Overview of VP-Next

A *Next Generation Open Video Codec*

Objectives & Progress

Adrian Grange
Chrome Media Group, Google

Introduction:
The WebM Project

- Goal:
  Dedicated to develop a high-quality, open video format for the web that is freely available to everyone.
- Open Source and User experience
  Google is dedicated to the concept of the open web platform, which leads to faster innovation and better user experience
- The project is a continuous effort to improve the Web platform -> next generation open video codec.

Introduction:
VP-Next

- Experimental branch at the WebM project
  - Started in earnest late in Q3 of 2011.
  - Start with basic VP8 building blocks, but everything is up for change.
  - Substantial progress made already in terms of compression efficiency.
- Focus
  - Scale better for larger images: HD is the new Sweet Spot
    - But don't neglect small resolutions and low bandwidth use cases.
  - Decoding Complexity should be reasonable for constrained devices.
  - Extended quality range from lossless to lossy.

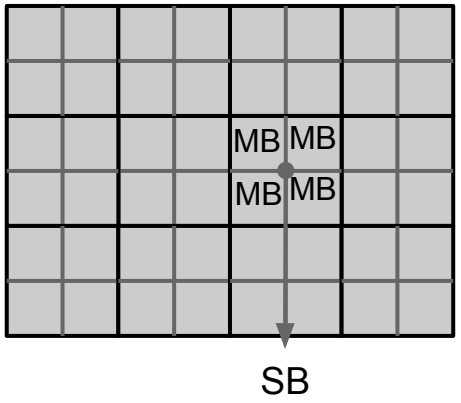
Overview:

List of Techniques

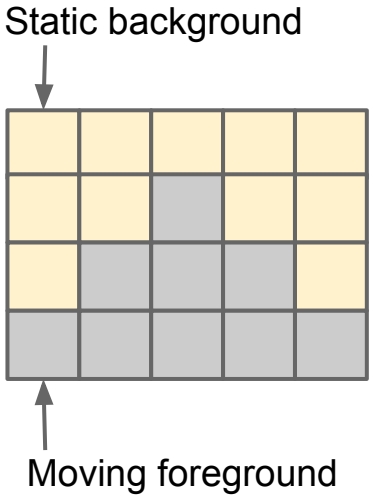
- Super-blocks
- Segmentation
- Enhanced Intra-Prediction
- Compound Prediction
- Prediction filtering
  - Non-interpolating
  - Interpolating for sub-pel motion
- New Transforms
  - DCT 8x8, 16x16
  - ADST 4x4, 8x8, 16x16
- Entropy Coding
- New MV references and coding
- New Loop filtering

Techniques:

- **Superblocks** (SB) are introduced:
  - 32x32 in progress, 64x64 possible.
  - Aggregate coding parameters.
  - Exploits temporal coherence better
  - Expect substantial further improvements for HD content.

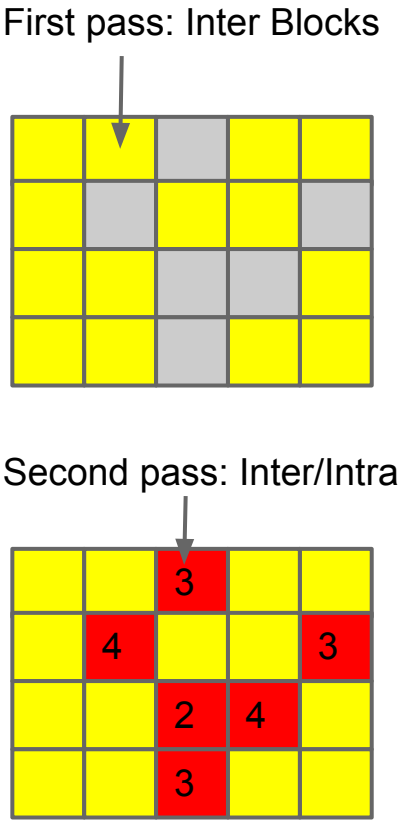


- **Segmentation** is significantly enhanced:
  - Group together MBs that share common characteristics into segments.
  - Encode segment at the MB level and control flags/features at segment level.
  - Differentially encoded from past frames.
  - Most benefit when the segmentation is temporally stable.
  - Unlocking the true potential requires a very smart encoder: Syntax provides a framework for innovation



Techniques:

- **Enhanced Intra**
  - Goal: to increase the quality of prediction for intra coded blocks within a mainly inter-coded frame.
  - 2-Pass encode:
    - First, encode all blocks that are clearly best with INTER modes.
    - Then, fill in INTRA macroblocks where they now have access to more "boundaries".
  - Pros: Good increase in quality for INTRA blocks
  - Cons: Decoding hardware more complex.



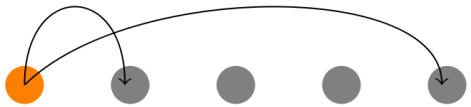
- **Compound prediction**
  - Combine two "single" predictors.
  - Currently, two inter predictors with the same mode and different reference frame are averaged to generate a new predictor.
  - Inter-intra & Intra-intra possibilities under investigation.

Techniques: Prediction Filters

- **Non-interpolating** A set of selectable filtering options:
  - No filtering vs Smoothing (Current implementation).
  - Bilateral filter to preserve edges.
  - Others under exploration (e.g. Rotation/Zoom).
  - Filter selection is indicated at the MB/SB level.
  - Upto 10% improvement with ~2% average gain for standard test sets.
- **Interpolating (sub-pel)** VP-Next introduces 8-tap filters:
  - 7-bit precision.
  - Regular & Sharp variants.
  - Selectable at frame and/or MB-level:
    - Frame: 8-tap Regular, 8-tap Sharp, 6-tap, Bilinear, *8-tap Switchable.*
    - MB: Flag switches between the 8-tap regular and 8-tap sharp.
  - High-precision MV:
    - 1/8th pel (frame-level flag selects).
    - useful mainly for slow motion.

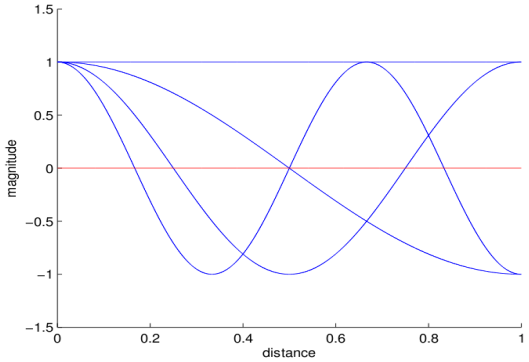
Techniques: New transforms

- **Larger DCT:** Critical for HD
  - 8x8 DCT (+ 2x2 2nd Order), 16x16 DCT.
  - Mode dependent transform selection (Initial Implementation).
  - RD Transform selection now implemented (4x4, 8x8, 16x16).
  - Sizeable improvement, especially on HD clips.
- **Asymmetric DST (ADST):**

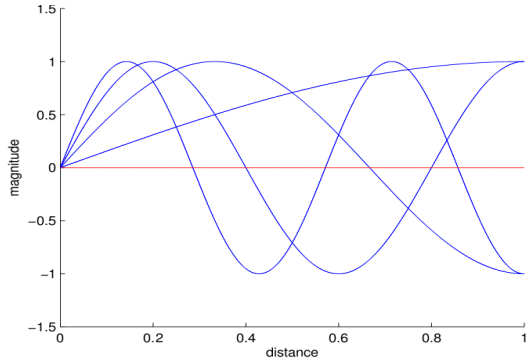


  - Basis functions optimal for intra prediction residuals given one-sided boundary.
  - Inherent smooth transition across prediction boundary reduces blocking artifacts.
  - ~6-15% bit-rate reduction for INTRA only, ~1-2% for video overall.

[Collaboration with UCSB]



Cosine bases.



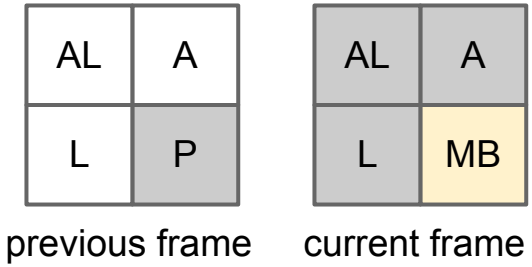
Asymmetric sine bases.

Techniques:

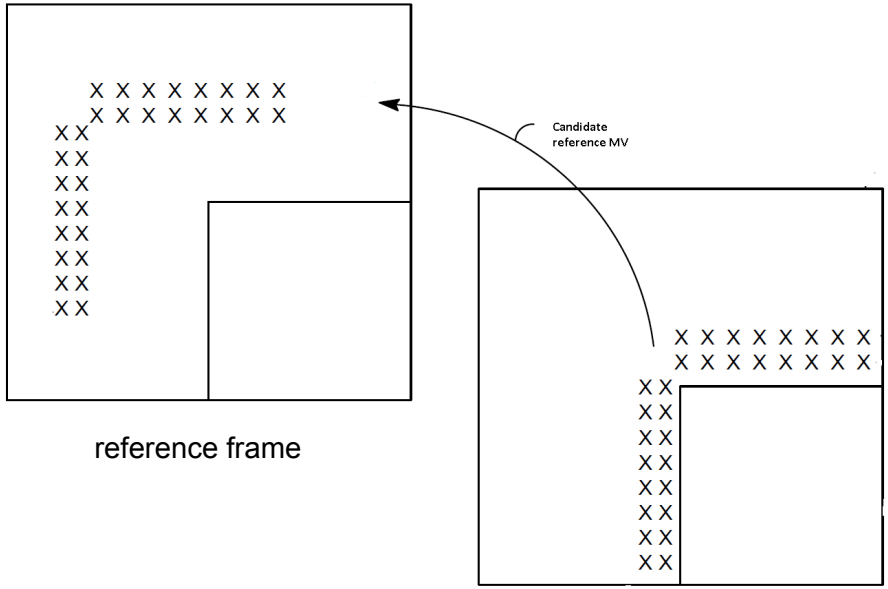
- **New transforms (Miscellaneous):**
  - Added support for a lossless mode using a reversible variant of the Walsh Hadamard transform.
  - Ongoing experiments:
    - Signaled transform size/type
    - Directional transforms
    - Prediction-dependent transform/coding
- **Entropy Coding**
  - Focus is on:
    - New predictive models and contexts.
    - More efficient updates / adaptation
  - A number of entropy coding improvements
    - Contextual MB skip coding
    - Reference frame contextual coding
    - Expanding the previous coef-contexts
    - Modifications to coding of explicit segment map (differential coding option and contexts)
    - Separate coding context for different frame types

Techniques:

- **MV Re-use**
  - VP-Next adds MV from co-located MB in Last frame to (Left, Above, Above-Left) as prediction options.



- **Ref MV Scoring**
  - Optimal reference MV selection.
  - Score candidate reference MVs using reconstructed pixels in close proximity of current MB.



Techniques:

- **MV differential encoding:**
  - Two part coding strategy:
    - A joint part where the (x, y) components are coded jointly up to a certain size.
    - Larger MVs are coded component-wise (x, y).
    - Integer and fractional pel part is separated.
- **New Loop Filter:**
  - Designed to reduce blocking
    - Cater for larger 8x8/16x16 transforms + ADST
    - Use filtering only across transform block boundaries
  - Flat block detection
    - use longer filter
  - To Do:
    - 16x16 filter, ADST

Conclusion

- Compression/Quality Improvements:
  - HD resolution: ~44%*.
  - CIF/SIF resolution: ~26%*.
  - *Compared to VP-Next baseline from Q4 2011.
  - Metric: Average % Bitrate saving at matching data-rate.
  - Initial limited tests indicate that, for HD material, VP-Next is ~ -7% compared to HEVC Main Profile (HEVC JM used for comparison).
- BUT, the bit-stream is far from finalized.
- Lots of potential for improvement:
  - New tools.
  - Optimizing existing tools.
- Goal: further reduce bit-rates at least 10-20% to move ahead of HEVC.
- Reference codebase is not yet ready for practical use:
  - New tools are being added & large-scale code refactoring in process.
  - The process of optimization has only very recently begun.