JCTVC-A114

Description of video coding technology proposal by NTT, NTT DOCOMO, Orange Labs, Panasonic and Technicolor

15-23 April 2010
1st JCT-VC meeting

NTT NTT docomo orange™ Panasonic ideas for life technicolor

Outline

- Proposal overview
- Coding tools
- Encoder control
- Performance & Complexity
- Conclusion

Basic design of CDCM

CDCM – Collaborative Design Coding Model

- Block-based predictive transform codec
  - New codec design, not an incremental extension of MPEG-4 AVC/H.264
- Main properties
  - Using simple and fundamental coding tools in the decoder
  - Optimized encoding algorithms
  - Implements multi-thread decoding capabilities

Performance

- Significant objective improvement / anchors - BD-rate gains:

JCfP version			
Anchor	BD gain Y	BD gain U	BD gain V
Alpha - CS1	31.6%	29.2%	30.0%
Beta - CS2	30.4%	10.6%	10.9%
Gamma -CS2	47.4%	34.1%	35.1%

Revised version			
Anchor	BD gain Y	BD gain U	BD gain V
Alpha -CS1	31.7%	34.5%	35.6%
Beta - CS2	30.6%	18.3%	18.1%
Gamma -CS2	47.6%	39.2%	40.0%

  - Performs equally well in all sequence classes, bitrates and constraint sets
  - Enhanced chroma reproduction

- Significant visual improvement compared to Alpha and Beta anchors
  - For all bitrates in both constraint sets
  - Sharper pictures, less blocking artifacts, better chroma representation

Content of the proposal

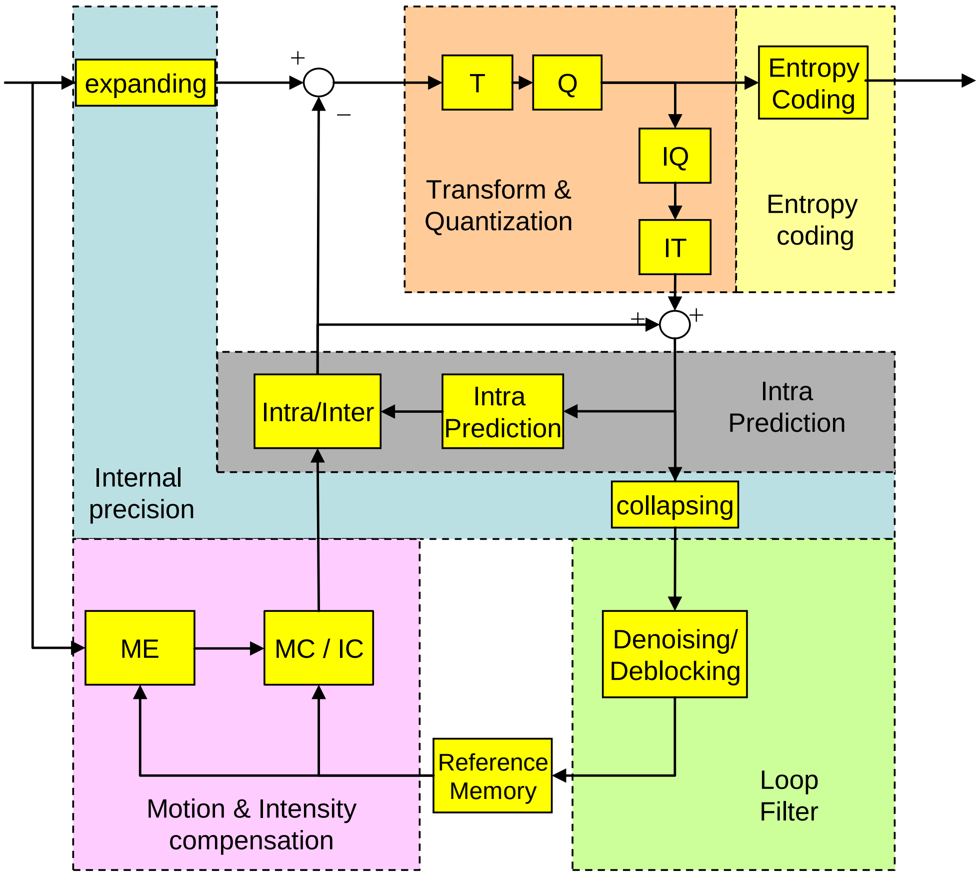
- Stable and robust encoder and decoder reference software
  - Proprietary, C++ based
  - Runs on multiple platforms, multiple compilers
  - Developed and tested by many companies
  - Systematic integration steps: valgrind, streams crosscheck, perf. verifying
  - Single and Multi-thread decoding capabilities

- Complete documentation
  - Decoder specification
  - Reference Encoder Model
  - User manual
  - Doxygen support

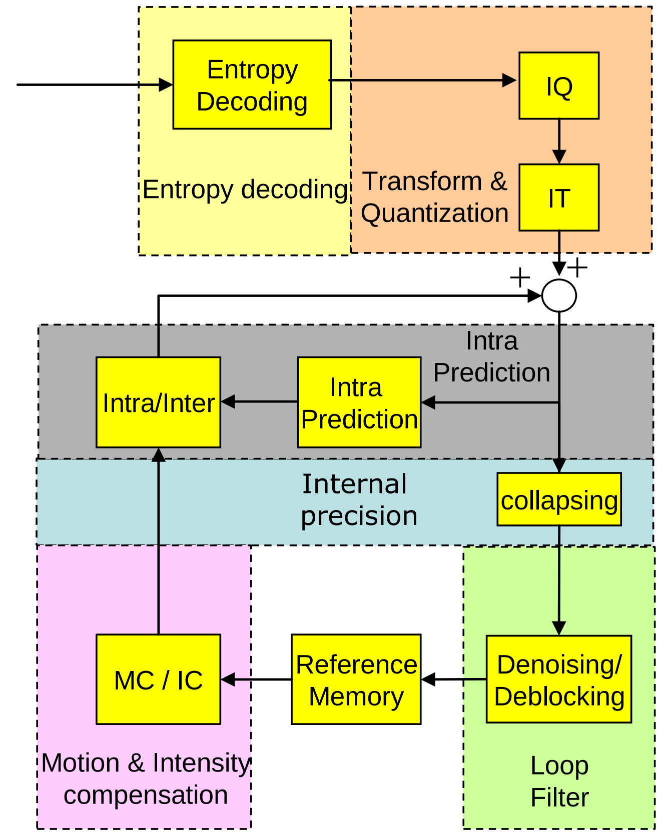
Outline

- Proposal overview
- Coding tools
- Encoder control
- Performance & Complexity
- Conclusion

CDCM Encoder and decoder block diagrams



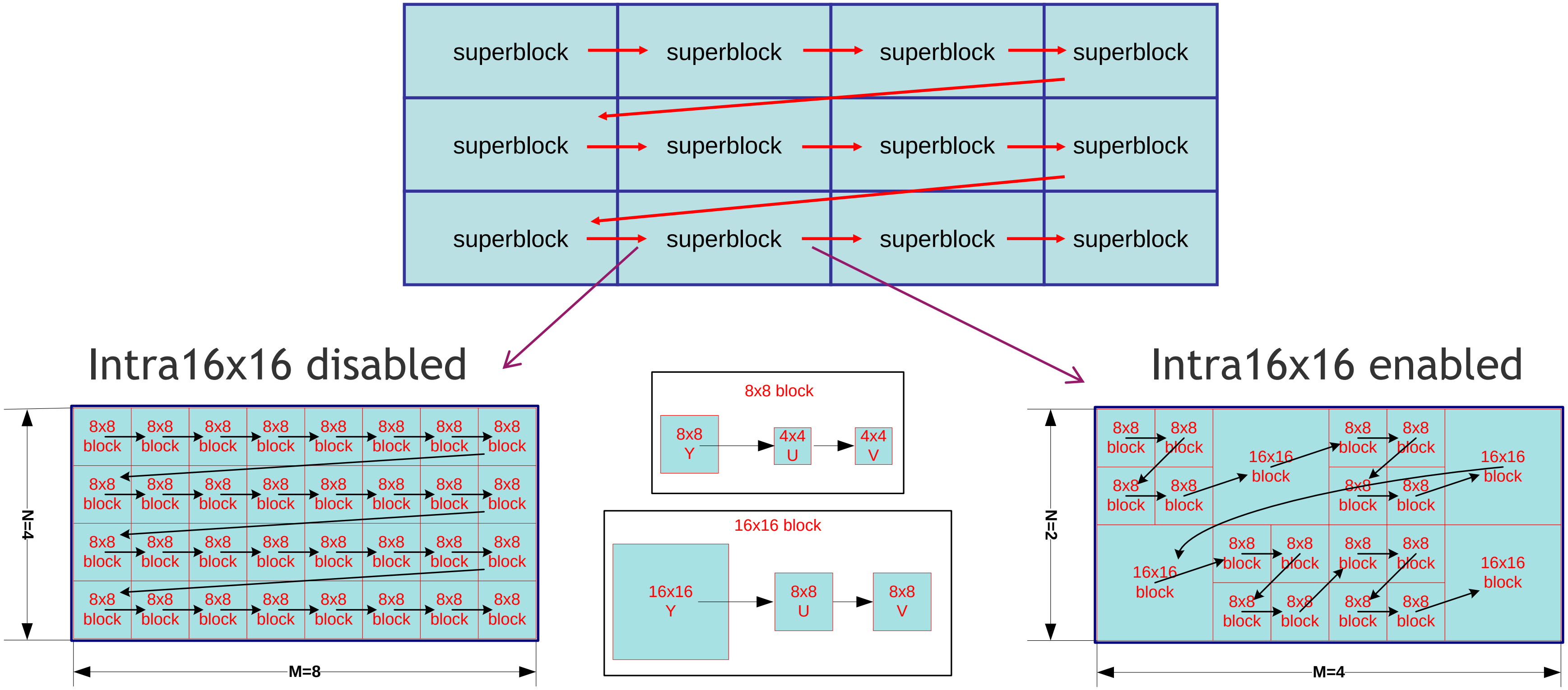
Encoder



Decoder

Basic picture structure

- Basic structure: Superblocks
- Basic Block-unit:
  - Inter: 8x8 Luma + 4x4 Chroma
  - Intra: 16x16 blocks & Transform can be enabled



Detailed list of tools

Motion Representation & Intensity Compensation	
Frame types	I-Frame, P-Frame (1 ref list), B-Frame (2 ref lists)
MV block size	8x8 for luma (4x4 for chroma)
MV resolution	-up to 1/8th pel in luma (1/16th pel in chroma) -selected at frame level
Motion sharing	motion block frontier can be displaced by 2 or 4 pixels (right or down)
Intensity Compensation	scale and offset signaled at the block level
MV Competition in P-pict	predictor index signaling – predictors set switchable at frame level

Luma/Chroma interpolation for Motion Compensation	
Luma 1/4 pel	-2 modes: 1. fixed AVC/H.264 interpolation – 2. SAIF (Wiener) -selected at frame level
Luma 1/8 pel	bilinear interpolation from 1/4th pel samples
Chroma	bilinear interpolation from full-pel samples, similar to MPEG-4 AVC/H.264 chroma interpolation (adaptation for 1/16 pel)

Differences with MPEG-4 AVC/H.264 are in red font

Detailed list of tools

Intra Prediction	
Intra pred block size (SIP type)	-Luma: 16x16 (I-picture only), 8x8, 4x4, 2x8 and 8x2 -16x16 block: each 8x8 block within can be split in 8x8, 4x4, 2x8 or 8x2 -Chroma: 8x8 and 4x4 implicitly inferred from luma SIP type
Prediction modes	- Consistent MPEG-4 AVC/H.264 like modes for intra4x4, intra8x8 - Generalisation MPEG-4 AVC/H.264 directional modes for intra16x16
Edge-Based Prediction Mode	DC mode can be replaced by a new directional mode, depending on the presence or not of an edge in neighboring reconstructed blocks
TM Averaging (TMA)	2 TMA modes for 16x16, 8x8, 4x4 - 1TMA mode for luma 2x8, 8x2
Low pass filter	-Luma: Same as MPEG-4 AVC/H.264 for intra8x8 and intra16x16 -Chroma: adaptive filtering implicitly inferred for chroma8x8

Differences with MPEG-4 AVC/H.264 are in red font

Detailed list of tools

Transform & Quantization	
Intra block transform size (inferred from SIP_type)	-Luma: can use 16x16 (intra blks), 8x8, 4x4, 2x8, 8x2 -Chroma: Intra chroma blocks can use 4x4 or 8x8 transforms
Inter block transform size (inferred from pred mode)	-Luma: can use 8x8 and 4x4 transforms -Chroma: use 4x4 transform
Adaptive transform selection (DCT-like & KLT)	-For luma, adaptive choice between 2 transforms for each block size -Signaling: For 16x16 Transf, flag explicitly signaled, Otherwise flag embedded in coeffs
Quantization	-qp resolution finer than that of the MPEG-4 AVC / H.264 -quantization matrices support

Entropy coding/decoding	
Zero tree coding	-In place of Signif. map coding of MPEG-4 AVC/H.264, consistent for Motion and Texture -recursive zero tree for 16x16 transform
CABAC	Similar to MPEG-4 AVC/H.264

In-loop filtering	
Non-linear Denoising (NDF)	-regularization based on a sparsity constraint on the signal -uses a thresholding-based approach in an over-complete transform domain
Frame Adaptive Denoising	frame-based Wiener filter using residual, prediction and reconstructed signals
Deblocking Filter	-Similar to MPEG-4 AVC/H.264 -possible to place deblocking filter before/after NDF

Differences with AVC/H.264 are in red

Outline

- Proposal overview
- Coding tools
- Encoder control
- Performance & Complexity
- Conclusion

Encoder control

- Multi-pass encoding
  - I, P or B frame type / SAIF on/off / 1/4th or 1/8th motion precision
  - To reduce complexity, combinations are constrained depending on frame type/level
- 3-steps motion estimation
  - 1.Initial ME: search performed in a quadtree structure, from 8x8 to up to 128x128
  - 2.Trellis ME: joint optimization of rows and columns of MVs
  - 3.Region ME: region growing process
  - To reduce complexity, some steps skipped depending on motion precision and interpolation filter type
- Intra-picture encoding
  - SIP type selection using a Quad-tree like selection
    - 2 more SIP types for luma (2x8 and 8x2)
  - Prediction modes tested
    - For SIP_16x16, 11 prediction modes instead of 4 in MPEG-4 AVC/H.264
    - For other SIP types, up to 2 TMA modes in addition to the 9 in MPEG-4 AVC/H.264
    - 1 additional mode on top of DC mode: Edge-Based Prediction Mode

Outline

- Proposal overview
- Coding tools
- Encoder control
- Performance & Complexity
- Conclusion

Performance & Complexity

Objective performance

- A revised version has been produced with bug fixing & improvts
  - Improvements: KLT function, Simplif λ setting, I16x16 support in P/B frames
  - Bug fixes: 5 encoder related, 1 decoder related (minor impact)

JCfP version			
Anchor	BD gain Y	BD gain U	BD gain V
Alpha - CS1	31.6%	29.2%	30.0%
Beta - CS2	30.4%	10.6%	10.9%
Gamma –CS2	47.4%	34.1%	35.1%

Revised version			
Anchor	BD gain Y	BD gain U	BD gain V
Alpha -CS1	31.7%	34.5%	35.6%
Beta - CS2	30.6%	18.3%	18.1%
Gamma -CS2	47.6%	39.2%	40.0%

- Comments
  - Performs equally well in all sequence classes, bitrates and constraint sets
  - Enhanced chroma reproduction

Performance & Complexity analysis

Encoding time analysis

- Tests conditions
  - Performed with a cluster using 64-bit Linux, with 64 cores
  - 2 types of machines:
    - 4 machines: 2 Quad-core Xeon E5450 @ 3.00GHz, 32GB RAM
    - 4 machines: 2 Quad-core Xeon E5540 @ 2.53GHz, 32GB RAM

	Class A	Class B	Class C	Class D	Class E
CS1 (sec/fr)	857	403	75	21	
CS2 (sec/fr)		379	71	19	145

Decoding time analysis

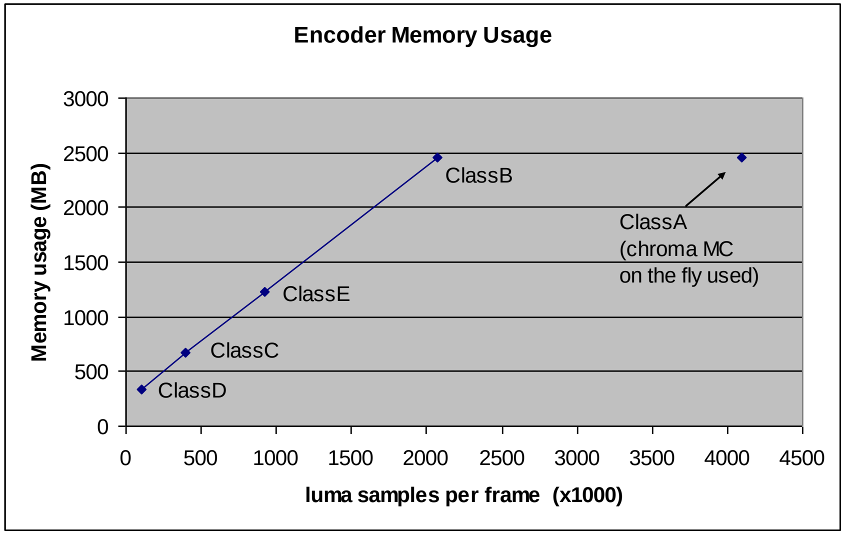
- Tests conditions: 1 Quad-core Xeon E5420 @ 2.5GHz
  - 1 Quad-core Xeon E5420 @ 2.5GHz, 4GB Memory, Windows XP - 32bit

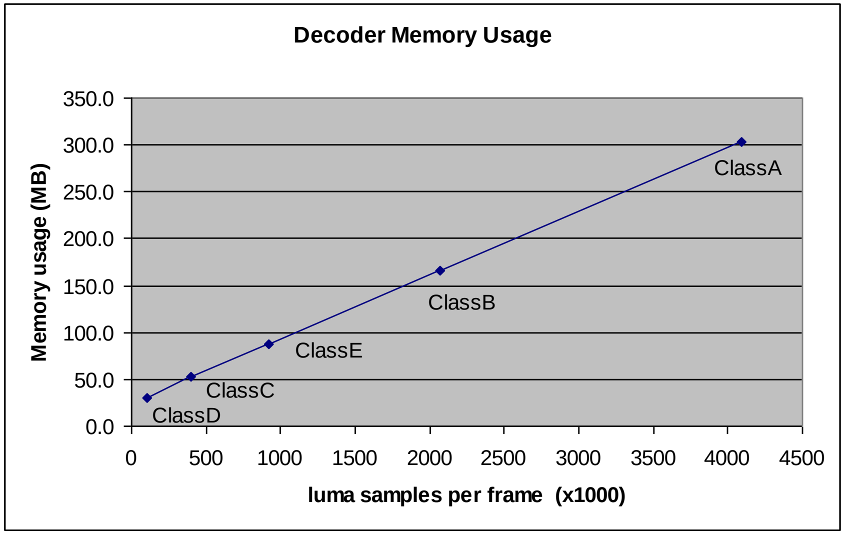
	Time ratio CDCM / JM
CS1 / alpha	11.1
CS2 / beta	14.7
CS2 / gamma	15.8

Performance & Complexity analysis

Expected encoder /decoder memory usage

- Memory usage were measured using linux “top” command





Outline

- Proposal overview
- Coding tools
- Encoder control
- Performance & Complexity
- Conclusion

Conclusion

- CDCM codec is proposed in response to the Joint CfP on Video Compression Technology
- High performance
  - Superior subjective quality (much better than JM, better than KTA)
  - Excellent objective quality (more than 30% compared to the JM anchors)
  - Better objective and subjective chroma performance observed
- Similar quality improvements proving that it is robust, adaptable and not content-specialized
  - for CS1, CS2
  - across all resolutions and sequence types
- The basic design of the CDCM codec leaves lots of potential for
  - other decoder tools to be added during the standardization phase
  - further encoder optimization after the standardization phase

Conclusion

- Attention on parallelizability
  - illustrated by the multi-threaded decoding capabilities of the software
- Well documented proposal
  - decoder specification,
  - user manual,
  - reference encoder,
  - Doxygen support of the source code
- Stable and robust encoder and decoder reference softwares
  - tested by many companies,
  - with many compilers,
  - using many platforms
- CDCM is a good candidate for the JCTVC Test Model

Annex1: Improvements

Annex1: Improvements

Improvements after JCfP bitstreams submission

- Improvement of KLT basis function
  - less blocky
- Simplification of Lambdas setting
  - 7% average BD-rate improvements for chroma components without any significant change for luma components
- Support of intra16x16 in P/B frames
  - improvement of chroma

Annex1: Bug fixes

Bug fixes

- Encoder only (not signaled in bitstream)
  - interpolation of (7/8th, 7/8th) position of 1/8th pel MC corrected
  - creation of candidate MVs in motion estimation process corrected
  - availability of 8x2 mode for intra prediction corrected
  - sip_type cost now considered when choosing between 4x4, 8x2, 2x8 and 8x8 modes for intra coding
  - bit-cost computation of block coding mode corrected

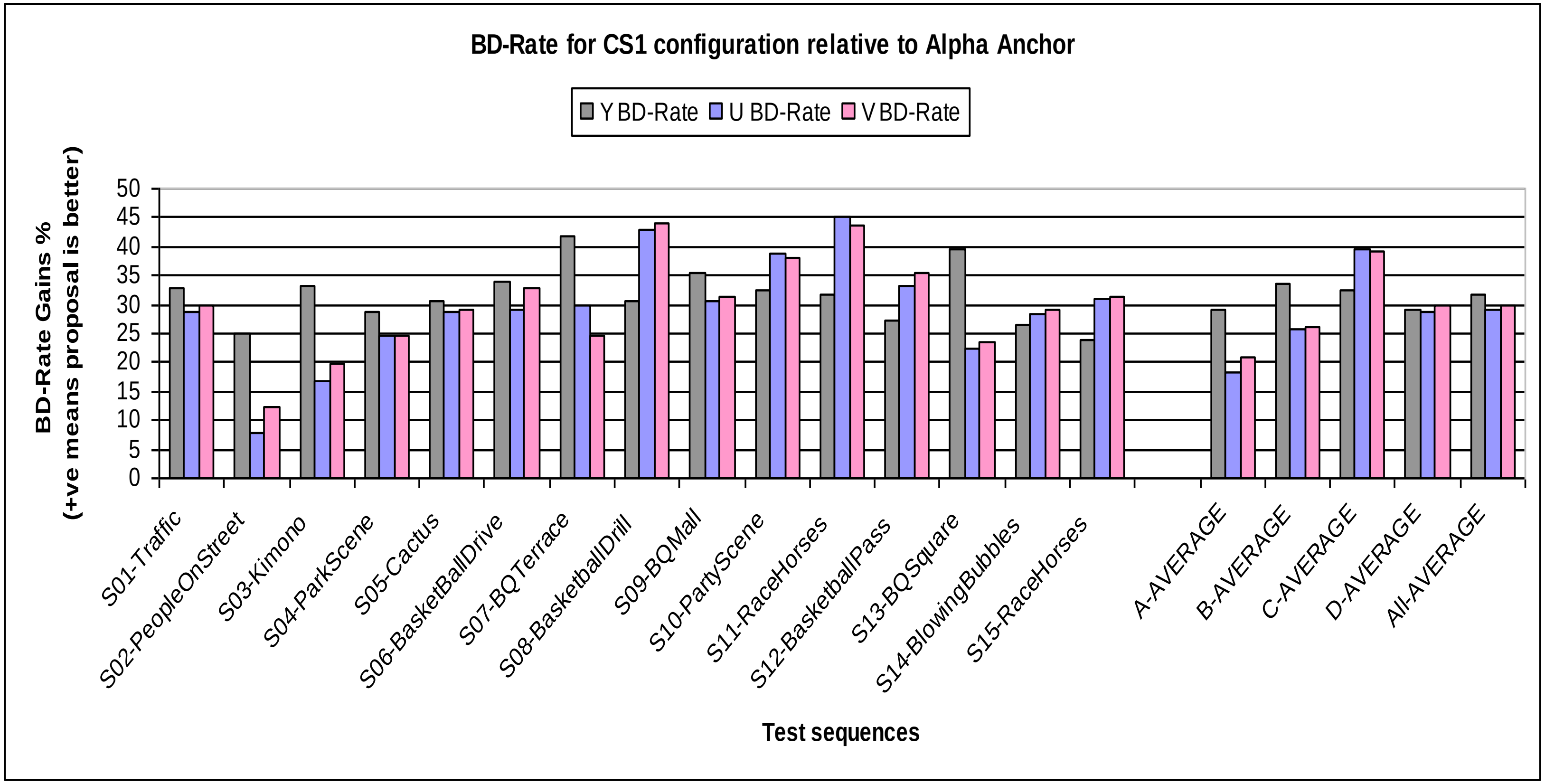
- Decoder related
  - A bug that causes an overflow after inverse quantization is corrected
    - Possible impact on decoded sequence

Annex2: detailed objective performance analysis

Annex2: detailed objective performance analysis

Objective performance

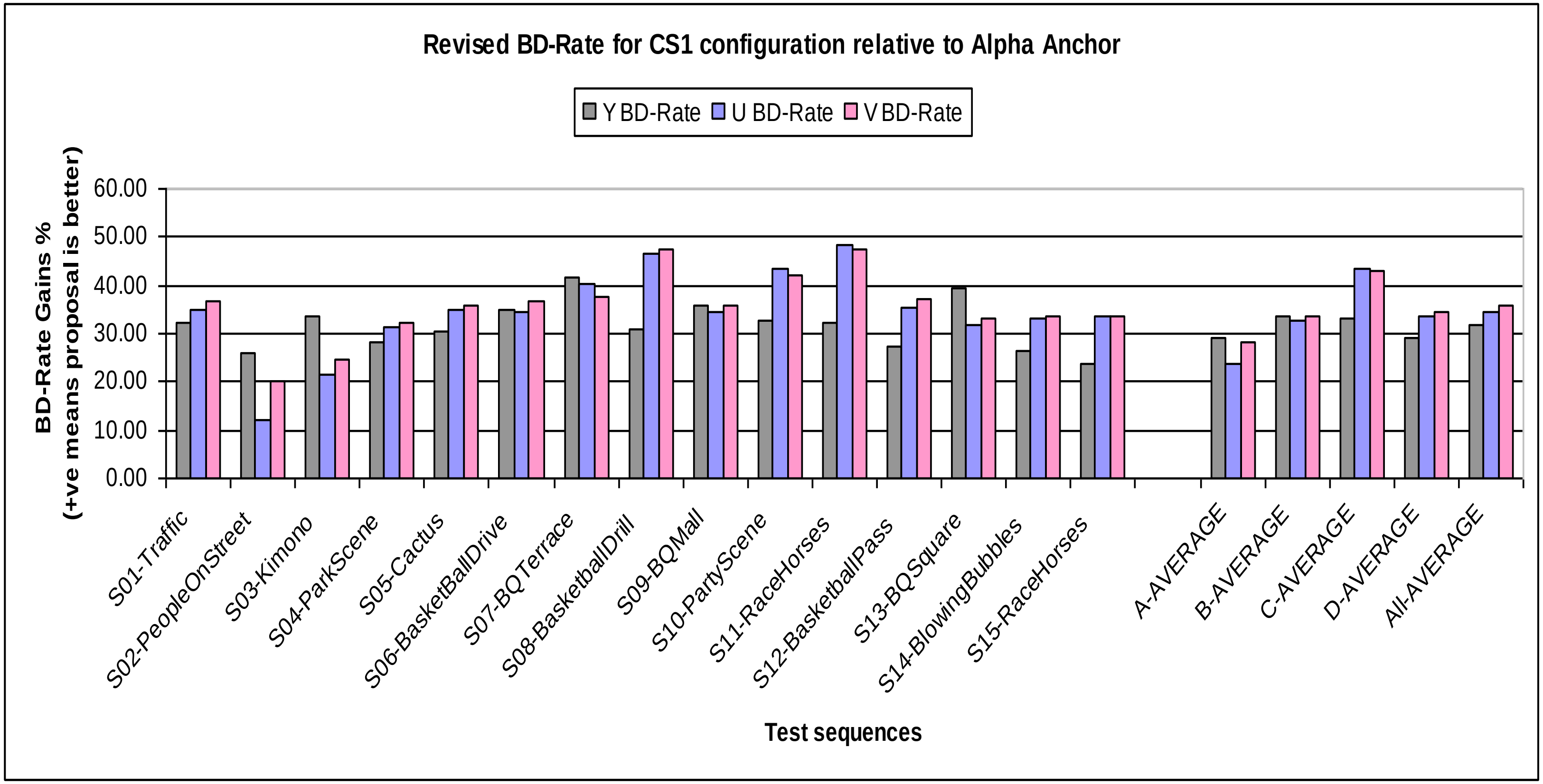
- BD-Rate for CS1 configuration relative to Alpha Anchor
  - Average: Y 31.6% - U 29.2% - V 30.0%



Annex2: detailed objective performance analysis

Objective performance

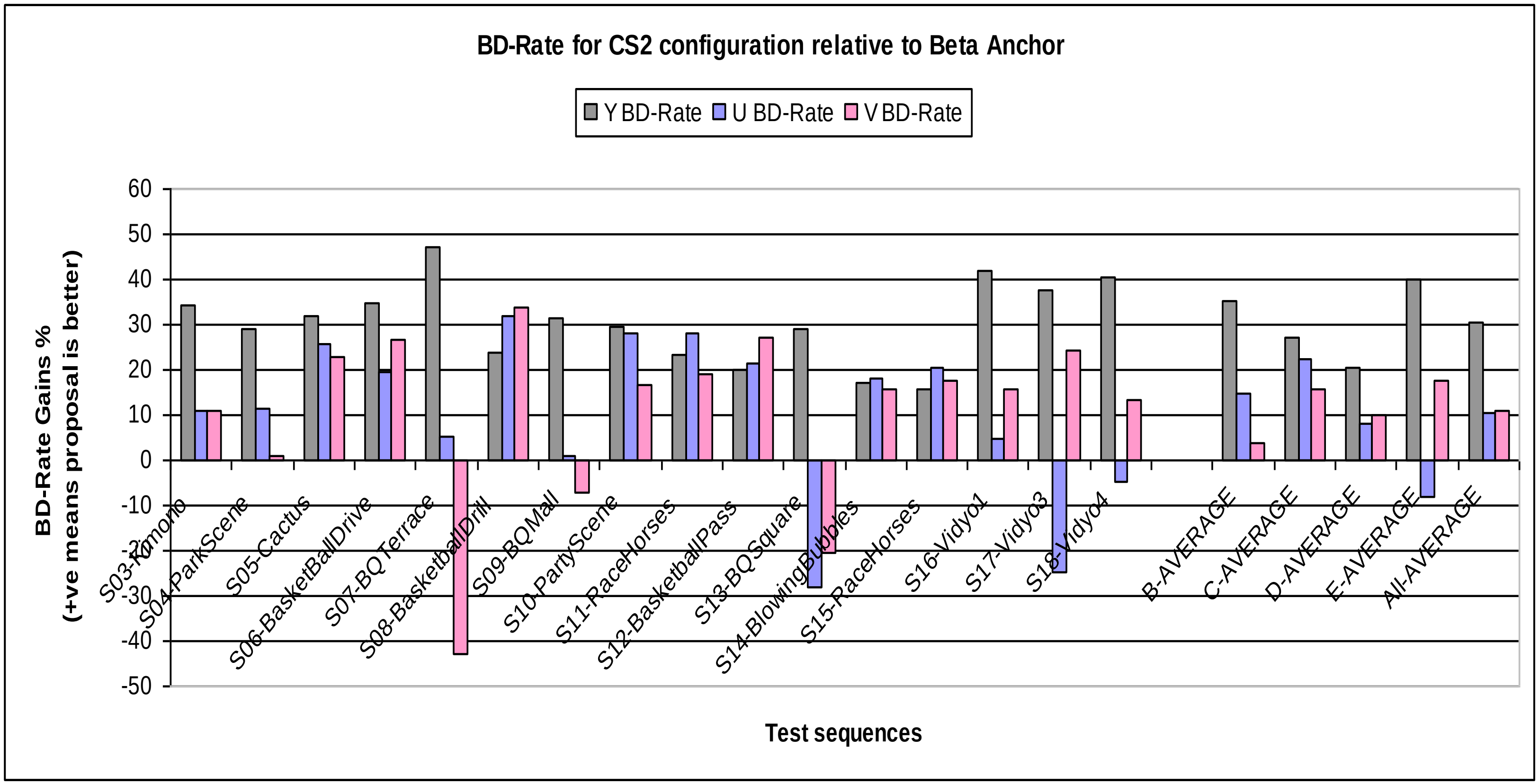
- BD-Rate for CS1 configuration relative to Alpha Anchor – New results
  - Average: Y 31.7% (31.6%) - U 34.5% (29.2%) - V 35.6% (30.0%)



Annex2: detailed objective performance analysis

Objective performance

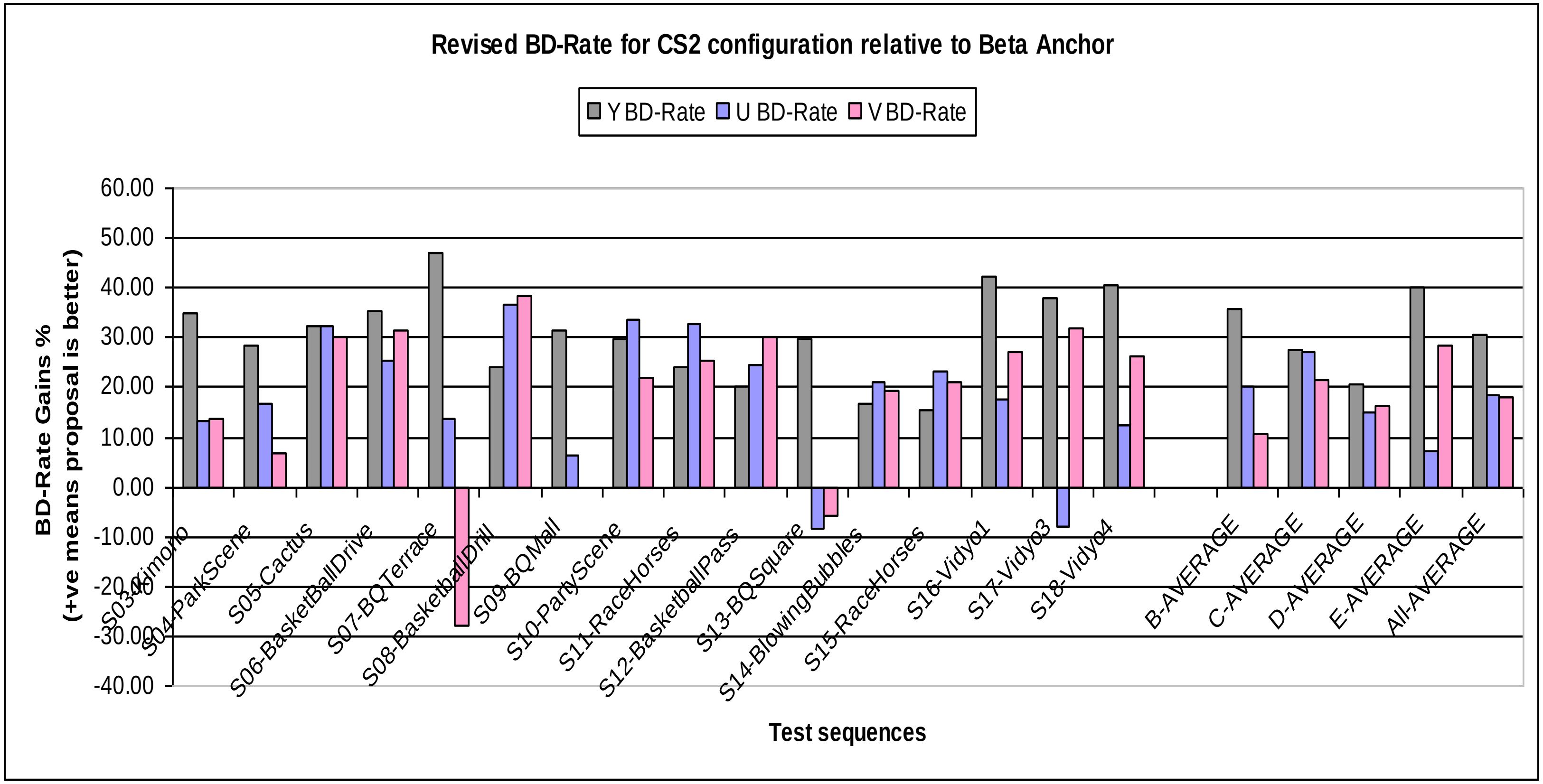
- BD-Rate for CS2 configuration relative to Beta Anchor
  - Y 30.4% - U 10.6% - V 10.9%



Annex2: detailed objective performance analysis

Objective performance

- BD-Rate for CS2 configuration relative to Beta Anchor – New results
  - Y 30.6% (30.4%) - U 18.3% (10.6%) - V 18.1% (10.9%)

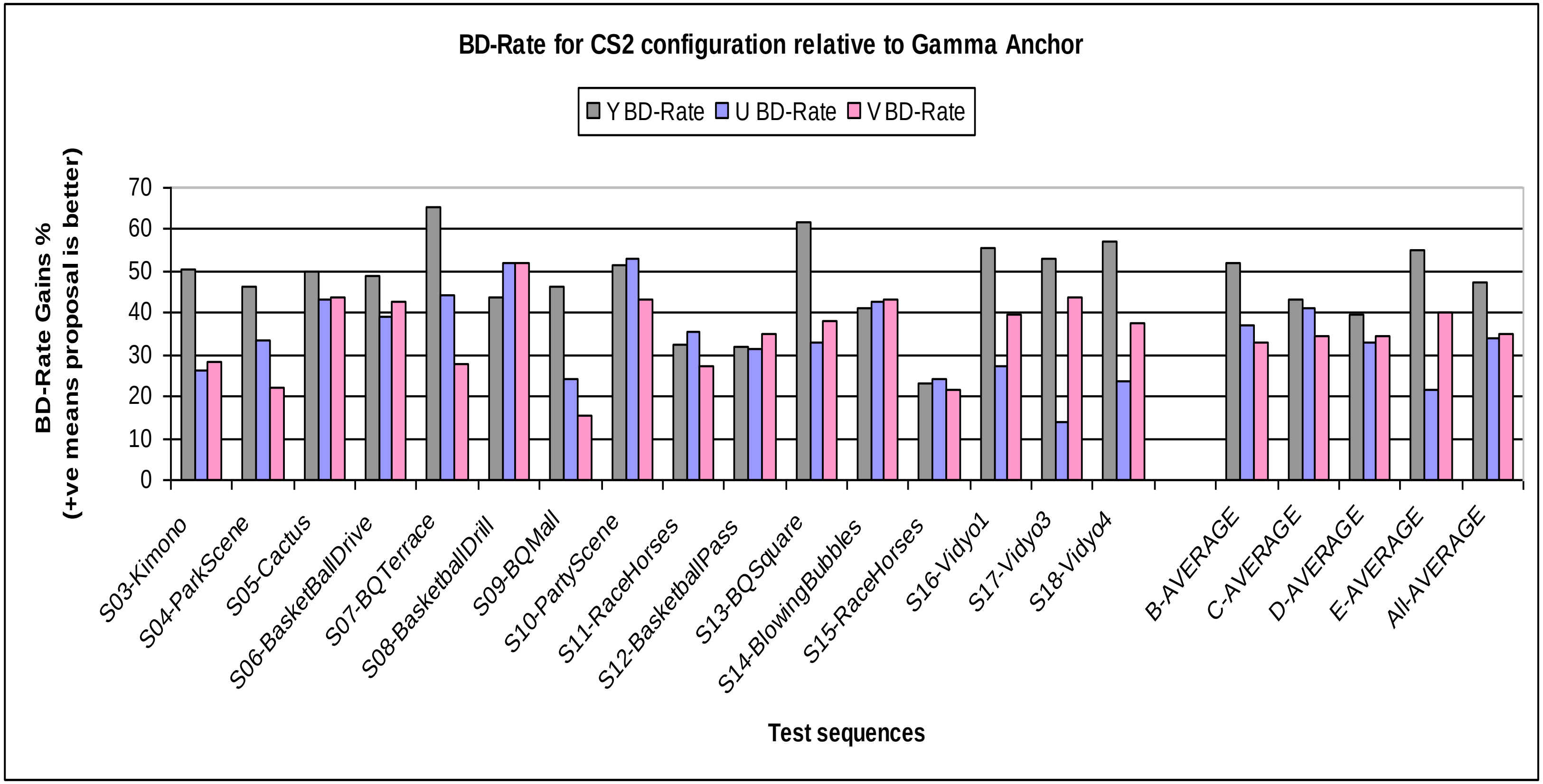


NTT docomo orange Panasonic ideas for life technicolor

Annex2: detailed objective performance analysis

Objective performance

- BD-Rate for CS2 configuration relative to Gamma Anchor
  - Y 47.4% - U 34.1% - V 35.1%



Annex2: detailed objective performance analysis

Objective performance

- BD-Rate for CS2 configuration relative to Gamma Anchor – New results
  - Y 47.6% (47.4%) - U 39.2% (34.1%) - V 40.0% (35.1%)