

Intra Coding in HEVC

Lainema, Jani, and Kemal Ugur. "**Angular intra prediction in high efficiency video coding (HEVC).**" *Multimedia Signal Processing (MMSP), 2011 IEEE 13th International Workshop on*. IEEE, 2011.

Manohar Kuse

mpkuse@connect.ust.hk

<http://ihome.ust.hk/~mpkuse>

General Introduction of HEVC (High Efficiency Video Coding)



What is HEVC?

- Video compression standard and successor to H.264/AVC
 - Jointly developed by 'ISO/IEC Moving Pictures Expert Group (MPEG)' & 'ITU-T Video Coding Expert Group (VCEG)'
 - MPEG & VCEG established 'Joint Collaborative Team on Video Coding (JCT-VC)' to develop HEVC
 - Also known as ISO/IEC 23008-2 (MPEG-H Part-2) and ITU-T H.265
-

Prominent Features

- Double compression rate for same video quality*
- Substantially better video quality for same bit rate*
- Support for 8K UHD resolution

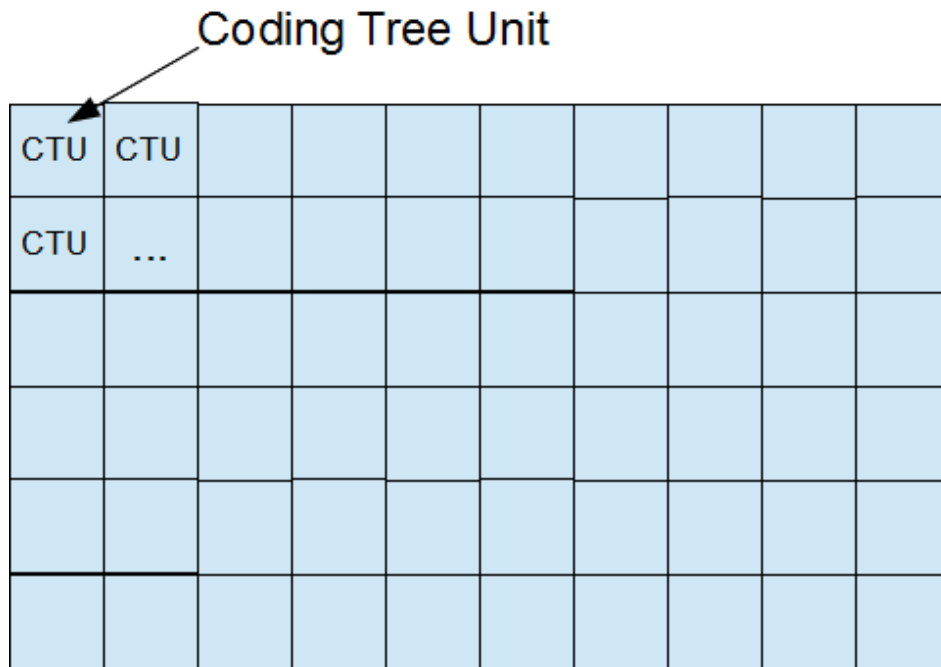
*Compared to H.264/AVC

The Standard Draft

- 1st Version of Standard draft published Early 2013
 - Download URL : <http://www.itu.int/rec/T-REC-H.265-201304-I/en>
- Extensions still under development (Early 2014)
 - Scalable Coding Extension
 - 3D Video Extensions
 - Screen Content Extension

HEVC – What are CTU, CU, CTB, CB, PB, and TB?

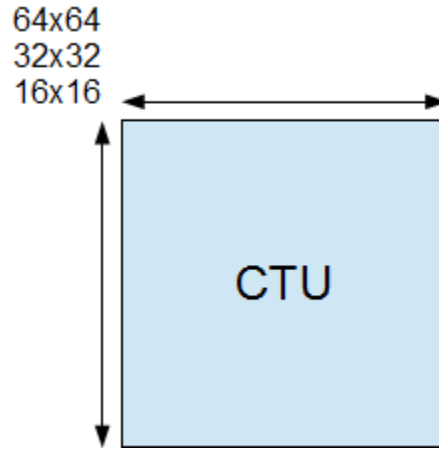
HEVC divides picture into CTUs*



CTU is an acronym for **Coding Tree Unit**

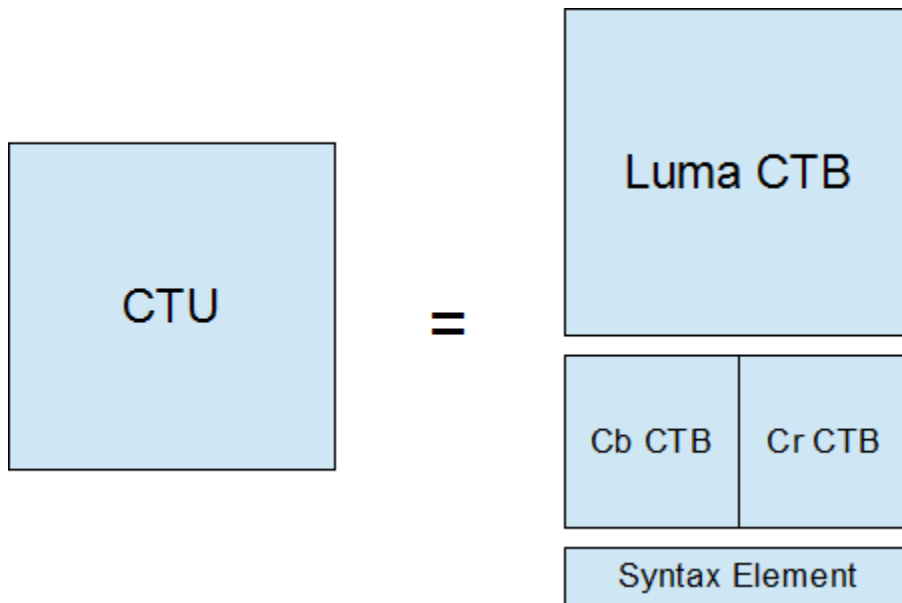
CTU (Coding Tree Unit) - Things to Remember

- Dimensions of CTU is fixed for the entire video sequence



xxxUnit Vs. xxxBlock

xxxUnit \Rightarrow Luma block, Chroma block, Syntax elements together



Coding Tree Unit (CTU)

Coding Unit (CU)

Prediction Unit (PU)

Transform Unit (TU)

Coding Tree Block CTB)

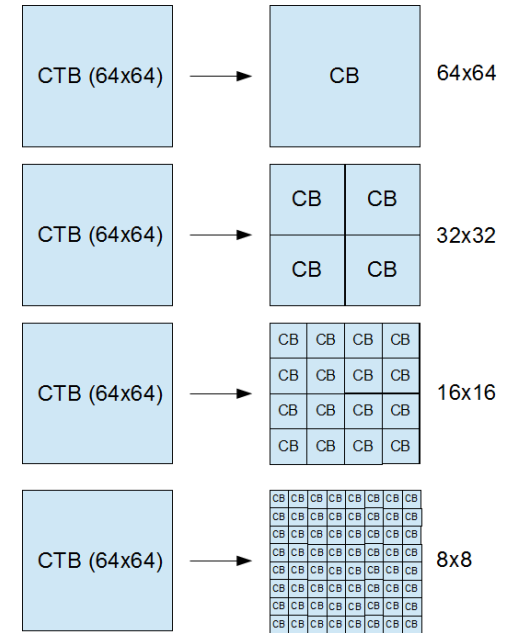
Coding Block (CB)

Prediction Block (PB)

Transform Block (TB)

CTB is split to form CB (Coding Block)

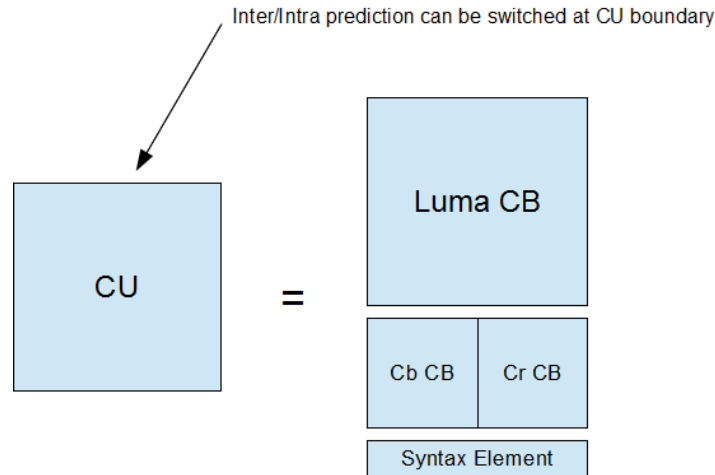
CTB too big to decide
whether to perform Inter or Intra prediction
(and select their prediction directions)



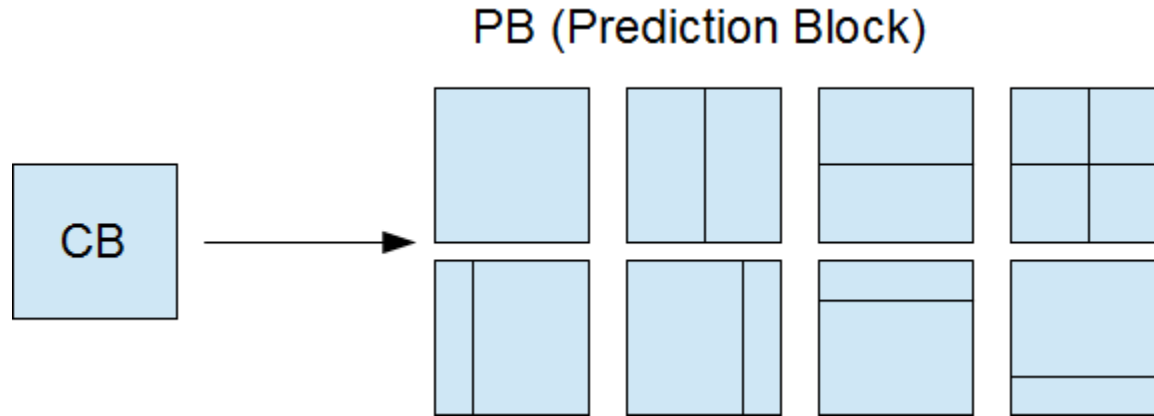
*Inter prediction only for P & B-frames

CU & CB

- Size of CU is good enough to switch between Intra- & Inter-prediction (P & B-frames only)
- Too large to store motion vectors (inter prediction) or intra prediction mode.



Prediction Block (PB)

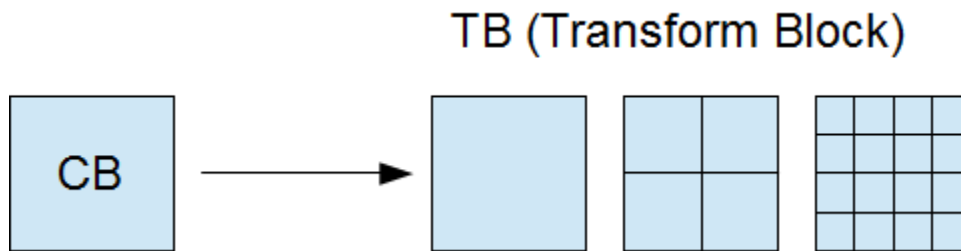


*Non square division of CB only for inter-prediction.

**Prediction direction is decided on per CU basis

Transform Block (TB)

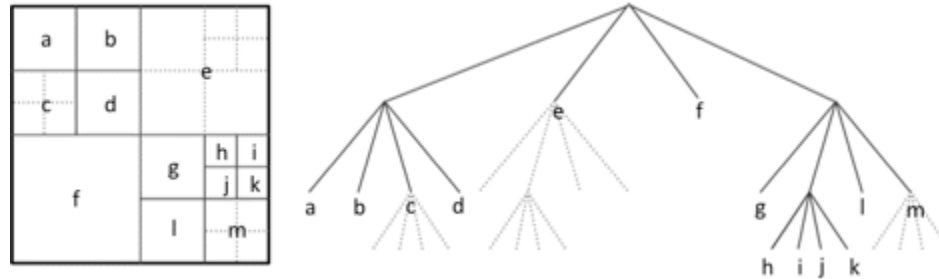
- Once prediction is made, the residue* need to be transformed with DCT-like transform
- CB may be too big, thus option to subdivide



*Difference between predicted block and actual block

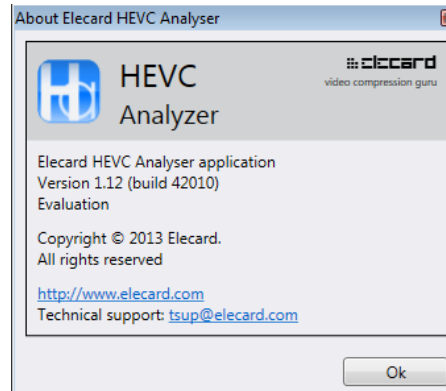
Quad-tree based block partitioning

- CTB \Rightarrow Bold lines
- CB \Rightarrow Solid lines
- TB \Rightarrow Dashed lines



Reference : <http://www.hhi.fraunhofer.de/fields-of-competence/image-processing/research-groups/image-video-coding/hevc-high-efficiency-video-coding/generic-quadtree-based-approach-for-block-partitioning.html>

Elecard HEVC Analyzer Demo



Elecard HEVC Analyzer Demo

win7 [Running] - Oracle VM VirtualBox

Elecard HEVC Analyzer v1.12 - C:\Users\Manohar\Desktop\HEVC_bin_files\mobile_intra.hevc

FILE REFERENCE VIEW OVERLAYS NAVIGATION HELP

stream

0 27

0/0 1/1 2/2 3/3 4/4 5/5 6/6 7/7 8/8

pixels

decoded

256 255 254 253 252 251 250 249 248 247 246 245 244 243 242 241 240 239 238 237 236 235 234 233 232 231 230 229 228 227 226 225 224 223 222 221 220 219 218 217 216 215 214 213 212 211 210 209 208 207 206 205 204 203 202 201 200 199 198 197 196 195 194 193 192 191 190 189 188 187 186 185 184 183 182 181 180 179 178 177 176 175 174 173 172 171 170 169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154 153 152 151 150 149 148 147 146 145 144 143 142 141 140 139 138 137 136 135 134 133 132 131 130 129 128 127 126 125 124 123 122 121 120 119 118 117 116 115 114 113 112 111 110 109 108 107 106 105 104 103 102 101 100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

256 255 254 253 252 251 250 249 248 247 246 245 244 243 242 241 240 239 238 237 236 235 234 233 232 231 230 229 228 227 226 225 224 223 222 221 220 219 218 217 216 215 214 213 212 211 210 209 208 207 206 205 204 203 202 201 200 199 198 197 196 195 194 193 192 191 190 189 188 187 186 185 184 183 182 181 180 179 178 177 176 175 174 173 172 171 170 169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154 153 152 151 150 149 148 147 146 145 144 143 142 141 140 139 138 137 136 135 134 133 132 131 130 129 128 127 126 125 124 123 122 121 120 119 118 117 116 115 114 113 112 111 110 109 108 107 106 105 104 103 102 101 100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

stream picture vps sps pps slice pixels prediction

hex viewer

Display: hexadecimal Text: ANSI Columns: Auto

0x00000001 00 00 01 40 01 0C 01 FF FF 01 60 00 00 03 00 00 ...8...99.....

0x00000011 03 00 00 03 00 00 03 00 00 F0 24 00 00 00 01 4286.....5

0x00000021 01 01 01 60 00 03 00 00 03 00 00 03 00 00 038.....

0x00000031 00 00 80 08 08 04 87 F9 7E 4B 1B 67 86 40 00 00 ...d...gH...

0x00000041 00 01 44 01 C1 90 95 B1 12 00 00 01 26 01 AF 19 ...D.A...s...

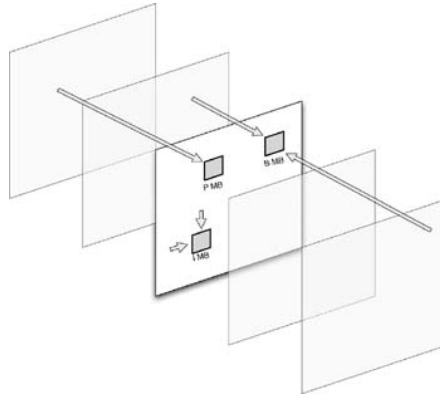
hex viewer stream viewer dpb buffer

index complete Strm 0 Disp 0 Type 1 Size 17002 Offset 0x00000001

11:21 PM 4/5/2014

Right Ctrl

Intra Prediction



General Motivation for Intra-prediction

- Neighbouring Pixel intensities are highly correlated
- Function of distance between the pixels

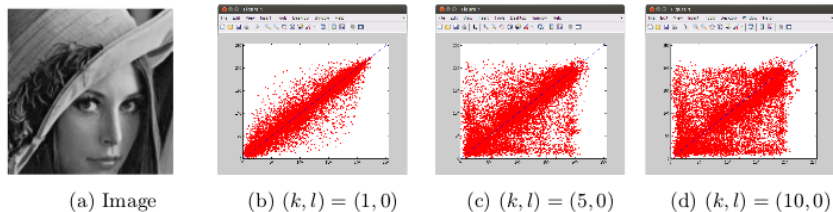


Figure 5: Scatter plots $(x_{i,j}, x_{i+k,j+l})$

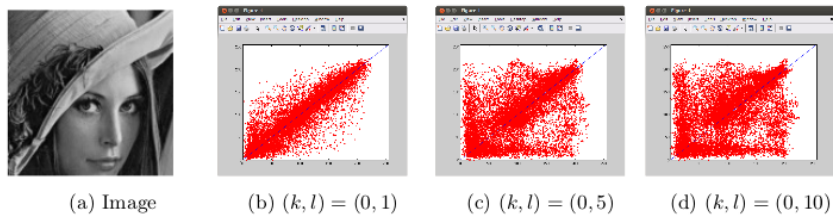


Figure 6: Scatter plots $(x_{i,j}, x_{i+k,j+l})$

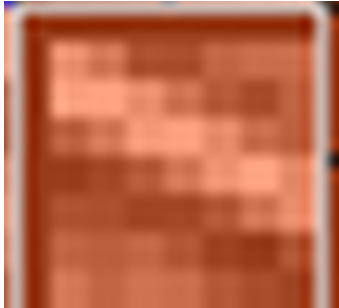
Horizontal & Vertical Correlation

Why difference coding result in better compression ?

1. Predict the unknown pixel intensities
 - a. from it's neighbours (this is justified due to correlation amongst the neighbours)
2. Code the difference

Motivation for having directional prediction

In presence of a dominant edge in a block,

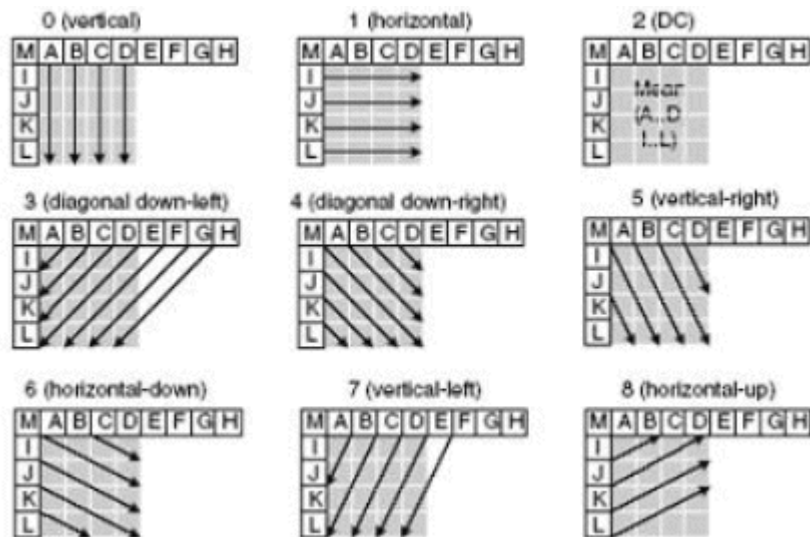


130	69	76	87	107	127	147	170
203	192	122	70	77	89	109	130
141	207	203	184	113	71	78	92
54	75	150	206	202	175	105	72
86	43	56	84	159	206	202	167
156	132	80	44	57	92	167	205
154	175	154	126	74	46	59	101
91	130	157	172	151	120	68	47

Highest correlation would be along the direction of the edge

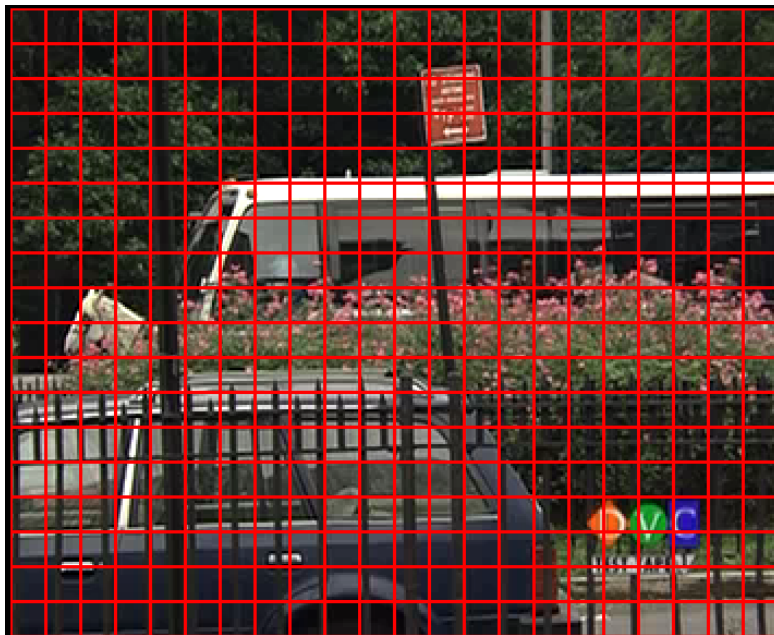
Intra-Prediction Modes in H.264

8 Angular Modes, 1 DC mode



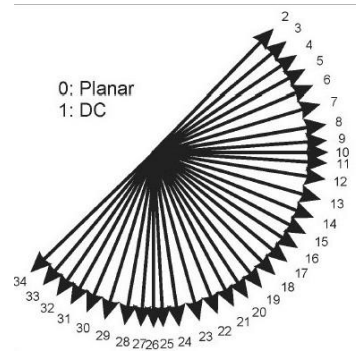
Not flexible enough

H.264 Macroblocks, 8 angular prediction modes



*With bigger CTB sizes in HEVC, having more prediction directions is imperative

Intra Prediction in HEVC & It's Efficient Computation

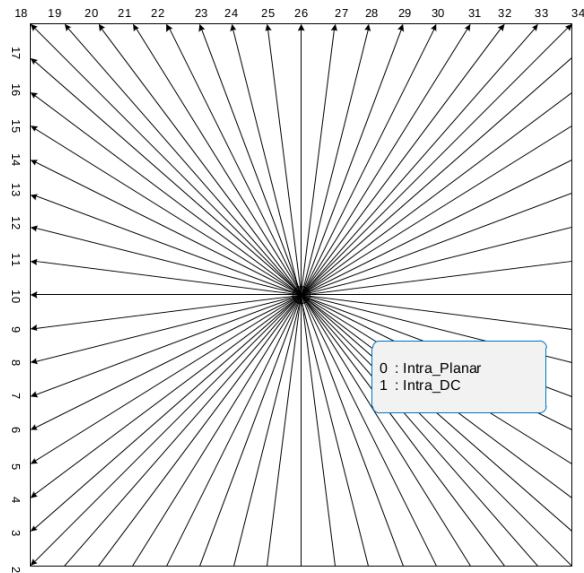


HEVC Intra-prediction

Introduces :

- 33 Angular modes
- 1 Planar mode
 - (was added later in 2012)
- 1 DC mode

with 1/32 pixel accuracy



*Figure above represents 1 pixel

Problem :

How to efficiently compute the predicted values?

Computing Predicted Samples

Predicted Sample : P_{xy}

Reference Sample : R_{xy}

“Predicted sample is obtained by
backward projecting its
location to the reference row
using selected prediction mode”

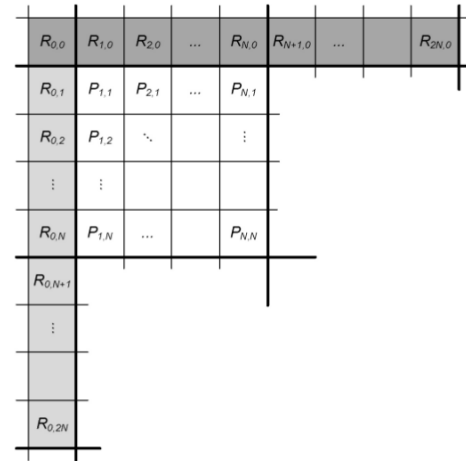


Fig. 1. Reference samples $R_{x,y}$ used in prediction to obtain predicted samples $P_{x,y}$ for a block of size $N \times N$ samples.

*Note: For efficiency purpose, HEVC uses only the reference row (for all vertical angular modes).
Also, HEVC uses only reference column (for all horizontal modes)*

Efficient Computation

$$P_{x,y} = ((32 - w_y) \cdot R_i + w_y \cdot R_{i+1} + 16) \gg 5$$

$$c_y = (y \cdot d) \gg 5$$

$$w_y = (y \cdot d) \& 31$$

$$i = x + c_y$$

$R_{0,0}$	$R_{1,0}$	$R_{2,0}$...	$R_{N,0}$	$R_{N+1,0}$...	$R_{2N,0}$
$R_{0,1}$	$P_{1,1}$	$P_{2,1}$...	$P_{N,1}$			
$R_{0,2}$	$P_{1,2}$	\ddots		\vdots			
\vdots	\vdots						
$R_{0,N}$	$P_{1,N}$...		$P_{N,N}$			
$R_{0,N+1}$							
\vdots							
$R_{0,2N}$							

Fig. 1. Reference samples $R_{x,y}$ used in prediction to obtain predicted samples $P_{x,y}$ for a block of size $N \times N$ samples.

*These equations valid only for all vertical predictions (mode 19-34)

Note-1 : '>> 5' left shift by 5 is divide by 32 operation

Note-2 : '&31' is for modulo division operation

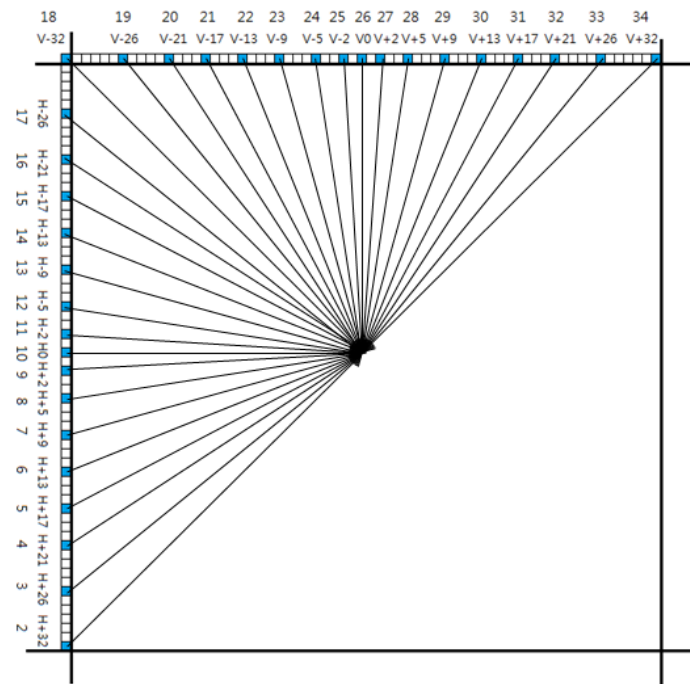
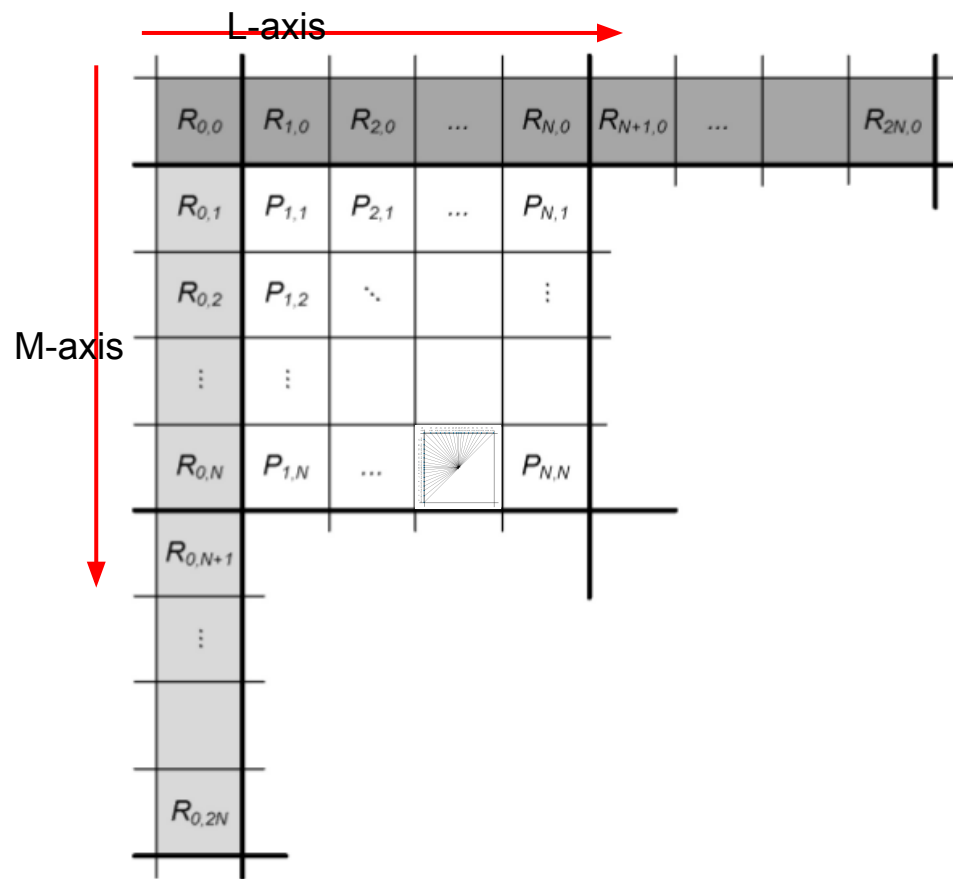
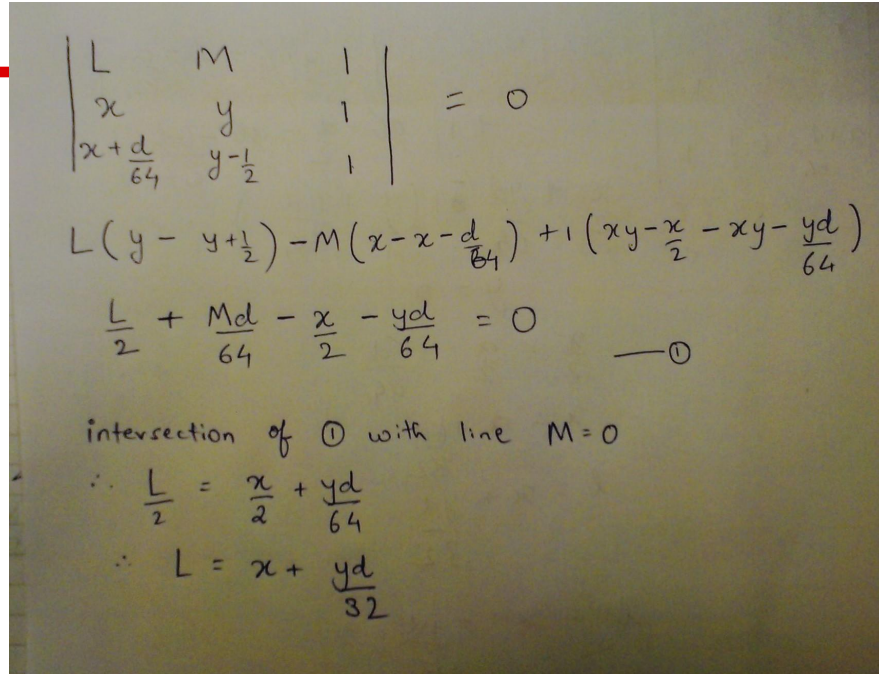


Fig. 1. Reference samples $R_{x,y}$ used in prediction to obtain predicted samples $P_{x,y}$ for a block of size $N \times N$ samples.

Deriving the equations



The image shows a piece of paper with handwritten mathematical work. At the top, a 3x3 determinant is set equal to zero. Below this, the determinant is expanded into a linear equation in terms of L. This equation is then simplified to a form labeled as equation ①. Finally, the intersection of this line with the line M=0 is found, leading to the expression for L.

$$\begin{vmatrix} L & M & 1 \\ x & y & 1 \\ x + \frac{d}{64} & y - \frac{1}{2} & 1 \end{vmatrix} = 0$$
$$L(y - y + \frac{1}{2}) - M(x - x - \frac{d}{64}) + 1(xy - \frac{x}{2} - xy - \frac{yd}{64})$$
$$\frac{L}{2} + \frac{Md}{64} - \frac{x}{2} - \frac{yd}{64} = 0 \quad \text{--- ①}$$

intersection of ① with line $M=0$

$$\therefore \frac{L}{2} = \frac{x}{2} + \frac{yd}{64}$$
$$\therefore L = x + \frac{yd}{32}$$

$$P_{xy} = (32 - w_y)/32 \cdot R_{\text{floor}(L)} + w_y \cdot R_{\text{floor}(L+1)}$$

Why the weighted average?

- 1- 'L' thus obtained may not be an integer. However, the available intensity values are only for integer values of 'L'.
 - 2- Thus, a solution could be to have a weighted average of 'floor(L)' and floor(L+1).
 - 3- Weights may be taken as the distance of L from floor(L) and floor(L+1) respectively.
 - 4- (3) can be realized with a modulo division operation.
- This justifies the equations show in HEVC standard document for intra-prediction.
-

Dealing with values outside reference rows

In this case, row is extended by projecting left reference column.

For performance reasons
copying rightmost value*

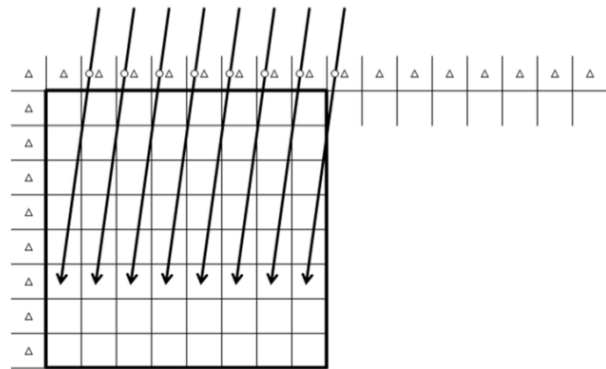


Figure 3. An example of angular prediction when operating on the sixth row of an 8x8 block with vertical prediction that utilizes a positive displacement value. Triangles indicate the reference pixels and circles indicate the projected fractional pixels at 1/32 pixel accuracy.

*Negligible effect for compression performance JCTVC-C046. [12]

Experimental Results

Experiment

Use H.264 modes (9) in HEVC Intra-main profile

vs

Use 33 Angular & 1 DC mode in HEVC Intra-main profile

Terminal

in.yuv

Türkiye Jokey Kulübü

9 10 11 12

```
Terminal
SUMMARY
-----
Total Frames | Bitrate   Y-PSNR   U-PSNR   V-PSNR
300          | 8559.8099 39.4652 41.5223 41.7081

I Slices-----
Total Frames | Bitrate   Y-PSNR   U-PSNR   V-PSNR
300          | 8559.8099 39.4652 41.5223 41.7081

P Slices-----
Total Frames | Bitrate   Y-PSNR   U-PSNR   V-PSNR
0            | -nan      -nan      -nan      -nan

B Slices-----
Total Frames | Bitrate   Y-PSNR   U-PSNR   V-PSNR
0            | -nan      -nan      -nan      -nan

RVM: 0.000
Bytes written to file: 13375914 (8560.585 kbps)

Total Time: 11832.290_sec.
[eeuser@roadrunner bin]$
```

```
Terminal
encoded 600 frames, 16.53 fps, 15482.41 kb/s
[eeuser@roadrunner x264-snapshot-20140406-2245]$ file ../x264_out
../x264_out: JVT NAL sequence, H.264 video @ L 40
[eeuser@roadrunner x264-snapshot-20140406-2245]$ ./x264 --keyint 1 ../hd_vid.yuv
--output ../x264_out.264 --dump-yuv ../x264_recon.yuv --input-res 1920x1080 -
-fps 25 ../hd_vid.yuv
yuv [info]: 1920x1080p 0:0 @ 25/1 fps (cfr)
x264 [info]: using cpu capabilities: none!
x264 [info]: profile High, level 4.0
x264 [info]: frame I:600 Avg QP:29.09 size: 77412
x264 [info]: mb I  I16..4: 13.1% 66.6% 20.3%
x264 [info]: 8x8 transform intra:66.6%
x264 [info]: coded y,uvDC,uvAC intra: 62.4% 65.6% 31.7%
x264 [info]: i16 v,h,dc,p: 58% 28% 6% 9%
x264 [info]: i8 v,h,dc,ddl,ddr,vr,hd,vl,hu: 31% 24% 12% 4% 5% 6% 7% 5% 7%
x264 [info]: i4 v,h,dc,ddl,ddr,vr,hd,vl,hu: 32% 29% 7% 4% 6% 6% 7% 4% 5%
x264 [info]: i8c dc,h,v,p: 49% 25% 21% 5%
x264 [info]: kb/s:15482.41

encoded 600 frames, 14.73 fps, 15482.41 kb/s
[eeuser@roadrunner x264-snapshot-20140406-2245]$
```

Subjective measure - demo

HE (High Efficiency Mode)
 Uses CABAC
 Uses ALF (Adaptive loop filters)
 Bit depth is 10bits

LC (Low Complexity mode)
 Uses CAVLC
 no ALF
 Bit depth of 8bits

Resolution	Sequence	Configuration	
		HE	LC
2560x1600	Traffic	-4.1	-5.2
	PeopleOnStreet	-4.5	-5.5
	Nebuta	-1.6	-1.5
	SteamLocomotive	-1.5	-2.0
1080p	Kimono	-2.4	-2.7
	ParkScene	-1.0	-1.6
	Cactus	-5.9	-6.9
	BasketballDrive	-7.6	-9.0
832x480	BQTerrace	-5.9	-7.3
	BasketballDrill	-8.6	-9.6
	BQMall	-4.0	-5.6
	PartyScene	-1.4	-1.8
416x240	RaceHorses	-3.3	-3.6
	BasketballPass	-4.5	-5.5
	BQSquare	-3.1	-3.4
	BlowingBubbles	-2.3	-2.8
720p	RaceHorses	-3.6	-4.0
	Vidyo1	-8.7	-10.5
	Vidyo3	-5.6	-7.5
	Vidyo4	-5.8	-7.2
Min/Max	Smallest gain	-1.0	-1.5
	Largest gain	-8.7	-10.5
Averages	4Kx2K	-2.9	-3.6
	1080p	-4.6	-5.5
	832x480	-4.3	-5.1
	416x240	-3.4	-3.9
	720p	-6.7	-8.4
	Average (all)	-4.3	-5.2

*All frames were intra coded

Improvement in coding efficiency in percentage (%).
 Calculated using Bjontegaard-delta measure (BD-bit rate)

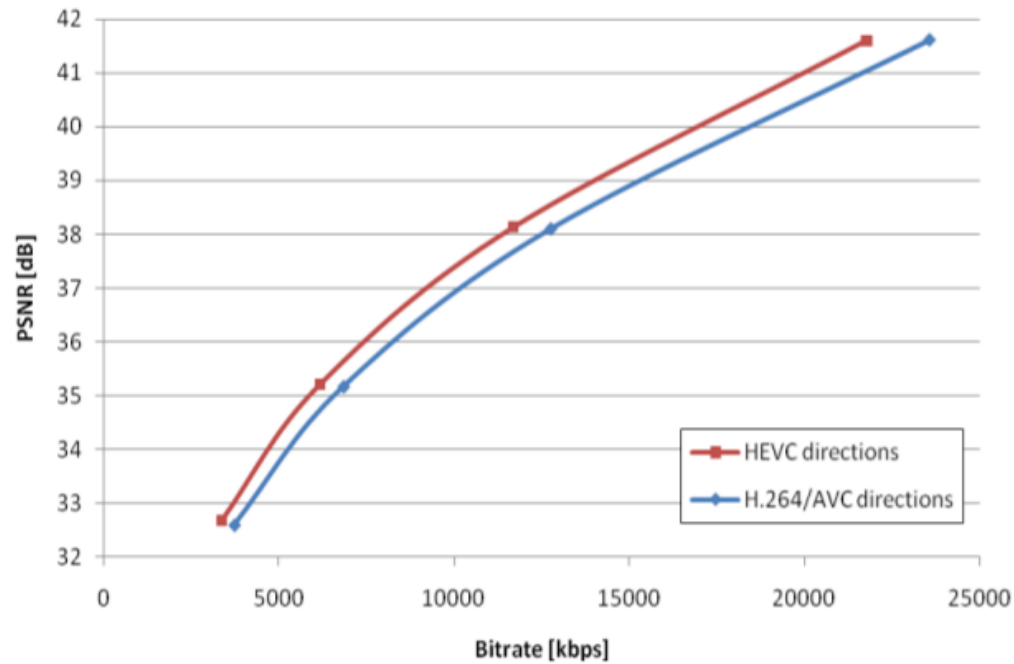


Fig. 4. Rate-distortion performance of the proposed method for the sequence Basketball Drill in low complexity configuration

1. For a fixed rate HEVC gives lesser distortion
2. For a fixed distortion HEVC gives smaller rate

References

Lainema, Jani, and Kemal Ugur. "Angular intra prediction in high efficiency video coding (HEVC)." *Multimedia Signal Processing (MMSP), 2011 IEEE 13th International Workshop on*. IEEE, 2011.

Lainema, Jani, et al. "Intra coding of the HEVC standard." *Circuits and Systems for Video Technology, IEEE Transactions on* 22.12 (2012): 1792-1801.

Section 8.4 of HEVC standard draft. <http://www.itu.int/rec/T-REC-H.265-201304-I/en>
