Xing-guo Zhu, Lu Yu, 2015. A reversibility-gain model for integer Karhunen-Loève transform design in video coding. *Frontiers of Information Technology & Electronic Engineering*, **16**(10):883-891. [doi:10.1631/FITEE.1500071]

A reversibility-gain model for integer Karhunen-Loève transform design in video coding

Key words: Integer transform, Karhunen-Loève transform (KLT), Integer-KLT, Transform coding, Video coding

Contact: Xing-guo Zhu
E-mail: 05xxgcsgx@zju.edu.cn
ORCID: http://orcid.org/0000-0002-6319-2471

Introduction

- Karhunen-Loève transform (KLT)
  - The optimal transform that minimizes distortion at a given bit allocation for a Gaussian source
  - Contains non-integers: need integerization

- Integer-KLT design

$$\boldsymbol{A} = \mathrm{int}\,(\boldsymbol{T} \cdot \alpha)$$

  - ***T*** is the original KLT with non-integers
  - α is a scalar factor
  - 'int' denotes integerization which could be rounding, floor, ceiling or other operations
  - ***A*** is the resulted integer-KLT

Traditional method of integer-KLT design

- Integer-KLTs with mathematical rounding

$$A(i,j)=\mathrm{round}(T(i,j)\cdot\alpha).$$

- Integer-KLTs with TCG

$$\boldsymbol{A}=\underset{\boldsymbol{C}\in\Phi}{\arg\max}\{\,\mathrm{TCG}(\boldsymbol{C})\,\}$$

  - Φ is the candidate set of integer-KLTs and ***C*** is a sample in the set
  - TCG(***C***) is the transform coding gain of ***C***

Integer-KLT design with R-G model

- Reversibility of a Transform ***A***

$$\begin{cases} R(\boldsymbol{A}) = \| \boldsymbol{U} \| \\ \boldsymbol{U} = \dfrac{\alpha^2 \boldsymbol{I} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}}{\alpha^2} \end{cases}$$

- TCG loss ratio of a transform

$$\mathrm{LR}(\boldsymbol{A}) = \frac{\mathrm{TCG}(\boldsymbol{T}) - \mathrm{TCG}(\boldsymbol{A})}{\mathrm{TCG}(\boldsymbol{T})} \times 100\%.$$

Integer-KLT design with R-G model

- Reversibility-gain (R-G) model

$$A = \arg\min_{C \in \Phi} \{R(C)\} \quad \text{s.t.}\ 0 \le \mathrm{LR}(C) \le \eta.$$

  - η is a threshold to control the TCG of ***A*** close to that of KLT.
  - The equation seems to be an optimization problem. We use an enumeration method to solve it since it is not convex.

Performance

- Given 16 KLT matrices, we generate three groups of 16 integer-KLTs by different integerization methods.
  - $\boldsymbol{A}_{\mathrm{r}}$: Integerization with rounding
  - $\boldsymbol{A}_{\mathrm{g}}$: Integerization with maximal TCG
  - $\boldsymbol{A}_{\text{R-G}}$: Integerization with R-G model

$$\begin{bmatrix} 0.322 & 0.538 & 0.592 & 0.506 \\ -0.741 & -0.351 & 0.311 & 0.481 \\ -0.557 & 0.623 & 0.180 & -0.519 \\ 0.194 & -0.446 & 0.721 & -0.493 \end{bmatrix} \qquad \begin{bmatrix} 41.26 & 68.90 & 75.77 & 64.75 \\ -94.83 & -44.88 & 39.82 & 61.58 \\ -71.24 & 79.78 & 23.02 & -66.44 \\ 24.80 & -57.08 & 92.34 & -63.12 \end{bmatrix}$$

(a) (b)

$$\begin{bmatrix} 41 & 69 & 76 & 65 \\ -95 & -45 & 40 & 62 \\ -71 & 80 & 23 & -66 \\ 25 & -57 & 92 & -63 \end{bmatrix} \qquad \begin{bmatrix} 41 & 69 & 76 & 65 \\ -96 & -44 & 39 & 63 \\ -72 & 79 & 24 & -67 \\ 25 & -57 & 92 & -62 \end{bmatrix} \qquad \begin{bmatrix} 42 & 70 & 75 & 64 \\ -94 & -45 & 41 & 62 \\ -72 & 79 & 23 & -67 \\ 26 & -57 & 92 & -63 \end{bmatrix}$$

(c) (d) (e)

Fig. 3 Row integer-KLTs for vertical prediction mode

(a) Float KLT matrix; (b) Scaled KLT matrix with α=128; (c) A_{r}; (d) A_{g}; (e) $A_{\text{R-G}}$

Performance

- Performance comparison on video coding
  - BD-bitrate: the bitrate increases of one method when compared with the anchor under the same PSNR
  - Anchor: video coding with $\boldsymbol{A}_r$
  - Experimental platform: KTA-1.9

- Results
  - $\boldsymbol{A}_{R\text{-}G}$ has the best rate distortion performance in video coding

Table 5 BD-rates of A_g and $A_{R\text{-}G}$ when compared with A_r

Sequence number	High bitrate (%)		Median bitrate (%)	
	A_g	$A_{R\text{-}G}$	A_g	$A_{R\text{-}G}$
0	2.74	−0.78	−0.06	0.03
1	3.60	−0.92	0.02	−0.11
2	1.98	−0.35	0.03	−0.08
3	1.80	−0.77	0.08	−0.05
4	2.30	−1.03	0.29	−0.11
5	2.19	−0.32	0.11	0.06
6	1.00	−0.25	0.00	0.02
7	2.23	−0.72	0.05	0.04
8	3.05	−0.91	0.05	0.01
9	0.83	−0.25	−0.06	−0.07
Average	2.17	−0.63	0.05	−0.02

Conclusions

- In this paper, we discuss integerization of KLT and focus on how to maintain useful properties (orthonormality and high TCG) of KLT for video compression.
- An R-G model that jointly considers the TCG and reversibility of a transform is proposed for integer-KLT design. The experimental results show effectiveness of R-G model.