16×16 Integer Cosine Transform for HD Video Coding

Jie Dong and King Ngi Ngan

Department of Electronic Engineering
The Chinese University of Hong Kong

Pacific-Rim Conference on Multimedia 2006
Nov. 2-4, Hangzhou, China

Outline

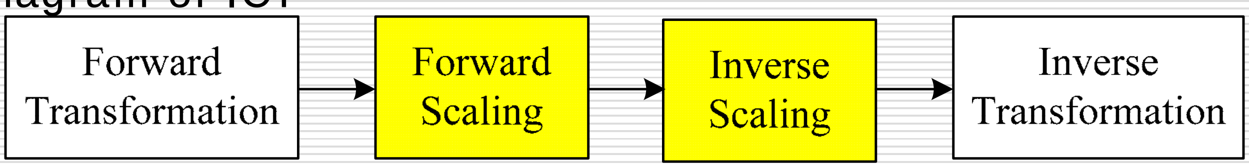
- Introduction
- The proposed 16×16 ICT
- Integrating into AVS
- Experimental results
- Conclusion

Introduction

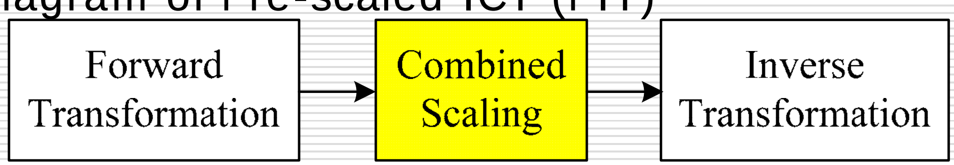
- HD video applications:
  - High density storage media, digital cinema, movie industry
- Considerations of HD video compression
  - High fidelity
  - Coding efficiency
  - Coding complexity
- Property
  - Rich details
  - Large homogeneous regions
- Using variable block-size transforms (VBT) according to the local activities

Review of Integer Cosine Transform (ICT)

- Generate an ICT
  - Replace the real-numbered elements of the DCT matrix with integers
  - Maintain the structure in the DCT matrix
    - relative magnitudes and signs
    - dyadic symmetry
    - orthogonality
- Advantages of ICT
  - Integer arithmetic implementation
  - Avoid mismatch between encoder and decoder
  - Good energy compaction capability if well-designed
- Flow diagram of ICT



- Flow diagram of Pre-scaled ICT (PIT)



The Proposed 16×16 ICT

- 16-bit implementation
- Modify the dyadic symmetries of the DCT matrix
- Good performance
- PIT, so simple for decoders
- Compatible with the 8×8 ICT in AVS
  - Compatible transformation
  - Compatible scaling

$$
\begin{bmatrix}
8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\
11 & 11 & 11 & 9 & 8 & 6 & 4 & 1 & -1 & -4 & -6 & -8 & -9 & -11 & -11 & -11 \\
10 & 9 & 6 & 2 & -2 & -6 & -9 & -10 & -10 & -9 & -6 & -2 & 2 & 6 & 9 & 10 \\
8 & 6 & 4 & 1 & -11 & -11 & -11 & -9 & 9 & 11 & 11 & 11 & -1 & -4 & -6 & -8 \\
10 & 4 & -4 & -10 & -10 & -4 & 4 & 10 & 10 & 4 & -4 & -10 & -10 & -4 & 4 & 10 \\
11 & 9 & -11 & -11 & -4 & -1 & 8 & 6 & -6 & -8 & 1 & 4 & 11 & 11 & -9 & -11 \\
9 & -2 & -10 & -6 & 6 & 10 & 2 & -9 & -9 & 2 & 10 & 6 & -6 & -10 & -2 & 9 \\
1 & 4 & -6 & -8 & 9 & 11 & -11 & -11 & 11 & 11 & -11 & -9 & 8 & 6 & -4 & -1 \\
8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\
4 & -1 & -8 & 6 & 11 & -9 & -11 & 11 & -11 & 11 & 9 & -11 & -6 & 8 & 1 & -4 \\
6 & -10 & 2 & 9 & -9 & -2 & 10 & -6 & -6 & 10 & -2 & -9 & 9 & 2 & -10 & 6 \\
11 & -11 & -9 & 11 & -6 & 8 & 1 & -4 & 4 & -1 & -8 & 6 & -11 & 9 & 11 & -11 \\
4 & -10 & 10 & -4 & -4 & 10 & -10 & 4 & 4 & -10 & 10 & -4 & -4 & 10 & -10 & 4 \\
9 & -11 & 11 & -11 & -1 & 4 & -6 & 8 & -8 & 6 & -4 & 1 & 11 & -11 & 11 & -9 \\
2 & -6 & 9 & -10 & 10 & -9 & 6 & -2 & -2 & 6 & -9 & 10 & -10 & 9 & -6 & 2 \\
6 & -8 & 1 & -4 & 11 & -11 & 9 & -11 & 11 & -9 & 11 & -11 & 4 & -1 & 8 & -6
\end{bmatrix}
$$

16-bit Implementation

- Forward ICT
  - Forward Transformation

$$F_{16\times16} = \left(T_{16} \times f_{16\times16} \times T_{16}^{T} + 2^{6}\right) >> 7$$

  - Combined Scaling

$$S_{16\times16} = (F_{16\times16} \otimes P_{16\times16} + 2^{15}) >> 16$$

- Inverse ICT
  - Horizontal transformation

$$b_{16\times16} = \left(S_{16\times16} \times T_{16} + 2^{2}\right) >> 3$$

  - Vertical transformation

$$f_{16\times16} = \left(T_{16}^{T} \times b_{16\times16} + 2^{6}\right) >> 7$$

Good Performance

- Good energy compaction capability especially for homogeneous regions, because the first three basis vectors maintain the structure exactly the same as those of DCT
- Apply KLT, DCT and the proposed ICT to highly correlated signals, e.g. first-order stationary Markov source with correlation coefficient ρ close to 1. And transform efficiency η of the 3 transforms are compared in the table below.

$$\eta = \frac{\sum_{i=1}^{15} \left| S_{16\times16}(i,i) \right|}{\sum_{i=0}^{15}\sum_{j=0}^{15} \left| S_{16\times16}(i,j) \right|}$$

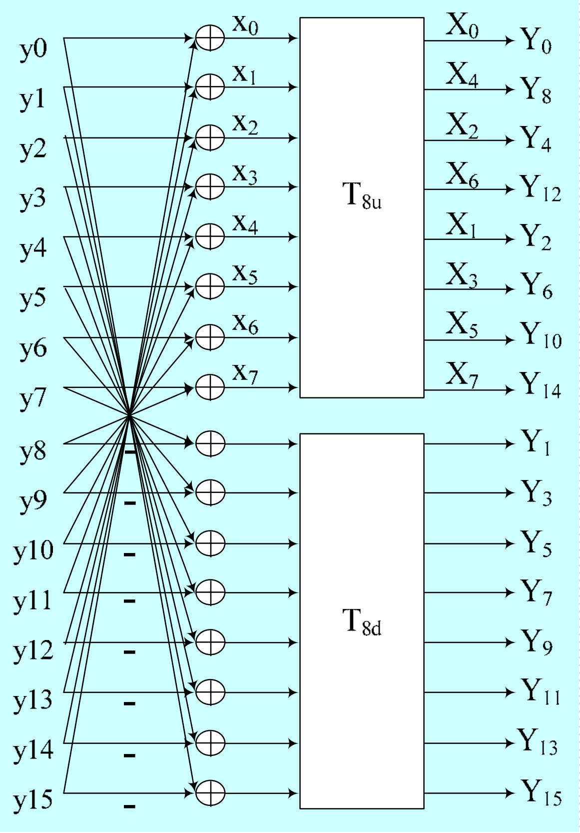
ρ	KLT	DCT	Proposed ICT
0.95	1.00	0.88	0.86
0.90	1.00	0.83	0.79
0.85	1.00	0.80	0.75

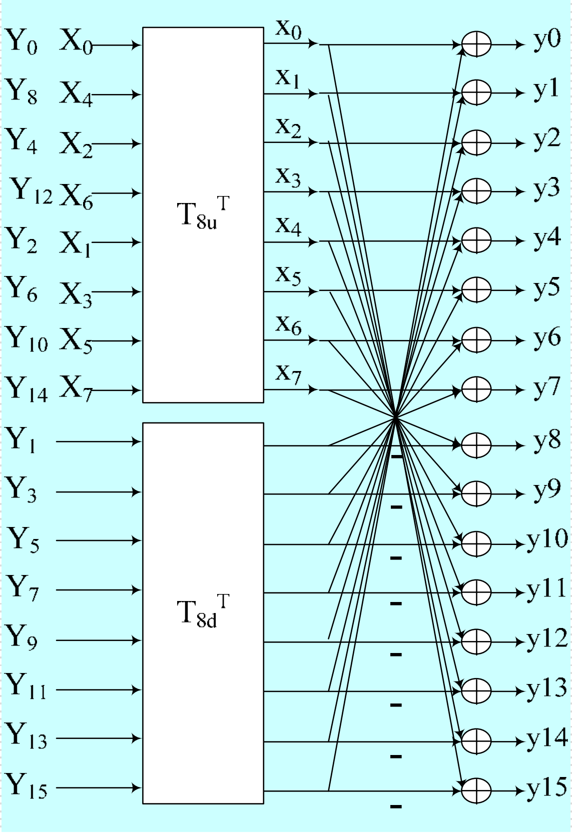
Compatibility with the 8×8 ICT

Compatible Transformation

$$T_{8u} = \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 10 & 9 & 6 & 2 & -2 & -6 & -9 & -10 \\ 10 & 4 & -4 & -10 & -10 & -4 & 4 & 10 \\ 9 & -2 & -10 & -6 & 6 & 10 & 2 & -9 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -10 & 2 & 9 & -9 & -2 & 10 & -6 \\ 4 & -10 & 10 & -4 & -4 & 10 & -10 & 4 \\ 2 & -6 & 9 & -10 & 10 & -9 & 6 & -2 \end{bmatrix}$$

$$T_{8d} = \begin{bmatrix} 11 & 11 & 11 & 9 & 8 & 6 & 4 & 1 \\ 8 & 6 & 4 & 1 & -11 & -11 & -11 & -9 \\ 11 & 9 & -11 & -11 & -4 & -1 & 8 & 6 \\ 1 & 4 & -6 & -8 & 9 & 11 & -11 & -11 \\ 4 & -1 & -8 & 6 & 11 & -9 & -11 & 11 \\ 11 & -11 & -9 & 11 & -6 & 8 & 1 & -4 \\ 9 & -11 & 11 & -11 & -1 & 4 & -6 & 8 \\ 6 & -8 & 1 & -4 & 11 & -11 & 9 & -11 \end{bmatrix}$$





Compatibility with the 8×8 ICT

Compatible Scaling

- The relationship between $P_{8\times8}$ and $P_{16\times16}$

$$P_{8\times8}(i, j) = P_{16\times16}(2i, 2j) \times 4$$

- $P_{16\times16}(2i, 2j)$ can be used as $P_{8\times8}(i, j)$, if N for the 8×8 scaling is 2 bits less than that for the 16×16 scaling.

Integrating into AVS

- AVS (Audio Video Coding Standard)
  - A national standard of China
  - The Zengqiang Profile is aiming at HD video coding
- Main features of the AVS Zengqiang Profile
  - Transform: 8×8 ICT
  - Intra prediction: 8×8
  - Motion compensation: from 16×16 down to 8×8
  - Arithmetic entropy coding

Transform Size Selection

- For every MB, try 2 types of transforms for luma residue
  - 1 16x16 transform
  - 4 8x8 transforms

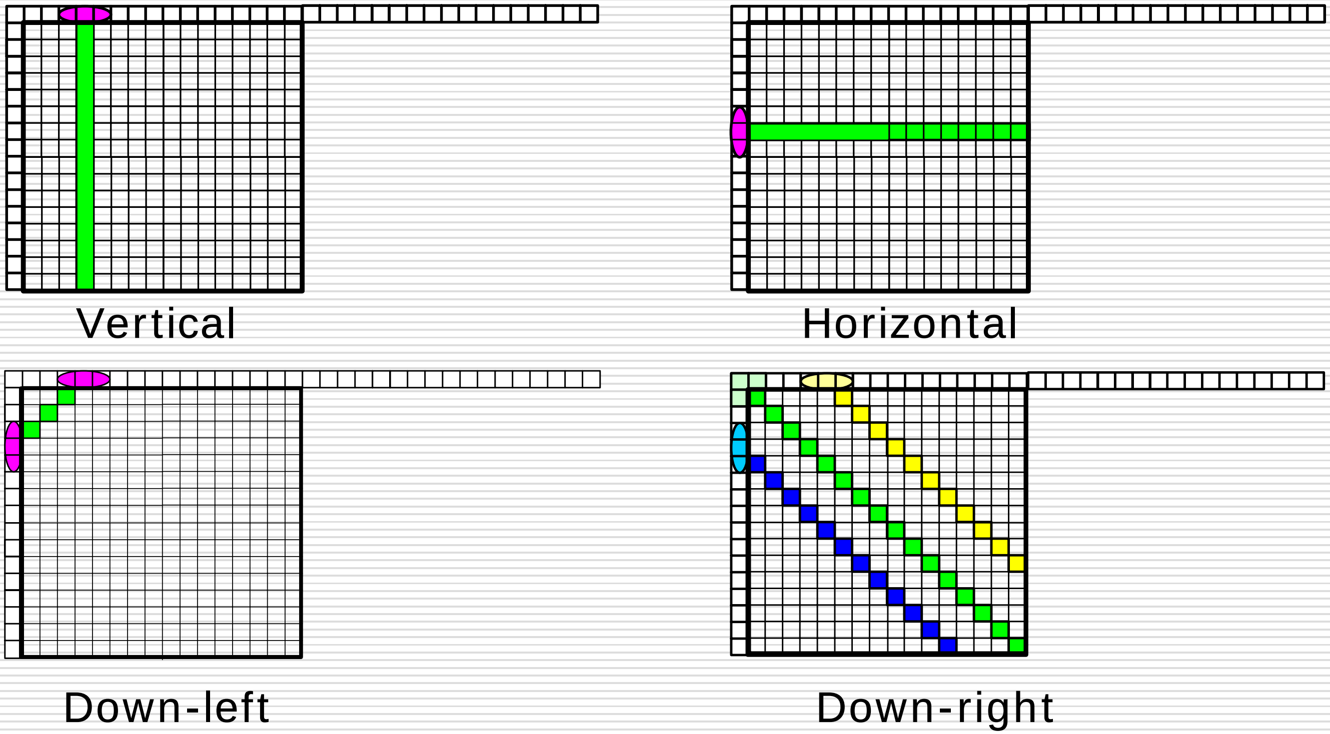
- Selecting the one with low cost

$$cost = SSD(MB) + \lambda\, bits(MB)$$

- Coding 1-bit binary signal indicating the transform type

16×16 Intra Prediction

- 5 modes: DC, Horizontal, Vertical, Down-left, and Down-right
- Pre-filtering: [1 2 1]/4



Entropy Coding of 16×16 blocks

- Binary arithmetic coding
  - Similar to coding 8x8 blocks
    - Inverse scan
    - (level,run) pair coding
    - Binarization
    - Probability model updating
  - Using special set of probability models
  - Arithmetic coding engine unchanged

Experimental Results

Test conditions

Sequence Structure	IBBPBBP....
Intra Frame Period	0.5 second
Entrop Coding	Arithmetic Coding
FME	ON
Deblocking Filter	ON
R-D Optimization	ON
QP	Fixed(22, 28, 34, 40)
Rate Control	OFF
Interlace Handling	PAFF
Reference Frame	2
Search Range	±32
Frame Number	150 (Progressive), 60 (Interlace)
Frame Rate	60 FPS (Progressive), 30 FPS (Interlace)
Resolution	1280x720 (Progressive), 1920x1088 (Interlace)

Experimental Results

	Test Sequence	PSNR Y Gain (dB)	Bit Rate Saving (%)	16x16 Transform Used (%)
Progressive	City	0.123	-4.39	72.02
	Crew	0.219	-9.81	73.07
	Optis	0.136	-5.21	70.58
	Riverbed	0.235	-4.77	80.88
	ShuttleStart	0.163	-6.87	75.34
Interlaced	Fireworks	0.162	-2.19	45.69
	Kayak	0.389	-5.75	66.41
	Flamingo	0.120	-2.26	50.05
	Average	0.193	-5.16	66.76

Conclusion

- A 16×16 ICT is proposed for HD video coding
  - Suitable for 16-bit implementation
  - Good performance
  - PIT, simple for decoders
  - Compatible with the 8×8 ICT in AVS
- Improving coding efficiency significantly
  - Average: 0.2dB, best case 0.4dB
- 16×16 ICT is a useful coding tool for HD video coding

Q & A

- References

[1] Cham, W.K., “Development of integer cosine transforms by the principle of dyadic symmetry”, Proc. IEE, I, 136,(4), pp. 276-282, 1989.

[2] AVS Video Group, “Information technology – Advanced coding of audio and video – Part 2: Video (AVSX WD 3.0)”, AVS Doc. AVS-N1242, Dec. 2005.

[3] Ci-Xun Zhang, Jian Lou, Lu Yu, Jie Dong, W. K. Cham, “The Technique of Pre-Scaled Integer Transform in Hybrid Video Coding”, IEEE ISCAS, 2005. Vol. 1, pp. 316-319, May 2005.

[4] T. Wiegand and B. Girod, “Lagrangian multiplier selection in hybrid video coder control”, in Proc. ICIP 2001, Thessaloniki, Greece, Oct. 2001.