



The Daala Video Codec: Research Update

Nathan Egge <negge@mozilla.com>
(Xiph.org, Mozilla)

FOSDEM: Open Media devroom
31 January 2015



Why Free Codecs Matter

...that's "Free" with a capital F

- “Free” refers to control, not [just] cost
- Encumbered codecs are a billion dollar toll-tax on communications tools
- Codec licensing is used as weaponry in competitive battles
 - Licensing regimes are universally discriminatory
- The success of the Internet was based on innovation without asking permission



Why Free Codecs Matter

(continued)

...or begging forgiveness

- Many applications can't tolerate any codec licensing costs at all
 - even the cost of just counting the users is too much
 - Ignoring the licensing creates risks that can show up at any time
 - a tax on success
 - Compatibility is usually the big cost, not CPU, bandwidth, etc.
-



HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)



<http://xkcd.com/927/>

...but that's missing the usual motivations behind new codecs!



More and More Codecs

- An organization can't license an encumbered codec when there's no acceptable license offered
- Building a new codec from scratch may cost less than licensing
- Adversarial licensing is a risk in a competitive market
 - FRAND is often none of Fair, Reasonable, or Non-Discriminatory



Changing the Game

- Creating good codecs isn't easy...
 - But we don't need many. Without weird competitive pressures the whole world can cooperate
 - Best implementations of the patented codecs are already often the free software ones
 - Where RF is established non-free codecs see no adoption. See: JPEG. Network effect decides
 - Unfortunately many different people care about many different things
 - Convincing everyone means being better in almost every way, not just one or two
-



Strategy is Essential

- Design alternatives to avoid the worst patent thickets
- Read and analyze patents, and publish the results
- Patent the new technology we develop
- Use a patent license that encourages adoption and discourages defection
- Target next-next-generation to avoid rushing to market
- Document, document, document!
 - *“the whole point of a Doomsday Machine is lost if you keep it a secret.”*



Strategy is Essential: These Parts Will Be Hard

- Be best-in-class or go home
- Woo competitors and critics
 - especially those who think they're allies
- Find new niches, uses, applications that are unoccupied and fill them
- Hardware Support



Next Generation Video: Daala

- Lets take some of the strategy that worked in Opus, and apply it to video:
 - Work in a *public process* in a recognized SDO with a *strong IPR disclosure policy* and Opus-like patent licensing
 - Question assumptions in the conventional structure of video codecs, no sacred cows
 - Target applications where high flexibility is essential
 - optimize for *perception* not *PSNR*



30 Second Introduction to Video Coding

Most video codecs use the same basic ideas:

- **Prediction:** Consider what you know about previous or typical content to predict future data
 - **Transformation:** Rearrange the information to make it more compressible
 - **Quantization:** Strategically lower the resolution of the transformed data
 - **Entropy coding:** Code the quantized data taking probability distribution into account
-



30 Second Introduction to Video Coding: Prediction

- **Intra-Prediction:** Predict portions of the current frame from already decoded portions of the current frame
- **Inter-Prediction:** Predict portions of the current frame from previous decoded frames
 - **Motion Compensation** to eliminate temporal redundancy



Input

\ominus



Reference frame

=

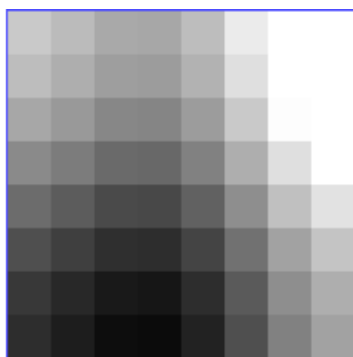


Residual



30 Second Introduction to Video Coding: Transformation

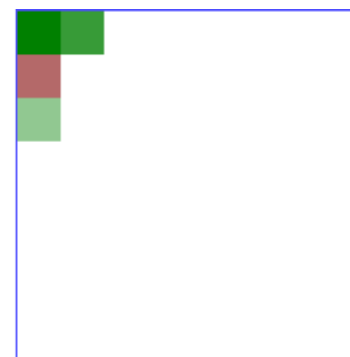
- Map spatial pixel values into some other more compressible representation via a 2D transform, usually the DCT.



Original pixel data

| | | | | | | | |
|-----|-----|-----|----|-----|-----|-----|-----|
| 114 | 108 | 100 | 99 | 109 | 129 | 152 | 166 |
| 109 | 102 | 95 | 94 | 104 | 124 | 146 | 161 |
| 99 | 93 | 85 | 84 | 94 | 114 | 137 | 151 |
| 86 | 80 | 72 | 71 | 82 | 102 | 124 | 138 |
| 73 | 66 | 58 | 57 | 68 | 88 | 110 | 125 |
| 60 | 53 | 46 | 45 | 55 | 75 | 97 | 112 |
| 50 | 43 | 36 | 35 | 45 | 65 | 88 | 102 |
| 45 | 38 | 31 | 30 | 40 | 60 | 82 | 97 |

DCT



DCT coefficient data

| | | | | | | | |
|------|-----|---|---|---|---|---|---|
| 700 | 200 | 0 | 0 | 0 | 0 | 0 | 0 |
| -150 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 110 | 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 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |



30 Second Introduction to Video Coding: Quantization and Coding

- **Quantization:** Compute the difference remaining after prediction, then lower its resolution.
 - This is the lossy part
- **Coding:** The quantized error signal is (hopefully) random numbers from some probability distribution.
 - Pack it efficiently into the bitstream



Daala Technological Differences

(so far)

- Lapped transforms rather than traditional DCT
 - Implemented via reversible lifting
- Multisymbol arithmetic encoding
- Perceptual vector quantization
- Chroma plane prediction from luma planes
- Overlapping-block motion compensation
- Time-frequency resolution switching



Recent Work / Updates

Technology demo pages document and explain many of these techniques in more detail:

Next generation video: Introducing Daala

Introducing Daala part 2: Frequency Domain Intra Prediction

Introducing Daala part 3: Time/Frequency Resolution Switching

Introducing Daala part 4: Chroma from Luma

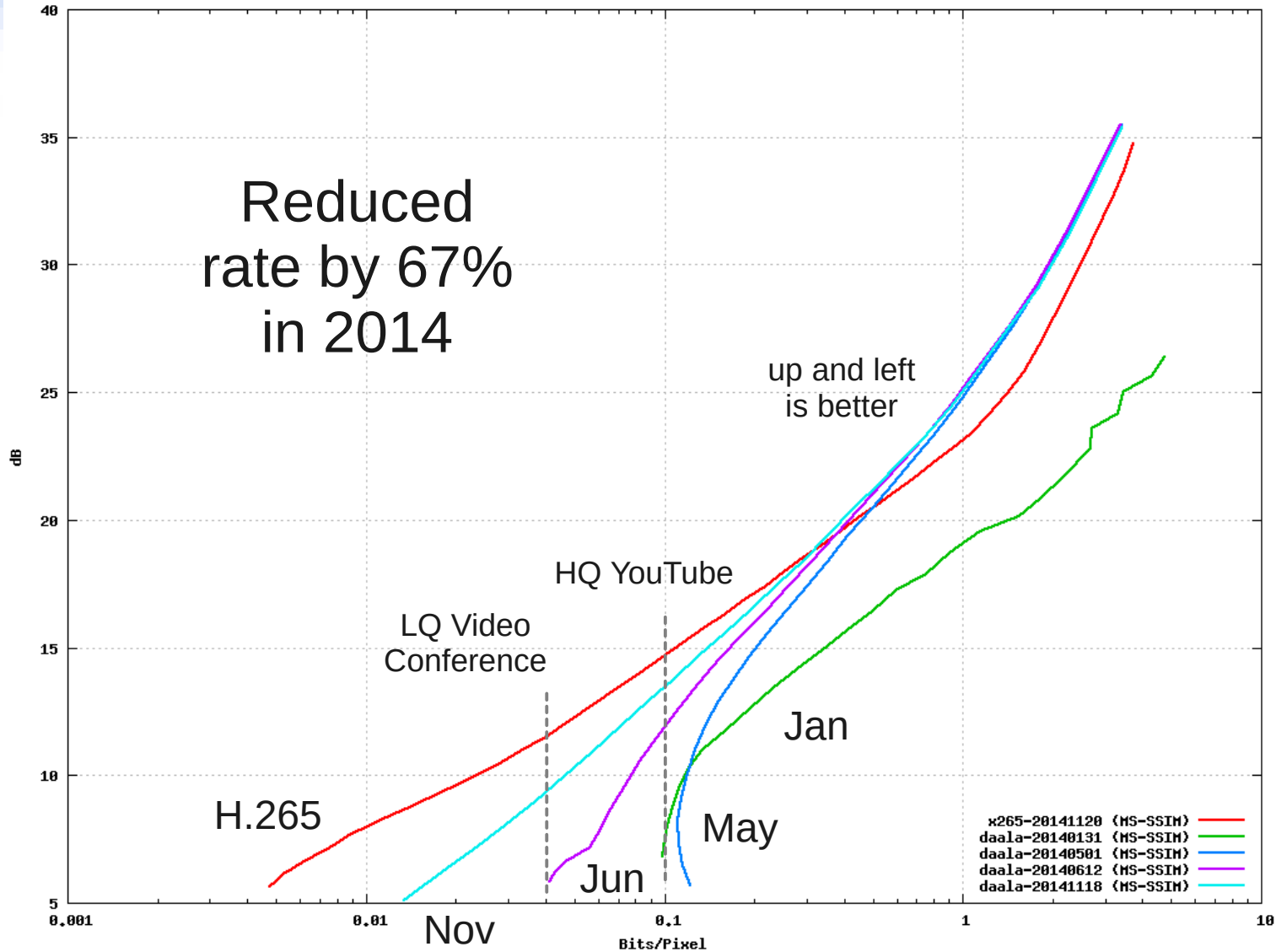
Daala demo 5: Painting Images For Fun (and Profit?)

Daala demo 6: Perceptual Vector Quantization (PVQ)





Daala Progress in 2014





Today's Formats Are a Long Way From Exhausting the Possible

How about unblending a cross-fade?



Spatial Sparsity-Induced Prediction for Images and Video: A Simple Way to Reject Structured Interference
Gang Hua and Onur G. Guleryuz (2011)



The Road Ahead

- The techniques we've been working with appear to work, but there is much to be done
- Industry is currently distracted figuring out how they're going to deploy HEVC (and VP9)
- Your participation is welcome!
 - <http://xiph.org/daala>
- Opus benefited from some applications served by no other audio codec.
 - Does something similar exist for video?



Daala: Additional Resources

- Wiki: <http://www.xiph.org/daala>
- Mailing list: daala@xiph.org
- IRC: #daala on irc.freenode.net
- Git repository: <git://git.xiph.org/daala.git>

Questions?