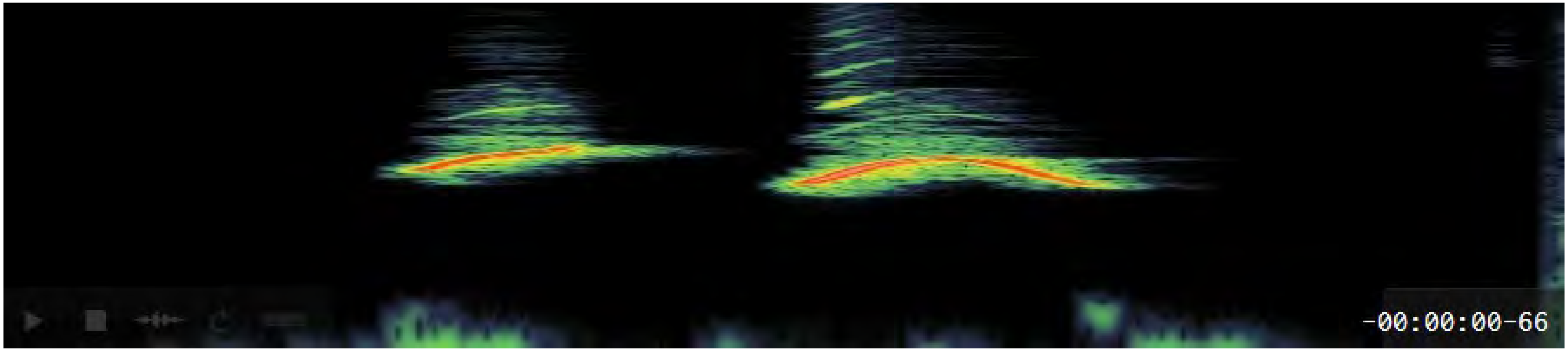
Audio Representation: Spectral

- Gives power in frequency bands as color
  - gray -> blue -> red -> white
- Shows if there are high or low frequencies present.



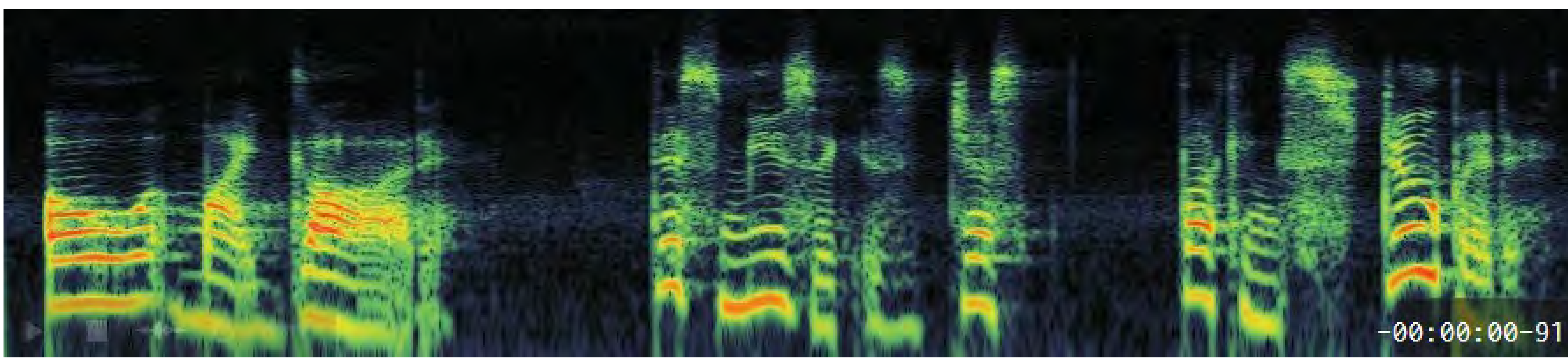
Audio Representation: Spectral

- Whistle
  - http://www.freesound.org/people/THE_bizniss/sounds/39548/



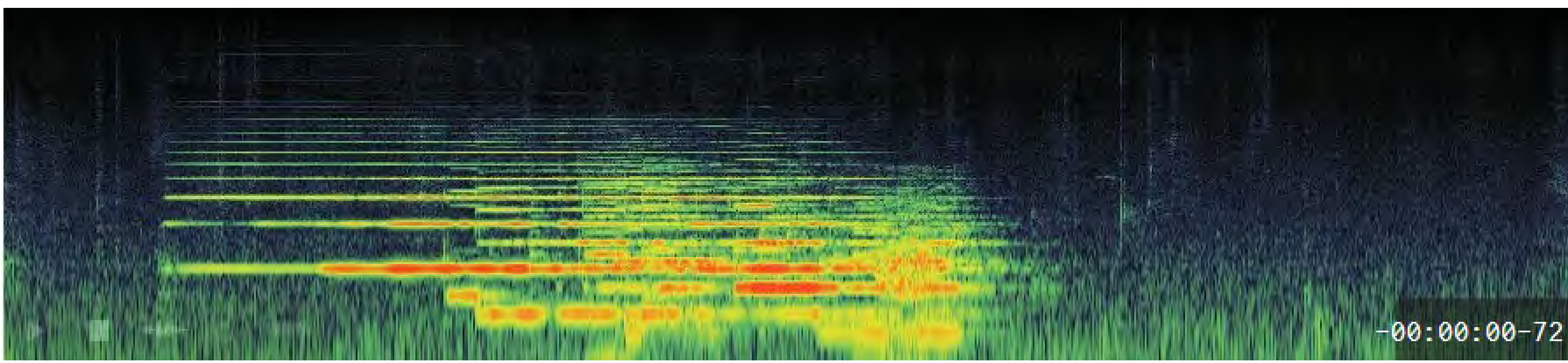
Audio Representation: Spectral

- Voice
  - http://www.freesound.org/people/epanody/sounds/107720/



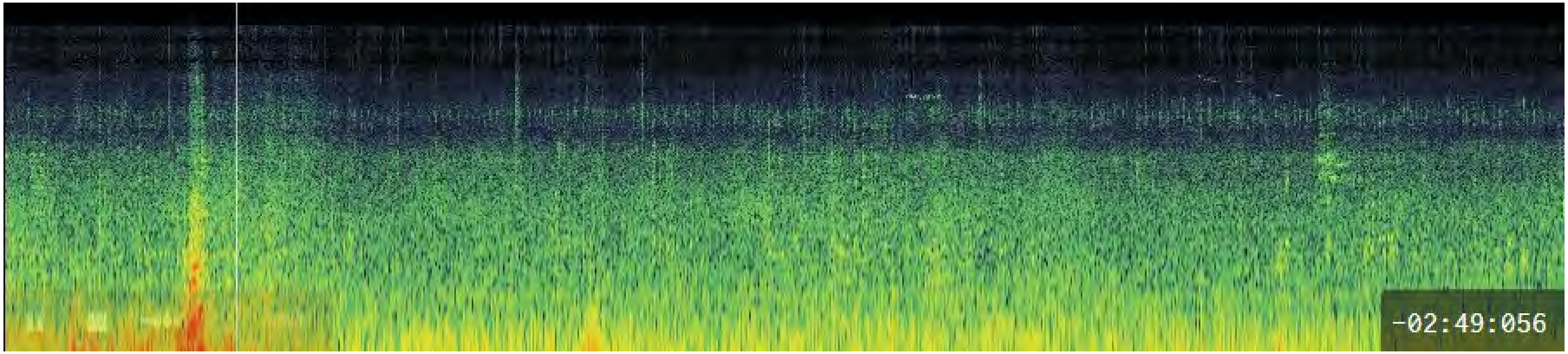
Audio Representation: Spectral

- Music
  - http://www.freesound.org/people/IEDlabs/sounds/82323/



Audio Representation: Spectral

- Environmental
  - http://www.freesound.org/people/sagetyrtle/sounds/36734/



Audio Processing

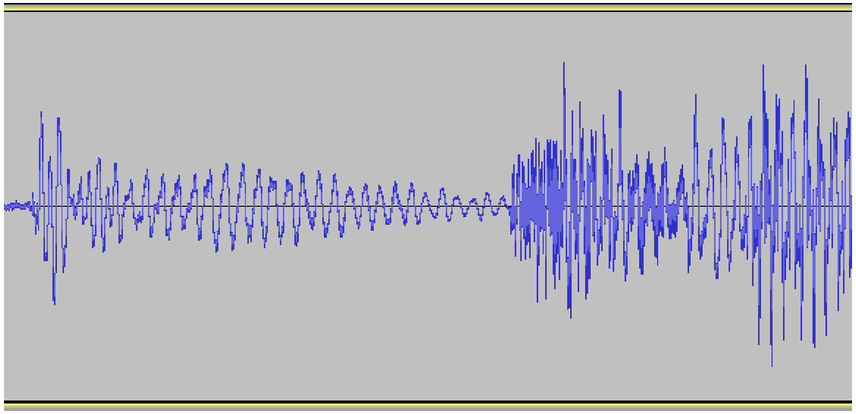
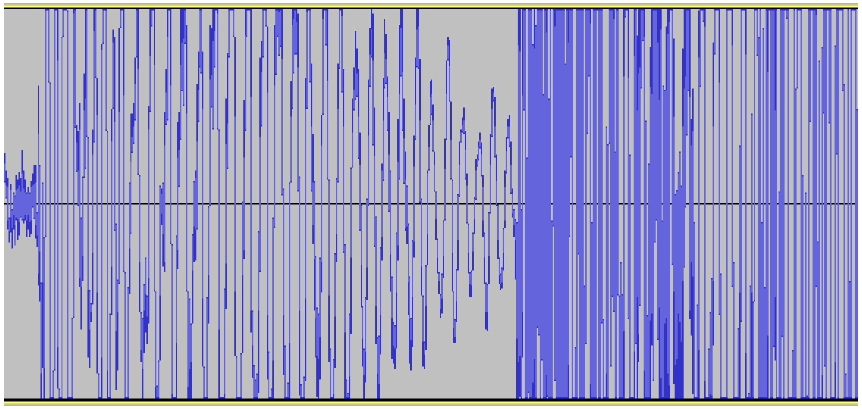
- High Pass Filter
  - Removes low frequencies
- Low Pass Filter
  - Removes high frequencies

- Demo: Audacity

Audio Processing

- Amplify
  - increase power of samples, ie. sample * x
  - may result in <u>clipping</u>

- Demo: Audacity

Audio Processing

- Change sample rate
  - makes sound faster or slower
  - changes frequencies

- Demo: Audacity, change speed filter.

Audio Processing

- Change pitch or tempo
  - sampling rate change does both
  - doing just one is complicated
- First compute frequency domain representation
  - ie. with Short Time Fourier Transform
- Then re-compute frequencies based with different time basis
- Apply inverse Fourier transform

Audio Processing

- Adding Echo
  - Add the same signal in a new track and reduce amplitude
  - Mix with original track
- Can be done online
  - ie. computer games, 3D sound

Sound Formats

- Waveform Audio Format
  - Container for several compression formats
  - Includes PCM, MP3, GSM, □-Law
- Musical Instrument Digital Interface
  - Control codes for instruments
  - Instruments can be "emulated"
- Compressed Audio Formats
  - MP3, OGG, AAC, …

Audio Compression

- Lossless compression enough?
  - does not work with audio too well
- Multiple options for compression
  - MPEG-1 Layer 3 (MP3), Ogg Vorbis
  - MPEG-4 AAC, Dolby Digital (AC3)
- In General: Psycho-Acoustic Modeling

MP3 Compression

- Developed by MPEG for Audio Compression (1992)
  - Part of MPEG-1 (Layer-3: Audio)
- Based on Psychoacoustic modeling
  - Developed on an empirical basis
  - Masking effects on 27 frequency bands
  - Model details which frequencies mask which

Critical Bandwidth

- If two sounds share the same critical bandwidth, the ear cannot distinguish between them.
- Sound separation is possible if the frequencies are "different enough"

Temporal Masking

- Sudden stimulus makes other sound inaudible.
- Backward masking
  - obscures a sound immediately preceding the masker
- Forward masking
  - obscures a sound immediately following the masker

MP3-Kompression

- Masking leads to actual compression
- Redundancies in stereo channels are exploited too
- Compression about 10 : 1

Introduction to Media Informatics: Metadata

Dr. Mathias Lux
Alpen-Adria Universität Klagenfurt

Contents

- **Introduction to Metadata**
- Metadata Formats
  - Media Production
  - Ontologies
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

What is Metadata?

Metadata is Data about Data

Meta^2 data is data about metadata

Metadata Applications

- Retrieval & Browsing
  - No need to download / view the whole video
- Management & Organization
  - Rights, Billing, Ordering, Classification
- Adaptation
  - Transformation to appropriate representation
- Service Description
  - Orchestration, Harmonization, Access
  - On technical and semantic level

Metadata Problems

- Interoperability
  - Complexity & power of metadata models
  - Integration in (different) applications & scenarios
- Preservation
  - Readability in 100, 1000 years
  - Description how to decode …
- Transmission
  - Synchronized, partially, etc.
- Timeliness
  - Changing with audiovisual content while editing?

Aspects of Metadata

- Content Description
- Administrative Aspects
- Quality Metadata
- Legal Metadata
- Technical Metadata

Aspects of Metadata: Content Description

- Agenda
  - Overview on sequence of information to particular topic
- Table of Contents
  - A list of all segments and their position
- Abstract
  - Describes the topic of a content within a few sentences
- Preface
  - Some words of the author …
- Structure
  - For consumption & navigation

Aspects of Metadata: Content Description

- Keywords & index
  - Content description and lookup of concepts
- Summary
  - Overview of the most important aspects
- Literature reference & footnotes
  - Additional material
- Comments
  - For interactive environments
- Categories
  - Conceptual classification in taxonomies (genre etc.)
- Languages
  - Which languages are used / available

Aspects of Metadata: Administrative Metadata

- Associated persons
  - Authors: who created the content
  - Contributors: who contributed to the content
- History of changes
  - Changes in content and metadata
  - with author, date, location and sort of action
- Unique identifier
  - e.g. URI or database id
- Versions
  - Versioning information … related to the history

Aspects of Metadata: Quality Aspects

- Weight
  - Prioritization of segments
- Expiration Date
  - Time period of validity of the content
- Recessions
  - Opinions, arguments from others
- Process description & history
  - Who corrected, translated and approved the content e.g. within an workflow
- Quality Assessment
  - Rating of the (e.g. visual) quality of the content

Aspects of Metatdata: Legal Metadata

- Copyright
  - Person or company legally permitted to sell or trade with the content
- Publishing Date
  - Date when the content has been released to public
- License Model
  - Defines how consumers are allowed to reuse the content

Aspects of Metadata: Technical Metadata

- Standards
  - Description of the standardized structure in which the content and the metadata are stored
- Application/System
  - Application the content and metadata can be / has been processed
- Resolution, compression of pictures or video clips
- Encryption Method
  - In case of encrypted content
- Storage Media
  - On which the content has been stored e.g. CDs, tapes, MO, paper etc.
- Logs
  - Technical history

Contents

- Introduction to Metadata
- **Metadata Formats**
  - Media Production
  - Ontologies
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

XML: Recapitulation

- Header identifying version & coding
- Tree-like structure
- Simple structuring elements
  - Tags & attributes (Markup)
  - Entities
- DTD and XML Schema for model definition
  - DTD is 'simple' and small
  - XML Schema is XML based and rather powerful

Contents

- Introduction to Metadata
- Metadata Formats
  - **Media Production**
  - Ontologies
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

Media Production: Dublin Core

- Aims to provide
  - Common denominator for metadata
  - Simple yet powerful schema
- Dublin Core Metadata Initiative defined
  - 15 elements (author, date, title, type, …)
  - Further refinements (creation date, extent, …)
- Dublin Core does not provide
  - A schema for storage
  - A schema for data types (e.g. dates)

Media Production: EBU P/Meta

- Aims to provide …
  - a universal standard for metadata exchange between professional media organizations
  - a definition of common meaning to the data fields and values that most broadcasters use in order to enable exchange
  - designed for use in a wide range of broadcasting activities
  - both language and system independent
  - a joint development by EBU (European Broadcasting Union) members on a not-for-profit basis
  - a scheme that makes use of other standards where possible, e.g. ISO country codes

Media Production: Other Standards

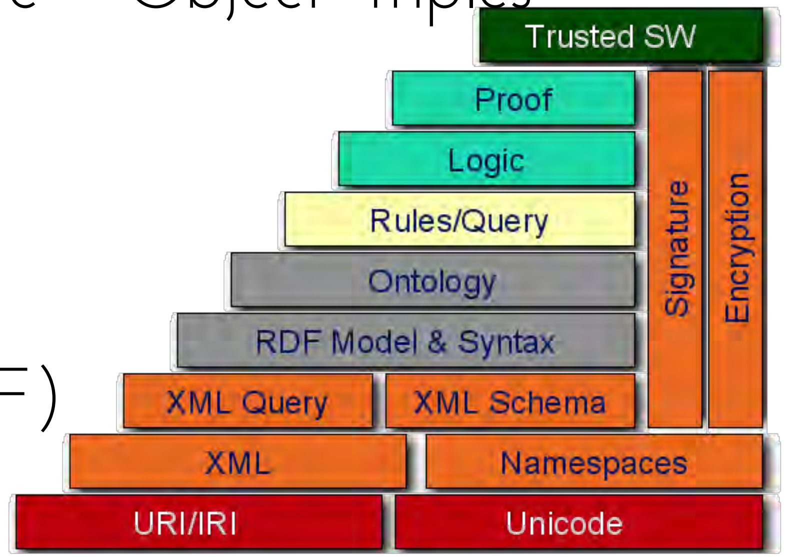
- SMPTE Metadata Dictionary
  - Society of Motion Picture and Television Engineers
    - Since 1916, 61 members
  - Standard for metadata exchange in TV
  - Defines set of attributes / fields
- MXF DMS-1
  - Metadata bundled with the Material Exchange Format (MXF)
  - Open format for the broadcasting area (SMPTE + EBU)
- Virtually 'no information' about these is available
  - Just for exchange for insiders
  - Might not be royalty free

Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - **Ontologies**
  - Home User
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

Ontologies: RDF

- Metadata Model published by the W3C
  - Reaction on the insufficiency of HTML metadata for search & inference
  - Based on "Subject – Predicate – Object" triples
  - URIs for identifying concepts
  - Spans a directed graph
  - Is used in conjunction with vocabularies (e.g. DC, FOAF)



Ontologies: SKOS

- Simple Knowledge Organization System
  - RDF Vocabulary for KOS
- Knowledge Organization Systems are
  - Taxonomies, Thesaurii, Classification Schemes, etc.
- Can be used to organize multimedia data

Ontologies MMSEM

- Multimedia Semantics : Incubator Activity of the W3C
  - Closed Aug. 2007

Deliverables:

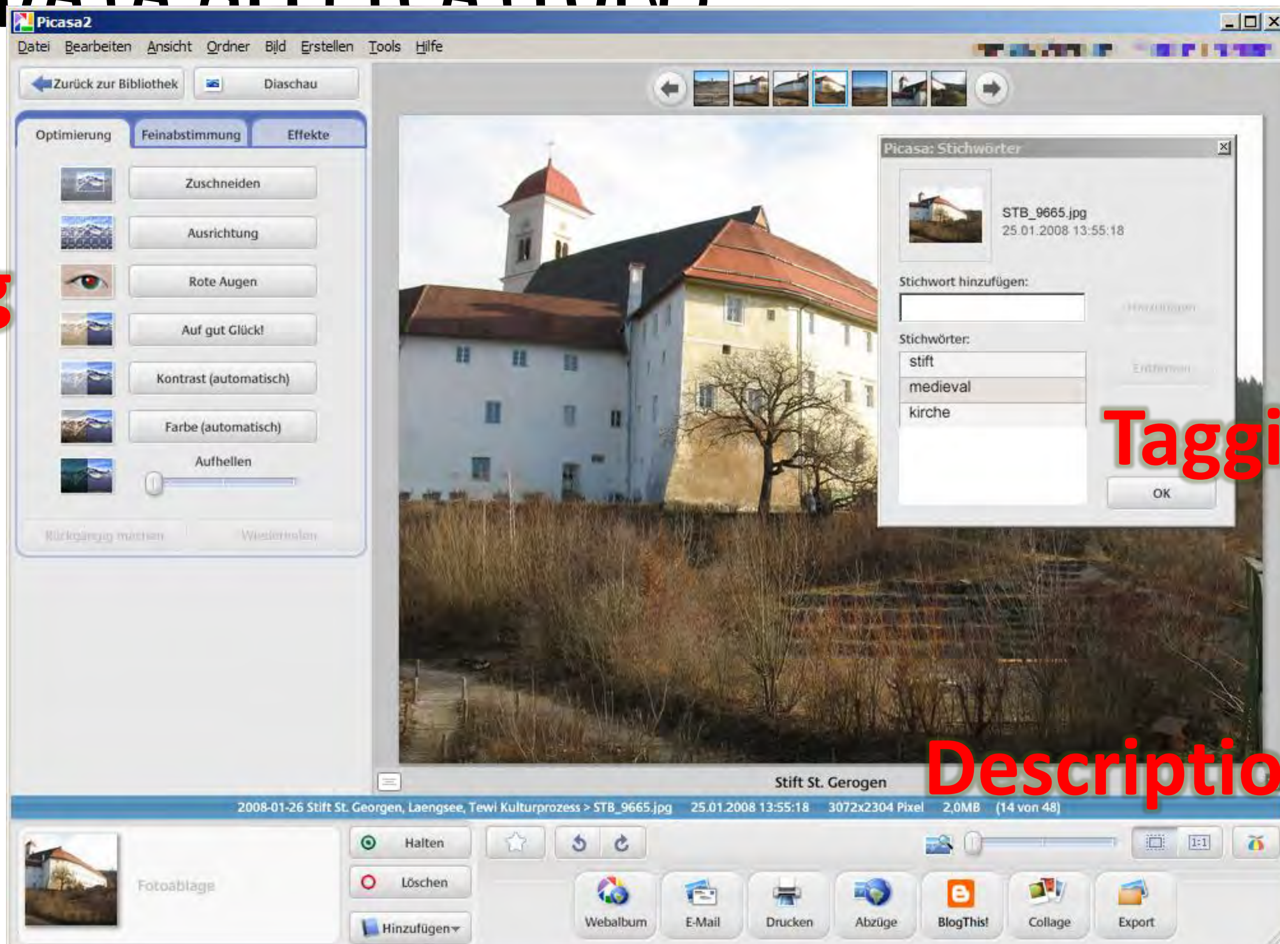
- Image Annotation on the Semantic Web.
  - use cases and general discussion about Semantic Web vocabularies and tools
- Multimedia Annotation Interoperability Framework.
  - a bottom-up approach to provide a simple extensible framework to improve interoperability
- MPEG-7 and the Semantic Web.
  - four current OWL/RDF proposals of MPEG-7, as well as a comparison of the different modeling approaches in the context of practical applications.

Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - Ontologies
  - **Home User**
- MPEG-7
- MPEG-21
- Metadata Generation & Annotation

Home User: Metadata Applications

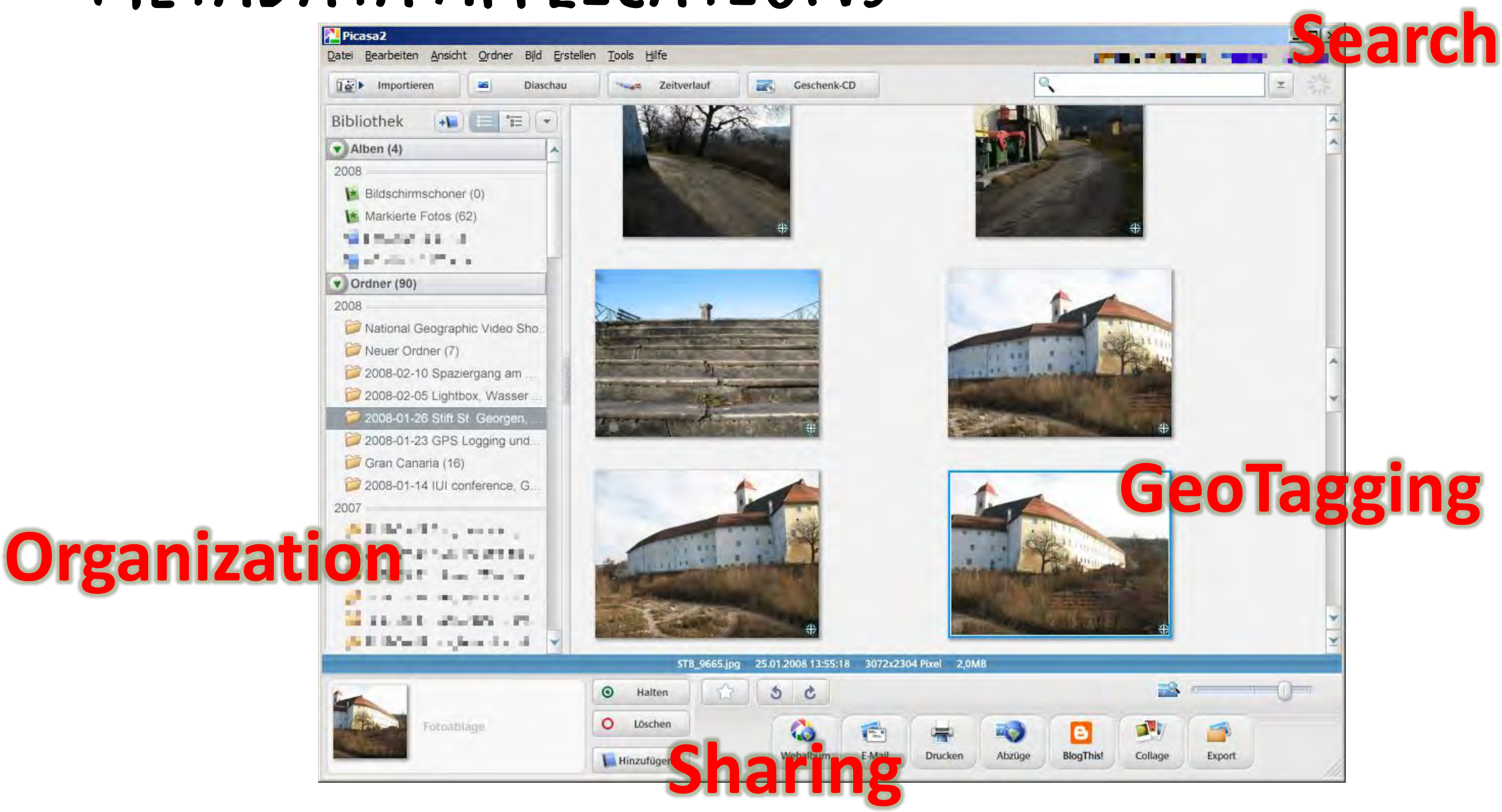
Editing



Tagging

Descriptions

Home User: Metadata Applications



Home User: EXIF

- Exchangeable Image File Format (EXIF)
  - Japan Electronic and Information Technology Industries Association (JEITA)
  - Extensive format for technical aspects
  - Settings and sensor readings at the time of recording
  - Mostly images from digital cameras

EXIF - Example

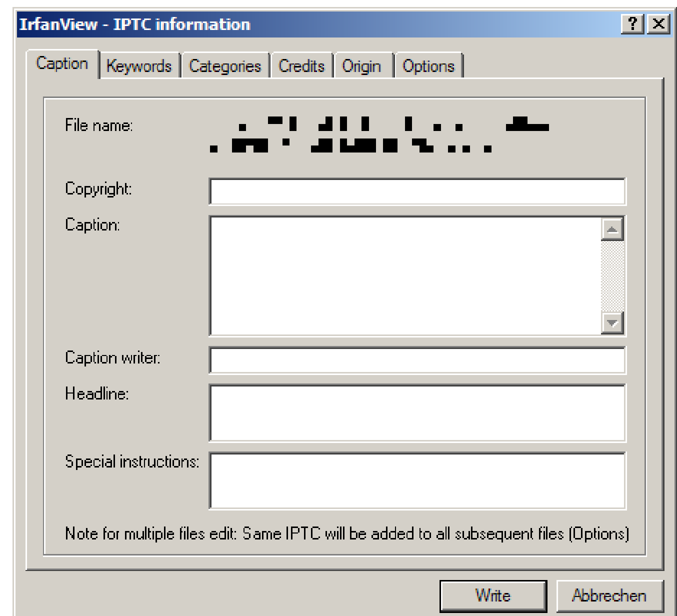
Make - Canon
Model - Canon PowerShot A620
Orientation - Top left
XResolution - 180
YResolution - 180
ResolutionUnit - Inch
DateTime - 2008:02:10 15:44:58
YCbCrPositioning - Centered
ExifOffset - 198
ExposureTime - 1/200 seconds
FNumber - 2.80
ExifVersion - 0220
DateTimeOriginal - 2008:02:10 15:44:58
DateTimeDigitized - 2008:02:10 15:44:58
ComponentsConfiguration - YCbCr
CompressedBitsPerPixel - 5 (bits/pixel)
ShutterSpeedValue - 1/202 seconds
ApertureValue - F 2.80
ExposureBiasValue - 0.00
MaxApertureValue - F 2.80

GPS information: -
GPSVersionID - 2.2.0.0
GPSLatitudeRef - N
GPSLatitude - 46 40 41.41
GPSLongitudeRef - E
GPSLongitude - 13 58 22.17
GPSAltitudeRef - Sea level
GPSAltitude - 503 m
GPSTimeStamp - 14 44 58

Maker Note (Vendor): -
Macro mode - Normal
Self timer - Off
Quality - Superfine
Flash mode - Auto + red-eye reduction
Sequence mode - Single or Timer
Focus mode - Single
Image size - Large
Easy shooting mode - Portrait
Digital zoom - None

Home User: IPTC

- IPTC Information Interchange Model (IIM)
  - Several elements to describe images (assets)
  - Rather common format
    - Adobe Bridge / Photoshop
    - Google Picasa
    - Irfanview …
  - Like a *predefined metadata form ->*



Home User

- eXtensible Metadata Platform (XMP)
  - Initiative from Adobe
  - Based on RDF, embedded in document
  - Also used in PDF, AI, PSD, etc.
- ID3
  - Metadata for MP3, spread by popular players
  - Two versions …
    - v1: 128 Byte block coding some fields at end of file
    - v2: Several optional tags inside stream

Broadcasting + iTV

- Electronic Program Guide (EPG)
  - In use in conjunction with DVB
  - Simple format in additional stream
- Multimedia Home Platform (MHP)
  - In use in Austrian DVB-T
  - Proprietary format for data + function
  - Based on Java

Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - Ontologies
  - Home User
- **MPEG-7**
- MPEG-21
- Metadata Generation & Annotation

MPEG-7

- ISO/IEC Standard: **Multimedia Content Description Interface**
- Moving Pictures Expert Group
  - Specification goes on ...
- It's based on XML (Schema)
  - Binary representations possible (BiM)
- Allows differing granularity of descriptions
  - Extensive to very simple

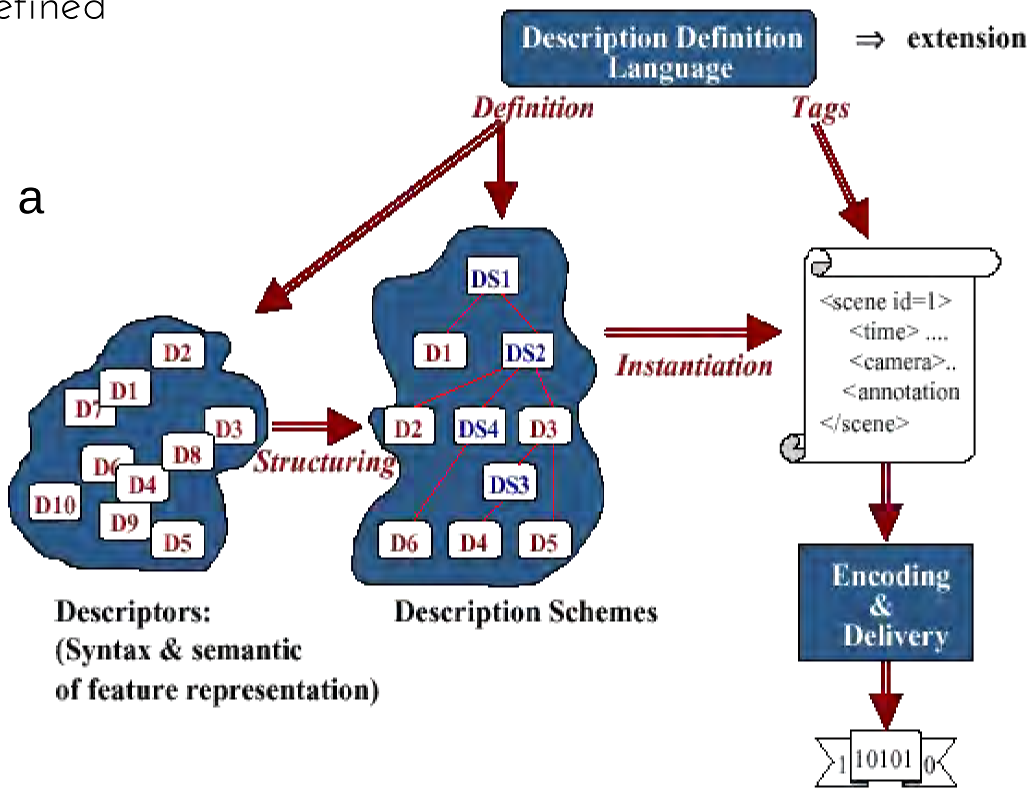
MPEG-7 History

- Call for Proposals: October 1998
- Evaluation: February 1999
- First version of Working Draft (WD): December 1999
- Committee Draft (CD): October 2000
- Final Committee Draft (FCD): February 2001
- Final Draft International Standard (FDIS): July 2001
- International Standard (IS): September 2001

MPEG-7 Basics

- Descriptors
  - Syntax and semantics of exactly one (low or high level) elementary feature
  - Also base data types are defined

- Description Schemes
  - Defines structures within a framework
- Description Definition Language (DDL)
  - Extension of XML Schemes
- Coding Schemes
  - Create and interpret descriptions in BiM



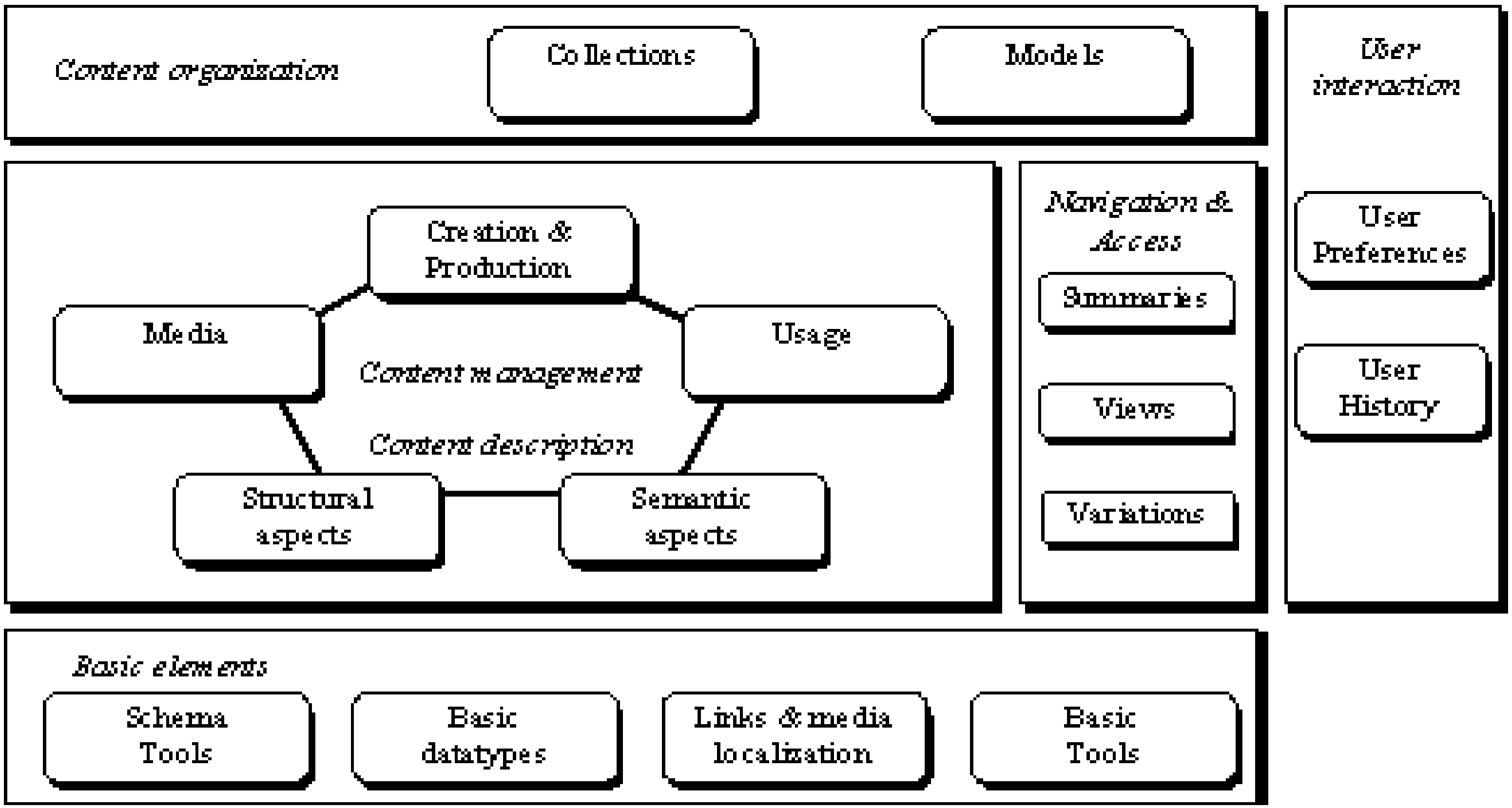
MPEG-7 Parts

1. MPEG-7 Systems
   - Tools needed to prepare MPEG-7 descriptions for efficient transport and storage and the terminal architecture.
2. MPEG-7 Description Definition Language
   - Language for defining the syntax of the MPEG-7 Description Tools and for defining new Description Schemes.
3. MPEG-7 Visual
   - Description Tools dealing with (only) visual descriptions.
4. MPEG-7 Audio
   - Description Tools dealing with (only) audio descriptions.
5. MPEG-7 Multimedia Description Schemes
   - Description Tools dealing with generic features and multimedia descriptions.

MPEG-7 Parts

6. MPEG-7 Reference Software
   - Implementation of relevant parts of the MPEG-7 Standard with normative status.
7. MPEG-7 Conformance Testing
   - Guidelines and procedures for testing conformance of MPEG-7 implementations
8. MPEG-7 Extraction and Use of Descriptions
   - Informative material about the extraction and use of some of the Description Tools.
9. MPEG-7 Profiles and levels
   - Provides guidelines and standard profiles.
10. MPEG-7 Schema Definition
    - Specifies the schema using the Description Definition Language

Scope of MPEG-7



from: http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

Basic Elements

Basic elements are fundamental constructs and used throughout the whole MPEG-7 description

- Basic datatypes
  - Time and date, relative and absolute
  - Numeric datatypes like matrices and vectors
- Links & Media Localization
  - Interconnections and content linking

Navigation & Access

- Descriptors for Browsing & Retrieval
  - Summaries
  - Partitions (time, space & frequency)
  - Decompositions (time, space & frequency)
  - Variations

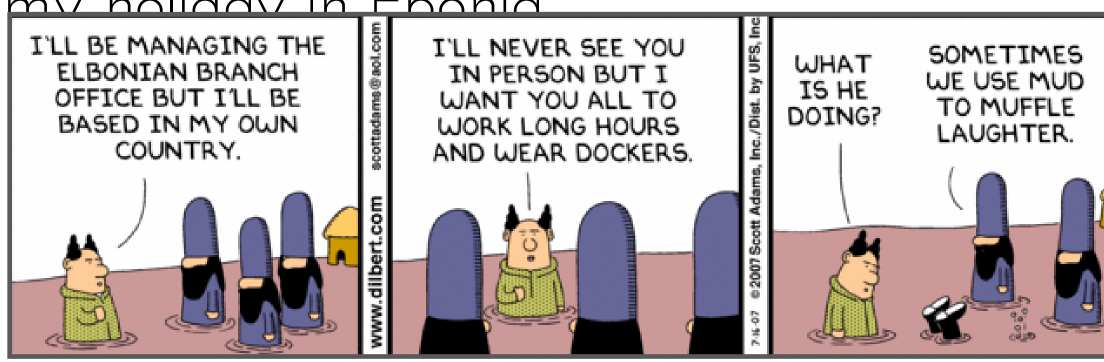
User Interaction

- Pertaining consumption of AV data
  - user preferences
  - usage history

- Meant to facilitate personalization
  - Matching User Interaction DS with content description
  - Is research topic @ ITEC

Content Organization

- Organization & modelling of collections
  - Audio-visual content, segments, events, and/or objects
    - E.g. pictures, scenes, music files, etc.
  - Allows collection description as a whole
    - E.g. "Pictures of my holiday in Elbonia"

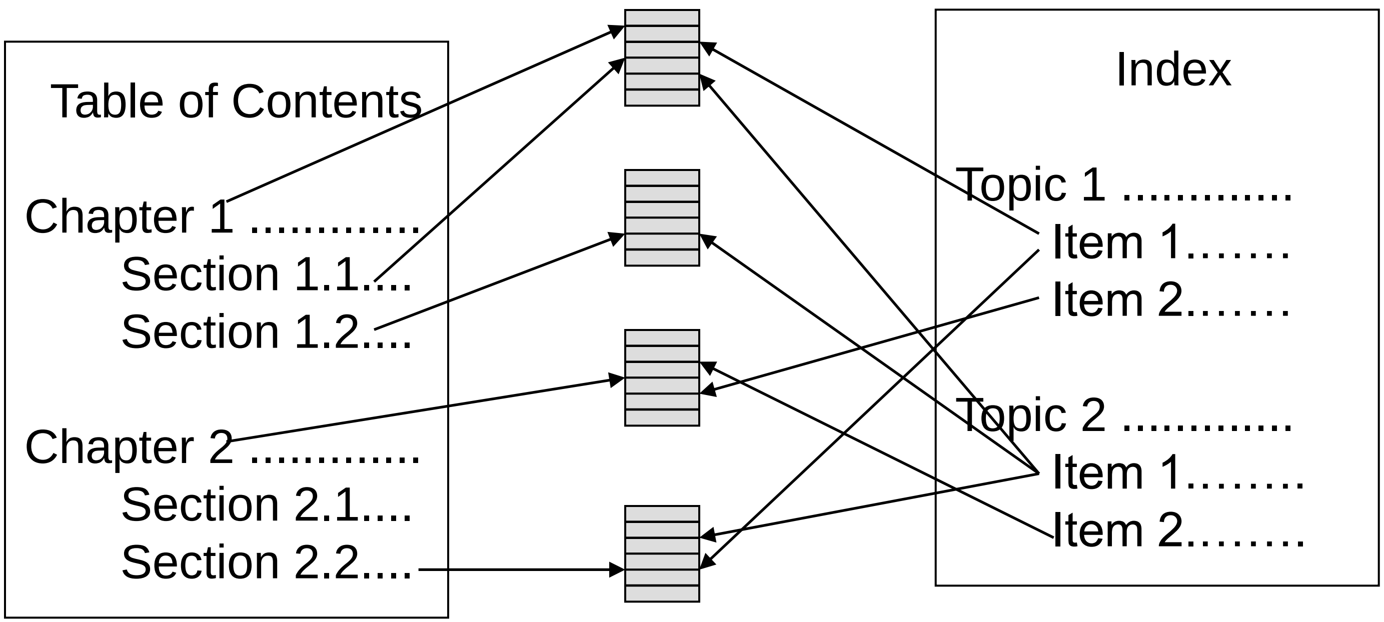


Content Management

- Creation & Classification
  - Title, textual annotation, creators, creation locations, and dates.
  - Categories such as genre, subject, purpose or language.
  - Review and guidance information: Age classification, parental guidance, and subjective review.
  - Related material information.
- Media coding, storage & file formats
  - Media profiles & master media
- Content Usage
  - Usage rights, usage record, and financial information

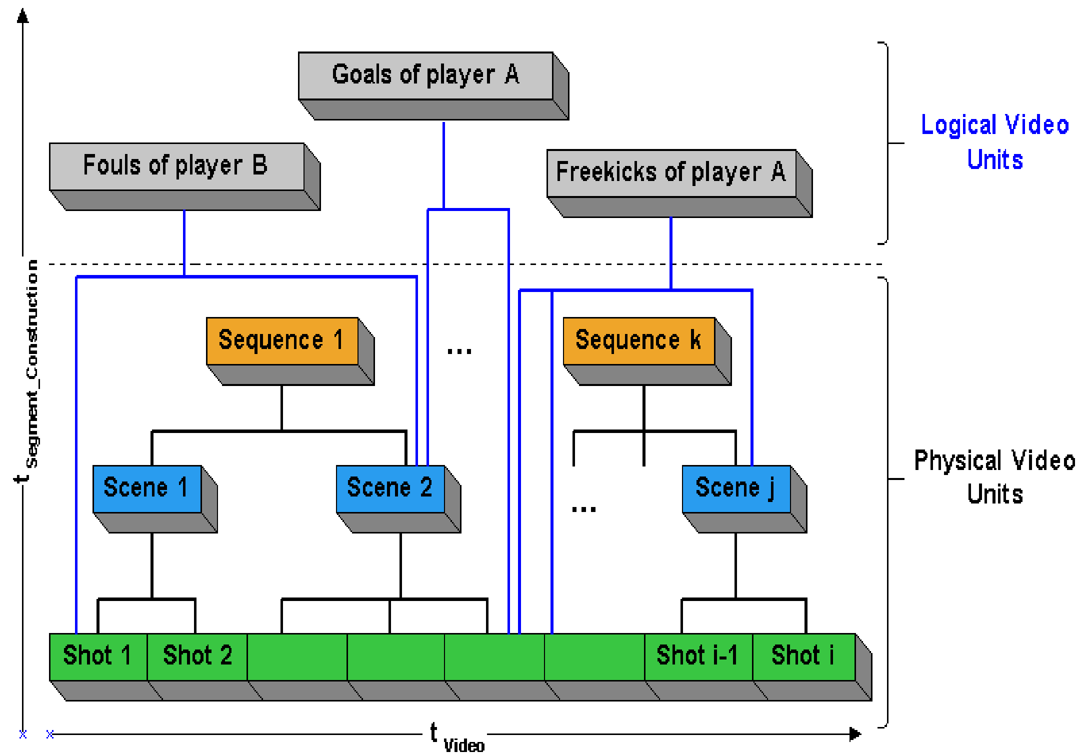
Content Description: Structural vs. Conceptual Aspects

- Program DS (in sense of TV program)
- Analogy to
  - Table of content – Region tree (linear partitioning)
  - Index – Object tree (non-linear structure)

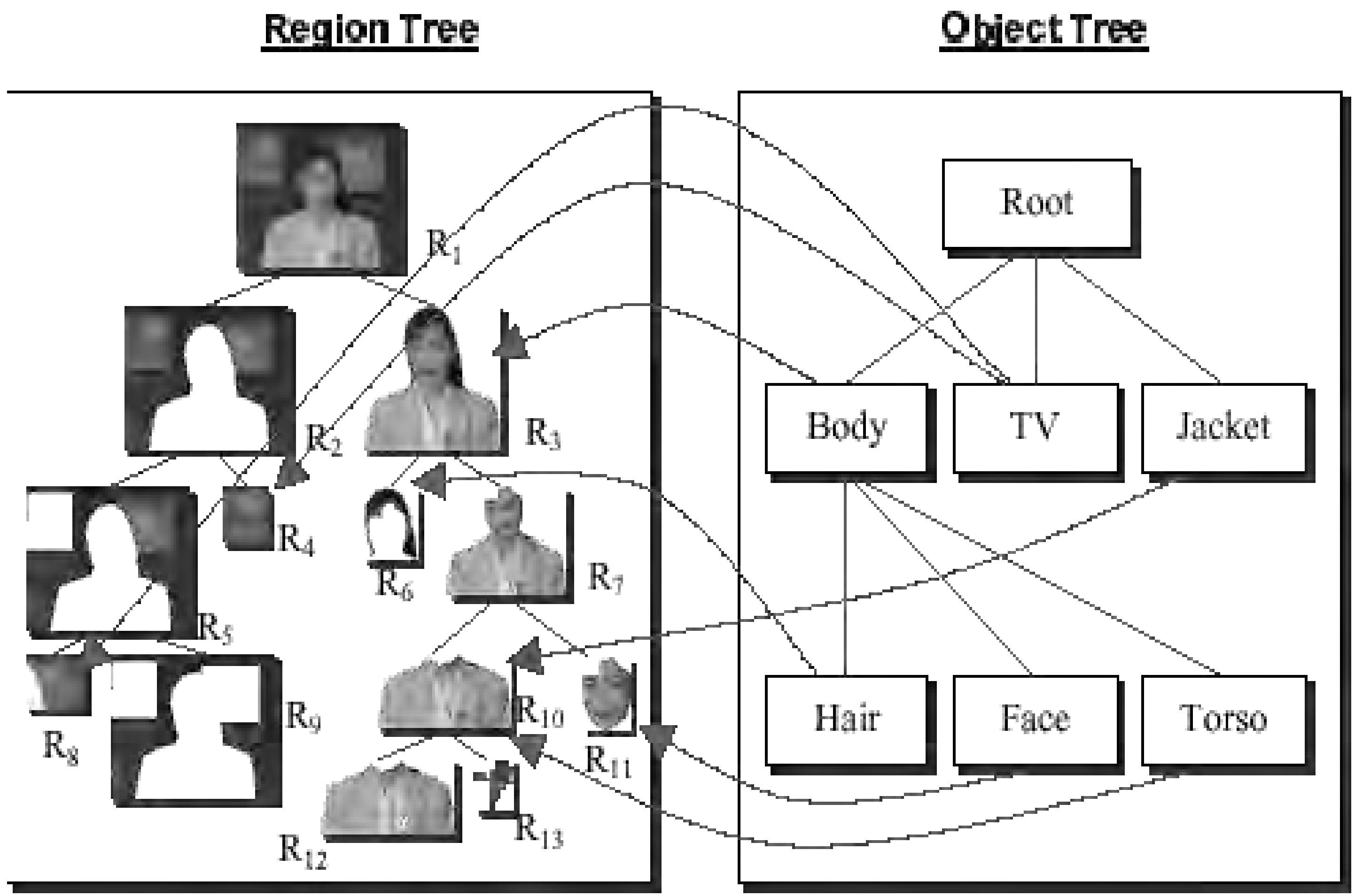


Content Description: Structural Aspects

- Divide a video stream into physical and logical video segments
- The higher the level of a physical video unit, the more semantic information is necessary
- Logical units are based on semantic content



Region and Object Trees



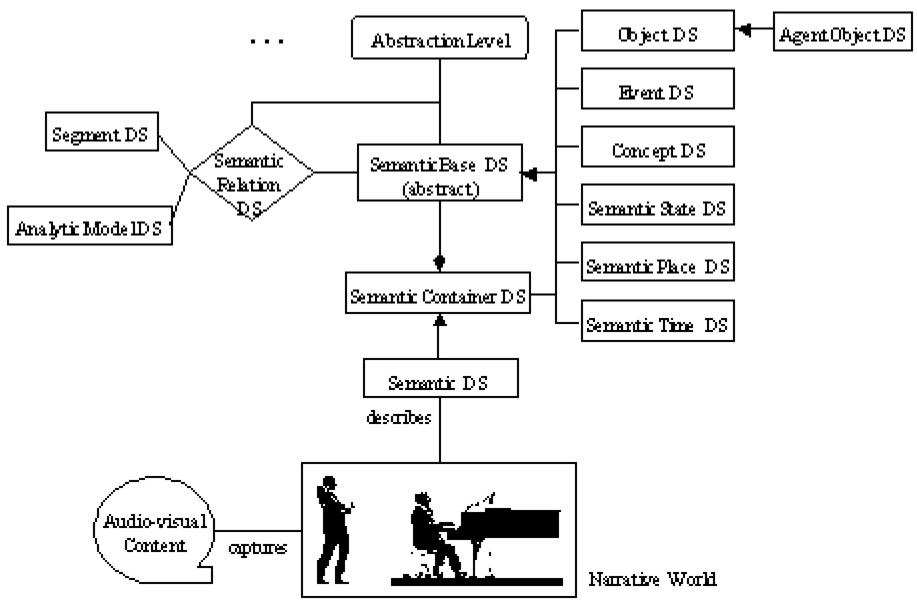
Content Description: Semantic Aspects

- Low Level Features
  - Extraction from Content
  - Descriptors for
    - Shape, color, texture (visual)
    - Timbre, rhythm (audio)
- High Level Features
  - Annotation
  - So called *semantic descriptors*
    - Textual information
    - Conceptual information

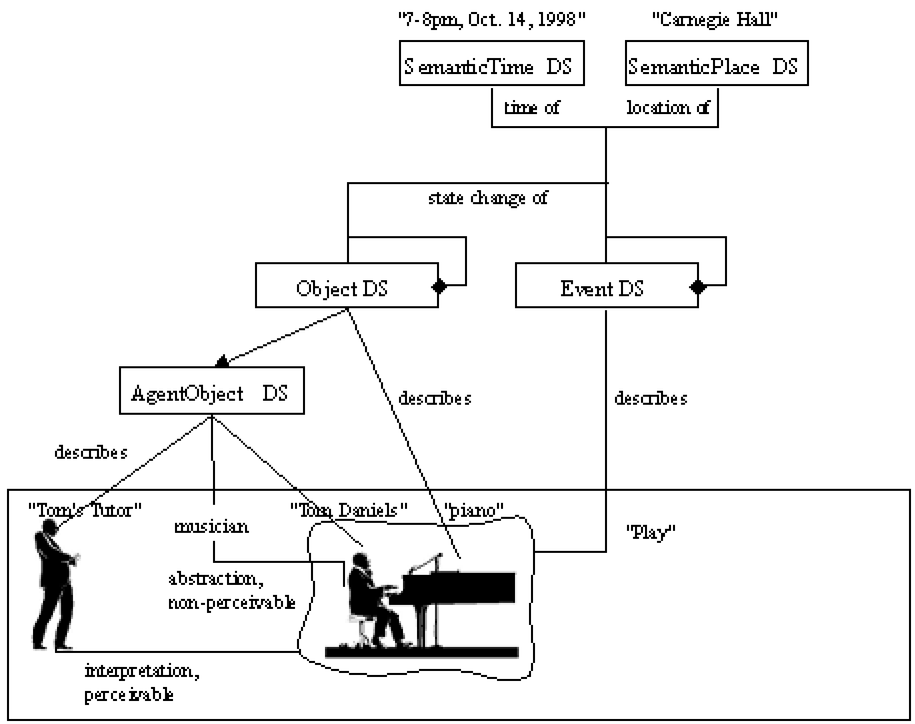
MPEG-7 High Level Descriptors

- Textual Descriptions
  - Text to describe temporal / spatial regions
- The W's
  - Structured way of textual descriptions
    - Who, Where, What Object, When, Why, How
- Instead of textual descriptions
  - Controlled Terms
    - Dictionaries, Taxonomies, Classifications Schemes
  - Semantic Description Scheme

MPEG-7 Semantic Description Scheme



Actual Description in MPEG-7

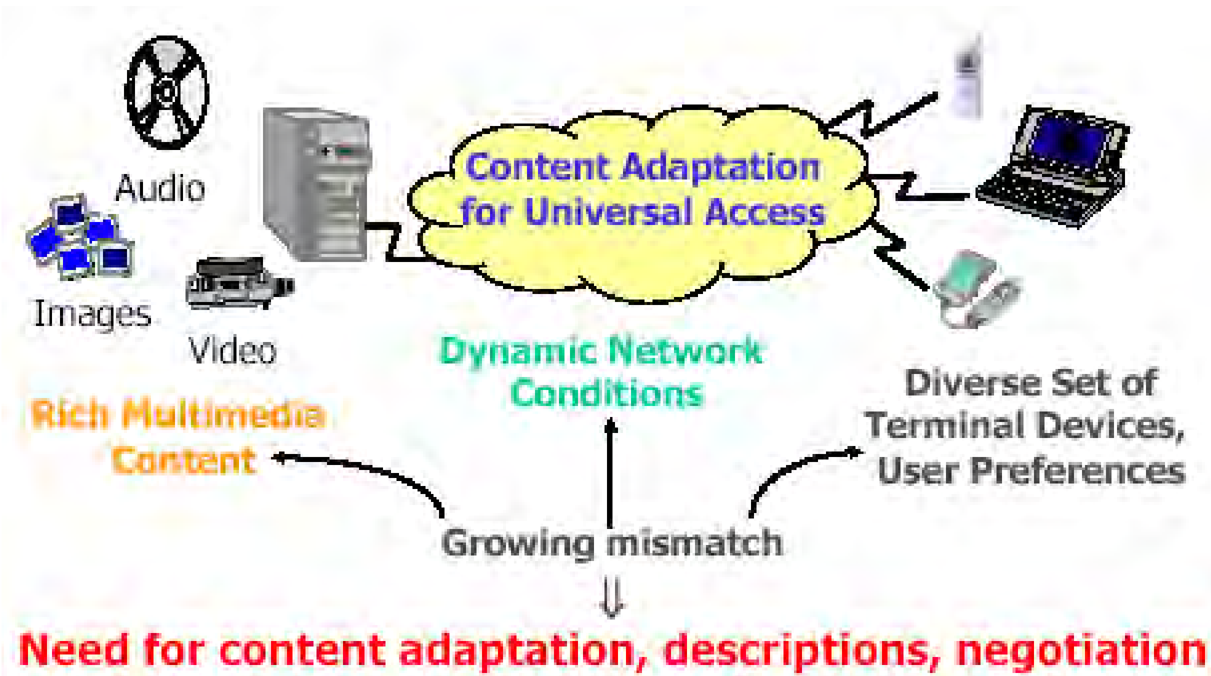


Narrative World

Contents

- Introduction to Metadata
- Metadata Formats
  - Media Production
  - Ontologies
  - Home User
- MPEG-7
- **MPEG-21**
- Metadata Generation & Annotation

MPEG-21 – MOTIVATION AND SCOPE



MPEG-21 Objectives

MPEG-21's goal is to create an *interoperable and integrated multimedia framework* in three steps:

1. **Develop** "**big picture**": understand how the components of the framework are related and identify where gaps in the framework exist
2. **Fill the gaps**: develop new standard specifications where needed
3. **Integrate**: achieve the integration of standards to support harmonized technologies for the management of multimedia content

MPEG-21 Digital Item

A **Digital Item (DI)** is a structured digital object with a standard representation, identification, and metadata within the MPEG-21 framework

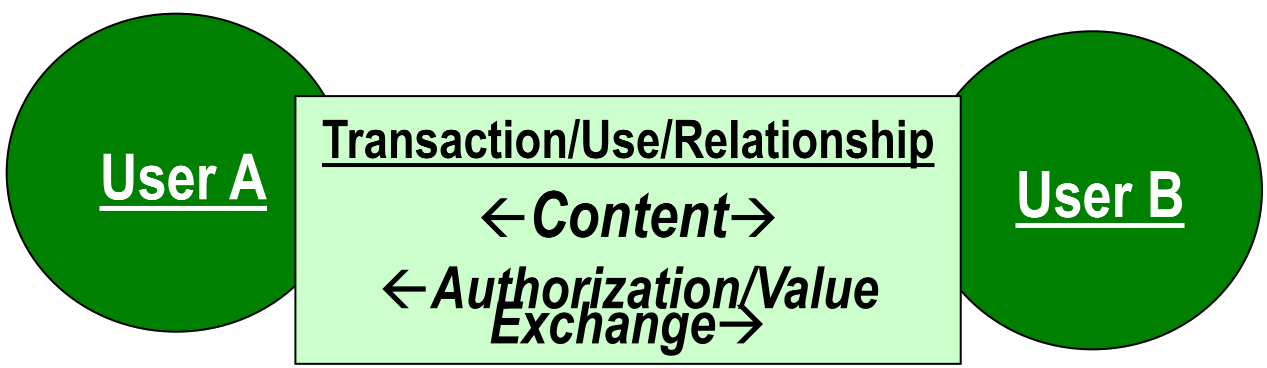
- Digital Items are "the content"
- DIs consist of
  - Resources (individual assets, distributed content),
  - Metadata (data about or pertaining the DI) and
  - Structure (relationships between parts of the DI)

Digital Item - Example

The DI is the fundamental unit for distribution and transaction within the MPEG-21 framework.

MPEG-21 User and User Interaction

- Any entity that interacts in the MPEG-21 environment or makes use of a Digital Item
- Users include individuals, organisations, corporations, consortia, governments, other standards bodies, etc.
- Roles including creators, consumers, rights holders, content providers, distributors, etc.
- Each User will assume specific rights and responsibilities according to their interaction with other users



Seven Architectural "Elements"

Digital Rights Management
- Pt. 4: **IPMP** Components
- Pt. 5: **R**ights **E**xpression **L**ang
- Pt. 6: **R**ights **D**ata **D**ictionary
- Amd.1: DII relationship types

Adaptation
- Pt. 7: **D**igital **I**tem **A**daptation
- Amd.1: Convers. And Permissions
- Amd.2: Dynamic and Distributed Adaptation

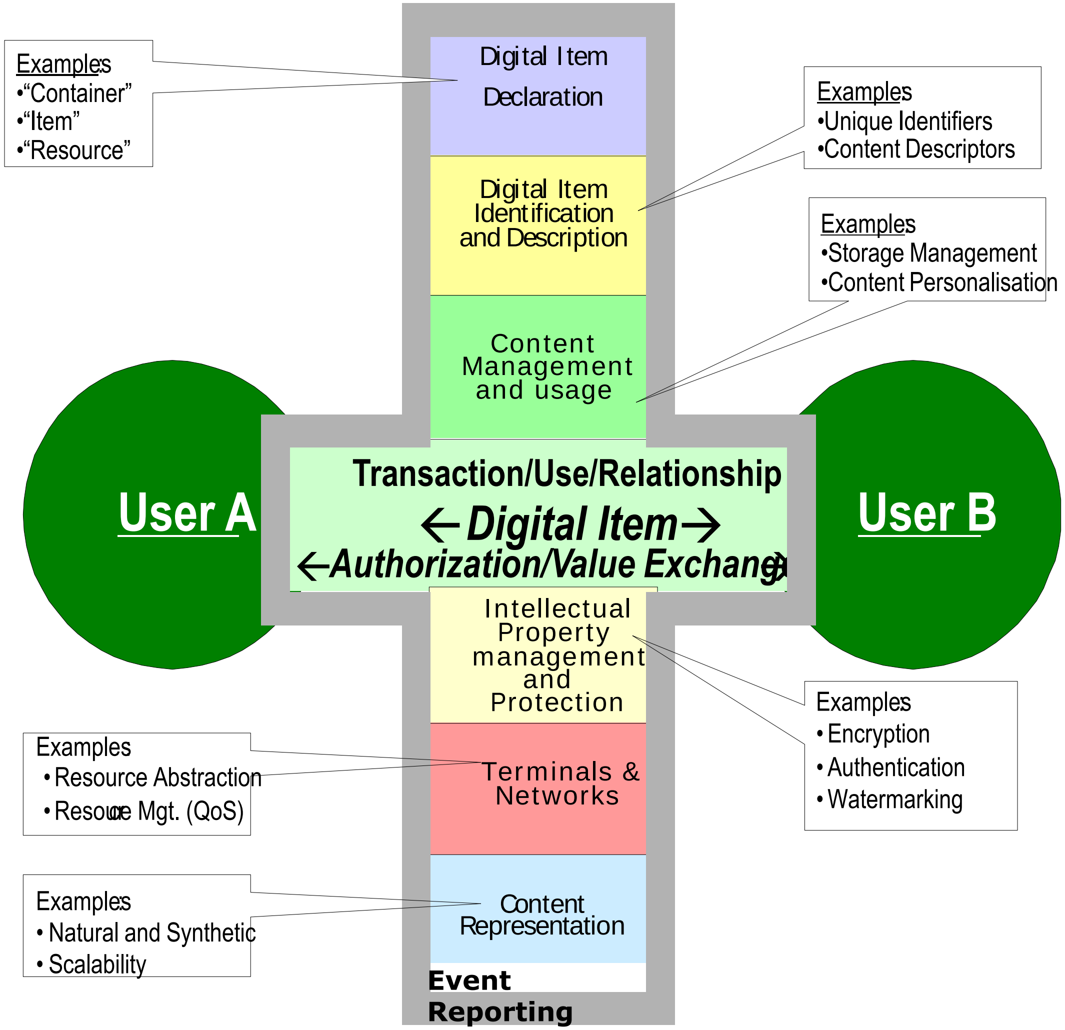
Processing
- Pt. 10: **D**igital **I**tem **P**rocessing
- Amd.1: Add‘l C++ bindings

Systems
- Pt. 9: File Format
- Pt. 16: Binary Format
- Pt. 18: **D**igital **I**tem **S**treaming

Misc
- Pt. 8: Reference Software
- Pt. 11: Persistent Association
- Pt. 12: Test Bed
- Pt. 14: Conform.
- Pt. 15: Event Reporting
- Pt. 17: Fragment Idenfication

Vision, Declaration, and Identification
- Pt. 1: Vision, Technologies and Strategy
- Pt. 2: **D**igital **I**tem **D**eclaration
- Pt. 3: **D**igital **I**tem **I**dentification

Roles of the Architectural Elements



Contents

Metadata Generation & Annotation

- Process of creating data about data
- Content has to be known
  - Watch & understand video / image collection
  - Listen and assess audio
- Metadata standard has to be known
  - What are the possible fields
  - What are the used classification systems

Evaluation (1/2)

- Goal: Identify the opinion of users on manual semantic annotation

- 5 Users with following (median) background:
  - 17 years of computer experience
  - Using a computer 50 h / week
  - 2 years experience with digital photo cameras
  - 4 years experience with imaging software

Evaluation (2/2)

- 2 Tasks were given:
  - Annotate a photo with a given description and an extensive prior introduction to semantic photo annotation with Caliph
    - video was shown,
    - concept was explained and
    - questions were answered
  - Annotate a photo fully on your own
  - After Tasks:
    - Questionnaire with several subjective questions
    - Evaluation Scale from: -3 (strongly disagree) to 3 (strongly agree)

Evaluation Results: General Questions

- The concept of meta data is very new to me: -2.6
- It was easy to understand the concept of semantic meta data while using Caliph: 1.8
- The visualization of the semantic meta data within Caliph is easy to understand and interpret: 2.2
- The annotation of images with textual descriptions can be done fast and easily: 1.4
- The annotation of images with semantic meta data can be done fast and easily: 1.2
- I can see an obvious benefit by using semantic meta data for image (multimedia) annotation: 1.4

Scale: (disagree) -3 to 3 (agree)

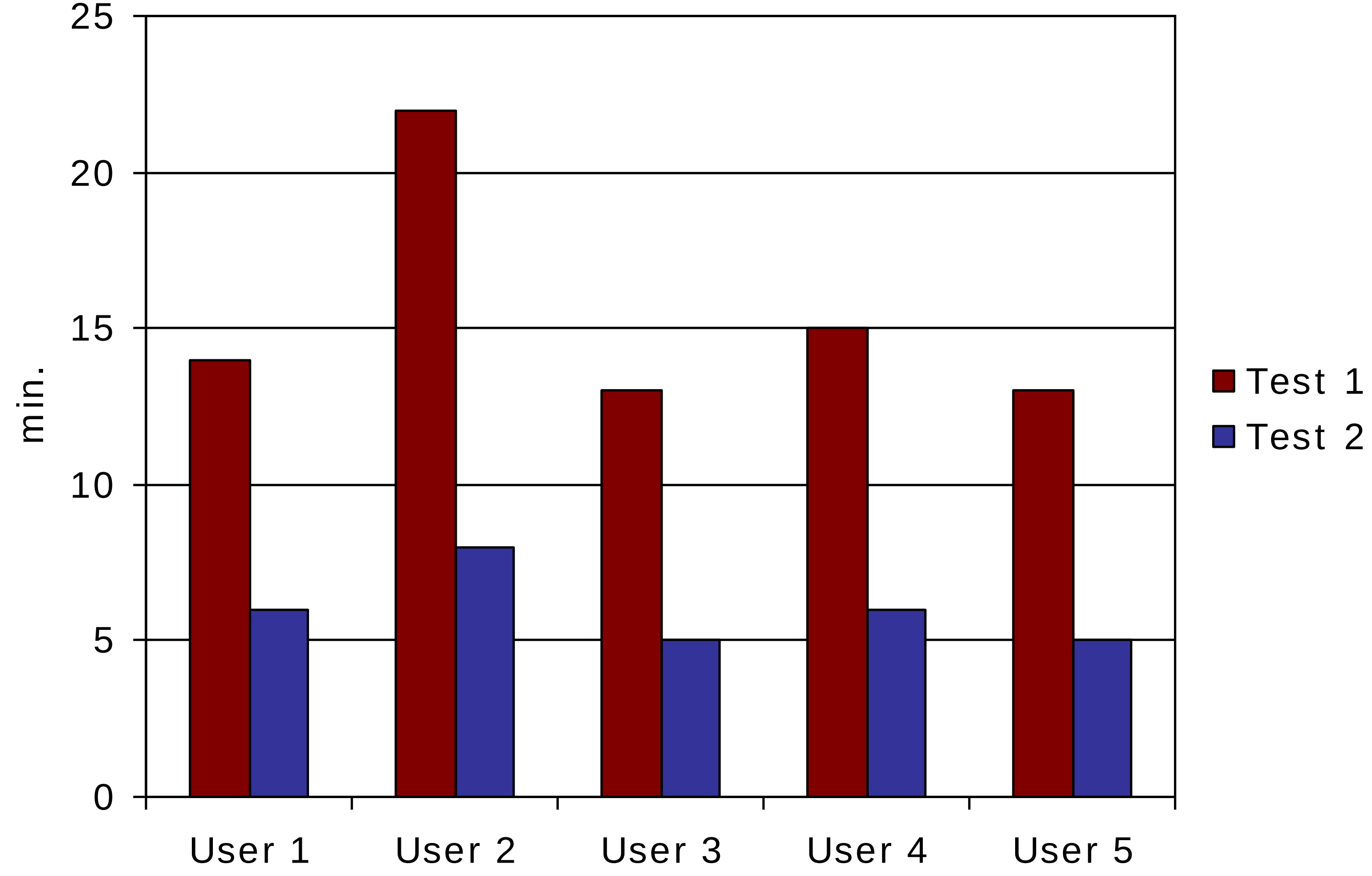
Evaluation Results: Scenario based questions

1. The complexity of semantic annotation is too high to be useful for organizing photos.
2. I would find it easy to annotate a large set digital photos (e.g. 100+).
3. I would recommend Caliph or a similar tool to annotate digital photos.
4. I can see an obvious benefit by using semantic meta data for the organization of photos.

Personal	Newspaper
-0.6	-1.8
-0.6	-0.2
0.8	1.4
1.4	2.2

Scale: (disagree) -3 to 3 (agree)

Evaluation Results: Annotation performance



Evaluation Results: Annotation performance

- Median times for annotation:
  - 15.4 minutes for the 1st test and
  - 6 minutes for the 2nd test
- Median time in a self test with 17 photos:
  - 1 minute and 53 seconds per photo
- That results in an approximate time of 10 h 27 min. for annotation of a set of 333 photos

Evaluation Results: Diversity of Annotations (2nd test)

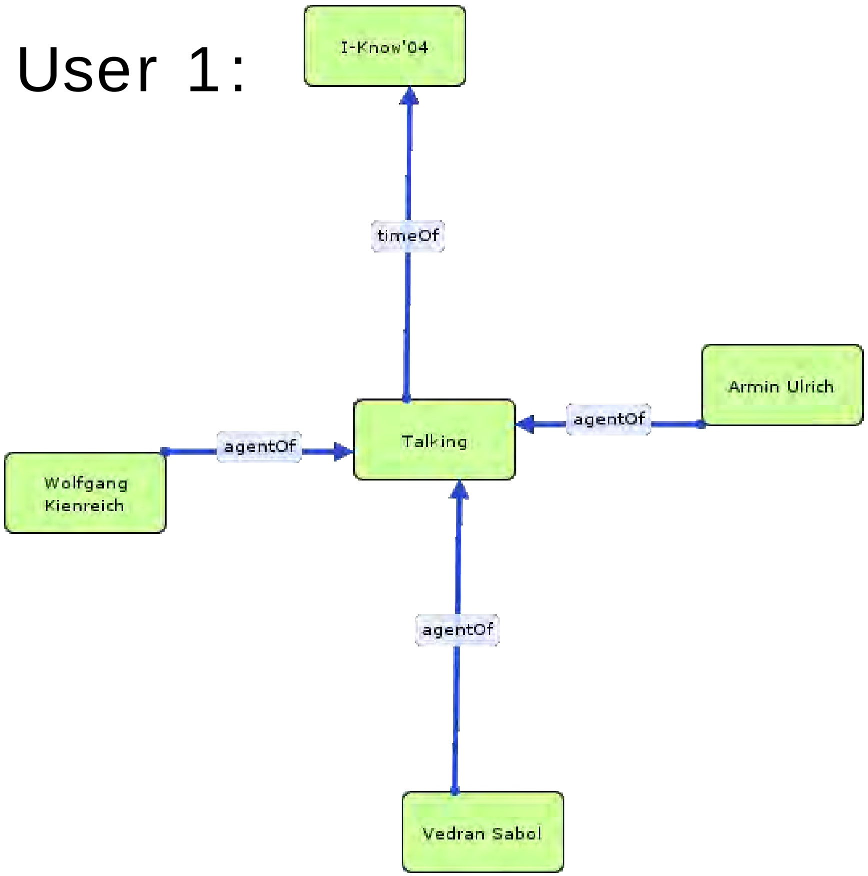
- Structured text annotation field "Who":
  1. Vedran, Wolfgang, Armin
  2. Wolf, Armin, Vedran
  3. Wolfgang Kienreich, Vedran Sabol, Armin Ulbrich
  4. wolfgang, armin, vedran
  5. W.Kienreich,A.Ulbrich,V.Sabol

Evaluation Results: Diversity of Annotations (2nd test)

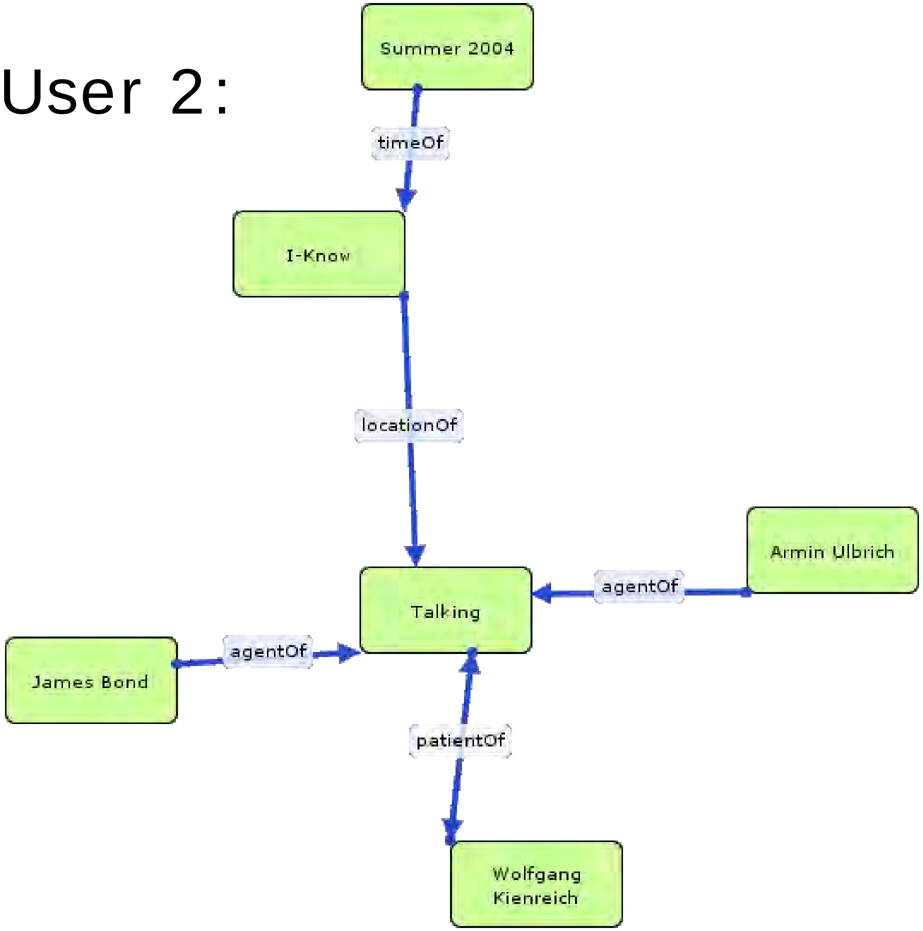
- Free text annotation:
  1. Stadthalle, Graz, Austria I-Know '04 Knowledge Managment Conference
  2. The three are sitting ...
  3. Wolfgang Kienreich, Armin Ulbrich und Vedran Sabol (v.l.n.r.) sprechen miteinander auf der I-Know.Wolfgang Kienreich, Vedran Sabol, Armin Ulbrich are at I-Know, Graz for Talking
  4. Stadthalle, Graz, Austria I-Know '04 Knowledge Managment Conference
  5. Wolfgang,Armin and Vedran talking to each other on I-Know 04 at Stadthalle Graz.

Evaluation Results: Diversity of Annotations (2nd test)

User 1:

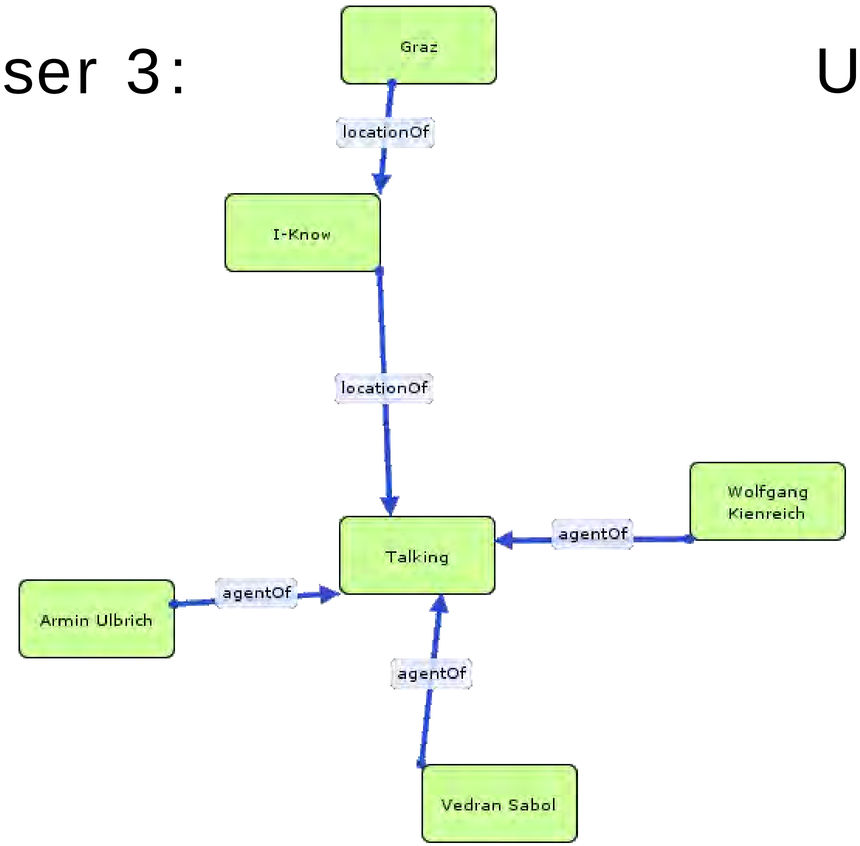


User 2:

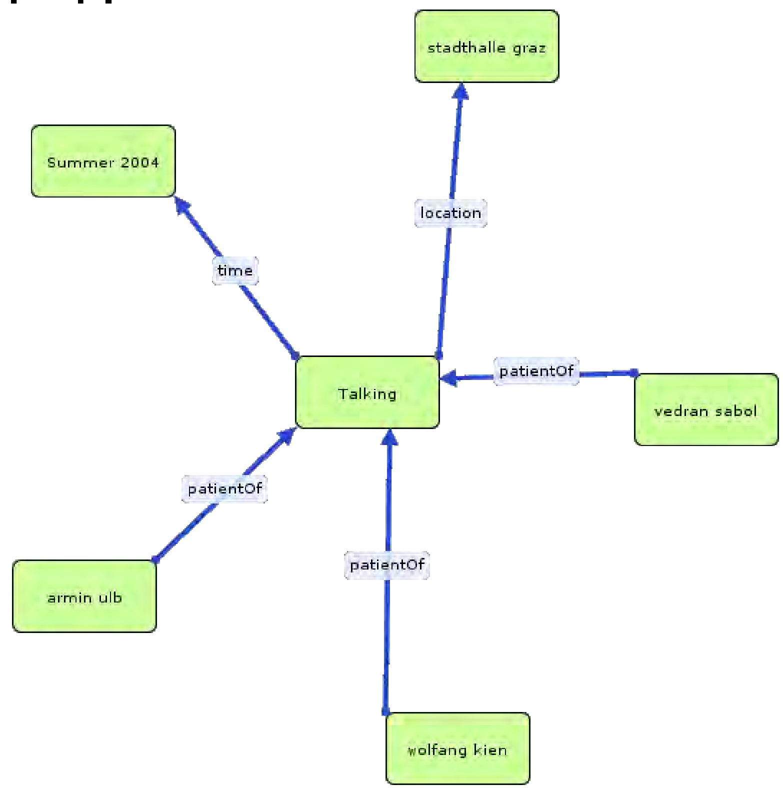


Evaluation Results: Diversity of Annotations (2nd test)

User 3:

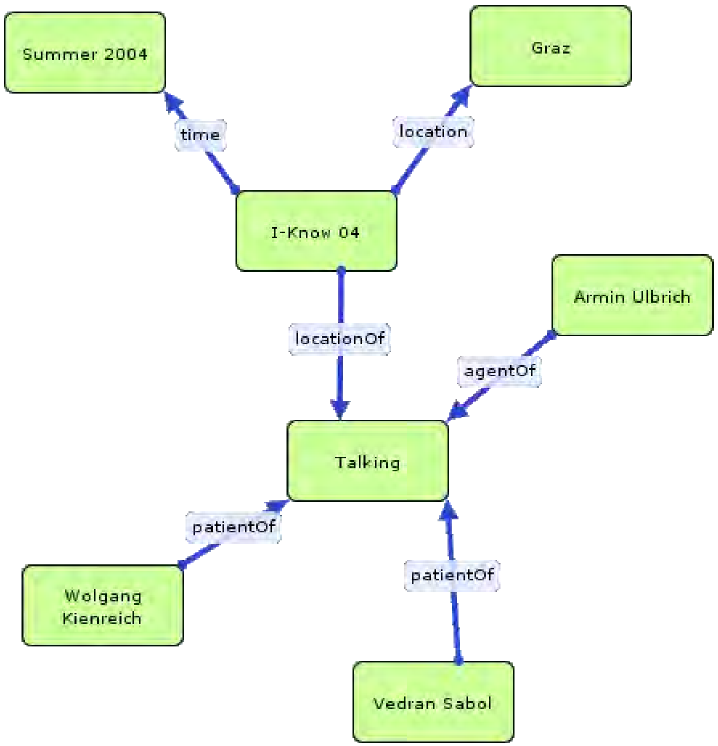


User 4:



Evaluation Results: Diversity of Annotations (2nd test)

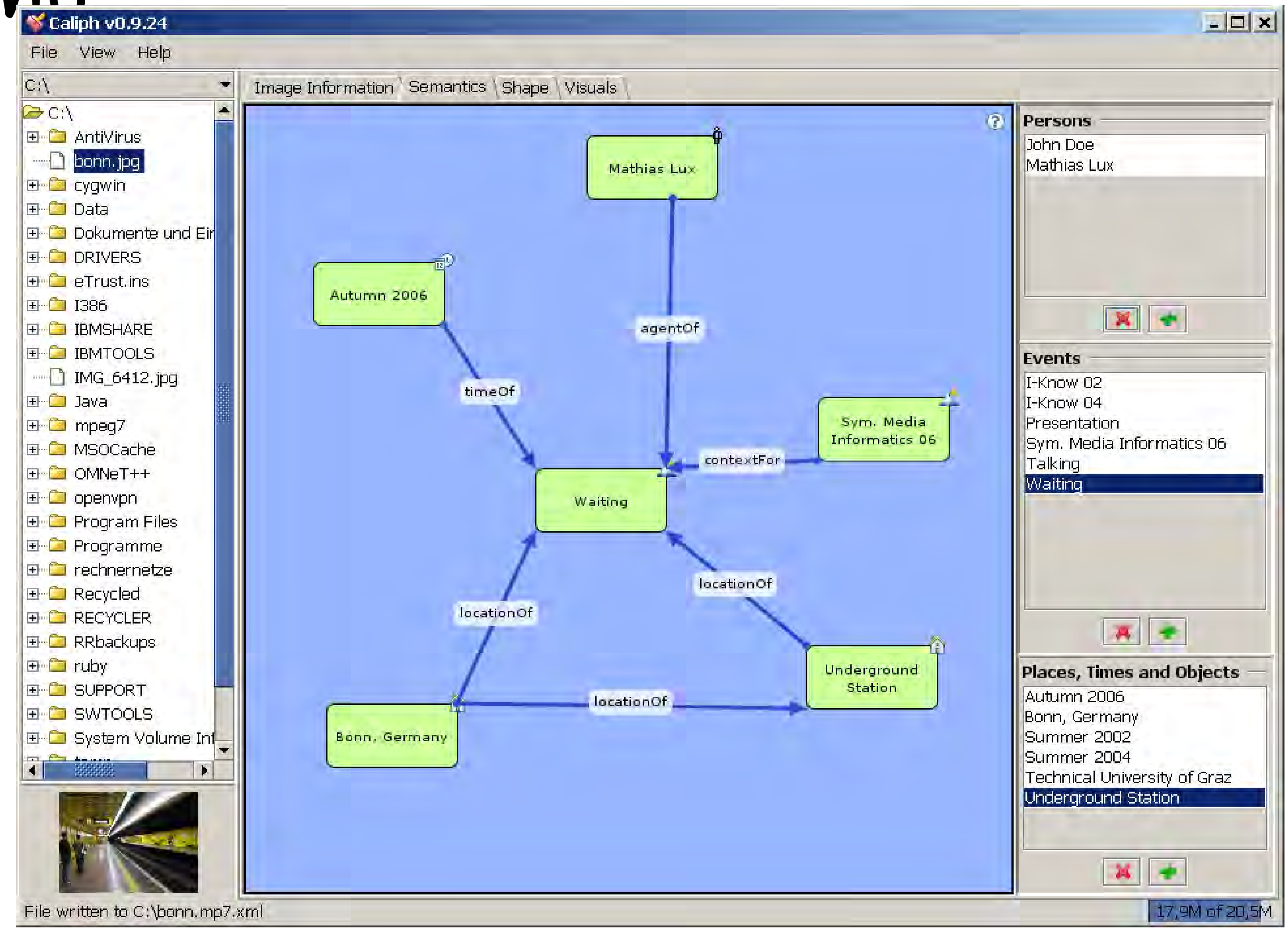
User 5:



Lessons Learned

- Users like the graphical annotations editor
- Users see semantic annotation in a professional (business) environment
- Semantic annotation is very time consuming
- The MPEG-7 nomenclature is not intuitive
  - Semantic agent / place / object & relations
  - Creator of image / description / quality rating
- Tagging with central tag repository …

Demo



Caliph v0.9.24
File View Help
C:\
AntiVirus
bonn.jpg
cygwin
Data
Dokumente und Ei
DRIVERS
eTrust.ins
I386
IBMSHARE
IBMTOOLS
IMG_6412.jpg
Java
mpeg7
MSOCache
OMNeT++
openvpn
Program Files
Programme
rechnernetze
Recycled
RECYCLER
RRbackups
ruby
SUPPORT
SWTOOLS
System Volume Inf
Image Information
Semantics
Shape
Visuals
Mathias Lux
Autumn 2006
agentOf
timeOf
Sym. Media Informatics 06
contextFor
Waiting
locationOf
locationOf
Underground Station
locationOf
Bonn, Germany
Persons
John Doe
Mathias Lux
Events
I-Know 02
I-Know 04
Presentation
Sym. Media Informatics 06
Talking
Waiting
Places, Times and Objects
Autumn 2006
Bonn, Germany
Summer 2002
Summer 2004
Technical University of Graz
Underground Station
File written to C:\bonn.mp7.xml
17,9M of 20,5M