NR

Recent Video Codec developments

MPEG-4, H.264, etc.

Eirik Maus, Wolfgang Leister

Norsk Regnesentral
Norwegian Computing Center

NR

Disclaimer

- Ingen av illustrasjonene er våre egne.
  - Kilde står oppgitt / Source is mentioned

Norsk Regnesentral
Norwegian Computing Center

NR

Recent Video Codecs

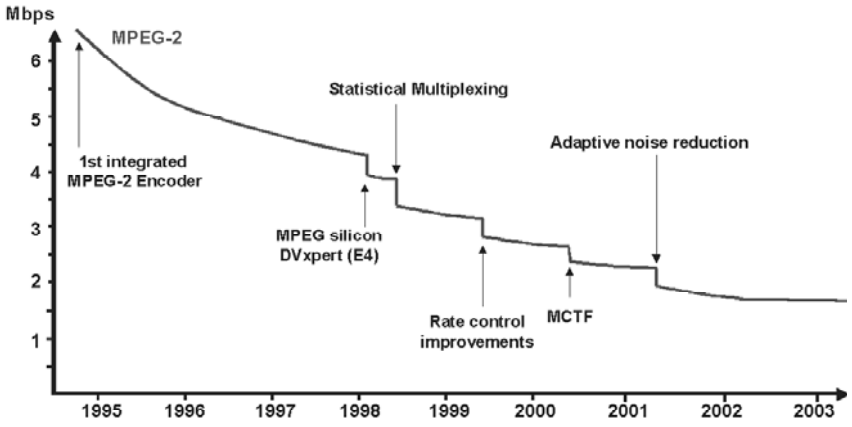
- MPEG 4 video codecs
  - Links to MPEG-4 Industry Forum,
  - but else not much information available ...
- WMV : Windows Media Video
  - Version 7, 8 based on MPEG-4
  - No descriptions found for version 9 (as of spring 2004)
- H.264 (H.26L, JVT, MPEG-4 AVC,...)
  - Future part of MPEG-4 (Annex 10)
  - See presentation by Thomas Wiegands:
  - http://bs.hhi.de/~wiegand/H264_03.pdf
- Other recent codec developments:
  - DivX 3, 4, 5 : based on WMV7 and MPEG-4
  - Xvid: MPEG-4 advanced simple profile, without b-frames(?)
  - On2 VP6, Real,..
  - H263++ (?)

NR

Comparison of recent codecs

- http://www.extremetech.com/print_article/0,1583,a=121163,00.asp
  - **Best: DivX 5.1.1**, **WMV9**, bad: QT/sorenson3, QT/MPEG-4
- http://www.doom9.org/codecs-103-1.htm (mai 2003)
  - 3ivX, DivX5, Dicas Mpegable, Nero Digital, Real RV9, Xvid 0.9
  - Not: WMV or quicktime
- http://www.doom9.org/codecs-203-1.htm (des 2003)
  - As above with newer versions + On2 VP6 and ffvfw
  - Xvid was liked best, VP6, RV9, divx 5

Codecs using the same standard are improving
Same file format, better



Source:
http://www.harmonicinc.com/stageone/files/harmonic/collateral/AVC_AAC_Techie_041503.pdf
(Tandberg presents a similar graph for their MPEG-2-products)

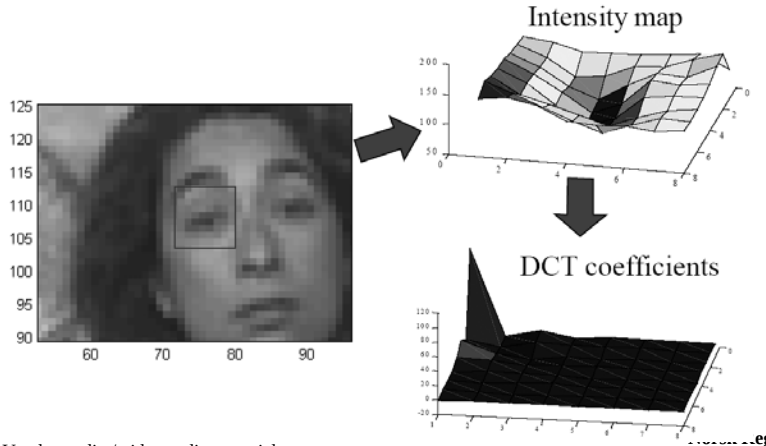
Main principles for image/video compression

1. Reduce size of numbers to be coded
   1. For video: Use redundancy between images in series
      - Find difference to area most similar in other frame
   2. Use redundancy within area/image
      - Organise data so that most important features in image/area are concentrated in few data units.
   3. «Predict» values for other pixels from these
   4. Calculate difference between prediction and real value, and code residual
      - Small numbers need fewer bits in encoding
2. Code residuals with shortest possible coding
   - Entropy-coding – several algorithms available

(see also video compression tutorial: http://www.vcodex.com/videocoding2b.pdf)

Concentrate ”energy” (main principle)

- Usually: transform from point-based to frequency based.



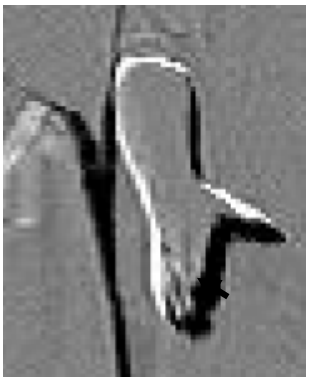
Kilde: Vcodex audio / video coding tutorial
http://www.vcodex.com/videocoding2b.pdf

Code difference values

Kilde: Vcodex audio / video coding tutorial
http://www.vcodex.com/videocoding2b.pdf

Detection of movements

- Compensate for movements between frames
  - chose reference area «to the left», «45 degrees up two pixels». ...



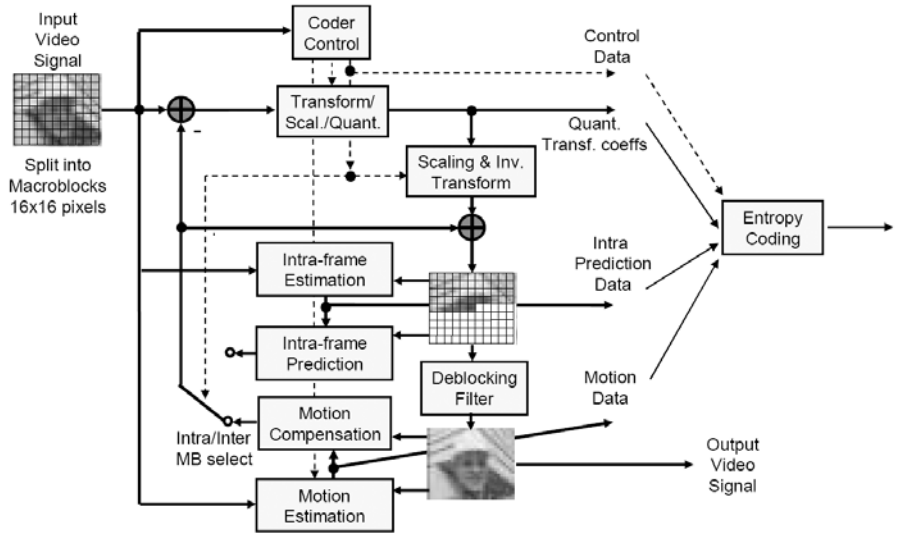
- Same hand
  - Moved up and to the left
- Much data (black/white) due to movements
- Using compensation for movement vector the amount of data would become less.

Kilde: Vcodex audio / video coding tutorial
http://www.vcodex.com/videocoding2b.pdf

How can we make compression better?

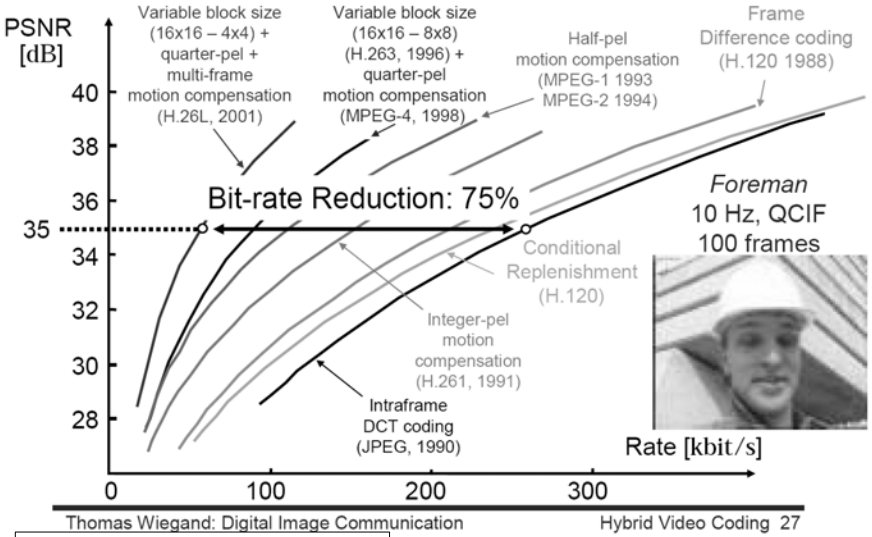
- Avoid coding values that are not very small ...
  - Minimize number of reference frames (intra-coded)
  - Better algorithms to extract main features
    - From other frames (inter-coding):
      Better choices of motion vector for matching areas
    - Pixels in same frame (intra-coding):
      Neighbours that are alike, e.g., flat surfaces
- Better compression: entropy coding
  - Sort data so that small values / similar values follow each other
    - Try to reduce possibility that previous data unit different
  - Chose how to code: shortest way to express series of values

Hybrid video codec basics



T. Wiegand and G. J. Sullivan: The H.264 | MPEG-4 AVC Video Coding Standard 16

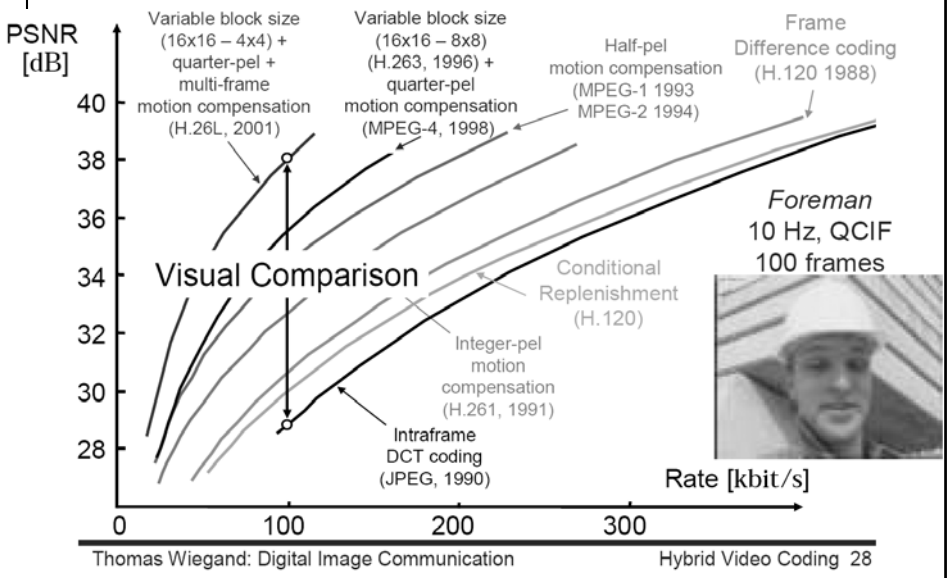
Milestones



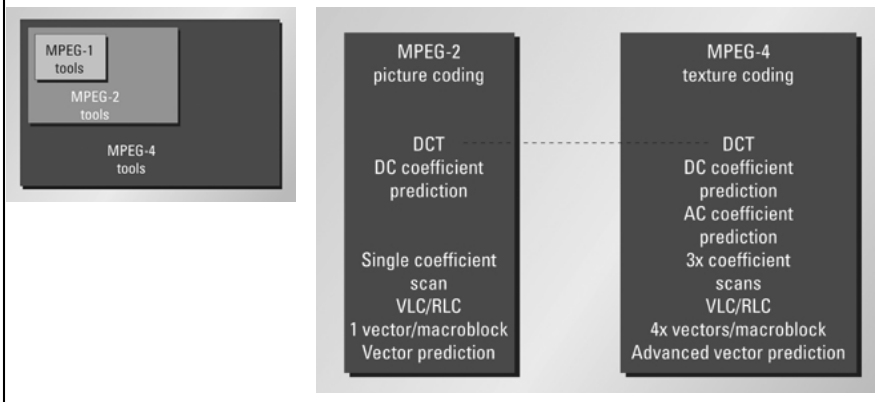
Thomas Wiegand: Digital Image Communication Hybrid Video Coding 27

Kilde: kursfoiler, Fraunhofer HHI:
http://bs.hhi.de/~wiegand/hybrid_video_coding_03.pdf

Milestones 2



Mpeg 4 tools



Kilde: Broadcast Engineering
http://broadcastengineering.com/ar/broadcasting_mpeg_secrets/index.htm

Mpeg-4-v1-visual objects, Profiles, tools, levels

Some MPEG-4 Visual Profiles

Visual Profile	Level	Typical Visual Session Size	Max bitrate (Kbit/s)	Max Objects	Object Types
Simple	L0	QCIF	64	1	Simple
	L1	QCIF	64	4	Simple
	L2	CIF	128	4	Simple
	L3	CIF	384	4	Simple
Core	L1	QCIF	384	4	Simple, Core
	L2	CIF	2000	16	Simple, Core
Main	L2	CIF	2000	16	Simple, Core, Main
	L3	D1	15000	32	Simple, Core, Main
	L4	1920 x 1088	38400	32	Simple, Core, Main
Adv. Simple	L0	QCIF	128	1	Simple, Adv. Simple
	L1	QCIF	128	4	Simple, Adv. Simple
	L2	CIF	384	4	Simple, Adv. Simple
	L3	CIF	768	4	Simple, Adv. Simple
	L4	½ D1	3000	4	Simple, Adv. Simple
	L5	D1	8000	4	Simple, Adv. Simple

Mpeg 4 Tools, Object types

MPEG-4 Visual Tools used by Visual Object Types	Visual Object Types															
Visual Tools	Simple	Core	Main	Advanced Simple	Fine Granularity Scalable	Simple Scalable	N-Bit	Animated 2-D Mesh	Basic Animated Texture	Still Scalable Texture	Simple Face	Advanced Real Time Simple	Advanced Coding Efficiency	Advanced Scalable Texture	Core Scalable	
Basic I-VOP P-VOP AC/DC Prediction 4-MV, Unrestricted MV	x	x	x	x	x	x	x	x				x	x		x	
Error Resilience: Slice Resynchronization Data Partitioning Reversible VLC	x	x	x	x	x	x	x	x				x	x		x	
Short Header	x	x	x	x	x		x	x				x	x		x	
B-VOP		x	x	x	2	x	x	x					x		x	
P-VOP with OBMC (Texture)																
Method 1 / Method 2 Quantization		x	x	x	x		x	x					x		x	
P-VOP based temporal scalability Rectangular Arbitrary Shape		x	x				x	x					x		x	
Binary Shape		x	x				x	x	x				x		x	
Grey Shape			x										x			
Interlace			x	1	1								x			
Sprite			x													
Temporal Scalability (Rectangular)						x									x	
Spatial Scalability (Rectangular)						x									x	
N-Bit							x									
Scalable Still Texture								x	x	x				x		
2D Dynamic Mesh with uniform topology								x	x							
2D Dynamic Mesh with Delaunay topology								x								
Facial Animation Parameters											x					
Body Animation Parameters																
Dynamic Resolution Conversion												x				
NEWPRED												x				
Global Motion Compensation				x									x			
Quarter-pel Motion Compensation				x									x			
SA-DCT													x			
Error Resilience for Visual Texture Coding														x		
Wavelet Tiling														x		
Scalable Shape Coding for Still Texture														x		
Object Based Spatial Scalability															x	
Fine Granularity Scalability					x											
FGS Temporal Scalability					x											

- Different object types allow different compression tools.
- Relevant types:
  - Simple
  - Advanced Simple
- Others are usually not used.

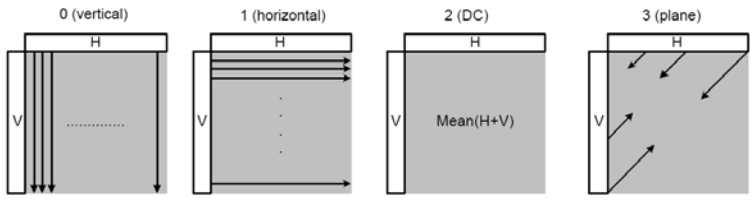
H.264 Complexity and application

- MPEG2 today is much better than before!
- Higher complexity can pay off
- Computing power has become cheap, anyway (?)

MPEG-4 AVC / H.264 Profiles	Target applications	Rough decoder complexity increase over MPEG-2	Preliminary estimates of efficiency improvements over MPEG-2
Baseline Profile	low delay applications, video phone, mobile ...	2.5 X more complex	1.5 x better
Extended Profile	mobile, streaming, ...	3.5 X more complex	1.75 x better
Main Profile	interlaced video applications, broadcast, packaged media, ...	4.0 X more complex	2.0 x better

Kilde Mpeg 4 industriforum:
http://www.m4if.org/public/documents/vault/m4-out-30035.zip

New tool: intra-prediction (1)

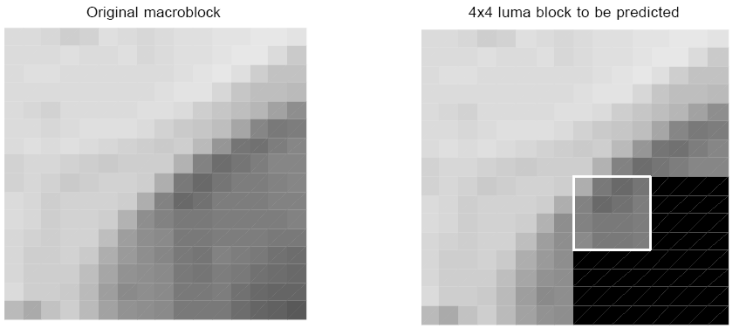


H.26L Intra 16x16 prediction modes (all predicted from pixels H and V)

- Prediction has basis in surrounding pixels
- However: only from blocks that have been decoded previously to the current one

Kilde: Vcodex h264/ mpeg-4-a-10 tutorials
http://www.vcodex.com/h264.html

New tool: intra-prediction (2)



Source: Vcodex h264/ mpeg-4-a-10 tutorials
http://www.vcodex.com/h264.html

New tool: intra-prediction (3)

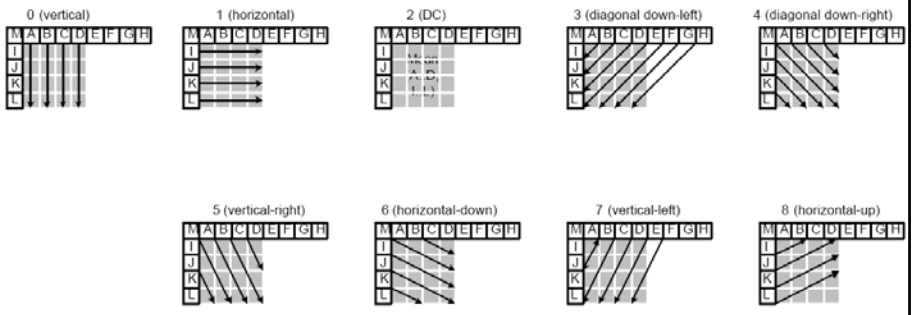
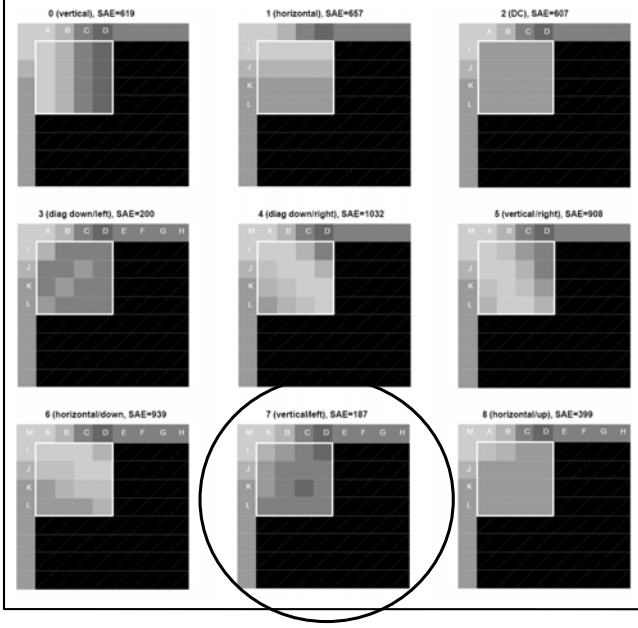


Figure 3 4x4 luma prediction modes

Source: Vcodex h264/ mpeg-4-a-10 tutorials
http://www.vcodex.com/h264.html

Intra prediction (4)



- Winner:
  metod nr 7
  - Sum av feil: 187
- The others:
  3: sae = 200
  4: sae = 1032

Source: Vcodex h264/ mpeg-4-a-10 tutorials
http://www.vcodex.com/h264.html

Macroblock subpartition (1)

- Makroblocks with movements/colour variations can be split upp into smaller parts (max. Twice):
  - 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4
- Each subblock has more equal values
- Results in tree-structure for motion vectors (inheritance + own)
  - Predicted from other / inherited values in order not to get too big.

Macroblock subpartition (2)

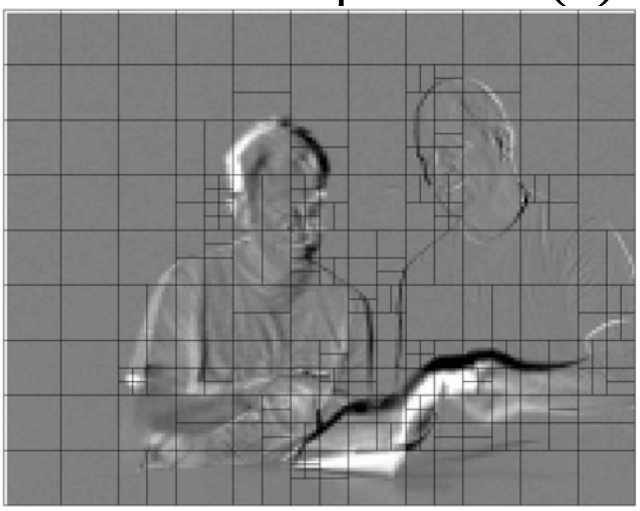


Figure 2-3 Residual (without MC) showing optimum choice of partitions

Norsk Regnesentral
Norwegian Computing Center

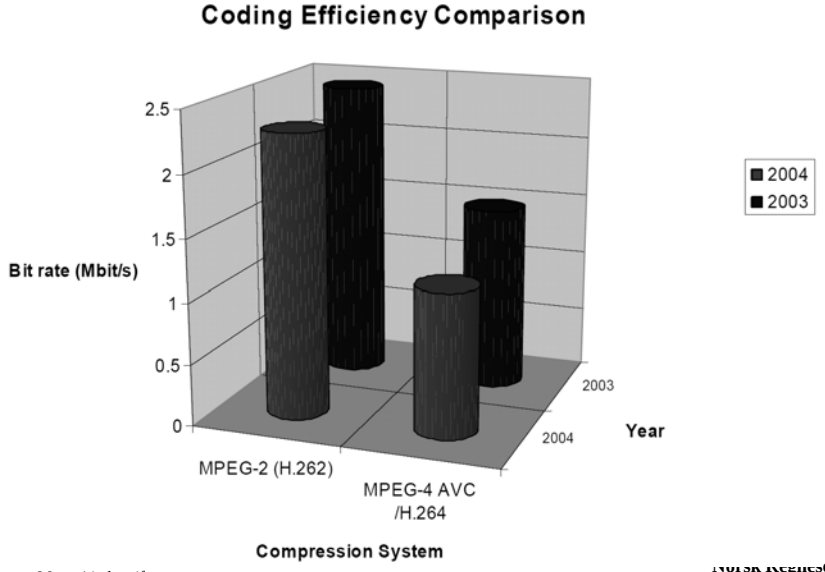
Kilde: Vcodex h264/ mpeg-4-a-10 tutorials
http://www.vcodex.com/h264.html

Other news in H.264

- Integer transform, not DCT
  - Nearly identical with factorised DCT
    - Tutorial: http://www.vcodex.com/h264.html
- Deblocking-filter for reconstruction is specified in standard
  - Can be used in motion prediction in encoder!
- Entropi-coding
  - Exp-Golomb codes (Variable length coding)
  - ELLER CABAC (not simple profile), like JPEG 2000
  - Cabac is best, but needs Ghz processor for decoding / presentation

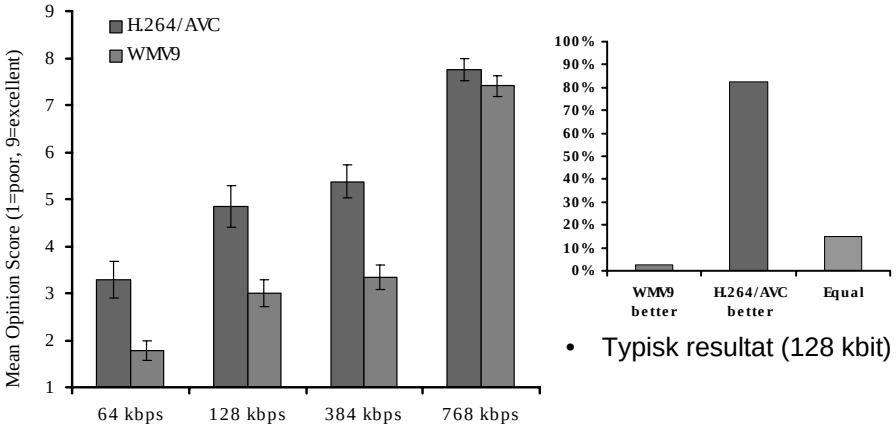
Norsk Regnesentral
Norwegian Computing Center

Quality: each year better algorithms



Source Mpeg 4 industriforum:
http://www.m4if.org/public/documents/vault/m4-out-30035.zip

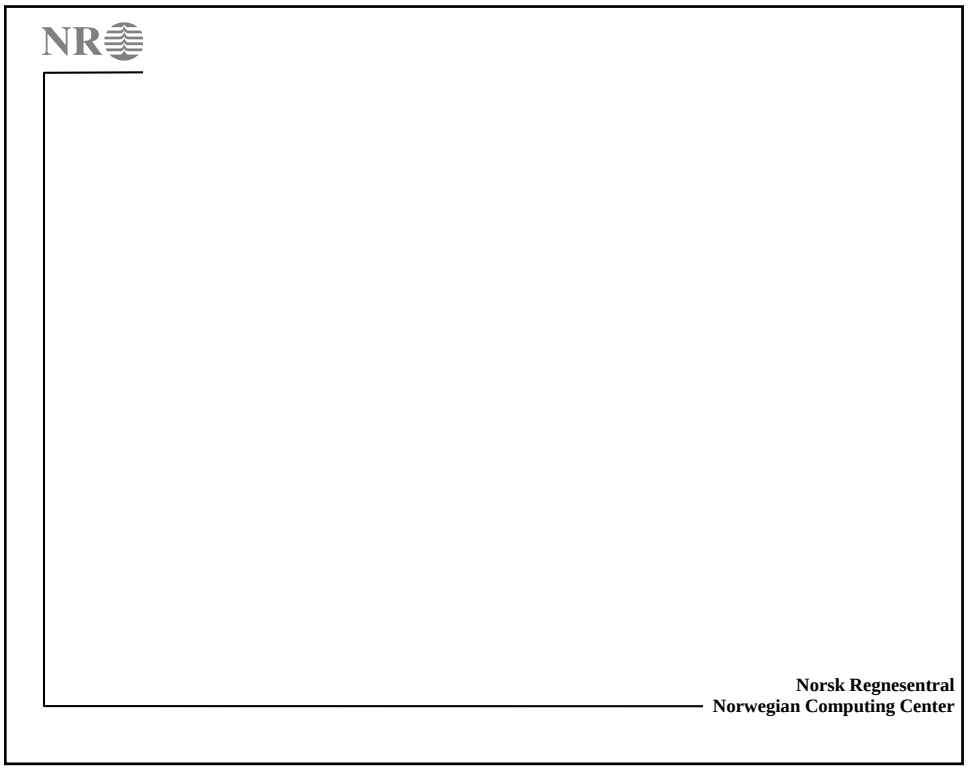
Quality: WMV9 vs. H.264



- Typisk resultat (128 kbit)

- Nokia user-experienced-quality test: H.264 is far the best
- But other tests say: «more noise», «larger errors», ...

Kilde: Nokia test for 3GPP:
http://www.3gpp.org/ftp/tsg_sa/WG4_CODEC/Ad-hoc_video_codec/Docs/AHVIC-004.zip


NR
Norsk Regnesentral
Norwegian Computing Center